THE IMPLICATIONS OF INTELLECTUAL PROPERTY RIGHTS FOR FIRM STRATEGY AND ORGANIZATIONAL CHOICE

by

Andrew Bowie Tepperman

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Economics
University of Toronto

Copyright © 2001 by Andrew Bowie Tepperman
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.
Abstract

The Implications of Intellectual Property Rights for Firm Strategy and Organizational Choice

Andrew Bowie Tepperman
Doctor of Philosophy
Graduate Department of Economics
University of Toronto
2001

This thesis consists of a set of essays analyzing the influence of intellectual property on different aspects of firm behaviour. Chapter 2 considers the effects of the degree of patent protection on firm organization. In bilateral arrangements, the threat of imitation influences the division of surplus and thus the incentive to make specific investments. Stronger property rights can make the imitation of lower-quality products less profitable. When investments influence expected product quality, this can lead to a disincentive to invest on the part of the patent-holder. For relatively low levels of patent protection, a strengthening of the property right may lead to results opposite those predicted by the traditional theory: contracting parties may also be made worse off by an increase in patent strength.

Chapter 3 characterizes the competitive effects of some common licensing contracts. I show that if firms compete in markets that are related to, but distinct from the market for a patented good, there may be a strong incentive to license the patented product with an intrabrand restraint on price or territory. Since firms will internalize the effects of demand for the patented good on the prices of related goods, contracts that fix the final price of the patented good are generally anticompetitive: the effects of exclusive territory contracts are less severe. I derive conditions under which simple royalty contracts also
can be used to influence prices in related markets for unpatented goods.

The final essay analyzes the effects of Canada's 1989 Patent Act reform on firm behaviour. A set of hypotheses based on pre-existing theoretical models of priority and disclosure policies are developed and tested using a unique sample of Canadian manufacturing firms. I find evidence to support the priority hypotheses: the move to the first-to-file priority rule appears to have induced an increase in R&D spending and patenting intensity. The disclosure hypothesis is not supported: smaller firms did not differentially decrease R&D following the introduction of public application disclosure.
Acknowledgements

I have been very fortunate to have had access to the expertise of a number of excellent economists during the preparation of this thesis. My sincere gratitude is extended to the members of my committee. Nancy Gallini, my principal supervisor, has been a constant, enthusiastic source of good ideas and sound advice. Ralph Winter, Li Hao, and Frank Mathewson provided a challenging audience for my work, while giving generously of their time and suggestions. Working with these individuals has been a pleasure as well as a wonderful learning experience. I am also grateful for the helpful comments on individual chapters provided by Martin Osborne, Carolyn Pitchik, and Jon Putnam.

Others have contributed less directly but no less importantly to my progress during the course of my years at the University of Toronto. Through example and their kind counsel, Dwayne Benjamin and Sam Hollander inspired me to pursue graduate work in economics: my abiding interest in the subject can be traced directly to their encouragement. My colleagues in the graduate program provided an environment that was both intellectually stimulating and relaxing as required. Finally, I thank my wife Leah for her support, patience, and understanding, her reserves of which were often called upon and never exhausted.
Contents

1 Introduction .................................................. 1

2 Intellectual Property Rights in a Theory of the Firm ........ 7
   2.1 Introduction ............................................. 7
   2.2 Basic Model ............................................. 13
      2.2.1 Product development ............................. 13
      2.2.2 Property rights .................................... 16
   2.3 Investment Incentives .................................. 18
      2.3.1 Nonintegration ..................................... 18
      2.3.2 Integration ......................................... 23
      2.3.3 Optimal ownership ................................. 25
   2.4 The Model with a Continuum of Valuations .............. 28
   2.5 Bargaining with Outside Options ........................ 32
   2.6 Further Extensions ..................................... 37
      2.6.1 Introducing costly litigation ..................... 37
      2.6.2 Third-party imitation .............................. 42
   2.7 Conclusion ............................................. 43

3 Licensing with Horizontal Restraints ....................... 51
   3.1 Introduction ............................................. 51
   3.2 Legal Background ....................................... 54
3.3 Monopoly and Duopoly Benchmarks .............................................. 58
3.4 The Price-Setting Contract ...................................................... 61
3.5 The Exclusive Territories Contract ............................................. 69
3.6 An Extension: Royalty Contracts ................................................. 75
3.7 Discussion and Conclusion ....................................................... 78
3.8 Appendix A: Proofs ............................................................... 82
3.9 Appendix B: Characterizing the Exclusive Territories Contract ........ 89

4 Do Priority and Disclosure Rules Matter? ..................................... 92
4.1 Introduction ........................................................................... 92
4.2 The 1989 Policy Changes and Theoretical Predictions .................. 96
4.3 Aggregate Evidence ............................................................. 103
4.4 Firm-level Evidence on R&D .................................................. 105
  4.4.1 Data and variable construction ........................................... 105
  4.4.2 Empirical results: simple model ........................................ 107
  4.4.3 Robustness of the simple specification .............................. 110
  4.4.4 Dynamics and simultaneity .............................................. 112
  4.4.5 Public disclosure and the “small firm” hypothesis ............... 115
4.5 Firm-level Evidence on Patenting ............................................ 117
  4.5.1 Empirical method ........................................................... 117
  4.5.2 Estimation results .......................................................... 120
4.6 Conclusion ........................................................................... 123

Bibliography ............................................................................. 140
List of Tables

3.1 Bertrand and RPM Results: Changing $Y/Z$ Substitutability $(c)$ .... 68
3.2 Bertrand and RPM Results: Changing Relative $Y$ Market Size $(e)$ .... 69
3.3 ET Results: Changing $Y/Z$ Substitutability $(c)$ .................... 72
3.4 ET Results: Changing Relative $Y$ Market Size $(e)$ .................... 73

4.1 Summary Statistics ....................................................... 131
4.2 Regression Estimates for R&D ........................................... 132
4.3 Dynamic Regression Estimates for R&D .................................. 134
4.4 Disclosure Hypothesis Tests ............................................. 136
4.5 Patent Production Function Estimates ................................. 137
4.6 Patent Estimates by Industry ............................................ 139
List of Figures

2.1 Game Structure ................................................. 47
2.2 Partitioning the Patent Strength Space (a) .................. 48
2.3 Effort Choice by S under Nonintegration .................... 49
2.4 Effort Choice by S under Nonintegration with Costly Litigation ..... 50

4.1 Canadian Manufacturing R&D .................................. 125
4.2 R&D in Selected Industries .................................... 126
4.3 Patenting by Canadian Inventors ............................. 127
4.4 Canadian Patent Propensity .................................... 128
4.5 Estimated R&D Year Effects ................................... 129
4.6 Estimated Patenting Year Effects ............................. 130
Chapter 1

Introduction

Intellectual property, fundamentally composed of ideas and knowledge, accounts for a significant proportion of firms' assets. Firms pursue a variety of means to derive value from their ideas. Innovative technologies can, for example, be embodied in new products that are sold directly to consumers; alternatively, a firm can choose to sell or license a technology to a separate firm. Legal mechanisms for protecting intellectual property, such as patents, trade secrets and copyrights, exist in order to encourage the creation of new ideas and products. This thesis consists of a set of essays discussing the effects of these intellectual property rights, in particular patents, on certain aspects of firm behaviour. One of the principal themes of these essays is that the availability of intellectual property rights influences firm behaviour to a significant degree. In this respect, the recent preoccupation of policymakers and academic researchers with the growing prevalence of patenting appears to be well-founded.

A patent can be thought of as a bargain between an inventive firm or individual and society as a whole: in return for allowing the inventor an exclusive right to use the new technology for the life of the patent the precise nature of the technology must be disclosed to the public. This allows the inventor to attempt to capture a return on his or her investment in research, thus rewarding the act of invention, while society benefits from
access to the increment in useful knowledge. The granting of an exclusive right limits
the diffusion of a new technology, and thus represents a cost to society. An implicit
premise of the patent system is, therefore, that this cost is worthwhile. In the absence of
patent protection, according to the prevailing theory, there would be insufficient incentive
to carry out costly research and development (R&D), since technical advances would be
freely available to all. This would lead to a slow rate of scientific progress and inadequate
economic growth.

One aspect of exclusive property rights works to mitigate the costs of limited diffusion:
they provide a basis for contracts between firms. Patents can be sold, licensed, or traded,
and these transactions often work toward an efficient allocation of resources. This is
evident when a patented technology ends up being used by the firm (or group of firms)
most able to make profitable use of it. In a dynamic economy comprising a wide range of
firms with different capabilities it may often be the case that the complementary skills or
equipment necessary to invent a technology and to market a product based on it reside
in separate firms. This type of situation arises naturally in industries in which innovative
start-ups are prevalent, such as biotechnology or software.

When this occurs, a number of decisions must be made regarding the best method
of exploiting the intellectual property. The most important are those that concern the
allocation of the property right to the patented technology. If the original innovating firm
retains the property right the arrangement resembles a licensing contract or strategic
alliance; if the firm in possession of necessary "downstream" assets acquires the right the
arrangement resembles a merger. In light of the wide variety of organizational structures
that are observed in high-tech industries, it is not clear that any one form dominates
the others. The characteristics of certain industries that induce firms to organize in a
particular way are likewise not obvious. What is the role of the enforceability of the
patent right to the ownership decision? In particular, do stronger patents facilitate one
form of organization relative to the other? Given the general perception that patents are
more strongly enforced today than in the past. This issue is of considerable interest to policymakers and scholars.

In Chapter 2 this question is addressed using a model based on standard theories of the firm. The model incorporates a key difference between intellectual property and the assets traditionally considered in these theories: ideas are nonrival, and hence can be used simultaneously by multiple parties. In a bilateral relationship involving incomplete contracts, the possibility of opportunistic imitation provides a backdrop to bargaining over the gains from trade. The threat of imitation therefore affects the incentive to make relationship-specific investments. Assuming that investments stochastically influence the quality of the final good that embodies the technology, I show that a strengthening of property rights from a certain low level eliminates the threat of imitation for lower-quality products. Thus, stronger patents can lead to a disincentive for the patent-holder to invest in improving product quality.

This analysis follows the canonical theory of the firm by assuming that organizational form is chosen so as to provide maximal incentives to invest. The fact that investment may respond negatively to a strengthening of property rights is a novel result, and it has two principal consequences. First, some standard predictions on the boundaries of the firm can be overturned in the region over which there exists a nonmonotonic relationship between investments and patent strength (i.e., at relatively low levels of patent protection). And second, a marginal increase in the protection accorded to patent rights might make the contracting parties worse off. The chapter thus extends the literature on the theory of the firm by incorporating a consideration of the intangible assets that are at the heart of many transactions in a modern economy.

The discussion in Chapter 2 sets out a simple “vertical” contracting environment in which the efficiency gains from cooperation between firms with complementary assets are clearly evident. In general, however, there is no reason to believe that all possible contracts involving patented technologies serve social welfare. How far should the freedom
to write contracts over this exclusive right be allowed to extend? Although this question encompasses a much broader set of issues than can be thoroughly considered here, it is useful to have at least some idea of the factors involved. This is an especially important issue for competition policy practitioners, since a valid patent may confer some degree of market power. While the possession of market power itself is not illegal, antitrust authorities are particularly suspicious of "horizontal" agreements (i.e., agreements between competing firms) that may enhance preexisting market power. These contracts have in fact often been found to facilitate anticompetitive practices. A patent licensing arrangement might provide an ideal environment for horizontal contracting that is detrimental to competition.

Chapter 3 provides a model exploring this issue. Whereas previous discussions of horizontal licensing contracts have largely assumed single-product firms, this essay starts with the observation that firms typically produce a number of products. When a firm with a patented product also competes with a rival in the market for differentiated imperfect substitutes, there can be a strong incentive to license the patented product. The licensee firm is given a share of the profits from the sale of the patented good by allowing it to serve some part of the market for that good. At the same time, cooperative pricing of the patented product is achieved by explicitly setting the price in the contract or by allocating exclusive territories. The licensee will then compete less aggressively in the market for unpatented substitute goods. I show that the extent to which such contracts are anticompetitive depends strongly on their precise form. Price-setting contracts are generally welfare-decreasing, since they allow for commitment to a very high patented good price. On the other hand, contracts containing an exclusive territory clause may in many cases increase welfare: a rationale for this finding using Ramsey pricing is presented. This chapter thus provides some justification for the asymmetric treatment of these types of contracts in the law: price-setting in licensing contracts is in practice regarded as highly suspect, while exclusive territory contracts are evaluated on a case-by-case basis.
Chapter 1. Introduction

The last chapter returns to the fundamental reasons for the existence of the patent system: to stimulate inventive effort and to encourage the disclosure of technological advances. While few would disagree that these are desirable goals, in practice there exist a variety of ways to influence the incentives to perform research and to encourage inventors to make the fruits of that research available to the public. The reform of Canada's Patent Act in 1989 provides an opportunity to test for the effects of certain specific aspects of patent policy on firm behaviour along these dimensions. This study fits into a small but growing literature that empirically analyzes particular aspects of patent policy.

The Canadian patent reform involved a number of important changes, among which two are paramount. When more than one inventor claims to have attained an invention first, some rule is needed to determine which is entitled to a patent. This "priority" rule was changed from first-to-invent (as in the U.S.) to first-to-file. In addition, where before the reform patent applications were kept secret until the grant date, under the new rules all applications are published 18 months after the application date. Understanding the effects of these provisions is important since Canada's experience can provide a guide to other countries, such as the U.S., that are considering adopting similar reforms to their patent policies.

Theoretical models exist that generate predictions on the effects of these two policy changes. To test these predictions I construct a unique dataset of Canadian manufacturing firms. I find results that are broadly consistent with the predictions of the priority theory: R&D spending increased modestly, as did the intensity of patenting. While the R&D result is somewhat sensitive to the specification used, the finding of increased patenting appears to be quite robust. I also find that the data does not support the predictions of the disclosure theory. This is in itself an interesting result, as it points to a particular aspect of the Canadian patent system which policymakers should be aware of. Canada's public disclosure policy might be less effective inasmuch as Canadian firms usu-
ally file for patents first in the U.S.: the U.S. examination system remained confidential through the 1990s.

In more general terms, the analysis in Chapter 4 illustrates empirically that relatively subtle alterations in patent policy can have effects on real economic behaviour. This too is an important finding. While large-scale changes to the enforcement of patent rights such as gradually occurred in the U.S. are likely to be uncommon in the future, many countries may have to make piecemeal changes to their national patent systems in accordance with the trend toward international harmonization. Evaluating the likely effects of these changes requires a more precise understanding of how different aspects of patent policy influence firms’ decisions. Today’s rapidly changing technological environment ensures that these issues will continue to occupy the attention of a large number of scholars and policymakers.
Chapter 2

Intellectual Property Rights in a Theory of the Firm

2.1 Introduction

The effects of weak or imperfect property rights on innovative activity have been acknowledged in a variety of contexts in the economics literature. Williamson (1985) argues that "generic," easily-imitated innovations are normally found in conjunction with firm assets that are unspecialized to any particular customer. Teece (1986) explores the appropriability problem in more detail, and suggests that innovating firms need to develop a strong position in certain complementary assets required for commercialization of the innovation in order to capture a return. More general facets of essentially the same phenomenon can be found in other environments. For example. Markusen (1995) notes that entry into a country by the multinational mode of production (as opposed to licensing) is more common when property rights over new or complex products are weak, and vulnerable to expropriation.

Few formal discussions link the strength of property rights with firm organization, despite this recognition in the literature of the fundamental relationship between the two
notions. In this chapter, I examine the importance of property rights for the incentive to invest and the optimal form of organization. In particular, I focus on the case of intellectual property (IP), which by its nature is associated with imperfect property rights. Consider the following example. A small firm, possibly a start-up, has made an innovation and protected it with a patent. The firm lacks the resources or ability to develop and commercialize a product based on the innovation and thus explores the possibility of negotiating a relationship with a larger downstream firm. Assuming a suitable partner firm exists, the two may then decide to proceed with a product development project that uses the patented technology as an input. The following question arises: when is it in the interest of both parties to develop the product as separate entities, and when does it pay the downstream firm to develop as an integrated firm, buying the technology outright? More fundamentally, what role does the patent itself play in encouraging development efforts and in organizational design?

Although this scenario is stylized, it captures the essence of a fairly common set of transactions. Strategic alliances and acquisitions involving innovative technology-holders have become increasingly frequent over the past two decades (Merges, 1995, 1998; Arora, Fosfuri and Gambardella, 1999). Innovation leading to the discovery of new technologies is often carried out by small firms that have only a limited ability to bring a product to market themselves (Acs, 1996; Almeida and Kogut, 1999). These innovative firms typically depend on licensing, forming an alliance with a larger firm, or being acquired altogether. In biotechnology, start-ups actively court partnerships with larger firms to overcome disadvantages related to scale in commercialization and production, while at the same time taking advantage of their particular skills and know-how (Pisano and Mang, 1993; Gomes-Casseres, 1999). Large producers sometimes also acquire biotechnology firms when they perceive a gap in their expertise in a particular area (Gambardella. 1999).

1 The Economist notes that "big firms are increasingly using their strong share price to acquire start-ups for their innovations... As a result, more and more entrepreneurs are starting enterprises with the express purpose of being bought out in due course." ("Easy way out." February 20, 1999).
1995). Broadly speaking, alliances and integration are alternative ways to take advantage of the complementary capabilities possessed by different firms.

Given the wave of policy changes purported to have strengthened IP rights, it is particularly important to examine how firms respond to changes in the degree of protection.\textsuperscript{2} The evidence as it stands suggests that firms in innovation-intensive industries have reacted to these changes in a variety of ways. Anand and Khanna (2000) provide empirical evidence that where patent protection is more effective, for example in the chemical industry, firms tend to be smaller, to license more, and to engage in more arm’s-length contracting.\textsuperscript{3} The chemical industry is also the focus of a study by Arora and Fosfuri (1998). They argue that entry by “upstream” firms specializing in chemical process design was stimulated in part by the availability of stronger IP protection.\textsuperscript{4} Based on this type of evidence, Merges (1995, 1998) speculates that in general, stronger IP rights enable the development of technologies within disintegrated structures such as joint ventures or other types of alliances.\textsuperscript{3}

This view becomes less convincing in light of evidence from other industries. In software, for example, there appears to be a significant (and growing) degree of backward vertical integration, in which larger technology-using firms acquire specialized start-ups.\textsuperscript{6}

---

\textsuperscript{2} Among the key policy changes are the creation in the U.S. of the Court of Appeals for the Federal Circuit (CAFC) in 1982, which apparently took a more “pro-patent” approach to infringement decisions: the Bayh-Dole Act of 1980, which allowed non-profit researchers to patent and license; and a 1980 court decision that allowed for the patenting of genetically-engineered organisms. See Warshofsky (1994). Kortum and Lerner (1998), and Jaffe (2000).


\textsuperscript{4} Similar evidence is provided by Merges (1999): until fairly recently, chemical and pharmaceutical firms did little outsourcing of fine chemical production. It is now common for small, patent-intensive firms to specialize in the production of particular compounds, while maintaining a relationship with downstream firms.

\textsuperscript{5} The argument contained in Merges (1995) has a strong normative component. He notes that IP rights “reduce the licensee’s opportunistic possibilities and thereby lower transactions costs.... The tantalizing possibility exists that policymakers might actually influence the profitability of these forms [i.e. joint ventures], and thereby indirectly influence industry structure in some cases” (pp. 1573-4, 1588).

\textsuperscript{6} — Marc Andreessen, the technical whiz behind Netscape, envisages a grim future in which innovations happen almost exclusively within small start-up firms, which are then systematically acquired by dominant companies that want to get their hands on the technology or prevent others from doing so.” The Economist. “Easy way out.” February 20, 1999.
Zeckhauser (1996) notes the growth in vertical mergers in the information technology industry. Large semiconductor firms such as Texas Instruments also frequently acquire smaller technology-rich firms. In 1999, for example, TI either acquired or announced its intention to acquire six firms, most of which were active patentees.\footnote{This information was obtained from TI’s 1999 News Release website, located at http://www.ti.com/corp/docs/press/company/pr99.shtml. One of the firms, Unitrode, had received over 50 patents during the 1990s according to a USPTO Patent Bibliographic Database search.}

This evidence challenges our understanding of the way firms react to increases in patent protection. Using the model presented in this chapter, I am able to shed some light on the impact that a strengthening of rights has on the incentive to invest and the organization of innovative activity. In my analysis, a patented input must be modified for use in a final product. The efforts of two managers, a buyer and a seller, customize the input and add value to the product. I assume that a protected basic innovation already exists at the time the parties begin their relationship, and that the input embodies the protected IP. The model applies equally well to the case where a basic patentable discovery is made during the time of collaboration, and the initial contract simply specifies the eventual ownership of the patent. Thus the case of the research joint venture is also encompassed in the formal model.

The influential property rights theory of the firm, developed by Grossman and Hart (1986). Hart and Moore (1990) and Hart (1995) (henceforth GHM), provides a natural starting point for the analysis (see Merges, 1999 for one approach). Traditional GHM theory contends that asset ownership confers “residual control rights”. In situations not explicitly covered by a contract, residual rights give the asset owner the power to decide the action to be taken. Asset ownership enhances the owner’s threat point in bargaining over the gains from trade, and therefore leads to higher investment. Therefore, an optimal ownership rule allocates assets to the party with the most important investment. Like Merges (1999), I argue that strong patents provide residual control rights that allow a patentee with a strong IP portfolio to prevent imitation by the licensee if a collaborative
arrangement breaks down. However, in contrast to Merges, I show that stronger IP rights do not always encourage additional investment by the IP-holder.

In the framework I discuss, managers' investments stochastically influence the quality of the final product: greater investment makes a higher realization of product quality more likely. Because contracts are incomplete, payoffs are determined by bargaining following the realization of quality. The distribution of the gains from trade in the bargaining stage strongly depends on whether the manager who does not own the IP can credibly threaten to imitate and produce alone. Due to the nonrival nature of IP assets, weak property rights afford an opportunity for imitation and use of the same technology by more than one firm. When such \textit{ex post} imitation is not a credible threat, the IP-owning manager captures a larger share of the bargaining surplus.

Opportunistic imitation of the basic technology is costly, and thus its credibility may depend on realized product quality as well as patent strength. When property rights are very weak, imitation is credible regardless of quality. In this sense, a marginal increment to investment (through its effect on the distribution of quality) attracts no additional imitation. For somewhat stronger property rights, however, \textit{ex post} imitation may only be a credible threat for high realizations of product quality. Thus a small increase in investment increases the likelihood of imitation by making a higher quality product more probable. This change in the strength of patent rights can thus \textit{discourage} quality-enhancing investment by the owner of the initial property right over the basic technology.

I find that the IP-holder's investment is increasing in the degree of property right protection only when patents are sufficiently strong. For lower levels of protection investments may be nonmonotonic in patent strength. This result is established first in a simple model, and it can carry over to a more general model in which the final value of the product can take any one of a continuum of possible values. There are two important implications of this result. First, the traditional GHM predictions regarding the optimal
choice of organizational form may not hold. Second, and perhaps more importantly, even when firm boundaries correspond to the GHM predictions, a small increase in the strength of patent protection can make the parties worse off. Finally, an additional source of investment nonmonotonicity arises when IP rights require costly litigation in order to be enforced. If the IP-holder initially litigates for only high-value products, and the IP right is then strengthened so that costly litigation is also desirable for low-value products, the low realization is made relatively more attractive. Investment, which depends on relative bargaining payoffs, may then fall.

This chapter is related to recent extensions of the GHM framework that call into question some of the key GHM conclusions. Using a fully-specified strategic bargaining framework that incorporates the "outside option principle," Chiu (1998) and De Meza and Lockwood (1998) demonstrate that a party can sometimes be given increased incentives to invest by relinquishing ownership of an asset. However, their results are a consequence of the specific bargaining solution, while I obtain similar results simply from the innate nature of IP. I do show, though, that my results extend qualitatively to a model incorporating their more rigorously-based strategic bargaining framework. Rajan and Zingales (1998) also show that ownership can be detrimental to investment incentives. When an investment specializes a piece of capital to a particular use, greater investment may reduce the payoff that the investor can receive elsewhere. The investor then captures a smaller share of the bargaining surplus, which leads to a disincentive to invest. This disincentive can be removed by allocating ownership to the noninvesting party.

A well-developed normative literature explores the implications of patent strength for the incentives to engage in research. The literature on optimal patent design primarily addresses the optimal mix of instruments for encouraging R&D: the length of patent protection, and the breadth of the protected set of products (Gilbert and Shapiro, 1990; Klemperer, 1990; Gallini, 1992). More recently, research has addressed markets in which
innovation occurs sequentially and by different firms. leading to a consideration of the types of licensing agreements firms will engage in (Green and Scotchmer, 1995; Maurer and Scotchmer, 1998). This normative literature for the most part assumes a fixed cost of R&D, and takes the ownership of property rights over the innovation as exogenous. In contrast, my model emphasizes the positive effects of patent rights on effort choices and optimal ownership.

The basic model and its implications for investment and organization appear in sections 2 and 3. In section 4 I explore a version of the model that allows for a continuous distribution of realizations of project quality. Section 5 provides a more detailed discussion of an alternative strategic bargaining model. Section 6 contains a discussion of the effects of costly litigation on investment. Lastly, section 7 summarizes and concludes.

2.2 Basic Model

2.2.1 Product development

The basic framework is related to the models of Baker, Gibbons and Murphy (2001) and Aghion and Tirole (1994). There are two managers, denoted B and S. The sequence of events follows. As of date 0, S has discovered a patented technology which can be used in the development of a final good. At date 0, subsequent to the development of the basic technology, the parties allocate the ownership right over the IP. Either the right remains with S, which is the nonintegrated ownership structure, or it is transferred to B, which is the integrated structure. At date 1, investment decisions are made that modify the technology and contribute to the value of the final product. Manager B’s effort is

---

4Green and Scotchmer (1995) derive a result that is similar to mine. They show that in a sequential innovation context, under some conditions an increase in patent breadth can stifle innovation. When patents are made too broad a second-generation innovator may be able to credibly threaten not to enter, which gives it greater bargaining power in ex ante negotiations. This reduces the initial innovator’s payoff, and can make innovation in the first-generation product unprofitable.

5As mentioned in the introduction, this stage can also represent a decision over which party should own the control rights to a patentable technology that is developed during the course of the alliance.
denoted \( E \in [0, \infty) \), and manager \( S \)'s is \( e \in [0, \infty) \). Both \( e \) and \( E \) are noncontractible and represent costs as well as efforts. At date 2, uncertainty over the value of the final good is realized and \( B \) and \( S \) bargain over the division of the surplus. The final product is assumed to be complex, and therefore its precise characteristics cannot be specified in an initial contract (see Hart, 1995). This is the standard incomplete contracting assumption that creates the need for bargaining \textit{ex post}.\(^{10}\) This game is represented in extensive form in Figure 1.\(^{11}\)

Product development can yield either a high-value or low-value product. When the product is produced by both \( B \) and \( S \) these values are denoted \( V_H \) and \( V_L \). These values are closely related to the revenues available from the product market, and are also assumed noncontractible. The high-value product is developed with probability \( P(e, E) < 1 \), which is increasing and concave in both arguments.

A version of the product can also be sold by a single manager, without access to the facilities or human capital of the other. In this case, if \( S \) sells the product alone, values are \( v_H \) and \( v_L \), and if \( B \) sells alone values are \( w_H \) and \( w_L \). These can be interpreted as spot market monopoly payoffs obtainable in the absence of the other manager. I assume that there are always gains from trade, and that access to the other manager's skills and facilities is important: in other words, \( B \) and \( S \) have complementary assets at date 0. Thus

\[
V_i > \max[v_i, w_i].
\]

\( i = H, L \). Denote \( \Delta V \equiv V_H - V_L \), \( \Delta v \equiv v_H - v_L \), and \( \Delta w \equiv w_H - w_L \), which are all by definition positive. The following assumptions guarantee that investments are more

\(^{10}\)For discussions of the difficulties posed by incomplete contracting over technology, see Teece (1988) and Zeckhauser (1996).

\(^{11}\)The notation \( \text{NB}(V_i, d(\alpha, i)) \) indicates that the Nash bargaining solution is used, and the surplus to be divided is given by \( V_i \) with disagreement outcomes \( d \).
specific than general:

\[ \Delta V \geq \Delta v. \]  
\[ \Delta V \geq \Delta w. \]  

When the conditions in (2.1) hold, an increase in the probability of realizing the high-value product is more valuable inside trade with the other manager than on the spot market. Put slightly differently, the increment to the total surplus from realizing a high-value product is higher than the increment to either manager's outside payoff. Thus effort has a greater marginal impact inside the specific relationship.

For convenience, suppose the probability of developing a high-value product is separable in effort. Then

\[ P(e, E) = p(e) + q(E). \]

where \( p'(e) \geq 0 \) and \( p''(e) < 0 \); similarly, \( q'(E) \geq 0 \) and \( q''(E) < 0 \). Because it is always efficient for \( B \) and \( S \) to cooperate, the first-best investments solve

\[ \max_{e, E} V_L + (p(e) + q(E)) \Delta V - e - E. \]

which gives

\[ p'(e^*)\Delta V = q'(E^*)\Delta V = 1. \]  

Finally, if both \( B \) and \( S \) produce noncooperatively, their substitute products compete in the spot market. The managers then derive values of \( v_i^c \) and \( w_i^c \) where \( v_i^c \leq v_i \) and \( w_i^c \leq w_i \) for \( i = H, L \). This value dissipation is a consequence of product-market competition. I assume further that

\[ \Delta v \geq \Delta v^c. \]  
\[ \Delta w \geq \Delta w^c. \]

so that the payoff increment from achieving the high-value state is larger when the pro-
Producer is a monopolist. A final natural assumption is that competitive payoffs in the high state exceed those in the low state, and thus $w_H^c \geq w_L^c$ and $v_H^c \geq v_L^c$.

In order to compete, the party without the property right must develop a noninfringing imitation of the basic technology. When $S$ initially owns the patent, the cost to $B$ of this endeavour is $K$. If $B$ has the property right, $S$'s imitation cost is $\tilde{K}$. I assume that imitation is profitable in the absence of infringement concerns:

$$w_L^c - K > 0.$$  \hspace{1cm} (2.4)

$$v_L^c - \tilde{K} > 0.$$  

These conditions ensure that if imitation is noninfringing, it will be worthwhile in both states.

### 2.2.2 Property rights

While the standard GHM setup considers ownership over physical assets, this model focuses on the IP right or patent as an asset. An ownership regime specifies initial property rights over the initial patent as well as any other relevant assets. As I have indicated, the two main ownership regimes considered in this chapter are integration, in which $B$ owns the patent (strictly speaking this should be denoted "$B$-integration"), and nonintegration, in which $S$ owns the patent. I do not deal with the possibility of $S$-integration: as I outlined above, in many cases alliances are motivated by the existence of technology owned by small upstream firms that typically need help commercializing a product. Larger downstream firms either fulfill this role by allying or by acquiring the upstream technology outright.\(^\text{13}\)

---

\(^{12}\)These inequalities are satisfied under both Bertrand competition and Cournot competition with linear demand. In the latter case product "quality" can be represented either by a rise in the intercept of the demand function or a fall in marginal cost.

\(^{13}\)This model is thus particularly well-suited to the case of biotechnology. Pisano and Mang (1993) argue that advances in the science of biotechnology altered the upstream competencies necessary to do productive R&D, while the downstream competencies involved in performing clinical trials and marketing
Let the parameter $\alpha \in [0, 1]$ represent the "strength" of the intellectual property right covering the input.\footnote{I will at times also refer to this as the strength of patent protection. While there may be substantial differences in the types of intellectual property protection firms opt for, such decisions are outside the scope of this chapter.} A number of researchers have noted that in practice, IP rights do not provide perfect protection from imitation (Aoki and Hu. 1999; Choi. 1998; Anand and Khanna. 2000). In the spirit of GHM models, this parameter influences the threat points of the two parties in the event that trade of the input does not take place.\footnote{The idea that the IP right influences threat points and thus the distribution of the gains from trade is similar to the setup in Green and Scotchmer (1995).} Under nonintegration, $B$ can sink the imitation cost $K^*$ and with probability $(1 - \alpha)$ discover a noninfringing technology. However, with probability $\alpha$, $B$'s imitation infringes, and he is then unable to produce noncooperatively. Similarly, when $S$ is an employee of $B$, she can threaten to imitate at cost $\tilde{K}$ and develop a noninfringing substitute technology with probability $(1 - \alpha)$. If imitation is successful, $S$ can defect and start a rival firm.

A more realistic treatment of the imitation subgame would allow for a license to be struck following the resolution of an infringement case won by the patent-holder. In the interest of simplicity, I ignore the possibility of this "ex post" licensing in my model (Green and Scotchmer. 1995). This omission entails a considerable gain in simplicity, and is made without a significant loss of generality as long as the basic technology accounts for a relatively large proportion of product value (in the sense that for example, $v_i$ is not much smaller than $V_i$).\footnote{For example, a biotechnology product that is already well on the way toward clinical trials at the time of licensing would fit this assumption. Many of the alliance contracts described by Pisano and Mang (1993) are in fact signed at a relatively late stage in product development.} Further remarks on ex post licensing are deferred to the concluding section.

The strength of patent protection is taken to be exogenous. The parameter $\alpha$ can be given two complementary interpretations. In the first, patent strength is a function of the current patent policy. Policy changes occasionally take place leading to either
stronger or weaker patent rights on average.\textsuperscript{17} Therefore $\alpha$ can represent the average degree to which patents are protected in the economy. Under the second interpretation, $\alpha$ is specific to the product or input being produced, or to the particular sector. Some projects naturally lead to results that can be better protected by formal IP rights, while others may not generate results that can be effectively protected by IP law.\textsuperscript{18} Different industries also have varying "appropriability conditions" in Teece's (1986) terminology, due in part at least to the degree to which the scientific outcome of research can be codified accurately. For example, patents are effective in protecting innovations in the chemical or biotechnology industries, but they are generally regarded as less useful in semiconductors or electronics (Anand and Khanna, 2000; Cohen, Nelson and Walsh, 2000). These appropriability conditions may also change over time, perhaps because of fundamental advances in science.\textsuperscript{19} A change in the parameter $\alpha$ can thus be justified by appealing either to the policy interpretation or to the technological interpretation.

\section*{2.3 Investment Incentives}

\subsection*{2.3.1 Nonintegration}

In the tradition of GHM, I assume initially that after the uncertainty is realized the gains from trade are split according to Nash bargaining.\textsuperscript{20} It is thus necessary to characterize the "threat points" or default payoffs at the time of bargaining. These payoffs will depend on the ability of the non-owner to threaten to imitate.

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{17} Following the establishment of the CAFC in the U.S., district court decisions holding a patent valid and infringed were significantly more likely to be upheld by a higher court than previously. Lerner (1995) cites a figure of 62\% of cases upheld between 1953 and 1978, compared to 90\% from 1982 to 1990.
\item \textsuperscript{18} In order to be patentable, innovations must satisfy utility, non-obviousness, and novelty requirements. Further, the strength of the IP right is likely to depend on whether the patent is found to be "subservient" to another patent, or at the other extreme, whether the patent opens up new avenues for future research (Merges and Nelson, 1990).
\item \textsuperscript{19} Arora and Fosfuri (1998) describe how developments in chemical engineering allowed firms to specify new chemical innovations more accurately. This in turn enabled firms to effectively enforce their property rights in court.
\item \textsuperscript{20} See section 5 for an extension of the model incorporating strategic bargaining.
\end{itemize}
\end{footnotesize}
I begin by considering the case of nonintegration. Supposing that bargaining breaks down, \( B \) threatens to imitate when the expected payoff from doing so is positive, given the realization of product quality. Thus,

\[
(1 - \alpha)w^c_i - K \geq 0
\]

must hold in order for \( B \) to imitate in state \( i \). This condition depends both on product quality and on \( \alpha \), the probability that the imitation infringes.

Condition (2.5) characterizes noncooperative payoffs over \( \alpha \in [0, 1] \). Rearranging this expression defines

\[ \hat{\alpha}_i = 1 - \frac{K}{w^c_i} \]

as the critical value of \( \alpha \) below which imitation is just profitable in state \( i = H, L \). Since \( w^c_H \geq w^c_L \), it is the case that \( \alpha_H \geq \hat{\alpha}_L \). In addition, the assumption in (2.4) above guarantees that \( \hat{\alpha}_L > 0 \).

The interval \([0, 1]\) can therefore be divided into three regions, depending on \( B \)'s incentive to imitate (see Figure 2). In region A, property rights are weak: \( 0 \leq \alpha \leq \hat{\alpha}_L \), so \( B \) can threaten to imitate regardless of the realized state. In region B, property rights are moderately strong, and \( \hat{\alpha}_L < \alpha \leq \hat{\alpha}_H \). In this range \( B \) will imitate in the high state but not in the low state. In region C, \( \hat{\alpha}_H < \alpha \leq 1 \) and property rights are strong. There will be no imitation by \( B \) in this range.

The Nash bargaining payoffs in state \( i \) and region \( j \) can be written

\[
s^i_j = \frac{1}{2} V_i + \frac{1}{2} [d^i_j - f^i_j].
\]

\[
b^i_j = \frac{1}{2} V_i + \frac{1}{2} [f^i_j - d^i_j]
\]

for \( S \) and \( B \), where \( d_i \) and \( f_i \) represent the default payoffs of \( S \) and \( B \) respectively.

Expected utilities \textit{ex ante} are thus

\[
U^i_S = s^i_L + P(e, E) (s^i_H - s^i_L) - e.
\]

\[
U^i_B = b^i_L + P(e, E) (b^i_H - b^i_L) - E.
\]
Bargaining payoffs (and therefore investment choices) differ across regions A, B and C solely because of differences in the default points $d_i$ and $f_i$.

Each party’s threat point in region A is simply his expected payoff in the continuation game in which $B$ imitates:

$$d_i^A = \alpha v_i + (1 - \alpha)v_i^c,$$

$$f_i^A = (1 - \alpha)w_i^c - K.$$

$i = H, L$. $S$ maximizes $U_S^A$ with respect to $e$. The resulting investment choice $e^A$ satisfies the following first-order condition:

$$\frac{1}{2}p'(e^A)[\Delta V + \alpha \Delta v + (1 - \alpha)\Delta v^c - (1 - \alpha)\Delta w^c] = 1.$$  \hfill (2.6)

for $0 \leq \alpha \leq \hat{\alpha}_L$. $B$’s chosen investment, $E^A$, satisfies

$$\frac{1}{2}q'(E^A)[\Delta V - \alpha \Delta v - (1 - \alpha)\Delta v^c + (1 - \alpha)\Delta w^c] = 1.$$  \hfill (2.7)

also for $0 \leq \alpha \leq \hat{\alpha}_L$. It follows from (2.1) and (2.3) that $e^A \leq e^*$ and $E^A \leq E^*$. Also because of (2.3). $S$’s investment is increasing in the degree of property right protection over region A, while $B$’s investment level is decreasing:

$$\frac{de^A}{d\alpha} \geq 0.$$

$$\frac{dE^A}{d\alpha} \leq 0.$$

This is a (locally) continuous restatement of the standard GHM assumption that investments increase in the degree of asset ownership. However, GHM consider a set of discrete assets, while this model concerns a single asset with a continuum of residual control rights.

Now consider region B. Since the only change is that $B$ can no longer threaten to imitate in the low-quality state, $d_H^A = d_H^B$ and $f_H^A = f_H^B$ in state $H$ and

$$d_L^B = v_L,$$

$$f_L^B = 0.$$
in state $L$. Effort choices in this region satisfy

$$\frac{1}{2} p'(e^B) [\Delta V + \alpha v_H + (1 - \alpha) v^*_H - v_L - (1 - \alpha) w^*_H + K] = 1$$

(2.8)

and

$$\frac{1}{2} q'(E^B) [\Delta V - \alpha v_H - (1 - \alpha) v^*_H + v_L + (1 - \alpha) w^*_H - K] = 1.$$  \hspace{1cm} (2.9)

for $\hat{\alpha}_L < \alpha \leq \hat{\alpha}_H$. Within region B, it is again the case that $e$ is increasing in $\alpha$, and $E$ is decreasing in $\alpha$.

The first main result shows that effort choices need not be monotonic (or continuous) in $\alpha$ when comparing effort choices across regions.

**Proposition 1** $e(\hat{\alpha}_L) \geq e(\hat{\alpha}_L + \varepsilon)$ for a small $\varepsilon > 0$. with a strict inequality when $v_L > v^*_L$.

**Proof:** Comparing $S$'s first-order conditions, the level of $e$ at $\hat{\alpha}_L$ is higher than that at $\hat{\alpha}_L + \varepsilon$ when the bracketed term in (2.6) is larger than that in (2.8). Simplifying the resulting expression (and using the fact that $(1 - \hat{\alpha}_L) w^*_L - K = 0$ by definition of $\hat{\alpha}_L$), the inequality becomes

$$(1 - \hat{\alpha}_L)(v_L - v^*_L) \geq 0,$$

which is satisfied. ■

There is thus a downward discontinuity in $S$'s effort choice at the boundary between regions $A$ and $B$. A similar upward jump occurs in $B$'s investment.

An intuitive justification for this proposition follows from the default points. At the border between regions $A$ and $B$, $B$'s threat points remain essentially the same on either side of $\hat{\alpha}_L$ (since $f^B_H = f_H^B$ and $f^A_L = (1 - \hat{\alpha}_L) w^*_L - K = f^B_L = 0$). However, $S$'s default payoff in state $L$ increases from $\hat{\alpha}_Lv_L + (1 - \hat{\alpha}_L)v^*_L$ to $v_L$ (while $d_H^A = d_H^B$). As a result, when $\alpha > \hat{\alpha}_L$, $S$ captures more of the surplus in bargaining in state $L$, both absolutely and relative to state $H$: the low-quality state is now relatively more attractive. Since effort choice depends on relative bargaining payoffs, it falls discontinuously for $\alpha > \hat{\alpha}_L$. 
CHAPTER 2. INTELLECTUAL PROPERTY RIGHTS IN A THEORY OF THE FIRM

This result emerges naturally in a model where property rights are represented by the degree of patent protection. Nevertheless, it is contrary to the notion of "property rights monotonicity" that figures as a key element of GHM models. In GHM, a buyer and seller bargain over the price of an input after investments have been made, much as B and S do here. If the buyer and seller do not agree on a price in the GHM model, the buyer can purchase a lower-quality input on the spot market and still realize some level of revenue. The more physical assets the buyer controls, the greater are her residual control rights, and hence her payoff outside trade with the seller (while the same holds for the seller). When the revenues available outside trade are increasing in investment, this implies that increasing the number of assets over which the buyer has control rights acts to increase her investment.\footnote{In Hart’s (1995) model, the buyer’s revenue from joint production is \( R(i) - p \), where \( p \) is the negotiated price paid to the seller. When trade does not occur, the buyer can get \( r(i; A) - \bar{p} \), where \( A \) is the set of assets over which she has control rights, and \( \bar{p} \) is the spot market price of a substitute input. By assumption, there are always gains from trade, and by Nash bargaining the buyer’s investment solves
\[
\frac{1}{2} r'(i; A) + \frac{1}{2} R'(i) = 1.
\] The key assumption regarding residual control rights is \( r'(i; a1, a2) \geq r'(i; a1) \geq r'(i; \emptyset) \), which implies from the first-order condition that the optimal choice of \( i \) is increasing in the number of assets.} The IP framework generates an endogenous violation of monotonicity in property rights. This arises because the manager not owning the asset may or may not threaten to imitate, which influences the owner’s threat points.

Finally, in region \( C \), B’s threat points are zero regardless of the state, while
\[
d_i^C = v_i,
\]
i = H, L. Investment levels in region \( C \) are constant over \( \hat{\alpha}_H < \alpha \leq 1 \):
\[
\frac{1}{2} \rho'(e^C) [\Delta V + \Delta v] = 1. \tag{2.10}
\]
and
\[
\frac{1}{2} q'(E^C) [\Delta V - \Delta v] = 1 \tag{2.11}
\]
From (2.1) S’s choice of \( e \) is still below the first-best, as is B’s. A comparison of (2.10) and (2.8) shows that S’s effort choice jumps upward at \( \hat{\alpha}_H \). The reason is analogous
to the explanation for the previous proposition: B's default payoffs are (nearly) zero in the neighbourhood of \( \hat{\alpha}_H \) regardless of the state. However, S's threat point in state H increases at \( \hat{\alpha}_H \) (as B can no longer threaten to imitate). Since the default in the low state is unchanged, the relative attractiveness of state H increases and thus the optimal \( e \) jumps up at \( \hat{\alpha}_H + \varepsilon \), for a small \( \varepsilon > 0 \). A symmetric downward jump occurs for \( F \).

The behaviour of S's effort choice over all \( \alpha \in [0, 1] \) is shown in Figure 3. The discussion above explains why effort at point b is greater than that at point a, and why there is a jump downward from b to c. Depending on the exact value of \( \hat{\alpha}_L \), effort at c may be higher than at point a.\(^\text{22}\) Similarly, when \( \hat{\alpha}_H \) is high relative to \( \hat{\alpha}_L \), effort at point d may be higher than at b and a.\(^\text{23}\) It is also possible to show that S's investment is necessarily larger at point e than at b.\(^\text{24}\) One key qualitative conclusion that can be drawn from this discussion of investment behaviour is that although \( e \) is not monotone in general, it is nondecreasing for all \( \alpha \) above \( \hat{\alpha}_L \).

### 2.3.2 Integration

Investment behaviour in the integrated structure (when B has control rights over the IP) is largely symmetric to the nonintegrated case, and thus does not require an equally

\(^\text{22}\)A comparison of the first-order conditions implies that the relevant inequality is

\[
\hat{\alpha}_L (v_H - v_H^c) - (v_L - v_L^c) + \hat{\alpha}_L w_H^c - w_L^c + K \geq 0.
\]

This will be satisfied for relatively high values of \( \hat{\alpha}_L \).

\(^\text{23}\)Effort is higher at d than at b when

\[
(\hat{\alpha}_H - \hat{\alpha}_L) (v_H - v_H^c) - (1 - \hat{\alpha}_L) (v_L - v_L^c + \Delta w^c) \geq 0.
\]

A weaker condition is sufficient for S's investment to be higher at d than at a:

\[
\hat{\alpha}_H (v_H - v_H^c) - (v_L - v_L^c + \Delta w^c) \geq 0.
\]

\(^\text{24}\)The sufficient condition is

\[
(1 - \hat{\alpha}_L) (v_H - v_H^c) - (1 - \hat{\alpha}_L) (v_L - v_L^c - \Delta w^c) \geq 0.
\]

which is met due to the assumption that \( \Delta u \geq \Delta u^c \) (see (2.3)).
detailed discussion. When the property right to the technology belongs to \( B \), \( S \) can threaten to defect and produce competitively by sinking an imitation cost \( \tilde{K} \). Because a noninfringing imitation is developed with probability \( \alpha \), \( S \) imitates in state \( i \) when

\[(1 - \alpha)\hat{c}_i^e - \tilde{K} \geq 0 \quad (2.12)\]

is met. This condition determines \( \hat{\alpha}_L \) and \( \hat{\alpha}_H \), where \( \hat{\alpha}_L \leq \hat{\alpha}_H \), which are analogous to \( \hat{\alpha}_L \) and \( \hat{\alpha}_H \) in the previous subsection.

With weak property rights (where \( 0 \leq \alpha \leq \hat{\alpha}_L \)), \( S \) threatens to defect and compete in both states. Expected utility maximization by \( S \) and \( B \) implies that the optimal investment levels \( \hat{e}^A \) and \( \hat{E}^A \) satisfy

\[
\frac{1}{2} \rho'(\hat{e}^A) [\Delta V - \alpha \Delta w - (1 - \alpha)\Delta w^e + (1 - \alpha)\Delta v^e] = 1 \quad (2.13)
\]

and

\[
\frac{1}{2} q'(\hat{E}^A) [\Delta V + \alpha \Delta w + (1 - \alpha)\Delta w^e - (1 - \alpha)\Delta v^e] = 1. \quad (2.14)
\]

By inspection of these first-order conditions, it follows that \( \hat{E}^A \) is now increasing in \( \alpha \) while \( \hat{e}^A \) decreases in \( \alpha \).

When \( \hat{\alpha}_L < \alpha \leq \hat{\alpha}_H \), so property rights are moderately strong, \( S \) can only threaten to imitate in state \( H \). Thus optimal investments solve

\[
\frac{1}{2} \rho'(\hat{e}^B) [\Delta V - \alpha w_H - (1 - \alpha)w^e_H + w_L + (1 - \alpha)v^e_H - \tilde{K}] = 1 \quad (2.15)
\]

and

\[
\frac{1}{2} q'(\hat{E}^B) [\Delta V + \alpha w_H + (1 - \alpha)w^e_H - w_L - (1 - \alpha)v^e_H + \tilde{K}] = 1. \quad (2.16)
\]

\footnote{For example, Levin (1982, p. 28) points out that in the semiconductor industry, "entry was facilitated in the early years by the ability of a few key employees to appropriate and transfer process and product design know-how sufficient for viable operation. Indeed, the typical new firm was a "spin-off" of an established business founded by a team of key technical personnel ..." Perhaps the most famous example of this movement of personnel was the departure of Robert Noyce and Gordon Moore from Fairchild to form Intel as a start-up in the late 1960's. This story is consistent with Anand and Khanna's (2000) observation that patents in computers and electronics are relatively ineffective at protecting the appropriation of rents from innovation.}
As before, $\hat{E}^B$ (locally) increases in the degree of IP protection.

Lastly, when property rights are strong and $\hat{a}_H < \alpha \leq 1$, $S$ does not imitate in either state. First-order conditions for effort choices are

$$\frac{1}{2} \rho'(\bar{e}^C) [\Delta V - \Delta w] = 1 \quad (2.17)$$

and

$$\frac{1}{2} q'(\hat{E}^C) [\Delta V + \Delta w] = 1. \quad (2.18)$$

The only threat $S$ can wield in this case is the withdrawal of her human capital. When $S$ is responsible for the initial innovation, this human capital can potentially be quite valuable, especially if the technology requires a significant degree of tacit know-how to use. In general, then, $\Delta w < \Delta V$. so (from (2.17)) $S$ exerts some positive level of effort even when property rights are strong.

### 2.3.3 Optimal ownership

Having characterized the optimal levels of investment under different property rights conditions, it is possible to turn to the optimal form of organization. I make the standard assumption that lump-sum transfers are feasible ex ante. The parties will then agree on the ownership arrangement that maximizes their total (expected) surplus, which is

$$TS = V_L + (p(e) + q(E))\Delta V - e - E.$$ 

The *ex ante* division of this surplus is immaterial. As in Hart (1995), any change in ownership structure that increases the investment of one manager while leaving the other unchanged is optimal. This follows from the fact that investments are always below first-best levels.

---

26 Arora (1995) formally discusses the importance of tacit know-how in technology transfer.

27 Aghion and Tirole (1994) implicitly assume that $\Delta w = \Delta V$. In their model $S$'s investment is zero when the property right to the innovation is assigned to $B$. 
The conventional (and intuitive) prediction based on the standard property rights model is that if S's investment is most important in determining the value of the final product, nonintegration will become relatively more attractive when property rights are strengthened. This is based on the idea that when investments are increasing in the degree of control rights, the party with the more important investment should be allocated property rights over the relevant assets.

Indeed, for quite high levels of IP protection, this model accords with the conventional story. It is clear from the foregoing discussion that when $\alpha$ is greater than both $\hat{\alpha}_H$ and $\hat{\alpha}_L$, S's investment is higher under nonintegration than integration, while the reverse is true for B's investment. Thus the standard GHM result goes through: a strong patent right should be allocated to the party with the most important investment.

However, with weaker property rights the predictions are slightly more complicated. In the previous two subsections, I have shown that investment choices do not always behave in the manner assumed by the canonical property rights theory. In particular, stronger property rights do not always induce higher investments. This results in a substantially weaker prediction regarding optimal ownership.

**Proposition 2** Let $q(E)$ be replaced by $\theta q(E)$, where $\theta$ is sufficiently small, and denote $\alpha_m = \max [\hat{\alpha}_L, \hat{\alpha}_L]$. Then if nonintegration is optimal for some $\alpha^* \in (\alpha_m, 1]$, it remains optimal for all $\alpha > \alpha^*$.

We can only be sure that S's effort choice is increasing in $\alpha$ under nonintegration and decreasing in $\alpha$ under integration when $\alpha$ is higher than $\alpha_m$. If $\alpha$ were in a lower range, then the discontinuities in $e$ induced by switching from one region to another could lead to a change in the desired form of organization.

To see that this is in fact possible, consider the following example.\(^{28}\) Suppose for simplicity that $\hat{\alpha}_L = \hat{\alpha}_L$. Thus the degree of IP protection above which the non-owner

---

\(^{28}\)This setup is not motivated by any view that these functional relationships are at all general or likely: it is merely intended to provide a counterexample to the standard GHM prediction.
only imitates in state $H$ is the same regardless of ownership. Assume also that only $S$ invests. Looking at (2.6) and (2.13) it is apparent that nonintegration is optimal at $\hat{\alpha}_L$. The following result demonstrates that the optimal organization may switch to integration when $\alpha$ increases:

**Proposition 3** When $\hat{\alpha}_L = \tilde{\alpha}_L$ and only $S$ invests, integration is optimal at $\alpha = \hat{\alpha}_L + \varepsilon$ (for small $\varepsilon$) when

$$v_L + w_L \geq \hat{\alpha}_L (v_H + w_H) + K + \hat{K}.$$  \hspace{1cm} (2.19)

**Proof:** The result follows directly from a comparison of (2.8) and (2.15). \Box

When moving from region A to region B, $S$'s effort choice falls discontinuously under nonintegration, and rises under integration. Condition (2.19) is sufficient for these jumps to be large enough such that investment is actually higher under integration than nonintegration. It is more likely to hold when the values in the low state are higher, and when the switching point $\hat{\alpha}_L$ is lower.\footnote{It is also necessary that $v_L + w_L > \hat{v}_H + \hat{w}_H$.} The opposite result holds when $B$ is the only investing party.

This switch in the optimal form of ownership depends on the existence of a large investment response under both forms of organization. In the more likely event that a change in $\alpha$ does not have this effect on ownership choice, there are straightforward consequences for the total surplus that the managers expect to generate. The following more general result on expected total surplus can be shown:

**Proposition 4** If only $S$ invests and $\hat{\alpha}_L \leq \hat{\alpha}_L$, a change in $\alpha$ from $\hat{\alpha}_L$ to $\hat{\alpha}_L + \varepsilon$ (for small $\varepsilon$) decreases expected total surplus when

$$v_L + w_L < \hat{\alpha}_L (v_H + w_H) + K + \hat{K}.$$
Chapter 2. Intellectual Property Rights in a Theory of the Firm

Proof: Immediate from investment comparative statics. ■

This condition (the reverse of (2.19)) ensures that S's investment is lower under integration for any \( \alpha > \hat{\alpha}_L \). So if S is the only investing party, nonintegration is optimal on either side of \( \hat{\alpha}_L \). Given that there is no organizational response to the change in \( \alpha \), it is clear that the lower investment decreases expected total surplus. In other words, when the conditions for the proposition are satisfied, an increase in the degree of patent protection makes the contracting parties worse off.

2.4 The Model with a Continuum of Valuations

The previous analysis might appear restrictive in the sense that there are only two possible valuations that a final product can take. In reality, it is likely the case that ex ante, an R&D project could take any number of possible valuations. In this section I investigate the restrictiveness of the two-value assumption by setting out a model in which the possible outcomes of the project are described by a continuously distributed random variable. Specifically, the model in this section will address the robustness of the nonmonotonicity of effort with respect to \( \alpha \).

I assume that only S invests, and that her investment increases the likelihood of realizing a higher-value outcome according to first-order stochastic dominance. Specifically, let \( v \) be the value of the final product when B and S cooperate, and suppose that \( v \) is distributed over the support \( V \equiv [v_L, v_H] \) according to the distribution function \( F(v, e) \). The corresponding density function is \( f(v, e) \), which is positive for all \( v \in V \). First-order stochastic dominance implies that \( F_e(v, e) < 0 \) for all \( v \in (v_L, v_H) \). I make the following

---

30 Note that this condition is sufficient, but not necessary for a decrease in expected total surplus. It is possible that even if (2.19) holds (so investment is higher under integration than nonintegration) for \( \alpha > \hat{\alpha}_L \) the level of investment may still be lower than at \( \alpha = \hat{\alpha}_L \).

31 If the distribution of valuations depends on both parties' investments, comparative statics of effort with respect to \( \alpha \) become considerably more difficult. This is because payoff functions are typically not supermodular in both \( e \) and \( E \), even though both efforts might be complementary instruments for shifting \( F \).
additional technical assumption:

$$\exists \nu^* \in V \text{ such that } f_c(\nu, e) > 0 \text{ for } \forall \nu > \nu^*. \quad (2.20)$$

Roughly speaking, an increase in effort shifts the probability density to the right. This assumption implies that high realizations of \( \nu \) are more likely following a marginal increase in \( e \).

Suppose that the noncooperative payoffs are simple linear functions of \( \nu \). Let \( S \)'s valuation of the final product be \( \sigma \nu \) when she produces alone, and \( \bar{\sigma} \nu \) when she competes with \( B \) in the product market, where \( \bar{\sigma} < \sigma < 1 \). Similarly, let \( B \)'s monopoly and competitive valuations be \( \check{\mathcal{J}} \nu \) and \( \bar{\mathcal{J}} \nu \) respectively, with \( \bar{\mathcal{J}} < \check{\mathcal{J}} < 1 \).

Consider the nonintegrated structure (because of the symmetry of the model integration is a straightforward extension). Fixing \( \alpha < 1 \), \( B \) would attempt to imitate if and only if

$$(1 - \alpha) \bar{\mathcal{J}} \nu - K \geq 0.$$  

Therefore there exists a critical realization of product value, \( \check{\nu} \), such that for \( \nu \) below \( \check{\nu} \) no imitation takes place. This switching point is defined by

$$\check{\nu} = \min \left[ \frac{K}{\check{\mathcal{J}}(1 - \alpha)}, v_H \right]. \quad (2.21)$$

Expression (2.21) defines a function \( \check{\nu}(\alpha) \) such that for interior values of \( \check{\nu} \).

$$\check{\nu}'(\alpha) = \frac{K}{\check{\mathcal{J}}(1 - \alpha)^2}.$$  

Thus the critical realization of product value is increasing in the degree of property rights protection.

For low realizations, where \( \nu \in [v_L, \check{\nu}] \), there is no threat of imitation and \( S \)'s payoff from \textit{ex post} bargaining is\footnote{I continue to assume that the gains from trade are split according to Nash bargaining where each manager receives half of the bargaining surplus in addition to his or her default point.}

$$s^4 = \frac{1}{2} (1 + \sigma) \nu.$$
For higher realizations such that $v \in [\hat{v}, v_H]$, $B$ threatens to imitate, and $S$'s payoff is

$$s^B = \frac{1}{2} (v + \alpha \sigma v + (1 - \alpha)\sigma v - (1 - \alpha)\sigma v + K).$$

Therefore, $S$'s expected utility ex ante is

$$U_S = \int_{v < \hat{v}} s^A dF(v, e) + \int_{v \geq \hat{v}} s^B dF(v, e) - e.$$

If the function $U_S(e, \alpha)$ has "decreasing marginal returns" (i.e., a negative cross-partial derivative) in some range of $\alpha$ then the optimally chosen $e$ is decreasing in $\alpha$ (see Edlin and Shannon, 1998). To see this, first differentiate with respect to $e$:

$$\frac{\partial U_S}{\partial e} = \int_{v < \hat{v}} s^A f_e(v, e) dv + \int_{v \geq \hat{v}} s^B f_e(v, e) dv - 1.$$

This follows from the fact that $e$ enters into the utility function only through the density $f(v, e)$. Differentiating this with respect to $\alpha$ and simplifying gives

$$\frac{\partial^2 U_S}{\partial e \partial \alpha} = (s^A - s^B) \hat{v} f_e(\hat{v}, e) + \int_{\hat{v}}^{v_H} \frac{\partial s^B}{\partial \alpha} f_e(v, e) dv. \quad (2.22)$$

The optimally chosen level of effort is decreasing in $\alpha$ when this expression is negative (Edlin and Shannon, 1998). Notice that $s^A > s^B$ when evaluated at $\hat{v}$, since $S$ receives a higher payoff in states in which imitation does not occur. Because of the fact that effort shifts the density to the right, $f_e$ may be initially negative, but at some $v$ becomes positive. To understand why the first term can be negative, note that an increase in $\alpha$ makes imitation less likely (i.e., makes $\hat{v}$ higher). By itself, this can decrease the incentive to invest since the possibility of realizing a higher-value product is now relatively less attractive.\(^{33}\) The second term shows the marginal effect of an increase in $\alpha$ in states where imitation does occur: this is positive, since the increase in $\alpha$ gives $S$ a larger share of the bargaining surplus and thus increases her incentive to invest.\(^{34}\)

\(^{33}\)This is akin to a "moral hazard" effect: since more low states are "insured" against, in the sense that imitation does not take place, there is less incentive on the margin to try to realize a high state. I thank Li Hao for suggesting this interpretation.

\(^{34}\)Both terms have counterparts in the two-state model analyzed in the previous section. However, because payoffs in that case are discontinuous at the critical values of $\alpha$, the second term is of only second-order importance.
The precise expressions for \( s^A \) and \( s^B \) can be substituted into (2.22) to get the following proposition:

**Proposition 5** A sufficient condition for \( e(\alpha) \) to be decreasing in \( \alpha \) over some range of \( \alpha < 1 \) is

\[
(\sigma - \hat{\sigma})\hat{v}^2 f_e(\hat{v}, e) + (\sigma - \hat{\sigma} + \bar{\beta}) \int_{\hat{v}}^{v_H} v f_e(v, e) dv < 0. \tag{2.23}
\]

This inequality will only be satisfied when the first term exceeds the second in absolute value: a necessary condition for this to happen is that \( f_e(\hat{v}, e) \) must be negative. It is apparent that for values of \( \alpha \) near 1, \( e \) must be increasing since \( \hat{v} \) gets close to \( v_H \) and therefore \( f_e \) eventually becomes positive.

In (2.23) the first term represents the marginal incentive to adjust effort arising from a shift in the density at a specific point, \( \hat{v} \). The second term is actually a weighted average of \( v \) taken over \([\hat{v}, v_H]\), where each value of \( v \) is weighted by the size of the same shift in the density. The inequality will hold with a distribution for which the density falls sharply at some points and for which this average \( v \) is not very high. Suppose that \( v \sim \mathcal{N}(e, \sigma^2) \): valuations are distributed normally with variance \( \sigma^2 \) and effort shifts the mean of the distribution to the right. Then it can be shown that sufficient condition (2.23) becomes

\[
\hat{v} \left[ (\sigma - \hat{\sigma}) \hat{v} \left( \frac{\hat{v} - e}{\sigma} \right) + (\sigma - \hat{\sigma} + \bar{\beta}) \right] + (\sigma - \hat{\sigma} + \bar{\beta}) \frac{1 - \Phi(\hat{v}, e)}{f(e, e)} < 0. \tag{2.24}
\]

where \( \Phi \) and \( f \) represent the normal distribution function and density respectively. Thus, when \( \hat{v} < e \) and \( \sigma \), the standard deviation, is small, (2.24) is satisfied and effort is decreasing in \( \alpha \). Because \( \hat{v} \) must eventually exceed \( e \), the mean of the distribution, as \( \alpha \) increases, it is clear that the inequality is eventually violated and effort increases in \( \alpha \) for relatively high levels of protection.

Why is this important? In the previous section it was demonstrated that an increase in the strength of patent protection could make the bargaining parties worse off. This
analysis extends that result by demonstrating that it holds under alternative, arguably more general conditions.

2.5 Bargaining with Outside Options

Thus far, I have used Nash bargaining to calculate payoffs for B and S in the final stage. This solution concept is consistent with Grossman and Hart (1986) and Hart (1995) (and also the axiomatic Nash solution), and seems to be a natural way to view the division of the surplus. However, an important branch of the recent literature on the theory of the firm characterizes in detail strategic bargaining games that can generate quite different results (De Meza and Lockwood, 1998; Chiu, 1998). In this section I explore the consequences of adopting a similar bargaining solution.

Although this literature shows that different bargaining solutions can generate different results, it is not immediately clear why one should be preferable to another: thus the appropriate choice should be dictated by the context of the model. The fundamental difference between the Nash solution (used in GHM) and the fully specified framework (used by Chiu and De Meza and Lockwood) is that in the former, outside options are considered to be threat points, while in the latter they are constraints. The choice of bargaining approach is made more transparent due to a recent result by Chiu and Yang (1999). They find that in fact both models are special cases of a more general bargaining game: when taking up an outside option leads to a permanent breakdown in bargaining, the constraint model applies: if instead bargaining stops for only a single period the Nash solution applies. In the IP context, it can readily be argued that if taking up the outside option involves a sunk expenditure on imitation (as well as a possible infringement case), the constraint model is more appropriate. It is therefore necessary to consider the model's implications in this other limiting case of the Chiu and Yang framework.

Consider again the simple case in which there are only two possible realizations of
product quality. After the realization of product quality managers $B$ and $S$ play a Rubinstein (1982)-type alternating-offers bargaining game. For simplicity, assume that in each period of this game, either $B$ or $S$ proposes an offer with probability $1/2$. The responder can either accept the offer, reject it, or terminate bargaining altogether in which case the managers receive their outside options, denoted $\omega^k_i$ for player $k = B, S$ in state $i = H, L$. Then it can be shown that in the limit as the discount factor approaches 1, the "outside option principle" is obtained: if $\omega^B_i \leq 1/2V_i$ for $k = B, S$ then each gets a payoff of $1/2V_i$; and if $\omega^k_i > 1/2V_i$ for some $k$, then that player gets $\omega^k_i$ while the other receives $V_i - \omega^k_i$.35

In the nonintegrated case, where $S$ has the property right, the outside options (like the default points previously) depend on the credibility of $B$'s threat of imitation. Thus, outside options are either

$$\left(\omega^S_i, \omega^B_i, \omega^C_i\right) = (\alpha v_i + (1 - \alpha) v^*_i, (1 - \alpha) w^C_i - K)$$

(2.25)

when $\alpha \leq \hat{\alpha}_i$, or

$$\left(\omega^S_i, \omega^B_i\right) = (v_i, 0)$$

(2.26)

when $\alpha > \hat{\alpha}_i$, where $\hat{\alpha}_i$ is the imitation cutoff point described above in section 3.

Assume initially that $v_i > 1/2V_i$ for $i = H, L$ so that $S$'s outside option is "binding" when there is no imitation expected. Thus once the product is developed, cooperation between $B$ and $S$ increases the (total) payoff available, but $S$ can still receive a fairly large share of this without cooperating. This assumption implies that, from (2.26), $S$'s outside option is binding for $\alpha > \hat{\alpha}_i$. Depending on the level of $\alpha$, it may also be binding for lower values. The cutoff level of $\alpha$ at which point $S$'s outside option is binding, denoted $\alpha^*_i$, is defined by

$$\alpha^*_i v_i + (1 - \alpha^*_i) v^*_i = \frac{1}{2}V_i.$$ 

35See, for example, Chiu (1998, Appendix E).
In order to keep the analysis as straightforward as possible, assume that

$$\hat{\alpha}_i \leq \alpha_i^c.$$ 

for \(i = H, L\). This implies that only \(B\)'s threat of imitation determines whether the outside option is binding.

Given these assumptions, and the fact that \(\hat{\alpha}_L \leq \hat{\alpha}_H\), bargaining payoffs can be written

$$s^A_i = b^A_i = \frac{1}{2}V_i$$

for \(i = H, L\) in region A:

$$s^B = \begin{cases} v_L & \text{in state } L \\ \frac{1}{2}V_H & \text{in state } H \end{cases}$$

$$b^B = \begin{cases} V_L - v_L & \text{in state } L \\ \frac{1}{2}V_H & \text{in state } H \end{cases}$$

in region B: and

$$s^C = v_i,$$

$$b^C = V_i - v_i$$

for \(i = H, L\) in region C (see Figure 2 for definitions of these regions).

Recall also from section 3 that investments are chosen to maximize \(U_S^j\) and \(U_B^j\), the expected utilities in region \(j\) for managers \(S\) and \(B\) respectively. It is then straightforward to show that the conclusions of section 3 extend qualitatively to the outside options case.

In particular, \(S\)'s investment in region A satisfies

$$\frac{1}{2}p'(e^A_S)[\Delta V] = 1.$$ 

while in region B, where the low-state outside option binds, effort is given by

$$p'(e^B_S)\left[\frac{1}{2}V_H - v_L\right] = 1.$$
Finally, in region C.

\[ p'(e_C^C) [\Delta \nu] = 1. \]

The effort level in region B is lower than that in region A, by comparison of the first-order conditions. The reason, as before, is that the low state is relatively more attractive. Although effort in region C, where \( \alpha > \alpha_H \), is higher than that in region B, it is ambiguous as to whether it is higher than that in region A. If, for instance, \( 1/2\Delta V^1 > \Delta \nu \), then the highest level of investment by S is obtained when IP rights are weakest. Investment by B behaves analogously: most notably, there is an increase in \( E \) between regions A and B. Finally, if \( w_i > 1/2V_i \), so that B's outside option is binding as a monopolist, the integrated structure is symmetric.

This analysis shows that, under the assumption of a binding outside option for S, the results obtained using a strategic bargaining model are qualitatively similar to those obtained using the Nash solution. One interesting and relevant scenario that can be explored fruitfully in the outside option context is the case where there is an asymmetry between the capabilities of S and B as independent firms. For instance, suppose that S manages a small research firm, with few facilities for production and marketing. On the other hand, B manages a large firm with considerable manufacturing capability. Then once the quality of the product is realized, B's outside opportunities are more valuable than S's.

To formalize the implications of this, suppose that \( v_i \leq 1/2V_i \), while \( w_i > 1/2V_i \) for \( i = H, L \). Then in the nonintegrated case, S's outside option never binds, so both B and S always receive a bargaining payoff of \( 1/2V_i \). However, in the integrated case, B's outside option is binding in the low state when \( \alpha > \alpha_L \), and in the low and high state when \( \alpha > \alpha_H \).\textsuperscript{36} The following proposition offers a strong prediction as to the optimal ownership structure.

\textsuperscript{36}See section 3.2 for a definition of these quantities.
Proposition 6 Suppose \( v_i \leq 1/2V_i \) and \( w_i > 1/2V_i \) for \( i = H, L \). Then:

1. If \( q(E) \) is replaced by \( \theta q(E) \) and \( \theta \) is small (i.e. only \( S \) invests), integration is optimal for \( 0 \leq \alpha \leq \hat{\alpha}_H \). and nonintegration is optimal for \( \alpha > \hat{\alpha}_H \) if and only if \( \Delta w > \frac{1}{2}\Delta V \): and

2. If \( p(e) \) is replaced by \( \varphi p(e) \) and \( \varphi \) is small (i.e. only \( B \) invests). nonintegration is optimal for \( 0 \leq \alpha \leq \hat{\alpha}_H \), and integration is optimal for \( \alpha > \hat{\alpha}_H \) if and only if \( \Delta w > \frac{1}{2}\Delta V \).

Proof: Consider part (1) of the proposition. As \( \theta \to 0 \), only \( S \)’s investment matters for the determination of product quality. Under nonintegration. \( S \)’s investment satisfies

\[
\frac{1}{2}p'(e^Y)[\Delta V] = 1. \tag{2.27}
\]

For the integrated structure, when \( \alpha \leq \hat{\alpha}_L \), so that \( B \)’s outside option does not bind in either state. \( S \) chooses the same level of investment. \( e^Y \). When \( \hat{\alpha}_L < \alpha \leq \hat{\alpha}_H \), so \( B \)’s outside option binds in the low-quality state. \( S \) chooses the investment that satisfies

\[
p'(e^L) \left[ \frac{1}{2}V_H - V_L + w_L \right] = 1. \tag{2.28}
\]

When \( \alpha > \hat{\alpha}_H \), so \( B \)’s outside option binds in both states. investment is given by

\[
p'(e^H) [\Delta V - \Delta w] = 1. \tag{2.29}
\]

Comparing (2.27) and (2.28), it is clear that since \( \frac{1}{2}\Delta V < \frac{1}{2}V_H - V_L + w_L \) (by the assumption that the outside option binds), \( e^L > e^Y \). so integration is superior for \( \alpha \leq \hat{\alpha}_H \): nonintegration is optimal for \( \alpha > \hat{\alpha}_H \) if and only if the term multiplying \( p'(e) \) in (2.27) exceeds the bracketed term in (2.29). or \( \frac{1}{2}\Delta V > \Delta V - \Delta w \). which gives the condition in the proposition. Part (2) of the proposition follows from similar steps using \( B \)’s first-order conditions for effort. 

Thus, the result shows that when \( S \) has the important effort choice. integration is at least weakly (and sometimes strongly) preferable for levels of IP protection up to \( \hat{\alpha}_H \).
When IP rights are stronger, nonintegration is optimal as long as $A_\Delta$ is relatively large. The proposition is driven by the fact that $B$ obtains a relatively high bargaining payoff in the low state for moderate levels of $\alpha$ (specifically, when $\alpha_L < \alpha \leq \alpha_H$). Manager $S$, as the non-owner, is given an additional incentive to invest so that the low state is not realized (as in Chiu. 1998). Of course, it diminishes $B$'s incentive to invest, so nonintegration is strictly preferable when the investment $E$ is more important.

This result rigorously justifies Merges's (1999) hypothesis that stronger property rights facilitate nonintegrated production. One of the interesting aspects of the result is that it emphasizes not only of the benefits of nonintegration but also the costs. The reason nonintegration is not optimal when IP rights are relatively weak is that $S$ actually is given better investment incentives in the integrated structure: the low state is undesirable, as $B$ would then capture a large share of the bargaining surplus. Nonintegration then becomes optimal for $\alpha > \alpha_H$ since $B$ captures too much of the surplus in both states when owning the IP.

2.6 Further Extensions

2.6.1 Introducing costly litigation

There is substantial evidence that the cost of patent litigation is a significant burden, especially for small firms. Survey evidence compiled by Cohen, Nelson and Walsh (2000) reveals that smaller firms are much more likely than large firms to cite the cost of defending a patent in court as a reason for not patenting. Similar findings are presented in

---

37 In the previous sections, the behaviour of $e$ was driven by changes in $S$'s threat point. Here, because of the assumption that the $S$ has only a limited outside option, the effort choice responds solely to changes in $B$'s outside payoff.

38 Sandburg (1998) notes that "while small companies often have no choice but to pursue alliances, exposure can be high if things don't work out as expected. Start-ups, in particular, frequently find the costs of seeking legal redress outweigh the potential returns." A case in point is provided by Go Corp. which was engaged in a joint venture with Microsoft to develop applications for its hand-held computer technology. Although Go alleged that Microsoft subsequently stole its trade secrets, it elected not to litigate.
Cordes, Hertzfeld and Vonortas (1999): a survey of about 200 fairly small high-technology firms in the U.S. indicates that the cost of enforcement is the most commonly cited disadvantage of patent protection. Lanjouw and Lerner (1997) show evidence that a change from the U.S. system of litigation, in which each party pays his own legal fees, to the British system in which the loser pays both sets of fees, substantially increases the value of patents. Lerner (1995) finds that litigation costs can also influence the research strategy of a firm: those firms with higher costs tend to stay away from research areas that have seen heavy patent activity from lower-cost firms.

In this section I return to the Nash bargaining solution. but now I assume a slightly different extensive form continuation game involving the possibility of costly litigation on the part of the patent-holder. Consider the nonintegrated structure (as before, the integrated story is symmetric). If no agreement is reached in the bargaining phase, \( B \) can decide to imitate, as before at a cost of \( K \). The difference here is that infringement is decided by a costly legal procedure. Following imitation, \( S \) can choose to sue \( B \) for infringement, at a cost of \( L \) (while the legal cost is \( l \) for \( B \)). The court then decides if the imitation infringes or not.

I make the following additional assumptions. First, imitation is profitable net of all costs for \( B \):

\[
 w^e_L - l - K \geq 0. \tag{2.30}
\]

Second, if \( S \) were assured of victory in the suit, it would be worthwhile to pursue the case rather than accommodate imitation:

\[
 v_L - L \geq v^e_L. \tag{2.31}
\]

(The fact that both of these conditions are satisfied in the low state implies that they also hold in state \( H \).) A further assumption is necessary, which is not precluded by any previous assumption:

\[
 \frac{w^e_L - l - K}{w^e_L} \geq \frac{L}{v_L - v^e_L}. \tag{2.32}
\]
This assumption ensures that all regimes of imitation and litigation are represented in the \( \alpha \)-space.

The threat points used in bargaining are found by working backwards. Knowing the (exogenous) probability that a court will find the imitation infringing, \( S \) will litigate when the net expected payoff from doing so exceeds the payoff from accommodating, or when

\[
\alpha v_i - (1 - \alpha)\bar{v}_i^c - L \geq v_i^c.
\]

for \( i = H, L \). Rewriting, litigation is credible in state \( i \) if

\[
\alpha(\bar{v}_i - v_i^c) \geq L. \tag{2.33}
\]

Thus, if either litigation costs are too high, or IP right strength too low, litigation will not be credible. Condition (2.33) implicitly defines cutoff points in \( \alpha \). denoted \( \hat{\alpha} \), such that for values of \( \alpha \) greater than \( \hat{\alpha}_i \), litigation is credible in state \( i \). Condition (2.31) ensures that \( \hat{\alpha}_i < 1 \) for \( i = H, L \). while (2.3) implies that \( \hat{\alpha}_H \leq \hat{\alpha}_L \).

In the previous stage, supposing (2.33) to be satisfied for \( i = H, L \). \( B \) would imitate if

\[
(1 - \alpha)w_i^c - l - K \geq 0. \tag{2.34}
\]

If (2.34) is violated, \( B \) does not enter. Thus (2.34) defines switching points \( \hat{\alpha}_L \leq \hat{\alpha}_H \) just as in the previous section. For a value of \( \alpha \) above \( \hat{\alpha}_i \), \( B \) would not imitate in state \( i \) if litigation was expected to be forthcoming. On the other hand, if (2.33) is not satisfied for \( i = H, L \). manager \( B \) expects no litigation, and thus will enter.

Assumption (2.32) ensures that \( \hat{\alpha}_L \leq \hat{\alpha}_L \) so that \( \alpha = [0, 1] \) contains all possible imitation and litigation regimes. There are thus five regions of \( \alpha \).

**Region 1: \( \alpha \in [0, \hat{\alpha}_H] \)** Here, IP rights are very weak: \( B \) threatens to imitate in either state and \( S \) cannot litigate. Threat points are thus given by

\[
d_1^1 = v_i^c.
\]

\[
f_1^1 = w_i^c - K.
\]
Nash bargaining payoffs and thus utilities are calculated as in section 3. Investments therefore satisfy

\[
\frac{1}{2}p'(e_1) [\Delta V + \Delta v^c - \Delta w^c] = 1. \tag{2.35}
\]

and

\[
\frac{1}{2}q'(E_1) [\Delta V - \Delta v^c + \Delta w^c] = 1 \tag{2.36}
\]

for S and B.

**Region 2:** \( \alpha \in (\bar{\alpha}_H, \bar{\alpha}_L) \) In this region S credibly litigates in state \( H \), while it is still profitable for B to imitate in both states. Default points are thus unchanged if the low state is realized, while in the high state they are

\[
d_H^2 = \alpha v_H + (1 - \alpha)v_H^c - L.
\]

\[
f_H^2 = (1 - \alpha)w_H^c - L - K.
\]

Thus effort choices in region 2 are given by

\[
\frac{1}{2}p'(e_2) [\Delta V + \alpha v_H + (1 - \alpha)v_H^c - L - v_L^c - (1 - \alpha)w_H^c + l + w_L^c] = 1. \tag{2.37}
\]

and

\[
\frac{1}{2}q'(E_2) [\Delta V - \alpha v_H - (1 - \alpha)v_H^c + L + v_L^c + (1 - \alpha)w_H^c - l - w_L^c] = 1. \tag{2.38}
\]

for S and B.

**Region 3:** \( \alpha \in (\bar{\alpha}_L, \bar{\alpha}_L) \) In region 3 S will litigate in both states, and B still imitates in both states. Therefore the threat points are

\[
d_H^3 = \alpha v_i + (1 - \alpha)v_i^c - L.
\]

\[
f_H^3 = (1 - \alpha)w_i^c - l - K.
\]

for \( i = H, L \). The first-order conditions defining \( e_3 \) and \( E_3 \) for S and B are identical to (2.6) and (2.7) respectively, since these threat points are just the same as in region A with the addition of constant legal costs.
Region 4: $\alpha \in (\hat{\alpha}_L, \hat{\alpha}_H]$  Here $S$ continues to litigate in both states, while the degree of IP protection is now high enough that it is only profitable for $B$ to imitate in the high state. Threat points are thus the same as for region 3 in state $H$, while in the low state they are now
\[
d_L^i = v_L,
\]
\[
f_L^i = 0.
\]
It is straightforward to show that $e_4$ and $E_4$ are defined by the first-order conditions (2.8) and (2.9).

Region 5: $\alpha \in (\hat{\alpha}_H, 1]$  In this last region, IP rights are at their maximal level, so litigation would occur in either state, and knowing this, it is not in $B$'s interest to imitate at all. Disagreement points are
\[
e_i^5 = v_i,
\]
\[
f_i^5 = 0
\]
in state $i = H, L$, and thus investments $e_5$ and $E_5$ are defined by (2.10) and (2.11) respectively.

Because the investment levels for both $S$ and $B$ in regions 3, 4 and 5 correspond identically to those in regions A, B and C in section 3, the previous monotonicity and comparative statics results hold here as well. In particular, there is a downward jump in $e$ between regions 3 and 4, and an upward jump between 4 and 5. When $S$ owns the IP right, $e$ is thus nondecreasing in the strength of the property right as long as $\alpha$ is above $\hat{\alpha}_L$.

However, there is an additional nonmonotonicity for lower levels of IP protection when litigation is costly. Consider a change in $\alpha$ from $\hat{\alpha}_L$ to $\hat{\alpha}_L + \varepsilon$ for some small $\varepsilon > 0$. Prior to the change, $B$ could imitate in the low state without worrying about litigation, obtaining a default payoff of $w_L^L - K$. At $\hat{\alpha}_L + \varepsilon$, it is now worthwhile for
S to litigate in the low state. so B's threat point falls to \((1 - \tilde{a}_L)\omega_L^c - l - K\). All of the other default payoffs are approximately unchanged. Since B now captures less of the bargaining surplus in the low state, while the division of surplus is the high state is unchanged, the low state is now relatively more attractive to S. This causes a drop in her effort level.

**Proposition 7** \( e(\tilde{a}_L) \geq e(\tilde{a}_L + \varepsilon) \) for a small \( \varepsilon > 0 \).

**Proof:** The level of \( e \) at \( \tilde{a}_L \) is higher than at \( \tilde{a}_L + \varepsilon \) when the bracketed term in (2.6) is larger than that in (2.37). Simplifying, and using the fact that \( \tilde{a}_L (v_L - v_L^c) - L = 0 \) by definition of \( \tilde{a}_L \), the inequality becomes

\[
\tilde{a}_L w_L^c + l \geq 0.
\]

As before, B's effort choice behaves in the opposite way. A summary of S's investment behaviour is shown in Figure 4. Because of the symmetry of the integrated problem, investment behaviour in that case is straightforward to characterize.

**2.6.2 Third-party imitation**

The analysis up to this point has assumed that the only firm capable of imitation is the other party to the collaboration. In reality it may be possible for outside firms to imitate a successful product *ex post*. This gives rise to another potential reason for B to have control rights over the IP, particularly if litigation is costly and the smaller upstream firm has few legal or financial resources available. Larger firms typically employ internal legal counsel, while smaller firms often have fewer legal resources available for a patent infringement dispute, and find patent enforcement prohibitively costly (see Lanjouw and Lerner, 1996; Cohen, Nelson and Walsh, 2000; Cordes, Hertfeld and Vonortas, 1999). Thus, if there is a possibility of imitation and entry by a third party, it may be optimal
to give the ownership right to the larger firm in order to protect the revenues of the product.

Suppose that some entrant, $E$, can imitate either a high- or low-quality product after the development stage, but before $B$ and $S$ bargain over the surplus. If patents are strong, $E$ only enters if there exists no possibility of an infringement suit. Bringing a suit is costly, and the decision to do so is left up to the party with the residual control rights over patent litigation. Thus when $S$'s litigation costs are very high it might be optimal to devolve the right to patent litigation to $B$, or even allocate the entire property right to $B$ if this is necessary to provide the incentive to litigate.\footnote{Lerner and Merges (1998) show that partial allocation of control rights is common in biotechnology alliances. In their sample the right to litigate infringers was allocated to the downstream firm about 25% of the time, while full ownership of the resulting patents was ceded to the customer in only 10% of cases.}

This idea can be interpreted in GHM terms. One of the key propositions of the property rights theory of the firm is that complementary assets ought to be owned together.\footnote{See Proposition 8 in Hart and Moore (1990) and Proposition 2 (D) in Hart (1995, p. 46).} If we interpret the IP right and the litigation technology as complementary assets, and note that from the perspective of the foregoing discussion $B$ inherently has a better litigation technology, it can be optimal for $B$ to own the IP right. In some cases, it might be possible for these assets to be so complementary that $S$ may actually increase her investment following integration. This would be true when entry by $E$ would dissipate the bargaining surplus substantially, so that the disincentive to invest arising from entry by $E$ is greater than that caused by relinquishing ownership of the patent. Thus in addition there need not be a disincentive to invest as a result of giving up ownership.

\section*{2.7 Conclusion}

This chapter shows that the effects of intellectual property rights on investment may be more complicated than the standard property rights theory of the firm would suggest.
In particular, for low levels of IP protection effort choices may be nonmonotonic with respect to changes in the degree of protection. There are two principal reasons why this can occur. First, when an increase in protection causes the non-owner to no longer imitate in the low-quality state, the relative attractiveness of the high-quality state to the owner-investor falls, leading to lower investment. Second, when the change in protection allows the owner to litigate in the low state, where previously this was unprofitable, the share of the bargaining surplus she can capture increases relative to the high state, again leading to lower investment. Both of these effects occur for relatively low levels of protection: thus, while the property rights theory provides an accurate characterization of investment under stronger protection, it may be less so when IP rights are weak.

While the investment results are interesting in their own right, they also have two important implications. It is possible that the GHM predictions on firm boundaries may not apply when patents are weak. In addition, under a wide range of conditions a strengthening of patent protection can lead to lower welfare for the contracting parties. This then represents a sufficient condition for a reduction in the incentive to invent the basic technology in the first place.

The first result is similar in spirit to one previously obtained by Chiu (1998). In his simplest model of bilateral trade, Chiu considers a case where only a single manager invests, and compares the investment level when the asset is owned by the investor to that where the asset is owned by the other party. In the latter case, if ownership of the asset enhances the noninvestor's "outside option" sufficiently, he must be guaranteed at least the value of this option in the bargaining game. This makes the investor the residual claimant to her investment, leading to first-best incentives. In other words, relinquishing ownership over the asset (posting a "hostage") can increase investment incentives. This violation of property rights monotonicity is obtained in Chiu's model by reinterpreting payoffs outside bilateral trade as outside options that can be taken up by terminating bargaining (see Chiu and Yang, 1999). I find a similar violation in the
context of standard Nash bargaining that arises from the effect of property rights on the out-of-equilibrium threat of imitation. Investment can increase with a decrease in property rights in my model when the threat of imitation in the low state makes it less attractive to the investor. This finding also extends, at least qualitatively, to the type of bargaining model considered by Chiu (1998) and De Meza and Lockwood (1998).

A useful and important extension of the model would take into account the desirability of negotiating an *ex post* license after an infringement suit in which the patent-holder prevailed (as in Green and Scotchmer, 1995). Allowing for this would change the bargaining payoffs during the stage where $S$ and $B$ split the gains from trade. As I have pointed out, my nonmonotonicity results continue to hold as long as the IP asset is relatively important. However, there exist some interesting additional effects which are not present in my simple treatment. It can be shown that my results as well as those of Green and Scotchmer (1995) are actually special cases of a more general model. If, in contrast to my assumption, the IP asset on its own contributes very little of the total surplus (i.e. $c_i$ is small relative to $V_i$), the results of Green and Scotchmer go through: an increase in patent strength from a *high* level can lower the owner's investment *ex ante*. This extension promises a more general understanding of the relationship between patent strength and investment.

Another interesting avenue for further research is the incorporation of financial constraints into the model. Aghion and Tirole (1994) argue that financially-constrained research firms might be unable to retain property rights over an innovation if they cannot afford to compensate the downstream firm *ex ante*. Lerner and Merges (1998) find empirically that such financial concerns are important determinants of the allocation of control rights in their sample of biotechnology alliances. Both of these papers implicitly

---

41In deriving their result Green and Scotchmer assume that either all of the value of a sequential innovation project resides in the second-generation product, or that there is no profit dissipation resulting from competing versions of the product. This is in fact very similar to my condition that the basic IP asset has little value on its own.
take the degree of intellectual property protection as fixed.\textsuperscript{42} Investigating the control decision when property rights are imperfect or changing is an interesting empirical and theoretical research area.

\textsuperscript{42}Anand and Galetovic (2000) analyze a model comparing the benefits of venture capital and corporate financing of research projects that incorporates property rights over an innovation. They emphasize the hold-up power that a researcher has over the financing firm when property rights are weak. However, they do not deal explicitly with the researcher's incentives to exert effort on the R&D project.
Figure 2.1: Game Structure

![Game Structure Diagram]

- **B, S**: effort
- **N**: state H
- **N**: state L

**NB(V_H, d(α, H))**

**NB(V_L, d(α, L))**

**NB(V_H, d'(α, H))**

**NB(V_L, d'(α, L))**
Figure 2.2: Partitioning the Patent Strength Space ($\alpha$)
Figure 2.3: Effort Choice by $S$ under Nonintegration
Figure 2.4: Effort Choice by S under Nonintegration with Costly Litigation
Chapter 3

Licensing with Horizontal Restraints

3.1 Introduction

Firms often license patented products or processes to other firms. In many instances licensing can be clearly efficiency-enhancing and generates large social benefits, especially when it allows firms to integrate complementary factors of production. In other cases, licensing can be motivated by a desire on the part of firms to influence the competitive environment in which they operate. Anticompetitive motives for licensing have been identified in both the economics and legal literatures; however, these analyses have largely been restricted to single-product firms. In the single-product case, licensing is not likely to raise serious antitrust concerns, especially when the alternative to licensing is monopoly control of the innovation by the patentee. For example, a license on a new product that restricts competition between the licensor and licensee will be no worse, from a social point of view, than exclusive production by the patentee.²

¹The literature on innovation and licensing assumes for the most part that the technology transferred in a licensing agreement is represented by or embodied in only a single product. See Gallini and Winter (1985), Katz and Shapiro (1985, 1987), Rockett (1990a), and Maurer and Scotchmer (1998).

²This point is explicitly recognized in the DOJ-FTC Antitrust Guidelines for the Licensing of Intellectual Property (United States, Department of Justice and Federal Trade Commission, 1995). However, there are exceptions: for instance, if the licensee is discouraged from further licensing or from using competing technologies by an exclusive dealing provision, further anticompetitive effects may arise.
In this chapter I show that the anticompetitive effects of licensing change importantly when the contracting parties are multiproduct firms. As suggested by Gallini and Trebilcock (1998), a contract between multiproduct firms restricting price competition in only the licensed product may also dampen price competition in related products sold by the firms. My analysis supports their intuition for a variety of contractual restrictions on price, territories, and royalties. In a principal-agent framework, I examine the case of a multiproduct firm deciding whether to license its intellectual property. So that Bertrand competition does not compete away the profits from the licensed good, some sort of intrabrandon restriction is necessary in the contract. Since the competitor now has a stake in the licensed good, it will internalize the effect that its other pricing decision has on the demand for the licensed good. This tends to raise the price of the competitor’s other product and sometimes the licensed product above those prices that would be realized in the absence of licensing. Thus, in this framework, licensing acts as a coordination device for increasing the firms’ joint profits on all products and, depending on the type of restraint imposed, may be strongly detrimental to consumers.

In a general sense, the idea behind this model is related to Bernheim and Whinston’s (1990) finding that multimarket contact can support collusive pricing. However, in their paper collusion in any market is sustained through the threat of severe punishments in all markets: here, coordination is achieved through a restraint on one of the products that allows firms to better account for the degree of substitutability among all products. The model is also related to the literature on the use of strategic licensing to reduce product market competition. In a paper by Eswaran (1994), cross-licensing of competing patents is shown to support collusion. By significantly increasing the punishment following a deviation from collusive pricing, cross-licensing (accompanied by a tacit agreement not to use the licensed technology) can be an effective facilitating device.\(^3\) Rockett (1990a,b)

---

\(^3\)Shapiro (2000), on the other hand, discusses the procompetitive benefits that arise when complementary patents are cross-licensed, a practice that is quite common in high-technology industries.
demonstrates that licensing can be used to crowd out the market from more competitive firms or to leave rivals with older, higher-cost versions of a technology. Other research has addressed the effects of licensing on competition in "innovation markets." Gallini (1984), for example, shows that a firm may have an incentive to license a potential rival in order to discourage the rival from developing a superior technology.

This chapter emphasizes the ability of firms to manipulate interdependent pricing decisions, and thus fits into the literature on contractual restraints. An important paper by Rey and Stiglitz (1995) outlines how territorial restrictions might dampen competition by making the perceived demand curve more inelastic. Whinston's (1990) model of anticompetitive tying shows how a firm with a monopoly position in one product can affect competition in a related product market. By bundling the two goods together, the firm commits to pricing aggressively and potentially forecloses the related market to the rival. In contrast to Whinston who argues that the monopolist may be able to exclude its rival, I look at the incentives for a monopolist to allow entry in order to raise the price in the differentiated market. Finally, Mathewson and Winter (1984) provide a model of restraints that is similar in that it allows for an integrated treatment of both price and territorial restrictions (see also Winter, 1993). Their focus, however, is on the vertical incentives for such contracts arising from the manufacturer-retailer relationship, while my analysis emphasizes only horizontal incentives among competing sellers.

In addition to these theoretical papers, the institutional literature documents the pervasive use of resale price maintenance (RPM) and exclusive territories (ET) in license agreements. Priest (1977) discusses a number of cases in which licensors fixed the final price of the licensed good. His main concern is evaluating whether a given use of RPM

---

4In Whinston's model the anticompetitive outcome is essentially forced on the firm in the tied market. Here, the competitor must be persuaded to accept the contract by offering profits that are at least as high as those the firm would get without licensing.

5In a licensing contract with an RPM provision, the licensor stipulates the final (or more generally, minimum) price at which the good is to be sold. An ET contract allows the licensee to be the sole seller of the good in a given geographic region.
constitutes an efficiency-enhancing arrangement or instead disguises a cartel. However, Priest frames his discussion in the single-product context, and thus omits the incentives considered in this chapter.

In the following section I discuss some of the evidence on the existence of these practices in actual licensing contracts, and the changing legal attitude toward them. Then in section 3 I introduce a simple benchmark model of two firms in which licensing with restraints is illegal, and therefore no licensing takes place. In section 4, I show that it may be profitable for the innovator to license the patented good to its competitor, stipulating the price at which the good must then be sold to consumers. Allocating exclusive territories in which each firm sells the patented good can achieve a similar effect, as I show in section 5. However, since the licensor cannot commit to charging a very high price for the patented good, profits tend to be lower than for the price restriction case. This asymmetry in the anticompetitive effects of these two restraints justifies their asymmetric treatment in the law, sometimes argued to be unfounded. In section 6 I consider an extension to the basic model in which the licensor and licensee produce differentiated versions of the patented good. Under some conditions, a simple royalty contract generates results similar to the vertical restrictions previously examined. Section 7 offers some ideas for further extensions and concludes.

3.2 Legal Background

A number of antitrust cases demonstrate the elements important in this analysis. One famous example is the General Electric case of 1926. In this case GE licensed patents covering the manufacture of superior tungsten bulb filaments to Westinghouse, among other smaller manufacturers, and in addition set final prices for wholesalers and retail-

---

6 See, for example, Mathewson and Winter (1984).
ers. Westinghouse was GE's principal rival in the market for various electrical products around the turn of the century: the two firms accounting for over 75% of electrical goods production in 1896 (Bright, 1949). As such, GE and Westinghouse apparently shared numerous related markets. The U.S. Supreme Court found these price restrictions to be permissible on the basis that the patentee was simply exercising its legally conferred monopoly power. GE's decision to rely largely on setting a final price as opposed to the traditional royalty has led to a variety of interpretations regarding the validity of such agreements. Hovenkamp (1998), for instance, argues that price-fixing (as opposed to setting a royalty) can be profitable when the patent is of dubious validity, or when the licensee is in a position to develop a superior technology. In either of these circumstances, the licensor is assured the monopoly outcome, while the licensee is saved the trouble of either challenging the patent or developing a substitute product. Such licenses would clearly be anticompetitive.

The "GE doctrine" that arose from this case appears to provide antitrust immunity to price-fixing licensors of a valid intellectual property right. Nonetheless, later cases limited the General Electric ruling, and government antitrust officials have requested the courts to overrule it. In U.S. v. Line Material, the Supreme Court found that price restrictions in a cross-licensing agreement were a violation of the Sherman Act. However, as Gallini and Trebilcock (1998) point out, in this case firms were in fact not competitors in the absence of the license, as the patents were in a blocking relationship. Licensing complementary IP rights to free up blocking patents is generally efficiency-enhancing.

---

8 The Court argued: "If the patentee goes further, and licenses the selling of the articles, may he limit the method of sale and the price? We think he may do so, provided the conditions of sale are normally and reasonably adapted to secure pecuniary reward for the patentee's monopoly." (p. 490).

9 In fact, GE did charge a royalty of 2 per cent of sales, to be subsequently reduced to 1 per cent (Priest, 1977, pp. 345-346). This is obviously quite low and apparently was not intended to allow GE to recoup any significant return on its technology.

10 Both of these arguments are quite similar to Gallini's (1984) demonstration of the incentive for a monopolist to deter entry by a potentially more efficient rival by pre-emptively offering a licensing contract.


fact which is now recognized in the U.S. and Canada (see below).

A later case more closely related to GE was the *Newburgh Moire* case of 1956.13 Newburgh developed a superior process for producing effects in cloth. The process was then licensed to two of the four other firms in the industry, with final prices fixed by Newburgh. The Third Circuit's ruling in this case restricted the GE rule to apply only to agreements involving a single licensee. The court argued that the patent law was not intended to allow a patentee to fix the final prices of numerous licensees. Like GE, this case involved a firm licensing a superior technology to other firms in the same industry.14

The 1965 *U.S. v. Huck*13 case appears to be the only example of a successful use of the GE doctrine as a defense. The facts of the case were quite similar to GE: Huck licensed a patented type of metal fastening device to a firm in the same industry stipulating that the licensee was to maintain the same prices as Huck, and in addition pay a nominal (5%) royalty. Both firms manufactured and sold the patented devices. Again, it is unclear whether the licensor and licensee competed by selling other (unpatented) products, but it was pointed out that there existed no agreement preventing the licensee from developing or producing independent, noninfringing fasteners. As in the GE case, no antitrust violation was found.

Recent developments in the antitrust treatment of licensing have tended to emphasize the potential efficiency motives for specific licensing provisions. The 1995 Department of Justice-Federal Trade Commission *Antitrust Guidelines for the Licensing of Intellectual Property* (hereafter referred to as the U.S. Guidelines) explicitly allow for these effi-

---

14 Hovenkamp (1998, ¶2041b) finds the distinction made by the *Newburgh Moire* Court puzzling: “Of course, there is a vast difference between an arrangement in which multiple licensees fix prices with each other, perhaps taking advantage of a provision in their patent licenses purporting to give them a GE immunity: and what appeared to be the *Newburgh Moire* situation, where the patentee negotiated with each licensee individually and they did not conspire or even communicate among themselves. Licensing multiple licensees would ordinarily be thought procompetitive. Further, even if restrained by their licenses from engaging in price competition, they would be able to compete aggressively in other ways.” (italics in original).
Vertical licensing agreements are the most straightforward from an antitrust perspective, as they tend to promote competition. Horizontal contractual restrictions among competing firms are also discussed. Resale price maintenance (RPM) in which the licensor sets the final price of a good is illegal per se, although some possible exceptions may be admitted. Exclusive territories (ET), another intrabrand restraint, is treated according to the rule of reason. The case law described above together with the discussion in the U.S. Guidelines suggests that per se illegality would not automatically extend to the types of restraints considered in this paper, so understanding their effects is a key issue.

In Canada, the Competition Bureau’s draft Intellectual Property Enforcement Guidelines (hereafter the IPEGs) give a specific definition of the anticompetitive effects that will be considered: “The Bureau will not consider licensing agreements anti-competitive unless they reduce competition to a level below that which would have existed in the absence of the license” (section 27). The horizontal effects of a particular licensing practice are of key interest to regulators: “A transaction or conduct must create horizontal effects for the Bureau to conclude it is anti-competitive” (section 61). The IPEGs further note that “IP licensing arrangements ... are inherently vertical, but can have horizontal effects, particularly if the licensor and licensee would have been actual competitors in the absence of the licensing arrangement” (section 62). These guidelines therefore suggest that cases involving restraints of the type considered in this chapter must be subjected to a detailed analysis of their competitive effects. My results provide some guidance toward this end.

---

16 The Agencies’ general approach in analyzing a licensing restraint under the rule of reason is to inquire whether the restraint is likely to have anticompetitive effects, and, if so, whether the restraint is reasonably necessary to achieve procompetitive benefits that outweigh these anticompetitive effects.” (U.S. Guidelines, § 3.4)

17 The U.S. Guidelines draw a distinction between the fixing of the resale price of a good, which is illegal per se, and the fixing of its first sale price, which is evaluated under the rule of reason. The GE doctrine applies in the latter case but not the former. See the U.S. Guidelines, § 5.2.
### 3.3 Monopoly and Duopoly Benchmarks

Suppose there are two goods, \( Y \) and \( Z \), that are related in consumers' demand functions. Each firm sells a differentiated version of good \( Y \) which can be thought of as a different brand: call these \( Y_1 \) and \( Y_2 \). Initially, only firm 1 can produce and sell good \( Z \), since this product is protected by a patent right. Differentiated-products Bertrand competition will take place for good \( Y \). Let the demand functions for \( Y \) facing each firm be symmetric for a given price of \( Z \) and thus \( D^1(p_1, p_2; p_Z) = D^2(p_2, p_1; p_Z) \). The demand function for \( Z \) is given by \( D^Z(p_Z; p_1, p_2) \). These demand functions are generated from a concave quasilinear utility function of the form \( U(q_1, q_2, q_Z) + \epsilon \), where \( \epsilon \) is expenditure on the numeraire good. The associated indirect utility function is \( u(p_1, p_2, p_Z) \). Different versions of the \( Y \) good are substitutes so \( \partial D^i/\partial p_i \leq 0 \) and \( \partial D^i/\partial p_j \geq 0 \). As well, good \( Z \) is a substitute for either type of \( Y \) so \( \partial D^i/\partial p_Z \geq 0 \) and \( \partial D^Z/\partial p_i \geq 0 \) for \( i = 1, 2 \). Suppose that the marginal costs of producing \( Y \) are the same for both types of the good, and are given by \( c^Y \). The marginal cost of \( Z \) is \( c^Z \).

To get a better idea of the externalities in the model, first consider the case where a monopolist produces and markets both versions of \( Y \) as well as \( Z \). The monopolist's profit is

\[
\pi^m = (p_1 - c^Y)D^1(p_1, p_2; p_Z) + (p_2 - c^Y)D^2(p_1, p_2; p_Z) + (p_Z - c^Z)D^Z(p_Z; p_1, p_2). \tag{3.1}
\]

The monopolist maximizes this expression with respect to \( p_1 \), \( p_2 \), and \( p_Z \). The own-price elasticity for \( Y_1 \) is \( \epsilon_{11} = -\partial \log D^1/\partial \log p_1 \) and that for \( Z \) is \( \epsilon_{ZZ} = -\partial \log D^Z/\partial \log p_Z \). The cross-price elasticity between goods \( Y_i \) and \( Z \) is \( \epsilon_{iZ} = \partial \log D^i/\partial \log p_Z \), with \( \epsilon_{iZ} \geq 0 \) for \( i = 1, 2 \). Finally, the "market" elasticity for both varieties of good \( Y \) is \( E^Y = \epsilon_{11} - \epsilon_{12} \), which includes both own-price and cross-price effects. Since

\[
\epsilon_{12} = \frac{\partial \log D^1}{\partial \log p_2}.
\]

Let \( p_i^m \) denote the monopoly price for product \( i \). In equilibrium \( p_1 = p_2 = p_2^m \) due to the symmetry of the \( Y \) demand functions, and therefore the first-order conditions can be
simplified to the usual expressions for the markups by a multiproduct monopolist (see Tirole, 1988):

\[
\frac{p^n_Z - c^Z}{p^n_Z} = \frac{1}{\epsilon_{ZZ}} + \frac{2(p^n_Z - c^Y)\epsilon_{YZ} D^Y}{\epsilon_{ZZ} p^n_Z D^Z}
\]  

(3.2)

for product Z, and

\[
\frac{p^n_Y - c^Y}{p^n_Y} = \frac{1}{E^Y} + \frac{(p^n_Z - c^Z)\epsilon_{ZY} D^Z}{E^Y p^n_Y D^Y}
\]

(3.3)

for product Y. In these expressions the elasticities and the demand functions are all evaluated at monopoly prices. In each of the markup equations, the second term on the right represents the externality imposed by the price of both sets of substitute goods. A monopoly internalizes this fully, since it can prevent consumers from switching to substitute goods by raising the price of the substitutes.

Now suppose that producers compete as duopolists in the Y market. Firm 1, which possesses an intellectual property right for good Z, maximizes the following expression with respect to \(p_1\) and \(p_Z\):

\[
\pi^1 = (p_1 - c^Y)D^1(p_1, p_2; p_Z) + (p_Z - c^Z)D^Z(p_Z; p_1, p_2).
\]

(3.4)

Firm 2, meanwhile, maximizes

\[
\pi^2 = (p_2 - c^Y)D^2(p_1, p_2; p_Z).
\]

(3.5)

Equilibrium duopoly prices are denoted by a "hat". The corresponding markups are

\[
\frac{\hat{p}_Z - c^Z}{\hat{p}_Z} = \frac{1}{\epsilon_{ZZ}} + \frac{(\hat{p}_1 - c^Y)\epsilon_{1Z} D^1}{\epsilon_{ZZ} \hat{p}_Z D^Z}
\]

(3.6)

for product Z, and

\[
\frac{\hat{p}_1 - c^Y}{\hat{p}_1} = \frac{1}{\epsilon_{11}} + \frac{(\hat{p}_Z - c^Z)\epsilon_{1Z} D^Z}{\epsilon_{11} \hat{p}_1 D^1}
\]

(3.7)

for product Y. Firm 2’s maximization yields

\[
\frac{\hat{p}_2 - c^Y}{\hat{p}_2} = \frac{1}{\epsilon_{22}}
\]

(3.8)
Again, the elasticities are evaluated at the equilibrium prices, and $\epsilon_{22}$ is defined analogously to $\epsilon_{11}$. If demand functions are not too convex, the relevant second-order conditions will be satisfied.\textsuperscript{18}

Even given the symmetry of the demand functions, the price that firm 2 offers for its version of good $Y$ will be different from that offered by firm 1. There is a rather complex relationship between the demand functions and the price markups, but some qualitative remarks particularly regarding the price of $Y$ are possible at this stage. First, unlike the monopoly case, individual firms do not take into account the effect of their own price for $Y$ on the demand for the other firm’s version. This is manifest in the difference between $E_Y$, the “market” elasticity for $Y$ which occurs in (3.3), and the “firm” elasticity $\epsilon_i$, which appears in (3.7) and (3.8). This will tend to lower the price of $Y$ in the competitive case from that which would hold under monopoly. In addition, there is an effect on the price of $Y$ which occurs through the demand for good $Z$. Obviously firm 2 does not internalize this effect under duopoly pricing, since it does not sell good $Z$. A low price for $Y2$ induces consumers to substitute away from both $Y1$ and $Z$, the prices of which then fall. This effect follows as a result of strategic complementarity between the two versions of $Y$: when firm 2 prices aggressively, firm 1’s best response is to lower its prices. Firm 1 would like to find a way to soften this aggressive pricing behaviour. In the following sections I outline some ways firm 1 might try to influence the price of $Y$ through the licensing of $Z$ with some sort of intrabrand restraint. These restraints soften competitive strategies in the price-setting game.\textsuperscript{19}

\textsuperscript{18}In firm 1’s case, this means that the determinant of the matrix $H = \begin{bmatrix} \pi_{11}^{1} & \pi_{12}^{1} \\ \pi_{21}^{1} & \pi_{22}^{1} \end{bmatrix}$ must be negative definite, where the cross-partial derivatives are evaluated at 1’s best response to 2’s price. Firm 2’s second-order condition is simply $2D_2^2 + (p_2 - c^Y)D_{22}^2 < 0$, also evaluated at the answer to its rival’s price.

\textsuperscript{19}Tirole (1988) describes a situation in which a firm accommodates an entrant by strategically “under-investing” in capacity to trigger a softer pricing response. He denotes this a “puppy dog” strategy. In the situation considered here, the underinvestment takes the form of the licensing contract, which allows the rival access to a related market.
3.4 The Price-Setting Contract

Firm 1 is able to prohibit its rival from producing good Z because of its (presumably valid) intellectual property right. However, as we have seen above, this exclusive right may cause each firm to ignore an important externality in its pricing decisions. Failing to internalize these externalities depresses prices and joint profits. If it chooses to do so, firm 1 can waive this right by contractually allowing firm 2 to produce and market the good. I assume that firm 2 already possesses the facilities and know-how necessary to produce Z, and is only lacking permission. This is likely to be true in many cases, particularly when firms are close competitors in related goods. The nature of competition in the market for good Z after the licensing contract is struck is key to the terms that might be included. To keep matters particularly simple, I assume that firms 1 and 2 will compete as homogeneous-product Bertrand duopolists in Z as long as both produce it and sell it in the same area. Thus there are no brand effects for Z. While this may appear restrictive, especially considering the differentiated nature of competition in Y, it provides a clearer view of the private incentive for licensing.\(^{20}\)

Suppose initially that firm 1 can choose to license good Z and set the retail price at which Z must be sold. Put differently, firm 1 licenses good Z with an RPM-type provision in the contract.\(^{21}\) Following the stage in which the licensing contract is agreed upon, the remaining retail prices are set by duopolistic competition as before. This contractual form captures the idea that firm 1 must somehow limit the degree of price competition in Z, as otherwise firm 2 will not take the demand for Z into account in setting its other price. Specifically, the common licensing practice of setting a per-unit royalty \(r\) on the sale of Z is not useful here, as the price for Z would just rise to \(r + cZ\) due to the strong

\(^{20}\) I consider the effects of relaxing this assumption in Section 6 below.

\(^{21}\) Given the qualifications to the GE doctrine discussed in the introduction, it is difficult to speculate on the current legal attitude toward this kind of arrangement. Aside from noting that it is not likely to be immediately classed as a per se violation. The U.S. Guidelines mention GE in a footnote, but it does not seem to be taken seriously as a possible defense. Nonetheless, trends in other areas of antitrust enforcement appear to be toward a more favourable view of RPM. See the concluding section below.
form of competition which I have assumed. Firm 2 would not receive any margin on its
Z sales, and would continue to ignore the effect of $p_2$ on the demand for Z. Accordingly,
I omit the choice of $r$ from the following analysis.  

The licensing decision is modeled by a parameter $\lambda \in [0, 1]$ which indicates the pro-
portion of the demand for Z that is allocated or "sold" to firm 2. The retail price for Z
will be the $p_Z > c$ that maximizes joint profits, subject to the "incentive compatibility
constraints" that $p_1$ and $p_2$ are subsequently set noncooperatively by Bertrand competi-
tion. Assume in addition that the licensor can charge a fixed fee $A$ in order to extract the
profit accruing to firm 2 in excess of what it would receive without the license. There-
fore, any licensing contract $\Gamma = \{\lambda, p_Z, A\}$ that increases joint profits over the Bertrand
equilibrium level will be acceptable to both firms, since firm 1 can hold firm 2 to its
reservation profit through the fixed fee. For a given $\lambda$, the licensing problem in its most
general form can be written as:

$$
\max_{p_1, p_2, p_Z} \pi^J = (p_1 - c_Y)D^1(p_1, p_2; p_Z) + (p_2 - c_Y)D^2(p_1, p_2; p_Z) + (p_Z - c)D^2(p_Z; p_1, p_2)
$$

subject to

$$
\begin{align*}
   p_1 &= \arg \max_{p_1} (\bar{p}_1 - c_Y)D^1(\bar{p}_1, p_2; p_Z) + (1 - \lambda)(p_Z - c)D^2(p_Z; \bar{p}_1, p_2) \\
   p_2 &= \arg \max_{p_2} (\bar{p}_2 - c_Y)D^2(p_1, \bar{p}_2; p_Z) + \lambda(p_Z - c)D^2(p_Z; p_1, \bar{p}_2).
\end{align*}
$$

The two constraints represent the "incentive compatibility constraints" facing the li-
censor: after the licensing agreement is made, firms compete as Bertrand duopolists in simultaneously setting the prices of $Y1$ and $Y2$. The solution to this maximization
problem gives the function $\tilde{\pi}^J(\lambda)$.

In analyzing the private incentive to license, the first thing to note is that the way
$p_Z$ is chosen under the RPM contract is different from the no-contract case. As the
specification above illustrates, here the price of good Z is chosen before the other prices.

---

22Recall also that the royalty rate played little to no part in the contract between GE and Westinghouse
discussed above.
to maximize joint profits rather than just the profit of firm 1. Even if \( \lambda = 0 \), this would tend to increase total profit: in particular, \( \hat{p}_Z \) could be chosen if desired. Thus, a finding that there is an incentive to set a positive \( \lambda \) is sufficient to demonstrate the private profitability of licensing.

Before further characterizing the private profitability of this licensing agreement, it is necessary to define some notation. Let

\[
CE_{11} \equiv (p_1 - c^Y)D^1_{11} + (1 - \lambda)(p_Z - c^Z)D^Z_{11},
\]
\[
CE_{12} \equiv (p_1 - c^Y)D^1_{12} + (1 - \lambda)(p_Z - c^Z)D^Z_{12},
\]
\[
CE_{21} \equiv (p_2 - c^Y)D^2_{21} + \lambda(p_Z - c^Z)D^Z_{21},
\]
\[
CE_{22} \equiv (p_2 - c^Y)D^2_{22} + \lambda(p_Z - c^Z)D^Z_{22}.
\]

These conditions represent the "curvature effects" that arise when firms' pricing decisions are constrained by the contract terms. The terminology is derived from the fact that all of these terms are zero when demand is linear.

Let the equilibrium prices under the constrained Bertrand-RPM contract be denoted \((p'_1, p'_2, \hat{p}_Z)\). The following result shows that given a fixed RPM price \( \hat{p}_Z > c^Z \), under some circumstances it is profitable for the licensor to set \( \lambda > 0 \).

**Lemma 1** Given \( \hat{p}_Z > c^Z \), the following condition is sufficient for a positive \( \lambda \) to be privately profitable:

\[
\frac{(p_2 - c^Y)\varepsilon_{12}D^2_{11}}{(p_1 - c^Y)\varepsilon_{21}D^1_{p_1D^1} + (p_Z - c^Z)\varepsilon_{Z2}D^Z_{p_2D^2}} < \frac{-2\varepsilon_{11}/p_1 + CE_{11}/D^1\varepsilon_{Z2}/p_2 + D^2/(p_1D^1)(\varepsilon_{21}/p_1 + CE_{21}/D^1)\varepsilon_{Z1}}{-2\varepsilon_{22}/p_2 + CE_{22}/D^2\varepsilon_{Z1}/p_1 + D^1/(p_2D^2)(\varepsilon_{12}/p_2 + CE_{12}/D^1)\varepsilon_{Z2}},
\]

where the elasticities and demands are evaluated at the constrained Bertrand-RPM prices \((p'_1, p'_2, \hat{p}_Z)\). and \( \lambda = 0 \).

\footnote{It should be emphasized that this result does not show that a joint change in \( \lambda \) and \( p_Z \) from their pre-licensing levels (zero and \( \hat{p}_Z \)) to their post-licensing levels is profitable under the given condition. This particular problem is not amenable to standard comparative statics techniques. Since although \( \lambda \) is continuous, \( p_Z \) tends to "jump" with the RPM contract. Explicit comparisons of profitability of this sort are postponed to the numerical computations.}
Proof: See Appendix A. ■

To understand this result, note that the effect of a change in $\lambda$ on joint profits, given an optimally set $\check{p}_Z$, is given by

$$\frac{d\hat{\pi}^J}{d\lambda} = \frac{\partial \pi^J}{\partial p_1} \frac{\partial p_1^*}{\partial \lambda} + \frac{\partial \pi^J}{\partial p_2} \frac{\partial p_2^*}{\partial \lambda} + \frac{\partial \pi^J}{\partial p_Z} \frac{\partial p_Z^*}{\partial \lambda}.$$  

The final term is zero by the envelope theorem. Therefore, in order for the derivative to be positive when evaluated at $\lambda = 0$, it must be the case that

$$\frac{\partial \pi^J/dp_1}{\partial \pi^J/dp_2} \bigg|_{\lambda=0} < \frac{-\partial p_Z^*/d\lambda}{\partial p_1^*/d\lambda} \bigg|_{\lambda=0}. \quad (3.10)$$

where the inequality has been reversed since $p_1^*$ falls with an increase in $\lambda$ (as shown in Appendix A). Rearranging this expression in elasticity form gives the proposition.

Looking at the left-hand side of expression (3.9), it appears that the relationship is more likely to be satisfied when $p_2$ is initially quite close to marginal cost, while the margins on $p_1$ and $p_Z$ are relatively high. These conditions are in line with the results of the somewhat informal discussion of the pre-licensing Bertrand equilibrium in the previous section. The right-hand side of (3.9) shows that both elasticities and curvature effects are important for the private agreement.24 The elasticity terms derive from the fact that as shown in Appendix A, when $\lambda$ increases, $p_1$ tends to fall and $p_2$ tends to rise, other terms remaining equal. These terms thus pick up the internalization of $Z$ demand in the setting of $Y$ prices that occurs as a result of the license. The profitability of a marginal change in $\lambda$ depends on how these price changes affect demand. The curvature terms, denoted $CE_{ij}$, capture the idea that convergence between the $Y$ prices may lead to diminished incentives to compete in the Bertrand pricing game. These terms derive from the degree of concavity of the demand functions. These effects will not make (3.9) harder to satisfy unless the smaller difference between the $Y$ prices significantly enhances profit-eroding competition.

24The third-degree price discrimination literature identifies similar elasticity and curvature incentives for private optimality (see Holmes, 1989; Winter, 1997).
Lemma 1 characterizes conditions under which some degree of licensing is privately desirable. For the straightforward case of linear demand (in which curvature effects are absent) it is possible to show that symmetry between the demand for \( Y_1 \) and \( Y_2 \) leads to \( \lambda^* = 1/2 \). Consider the system of demand functions given by

\[
D^1(p_1, p_2; p_Z) = a - dp_1 + b p_2 - cp_Z \\
D^2(p_1, p_2; p_Z) = a - dp_2 + b p_1 + cp_Z \\
D^Z(p_Z; p_1, p_2) = a_Z - cp_Z + cp_1 + cp_2.
\]

These demand functions are associated with an indirect surplus function of the form

\[
S(p_1, p_2, p_Z) = z - ap_1 - ap_2 - aZp_Z + \frac{1}{2}dp_1^2 + \frac{1}{2}dp_2^2 + \frac{1}{2}ep_Z^2 \\
- bp_1p_2 - cp_1p_Z - cp_2p_Z.
\]

where \( z \) is an arbitrary constant. The parameters \( c \) (not to be confused with costs of production \( c^Y \) and \( c^Z \)) and \( e \) are important for the following discussion. The first represents the degree of substitutability between \( Y \) goods and product \( Z \), while the second is inversely related to the market size for product \( Z \). With this demand system the right-hand side of (3.9) equals one, and Proposition 8 follows:

**Proposition 8** With linear demand given by system (3.11), \( \lambda^* = 1/2 \).

**Proof:** See Appendix A.

The proof illustrates that when \( \lambda < 1/2 \), joint profit is increasing in \( \lambda \). Prices \( p_1 \) and \( p_2 \) converge as \( \lambda \) nears \( 1/2 \) and the incentive to raise \( \lambda \) disappears. A similar result will hold for general demand functions, provided the curvature effects are not too strong.

To fully understand the effects of licensing with the RPM provision, it is necessary to investigate the optimal \( \bar{p}_Z \). Because this price is chosen at the outset, and is taken as given during the stage where the firms make their \( Y \) pricing decisions, the \( Z \) price choice is qualitatively similar to the Stackelberg game of quantity precommitment (see
Tirole, 1988; Church and Ware, 2000). More formally, the Bertrand equilibrium prices for \( Y_1 \) and \( Y_2 \) are functions of \( p_Z \). Because goods \( Y_i \) and \( Z \) are demand substitutes, strategic complementarity implies that the Bertrand prices will be increasing in \( p_Z \) when \( \lambda^* = 1/2 \). This gives the following proposition:

**Proposition 9** For \( \lambda^* = 1/2 \), the Bertrand equilibrium prices \( p_1^* \) and \( p_2^* \) are increasing functions of \( p_Z \).

**Proof:** See Appendix A. ■

Rewriting the derivative in Appendix A gives

\[
\frac{dp_1}{dp_Z} = \frac{\varepsilon_2 \cdot \varepsilon_1 \cdot \varepsilon_2}{2\varepsilon_1\varepsilon_2 - \frac{d^2}{dp_1} \cdot [(p_1 - c') (D_1 + D_2) + \frac{1}{2}(p_Z - c^2) (D_1^2 + D_2^2)]}.
\]

(3.13)

A symmetric condition holds for \( p_2 \). This equation can be decomposed to illustrate that both elasticity and curvature effects are also at work in determining the impact of \( p_Z \) on \( Y \) prices. Naturally, a change in \( p_Z \) alters the pricing incentives for \( Y_1 \) and \( Y_2 \) through its direct effect on the demand for these products. The elasticities in both the numerator and the denominator capture this. In addition, this price change could reduce the effectiveness of Bertrand competition in lowering \( p_1 \) and \( p_2 \). The bracketed terms highlight this possibility. In the linear demand case, curvature effects are not present, and it can be shown that this derivative is larger the higher is \( c \), the degree of substitutability between \( Y \) and \( Z \). The reason why this should hold is intuitive: when \( Y \) and \( Z \) are more substitutable, a higher \( p_Z \) induces consumers to switch to purchasing \( Y \) goods, which firms can then price higher. This result also demonstrates the complicated effects of licensing on the equilibrium \( Y \) prices. Holding \( p_Z \) constant, licensing allows \( p_1 \) and \( p_2 \) to converge. But if RPM licensing contracts are allowed to be made before the

\[25\] For example, consider the constant elasticity specification with \( D^1 = p_1^{-\sigma} p_2^{-\gamma} p_Z^{-\gamma} \). This typical nonlinear demand function generates positive own- and cross-price second derivatives and therefore the bracketed terms in both the numerator and the denominator will be positive. The curvature effects will then increase \( \frac{dp_1}{dp_Z} \).
competitive stage, then firm 1 will be able to take into account the effect of \( p_Z \) on the other prices, which as a result will be higher in equilibrium.

Firm 1’s first-order condition for setting the optimal \( p_Z \) is determined by maximizing joint profits with respect to \( p_Z \), given that \( (p_1, p_2) = (p_1^*, p_2^*) \). This yields

\[
\frac{\partial \pi^f}{\partial p_Z} + \frac{\partial \pi^f}{\partial p_1} \frac{\partial p_1^*}{\partial p_Z} + \frac{\partial \pi^f}{\partial p_2} \frac{\partial p_2^*}{\partial p_Z} = 0. \tag{3.14}
\]

The first term in this expression represents the “direct effect” of a change in \( p_Z \) on joint profit, while the next two terms give the “indirect effect” arising from the interaction between the price of \( Z \) and the Bertrand equilibrium prices of \( Y_1 \) and \( Y_2 \). Given that firm 1 maximizes joint profits, as opposed to its single-firm profit, and that it takes into account the effect of \( p_Z \) on the other prices, it is not surprising to find that the \( p_Z \) chosen here is higher than the Bertrand equilibrium price shown in the previous section.\(^{27}\) Recall, however, that \( p_Z \) is chosen here to maximize joint profits rather than only 1’s profit. Because of this basic difference between the two problems, a linear-demand computational exercise is performed, the results of which are given below.

The \textit{a priori} intuition regarding the effect of price changes is ambiguous: a higher price of \( Z \) might not reduce \( Y \) output and consumer surplus, since consumers substitute away from \( Z \) to \( Y \). The numerical comparative statics results shown in Table 1 illustrate

\(^{26}\)Strictly speaking, an additional constraint is the requirement that \( \bar{p}_Z \leq p_Z^{\text{choke}} \), where the latter is the price at which demand for good \( Z \) falls to zero. For the analytical presentation in the text I assume that this condition is satisfied. However, for the computational exercise that follows, it is crucial to ensure that it is satisfied, as for some parameter values firm 1 would actually like to make the demand for \( Z \) negative. In these calculations I restrict attention to the combinations of parameters for which the choke condition is nonbinding.

\(^{27}\)In fact, it is possible to show that the RPM price for \( Z \) is higher than the monopoly price (conditional on the realized \( p_1 \) and \( p_2 \)). To see this note that condition (3.14) can be rewritten as

\[
D^Z + (p_Z - c^Z)D^Z_2 + 2(p_1 - c^Y)D^1_2 + 2 \frac{dp_1}{dp_Z} \left[ (p_2 - c^Y)D^1_2 + \frac{1}{2}(p_Z - c^Z)D^2_2 \right] = 0.
\]

Term 1 represents the standard monopoly pricing first-order condition, while term 2 (which is positive as long as \( p_1 \) and \( p_2 \) are below the monopoly prices) indicates the additional effect of an increase in \( p_Z \). In order that the first-order condition is satisfied term 1 must be made negative by raising \( p_Z \) above the monopoly level.
the dependence on c, the degree of substitutability between \( Y \) and \( Z \) showing that this reasoning does not appear to hold. This analysis is performed using the linear demand system given in (3.11), with \( a = a_Z = 1, c^Y = c^Z = 0, b = 0.5, \) and \( d = e = 2.28 \). I also set \( z = 3 \) in order to fix consumer surplus (measured by (3.12)) at an arbitrary level corresponding to zero prices.

For all values of c it is privately profitable to implement the RPM licensing agreement, but social welfare, defined as the sum of consumer surplus and profits, falls. Initially the licensing agreement causes \( p_1 \) to fall and \( p_2 \) to rise. As c is raised, all RPM prices and profits increase, and licensing remains profitable in each case. However the divergence between the social and private desirability of the licensing contract increases as c grows. For example, when c = 1, the price of \( Z \) is raised in the RPM contract sufficiently high to choke off almost all \( Z \) demand. Consumers then indeed are forced to substitute to \( Y \), but because of the very high degree of strategic complementarity between \( p_Z \) and the \( Y \) prices, \( p_1 \) and \( p_2 \) are induced to be so high that consumer surplus falls precipitously (see equation (3.13)).

\[ 28 \text{This set of parameters satisfies both the second-order conditions and the choke price constraint. The latter constraint first binds at a level of c just slightly above 1.} \]
Table 3.2: Bertrand and RPM Results: Changing Relative Y Market Size (e)

<table>
<thead>
<tr>
<th></th>
<th>e = 2</th>
<th></th>
<th>e = 4</th>
<th></th>
<th>e = 6</th>
<th></th>
<th>e = 8</th>
<th></th>
<th>e = 10</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B RPM</td>
<td></td>
<td>B RPM</td>
<td></td>
<td>B RPM</td>
<td></td>
<td>B RPM</td>
<td></td>
<td>B RPM</td>
<td></td>
</tr>
<tr>
<td>$p_1$</td>
<td>0.391</td>
<td>0.388</td>
<td>0.336</td>
<td>0.333</td>
<td>0.319</td>
<td>0.316</td>
<td>0.310</td>
<td>0.308</td>
<td>0.305</td>
<td>0.304</td>
</tr>
<tr>
<td>$p_2$</td>
<td>0.348</td>
<td>0.388</td>
<td>0.315</td>
<td>0.333</td>
<td>0.305</td>
<td>0.316</td>
<td>0.300</td>
<td>0.308</td>
<td>0.297</td>
<td>0.304</td>
</tr>
<tr>
<td>$p_Z$</td>
<td>0.390</td>
<td>0.478</td>
<td>0.186</td>
<td>0.220</td>
<td>0.123</td>
<td>0.143</td>
<td>0.091</td>
<td>0.106</td>
<td>0.073</td>
<td>0.084</td>
</tr>
<tr>
<td>$\Sigma \pi^t$</td>
<td>0.701</td>
<td>0.716</td>
<td>0.501</td>
<td>0.506</td>
<td>0.441</td>
<td>0.443</td>
<td>0.411</td>
<td>0.413</td>
<td>0.394</td>
<td>0.395</td>
</tr>
<tr>
<td>$S$</td>
<td>2.085</td>
<td>2.015</td>
<td>2.330</td>
<td>2.304</td>
<td>2.406</td>
<td>2.391</td>
<td>2.444</td>
<td>2.432</td>
<td>2.465</td>
<td>2.457</td>
</tr>
<tr>
<td>$W$</td>
<td>2.786</td>
<td>2.731</td>
<td>2.832</td>
<td>2.810</td>
<td>2.847</td>
<td>2.834</td>
<td>2.855</td>
<td>2.845</td>
<td>2.859</td>
<td>2.852</td>
</tr>
</tbody>
</table>

In Table 2 I present the results of a similar exercise in which the parameter $e$ varies from 2 to 10. A higher $e$ indicates that the $Y$ market is relatively dominant. The other parameters remain the same with the exception of $c = 0.5$, which is an intermediate value from the previous set of assumptions.\(^{29}\) Unlike the previous table, $p_1$ falls and $p_2$ rises for all values of $e$. As $e$ is increased, the $Z$ market becomes relatively less important and there is therefore less scope for influencing $Y$ prices with the horizontal restraint. In this set of simulations the welfare difference between Bertrand and RPM decreases as $e$ grows, and the RPM contract remains privately profitable. Nevertheless, welfare is lower than in the absence of licensing.

### 3.5 The Exclusive Territories Contract

In the previous section, the licensee, firm 2, was induced to take the demand for $Z$ into account when pricing $Y_2$ by creating a markup over cost in $Z$. This markup was ensured by the RPM provision specified in the contract. Another possible way which firm 1 might produce this wedge is to allow firm 2 an exclusive territory in which to

\(^{29}\)These results do not depend on the assumption that $b$ and $c$ are equal.
sell $Z$, and then as before to extract firm 2's excess profit with the fixed fee $A$. This is a particularly important case to consider for two principal reasons. First, competition authorities' hostility toward the inclusion of price-fixing arrangements in licenses may lead firms to seek substitute contractual devices instead of risking a costly antitrust suit. Second, and most importantly, territorial division is explicitly permitted under the U.S. Patent Act (Hovenkamp. 1998).30

The ET contract is in many respects similar to the RPM contract discussed in the previous section. However, while the final price of $Z$ was previously specified in the licensing contract, this price is now determined by each firm at the same time as the other prices. Let $\lambda$ represent the fraction of the demand for $Z$ satisfied by firm 2. Then the ET contract is denoted $\Gamma^* = \{\lambda, A\}$. In order for equilibrium prices to be well-defined in this model, assumptions on arbitrage are necessary. I assume that consumers cannot arbitrage $Z$, so that there are essentially two relevant prices for good $Z$: $p^1_Z$ in firm 1's territory, and $p^2_Z$ in 2's territory. This partitions the demand for $Y$ goods as well, since for example some of the demand for $Y1$ will come from consumers who reside in firm 2's $Z$ territory.

For a given $\lambda$, firm 1 solves

$$\max_{p^1_Z, p^2_Z} \pi^1 = (p_1 - c^Y) \left[ (1 - \lambda)D^1(p_1, p_2; p^1_Z) + \lambda D^1(p_1, p_2; p^2_Z) \right] + (1 - \lambda)(p^1_Z - c^Z)D^2(p^1_Z; p_1, p_2).$$

and firm 2's problem is symmetric:

$$\max_{p^2_Z} \pi^2 = (p_2 - c^Y) \left[ (1 - \lambda)D^2(p_1, p_2; p^2_Z) + \lambda D^2(p_1, p_2; p^1_Z) \right] + \lambda(p^2_Z - c^Z)D^2(p^2_Z; p_1, p_2).$$

The solution to this problem and its comparative statics with respect to $\lambda$ are considered in more detail in Appendix B. Here, to maintain comparability with the RPM contract, I

30Hovenkamp (1998. 72044a) notes: “Further, in those cases in which the patentee also makes the manufactured product in a territory, the statute explicitly authorizes a form of “horizontal” territorial division that would be illegal per se if done in the absence of an intellectual property license... Because the territorial provision is express in the patent statute, the rule permitting territorial restraints rests on a stronger footing than the General Electric rule pertaining to price fixing.”
focus on the special case of the symmetric ET contract which sets $\lambda^* = 1/2$. The equilibrium price markups for $Y_1$ and $Y_2$ can be rearranged from the first-order conditions to give

$$\frac{p_i^r - c^r}{p_i^r} = \frac{1}{\epsilon_{11}} + \frac{1}{2} \cdot \frac{(p_Z^r - c^r)\epsilon_{12} D^Z}{\epsilon_{11} p_i^r D^1}. \quad (3.15)$$

Because of the symmetry of the demand functions, $p_1^r = p_2^r$. The common equilibrium price of $Z$ then satisfies $p_Z^1 = p_Z^2 = p_Z^r$. with

$$\frac{p_Z^r - c^r}{p_Z^r} = \frac{1}{\epsilon_{22}} + \frac{(p_1^r - c^r)\epsilon_{12} D^1}{\epsilon_{22} p_Z^r D^2}. \quad (3.16)$$

A comparison between the symmetric ET prices and the RPM prices leads to the following proposition:

**Proposition 10** *Prices in the RPM contract are at least as high as those arising from the symmetric ET contract: $p_i^r \leq p_i^r$ for $i = 1, 2$, and $p_Z^r \leq p_Z^r$.*

**Proof:** See Appendix A. ■

This result follows directly from the fact that in the RPM case firm 1 can commit to a price for $Z$ before both $p_1$ and $p_2$ are determined, since it is set in the stage prior to price competition. In the symmetric ET contract this commitment is no longer possible. Further, the price of $Z$ is now set only to maximize individual profits rather than joint profits: firms individually set $p_Z$ as local monopolists at the same time as they compete with each other in $Y$ prices.

Comparing joint profits across regimes shows that the licensor would prefer, if possible, to set the final price in the initial contract:

**Proposition 11** *Joint profits in the RPM regime are at least as high as joint profits under the symmetric ET contract.*

---

31In the linear demand examples discussed below, joint profit does indeed attain a maximum at $\lambda = 1/2$. 

---
Table 3.3: ET Results: Changing Y/Z Substitutability (c)

<table>
<thead>
<tr>
<th></th>
<th>c = 0.1</th>
<th>c = 0.25</th>
<th>c = 0.5</th>
<th>c = 0.75</th>
<th>c = 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>p_1</td>
<td>0.297</td>
<td>0.319</td>
<td>0.369</td>
<td>0.447</td>
<td>0.579</td>
</tr>
<tr>
<td>p_2</td>
<td>0.297</td>
<td>0.319</td>
<td>0.369</td>
<td>0.447</td>
<td>0.579</td>
</tr>
<tr>
<td>p_Z</td>
<td>0.272</td>
<td>0.310</td>
<td>0.388</td>
<td>0.501</td>
<td>0.684</td>
</tr>
<tr>
<td>Σπ^i</td>
<td>0.486</td>
<td>0.549</td>
<td>0.703</td>
<td>0.965</td>
<td>1.485</td>
</tr>
<tr>
<td>S</td>
<td>2.324</td>
<td>2.252</td>
<td>2.086</td>
<td>1.820</td>
<td>1.337</td>
</tr>
<tr>
<td>W</td>
<td>2.809</td>
<td>2.801</td>
<td>2.788</td>
<td>2.785</td>
<td>2.821</td>
</tr>
</tbody>
</table>

Proof: See Appendix A.

A legal prohibition on RPM contracts may therefore force firms to substitute to a less profitable symmetric ET contract. Although profit falls, welfare will tend to increase when one restraint is substituted for the other. Table 3 shows the effects of a symmetric ET contract on prices, consumer surplus, profit, and welfare, for the linear demand model with the same parameter combinations as Table 1. A more general treatment of welfare effects is presented in Appendix B.

Comparing Tables 1 and 3, first note that the licensing contract is profitable in each case. The territorial restraint still allows p_2 to rise above the no-license price in each case, while p_1 falls in all cases. Interestingly, p_Z is actually lower with the ET contract than without licensing, suggesting that the effect on joint profit from raising p_2 is strong enough to outweigh the inability to set a higher Z price. The result that for high c the Z price is sufficiently high as to choke off demand, which holds under RPM, does not hold with the ET contract. It is more profitable for firms to concentrate in Y, as a comparison of profit levels in the two regimes shows, but Z prices under ET cannot be sustained to a high enough degree. This point shows that the Stackelberg character of the RPM game allows firms a degree of commitment to a high Z price which is simply not feasible with the ET contract. Finally, note that in all cases welfare is at least as high with the ET
contract as in the absence of licensing. This can be attributed to the fall in the patented good price (relative to no licensing) in connection with the convergence of the unpatented good prices. Below I provide a Ramsey pricing explanation for this finding.\footnote{More general profitability and welfare analysis is carried out in Appendix B.}

To allow comparison with Table 2, now suppose $c$ is fixed at 0.5 and let $e$ vary. Table 3.4 presents the results of this exercise. Again, all values of $p_1$ fall relative to the Bertrand prices, while $p_2$ rises. As $e$ is increased, RPM and ET profits converge, since there is less scope for extending market power. Consumer surplus is always higher with the ET contract than without the license, although by a very small amount.

Because the linear demand example abstracts away from potentially important curvature effects, it is merely illustrative of some of the effects which should guide horizontal restraints policy. The analysis of the previous section emphasized that licensing can be driven by elasticity incentives. But the curvature incentives, ignored in the computations, can also lead firms to engage in licensing in order to dull the intensity of price competition. The natural question to ask is whether one type of contract tends to harm welfare more than the other. Proposition 10, which applies in both the linear and nonlinear cases, indicates that regardless of the degree to which the curvature effect is driving the
change in prices, all prices will be lower with the ET contract. Therefore the potential harm to competition arising through collusive pricing is mitigated if RPM contracts are not allowed.

Welfare comparisons between the RPM and ET contracts can be addressed more generally by looking at a Ramsey pricing exercise. Denote consumer surplus by \( S(p_1, p_2, p_Z) \). Then by Roy's identity, \( S_i = -D_i \). The optimal Ramsey prices maximize social welfare (consumer plus producer surplus) subject to the constraint that profits are set at some positive level.\(^{33}\) Straightforward calculations yield the following Ramsey expression:

\[
\left( \frac{p_1 - c^Y}{p_1} \right) \left( \epsilon_{11} - \epsilon_{21} + 2\epsilon_{1Z} \frac{p_1 D^1}{p_Z D^Z} \right) = \left( \frac{p_Z - c^Z}{p_Z} \right) \left( \epsilon_{ZZ} + \epsilon_{Z1} \frac{p_Z D^Z}{p_1 D^1} \right). \tag{3.17}
\]

and in addition \( p_1 = p_2 \).

Consider first the symmetric ET contract. Conditions (3.15) and (3.16) can be rearranged to allow for easier comparison with the Ramsey optimum. We then have \( p_1^v = p_2^v \) and

\[
\left( \frac{p_1^v - c^Y}{p_1^v} \right) \left( \epsilon_{11} + \epsilon_{1Z} \frac{p_1^v D^1}{p_Z D^Z} \right) = \left( \frac{p_Z^v - c^Z}{p_Z^v} \right) \left( \epsilon_{ZZ} + \frac{1}{2} \epsilon_{Z1} \frac{p_Z^v D^Z}{p_1^v D^1} \right). \tag{3.18}
\]

There are several differences between the \( p_1/p_Z \) price ratio in this expression and the optimal price ratio in (3.17). First, firms still do not internalize the \( Y \) externality, so the cross-price elasticity \( \epsilon_{21} \) does not appear on the left-hand side of (3.18). Second, firms internalize the effect of their \( Z \) price on \( Y \) consumers in only their own territory, so the \( YZ \) elasticity term is weighted by unity instead of 2. These two forces work in opposite directions. Last, since in the symmetric ET equilibrium each firm serves only one half of the \( Z \) market, firms do not fully internalize these price effects and the second term on the right-hand side is multiplied by \( \frac{1}{2} \). This effect tends to decrease the equilibrium price ratio, although the combined impact on the price ratio is ambiguous.

\(^{33}\)The typical Ramsey pricing exercise involves maximizing total welfare subject to the constraint that the profits of a natural monopoly must be nonnegative. In the problem considered in this chapter, profits are bounded away from zero by the nature of differentiated-products Bertrand competition. The same second-best analysis thus applies.
Now consider the RPM contract in Ramsey form. Recall that the optimal $\bar{p}_Z$ is found by solving (3.14), where the effect of the specified $Z$ price on Bertrand competition is given by (3.13). Calculations yield

$$\left(\frac{p'_1 - c^Y}{p'_1}\right) \left(\epsilon_{11} + 2\epsilon_{1Z} \frac{p'_1 D^1}{\bar{p}_Z D^Z} + 2(\epsilon_{11} - \epsilon_{21}) \frac{D^1}{D^Z} \frac{d p_1}{d p_Z}\right) + 2 \frac{D^1}{D^Z} \frac{d p_1}{d p_Z} \quad (3.19)$$

and $p'_1 = p'_2$. One key difference between the symmetric ET and RPM contracts is the commitment to a higher price in the latter. Because the derivative terms are positive, the bias away from the Ramsey optimum is more pronounced with the RPM restraint: the $p_1/p_Z$ price ratio is too low. When cross-price effects between $Y$ and $Z$ are important (see Table 1) this bias becomes quite severe, in particular because the derivative becomes larger as cross-price effects increase. Even in the case related to Table 2 where own-elasticity effects are strong, the RPM contract does not necessarily lead to a close approximation of the Ramsey optimum, although the damage to welfare does become increasingly attenuated.

### 3.6 An Extension: Royalty Contracts

In many realistic cases it is likely that once licensing takes place, the licensee will produce a slightly different product than the licensor. Therefore, it is important to consider the possibility that competition in $Z$ is not of the pure Bertrand variety, as assumed to this point, but rather differentiated-products Bertrand competition. With differentiated $Z$ products, a role for the royalty rate emerges. Intuition suggests that a high royalty might be able to reproduce some of the same effects as, say, an RPM contract, provided that price competition does not entirely erode each firm’s margin on its $Z$ product. In this section I present a comparative static result indicating conditions under which a simple two-part pricing contract can be used to increase the price of $Y$, even in the absence of any other contractual terms.
Suppose that now the licensing contract is completely characterized by $\Gamma^d = \{r, A\}$. where $r$ is the royalty charged by firm 1 on firm 2's sales of its version of $Z$, and $A$ is a fixed fee paid from firm 2 to firm 1. $Y$ prices are denoted $p_i, i = 1, 2$ as before, and now $Z$ prices are $q_i, i = 1, 2$. Marginal costs are constant, and are denoted $c^Y$ and $c^Z$. The symmetric demand functions are $y^1(p_1, p_2; q_1, q_2)$ and $y^2(p_1, p_2; q_1, q_2)$ for $Y1$ and $Y2$ respectively, and $z^1(q_1, q_2; p_1, p_2)$ and $z^2(q_1, q_2; p_1, p_2)$ for $Z1$ and $Z2$. All own-price derivatives are negative, and all cross-price derivatives are positive (i.e. all goods are gross substitutes). Profit functions are then

$$\pi^1 = (p_1 - c^Y)y^1(p_1, p_2; q_1, q_2) + (q_1 - c^Z)z^1(q_1, q_2; p_1, p_2) + r \cdot z^2(q_1, q_2; p_1, p_2) + A,$$

$$\pi^2 = (p_2 - c^Y)y^2(p_1, p_2; q_1, q_2) + (q_2 - c^Z - r)z^2(q_1, q_2; p_1, p_2) - A.$$

In general, the optimal licensing contract from the perspective of firm 1 is the solution to

$$\max_{r, A} \pi^1(p_1, p_2; q_1, q_2)$$

subject to

$$(p_1, q_1) = \arg \max_{\hat{p}_1, \hat{q}_1} \pi^1(\hat{p}_1, p_2; \hat{q}_1, q_2).$$

$$(p_2, q_2) = \arg \max_{\hat{p}_2, \hat{q}_2} \pi^2(p_1, \hat{p}_2; \hat{q}_1, q_2).$$

$$\pi^2 \geq \bar{\pi}^2.$$

The first two conditions represent the incentive compatibility constraints, which require that prices constitute a differentiated-products Bertrand equilibrium. The last constraint is the individual rationality condition, stating that firm 2 must be induced to accept the contract by receiving no less than the profits absent the license.

Since the purpose of this investigation is to find conditions under which the licensing contract profitably influences equilibrium $Y$ prices, it is not necessary to provide a full characterization of the solution to this problem. Instead, I perform a comparative static
exercise on \( p_2 \) to see whether in equilibrium it is increasing in \( r \). which is a necessary condition for the contract to satisfy the goals of firm 1. Let subscripts \( i \) and \( j' \) denote derivatives with respect to \( Y_i \) and \( Z_j \) respectively.

**Proposition 12** A sufficient condition for \( \frac{dp_2}{dr} \geq 0 \), assuming a unique, stable Bertrand equilibrium, is:

\[
-z_2^2 + \left( \frac{-\pi q_2}{\pi_1} \right) z_1^2 + \left( \frac{-\pi q_2}{\pi_1'} \right) z_1' \cdot \left( \frac{\pi q_2}{\pi_2'} \right) z_2' \geq 0.
\] (3.20)

**Proof:** See Appendix A. □

Notice that the first term is negative, while the remaining terms are positive: this follows from strategic complementarity and the fact that goods are gross substitutes. The first term arises from the "direct" effect that an increase in \( r \) has on the choice of \( p_2 \): because \( r \) squeezes the margin that firm 2 can achieve on \( Z \) there is thus less scope for \( Z \) demand to influence the equilibrium \( Y \) price. The remaining terms indicate the effect that the other prices have on \( p_2 \). An increase in the royalty rate tends to increase all other prices. If these indirect effects outweigh the direct effect then the net impact of \( r \) on \( p_2 \) is positive, and therefore the strategy may be profitable.

This analysis is suggestive, but it is not precise: a good deal of information, particularly regarding the curvature of the demand functions, is subsumed in the parenthesized terms.\(^{34}\) It is again instructive to look at a linear demand example, in which effects arising from the relative curvature of the demand functions are not present. Consider the following modification of the demand system in (3.11):

\[
y^1(p_1, p_2; q_1, q_2) = a - p_1 + b p_2 + c q_1 + c q_2
\] (3.21)

\[
y^2(p_1, p_2; q_1, q_2) = a - p_2 + b p_1 + c q_1 + c q_2
\]

\[
z^1(q_1, q_2; p_1, p_2) = a z - e q_1 + f q_2 + c p_1 + c p_2
\]

\[
z^2(q_1, q_2; p_1, p_2) = a z - e q_2 + f q_1 + c p_1 + c p_2
\]

\(^{34}\)It is possible, although not particularly instructive, to rewrite the terms in parentheses to indicate their dependence on elasticity and curvature effects, as demonstrated in previous sections. The complexity of the resulting expressions limits the usefulness of this exercise.
CHAPTER 3. LICENSING WITH HORIZONTAL RESTRAINTS

In (3.21), the parameter \( b \) represents the degree of substitutability between \( Y_1 \) and \( Y_2 \). while \( e \) indicates the relative degree of \( Y \) market dominance: a high \( e \) means that the \( Z \) market is relatively small. The new parameter \( f \) now measures the substitutability of \( Z_1 \) and \( Z_2 \). Condition (3.20) can readily be reduced to

\[
\left( \frac{1}{e} - 1 \right) c + \left( \frac{b}{2} \right) c + \left( \frac{c}{2e} \right) f \geq 0.
\]

\textit{Ceteris paribus}, a higher \( f \) certainly increases the chance that this expression will be satisfied, which is in line with the intuition presented above. When \( f \) is high an increase in the price of \( Z_1 \) feeds back into the price of \( Z_2 \). The parameter \( b \) works similarly, but for \( Y \) prices. Note that when \( e = 1 \), so the \( Y \) and \( Z \) markets are the same "size", the condition is satisfied. As \( e \to \infty \), (i.e. as the \( Z \) market becomes tiny) the condition is violated.\(^{35}\) This contrasts with the linear demand findings in previous sections, in which the licensing contract was profitable for any market size.

3.7 Discussion and Conclusion

The model presented in this chapter provides a rationale for horizontal restraints in licensing contracts which is based on the licensor's desire to affect the competitive pricing of substitute goods. Antitrust policy toward IP licensing seems largely to be following a trend of increasing lenience, inasmuch as authorities are willing to consider a variety of efficiency defenses for contractual practices. Policy regarding horizontal contractual restrictions appears to be following this trend, at least for cases that do not have an explicit IP dimension.\(^{36}\) This chapter suggests that with regard to multiproduct firms, the suspicion with which price-fixing restraints in particular have historically been greeted

\(^{35}\)This follows because \( 2 - b > 0 \) in order to satisfy the second order conditions.

is largely justified. More generally, for the reasons noted in section 2, in a case involving the types of licensing practices discussed here both the U.S. Guidelines and the Canadian IPEGs would require an explicit analysis of competitive effects. The results presented in this chapter represent a step toward a coherent analytical treatment of such contracts.

A contract in which the licensor sets the final price of the patented good increases total profits because it induces the licensee to internalize an externality which otherwise would go unnoticed. This is the elasticity incentive for licensing. The simulations presented above reveal that the elasticity incentive alone is generally enough to render these agreements anticompetitive. For nonlinear demands, there is an additional effect arising from the curvature of demand functions which may further limit competition.

An alternative contract in which the licensor offers an exclusive territory for the sale of the patented good can achieve a similar result. One difference is that the high price for Z in the RPM contract cannot be supported in the ET contract, where all prices are determined simultaneously. In this model, unlike some other treatments of contractual restraints (e.g. Mathewson and Winter, 1984; Tirole, 1988; Winter, 1993), RPM and ET are not close substitutes. This result arises because of the timing associated with the various contracts: Bertrand competition which takes place with a fixed price for one good (here, Z) in the background is quite different from competition in which all prices are choice variables. Rey and Stiglitz (1995) note the general point that by choosing the timing of pricing decisions appropriately, cooperative behaviour can be facilitated. In the context of this chapter, a firm with a patented product can more easily extend its monopoly power into a related market when the price is initially fixed than otherwise.

Although the model I have presented does not explicitly incorporate an R&D stage, it is straightforward to introduce one. Because horizontal restraints tend to increase the

---

37 This analysis also makes it clear that there is no need in the multiproduct context to consider coordinated or conspiracy-type effects in the event that there are multiple licensees. In U.S. v. Masonite Corp., 316 U.S. 265 (1942), the U.S. Supreme Court held that when each licensee becomes aware of the price-fixing restrictions on each other, a conspiracy can be found. Part of Masonite's defense was its insistence that it had licensed each firm independently.
profit available to the winner of the patent race for $Z$, R&D expenditures will be higher when such agreements are allowed. Suppose the R&D stage is modeled as a game in which identical firms 1 and 2 choose their research intensity, $x_i$. This in turn affects the probability of developing product $Z$ according to the increasing and concave function $p(x_i)$. It can then easily be shown that research intensity is increasing in *ex post* profits. so that if, for example, an RPM contract is legal, the probability that at least one firm develops product $Z$ is higher than if such a contract were prohibited. From this *ex ante* perspective, the issue of the legality of horizontal restraints therefore depends on the difficult tradeoff between consumers' expected additional surplus from the introduction of $Z$, and the social harm caused by the licensing contract. Even in this simple setup it is apparent that a sensible antitrust policy would have to take into account the elasticity of the research production function $p(x_i)$, which might empirically be very difficult.

One further extension to the model in this chapter involves a licensor that does not actually produce the patented product, but nonetheless fixes the price of the good to influence competitive pricing in other markets. This extension of the model would help explain observed contractual pricing provisions that would seem to be otherwise privately undesirable. For example, in *Royal Industries v. St. Regis Paper*, the plaintiff alleged that St. Regis, the licensee for patented plastic ties, violated the terms of an oral agreement to set the price at a minimum level. Although direct evidence is unavailable, the plaintiff and the defendant may have been competitors in other markets prior to the license. St. Regis did not compete directly with Royal in the plastic ties market, but the reasoning above explains why the Royal may have had an interest in fixing the price. *LucasArts* is a similar case involving the licensing of a copyright rather than a patent. LucasArts, a software firm, licensed a software utility called SCUMM, and set the price at which any software developed with SCUMM could be sold. The contract was upheld.

---

under the *GE* rule. In this case the licensor was attempting to influence the price of final products in which the protected subject matter was used, rather than seeking to obtain a reasonable royalty on the use of the code itself. My model suggests one possible reason why the licensor might care about the price of the final product.
3.8 Appendix A: Proofs

Proof of Lemma 1: The constrained joint profit maximization problem is given by

$$\max_{p_1, p_2, p_Z} \pi^J = (p_1 - c^Y) D^1(p_1, p_2, p_Z) + (p_2 - c^Y) D^2(p_1, p_2, p_Z) + (p_Z - c^Z) D^Z(p_Z, p_1, p_2).$$

subject to the constraint that the prices of $Y$ must be determined by Bertrand competition:

$$p_1 = \arg \max_{p_1} (\bar{p}_1 - c^Y) D^1(\bar{p}_1, p_2, p_Z) + (1 - \lambda)(p_Z - c^Z) D^Z(p_Z, \bar{p}_1, p_2).$$

$$p_2 = \arg \max_{p_2} (\bar{p}_2 - c^Y) D^2(p_1, \bar{p}_2, p_Z) + \lambda(p_Z - c^Z) D^Z(p_Z, p_1, \bar{p}_2).$$

The first-order conditions arising from the constraints are

$$D^1 + (p_1 - c^Y) D^1_1 + (1 - \lambda)(p_Z - c^Z) D^Z_1 = 0. \quad (3.22)$$

$$D^2 + (p_2 - c^Y) D^2_2 + \lambda(p_Z - c^Z) D^Z_2 = 0.$$

These implicitly define reaction functions, which will be denoted $R_1(p_2, \lambda)$ and $R_2(p_1, \lambda)$. Note that the best-response functions hold $p_Z$ and $\lambda$ constant. The corresponding second-order conditions are

$$SOC_1 = 2D^1_1 + (p_1 - c^Y) D^1_{11} + (1 - \lambda)(p_Z - c^Z) D^Z_{11} < 0.$$

$$SOC_2 = 2D^2_2 + (p_2 - c^Y) D^2_{22} + \lambda(p_Z - c^Z) D^Z_{22} < 0.$$

These conditions ensure a unique $(p_1^*, p_2^*)$. Also, by virtue of the strategic complementarity between $p_1$ and $p_2$, denote

$$G_1 = D^1_2 + (p_1 - c^Y) D^1_{12} + (1 - \lambda)(p_Z - c^Z) D^Z_{12} \geq 0.$$

$$G_2 = D^2_1 + (p_2 - c^Y) D^2_{21} + \lambda(p_Z - c^Z) D^Z_{21} \geq 0.$$

Finally, the Bertrand pricing equilibrium is stable if

$$\left| \frac{\partial R_i(p_j, \lambda)}{\partial p_j} \right| < 1.$$
This implies that

$$\left| \frac{G_i}{SOC_i} \right| < 1.$$  \hspace{1cm} (3.23)

for \( i = 1, 2 \).

Totally differentiating the first-order conditions,

$$\left[ SOC_1 + G_1 \frac{\partial R_2}{\partial p_1} \right] dp_1 + \left[ -(p_Z - c^Z)D_1^Z + G_1 \frac{\partial R_2}{\partial \lambda} \right] d\lambda = 0.$$  \hspace{1cm} (3.24)

A symmetric condition holds for the other first-order condition. By the implicit function theorem,

$$\frac{\partial R_2}{\partial p_1} = \frac{-G_2}{SOC_2},$$

and

$$\frac{\partial R_2}{\partial \lambda} = \frac{-(p_Z - c^Z)D_2^Z}{SOC_2}.$$  

Inserting these into (3.24) gives

$$\frac{\partial p_1}{\partial \lambda} = \frac{(p_Z - c^Z) \left[ SOC_2 D_1^Z + G_1 D_2^Z \right]}{SOC_1 SOC_2 - G_1 G_2}.$$  \hspace{1cm} (3.25)

A similar set of substitutions gives

$$\frac{\partial p_2}{\partial \lambda} = \frac{-(p_Z - c^Z) \left[ SOC_1 D_2^Z + G_2 D_1^Z \right]}{SOC_1 SOC_2 - G_1 G_2}.$$  \hspace{1cm} (3.26)

Note that by using the two stability conditions from (3.23), we know that the denominators must be positive. Under reasonable conditions, demand functions will be sufficiently concave such that the terms in the numerator contained within the square brackets will be negative. (For example, with linear demand, the mild restriction \( SOC_1 D_2^Z - G_2 D_1^Z < 0 \) works out to be equivalent to the equilibrium stability condition shown above in (3.23).)

Assuming that this restriction holds, an increase in \( \lambda \) lowers \( p_1 \) and raises \( p_2 \), holding all other variables constant.

Let \( \bar{\pi}'(p_1', p_2', p_Z) \) be the maximized joint profit evaluated at the Bertrand maximizers for \( p_1 \) and \( p_2 \), and the \( p_Z \) that solves the first-stage problem. Then by the envelope theorem, to find the effect of a change in \( \lambda \) on \( \pi^* \), we need to find

$$\frac{d\bar{\pi}'}{d\lambda} = \frac{\partial \bar{\pi}'}{\partial p_1} \frac{\partial p_1'}{\partial \lambda} + \frac{\partial \bar{\pi}'}{\partial p_2} \frac{\partial p_2'}{\partial \lambda}.$$  \hspace{1cm} (3.27)
The effects of $\lambda$ on prices are known from (3.25) and (3.26). It is straightforward to show that

$$A \equiv \frac{\partial \pi}{\partial p_1} = D^1 + (p_1 - c^1)D_1^1 + (p_2 - c^1)D_1^2 + (p_2 - c^2)D_2^2,$$

$$B \equiv \frac{\partial \pi}{\partial p_2} = D^2 + (p_2 - c^1)D_2^1 + (p_1 - c^1)D_2^1 + (p_2 - c^2)D_2^2.$$

Putting these together and simplifying yields

$$\frac{d\pi}{d\lambda} = \frac{(p_2 - c^2)}{SOC_1SOC_2 - G_1G_2} \{A [SOC_3D_1^2 + G_1D_2^2] - B [SOC_1D_2^2 + G_2D_2^2]\}.$$

The term outside the brackets is positive for any $p_2 > c^2$. Then

$$\left.\frac{d\pi}{d\lambda}\right|_{\lambda=0} > 0$$

if and only if

$$\frac{A}{B} < \frac{SOC_1D_2^2 + G_2D_2^2}{SOC_3D_2^1 + G_1D_2^2}. \quad (3.28)$$

Note that the inequality is reversed since the bracketed terms are both negative. Also note that the RHS can readily be shown to be equal to 1 in the case of linear symmetric demands.

The LHS of (3.28) can be translated into elasticity terms, and then simplified using the Bertrand conditions from (3.22). This gives

$$LHS = \frac{D^1}{D^2} \cdot \frac{(p_2 - c^1)\epsilon_{12}/p_1D^2 + \lambda(p_2 - c^2)\epsilon_{21}/p_1D^1}{(p_1 - c^1)\epsilon_{21}/p_2D^1 + (1 - \lambda)(p_2 - c^2)\epsilon_{22}/p_2D^2}. \quad (3.29)$$

Recall the curvature effects, as given above in the text. Then the RHS can be expressed in elasticity terms as

$$RHS = \frac{D^1}{D^2} \cdot \frac{(-2\epsilon_{11}/p_1 + CE_{11}/D^1)\epsilon_{22}/p_2 + D^2/(p_1D^1)(\epsilon_{21}/p_1 + CE_{21}/D^2)\epsilon_{21}}{(-2\epsilon_{22}/p_2 + CE_{22}/D^2)\epsilon_{21}/p_1 + D^1/(p_2D^2)(\epsilon_{12}/p_2 + CE_{12}/D^1)\epsilon_{22}}. \quad (3.30)$$

Rewriting (3.28), in order for an increase in $\lambda$ to be privately profitable at $\lambda = 0$,

$$\frac{(p_2 - c^1)\epsilon_{12}/p_1D^2}{(p_1 - c^1)\epsilon_{21}/p_2D^1 + (p_2 - c^2)\epsilon_{22}/p_2D^1} < \frac{(-2\epsilon_{11}/p_1 + CE_{11}/D^1)\epsilon_{22}/p_2 + D^2/(p_1D^1)(\epsilon_{21}/p_1 + CE_{21}/D^2)\epsilon_{21}}{(-2\epsilon_{22}/p_2 + CE_{22}/D^2)\epsilon_{21}/p_1 + D^1/(p_2D^2)(\epsilon_{12}/p_2 + CE_{12}/D^1)\epsilon_{22}}.$$
with all prices and demands evaluated at $\lambda = 0$.

**Proof of Proposition 8:** Consider the linear symmetric demand setup given by system (3.11). Bertrand equilibrium prices (given $p_Z$) are

\[
\begin{align*}
p_1 &= \frac{(2d + b)(a + c') + (2d + b)cp_Z + [b\lambda + 2d(1 - \lambda)]c(p_Z - c^2)}{4d - b^2}, \\
p_2 &= \frac{(2d + b)(a + c') + (2d + b)cp_Z + [b(1 - \lambda) + 2d\lambda]c(p_Z - c^2)}{4d - b^2}.
\end{align*}
\]

Differentiating with respect to $\lambda$,

\[
\frac{\partial p_1}{\partial \lambda} = \frac{-c(p_Z - c^2)}{2d + b}, \\
\frac{\partial p_2}{\partial \lambda} = \frac{c(p_Z - c^2)}{2d + b}.
\]

Note also that

\[
\begin{align*}
\frac{\partial \pi}{\partial p_1} &= a - 2dp_1 + dp_2 + cp_Z + c'H(p_2 - c') + c(p_Z - c^2), \\
\frac{\partial \pi}{\partial p_2} &= a - 2dp_2 + dp_1 + cp_Z + c'H(p_1 - c') + c(p_Z - c^2).
\end{align*}
\]

Using the previous four expressions in (3.27), $\frac{\partial \pi}{\partial \lambda} > 0$ if and only if

\[b(p_1 - p_2) + (1 - 2\lambda)c(p_Z - c^2) > 0.\]

which is equivalent to

\[
\left(\frac{b}{2d + b} + 1\right)(1 - 2\lambda)c(p_Z - c^2) > 0.
\]

which holds when $(1 - 2\lambda) > 0$. Therefore joint profits attain a global maximum at $\lambda^* = 1/2$.

**Proof of Proposition 9:** Assume $\lambda = 1/2$, and consider $\frac{dp_1}{dp_Z}$ (results for $p_2$ are analogous). If $p_i$ and $p_Z$ are strategic complements, then

\[H_i = D_2 + (p_i - c')D_{1Z} + \frac{1}{2}D_i^2 + \frac{1}{2}(p_Z - c^2)D_{1Z}^2 \geq 0.\]

for $i = 1, 2$. Totally differentiating firm 1's first-order condition from (3.22),

\[
\left[SOC_1 + G_1 \frac{\partial R_2}{\partial p_1}\right] dp_1 + \left[H_1 + G_1 \frac{\partial R_2}{\partial p_Z}\right] dp_Z = 0. \tag{3.31}
\]
Using the implicit function theorem on firm 2’s first-order condition, we know that
\[ \frac{\partial R_2}{\partial p_Z} = \frac{H_2}{SOC_2}. \]
Simplifying (3.31),
\[ \frac{\partial p_1}{\partial p_Z} = -\frac{[H_1 SOC_2 - G_1 H_2]}{SOC_1 SOC_2 - G_1 G_2}. \]
The proof of Lemma 1 establishes that the denominator of this expression is positive.
Noting that \( SOC_2 \) in the numerator is negative, while the other terms are positive, the whole expression must be positive. In particular, when \( \lambda = 1/2 \), \( SOC_1 = SOC_2 \), \( G_1 = G_2 \), and \( H_1 = H_2 \), all by symmetry, and the derivative simplifies to
\[ \frac{\partial p_1}{\partial p_Z} = -\frac{H_1}{SOC_1 + G_1} \geq 0. \]
Expressing this derivative in elasticity form gives (3.13) in the text. ■

Proof of Proposition 10: Suppose first that \( p_Z \) is the same in both the ET and RPM contexts, and compare the incentives for setting \( p_1 \) and \( p_2 \), when \( \lambda = 1/2 \). The first-order conditions for both the ET and RPM contracts are
\[
\begin{align*}
D^1 + (p_1 - c^c)D^1_1 + & \frac{1}{2}(p_Z - c^c)D^1_2 = 0, \\
D^2 + (p_1 - c^c)D^2_1 + & \frac{1}{2}(p_Z - c^c)D^2_2 = 0.
\end{align*}
\]
The difference is that \( p_Z \) is a parameter under RPM, while it is determined jointly with the other prices under ET. Since we know that \( p_1 \) is increasing in \( p_Z \) from above, if it can be shown that \( p_Z \) is higher under RPM then all prices will be higher. Under ET, the first-order conditions for setting \( p_Z \) are
\[ D^2 + (p_Z - c^c)D^2_2 + (p_1 - c^c)D^2_1 = 0. \]
for both firms 1 and 2 (since \( p_1 = p_2 \)). Compare this with the analogous condition determining \( p_Z \) under RPM:
\[ D^2 + (p_Z - c^c)D^2_2 + 2(p_1 - c^c)D^1_2 + A \frac{dp_1}{dp_Z} + B \frac{dp_1}{dp_Z} = 0. \]
where $A$ and $B$ are defined in the proof to Lemma 1 above. Because the last two terms are positive (and the maximization problem is concave), it follows that $\tilde{p}_Z \geq p'_Z$. Finally, since equilibrium prices are increasing in $\tilde{p}_Z$ (by strategic complementarity), $p'_i \geq p'_Z$ for $i = 1, 2$. ■

Proof of Proposition 11: Let $p' = (p'_1, p'_2)$. Note that because $p_Z$ is freely chosen by firm 1 in the RPM case, by revealed preference it must be the case that $\pi'_J(p', \tilde{p}_Z) \geq \pi'_J(p^*, p'_Z)$, as in particular firm 1 could, if it wanted to, choose $p'_Z$ and replicate the ET solution. ■

Proof of Proposition 12: The first-order conditions for the Bertrand equilibrium are:

$$
\begin{align*}
&y^1 + (p_1 - c^Y)y^1_1 + (q_1 - c^Z)z^1_1 + r z^2_1 = 0. \\
&(p_1 - c^Y)y^1_{1r} + z^1_{1r} + (q_1 - c^Z)z^1_{1r} + r z^2_{1r} = 0. \\
&y^2 + (p_2 - c^Y)y^2_2 + (q_2 - c^Z - r)z^2_2 = 0. \\
&(p_2 - c^Y)y^2_{2r} + z^2_{2r} + (q_2 - c^Z - r)z^2_{2r} = 0.
\end{align*}
$$

These conditions define the reaction functions, which will be denoted $R^1$, $R^1'$, $R^2$ and $R^2'$ respectively, each of which is a function of the remaining prices as well as $r$. To find the effect of $r$ on $p_2$, totally differentiate the third condition:

$$
\left[ \begin{array}{c}
\pi^2_{22} + \pi^2_{21} \frac{\partial R^1}{\partial p_2} + \pi^2_{21} \frac{\partial R^1'}{\partial p_2} + \pi^2_{22} \frac{\partial R^2}{\partial p_2} \\
-\pi^2_{21} \frac{\partial R^1}{\partial r} + \pi^2_{21} \frac{\partial R^1'}{\partial r} + \pi^2_{22} \frac{\partial R^2}{\partial r}
\end{array} \right] dp_2 + \left[ \begin{array}{c}
-\pi^2_{21} \frac{\partial R^1}{\partial r} + \pi^2_{21} \frac{\partial R^1'}{\partial r} + \pi^2_{22} \frac{\partial R^2}{\partial r}
\end{array} \right] dr = 0.
$$

Uniqueness of the Bertrand equilibrium (see Vives, 1999) implies that the bracketed term on the first line of (3.32) must be negative. Therefore, $\frac{dp_2}{dr} \geq 0$ if the bracketed term on
the second line is nonnegative. Noting that

\[ \frac{\partial R^1}{\partial r} = \frac{-z_1^2}{\pi_{11}}, \quad \frac{\partial R^r}{\partial r} = \frac{-z_1^2}{\pi_{11}}, \quad \frac{\partial R^y}{\partial r} = \frac{z_2^2}{\pi_{22}}. \]

all of which are positive and follow from the implicit function theorem. Substitution yields equation (3.20) in the text. ■

---

40 A more straightforward method of proving the proposition, which does not depend on stability and uniqueness, follows from the supermodularity of the price-setting game. In this case the proposition follows directly if the bracketed term on the second line of (3.32) is positive, based on the properties of supermodular games. If the equilibrium is not unique, then the result applies to the highest and lowest equilibria (Vives, 1999).
3.9 Appendix B: Characterizing the Exclusive Territories Contract

As shown in section 5, the maximization problems facing the two firms are

\[ \max \pi_1 = (p_1 - c_1^Y) \left[ (1 - \lambda)D_1^1(p_1, p_2; p_1^1) + \lambda D_1(p_1, p_2; p_1^2) \right] + (1 - \lambda)(p_1^1 - c_1^Z)D_1^2(p_1^1; p_1, p_2) \]

for firm 1 and

\[ \max \pi_2 = (p_2 - c_1^Y) \left[ (1 - \lambda)D_2^1(p_1, p_2; p_2^1) + \lambda D_2(p_1, p_2; p_2^2) \right] + \lambda(p_2^1 - c_2^Z)D_2^2(p_2^1; p_1, p_2) \]

for firm 2. Denote by \( D_i \) the demand for product \( i \) from consumers residing in firm 2’s exclusive territory. Then the first-order conditions for the choice of \( p_1 \) and \( p_2 \) for firm 1 are

\[ (1 - \lambda)D_1 + \lambda \hat{D}_1 + (p_1 - c_1^Y) \left[ (1 - \lambda)D_1^1 + \lambda \hat{D}_1^1 \right] + (1 - \lambda)(p_1^1 - c_1^Z)D_1^2 = 0 \]  
(3.33)

and

\[ D_2^2 + (p_2^1 - c_2^Z)D_2^2 + (p_1 - c_1^Y)D_2^1 = 0 \]  
(3.34)

(after dividing by \( (1 - \lambda) \)). The analogous first-order conditions for firm 2 are

\[ (1 - \lambda)D_2 + \lambda \hat{D}_2 + (p_2 - c_1^Y) \left[ (1 - \lambda)D_2^2 + \lambda \hat{D}_2^2 \right] + \lambda(p_2^1 - c_2^Z)D_2^1 = 0 \]  
(3.35)

and

\[ D_2^1 + (p_2^1 - c_2^Z)D_2^1 + (p_2 - c_1^Y)D_2^2 = 0. \]  
(3.36)

To find the private incentive to raise \( \lambda \) above zero, we need to know whether

\[ \left. \frac{d\bar{\pi}^J}{d\lambda} \right|_{\lambda=0} > 0 \]

is satisfied, where \( \bar{\pi}^J \) represents joint profit. Differentiating the sum of profits and using the Bertrand first-order conditions, it can be shown that

\[ \left. \frac{d\bar{\pi}^J}{d\lambda} \right|_{\lambda=0} = \left. \frac{\partial \bar{\pi}^J}{\partial \lambda} \right|_{\lambda=0} + (p_2 - c_1^Y)D_1^1 \left. \frac{\partial p_1}{\partial \lambda} \right|_{\lambda=0} + (p_2 - c_1^Y)D_2^2 \left. \frac{\partial p_1}{\partial \lambda} \right|_{\lambda=0} \]

\[ + \left. [(p_1 - c_1^Y)D_1^2 + (p_1^1 - c_1^Z)D_2^1] \frac{\partial p_2}{\partial \lambda} \right|_{\lambda=0}. \]  
(3.37)
A sufficient condition for private profitability is that this expression is positive.

First note that

\[
\frac{\partial \pi^I}{\partial \lambda} = (p_1 - c^Y) \left( \tilde{D}^1 - D^1 \right) + (p_2 - c^Y) \left( \tilde{D}^2 - D^2 \right) - (p_2^Z - c^Z) D^Z + (p_2^Z - c^Z) \tilde{D}^Z.
\]

The sign of this expression is fundamentally ambiguous: at \( \lambda = 0 \), \( p_2^Z \geq p_2^Z \) since prices are strategic complements and \( p_2 < p_1 \). Therefore \( \tilde{D}^i \leq D^i \), so the first three expressions will be negative, while the last term is positive. When demand is linear and the Z market is small, \( D^Z \approx \tilde{D}^Z \), so the whole expression is negative.

Conditions (3.33) and (3.35) define reaction functions for \( p_1 \) and \( p_2 \) choices which are functions of the other prices and \( \lambda \): denote these by \( R^1 \) and \( R^2 \). Differentiate these reaction functions with respect to \( \lambda \):

\[
\frac{\partial R^1}{\partial \lambda} = - \frac{(\tilde{D}^1 - D^1) + (p_1 - c^Y) \left[ \tilde{D}^1 - D^1 \right] - (p_2^Z - c^Z) D^Z}{SOC_1},
\]

\[
\frac{\partial R^2}{\partial \lambda} = - \frac{(\tilde{D}^2 - D^2) + (p_2 - c^Y) \left[ \tilde{D}^2 - D^2 \right] + (p_2^Z - c^Z) D^Z}{SOC_2}
\]

where \( SOC_1 \) is the (negative) second-order condition with respect to \( Y_i \). With linear demand, for example, the first expression is negative and the second is positive as long as \( c^Z \) is small.

Differentiating first-order condition (3.33) gives

\[
\frac{\partial p_1}{\partial \lambda} = - \frac{\left[ (\tilde{D}^1 - D^1) + (p_1 - c^Y) \left[ \tilde{D}^1 - D^1 \right] - (p_2^Z - c^Z) D^Z + \pi_{12}^R \frac{\partial R^1}{\partial \lambda} \right]}{J_1}
\]

where \( J_1 < 0 \) is the Jacobian of (3.33) and \( \pi_{12}^R \geq 0 \) is the cross-partial with respect to \( p_1 \) and \( p_2 \). This is negative for linear demand for the same reason that \( \frac{\partial R^1}{\partial \lambda} \) is negative.

Differentiating (3.35),

\[
\frac{\partial p_2}{\partial \lambda} = - \frac{\left[ (\tilde{D}^2 - D^2) + (p_2 - c^Y) \left[ \tilde{D}^2 - D^2 \right] + (p_2^Z - c^Z) D^Z + \pi_{21}^R \frac{\partial R^1}{\partial \lambda} \right]}{J_2}
\]

This expression will be positive provided again that \( c^Z \) is not high and that goods \( Y \)
and \( Y2 \) are not too closely substitutable. Finally, differentiating (3.34),
\[
\frac{\partial p_2^1}{\partial \lambda} = -\left[ \pi_1 \frac{\partial p_1^1}{\partial \lambda} + \pi_2 \frac{\partial p_2^2}{\partial \lambda} \right] / J_1,
\]
where \( 1' \) represents a derivative with respect to \( p_2^1 \). This derivative is negative for linear demand: goods \( Y1 \) and \( Z1 \) are stronger strategic complements than are \( Y2 \) and \( Z1 \).

Taking all of these expressions together, and assuming that \( \frac{\partial p_1^1}{\partial \lambda} \leq 0 \), it is privately profitable to raise \( \lambda \) if (looking at (3.37)) the margin on \( p_2 \) is relatively small, while the margins on \( p_1 \) and \( p_2 \) are initially high. Thus, rearranging, a necessary condition for private profitability is
\[
\frac{\partial p_2}{\partial \lambda} \geq \frac{(p_2 - c^1)D_1^2}{[(p_1 - c^1)D_1^1 + (p_2 - c^2)D_2^2]} \left( -\frac{\partial p_1}{\partial \lambda} \right) + \frac{(p_2 - c^1)D_2^2}{[(p_1 - c^1)D_1^1 + (p_2 - c^2)D_2^2]} \left( -\frac{\partial p_2}{\partial \lambda} \right).
\]
Each of these terms (evaluated at \( \lambda = 0 \)) are positive under the natural conditions discussed above.

In the simulations, consumer surplus increased with the imposition of the ET restraint. As a step toward more general welfare analysis, the necessary condition for private profitability can be compared with a sufficient condition for some measure of welfare to improve (Winter, 1997). Here it is easier to work with consumer surplus, \( S \): of course a positive change in \( S \) implies a welfare increase, assuming private profitability is satisfied. It can be shown that a sufficient condition for \( S \) to increase with \( \lambda \) is
\[
\frac{\partial p_2}{\partial \lambda} \leq \frac{D_1^1}{D_2^2} \left( -\frac{\partial p_1}{\partial \lambda} \right) + \frac{D_2^2}{D_2^2} \left( -\frac{\partial p_2}{\partial \lambda} \right).
\]
Combining these conditions, for welfare to increase it is sufficient that
\[
\frac{D_1^1}{D_2^2} \geq \frac{(p_2 - c^1)D_1^2}{[(p_1 - c^1)D_1^1 + (p_2 - c^2)D_2^2]}
\]
and
\[
\frac{D_2^2}{D_2^2} \geq \frac{(p_2 - c^1)D_2^2}{[(p_1 - c^1)D_1^1 + (p_2 - c^2)D_2^2]}.
\]
When these are satisfied, \( p_2 \) increases enough for the arrangement to be privately profitable, but not so much that consumer surplus is harmed.
Chapter 4


4.1 Introduction

Given the importance of the innovation and diffusion of new technologies in stimulating economic growth, the extensive recent development of the literature on patent policy is not surprising. A large part of this literature is theoretical in nature, and attempts to provide results on optimal patent design to guide policymakers. But a number of important changes in national patent systems over the past two decades have also provided excellent opportunities for empirical analyses of the effects of patent policy in practice. Such studies test and refine our theoretical understanding of the way patent policy works, and in addition point out other areas of potential concern.

In this chapter I exploit one such policy change: the major reform of Canada's *Patent Act* which went into effect in October of 1989. This reform consisted of two major
elements among a list of less important changes. First, the "priority rule" was changed. The priority rule provides a way of resolving disputes among a number of inventors over who is entitled to a patent for a given invention. Under the "first-to-invent" rule, an inventor needs to prove that she discovered the invention earlier than the others (regardless of whether a patent application has been filed); under the "first-to-file" rule, the first inventor to apply for a patent is deemed to have priority. The 1989 reforms included a transition from the first-to-invent to the first-to-file rule. Second, the rules governing application disclosure were altered. Before 1989, patent applications were kept secret prior to the grant date, while after the reform applications were published 18 months after filing.

It is important to gain a fuller understanding of both of these aspects of patent policy, as the implications of these changes extend well beyond Canada. For instance, the U.S. continues to employ the first-to-invent priority rule, although it is now among the last industrialized countries in the world to do so. Since the Uruguay Round of the General Agreement on Tariffs and Trade, the U.S. has been under pressure to change this provision to better conform with international practices (Jaffe, 2000). Partly as a response to the trend toward international harmonization, the U.S. has also recently implemented an application disclosure policy.

I derive a set of hypotheses based on existing results in the relatively mature theoretical literature. These hypotheses are then tested using a unique dataset consisting of a sample of Canadian manufacturing firms. I find evidence of a modest increase in R&D spending following the reform, which is consistent with the priority theory. In addition, there appears to be an increase in the patenting intensity of these firms, holding R&D constant, which also accords well with the theory. On the other hand, I find little evidence

---

1 The precise content of the package of policy reforms is discussed in more detail in the following section.

2 The relevant legislation is the American Inventors Protection Act, enacted in 1999. See Johnson and Popp (2001) and the discussion in section 2 below.
to support a hypothesis derived from a model of disclosure policy. Together, these results imply that the package of reforms implemented in Canada had an observable effect on firms' behaviour. They also suggest that while our theoretical understanding of priority policy may be reasonably sound, further theoretical and empirical work is needed on the impact of application disclosure.

I thus contribute to the growing empirical literature on the effects of various aspects of patent policy. Much of this literature was stimulated by the perceived strengthening of patent protection in the U.S. following the formation of the Court of Appeals for the Federal Circuit in 1982 and the important court decisions thereafter (see Jaffe, 2000 for a general overview). Kortum and Lerner (1998) investigate whether the dramatic surge in patenting in the U.S. over the late 1980s and 1990s was due to the more favourable legal environment for patents provided by the new court. They use primarily aggregate data on patenting in several different countries and find that the increase in patenting is not due to the strengthening of patent rights. This conclusion is based largely on the finding that the U.S. appears to be a more important source of patented inventions, but not necessarily a more important destination for foreign applications. Kortum and Lerner conclude that the patent surge can be attributed to an increase in the productivity of R&D which was independent of the change in patent policy.

Hall and Ziedonis (2001) address the same fundamental question as Kortum and Lerner, but from a different angle. They use firm-level patenting data from the semiconductor industry, in which patent intensity has grown rapidly in spite of the fact that managers in this industry claim patents are not an effective way of protecting their R&D investments. Hall and Ziedonis find evidence that the stronger patent regime did have an impact on patenting in this industry. Since the formation of the CAFC, capital-intensive manufacturers have patented particularly intensively in order to use patents as currency when negotiating cross-licenses, thus avoiding costly infringement litigation. This patent

3Survey evidence collected by Hall and Ziedonis shows that managers use patent rights as bargaining
accumulation is largely defensive, and does not represent an increase in innovation.

High-tech firms are also the focus of Bessen and Maskin (2000). Using a simple model, they show that for technological areas in which firms' R&D efforts are complementary (in that more firms researching implies a greater likelihood of a sequence of future innovations), the introduction of patent protection may decrease R&D. Bessen and Maskin suggest that the software industry corresponds to their assumptions: today's breakthroughs stimulate future innovations, and the more firms there are performing R&D, the more future breakthroughs there will be. They provide evidence that the availability of patent protection for software in the U.S. beginning in the 1980s did not cause an increase in software R&D, and may even have led to a decrease.

Possible consequences of adopting an application disclosure regime in the U.S. are explored by Johnson and Popp (2001). They use citation analysis on all U.S. domestic patents from 1976 to 1996 and show that more “important” patents (those with a large number of subsequent citations) tend to have longer lags between application and grant dates. Therefore these will be affected by a disclosure policy that mandates publication 18 months after the application date. Johnson and Popp also present evidence that the rate of knowledge diffusion will be positively affected by disclosure. Simulation results suggest that this could translate into a sharp increase in patented innovations over the short to medium term.

Changes to patent policy outside the U.S. have also received some empirical attention. Sakakibara and Branstetter (2001) test the hypothesis that broader patent protection introduced in Japan in 1988 stimulated additional innovation, as economic theory suggests it should. Using careful econometric techniques on a large panel sample of firms, they find no evidence of additional innovative effort or output. Scherer and Weisburst (1995) test a similar hypothesis on a more limited scale. They examine whether the introduc-
tion of product patent protection for pharmaceuticals in Italy in 1978 had an effect on innovation by domestic producers. Analysis of the aggregate R&D and patent data indicates that product patent introduction had no discernible impact on drug R&D or new products, although Italian firms were more inclined to file patent applications in the U.S. afterwards.

In the next section I discuss the reform of Canada's Patent Act in more detail, and identify three principal hypotheses to be tested. Section 3 provides a brief look at aggregate Canadian data on patenting and R&D. In section 4 I discuss the firm-level dataset and present a set of results from hypothesis tests relating to R&D spending. This involves a consideration of some statistical issues that arise due to the nature of the estimation. In section 5 I turn to the effects of the policy change on the patenting behaviour of the sample firms. Section 6 concludes and speculates on interesting avenues for future research.

4.2 The 1989 Policy Changes and Theoretical Predictions

A number of significant changes to Canada's Patent Act went into effect on October 1, 1989. As summarized by Rafiquzzaman (1999), the most significant included: (i) a change in the priority rule from first-to-invent to first-to-file; (ii) a change in the disclosure rule, so applications are published (or "laid open") 18 months after the application priority date; (iii) a change in the maximum patent duration from 17 years after the patent grant to 20 years after the application priority date; and (iv) a move from automatic examination of patent applications to examination by request only.¹

¹The reforms also included other elements which appear (arguably) less important. The novelty requirement was modified so that where previously applications for inventions filed within two years of their first public appearance were acceptable, the new Act requires a maximum one year interval. Additionally, the language for claims was altered to require a more precise specification of the invention to be protected. See Rafiquzzaman (1999).
This complex package of reforms is interpreted by Rafiquzzaman (1999) as enhancing the “scope” of patent protection in Canada, and presumably thereby increasing the value of holding a Canadian patent. While the theoretical literature on optimal patent design has offered a variety of interpretations of patent scope, it is not straightforwardly apparent that the package of Canadian reforms described above corresponds directly with any of them. Fortunately, however, the theoretical literature does provide some guidance relating to specific aspects of the reforms.

Consider first the change in the priority rule. Scotchmer and Green (1990) provide a framework for interpreting this policy change that yields some clear predictions. They consider a model of sequential innovation in which two firms race to develop a technology that requires two separate R&D successes. These can be considered first- and second-generation advances over a base technology. There are three stages in their dynamic game. First, firms decide whether to begin the R&D race or not. Next, the innovator that successfully discovers the first-generation technology (the “leader”) decides whether or not to patent. Last, if the initial innovation is kept secret, the “lagging” firm must decide whether or not to drop out of the race for the next-generation technology.

The priority rule is important in the event that the leader suppresses her discovery. If the lagging firm then independently duplicates the first-generation technology, there will be a dispute over its ownership. Under the first-to-invent rule, the original innovator is able to establish priority, while under first-to-file the lagging firm may be able to assert a claim over the technology if it is patented immediately.

5The scope, or breadth of a patent indexes the extent to which other (competing) technologies must be different from the patented technology in order to avoid infringement. Gilbert and Shapiro (1990) and Klemperer (1990) analyze the optimal mix of patent scope and length to offset research costs, and argue for narrow patents based on a deadweight loss criterion. Gallini (1992) interprets scope in terms of imitation costs, and shows that broader scope deters inefficient imitation. Matutes, Rogibeau, and Rockett (1996) view scope as an instrument for giving breakthrough innovators a head start in developing applications of their basic technology. Green and Scotchmer (1995) and Denicolo (2000) are concerned with innovation as a sequential process, in which broader scope rewards current inventors at the expense of future inventors.

6The parameters of the model are the cost of doing R&D for one period and the Poisson “hit rate” of innovative success.
Scotchmer and Green show that the first-to-file rule encourages patenting in a larger region of the parameter space than the first-to-invent rule, conditional on achieving an R&D success in the race for the first-generation technology. In other words, the "propensity to patent" is greater under first-to-file. The reason for this is quite intuitive. Patenting entails a cost and a benefit. The cost is that valuable technological information is revealed to the lagging firm upon the granting of the patent, so the two firms then begin racing for the second-generation technology on an equal footing. The benefit is that the first innovator asserts a claim over the intermediate innovation. The costs to patenting are the same under the two priority regimes, while the benefit is smaller under first-to-invent, since the leader retains the property right even if the lagging firm subsequently innovates.

The model also predicts that R&D spending is higher under first-to-file, perhaps even to a socially excessive degree. The reasons for this are also intuitive, although more subtle. It is convenient to classify the differential incentives to engage in R&D as \textit{ex ante} and \textit{ex post} incentives (relative to the discovery of the first-generation technology); these both operate in the same direction. The \textit{ex post} incentive concerns the lagging firm's decision to remain in the R&D race after an unpatented discovery by the leader. This incentive is weaker under first-to-invent, since even if the lagging firm overtakes the leader and discovers the second-generation technology, its profits will be dissipated by competition from the intermediate product (to which the leader retains the property right). Thus the lagging firm is "shaken out" to a greater degree under first-to-invent.

The \textit{ex ante} incentive concerns the decision to enter the R&D race in the first place. Recall that under a certain set of parameters a leading firm will patent under first-to-file but not under first-to-invent. Foreseeing this patenting behaviour at the outset, firms

---

7See Gallini, Putnam and Tepperman (2001) for a detailed discussion of the theoretical foundations and empirical literature concerning the propensity to patent.

8A key assumption of the model is that the successful first-generation innovator can send a credible signal that a discovery has been made without revealing the nature of the technology.
will thus be more inclined to enter the race under first-to-file, since even if they lose the first R&D race they will immediately start the second on even terms. This provides a greater expected return than first-to-invent, under which the discovery is kept secret and the second R&D race is highly asymmetric.

There are thus two empirically testable predictions arising from the analysis of Scotchmer and Green. These are denoted by Hypotheses 1 and 2 below:

**Hypothesis 1**  
*R&D spending is higher under the first-to-file priority rule than under first-to-invent.*

**Hypothesis 2**  
The propensity to patent (patenting conditional on R&D) is higher under the first-to-file priority rule than under first-to-invent.

Now consider the automatic disclosure of patent applications. Prior to Canada’s patent reform, Canada and the U.S. were the only G-7 countries not requiring disclosure after 18 months. Motivated by considerable public debate in the U.S. over the merits of pre-grant publication, Aoki and Spiegel (2000) offer a formal analysis of the effects of disclosure policy. They address the informal argument that small inventors are particularly vulnerable to exploitation by large multinationals if application disclosure is required, which has been a mainstay of those opposed to the reform. Aoki and Spiegel show that as this argument suggests, public disclosure of applications tends to weaken the overall incentive to innovate, and, more importantly, dulls the incentive for a successful initial innovator to expend resources on developing a basic technology. This occurs because disclosure of the application decreases the head start the initial innovator has over competitors in developing the technology.

---

9This is a different effect from those discussed above Hypotheses 1 and 2, and occurs for a different reason. In the Scotchmer and Green model a change to first-to-file induces both more R&D and more patenting in equilibrium; thus fewer inventions are kept secret and instead become public knowledge after the patent is granted. Conversely, in the Aoki and Spiegel framework, the increase in public disclosure is legislated exogenously, and while firms can still decide whether to patent or not, the action in the model is driven by the shorter time period between application and publication after the reform.
With respect to their model's predictions, Aoki and Spiegel emphasize the "small firm" interpretation (see also Rafiquzzaman, 1999): if an initial invention is discovered by a small firm with limited resources, public disclosure tends to benefit large corporations conducting follow-on development work, due to assumed economies of scale in the use of information generated by disclosure. Thus the model predicts that holding other factors constant, a change from a confidential filing system to a public disclosure system would lead to a differential R&D response by small firms and relatively large firms, with the former investing relatively less. The hypothesis to be tested is:

**Hypothesis 3**  
*The introduction of public disclosure of patent applications leads to a decrease in R&D by small firms relative to large firms.*

This model also generates a prediction on the propensity to patent that apparently conflicts with Hypothesis 2. Aoki and Spiegel show (in Proposition 4) that when patent protection is weak (in the sense that patents are unlikely to be enforceable in court), a successful innovator will patent under neither the confidential filing nor the public disclosure systems: when protection is strong she will patent under both systems; and when protection is in a specific intermediate range she will patent only under the confidential filing system. Thus, because of the information transmitted to the rival through a patent disclosure, the propensity to patent will (weakly) fall as a result of a move to the public disclosure system. And since the degree of patent protection differs by industry, propensity is predicted to strictly fall in some industries and to remain unchanged in others.

To what extent does this complicate the testing of Hypothesis 2? As I describe below, I rely on U.S. patent data to estimate the model of patenting. Because it is unlikely that the propensity to patent in the U.S. (which retained a confidential filing system) changed, Hypothesis 2 can still be tested cleanly. This prediction of the Aoki and Spiegel model is mentioned primarily for completeness, and as a hypothesis that could be tested using
Canadian patent data.

Further, while the hypotheses generated by the Aoki and Spiegel model might be reasonable for a very large country such as the U.S., it is arguable whether the Canadian case satisfies the implicit assumptions required. More specifically, for over a decade following the patent reform in Canada, Canadian firms had unrestricted access to the confidential U.S. system. The majority of patent applications by Canadian inventors are filed first in the U.S., so these inventors might be less sensitive to a change in disclosure policy at home. Thus, a failure to find evidence supporting the Aoki and Spiegel model in the data would be consistent with the idea that some aspects of Canada’s domestic patent policy are less relevant to Canadian inventors than the corresponding aspects of U.S. patent policy.

The theoretical literature has thus far not considered the remaining two principal provisions of the Canadian patent reform: the change to an application-based patent term and the examination-by-request requirement. Under the new patent term rules, it is possible to sue for infringement that occurs between the laying-open and the granting of a patent. Thus, there are a maximum of 20 years less 18 months of effective patent protection available, which exceeds the previously available 17 years. It is difficult to know a priori the class of inventors or industries that are most likely to be affected. Particularly since not all patents are renewed to their term limit; but there appears to be no reason to anticipate that this change would distort the observable implications of the other reforms.

---

10 The American Inventors Protection Act of November 1999 provided for a limited degree of disclosure: patent applications can only remain confidential at the inventor’s request and require certification that the invention will not also be the subject of a patent application abroad.

11 A search on the U.S. Patent and Trademark Office (USPTO) on-line patent database indicates that of the 3056 subsequently granted 1994 applications by Canadian inventors, only 404 were previously filed in Canada. This ratio of Canadian-priority applications to total successful U.S. applications also seems to be falling over the last decade.

12 An additional implication of this change is that “submarine patents” (old applications kept alive in the hopes of extracting future royalties) are no longer a concern (Jaffe, 2000). While there are some infamous examples of such applications, these likely comprise only a small fraction of patents.
As for the move away from automatic examination to examination by request, a rich set of potential incentive effects might be imagined, at least in theory. For instance, under the new rules an initial patent application is analogous to the purchase of a "real option" in the following sense. Currently, an inventor or a firm can pay an initial application fee, setting the priority date for the application. Then if at any time within seven years market conditions become more favourable or rivals appear to be encroaching on the applicant's core technology areas, the applicant can pay the "strike price" to request examination, thus exercising the option. In practice, however, the fact that Canadian firms tend to apply first in the U.S. and thus to rely on (automatic) examination by USPTO examiners mitigates any likely incentive effects of this change. For these reasons, I ignore the possible implications of the last two aspects of the patent reform, and concentrate in this chapter on empirically identifying consequences of the priority and disclosure amendments.

When might these effects be observable in the data? Although the reforms were officially implemented on October 1, 1989, they were first publicly discussed in June of 1986, and were passed into law in November, 1987. It is therefore possible that firms began adjusting their behaviour in anticipation of the implementation of the policy changes. In the analysis below I adopt a flexible approach that allows me to identify the likely starting date of any reform-induced adjustments, and thus I can accommodate the possibility that firms began their response prior to 1989.  

---

13Conversations with Canadian patent agents suggested that they and their clients do not in fact view a Canadian patent application as a real option purchase. Instead, clients are encouraged to use U.S. examination as a sounding-board for an invention's worth, and then apply (and request examination) in Canada if the U.S. feedback is promising.  
14These dates are drawn from the Patent and Trademark Institute of Canada's "Message from the President" of 1987 and 1989, both published in the Canadian Intellectual Property Review (volumes 3 and 5).
4.3 Aggregate Evidence

A natural first step in the analysis of a particular policy change is an investigation of aggregate trends. Motivated by the theoretical predictions of the previous section, this section provides an informal look at some aggregate data. Figure 1 shows total Canadian R&D spending in the manufacturing sector, which is the focus of this chapter. The data were originally expressed in nominal Canadian dollars and were deflated to 1992 dollars using a price deflator constructed as an average of the manufacturing-sector price index (CANSIM's Industrial Product Price Index for manufacturing) and an index of unit labour costs in manufacturing (see Hall, 1990 and Coe and Helpman, 1995 for further discussion of this construction).  

The vertical line in Figure 1 indicates the introduction of the reforms in October of 1989. There is no apparent change in the trend or intercept of R&D spending as a result of these reforms. R&D investment does appear to have a moderate cyclical component: the recessions of the early 1980s and 1990s seem to coincide with roughly horizontal segments of the graph. There is a distinct upward turn in the trend in approximately 1992.

Figure 2 shows R&D expenditures in three key Canadian industries: communications, transportation, and drugs. These are among the most patent- and R&D-intensive industries in Canada, so if the patent reforms were to have an observable effect on aggregate R&D expenditure it should be noticeable here (Trajtenberg, 2000). The reforms do not seem to have an impact on R&D in any of these three industries. Communications R&D appears to grow modestly between 1985 and 1992, and then very rapidly thereafter. R&D investment in transportation shows a steady upward trend over the entire period. Drug R&D noticeably accelerates prior to the 1989 reforms, most likely due to the scaling back

---

15 This deflator implicitly assumes that R&D costs consist of capital and labour costs in equal proportions. Experimentation with alternative deflators suggested that the results are not particularly sensitive to this specific choice.
of the compulsory licensing provision for patented pharmaceuticals in 1987 (Pazderka, 1999). These two figures thus provide scant evidence of any impact of the 1989 reforms on R&D spending in manufacturing (Hypothesis 1).

The other principal hypothesis identified by the model of priority rules concerns the extent to which technologies are disclosed through the patent system (Hypothesis 2). Because of the tendency of Canadian inventors to file first in the U.S., as well as the practical difficulty in identifying firm-level application information in Canadian patent data prior to 1989, I use patent filings in the U.S. to capture this decision. Figure 3 shows the total number of patent applications filed by Canadian inventors in the U.S. between 1973 and 1997. Applications are steady until about 1985, after which they grow much more quickly. There is no clear change during or after 1989.

The priority model of Scotchmer and Green (1990) identifies the propensity to patent as the primary variable of interest. Counts of applications may be a misleading measure of patent propensity because these are affected both by rates of innovation and by the decision to patent or suppress an invention (Gallini, Putnam and Tepperman, 2001). To better isolate the latter effect, a measure of patent propensity is shown in Figure 4 which attempts to roughly control for the rate of innovation. This is simply total patent applications divided by total business R&D spending in Canada (expressed in 1992 dollars). The propensity to patent shows a steady decline until about 1985, after which it increases. Propensity is roughly constant between 1991 and 1997. Again, the aggregate data are not consistent with the predictions of the model as the increase in patent propensity seems to substantially pre-date the 1989 reforms.

The discussion thus far has concentrated on aggregate evidence for the hypotheses

---

16 In constructing the firm-level patent portfolios used in section 5 below, Canadian data were found to be less useful. Before 1989 only granted patents are available in the Canadian Intellectual Property Office database, while after 1989 granted and non-granted applications are displayed. The latter group includes both those for which examination has been requested and unexamined applications. In contrast, the U.S. data are consistent over the same time period.

17 Total (rather than manufacturing) R&D spending is used as the denominator here, so the data are adjusted using the GDP deflator rather than the average price index discussed previously.
concerning the change in the priority rule. It is more difficult to investigate at an aggregate level the "small firm" hypothesis (Hypothesis 3) generated by the application disclosure model of Aoki and Spiegel (2000), as a discussion of firm size requires firm-level data. Accordingly, I postpone a fuller treatment of this hypothesis until the next section.

4.4 Firm-level Evidence on R&D

4.4.1 Data and variable construction

This study uses a unique firm-level dataset to examine the impact of the reforms. Firm data are drawn from the Canadian File of the COMPUSTAT database and cover the time period from 1980 to 1999. As previous empirical studies on patenting an innovation have done, I concentrate on the manufacturing sector only (SICs 2000 through 3999).

Since R&D data are necessary both for the analysis in this section and the patenting model in the next section, I select all firms with at least three years of R&D data prior to 1996. The use of publicly-traded firms with a minimum amount of R&D data necessarily implies that the sample is not representative of all Canadian manufacturing firms: the mean sample firm is larger and more R&D-intensive than the "average" Canadian firm. In examining whether patent policy has an observable impact at all, a sample of this sort is useful because these firms are more likely to be sensitive to such policy changes.

The resulting sample consists of an unbalanced panel of 84 firms from a variety of

---

18 Nonetheless, Rafiquzzaman (1999) presents some aggregate evidence indicating that patent applications by small firms grew at a faster rate than those by large firms between 1990 and 1997, but applications by small firms were disproportionately unsuccessful in being granted. He interprets this as evidence that in a first-to-file environment small firms need to rush to the patent office with low-quality inventions to hold off larger competitors. This does not really address the hypothesis suggested by Aoki and Spiegel (2000), which turns on the disclosure requirement's effect on the comparative rate of innovative effort between small and large firms.


20 There can be a lag between patent application and grant dates of up to six years in some cases. This selection procedure largely eliminates any such truncation bias, at the expense of sample size.
industries. After accounting for missing values, there are on average around 10 years of data per firm for the R&D regression analysis, and 8 years for the patent analysis because the sample period is shortened to accommodate grant lags. This sample size is somewhat smaller than, but roughly comparable to, other firm-level studies, including Hausman, Hall and Griliches (1984) (128 firms over 7 years); Hall and Ziedonis (2001) (95 firms over 10 years on average); and Branstetter (2001) (205 firms over 5 years).

As before, the R&D data are deflated by the average of the Industrial Product Price Index for manufacturing and an index of manufacturing unit labour costs. I use sales data to proxy for firm size. These data are deflated using the manufacturing product price index. As an additional control variable for the R&D analysis, I collect the inputs to a measure of Tobin’s “average q”. In the basic regressions I assume this variable is exogenous: I later consider relaxing this assumption. Average q here is defined as

\[
q_{it} = \frac{(\text{market value})_{it}}{(\text{replacement cost})_{it}}
\]

for firm \(i\) at time \(t\). I construct the numerator of this expression as the sum of the market value of equity (shares outstanding multiplied by end-of-year price), the book values of long-term and short-term debt, and the book value of preferred stock. To construct the replacement cost of firm assets for the denominator I follow the procedure in Perfect and Wiles (1994), which involves adjusting inventories for inflation depending on the firm’s valuation method, and calculating a recursive measure of the real value of the capital stock (see Perfect and Wiles, 1994 for details). These measures then replace the book values of inventories and fixed capital in the firm’s total assets.

For the analysis in the next section, I find the number of subsequently granted U.S.
Chapter 4. Do Priority and Disclosure Rules Matter?

patent applications for each firm in each year for which R&D data are available. The date of first application (i.e. the "priority" date) is used, which correctly accounts for applications filed previously in another country. Accurately assembling patent portfolios involves identifying firms’ name changes and subsidiaries, as the available patent data are not organized uniquely by firm entity, but rather according to the name of the assignee firm at the time of application. Details on name changes and subsidiaries were obtained from a number of sources, including industrial reference resources such as the CorpTech Directory of Technology Companies and Dun and Bradstreet’s Who Owns Whom, and firm resources (primarily websites and annual reports). This procedure also verified that the sample firms were indeed based in Canada (although some of their subsidiaries were not). Patent data were then obtained from the USPTO’s patent database, with patents awarded to subsidiaries counting toward the total patents of the sample firm.

Summary statistics of the firm-level variables are shown in Table 1. Since this is a panel data set, each observation represents a firm-year. R&D and sales are expressed in millions of 1992 Canadian dollars, while employees are measured in thousands. The average sample firm spends $5.4 million on R&D per year, generates almost $220 million in sales, and has just over 2300 employees. Just over 5 successful patent applications are filed per year on average. However, a large number (almost half) of the patent observations are zero in a given year, and in only 10 percent of cases did firms file 10 or more successful applications in a given year.

4.4.2 Empirical results: simple model

To investigate the impact of the 1989 patent reforms on R&D at the firm level (and thus test Hypothesis 1), I use a variant of the empirical model of Sakakibara and Branstetter (2001). Let the logarithm of real R&D spending and sales by firm \( i \) at time \( t \) be respectively denoted \( r_{it} \) and \( s_{it} \). For a number of observations (primarily small pharmaceutical firms) sales are equal to zero. I set \( s_{it} = 0 \) for such cases and create a dummy variable
\( n_{oit} \) which equals one if sales equal zero, and zero otherwise. Then the following simple relationship is initially hypothesized:

\[
    r_{it} = \beta_0 + \beta_1 q_{it} + \beta_2 s_{it} + \beta_3 n_{oit} + \eta_{it}. \tag{4.1}
\]

where \( \eta_{it} \) is a stochastic disturbance term. Thus R&D spending depends on the size of the firm through the sales term, and the firm’s investment opportunities and expected future profitability at time \( t \) through \( q_{it} \).\(^{25}\) The signs of \( \beta_1 \) and \( \beta_2 \) should be positive, while \( \beta_3 \) is expected to be negative or insignificant.

I exploit the longitudinal nature of the data and assume the error takes the form

\[
    \eta_{it} = \sigma_i + \gamma_t + \epsilon_{it},
\]

where \( \sigma_i \) is a firm-specific R&D effect, \( \gamma_t \) are “year effects” (typically dummy variables) capturing exogenous factors impacting R&D over time that are common to all firms, and \( \epsilon_{it} \) is an i.i.d. error term. This specification allows for firm heterogeneity in research intensity, and captures effects of external policy changes through the \( \gamma_t \) terms. In some specifications I also add a set of industry dummy variables to control for variation in research intensity across industries.

Scotchmer and Green’s (1990) model of priority predicts that a move to a first-to-file system should be associated with higher levels of R&D for both \( \text{ex ante} \) and \( \text{ex post} \) reasons. If the \( \gamma_t \) terms capture only the effects of patent policy (and no other aggregate economic phenomena), then according to Hypothesis 1 a sharp increase in the magnitude of the year dummies should be observed around the time of the patent reform.

As a first step, I estimate (4.1) both by random effects and fixed effects or “within” estimation. The random effects estimator treats \( \sigma_i \) as a random variable, and is consistent and efficient if these terms are truly uncorrelated with the other regressors. The fixed

\^{25}\)There is almost certainly a considerable amount of measurement error in the \( q \) variable. As Sakakibara and Branstetter (2001) point out, though, interest does not center on a precise, unbiased estimate of the \( q \) coefficient, but rather this variable is included as a control for the firm’s growth opportunities at a given time.
CHAPTER 4: Control for Industry Effects

Columns (1) and (2) of the table provide estimates of the control effects on the coefficients. In addition to the control effects, I also include a set of industry-level controls for industry-level factors that influence the estimated coefficients. In general, it is considerably higher in the sample than the coefficient. The elasticity of R&D spending with respect to the coefficient is significantly higher in the sample than the coefficient. The elasticity of R&D spending with respect to the coefficient is significantly higher in the sample than the coefficient.

Of primary interest are the estimated coefficient that is significantly different from zero. These coefficients are plotted in Figure 5. There is a distinct pattern of R&D spending, which appears to be steady at the time of the patent reform. In further analysis, I re-estimate the fixed effects specification with a time trend.

I include a time trend and a post-1989 dummy variable in the model. These estimates too suggest that R&D spending increases over time.
effects estimator treats the firm-specific terms as fixed over time, and is consistent if the $\sigma_i$ are correlated with sales and $q$ (see Greene. 1993; Branstetter. 2001). If the firm effects are in fact correlated with the regressors, there will be a systematic difference between the coefficients estimated by the two methods, and a Hausman test will reject the random effects model.

Columns (1) and (2) of Table 2 show results from these two estimation techniques, using a full set of year dummy variables to represent $\gamma_t$. In the random effects specification I also include a set of dummy variables for broad industrial clusters in order to control for industry-level factors that impact on the incentive to engage in R&D.\textsuperscript{26} The pattern of the estimated coefficients in the two models is similar, although the sales elasticity is considerably higher in the random effects specification. A Hausman test strongly rejects independence of the $\sigma_i$, so I concentrate on the fixed effects specification for interpretation.\textsuperscript{27} The elasticity of R&D with respect to sales is 0.3 and is significantly different from zero. The coefficients on $q$ and the dummy for no sales are signed as expected but insignificant.

Of primary interest are the estimated year effects. These are positive and, notably, significantly different from zero beginning in 1990. For easier interpretation, the coefficients are plotted in Figure 5. There is a distinct upward trend in the time path of R&D spending, which appears to be steeper after 1989. However, there is no dramatic increase at the time of the patent reform. To further interpret the year effects, in column (3) I re-estimate the fixed effects specification, but instead of a full set of year dummies here I include a time trend and a post-1989 dummy which represents a shift in the intercept. These estimates too suggest that R&D spending increased over time but without a sud-

\textsuperscript{26}These dummies are swept out of the fixed effects specification by the differencing procedure. The eight clusters are: basic manufacturing; chemicals and petroleum; drugs; metals and stone; fabricated metals and machinery; communications and electronics; electrical, computers and instruments; and transportation and aircraft. These roughly follow the aggregation of narrower industrial classes employed by Sakakibara and Branstetter (2001).

\textsuperscript{27}The chi-square statistic is 42.22 with 22 degrees of freedom which is significant at less than 1%.
den change at the onset of the patent reforms. Overall, the trend estimate implies that real R&D spending grew at approximately 6 percent per year, controlling for firm size. Therefore this static model does not provide strong evidence in favour of Hypothesis 1.

### 4.4.3 Robustness of the simple specification

The estimation results of the previous model are interesting, but inference based on the estimated year effects from the simple model is complicated by two main factors. First, the year effects are assumed to reflect only the impact of patent policy and not other important aggregate economic forces; and second, the model's specification is inherently static, while in reality it may be more appropriate to consider R&D spending as a dynamic process. In addition, to this point I have considered explanatory variables to be strictly exogenous, whereas a more realistic model would allow for endogeneity of a firm's sales and average $q$. I consider the first problem in this subsection and the second set of issues in the following subsection.

In principle, the year effects could be picking up other factors common to all firms that influence R&D spending. Macroeconomic fluctuations are a possible concern in this regard: if R&D spending is procyclical, a downturn in the economy could lead to lower estimated year effects, but would obviously be unrelated to patent policy. To see whether this is a concern, or whether instead the other explanatory variables (notably sales and $q$) might be adequately capturing and controlling for cyclical effects, I identify business cycle peaks and troughs between 1980 and 1999 from CANSIM's industrial capacity utilization series for manufacturing.\(^{28}\) According to these data, business cycle troughs occurred in 1982Q4 and 1991Q1, while peaks occurred in 1988Q1 and 1995Q1. The estimated year effects in Figure 5 do not appear to match the pattern that would be expected if they were influenced by cyclical fluctuations. In fact, sharp downturns are not evident at all.

\(^{28}\)These data are from CANSIM series D883647. The cyclical effects identified in this series are qualitatively very similar to those of the manufacturing sector GDP series shown in Treffer (2000).
during the trough periods, nor are peaks characterized by sharp upturns. Thus the year coefficients are likely not significantly affected by cyclical fluctuations.

Another factor potentially confounding interpretation of the year coefficients is the Canada-U.S. Free Trade Agreement (FTA), which was implemented on January 1, 1989. This called for a lowering of tariff protection in a variety of Canadian industries. If firms in these industries responded by strongly increasing their R&D efforts to enhance efficiency (perhaps in the expectation of more vigorous competition), the post-1989 year effects might overstate the impact of patent policy. Treffler (2001) identifies 71 4-digit industries that were most affected by the FTA, in terms of the mandated tariff concession to the U.S. relative to the rest of the world. I re-estimate the fixed effects regressions from Table 2, dropping observations from any firms in these industries.\(^{29}\) The results of these unreported regressions are extremely similar to the full sample results. In particular, the estimated year effects exhibit an identical pattern of significance. I thus conclude that the contemporaneous introduction of the FTA does not affect the estimation.

It is also possible that "global" factors such as patent reform in the U.S. influenced the behaviour of Canadian firms. A number of court decisions during the 1980s indicated that U.S. courts were more "pro-patent" than previously (Jaffe, 2000; Hall and Ziedonis, 2001). If Canadian firms enhanced their R&D as a response to these events, it would be difficult to accurately identify the extent to which these firms were responding specifically to the Canadian reforms. Ideally, a control variable could be constructed that allowed global effects to be taken into account in the empirical analysis. Although practically speaking this is difficult to do (see the concluding section for further discussion). One point which might dispel some of the concern over this issue is that these court cases appeared to gradually strengthen the perceived degree of patent protection (Hall and Ziedonis, 2001). whereas in this chapter I am testing for a sudden sharp response in R&D

\(^{29}\)The COMPSTAT data attach a primary 4-digit U.S. SIC code to each firm, which I compare to the 4-digit Canadian SIC codes in Treffler (2001). There are 8 dropped firms accounting for a total of 83 observations.
and patenting. It would be difficult to attribute any finding of such a sharp response to the incremental changes occurring in the U.S.

4.4.4 Dynamics and simultaneity

The second major class of specification issues concerns the dynamic nature of R&D. and the related question of the exogeneity of the regressors. R&D investment may be costly for firms to adjust, and may therefore only slowly approach its desired long-run level. Hall, Griliches and Hausman (1986) show evidence of sluggish R&D adjustment: in a large sample of U.S. manufacturing firms they find that R&D autoregressions produce a coefficient near one on the first lag and a small and barely significant coefficient on the second. This implies that a consideration of the dynamic adjustment properties of R&D is important in practice. In order to discern the immediate impact of the policy change in question, it is necessary to somehow control for this dynamic pattern.

Consider then the following simple R&D specification as a variant of (4.1):

\[ r_t = \alpha r_{t-1} + \beta_0 + \beta_1 q_t + \beta_2 s_t + \beta_3 n_{o,t} + \eta_t. \]  

(4.2)

which is just the previous specification with a lagged R&D term (see also Sakakibara and Branstetter, 2001). This model can be justified by assuming that the long-run desired level of R&D spending is related to sales and q. Dynamic adjustment toward this long-run level then takes place according to (4.2). If the exogenous patent reforms affected longer-run R&D spending, we would expect to see a distinct increase in the year coefficients in the post-1989 period, or alternatively a positive and significant coefficient

---

30I also experimented with a variety of alternative dynamic specifications in the estimation that follows, allowing for more lags of R&D and lagged values of the explanatory variables. As none of these alternative specifications appeared to provide any additional benefit, I report only estimates of (4.2).

31Mulkay, Hall, and Mairesse (2000) consider a more general dynamic structural specification for R&D investment in which a long-run relation between the R&D "capital stock" and sales is hypothesized. In their specification, the firm's stock of knowledge capital depreciates and is added to by R&D investment each year until the desired level is reached. In order to operationalize their specification, a good estimate of knowledge capital in the initial year of data is required, involving either educated guesswork or a significant amount of pre-sample information. I sidestep this problem by using the simple specification in (4.2), which incidentally is very similar to the strategy of Sakakibara and Branstetter (2001).
on the post-1989 intercept dummy. It would also be unsurprising to find that much of 
the time trend effect identified in the previous static estimation is absent here, as the 
empirical model now accounts for within-firm dependence over time.

Traditional fixed effects or within estimation of (4.2) successfully eliminates firm ef-
fects which might be correlated with the other regressors. However, some additional 
problems can arise in this context. A recent literature focuses on consistent and effi-
cient estimation of dynamic models such as (4.2) using panel data (e.g. Arellano and 
latter two studies suggest that fixed effects estimation is reasonably successful at min-
imizing biases related to measurement error and simultaneity between regressors and 
past disturbances when the time period under study is sufficiently long.\footnote{Mul-
key, Hall. and Mairesse. (2000) claim that the bias due to past-disturbance simultaneity is very 
small with 12 years of data per firm. Since I have up to 20 years of data per firm (and an average of 
over 10 years) the same conclusion is likely to hold here.} However, the 
within estimates may yet be biased and inconsistent due to simultaneity between cur-
rent disturbances and the regressors. To address this problem, these studies also employ 
alternative efficient instrumental variables estimators using the generalized method of 
moments (GMM).

The GMM technique accounts for unobserved heterogeneity by first-differencing (4.2). 
using lagged levels as instruments for differenced endogenous variables (see Arellano and 
Bond. 1991). Unfortunately, this approach, although efficient among single-equation 
instrumental variables methods, can perform poorly in relatively small samples (Mulkay. 
Hall. and Mairesse. 2000). In Table 3, I therefore present results from estimating (4.2) 
using both fixed effects and GMM estimation.

Column (1) shows the estimated coefficients from the standard fixed effects model and 
a full set of year dummy variables. The lagged R&D coefficient is significant, confirming 
R&D indeed adjusts slowly. Sales retains its positive impact, although the coefficient is 
lower than in the static model. Turning to the year effects, there is no evident trend
over time, nor is there any dramatic change after 1989. In column (2) I re-estimate the fixed effects equation with a trend and a post-1989 dummy. This confirms that the trend component is removed by allowing for lagged R&D. In addition, there is evidence that the long-run level of R&D increased after 1989, as the coefficient on the dummy variable for the regime change is positive and significant at 5 percent.

In columns (3) and (4) I retain the same specification, but instead use the GMM estimator of Arellano and Bond (1991) rather than the fixed effects estimator. Column (3) shows estimation results when variables other than lagged R&D are treated as exogenous. The coefficients are for the most part similar to the fixed effects estimates, although the lagged R&D coefficient is lower, as is the post-1989 year dummy. The Sargan test for this estimation rejects the model's overidentifying restrictions. This is not uncommon in small-sample estimation, and might be because the remaining variables are actually endogenous, or possibly because of heteroskedasticity in the error term. I therefore re-estimate the same equation in column (4), allowing for endogeneity of q and sales by including up to four lags of these variables as instruments, and report heteroskedasticity-consistent standard errors. While the coefficients in column (4) themselves are not much different, standard errors are larger. Overall, the similarity between the within and GMM estimates suggests that simultaneity is not likely to be a major problem: however, since the sample is small, the instrument set is weak and this conclusion must be tentative.

These estimates suggest that R&D may have responded to the patent reform as Hypothesis 1 predicts. However, there may yet be hidden effects influencing the estimation which are driving this result. One additional point that is convenient to consider at this stage is the impact of the repeal of compulsory licensing of pharmaceuticals. As men-

---

33] also estimated models with a full set of year dummies using GMM. These coefficients were all insignificant and showed no particular trend.

34] The chi-square statistic is 288.1, with 170 degrees of freedom (p < 0.001). Other diagnostic tests are also performed which yield results more favourable to the specification: the data do not reject the existence of first-order autocorrelation in the differenced residuals, but do reject second-order autocorrelation.
tioned previously, this policy change was enacted in 1987 and had significant effects on
the R&D expenditures of large drug firms in Canada (see also Pazderka, 1999). Drug
firms in this sample may still be reacting to this regime change in 1990, particularly
those relatively small research-intensive firms that depend on alliances with larger drug
manufacturers. This would influence the estimated coefficient of the post-1989 dummy.
In column (5) I therefore re-estimate the fixed effects specification of column (2), allowing
for different post-1989 intercepts for drug firms and non-drug firms. As suspected, the
drug coefficient is large and significant: these firms did indeed increase their R&D after
1989. While some of this increase may have been due to the priority change, much of it
was likely caused by the repeal of compulsory licencing. The non-drug coefficient is pos-
itive but imprecisely estimated. This suggests that non-drug firms may have increased
their R&D in line with the hypothesis, although the sample size does not allow any more
precise inference to be made.

The dynamic estimates presented in this subsection thus provide evidence on two
points. First, after controlling for within-firm R&D adjustment over time, the trend
effect visible in Figure 5 is insignificant. Second, consistent with Hypothesis 1, the
estimates point to perhaps a modest increase in R&D following the patent reforms. For
example, with an estimated regime change coefficient of 0.1, the immediate increase in
R&D spending is $\exp(0.1) - 1 = 10.5$ percent.

4.4.5 Public disclosure and the “small firm” hypothesis

To this point the empirical analysis has focused on Hypothesis 1, derived from Scotchmer
and Green (1990). As described in section 2, Aoki and Spiegel (2000) predict that small
firms should decrease their R&D investment relative to large firms in response to a
change from a confidential filing to a public disclosure system (Hypothesis 3). I test this
by estimating the following specification, which allows for an interaction between the
Chapter 4. Do Priority and Disclosure Rules Matter?

Patent reform dummy and firm size:

\[ r_{it} = \alpha r_{it-1} + \beta_0 q_{it} + \beta_2 s_{it} + \beta_3 no_{it} + \gamma D_{89} \ast s_{it} + \delta_1 D_{89} + \delta_2 \text{time} + \sigma_i + \epsilon_{it}. \]

where \( D_{89} \) is a dummy variable indicating the post-1989 time period, and \( \text{time} \) is a calendar time trend. Hypothesis 3 predicts that \( \gamma \) should have a positive sign, as smaller firms decrease their R&D relative to large firms in the post-reform period.

Fixed effects estimates of this equation are shown in column (1) of Table 4. Contrary to Hypothesis 3, the estimated coefficient on the interaction term is negative and significant at the 10 percent level. Thus smaller firms actually have relatively higher R&D intensities in the post-reform period. To get an idea of the size of this effect, the impact of the reform on relatively small and large firms can be compared. For firms in the 25th percentile of the size distribution (as measured by sales) the introduction of the reforms implied an immediate increase in R&D of approximately 25 percent, significant at the one percent level. Firms in the 75th percentile of sales, on the other hand, saw a statistically insignificant increase of 7.6 percent.

Since there is bound to be a considerable degree of heterogeneity across industries in the effects of firm size on R&D, I re-estimate the above equation allowing \( \gamma \) to vary by industry cluster. By this method it is possible to tell whether particular industries are driving the small firm result, or whether instead a more general phenomenon seems to be at work. The results including industry-specific interactions are shown in column (2). Most of the relevant coefficients remain negative, although only the chemicals coefficient is significantly different from zero. The absence of a positive coefficient for any industry strongly indicates that rather than small firms engaging in less R&D, they may in fact do relatively more after the introduction of the reforms.

Why might such a result be obtained? It is possible that as suggested in section 2, Canadian firms do not care about domestic public disclosure rules, as they are still able

\[ \text{Footnote: I also experimented with a version that excludes the lagged R&D term, which generates qualitatively similar results.} \]
to use the confidential U.S. system. However, this alone does not explain why small firms could be *more* R&D-intensive in the post-reform period. Several more substantive explanations are plausible. At a theoretical level, there may be reasons why smaller firms could benefit from an introduction of public disclosure that are not considered by Aoki and Spiegel. For instance, disclosure could reduce informational asymmetries or uncertainty regarding rivals' research programs. R&D typically involves undertaking a considerable amount of risk, which affects small firms with less diversified project portfolios more acutely: any policy which lessens the risk would help stimulate additional spending. Alternatively, knowing more accurately how close a rival is to reaching a research goal could help stimulate R&D, in that a small firm can be more certain of receiving a return on its expenditures. A more empirical explanation might involve a decrease in the fixed costs of research, particularly in the 1990s, which allowed small firms to enter R&D races on a larger scale than was previously possible. Further theoretical and empirical research is needed to sort out these hypotheses.

### 4.5 Firm-level Evidence on Patenting

#### 4.5.1 Empirical method

In this section I turn to estimation of the "patent production function" in order to test Hypothesis 2. The patent production function literature has historically focused on the empirical relationship between patenting and R&D expenditures, with other explanatory variables (such as firm size) sometimes included. One approach to estimation specifies the logarithm of patents as a function of the logarithm of current and lagged R&D and other variables (Pakes and Griliches, 1980; Branstetter, 2001; Sakakibara and Branstetter).  

---

36 In addition, the extension of patentability during the 1980s to new subject matter such as biotechnology and software coincided with widespread entry by small start-ups with few physical assets. See Jaffe (2000) and Kortum and Lerner (1998).
2001). These papers deal with zero values in the dependent variable by adding one and sometimes including a dummy variable representing "no patents."

In patent data in general, and particularly in my data, a large number of observations are zero or one. So count data methods such as Poisson regression are usually more appropriate. These methods have been developed in a related branch of the patent literature (see Hausman, Hall and Griliches, 1984; Hall, Griliches and Hausman, 1986; Hall and Ziedonis, 2001). In the basic (pooled) Poisson model, there is a hypothesized relationship between the characteristics of the firms (for example R&D spending or size), denoted $X_{it}$, and the expected number of patents per period:

$$E[p_{it} | X_{it}] = \lambda_{it} = \exp(X_{it} \beta + \gamma_t),$$

where as before the $\gamma_t$ terms are year effects. The random error comes from the fact that the $p_{it}$ are assumed to be Poisson distributed, with mean and variance given by $\lambda_{it}$.

In panel data it is possible to allow for permanent unobserved differences in patent intensities across firms: let the expected number of patents be $\lambda_{it} = \alpha_i \lambda_{it}$, where $\alpha_i$ is a firm-specific effect that is typically specified as $\alpha_i = \exp(\omega_i)$. Thus, as in the previous section, the $\omega_i$ terms can be thought of as representing firm-level heterogeneity in patent intensity that does not change over time, while the $\gamma_t$ terms capture factors such as patent policy that impact on all firms' patenting over time (see Hausman, Hall and Griliches, 1984; Hall and Ziedonis, 2001). These two terms encapsulate the propensity to patent, and this section focuses on the year effects, which are common to all firms.

It is worth emphasizing that the analysis presented here is not intended to demonstrate the effects of patent policy on innovation. Patent counts are in any case an imperfect measure of innovative output, as many researchers have acknowledged. Instead, I attempt to control for innovative effort by including R&D spending and firm size as independent variables. Then any remaining independent variables having explanatory

---

37 See Griliches (1990) for an authoritative overview of the pros and cons of patent data in general.
power are actually capturing the incentive to disclose an invention through the patent system, which is precisely the hypothesis to be tested.\textsuperscript{38}

Fixed effects estimation in the context of count data is different from standard within estimation as it is necessary to condition on the total patents awarded to the firm over the sample period. Firms which do not have any patents in the sample do not provide any additional information toward the conditional estimation, and are thus dropped. On the other hand, random effects estimation includes all firms, and additionally allows for the variance and the mean of the distribution to differ from each other in a more realistic way (Hausman, Hall and Griliches. 1984). The standard Hausman test can be used to determine whether firm effects are correlated with the regressors, in which case inference should be based on the consistent conditional estimates (Cameron and Trivedi, 1998).

I include in $X_{it}$ the explanatory variables identified by the patent production function literature as being most relevant. The log of contemporaneous R\&D, $r_{it}$, controls for the fact that firms with larger R\&D budgets are likely to apply for more patents in any given period. Hall, Griliches and Hausman (1986) examine whether additional lags of R\&D should also be included: there is some evidence that R\&D influences patenting over quite a long term, although these lags are poorly identified and current-period R\&D remains the best predictor. A possible criticism of the use of current-period R\&D in the present context might be that when the policy change takes place, it would take some time for even immediate adjustments in R\&D spending to work their way into patenting, and thus past R\&D is more appropriate. However, this is not strictly accurate, since firms may have a stock of unpatented inventions at the time of the reform which they can choose to patent. I therefore retain current R\&D as a control. I also include the log of sales, $s_{it}$, (as well as the "no sales" dummy) to control for other size factors. The coefficients on R\&D and sales are interpreted as elasticities.

\textsuperscript{38}This approach to measuring the propensity to patent is the same as that used by Hall and Ziedonis (2001). See Gallini, Putnam and Tøpperman (2001) for a discussion of different methods of disentangling the innovation and disclosure decisions using patent data.
4.5.2 Estimation results

Table 5, column (1) shows the results of a Poisson random effects regression of the yearly count of patents on the control variables and the full set of year dummies. Only the random effects coefficients are reported, as a Hausman test does not reject this more efficient model in favour of the conditional fixed effects model. Note that there are fewer observations and years for this estimation than for the R&D equation: patent data were only collected up to 1995 so as not to improperly exclude pending applications. As expected, the R&D elasticity is positive and significant, at about 0.24. The coefficients on the sales variables are improperly signed and insignificant, suggesting that after controlling for R&D spending firm size is not an important predictor of patenting for these firms.

The estimated year coefficients exhibit an interesting pattern: they are far the most part insignificantly different from zero before 1988, after which they are largely positive and significant. Figure 6 plots the year coefficients and shows the sharp increase in patent propensity in 1988. These estimates imply that the propensity to patent increased by \( \exp(0.357 - 0.113) - 1 = 28 \) percent from 1987 to 1988, which is significant at the 1 percent level. The magnitude, if not the significance of this change decreases slightly when the full set of year dummies are replaced (in an unreported regression) by a time trend and an intercept dummy variable indicating 1988 and after. In this case the estimated increase in patent propensity is about 16 percent, significant at less than the 5 percent level. Since the patent reforms were not officially implemented until 1989, this timing implies that firms began responding in advance of the policy change. A firm with a stock of unpatented inventions might decide to begin patenting prior to the policy change if it expected its rivals to do likewise with their similar inventions. Thus, firms might try to preempt their competitors by patenting early.

How robust is this finding? The Poisson model captures the count nature of the data well, but it also imposes a strong distributional assumption on the generation of patentable inventions. One alternative which has been used in the literature is to assume
the data follow a negative binomial distribution, while still allowing for firm-specific effects. Such a model is derived in Hausman, Hall and Griliches (1984). It has the benefit that each firm’s random effect has a different distribution, which takes on a realization each year (in contrast to the Poisson model with random effects in which firms’ random intercepts are constant over years). Thus the negative binomial model is somewhat more flexible in allowing for unobserved heterogeneity.\textsuperscript{39} The coefficients from the negative binomial random effects estimation are shown in column (2). There is still a significant difference between the pre-1988 and post-1988 year coefficients which is if anything more dramatic than the Poisson estimates imply. In addition the R&D elasticity is slightly larger, and the sales elasticity is positive and significant.

Another alternative is to drop the distributional assumption and estimate a standard log-log random effects model. This model is appealing in that no strong assumptions are needed to justify it; however, biases in the estimates can occur when a large number of patent observations are zero. To create the dependent variable in this case I simply add one to each patent count and take the log (see Branstetter. 2001). The estimates are presented in column (3). Again, there is a significant change in patent propensity beginning in 1988. The R&D elasticity is smaller than the previous estimates, as might be expected since the transformation of the dependent variable induces a smaller range of variation. The sales elasticity is again positive and significant, although quite small.

These alternative specifications strengthen the finding of an increase in patent propensity around the time of the reform. One further test of robustness is possible given the cross-sectional variation in the data. Scotchmer and Green (1990) show that in theory patenting increases with a move from the first-to-invent system because firms have a diminished incentive to strategically suppress a technological breakthrough. The strategic disinclination to patent under first-to-invent depends crucially on the inability of rivals

\textsuperscript{39}This flexibility comes at a cost, as the distributional assumption may be even more restrictive than the Poisson model. If the data do not follow a negative binomial distribution the estimates will be biased and inconsistent (Branstetter. 2001 provides more discussion of this point).
to learn about the technology in the absence of a patent. or in other words that the invention can be kept secret. In practice. secrecy works better in some industries than others. as the survey results of Cohen. Nelson and Walsh (2000) indicate. They show that firms rely heavily on secrecy to protect their investments in R& D in food. textiles. plastics. chemicals (including drugs). metals. transportation. and some electrical industries. In contrast. secrecy is less effective in communications equipment. computers. machinery and instruments (see also Grindley and Teece. 1997). Thus. if the Scotchmer and Green (1990) model has predictive power in explaining the patenting effects of a move to first-to-file. we would expect to see firms in the former group of industries reacting more strongly (in terms of the propensity to patent) than those in the latter.

I therefore estimate the patent production function using a time trend and industry cluster-specific post-1988 dummies. These results are shown in Table 6. The estimated industry-specific interactions are broadly consistent with the theory: coefficients for basic manufacturing. chemicals. drugs. metals and transportation are all significant and positive. while those for communications. electronics and instruments are either insignificant or negative. One surprising finding is the strength of the negative effect for communications equipment. since the theory does not predict a decrease in patenting for any industry. This is particularly puzzling in light of the fact that U.S. high-tech firms in similar industries have been stepping up their patenting over the same time period (Hall and Ziedonis. 2001). Whether this is a peculiarity of Canadian high-tech firms patenting in the U.S. or is due instead to some aspect of the industry not adequately captured by the theoretical model is an issue that awaits further investigation.

\footnote{These results refer specifically to product innovations. Process innovations are more easily kept secret. and accordingly firms in most industries tend to rely primarily on secrecy for protection of processes. Firms in the latter set of industries tend to rely more on lead time and complementary manufacturing capabilities to protect product innovations. Secrecy is less effective in these industries because. as Grindley and Teece (1997) point out. firms generally have a good idea of the research orientation of their competitors. Reverse engineering of technologies is not uncommon as a method of learning about rivals' innovations.}
4.6 Conclusion

In this chapter I have investigated the empirical consequences of Canada’s 1989 patent reforms using as a foundation prior theoretical work on priority rules and public disclosure policy. The results provide some support to the hypotheses derived from the theoretical models. With regards to the change in the priority rule from first-to-invent to first-to-file, I find a moderate (but imprecisely estimated) effect on firms’ R&D investments. Also in line with the theory of Scotchmer and Green (1990), I find a significant change in the propensity to patent at approximately the time of the reforms. This suggests that firms reacted to the implementation of the new priority rule by disclosing more of their technological advancements through patenting, just as the theory predicts. Last, I find no evidence in favour of the public disclosure model of Aoki and Spiegel (2000). Small firms do not appear to decrease their R&D in response to the new disclosure rule, contrary to the model’s prediction.

For this empirical analysis I have avoided taking a stand on whether the changes to the patent law in 1989 “strengthened” or “broadened” patent protection in Canada, in favour of a more agnostic approach that involves testing specific hypotheses. It remains an open question whether these changes actually increased the value of patent protection as Rafiquzzaman (1999) suggests. But given the maturity of the theoretical literature on the impact of patent policy, it is in many cases not necessary to make such an assumption. As I have attempted to illustrate, theoretical predictions from models of specific aspects of patent policy can be useful in grounding empirical analyses without resorting to judgements over patent strength.

Although this analysis provides some evidence of a strategic R&D response following the introduction of first-to-file, this must be interpreted as a tentative finding. There exist additional ways to increase confidence in these results that can be considered in extensions. One possibility would be to control for events outside of Canada by including industry-level U.S. variables in the R&D and patenting specifications. A more ambitious
extension along the same lines would include a sample of U.S. firms to act as a control group. These would ensure that any effects that are picked up in the year dummies are indeed related to Canada.

Further, I have followed the recent microeconometric literature on the effects of patent policy in using firm-level data: however, it remains conceivable that even this level of disaggregation is not fine enough to accurately pick up the hypothesized strategic R&D responses. Recall that the theory is developed at the level of the individual R&D race for a particular technology. In contrast, my data includes R&D expenditures by the firm on all items, not all of which are directly related to actual R&D competition with rival firms. Thus the next logical step would be to test Hypothesis 1 at the project level of disaggregation. While this requires an ambitious data collection effort, some previous studies have used project- or technology-level data on R&D and innovation with impressive results (see Henderson and Cockburn, 1996; Cockburn, Henderson and Stern, 1999; Branstetter and Sakakibara, 2000).

The results of tests of the public disclosure hypothesis point to an additional important area for future research on the effects of patent policy: the extent to which Canadian inventors depend on the U.S. patent system. Given the large U.S. market it is not surprising that Canadian firms patent extensively in the U.S., but this does not necessarily imply that Canadian patent policy is ineffective in influencing the behaviour of these firms. As a general point, though, it remains the case that Canadian patent policy might be constrained by the ability of firms to resort to a dominant patent system beyond the control of Canadian authorities. Obtaining a more precise idea of the degree to which this occurs is an interesting research avenue with significant policy implications.

\[41\] Of course, the influence of the public disclosure rule itself, at least in theory, might be perceived as socially undesirable.
Figure 4.1: Canadian Manufacturing R&D

Source: OECD, AMBERD database
Figure 4.2: R&D in Selected Industries

Source: OECD, ANBERD database.
Figure 4.3: Patenting by Canadian Inventors

Source: OECD, Basic Science and Technology Statistics; WIPO
Figure 4.4: Canadian Patent Propensity

Source: OECD, Basic Science and Technology Statistics, ANBERD database, WIPO
Figure 4.5: Estimated R&D Year Effects

Source: Column (2) of Table 2
Figure 4.6: Estimated Patenting Year Effects
Table 4.1: Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real R&amp;D ($MM)(^a)</td>
<td>963</td>
<td>5.428</td>
<td>1.912</td>
<td>0.001</td>
<td>5203.465</td>
</tr>
<tr>
<td>Real sales ($MM)(^a)</td>
<td>1133</td>
<td>218.855</td>
<td>2.676</td>
<td>0</td>
<td>25817.78</td>
</tr>
<tr>
<td>D(sales = 0)</td>
<td>1133</td>
<td>0.031</td>
<td>(N = 35)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average (q)</td>
<td>1019</td>
<td>1.83</td>
<td>2.407</td>
<td>0.203</td>
<td>25.413</td>
</tr>
<tr>
<td>Employees (M)(^a)</td>
<td>756</td>
<td>2.302</td>
<td>2.144</td>
<td>0.004</td>
<td>80.627</td>
</tr>
<tr>
<td>Patents</td>
<td>707</td>
<td>5.348</td>
<td>15.467</td>
<td>0</td>
<td>118</td>
</tr>
</tbody>
</table>

\(^a\)Geometric mean and standard deviation of log reported.
Table 4.2: Regression Estimates for R&D

<table>
<thead>
<tr>
<th></th>
<th>Coefficients</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Random Effects</td>
<td>Fixed Effects</td>
<td>Fixed Effects</td>
<td></td>
</tr>
<tr>
<td>Log(sales)</td>
<td>0.411***</td>
<td>0.294***</td>
<td>0.310***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.040)</td>
<td>(0.040)</td>
<td></td>
</tr>
<tr>
<td>D(sales = 0)</td>
<td>-0.014</td>
<td>-0.038</td>
<td>-0.089</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.262)</td>
<td>(0.275)</td>
<td>(0.273)</td>
<td></td>
</tr>
<tr>
<td>Average q</td>
<td>0.002</td>
<td>0.005</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.014)</td>
<td>(0.014)</td>
<td></td>
</tr>
<tr>
<td>Time trend</td>
<td></td>
<td></td>
<td>0.058***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.009)</td>
<td></td>
</tr>
<tr>
<td>D(1990+)</td>
<td></td>
<td>-0.030</td>
<td>(0.100)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1981</td>
<td>0.160</td>
<td>0.164</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.237)</td>
<td>(0.232)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1982</td>
<td>0.127</td>
<td>0.133</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.223)</td>
<td>(0.218)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1983</td>
<td>0.238</td>
<td>0.245</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.224)</td>
<td>(0.220)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1984</td>
<td>0.180</td>
<td>0.196</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.223)</td>
<td>(0.218)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1985</td>
<td>0.402*</td>
<td>0.435**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.219)</td>
<td>(0.214)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1986</td>
<td>0.268</td>
<td>0.306</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.218)</td>
<td>(0.213)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1987</td>
<td>0.247</td>
<td>0.286</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.212)</td>
<td>(0.207)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1988</td>
<td>0.229</td>
<td>0.274</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.209)</td>
<td>(0.205)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1989</td>
<td>0.055</td>
<td>0.125</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.206)</td>
<td>(0.202)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table continues on next page
Table 4.2 continued

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>0.318</td>
<td>0.397**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.204)</td>
<td>(0.200)</td>
<td></td>
</tr>
<tr>
<td>1991</td>
<td>0.358*</td>
<td>0.434**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.206)</td>
<td>(0.202)</td>
<td></td>
</tr>
<tr>
<td>1992</td>
<td>0.405**</td>
<td>0.480**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.204)</td>
<td>(0.200)</td>
<td></td>
</tr>
<tr>
<td>1993</td>
<td>0.484**</td>
<td>0.557***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.199)</td>
<td>(0.195)</td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>0.527***</td>
<td>0.623***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.200)</td>
<td>(0.196)</td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>0.661***</td>
<td>0.757***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.200)</td>
<td>(0.196)</td>
<td></td>
</tr>
<tr>
<td>1996</td>
<td>0.735***</td>
<td>0.838***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.200)</td>
<td>(0.197)</td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>0.894***</td>
<td>1.001***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.200)</td>
<td>(0.198)</td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>0.962***</td>
<td>1.074***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.202)</td>
<td>(0.198)</td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>0.955***</td>
<td>1.064***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.210)</td>
<td>(0.206)</td>
<td></td>
</tr>
<tr>
<td>Industry dummies</td>
<td>Sig. At 1%</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.831***</td>
<td>-0.261</td>
<td>-0.492</td>
</tr>
<tr>
<td></td>
<td>(0.472)</td>
<td>(0.261)</td>
<td>(0.201)</td>
</tr>
<tr>
<td>N</td>
<td>861</td>
<td>861</td>
<td>861</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.528</td>
<td>0.352</td>
<td>0.340</td>
</tr>
</tbody>
</table>

Notes: Standard errors reported in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels respectively.
Table 4.3: Dynamic Regression Estimates for R&D

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fixed</td>
<td>Fixed</td>
<td>GMM</td>
<td>GMM</td>
<td>Fixed</td>
</tr>
<tr>
<td></td>
<td>Effects</td>
<td>Effects</td>
<td>GMM</td>
<td>GMM</td>
<td>Effects</td>
</tr>
<tr>
<td>Lagged Log(R&amp;D)</td>
<td>0.689***</td>
<td>0.698***</td>
<td>0.566***</td>
<td>0.607***</td>
<td>0.674***</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.030)</td>
<td>(0.038)</td>
<td>(0.085)</td>
<td>(0.031)</td>
</tr>
<tr>
<td>Log(sales)</td>
<td>0.121***</td>
<td>0.121***</td>
<td>0.148***</td>
<td>0.172**</td>
<td>0.125***</td>
</tr>
<tr>
<td></td>
<td>(0.033)</td>
<td>(0.032)</td>
<td>(0.039)</td>
<td>(0.083)</td>
<td>(0.032)</td>
</tr>
<tr>
<td>D(sales = 0)</td>
<td>-0.100</td>
<td>-0.070</td>
<td>-0.184</td>
<td>-0.132</td>
<td>-0.140</td>
</tr>
<tr>
<td></td>
<td>(0.222)</td>
<td>(0.218)</td>
<td>(0.216)</td>
<td>(0.131)</td>
<td>(0.218)</td>
</tr>
<tr>
<td>Average q</td>
<td>0.020</td>
<td>0.017</td>
<td>0.010</td>
<td>-0.002</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.012)</td>
<td>(0.012)</td>
<td>(0.012)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>Time trend</td>
<td>-0.006</td>
<td>-0.008</td>
<td>-0.011</td>
<td>-0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.013)</td>
<td>(0.008)</td>
<td></td>
</tr>
<tr>
<td>D(1990+)</td>
<td>0.160**</td>
<td>0.128*</td>
<td>0.120</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.079)</td>
<td>(0.074)</td>
<td>(0.100)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug D(1990+)</td>
<td>0.697***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.189)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other D(1990+)</td>
<td>0.093</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.081)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1981</td>
<td>0.030</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.171)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1982</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 4.3 continued

<table>
<thead>
<tr>
<th>Year</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1988</td>
<td>0.048</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.151)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1989</td>
<td>-0.139</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.150)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>0.096</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.149)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1991</td>
<td>0.089</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.149)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1992</td>
<td>0.174</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.149)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1993</td>
<td>0.136</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.148)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>0.053</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.144)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>0.029</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.146)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1996</td>
<td>0.073</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.147)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>0.169</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.148)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>0.159</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.149)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>-0.021</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.157)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.095</td>
<td>-0.054</td>
<td>NA</td>
<td>NA</td>
<td>-0.117</td>
</tr>
<tr>
<td></td>
<td>(0.203)</td>
<td>(0.164)</td>
<td></td>
<td></td>
<td>(0.164)</td>
</tr>
<tr>
<td>N</td>
<td>775</td>
<td>775</td>
<td>683</td>
<td>683</td>
<td>775</td>
</tr>
</tbody>
</table>

Notes: Column (4) estimation treats Log(sales) and q as endogenous and reports robust heteroskedasticity-consistent standard errors.
Table 4.4: Disclosure Hypothesis Tests

Dependent variable: log(R&D)

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Fixed Effects)</td>
<td>(Fixed Effects)</td>
<td></td>
</tr>
<tr>
<td>Lagged Log(R&amp;D)</td>
<td>0.686***</td>
<td>0.674***</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.032)</td>
</tr>
<tr>
<td>Log(sales)</td>
<td>0.153***</td>
<td>0.146***</td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td>(0.053)</td>
</tr>
<tr>
<td>D(sales = 0)</td>
<td>-0.098</td>
<td>-0.080</td>
</tr>
<tr>
<td></td>
<td>(0.218)</td>
<td>(0.217)</td>
</tr>
<tr>
<td>Average q</td>
<td>0.014</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>Time trend</td>
<td>-0.003</td>
<td>-0.003</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>D(1990+)</td>
<td>0.371***</td>
<td>0.277*</td>
</tr>
<tr>
<td></td>
<td>(0.135)</td>
<td>(0.143)</td>
</tr>
<tr>
<td>D(1990+)*Log(sales)</td>
<td>-0.039*</td>
<td>-0.032</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>Chemicals, petroleum</td>
<td>-0.048**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td></td>
</tr>
<tr>
<td>Drugs</td>
<td>-0.068</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.068)</td>
<td></td>
</tr>
<tr>
<td>Metals, stone</td>
<td>-0.029</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td></td>
</tr>
<tr>
<td>Fabricated metals, machinery</td>
<td>0.047</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td></td>
</tr>
<tr>
<td>Communications, electronics</td>
<td>-0.009</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
<td></td>
</tr>
<tr>
<td>Other electrical, computer, instruments</td>
<td>-0.026</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td></td>
</tr>
<tr>
<td>Transportation, aircraft</td>
<td>-0.012</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.253</td>
<td>-0.169</td>
</tr>
<tr>
<td></td>
<td>(0.193)</td>
<td>(0.261)</td>
</tr>
<tr>
<td>N</td>
<td>775</td>
<td>775</td>
</tr>
</tbody>
</table>

Notes: Shown in (2) are industry-specific interaction coefficients. The omitted category is Basic manufacturing; its interaction is given by D(1990+)*Log(sales).
## Table 4.5: Patent Production Function Estimates

<table>
<thead>
<tr>
<th></th>
<th>Coefficients</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td></td>
<td>Poisson</td>
<td>Negative</td>
<td>Random</td>
</tr>
<tr>
<td></td>
<td>Binomial</td>
<td>Effects</td>
<td></td>
</tr>
</tbody>
</table>

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Log(R&amp;D)</td>
<td>0.238***</td>
<td>0.316***</td>
<td>0.161***</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.056)</td>
<td>(0.026)</td>
</tr>
<tr>
<td>Log(sales)</td>
<td>-0.008</td>
<td>0.100**</td>
<td>0.080***</td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td>(0.051)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>D(sales = 0)</td>
<td>0.253</td>
<td>-0.167</td>
<td>-0.019</td>
</tr>
<tr>
<td></td>
<td>(0.368)</td>
<td>(0.540)</td>
<td>(0.178)</td>
</tr>
<tr>
<td>1981</td>
<td>-0.008</td>
<td>-0.052</td>
<td>-0.047</td>
</tr>
<tr>
<td></td>
<td>(0.121)</td>
<td>(0.258)</td>
<td>(0.168)</td>
</tr>
<tr>
<td>1982</td>
<td>0.112</td>
<td>0.176</td>
<td>0.194</td>
</tr>
<tr>
<td></td>
<td>(0.115)</td>
<td>(0.236)</td>
<td>(0.156)</td>
</tr>
<tr>
<td>1983</td>
<td>0.057</td>
<td>0.079</td>
<td>0.098</td>
</tr>
<tr>
<td></td>
<td>(0.117)</td>
<td>(0.241)</td>
<td>(0.157)</td>
</tr>
<tr>
<td>1984</td>
<td>0.130</td>
<td>0.131</td>
<td>0.167</td>
</tr>
<tr>
<td></td>
<td>(0.114)</td>
<td>(0.237)</td>
<td>(0.155)</td>
</tr>
<tr>
<td>1985</td>
<td>0.262**</td>
<td>0.269</td>
<td>0.272*</td>
</tr>
<tr>
<td></td>
<td>(0.112)</td>
<td>(0.228)</td>
<td>(0.153)</td>
</tr>
<tr>
<td>1986</td>
<td>0.175</td>
<td>0.204</td>
<td>0.264**</td>
</tr>
<tr>
<td></td>
<td>(0.114)</td>
<td>(0.231)</td>
<td>(0.151)</td>
</tr>
<tr>
<td>1987</td>
<td>0.113</td>
<td>0.141</td>
<td>0.242</td>
</tr>
<tr>
<td></td>
<td>(0.116)</td>
<td>(0.234)</td>
<td>(0.149)</td>
</tr>
<tr>
<td>1988</td>
<td>0.357***</td>
<td>0.490**</td>
<td>0.485***</td>
</tr>
<tr>
<td></td>
<td>(0.112)</td>
<td>(0.221)</td>
<td>(0.147)</td>
</tr>
<tr>
<td>1989</td>
<td>0.436***</td>
<td>0.528***</td>
<td>0.532***</td>
</tr>
<tr>
<td></td>
<td>(0.111)</td>
<td>(0.219)</td>
<td>(0.145)</td>
</tr>
<tr>
<td>1990</td>
<td>0.340***</td>
<td>0.364</td>
<td>0.359**</td>
</tr>
<tr>
<td></td>
<td>(0.113)</td>
<td>(0.224)</td>
<td>(0.145)</td>
</tr>
<tr>
<td>1991</td>
<td>0.159</td>
<td>0.236</td>
<td>0.327**</td>
</tr>
<tr>
<td></td>
<td>(0.117)</td>
<td>(0.226)</td>
<td>(0.145)</td>
</tr>
</tbody>
</table>

*Table continues on next page*
Table 4.5 continued

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
<td>0.321***</td>
<td>0.302</td>
<td>0.356**</td>
</tr>
<tr>
<td></td>
<td>(0.114)</td>
<td>(0.225)</td>
<td>(0.144)</td>
</tr>
<tr>
<td>1993</td>
<td>0.411***</td>
<td>0.345</td>
<td>0.331**</td>
</tr>
<tr>
<td></td>
<td>(0.112)</td>
<td>(0.220)</td>
<td>(0.143)</td>
</tr>
<tr>
<td>1994</td>
<td>0.338***</td>
<td>0.402*</td>
<td>0.413***</td>
</tr>
<tr>
<td></td>
<td>(0.116)</td>
<td>(0.222)</td>
<td>(0.144)</td>
</tr>
<tr>
<td>1995</td>
<td>0.382***</td>
<td>0.325</td>
<td>0.329**</td>
</tr>
<tr>
<td></td>
<td>(0.118)</td>
<td>(0.224)</td>
<td>(0.144)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.198</td>
<td>-0.604*</td>
<td>-0.157</td>
</tr>
<tr>
<td></td>
<td>(0.343)</td>
<td>(0.352)</td>
<td>(0.180)</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-1377.78</td>
<td>-1203.13</td>
<td>NA</td>
</tr>
<tr>
<td>N</td>
<td>707</td>
<td>707</td>
<td>707</td>
</tr>
</tbody>
</table>

Notes: Standard errors reported in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels respectively. Random effects methods for count data are used to estimate (1) and (2). Dependent variables are patent counts for (1) and (2) and log(patents+1) for (3).
Table 4.6: Patent Estimates by Industry

<table>
<thead>
<tr>
<th>Dependent variable: count of patents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poisson Coefficients</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log(R&amp;D)</td>
<td>0.288***</td>
<td>(0.052)</td>
</tr>
<tr>
<td>Log(sales)</td>
<td>0.106</td>
<td>(0.066)</td>
</tr>
<tr>
<td>D(sales = 0)</td>
<td>0.463</td>
<td>(0.367)</td>
</tr>
<tr>
<td>Time trend</td>
<td>0.001</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Basic manufacturing</td>
<td>1.227***</td>
<td>(0.130)</td>
</tr>
<tr>
<td>Chemicals, petroleum</td>
<td>0.284*</td>
<td>(0.156)</td>
</tr>
<tr>
<td>Drugs</td>
<td>2.125***</td>
<td>(0.742)</td>
</tr>
<tr>
<td>Metals, stone</td>
<td>0.344***</td>
<td>(0.098)</td>
</tr>
<tr>
<td>Fabricated metals, machinery</td>
<td>0.371*</td>
<td>(0.224)</td>
</tr>
<tr>
<td>Communications, electronics</td>
<td>-0.272***</td>
<td>(0.076)</td>
</tr>
<tr>
<td>Other electrical, computer, instruments</td>
<td>0.031</td>
<td>(0.187)</td>
</tr>
<tr>
<td>Transportation, aircraft</td>
<td>1.095***</td>
<td>(0.275)</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.213*</td>
<td>(0.674)</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-1282.15</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>707</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Industry-specific 1988+ dummy coefficients are shown. Estimation is by Poisson random effects. Regression also includes industry dummies significant at 10 percent.
Bibliography


Arora. Ashish. Andrea Fosfuri and Alfonso Gambardella (1999) "Markets for technology: Why do we see them, why don’t we see more of them, and why we should care." mimeo. Carnegie Mellon University. Heinz School of Public Policy and Management.


BIBLIOGRAPHY


Bibliography


