The positive effect of entry of rivals on incumbents' profits: the see-saw effect

Gherardo Gino Giuseppe Girardi

Thesis submitted for the degree of Ph.D. in Economics
(second version including changes specified by examiners)

London School of Economics and Political Science

Abstract

The thesis is a theoretical study of different mechanisms by which an increase in the number of rival firms leads to an increase in profits of an incumbent. Consequently, an incumbent may find it profitable to invite rivals, as is sometimes observed, for example in the semiconductor industry.

Chapter 2: We expand the work by Farrell and Gallini (1988), in which a monopoly invites rivals to commit in period1 to low prices in period 2. Instead of assuming Bertrand competition as Farrell and Gallini do, we assume Cournot competition, and find the optimal number of rivals from an incumbent's point of view. Furthermore, the incumbent may invite rivals to enter already in period 1 so as to 'share' any losses incurred in the first period of production.

Chapter 3: We propose a new mechanism by which an incumbent firm may want to invite rivals, based on the existence of capacity constraints, which we model as decreasing returns to scale. An incumbent may attract rivals in period 2 to increase consumer surplus in the same period above the level which it finds cost-effective to achieve on its own, enabling consumers to tolerate a higher price in period 1.

Chapter 4: A multinational which trains workers may enjoy higher profits when more of its workers are poached by rival firms if it is compensated by a sufficiently lower wage during training. More workers are poached when there are more non-training firms and when training is more general. The result depends on the assumption of decreasing returns to scale, so that an increase in the number of poaching firms, or training that is more general, expands worker surplus beyond the level that the multinational could achieve on its own. The multinational can then pay trainees a lower wage.

Contents

List of figures (diagrams and tables)

Acknowledgements

I am mainly thankful to God for enabling me to come closer to Him whilst writing this thesis, which I believe is why He made it possible for me to write it. I am also grateful to all those members of the celestial family who provided assistance, including our daughter Constanza.

I would like to thank Professor Anthony Venables for his supervision. I would also like to thank several members of staff in the Economics Department at the LSE for their suggestions: Professor Alan Manning, Professor Charles Bean, Professor Steve Pischke, Dr Mark Schankerman and Professor John Sutton. Among the research students at the LSE, I would like to thank in particular Thomas Buettner, Leo Ferraris, Paloma Lopez-Garcia, Barbara Petrongolo, Kwok Tong Soo and Roberto Veneziani. I am grateful to all government officials, trade union officials and industrialists in Mexico with whom I had informal talks. I would also like to thank the ESRC for providing full financial support. My family and friends, in particular my wife Ana Elvia, my parents and Chris Tandy also helped me. Our newly born Santiago helped with his presence and cheerfulness.

On the ethics of this thesis: a warning

Modern economic theory is built on the assumptions that people are rational, meaning that they take decisions by comparing costs and benefits, and self-interested. Broadly speaking, I agree with these assumptions. However, I also believe that most individuals, including those acting in this thesis, are deceived: they do not understand what is best for them, although they believe that they do. They, or rather we, since I often find myself behaving as they do, believe that maximizing our economic surplus without considering the effects that this has on other people benefits us. This self-interest is, I am convinced, a form of immaturity, reflecting our inability to see that it is the common good which we ought to be primarily concerned with, and that there is no conflict at all between helping one's neighbour and one's self interest, rather only complementarity. The accuracy of this statement cannot be mathematically demonstrated; rather, it can be verified only through individual experience, normally over the course of a number of years.

The individuals who act in this thesis are fortunate, however, on those occasions when the self-interest outcome coincides with the common good outcome. Sometimes, when a monopolist decides to invite a rival, all agents weakly benefit, that is, all firms and consumers either benefit or are at least unaffected. Thus, the negative consequences of the agents' myopia are at times lessened. This offers some consolation for the author, who does not wish to deceive readers into believing that it is good to pursue one's self-interest, from anyone's point of view.

Things are, however, much more serious than that. The monopolist who invites rivals does so with the aim of maximizing profits. In the thesis, when a monopolist's profits increase, we say that it benefits, or enjoys higher profits. These are highly contestable statements. If wealth accumulation is not used for the sake of benefiting others, the owner or owners of monopolist will be hurt. For example, he may come to rely on financial success in order to feel secure. His conscience, whether or not he is aware of it, will be burdened. His egocentric vision will prevent him from extending beyond himself and to experience freedom from worries and fear. It is possible for him to be stuck in a bad equilibrium in which he never discovers what is truly advantageous for him.

I would like to apologize for adding these considerations as an appendix, rather than giving them priority by embodying them into the thesis. The reason for this is that, from a professional point of view, I am just beginning to explore these issues. Also, challenging the mainstream wisdom requires courage and hard work, and I could do with a bit more of both.

A final point I would like to make is that the work in this thesis is the product of the modern approach to economics, in which most of us are trained, and which artificially separates scientific from ethical considerations, reflecting yet another myopia. It relegates the discussion of ethics to a secondary position, in this way suppressing a characteristic of man which distinguishes him from animals, and making the subject rather lifeless.

Chapter 1

Introduction

1 Overview

Why do we observe firms inviting rivals without asking for compensation such as royalties?

Why do firms design so-called open architecture products that are easy to imitate?

Why do firms provide general training when there is a significant probability that trained workers will be poached by rivals?

These are some of the key questions addressed in this thesis. We study 2-period games in which an incumbent's profits increase when the number of rivals increases. The first two chapters are on industrial organization: we focus on the practice of second-sourcing of production, and consider an incumbent's incentive to invite rivals. The third chapter is on labour economics: we consider how the incentive to train workers is affected by the presence of rival firms which poach trained workers, and how a higher degree of generality of training affects the incentive to train.

In all three chapters, we find that there is a price movement in the second period which hurts the incumbent, but there is a price movement in the first period in the opposite direction which more than compensates for that in the second, so that the incumbent is on the whole better off. We call the inverse relationship between prices in different periods the see-saw effect.

Although the models of all the chapters display the see-saw effect, those in chapters 2 and 3 do so for different reasons, and that in chapter 4 considers a novel application. In chapter 2, an incumbent has an incentive to invite rivals so as to commit in period 1 to a large quantity at a low price in period 2. In chapter 3, an incumbent may invite rivals so as to increase industry capacity, which is captured by the assumption of decreasing returns to scale. In chapter 4, the decreasing returns to scale argument is applied to the labour market.

Before we begin to discuss these theories, we briefly describe two other mechanisms present in the literature which explain how more competition may increase profits of individual firms, both pertaining to the study of firms location.

In economic geography, if an increase in the number of firms at a given location raises the return to other firms, then agglomeration of economic activity in one area will occur. As Venables (1996) shows, this may occur in industries that are vertically linked through an input-output structure: in the presence of trade costs, upstream firms locate where there are lots of downstream firms so as to enjoy higher demand. Furthermore, downstream firms enjoy cost advantages by locating where there are lots of upstream firms. Putting these forces – known as linkages - together generates a force for agglomeration in a single location.

Another plausible explanation as to why firms may benefit from the presence of rivals is found in Schulz and Stahl (1996). With consumers uncertain about the characteristics of the products on offer and having to incur costly searching, a firm which locates near other firms will attract more consumers to the cluster by enabling them to take advantage of economies of scope in searching. The standard competition effect is more than offset by the increase in total demand. This is the typical trade-off found in the economics of networks.

We now consider separately the see-saw effect in the context of industrial organization, with particular application to the practise of second-sourcing (chapters 2 and 3), and in the context of labour economics, with particular application to training (chapter 4). We describe the phenomena that the theories try to explain, set out the main questions to be addressed, provide intuitions of the main results of each chapter, discuss the key assumptions, and review the literature.

2 Industrial organization: why do incumbents invite rivals to enter? The practice of second-sourcing in the semiconductor industry

The practice of second-sourcing involves two agents, a so-called original source supplier and a second-source supplier.

The original-source supplier is a supplier that produces a new product[1]. The second-source supplier is any manufacturer that produces the same product. The second-source's product need not be identical, but it must be technologically interchangeable, meaning that it must have the same form, fit and function (Rosenblaum (1985)). Thus, given that the products are interchangeable, the buyer is indifferent between the two products, at least from a technological point of view. The original and second-source producer(s) compete on the same market.

Since a great deal of the empirical work on second-sourcing is about the semiconductor industry, that is the industry on which we shall focus in order to draw examples. The industry is characterized by the presence of many firms which copy existing products. It is well established that the cost of imitation in the semiconductor industry is significantly less than the cost of innovation. In the mid 70's, it was estimated that the cost of developing a second-source copy of a microprocessor, was perhaps only 10% - 20% of the cost developing the original product (Swann(1987), who quotes from Electronics, 15 May 1975, p74).

The second-source supplier may be licensed under the original supplier's patent, or may be unlicensed. Sometimes the original source supplier will approach second-source producers and invite them to produce the same product; sometimes a second-source producer will have started to produce the product before contact with the original-source supplier is established, by engaging in reverse-engineering[2]. In the 1980's at least, generally speaking companies could proceed to make copies in the confident knowledge that they would not be infringing copyright laws. The original source producer would then approach them and propose to them a retrospective license.

It appears that many of the second-source producers which began producing without the explicit consent of original sources nevertheless enjoyed their tacit consent (Swann(1987)). Thus, it appears that benefits have accrued to original source producers even when they did not explicitly invite second-source suppliers. The phenomenon that is more striking, however, is that of incumbents inviting rivals, in which case the (perceived, at least) benefit of facing competition is not evident. Original source producers court some potential second source producers and not others, so that some

[1] Usually the original source is also the developer of the product, but this is not relevant to our analysis.

[2] Those who are opposed to reverse-engineering call this process 'piracy'. The fact that the former term, rather than the second, is very wide-spread in the literature on second-sourcing suggests that commentators' overall impression of this activity is positive.

imitators enter uninvited (Swann(1987)). To those who accept the offer, the original source gives the masks with which to manufacture the particular design. Uninvited firms, on the other, are given copying rights, but not details of the technology to produce the particular design: they obviously already have a technology to do so, albeit not necessarily the same as that of the original source producer.

The data on second sourcing is most complete for the period between the mid 70's and the mid 80's, since this is when the phenomenon really took off. Consider, as an example, microprocessors. The figures below give an idea of the increase of the phenomenon of second-sourcing of original designs of microprocessors.

Let us begin by considering the strategies of producers, in particular how many firms produced their own designs only, how many second-sourced only, and how many did both:

Figure 1.1: Strategies employed by microprocessor producers in different years

	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985
Own design only	10	11	14	10	10	8	6	7	6	4	4	5
Second source only	3	5	13	14	15	16	18	22	25	27	26	32
Both	0	5	4	10	11	10	13	10	12	10	10	9
Total	13	21	31	34	36	34	37	39	43	41	40	46

Source: Swann(1987), who compiles the table from EDN and Electronic Design, various issues, and from Swann (1985)

The most striking features of the table are the increase in producers acting purely as second source suppliers over the period 1974-1985, and, perhaps to a lesser extent, the fall in the number of producers supplying their own designs only over the same period. This suggests that firms have become more specialized over time, with a few firms developing and supplying their product on their own, and most firms copying.

Next, let us shift our focus away from the number of producers engaging in second-sourcing to the number of products that were second-sourced:

Figure 1.2: Microprocessor products available in different years

	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985
Original designs	12	21	32	39	43	43	44	43	47	43	40	40
Designs with a second source	3	7	11	17	22	31	29	33	35	34	31	30
Second source products (total)	3	11	20	41	51	69	83	103	117	111	107	109

Source: As for Figure 1.1

This second table shows that, over the period 1974-1985, the proportion of original designs that were second-sourced grew significantly, from 25% (3 out of 12) in 1974 to 75% (30 out of 40) on 1985. The total number of second source products grew enormously, from 3 to 109. These trends confirm those found in the previous table.

Why would firms which are in possession of original designs wish to second-source? The literature on this topic stresses a number of possible reasons[3]:

1) To obtain access to rivals' technologies as part of what is known as a cross-license agreement[4]. Companies may own thousands of patents, used in tens of thousands of products, and add hundreds each year. Companies cross-license portfolios of patents in a field-of-use, without making reference to specific patents, since it would be cumbersome and costly to license individual patents (Grindley and Teece (1987)). This is an important reason for second-sourcing, and may indeed account for the modest levels of royalty payments (we may think of the two suppliers as each paying the other equal amounts).

2) To obtain royalty payments. These were very low until the mid-80's (Rosenblaum(1995)), and then increased significantly (Grindley and Teece (1997)). It would appear that this increase is due to the introduction of the US Semiconductor Chip Protection Act in 1984[5], which signalled a move by the judicial system in favour of more strict upholding of patent holders' rights[6]. At the same time, Texas Instruments stood out from the pack by becoming more aggressive in seeking market returns on their intellectual properties, causing an uproar as its strategy shifted away from what had been the general 'mindset' of the industry until then (Grindley and Teece (1997))[7].

---

[3] There is a further reason, specific to military procurement, and not applicable to our analysis. Here, following the design of a new technology by a firm, the government, acting as a monopsonist, conducts an auction, in which the original developer of the technology also participates, to decide which firm will produce the product in question. It is believed that the auction pushes the price paid by the government down to its true, private cost (Anton and Yao (1987)).

[4] The purpose of cross-licenses is usually simply to gain legal access to a rival's patented technology, rather than to transfer technology.

[5] The shift in government policy is attributable to the perception that the American economy was becoming more vulnerable to foreign, and especially Japanese, competition. The previous policy of permitting access to foreigners of US R&D knowledge in exchange for very low royalty fees was deemed to place American producers at a disadvantage relative to foreign firms, whose ability to exploit advances in US R&D laboratories had increased (Hall and Ham (1999)). Interestingly, the motivations for inviting rivals explored in this thesis suggest that such low royalty fees might have been chosen by American producers to attract rivals into entering.

[6] In the context of second-sourcing, it is possible that in a weak copyrights environment more imitators entered than innovators had wished.

[7] Grindley and Teece (1997) consider the significant increase in the number of patents taken out since the mid 80's. They argue that companies previously patented less because product life cycles were shorter and so would simply not bother to patent inventions, since products would quickly become obsolete.

3) To insure buyers against opportunistic behaviour by a monopolist after buyers sink a product-specific fixed cost[8]. In other words, after buyers have incurred such a cost, a monopolist may impose a high price (and possibly a low quantity). When Xerox developed the local-area-network product Ethernet, it offered open licenses at a nominal price. According to Farrell and Gallini (1988), had Xerox not opened up its technology, semiconductor firms might not have made the specific investment to develop the dedicated semiconductor chip. This is because semiconductor firms may have feared that customers would have been "reluctant to buy such a chip for fear that Xerox would [have] behave[d] opportunistically and cut off the supply sometime in the future, or change[d] exorbitant prices once the chip had proved successful" (quotes from Sirbu and Hughes (1988)). Once assured that Xerox would license openly, Intel and others developed the dedicated chip[9].

4) To insure the buyer against production shocks affecting microchip producers. As microchips became more sophisticated in the 70's and 80's, the probability that something went wrong in the production process and that a whole batch of them had to be discarded increased. Ensuring that there are more than one producer insures the buyer against the possibility that he would not be able to obtain a sufficiently large number of chips[10] (UN (1986)).

5) To signal to the buyer that the microchip is or will become an industry standard, with the associated benefits of network externalities. There is a vast literature on this topic, pioneered by economists like Katz and Schapiro (see Katz and Schapiro (1986)). Products which attract second sources early in their life-cycle will be perceived as likely industry standards, and will tend to continue to attract further second source producers.

6) To increase the demand for 'supporting products' produced by the original source. Often, microchips come in groups whose members are in some way technologically complementary, so purchasing one microchip reduces the investment cost of another microchip in the same group (UN (1986)). Swann (1987) finds evidence suggesting that original sources will only authorize second sources to produce a subset of the range of products within such groups of products.

---

[8] It is assumed that no other credible mechanism to commit ex ante to low prices in the future is available, such as long term contracts.

[9] Dutta and John (1995) find evidence in favour of the Farrell and Gallini (1988) findings. By setting up a laboratory experiment, they show that competition from a licensee suppresses the price increase that a monopolist could impose so as to exploit locked-in customers.

[10] This explanation assumes that production shocks among producers are (sufficiently) independent of each other.

7) To increase industry capacity. Quite simply, buyers may require a minimum amount of produce to use as input in their own production, and original source suppliers may not have the required capacity. Dick (1992) considers firms which must choose capacity before uncertainty is resolved. In his model, firms sign a contract: whichever firm successfully innovates, it second sources production at a pre-agreed price. This enables each firm to select an ex-ante level of capacity that is lower and more profitable than that without contracting.

8) To comply with regulation, in particular the 1956 AT&T and IBM decrees, that obliged these firms to divulge a whole range of basic semiconductor and telecommunications technology for next to nothing to domestic and foreign firms[11]. However, the effect of these degrees was merely to formalize policies that were already at least in part being implemented by the two firms (Grindley and Teece (1997)).

This thesis is concerned with explanations (3), (6) and (7) in the list above, namely, the commitment problem faced by a monopolist, and the complementary nature of technologically advanced products.

In chapter 2, we expand the model of Farrell and Gallini (see item (3) in the list above), replacing their assumption of Bertrand competition with Cournot competition. We can thus identify how many rivals a monopoly incumbent will invite in period 2. We also find that the incumbent may invite rivals to enter already in period 1, to 'share' losses in that period. This constitutes a significant development in the light of Farrell and Gallini(1988)'s statement, that "Obviously, second-sourcing that enables other firms to compete with the innovator on equal terms immediately (and without royalty payments) is not profitable [for the innovator]". This statement is true given their assumption of Bertrand competition, but we show that it is false in a Cournot competition environment. Multiple equilibria due to self-fulfilling expectations are possible whereby, if entrant firms expect that not enough buyers will enter, then they will not enter, and this will cause their expectations to be confirmed.

As Farrell and Gallini (1988) point out, the set-up costs incurred by buyers effectively makes the good sold in period 1 complementary with the same good sold in period 2, in the sense that a higher expected second period price reduces first-period demand. We

[11] Grindley and Teece (1997) believe that "there is no doubt that [such a liberal policy] provided a tremendous contribution to world welfare. It remains as one of the most unheralded contributions to economic development – possibly far exceeding the Marshall Plan in terms of the wealth generation capability it established abroad and in the United States."

can therefore apply the model to a whole range of complementary goods, such as computer hardware and software, video game consoles and video game cartridges, printers and ink cartridges, video recorders and video cassettes, DVD's and DVD players, etc.. In the case of the semiconductor industry, we can apply the model to complementary products within the same group, as in (6) above. An example is support microchips that can be used with a variety of microprocessors (Swann(1987)).

In chapter 3, we propose a new mechanism by which an incumbent may want to invite rivals in period 2, based on the incumbent's desire to relax the capacity constraint at the industry level; this increases consumer surplus in period 2 and consumers are therefore more willing to tolerate a higher price in period 1. We model capacity constraint as decreasing returns to scale, which, as Tirole (1988) points out, is a smooth of form of capacity constraints. We show that an incumbent invites rivals if in period 2 price is less than its marginal cost. This model is therefore concerned with item (7) in the list above, and also with item (6) by virtue of its implications for markets for complementary products.

The models of chapter 2 and chapter 3 address not only the question of whether or not an incumbent has an incentive to invite rivals, but also the question of how many. Swann (1987) states that, paradoxically, it may be in the original producer's interest to attract some second source producers, but only some, for beyond a certain point the original firm finds itself against "unwanted competition". The models the chapters 2 and 3 identify when this critical point is reached.

Rosenblaum (1985) and Swann (1987) have suggested that some second-sourcing agreements are little more than an attempt by original sources to get whatever little royalty payments they can obtain from firms which have already entered, copied and started producing the original source's product. They cite as supporting evidence weak copyright laws, especially prior to the introduction of the US Semiconductor Chip Protection Act of 1984[12], and the high mobility of technical personnel between firms, which may well reduce the innovator's ability to prevent its technology from being divulged.

[12] The US Semiconductor Chip Protection Act, whilst prohibiting the practice of dissecting a chip to reproduce the masks that are used as tooling for producing a competitive product, permitted the practice of using the dissection information in the independent development of an equally competitive product. What was required of the imitating firm engaging in reverse engineering was that it left behind a 'paper trail' evidencing independent effort, including computer simulations and time records (Swann (1987)).

This evidence may support the models of chapters 2 and 3, in which the innovator wants its technology to be divulged, even freely. Weak copyright laws and high mobility of personnel may help to persuade buyers that there will be many firms copying the product. This would fulfil the information requirement of the models, whereby buyers are informed in advance of whether firms enter or not in the second period. The point of these models is not that original source producers seek royalties; on the contrary, it is to show that innovators may benefit from entry of rivals even without receiving any royalties.

How realistic are the main assumptions of the models in chapter 2 and 3, with particular reference to the semiconductor industry? There are 3 fundamental assumptions which we make: (a) The informational requirement that all agents, and buyers in particular, accurately predict the number of suppliers of the product in question; (b) The assumption that set-up cost faced by buyers be sufficiently large; and (c) The assumption that entry of rivals be observed to take place in both periods, that is, both at the time the product is launched, and afterwards.

Two factor justification the assumption that all agents, and buyers in particular who are concerned about high prices and low quantities in period 2, are well informed from the start. The first is that it is in the interest of original and second-source producers to ensure that their agreements are made evident to buyers. Consider the case of Xerox. As we mentioned above, had Xerox not opened up its technology, semiconductor firms might not have made the specific investment to develop the dedicated semiconductor chip for fear that customers would have been "reluctant to buy such a chip for fear that Xerox would [have] behave[d] opportunistically and cut off the supply sometime in the future, or change[d] exorbitant prices once the chip had proved successful". This argument, found in Farrell and Gallini (1988), strongly suggests that Xerox wanted all parties involved, and customers in particular, to know about its decision to open up its technology. The second factor is that the practice of second-sourcing was very well established in the semiconductor industry in the seventies and eighties, so that buyers could expect with a high probability that an original product would be second-sourced.

The second critical assumption in the two models is the fact that microchip buyers face significant investment costs in purchasing a new microchip. An example of a product with such high set-up costs incurred by buyers is microprocessors. According to Dutta and John (1995), microprocessors involve a heavy commitment of resources to design hardware and software around the specific requirements of the particular microprocessor. They point to estimates (reported in Electronic News, 1980) suggesting that the choice of any particular 16-bit microprocessor family in a system is usually a 5 to 10 year commitment to that family and entails an irrecoverable investment of between 10 and 20 million dollars.

The third key assumption refers to the timing of the entry of second-source producers. In chapter 2, we require that they be able to enter in either period, whilst in chapter 3 we require that they be able to enter in period 2. The empirical literature indicates that there is a lot of variation in the timing of entry.

Farrell and Gallini (1988) describe the practise of delayed second-sourcing as 'common'. Still on delayed entry, Swann (1987) says "It is common for new entrants in a particular segment to act as "second source" and copy leading products [...]. There are obvious advantages of this for the new entrant – principally that he enjoys much reduced development costs and that there are in *already in existence* [our emphasis] a network of supporting products and software that can be used in conjunction with the well established designs". At the same time, we also observe second-source agreements in which the original and second source producers enter into an agreement prior to the launch of the product (Dick 1992).

We assume in chapter 2 that contacts cannot be signed, perhaps because they are not enforceable, and focus instead on arm-length market transactions. If the monopoly could sign a contract committing it to a large quantity and a low price in the future, it would do so, removing the incentive to attract rivals. Indeed, because inviting rivals would involve a loss in market share, it would always be optimal for the incumbent to sign such contracts. In chapter 3, on the other hand, we assume that such contracts can be signed, so as to focus on a different motivation for inviting rivals, namely to overcome an incumbent's capacity constraint.

A simplifying assumption, which we make in all three chapters, is that time is not discounted. If it was, then the see-saw effect would be quantitatively smaller, though qualitatively unaffected: future price movements due to more competition tomorrow would be less significant today, so that any offsetting price change today would also be less significant. Finally, we simplify the analysis further by disregarding mixed strategy equilibria, concentrating instead on pure strategy equilibria.

Having discussed the empirical literature on second-sourcing, and the extent to which the assumptions of the models of chapters 2 and 3 of this thesis are consistent with them, we now summarize the relevant theoretical literature.

The most important paper as far as this thesis is concerned, and on to which we have referred to already, is that published in 1988 by Joseph Farrell and Nancy Gallini in the Quarterly Journal of Economics. They showed, in a Bertrand competition framework, that a monopoly can overcome its inability to commit to low prices by inviting rivals. We have already spoken about this paper in the discussion above, so we proceed to discuss other relevant literature.

In Shephard (1987), an innovator licenses rivals in order to commit to high product quality, making it quite similar to Farrell and Gallini (1988). Given that buyers' surplus increase with quality, quality plays a similar role to that which price does in Farrell and Gallini(1988), but with the opposite sign. Shephard (1987), unlike Farrell and Gallini (1988), assumes that contracts are available to keep competition in check and to redistribute profits back to the patent-holder.

Kende has a number of models describing the interaction between hardware and software provision which are driven by the commitment mechanism of Farrell and Gallini(1988). In Kende (1995a), two common strategies used by monopoly sellers of hardware are examined. The first is licensing of the hardware technology to third-party firms. This strategy is profitable because (a) it enables the hardware manufacturer to commit to low prices in the future, and (b) licensing generates royalties expressed as a percentage of revenues per unit sold. For sufficiently low royalty fees, the final price of the hardware will be low. Because hardware and software are complements, demand for software increases, so that software firms have an incentive to produce software. This in turn raises demand for the hardware, which more than compensates the incumbent for the loss in market share and the reduction in price. The trade-off here, which is generally found in discussions of networks and of competition between standards, is between a higher total demand on the one hand, versus a lower market share on the other hand. This trade-off is also present in all models of this thesis. The second strategy in Kende (1995a) involves the hardware producer integrating into software production, thus unilaterally committing to the provision of software.
Comparing Kende (1995a)'s model with chapter 2 of this thesis, the original source supplier of chapter 2 of this thesis is both a hardware and a software producer, and licenses either software production only, or both hardware and software production.

Comparing Kende (1995a) with that of chapter 3 of this thesis, the original source supplier of chapter 3 of this thesis is both a hardware and a software producer, and licenses software production only.

In Kende (1995b), a hardware manufacturer again licenses its technology to second-sources in order to commit to lower hardware prices, so as to stimulate software production and thus hardware sales. This time, Kende considers a battle between 2 standards. He finds that low software costs make licensing of hardware profitable, since lots of software will be forthcoming.

In Kende (1998), a monopolist selling a system consisting of a main component (e.g. a video game console) and a range of differentiated secondary components (e.g. video games cartridges) can enjoy increased profits by allowing competition in the aftermarket for the secondary components, again for the reason that the presence of competitors enables the monopolist to credibly commit ex ante to low aftermarket prices. Kende shows that the lower the elasticity of demand for the system as a whole, the more profitable it is to keep market power over all components, since the fall in price of secondary components caused by opening the system will not generate a sufficiently large inflow of buyers. In chapters 2 and 3 of this thesis, we have a similar result, namely that the faster the set-up cost of buyers increases with the number of buyers, the less the increase in the number of buyers when period 2 prices fall as a result of increased competition, and therefore the lower the incentive to invite rivals.

Still in Kende (1988), the more substitutable are secondary components, the greater the incentive to keep the system closed, since allowing a competitor to sell its own brand of secondary component would take too much demand away from the incumbent's own secondary components. Finally, the greater the share of the budget spent on the secondary components, the lower the incentive to invite rivals.

Lichtman (2000) develops a theoretical model showing that profits of all firms, both of a platform product and of peripheral devices, could increase if agents co-ordinated and lowered the prices of peripherals. This is because lowering the price of a peripheral has a positive effect on demand for both the platform product and for the other peripherals. Lichtman goes on to argue intellectual property law should facilitate price co-ordination in emerging technologies, since this would benefit both producers and consumers.

Conner (1995) shows that the entry of clones of lower quality may be valuable assets to an innovating firm. Suppose that there are positive network externalities, so that consumers value a product more the greater the number of other consumers who use it. Clones making a product of lower quality compared to the innovator's product would bring lower-valuing consumers into the market, which in turn would make the innovator's product more valuable to medium- and high-valuing consumers, increasing the user base for the innovator's technology. Conner uses the example of Sun Microsystems, which encouraged cloning of its computer workstations (see Business Week, July 24 1989), to illustrate this possibility, and contrasts it with IBM's decision when launching the PS/2 computer line to give up compatibility with the older PC models, thus making cloning of the new line less attractive to potential cloners.

An economic geography paper incorporating the same principle as that of Farrell and Gallini (1988) is that of Rotenberg and Saloner (2000). They explain regional agglomeration of production and trade stemming, in part, from the fact that competition for trained workers between firms located near to each other: in doing so, they prevent the existence of a monopsony, which would pay a low wage after workers have trained, and thus discourage workers from undertaking industry-specific human capital investment. As more firms gather, and the after-training wage increases, workers are more willing to incur costly training, resulting in the emergence of the industry.

Just like entry of rivals would appear to reduce profits but in fact may do the opposite and increase them, either because of Farrell and Gallini (1988)'s commitment explanation or because of network effects, software piracy may only apparently be damaging to software producers. Slive and Bernhardt (1998)[13] argue that, in the presence of significant network externalities, it may be optimal for software manufacturers to tolerate piracy by home consumers, whose willingness to pay is very low, and who are more difficult to catch and punish than business buyers. Network externalities could be generated by the fact that knowledge which an individual user accumulates at home can be costlessly transferred between home and the work environment; at work, a pirate's knowledge can then be applied to the legal software market.

[13] Conner and Rumelt (1991) develop a similar model in which some customers who pirate would not buy if protection of patent rights was strengthened. Thus, tolerating piracy increases the total number of network users.

Whilst the literature and networks and complementary products is huge[14], the work by Economides (1996a) is particularly relevant in that he discusses invitations extended by incumbents to rivals to enter. Economides (1996a) builds a model with direct network externalities[15], so that the price that a buyer is willing to pay is increasing with the number of buyers. He shows that if the network effect is strong enough, it is in the interest of a quantity leader to invite rivals to enter and license his technology without charge. The reason is that suggested by Farrell and Gallini (1988), namely in order to overcome his inability to commit in the first period to low prices in the second. If the network externality is very strong, the incumbent may even find it profitable to subsidize entry by paying royalty fees to entrants.

There is a close connection between the model of Farrell and Gallini (1988) and the models pertaining to the literature on switching costs (see Klemperer(1995) for a survey). This connection is not surprising, since these models consider the cost faced by buyers of locking into a particular brand, whilst Farrell and Gallini (1988) consider the cost of locking into a particular product, which may be produced by more than one firm. In both cases, we observe an inverse relationship between prices in different periods; however, the period in which competition occurs differs. In the switching cost literature, the higher the second-period price that a monopoly can charge, the lower the first period price will be as result of competition in that period to acquire the monopoly. In Farrell and Gallini (1988) and in the model of chapter 2, the greater competition is in period 2, the lower the price in period 2, the higher the monopolist's price which consumers tolerate in period 1.

The result obtained in the switching cost literature that the first period price is lower due to switching costs depends, as in Farrell and Gallini (1988), on the assumption that sellers cannot commit to a low second-period price. If they could commit, the model would collapse to a 1 period model with Bertrand competition (Varian (2000)).

---

[14] Two surveys are Economides (1996b) and Holler and Thisse (1996).

[15] Economides (1996a) provides an indirect network interpretation of his model, with 2 complementary goods, making it comparable to the model of chapter 2 of this thesis. There is, however, an important difference between the two models. In Economides (1996a), each firm in market 1 chooses its optimal quantity taking as given the output of firms in market 2. In chapter 2 of this thesis, on the other hand, firms recognize that their choice of quantity in period 1 (market 1 in the terminology of Economides(1996a)) affects the number of buyers who enter in period 1 and who also buy in period 2 (market 2 in the terminology of Economides(1996a)); thus the second period quantity of each firm is a function of each firm's quantity in period 1. We believe that allowing firms in one market to take into account the effect of their actions on the number of buyers in another market is realistic in view of the complementary nature of the two markets and adds value to the analysis.

There are a number of papers which study the effect of market power in proprietary aftermarkets (see Shapiro (1995), Shapiro and Teece (1994), and Borenstein at al. (1995,1996)). As Kende (1998) explains, this literature is motivated by antitrust cases such as Eastman Kodak Co and Image Technical services. Here, the Supreme Court ruled that Kodak could have market power in the aftermarket even though it had none in the equipment market. These papers show that market power in the aftermarket is exploited in so far as the lock-in effect whereby buyers are locked into specific brands dominates the reputation effect which affects new customers.

Having discussed the role that competition plays in increasing profits of incumbents in an industrial organization environment, let us now discuss the role of competition in the context of labour economics, in particular with application to the issue of training.

3 Labour economics: why does poaching of trained workers not necessarily deter training?

We wish to explore how the see-saw effect might operate in the labour market. For this purpose, we consider how increased competition for trained workers affects the incentive to train workers. The simplest explanation as to why some firms may poach trained workers is that they lack the necessary training facilities to do their own in-house training. This scenario seems most plausible in the context of developing nations, where multinationals may train workers which cannot be trained by local firms, because they simply do not have the know-how which multinationals have and pass on to workers. The model of chapter 4 studies precisely this situation. Let us begin by considering the empirical literature on the interaction between training of local workers by multinationals in developing nations and poaching by local firms.

The body of empirical work on the impact of multinationals on LDC development via training of local labour suggests that it is broadly positive. Gershenberg's studies of Kenya (1987) and Jamaica (1990) show that multinationals provide more training than local private firms (though not more than public-owned firms). Numerous other studies confirm the finding that multinationals offer more and better training than local firms: Goncalves (1986) for Brazil; Iyanda and Bello (1979) for Nigeria; Yong (1988) and Sibunruang and Brimble (1988) for Thailand; see also Chen (1983). Grosse (1992), working for the UNCTC, reports that the impact of multinationals on LDC's is most marked in the training of skilled personnel[16]. Walkenhorst (2000) finds that foreign-owned sugar producers in Poland and Slovakia in the post-communist transition period focused heavily on technical training of factory personnel and sugar beet growers "during the early stages of their engagement".

Furthermore, Pack (1993) reports that labour mobility in Korea from multinational to local firms is significant and that it is not unusual for managers trained by multinationals to leave these and set up their own business. These positive assessments of the impact of multinationals on local firms is consistent with a substantial body of evidence suggesting that local firms located next to multinationals enjoy productivity spillovers (see Blömstrom and Kokko(1998) for a review of the literature)[17].

However, in the same studies mentioned above, Gershenberg finds that "a much smaller proportion of all managers trained in multinational firms move to non-multinationals than is true of managers employed by other kinds of firms". Gershenberg's work therefore casts some doubt on the "'spread effect' (the movement of trained managers from multinational to local firms) which is generally assumed to constitute the most significant contribution of multinational firms to the development of an indigenous cadre of managers".

In chapter 4, we present a theory of economic development which explains how an LDC may benefit when local firms poach workers trained by a multinational, the latter being the only source of foreign technology. At the same time, we wish to study the implications of the see-saw effect, the inverse relationship between prices in different periods, when applied to the labour market. The theory is more general than its application suggests; indeed, it could be regarded as a theory of apprenticeships. However, casting it in an LDC setting gives it concreteness and seems particularly apt in that the model assumes that some firms cannot train. It makes sense to think of these firms as local firms which do not have the technology of the multinational.

[16] For a brief review, see Dunning (1992).
[17] See Noorbakhsh, Paloni and Youssef (2001) for a brief survey of the evidence suggesting that the skill and education level of workers affects the volume of FDI inflows and the activities undertaken by multinational firms in a country.

Let us begin by giving an overview of the theory and of the intuition behind the main result. We consider a multinational[18] which decides to produce in an LDC, bringing with it new technology that is of use to local LDC firms. This technology is embodied in the workers by way of training in the first period, and is transferred to local firms when workers are poached in the second period. Local firms are assumed not to possess the multinational's training technology. This is a reasonable assumption given that training firms in developing nations are often unable to train adequately, as Katz and Ziderman(1999) explain. Where enterprises are small and trade associations underdeveloped, firms may not lack the required managerial ability, and, if there are increasing returns to scale in training, it may not be able to finance the required investment in training capacity[19].

The most important result in this paper is that the multinational's profits rise when competition for its trained workers increases. The increased competition may be due either to training being of a more general nature (i.e. its usefulness to local firms is greater), or to there being more non-training firms demanding the services of trained workers. Consider an increase in the generality of training. When we say that training is more general, we mean either that the productivity of a poached worker in local firms is higher, or that there are more non-training firms which can benefit from poaching trained workers[20]. As a consequence, when training is more general, local firms attract more labour from the multinational. Contrary to standard intuition, the multinational's profits increase.

This result hinges on the multinational's ability to lower the first period wage to the lowest possible level consistent with workers leaving the agricultural sector in the first period to work in industry; in particular, whenever the second period wage increases, the multinational can recoup this cost by reducing the first period wage (the see-saw effect), maintaining the sum of wages over two periods which workers demand in industry at a level equal to the sum of wages in agriculture. We show that the multinational is more than compensated by the reduction in the first period wage, so that its profits over both

[18] There is nothing in the model which obliges us to interpret this firm as a multinational, such as the existence of a branch (headquarters or subsidiary) in another country. Thus the theory is more general than the example suggests.

[19] Katz and Ziderman (1999) suggest that firms in LDC's may lack 'managerial foresight' to organize in-house training. Whilst training certainly requires the behaviour of the agents involved to be forward-looking, it might be that employers and/or employees in developing nations discount time more than those in developed nations, making training unattractive.

[20] In the terminology of Stevens (1994), it is the degree of transferability that we are allowing to vary, general training being an extreme case of training that it is entirely transferable.

periods increase. A similar argument applies in the case of an increase in the number of local firms.

The result that more general training results in higher profits of the incumbent does not require that in the first period there be only one incumbent. Indeed, there could be many incumbents, as we show in the section where we relax the assumptions. Thus the model of chapter 4 provides an extension of the classic, benchmark model of diminishing returns (or decreasing returns to scale with only one input in production) from a 1 period setting to a 2 period setting, and shows that in so doing it becomes possible for profits of an incumbents firm to increase when the number of rival firms increases.

We also show that, when firms face a fixed cost of entry, multiple equilibria due to self-fulfilling expectations are possible, whereby if period 2 entrant firms believe in period 1 that there will not be enough workers available in period 2, then they will not enter, the multinational firm in period 1 will train less workers, and the expectations of the entrants will be confirmed. This result is analogous to those in chapters 2 and 3.

The result that profits of the multinational rise with competition for trained workers depends on the assumption of decreasing returns to scale. This is because an increase in the number of poaching firms in period 2, or training that is more general, expands worker surplus in that period beyond the level that the multinational would fid cost-effective to achieve on its own. The multinational can then pay workers a lower wage in period 1. Chapter 4 thus contains the same principle as chapter 3, but applied to a new setting, that of the labour market.

The theory assumes that workers can sustain a wage cut in the first period. At first sight, it would appear that the theory is vulnerable to the criticism that, if the wage cut is sufficiently large, it may not be possible for the worker to survive. If a worker lacks savings, or cannot resort to (sufficiently cheap) borrowing, or if he does not have a family that will support him[21], then he will not be able to sustain the wage cut. This implies that, when competition increases in period 2, or when training becomes more general, so that the period 2 wage increases, the multinational in period 1 will not be able to pay a lower wage. This would then destroy the see-saw mechanism, so that an

[21] In addition, if workers and firms cannot sign contracts whereby workers make delayed payments to firms to cover the costs of training, possibly via deductions from workers' wages over time, then firms may not be able to forego wage reductions while imparting training. This is the problem of firms not being able to obtain property rights over the human capital sold to workers (Katz and Ziderman (1999)).

increase in competition in period 2 will result in the multinational incurring losses over both periods. However, this is only true in the simplest version of our model in which workers do not require a wage premium over what they earn in agriculture, i.e. they do not face any cost of moving from agriculture to industry. When they do face such a cost, the wage in period 1 need not fall below the level which is required by workers to survive, so that the above analysis goes through. What matters is not the level at which the sum of wages over the two periods is fixed, but simply that it is fixed. By the same argument, the first period wage need not fall below the wage earned in agriculture, thus fitting the facts.

Capital market imperfections, in addition to reducing the amount of training, may induce firms to provide more firm-specific training, since this reduces turnover compared with general training, so that firms will be more willing to incur training costs. At the same time, poorly-functioning capital markets imply that risk-averse workers cannot invest in assets which are inversely correlated with their acquired skills so as to diversify their portfolio of assets. Consequently, they may seek to become 'Jacks of all trades', and demand more general training, rather than more specialized training. Whilst large families may compensate for this by having each member undertake a different specialization, this may not be a viable solution because economic growth has the deleterious effect of reducing family ties and cooperation within the family (Katz and Ziderman(1999)).

Another reason why the first period wage may not fall is that there may be minimum wage legislation in place. The effect on the multinational is the same as that of workers not being able to sustain a wage cut: the see-saw mechanism is broken, so that the multinational cannot benefit from an increase in competition in period 2. For the same reason, the multinational may not be able to benefit from more general training. In addition, the first period wage may not fall because of trade union pressure applied to all wages irrespective of whether they are incurring training or not.

The assumption that all agents perfectly predict the future is common to all three chapters. This assumption seems to be a valid approximation in chapters 2 and 3, in which the semiconductor industry is the main example used, to illustrate the models, and in which industry buyers tend to be sophisticated firms which spend considerable resources on predicting future market trends. The assumption of perfect foresight is less realistic in chapter 4, where workers in developing nations switch from agriculture into a new sector, namely industry. This information requirement is probably less stringent than it appears considering that workers are basically expected to predict in period 1 that more firms in period 2 imply higher period 2 wages. Nevertheless, given that we are considering a developing nation in which a forward looking mentality may be less pronounced than in a so-called developed nation, the assumption might still be a demanding one.

The commitment problem of chapter 2 is absent in chapter 4. This is because each buyer's demand must be elastic for there to be a commitment problem, and this is the case in chapter 2, but not in chapter 4, where the model assumes that each worker supplies his work inelastically[22] (workers in chapter 4 play the same role that buyers play in chapter 2). If demand was inelastic, the incumbent in chapter 2 would not care about the level of each individual price, but only about the sum of prices. As demand is elastic, the incumbent would like prices to be same over the two periods, so that quantities are also the same and profits are maximized. However, because he cannot commit to not behave opportunistically in period 2, prices in different periods will differ, and overall profits will be lower.

In chapter 4, instead of there being a commitment problem, decreasing returns to scale drive the result that more competition results in higher prices, which cannot happen with constant returns to scale, at least not with equally efficient firms. Chapter 2, on the other hand, assumes constant returns to scale, so that the decreasing returns motivation to invite rivals of chapter 4 is absent.

The role that the set-up cost plays in chapter 2 in enabling buyers to consume in period 2 is similar to that which the foregone wage in agriculture plays in chapter 4. The set-up cost must be positive in chapter 2 to generate the commitment problem. In chapter 4, we have the equivalent of such a fixed cost, but this is not enough to generate a commitment problem: an elastic supply of work by each worker is also needed, as we explained above, which we do not have.

Another difference between chapters 2 and 4 is that in chapter 4, we consider an extension where firms hire workers not only in period 1, but also in period 2, whilst in chapter 2 it is not possible for more buyers to enter in period 1. The difference is due to the fact that training creates 2 groups of workers in period 2, those who have been

[22] We later relax this assumption and consider workers whose supply work is elastic, but this is not necessary for the results of chapter 4 to hold, in particular the result that more competition for trained workers can raise profits of training firms.

trained in period 1, with higher productivity, and those who have not been, with lower productivity; those who have already been trained are distinguishable from those who have not and receive a higher wage in period 2. In contrast, if new buyers were to enter in period 2, there would not be two separate groups of buyers, since buyers would purchase the same quantity at any given price regardless of whether they enter in period 1 or period 2.

There is some evidence which supports the main results of chapter 4. For example, in a CBI survey in 1989, only 38% of employers claimed that the possibility that workers may leave deters them from training. If we take poaching per se to be damaging for training firms, then the above empirical finding would make sense provided that training firms enjoy some kind of compensation for the loss of trained workers and thus revenues. The model of this paper claims that the compensation may exceed the loss, hence the result that more poaching is associated with increased profits of the multinational. It does not claim that poaching per se is beneficial to training firms.

We now summarize the recent theoretical literature which studies the relationship between training by multinationals in LDC's and poaching by local firms.

Fosfuri, Mott and Rønde (2001) develop a model where technological spillovers arise as a result of workers trained by multinationals being poached by local firms. However, even without the workers being poached, an LDC can enjoy pecuniary spillovers due to the fact that, in order to retain workers, the multinational needs to pay higher wages. If a multinational does not compete in the product market with a local firm, it has a lower incentive to retain a worker, so that technological spillovers (poaching of trained workers) are more likely. Furthermore, turnover of trained workers is higher the more general is the training given by the multinational, in line with the literature on training.

Glass and Saggi (2001)'s theoretical findings are similar to those of Fosfuri, Mott and Rønde (2001). They find that technological spillovers, when they occur, benefit local firms at the expense of workers, whose wages fall since multinationals no longer attempts to curtail the diffusion of their technology by paying high wages. Hence, welfare when technology diffusion takes place may be less than welfare when technology diffusion does not take place, and policies designed to encourage technology transfer may not be welfare-improving.

In Ethier and Markusen (1996), the workers of a multinational firm can leave and set up local firms. The authors consider different strategies which a firm can adopt to curtail technological diffusion, including foreign direct investment, exporting and licensing. If the firm produces abroad, it will suffer from quick dissipation of its knowledge, but it will be able to avoid exporting costs. They find that all arrangements are possible, depending on underlying parameter values. In particular, foreign direct investment is chosen when exporting cots are medium, so that employees have an inventive not to leave and face competition from exports.

To the best of my knowledge, no other theoretical mechanism has been proposed hypothesizing that a training firm benefits when an increase in competition for its trained workers takes place. The theoretical literature on training has largely ignored the question of how more competition for trained workers affects training firms and thus the quantity of training, focusing instead on the question of whether firms will pay for training that is of a general nature. There are, to the best of our knowledge, only 2 papers in the training literature that deal with an increase in competition for trained workers. The first is Autor (2000) and the second is Stevens (1994).

In Autor (2000), so-called temporary help firms offer general training to good and bad workers, but only good workers accept it since it is more valuable to them by virtue of their higher productivity. Thus temporary help firms are valued by their customers for separating good from bad workers. An increase in market competition is here modelled as a change in market structure from a given number of firms environment to one with free-entry. More competition is assumed to lower profits to zero, in sharp contrast with the main result of this paper. Training is increased as firms must, in a competitive environment, ensure that workers' utility is maximized, and increased training achieves this. Like Autor (2000), the model of chapter 4 shows that training increases with increased competition, though for different reasons.

In Stevens (1994), a random shock process determines the productivity of firms after training has been carried out and each firm pays a given worker only what its closest rival would pay him, so that the wage is less than the marginal product. This implies that there is imperfect competition in the labour market. A greater number of firms competing for a worker results in lower expected profits for the training firm by virtue of the probability of poaching being greater. Specific (though not entirely specific) training becomes less valuable, as the worker is more likely to leave. Thus more competition in

Stevens (1994) is associated with lower profits, as in Autor(2000) but in contrast with our model, and less training, unlike in both Autor(2000) and here.

Comparing the theory of this paper with that of Stevens (1994) helps to better understand how decreasing returns to scale relate to the existence of a training externality. Stevens argues that, in situations where training is not entirely general, an externality arises, that is, non-training firms benefit from training by other firms. Since training that is not entirely general is of use to only a limited number of firms, competition for labour is imperfect; consequently, the wage falls short of the marginal product, generating the externality. In the model of chapter 4, however, the second period market is competitive, so that poaching firms pay a wage equal to the marginal product, and yet the externality still exists. That is, local firms still benefit from poaching workers which the multinational trains.

The results of chapter 4 and the work by Stevens can be reconciled by noting that she assumes workers' productivities to be independent of each other, whilst in our model they are assumed to be characterized by diminishing returns. Consequently, local firms make positive profits on inframarginal workers. This means that there is a training externality with equality between the wage and the marginal product of local firms. This discussion highlights the important role played by decreasing returns to scale in generating the training externality.

Although Becker's (1964) benchmark human capital model does not address the question of how more competition affects training firms, certain implications concerning this question can nevertheless be drawn. Becker's model is not suitable for studying the effects of more general training, since he assumes either completely firm-specific training or completely general training, with no intermediate case. We can, nevertheless, consider what happens to the training firm's profits in Becker's model as training changes discontinuously from completely firm-specific to completely general. In the case of completely firm-specific training, the firm and the worker share the net return to training, so that there is a profit to be enjoyed by the training firm.

In discussing completely general training, Becker used a stylized model which assumes that, once trained, a worker's productivity is the same both in the firm that trained him and in the firm which may poach him. In the period subsequent to that in which training takes place, the wage rises in line with the worker's marginal product, so that other firms are indifferent as to whether or not to employ the worker (as is the training firm). In a fully competitive environment, the training firm can do no better than to recoup its training costs by reducing the wage while training the worker, so that it breaks even on training the worker. Thus none of the benefits from training accrues to firms: they accrue entirely to the worker.

Hence, in going from completely firm-specific to completely general training, the training firm's profits in Becker's model change from some positive value to zero. This is in contrast with our model, in which more general training results in higher profits for the training firm (the multinational).

It is instructive to compare the theory of this paper with the recent literature on imperfect labour markets for trained workers (see Acemoglu and Pischke (1999) for a survey). Suppose, for example, that a current employer has superior information about a worker's skill level compared to potential poaching firms. Then, if the employer provides training, he will enjoy an increase in productivity which, because of the difficulty faced by outsiders in judging the level of training, will be accompanied by a less than proportional increase in the wage. Thus there is an incentive to engage in training even if this is of a general nature. Informational asymmetry turns training that is general in nature into what is effectively firm-specific training. This results in imperfect competition in the labour market, so that the wage and productivity do not step up in line following training. Acemoglu and Pischke (1999) use the theory of imperfect labour markets to explain the evidence showing that firms pay for general training (see Steedman (1993), Soskice (1994) and Harhoff and Kane (1997) for studies of the German apprenticeship system).

These findings are in sharp contradiction with the prediction of Becker's (1964) model, which says that firms would not pay for general training. In contrast with the literature on imperfect competition and training, the multinational of chapter 4 which trains workers does not pay for general training, in the sense that it lowers the first period wage to (more than) cover for the cost of training, in a way similar to Becker (1964)'s.

Chapter 2

Why firms invite rivals: the commitment problem

1 Introduction

There are well-documented cases of firms, mainly technological innovators, which give away their technology for free in an effort to invite competitors. For example, Xerox, the developer of the LAN Ethernet, offered open licenses at a nominal charge (Sirbu and Hughes (1986)). Similarly, Apple Computers in 1994 decided to start licensing its proprietary operating system (Kende (1995a)). These examples suggest that it is sometimes in the interest of incumbents to face competition for the same buyers[1].

Farrell and Gallini (1988) have suggested that firms employ such a strategy to overcome a commitment problem relating to their monopoly status: a monopoly cannot commit in the first period to charge a low price in the second; consequently, buyers, who must incur a product-specific set-up cost in the first period, demand to be compensated via a low first period price[2]. If the required first period price is negative, and negative prices are ruled out, then the monopolist sells nothing. Additionally, if each buyer's demand is elastic, the price path without commitment is more "spread out" than that with commitment; since this spread is undesirable from the incumbent's point of view, there is a further incentive to invite rivals.

In the same spirit, Kende (1995a) considers a hardware manufacturer which licenses its hardware technology to commit to lower hardware prices and so attract more software developers, software and hardware being complementary goods. This in turn raises hardware purchases and profitability. Kende's result relies on the same monopoly commitment problem of Farrell and Gallini (1988). Economides (1996a) shows that, in the context of direct network externalities, a hardware producer, which licenses rivals as

[1] Anther examples is Sun Microsystems, which encouraged cloning of its computer workstations (see Business Week, July 24 1989).
[2] Clearly, if the incumbent could commit via a price guarantee, this would be less costly than inviting rivals, and would therefore be a superior alternative from the incumbent's point of view. In practise, there are also other forms of commitment which are sometimes used, such as free upgrading policies.

a way of committing to low prices, may charge negative royalties if this will attract enough buyers of his product.

In this paper, we develop the commitment problem analysis of Farrell and Gallini (1988). We do so primarily by assuming that firms compete in quantities, rather than prices. This is for two reasons. Firstly, to provide a theoretical benchmark: knowing the results for the case of Bertrand competition through the work of Farrell and Gallini (1988), it is important that we also know the results for the 'alternative' case of Cournot competition. This paper shows that assuming Cournot competition does not merely lead to a 'smooth-version' of the results under Bertrand competition, but generates new results, such as the possibility that an incumbent may invite rivals to enter in period 1, rather than wait until period 2. Secondly, we observe firms which license their technology to many producers (Dick (1992), Swann (1987)). Bertrand competition implies, when goods are homogenous, that we cannot say anything about how many rivals a firm might invite, since profits collapse to zero with any number of rivals greater than 1. Cournot competition, on the other hand, enables us to make predictions about this number[3].

What generates the commitment problem? It is the combination of the fixed-cost which buyers have to incur at the start of period 1 and the fact that each buyer's demand is elastic. Consider buyers who are identical in always and in particular in their fixed cost of entry, so that the analysis is an analysis per buyer. Given each buyer's demand is elastic, profits would be highest if prices were the same in both periods. However, an incumbent operating alone in both periods will charge a high price in the second period, thus exploiting the fact that buyers have incurred a sunk investment cost, and would therefore need to compensate buyers by offering them a sufficiently low price in the first period. Inviting competitors to enter in the second period will lower the price in period 2; because buyers foresee this in period 1, they tolerate a higher price in period 1. Thus prices come closer together, benefiting the incumbent, though at the cost of a loss in market share in the second period.

We call the inverse relationship between the first and second period prices the 'see-saw effect'. Given that the whole mechanism is about the incumbent wanting to attract rivals so as to credibly commit to buyers, we call this the 'commitment mechanism', to distinguish it from the capacity constraint mechanism of chapter 3.

[3] Cournot games are sometimes seen as analytically tractable versions of more sophisticated Bertrand games with differentiated products.

The incumbent can control how many rivals it attracts by limiting access to its technology, assuming that it has this ability, e.g. by means of patent rights. The model assumes that the incumbent offers this technology freely, since the introduction of royalties would not add anything central to our analysis; rather, it would take away from the focus of the paper, which is that an incumbent may attract rivals even in the absence of royalties.

Now consider buyers who are heterogeneous in their fixed-cost of entry, which sets a limit on how many buyers enter at the start of period 1. This adds a further incentive for an incumbent to invite rivals: to attract more buyers. If an incumbent firm can invite rivals in the period 2, so that the period 2 price falls, the incumbent will be able to attract more buyers, whose demand it will have to 'share' with the entrants in period 2, but clearly not in period 1. This effect may more than compensate for the increased competition in period 2, so that the incumbent has an overall incentive to attract competition.

We find situation sin which profits of the incumbent rise with the number of entrants only if a sufficient number of buyers enter as a result of the lower second period price. In other words, demand must be sufficiently elastic. This is so provided that buyers' set-up costs do not rise too steeply with the number of buyers, i.e. that buyers are sufficiently homogeneous.

We also find that the profits of entrants can rise with the number of entrants, since an increase in the number of entrants attracts more buyers. This result has a number of implications. Firstly, given that the incumbent benefits from entry of rivals, it implies that entry may be advantageous for all firms. Since, in addition, it is advantageous for buyers, it may be advantageous for all agents. Thus, at least in theory, increased competition can benefit all economic agents (without any redistribution among them). This cannot happen in a 1-period model, since there is no commitment problem.

Secondly, the result may help to explain the phenomenon of sub-licensing. Sub-licensing is observed in second-source agreements in the electronics industry, in particular in the semiconductor industry (Swann 1987), at very modest royalty fees. The commitment explanation of this paper may explain this phenomenon: the incumbent may allow sub-licensing so as to achieve the number of rivals it finds most profitable. The incumbent may allow sub-licensing instead of approaching sub-licensing firms directly because it may not have had contacts with the sub-licensing firms in the past, whilst a licensee might have; thus the transaction cost of licensing may fall via sub-licensing.

In a key extension of the main model, we find the incumbent may want to invite rivals to enter in period 1 already, i.e. without delay. This is because the commitment problem has repercussions in the first period: to induce a sufficiently large entry of customers, the incumbent may need to produce at price below marginal cost in period 1. Inviting rivals in period 1 enables the incumbent to 'share out' some of the period 1 losses that it would incur were it to try and attract the same number of buyers as would enter with a duopoly in period 1.

This result is important in view of the statement which Farrell and Gallini (1988) made: "Obviously, second-sourcing that enables other firms to compete with the innovator on equal terms immediately (and without royalty payments) is not profitable: we consider a less drastic form of second-sourcing, in which the technology is given away with a lag". The authors' conclusion is accurate because they assume Bertrand competition, so that inviting a rival in period 1 causes profits to be zero in each period. However, with Cournot competition, profits in period 1 do not fall to zero with a finite number of rivals. If the loss-sharing motive of period 1 is sufficiently strong, it may be profitable for a monopolist incumbent to invite rivals to enter already in period 1.

The result that an incumbent may want to invite rivals to enter in period 1 is significant, as it helps to explain a number of real life phenomena. First of all, it helps to explain those second-source agreements in which the original source producer licenses a second source producer to compete against it already from the moment that the product is launched, and not later (Dick (1992)). It is common to find that innovators incur a loss in the initial phases just after the product has been launched. The analysis of this paper suggests that an incumbent may attract rivals to share these initial losses.

A related case is that of Apple Computers which, in 1994, chose to start licensing its proprietary operating system (Kende(1995a), who refers to Business Week, September 26, 1994). Some commentators argued that Apple ought to have taken this step earlier, to expand the client base more and at an earlier time, so as to encourage software producers to write software for Apple, rather than for Microsoft Windows. The results of our model may be interpreted to suggest that who was correct depends on the size of the fixed cost faced by buyers: inviting rivals from the beginning would be optimal if the set-up cost of buyers was large enough, as this would result in large enough losses in period 1 for Apple if operating without competitors.

As Farrell and Gallini (1988) point out, the set-up costs incurred by buyers effectively makes the good sold in period 1 complementary with the same good sold in period 2, in the sense that a higher expected second period price reduces first-period demand. We can therefore apply the model to a whole range of complementary goods, such as computer hardware and software, video games and video game cartridges, printers and ink cartridges, video recorders and video cassettes, etc.. In the case of the semiconductor industry, we can apply the model to complementary products which belong to the same 'technological group'. An example is support microchips that can be used with a variety of microprocessors (Swann(1987)).

The chapter is structured as follows.

In section 2, we set up the model. This section describes the see see-saw effect, the mechanism by which prices in different periods move in opposite directions when the number of rivals increases, and which comes from the behaviour of buyers.

In section 3, we consider the incumbent's decision to invite rivals to enter in period 2. This is the Cournot equivalent of the Farrell and Gallini (1988) model, which assumes Bertrand players.

In section 4, we explore the possibility that an incumbent may want to invite a rival in period 1 already, i.e. without delay.

In section 5, we consider the social optimum, and show that the socially optimal number of entrant firms may be finite.

Finally, in section 6 we draw some conclusions.

2 Model

Buyers are assumed to be price-takers. We assume buyers' utility to be $V = U(q_{A,t}) + y_{A,t}$, where $q_{A,t}$ denotes the quantity chosen by buyers in time period $t = \{1,2\}$ and $y_{A,t}$ is a numeraire good. Given income $I = p_t q_{A,t} + y_{A,t}$, where p_t is the price in period t, the maximization problem $\underset{q_{A,t}, y_{A,t}}{Max}\ U(q_{A,t}) + y_{A,t} - \Lambda(I - p_t q_{A,t} - y_{A,t})$ is equivalent to

$$\underset{q_{A,t}}{Max}\, U(q_{A,t}) - p_t q_{A,t} \tag{1}$$

by virtue of the fact that $\Lambda=1$. The maximization problem is intuitively appealing in that it simply involves maximizing the gap between benefits and costs. Buyers can be either firms of consumers. If they are firms, then they buy an input in production, for example a computer hardware manufacturer may buy a semiconductor component.

The resulting demand function is

$$\frac{dU}{dq_{A,t}} = p_t \tag{2}$$

Critically, each buyer faces a fixed cost of entry. If buyers are firms and the good in question is an input in production, then this cost could be an adaptation cost of adjusting the production process to accommodate the new input. If the good in question is a consumer good, it could be the cost of learning how to use the new product, or of having to buy complementary equipment so as to use the new product.

Buyers can be either homogeneous or heterogeneous in their fixed cost of entry, F_i. Suppose buyers are heterogeneous. Arranging buyers in order of increasing fixed cost of entry, we assume

$$F_i = F_0 + F_1 r^{\gamma} \tag{3}$$

where F_i is the fixed cost of buyer i, and r is his rank number in the sequence of buyers. In equilibrium, the marginal buyer's fixed cost F is

$$F = F_0 + F_1 n_A^{\gamma} \tag{4}$$

where n_A is the number of buyers, which is determined in period 1. The marginal buyer just breaks even, i.e.:

$$CS_1 + CS_2 - F = U(q_{A,1}) - p_1 q_{A,1} + U(q_{A,2}) - p_2 q_{A,2} - F = 0 \tag{5}$$

where CS_t is consumer (buyer) surplus in period t.

If buyers are homogeneous (F_1=0), then (5) applies to all of them as they all face the same fixed cost. In this case, F_0 needs to be sufficiently large for (5) to hold[4]. As long as consumer surplus in both periods is weakly positive, then buying takes place in both periods. Now, because (5) says that buyers must purchase in both periods to break even, it follows that a buyer would incur a net loss if he purchased in only one period. Thus buyers must enter in period 1. This implies that no buyers enter in period 2, simplifying the analysis considerably.

Totally differentiating (5) enables us to identify the see-saw effect:

$$\frac{dU}{dq_{A,1}} dq_{A,1} - p_1 dq_{A,1} - q_{A,1} dp_1 + \frac{dU}{dq_{A,2}} dq_{A,2} - p_2 dq_{A,2} - q_{A,2} dp_2 - \frac{dF}{dn_A} dn_A = 0 \tag{6}$$

Applying the first order conditions (2) to (6), some terms cancel out and we have the central equation of this chapter:

[4] With logarithmic utility, (5) holds in the case of homogenous buyers even for a small F_0 because the elasticity of demand is constant at -1 and so the incumbent focuses on minimizing costs by reducing quantity to the bare minimum (this being given by (5)).

$$-q_{A,1}dp_1 - q_{A,2}dp_2 - \frac{dF}{dn_A}dn_A = 0 \qquad (7)$$

The see-saw effect is most clearly seen when buyers are homogenous. In this case $dF/dn_A=0$, and therefore

$$\frac{dp_1}{dp_2} = -\frac{q_{A,2}}{q_{A,1}} \qquad (8)$$

Thus the relationship between prices in different period is negative. This is of critical importance: for two reasons. First, it is the via the see-saw effect that, for any given number of buyers, more competition in period 2 raises the price in period 1 and causes profits of the incumbent to increase. Second, when more competition in period 2 makes the price fall in the same period, demand in period 2 increases; provided that buyers are heterogeneous in their fixed cost of entry, more buyers enter, so demand increases in the first period, too. In other words, the period 1 and period 2 markets can be thought of two complementary products, and this again is via the see-saw effect.

In the specific examples which we discuss below, we wish to consider the classic case of linear demand. This implies that we assume the quadratic utility function

$$U = aq_{A,t} - \frac{\lambda}{2}q_{A,t}^{\;\;2} \qquad (9)$$

The use of a quadratic utility function makes the period 1 optimization algebraically cumbersome, as we shall see. We experimented with different variants of quadratic utility with the aim of making the algebra as simple as possible, but could not find a superior alternative to (9). Inserting (9) into (1), we obtain that buyers in each period t solve the problem

$$\underset{q_{A,t}}{Max} \quad CS_t = aq_{A,t} - \frac{\lambda}{2} q_{A,t}{}^2 - p_t q_{A,t} \tag{10}$$

The resulting inverse demand function is

$$p_t = a - \lambda q_{A,t} \tag{11}$$

Note that a higher λ implies a less elastic demand schedule.

Firms face a constant unit cost of production c. Each firm in period 2 solves

$$\begin{aligned} &\underset{q_2}{Max} \quad (p_2 - c) q_2 \\ &s.t. \quad (11), q_{A,2} = \frac{q_2 + s_2}{n_A} \end{aligned} \tag{12}$$

where s_2 is the combined output of rivals and n_A is the number of buyers, which determined in period 1 and so is exogenous in period 2[5]. Let the number of firms which enter in period 1 (including the incumbent) be m and the number of firms which enter in period 2 be n. This implies that in equilibrium $s_2=(m+n-1)q_2$, and so the equilibrium value of q_2 is

$$q_2 = \frac{(a-c)n_A}{\lambda(m+n+1)} \tag{13}$$

[5] Note that a firm in period 2 will not find it profitable to deviate from the Nash equilibrium in order to attract more buyers (by increasing quantity and thus make the price fall so much that new buyers enter with fixed costs larger than those of existing buyers). This is because the incentive which a firm in period 2 faces is smaller than the incentive which the incumbent in period 1 faces when deciding how many buyers to attract: the incumbent in period 1 sells to the same buyers over both periods, not just in the second.

In period 1, there may be 1 or more firms. If there are more than 1 firm, each firm plays Nash-Cournot taking into account that its output, by affecting p_1, affects the number of buyers who enter. After firms have chosen output, buyers incur their fixed cost of entry and enter. Each of the m firms faces the following problem:

$$\max_{q_1} \quad \Pi = (p_1 - c)q_1 + (p_2 - c)q_2$$
$$s.t. \quad (5), (11), (13), q_{A,1} = \frac{q_1 + s_1}{n_A}, q_{A,2} = \frac{(n+1)q_2}{n_A} \tag{14}$$

where s_1 is the combined output of rivals in period 1. In equilibrium, $s_1=(m-1)q_1$. The endogenously determined number of buyers n_A turns out to be essentially a cubic function. Let

$$z \equiv \sqrt{\frac{F_1}{\lambda} n_A^{\gamma} + \beta}, \quad where\ \beta = \frac{F_0}{\lambda} - \frac{(m+n)^2(a-c)^2}{2\lambda^2(m+n+1)^2} \tag{15}$$

Then

$$z^3\left\{\frac{\lambda}{2}\left[\frac{(m-1)\gamma)}{m}\right] - 2(\gamma+1))\right\} + z^2\left\{\frac{(a-c)(2+\gamma)}{2}\right\} +$$
$$+ z\left\{-\beta\lambda\left[\frac{(m-1)\gamma)}{m} - 2(\gamma+1)\right] - \beta'\right\} - (a-c)\beta\gamma = 0 \tag{16}$$
$$where \quad \beta' = 2F_0 - \frac{\left[(m+n)^2+1\right](a-c)^2}{\lambda(m+n+1)^2}$$

Profits in (14) simplify to some extent to the following expression:

$$\Pi = n_A\left\{\frac{1}{m}\left[(a-c)q_1 - \lambda q_1^{\,2}\right] + \frac{(a-c)^2}{\lambda(m+n+1)^2}\right\}$$

$$where \quad q_1 = \sqrt{2\left[\beta + \frac{F_1 n_A}{\lambda}\right]} \tag{17}$$

In words, (17) says that total profits are given by the sum of profits per buyer in period 1 and profits per buyer in period 2, multiplied by the number of buyers.

Note that the parameters a and c always occur together as the term (a-c). So we can just set c=0 and focus exclusively on revenues. This is a consequence of assuming linear demand.

3 Inviting rivals to enter in period 2

Set m=1, so that there is 1 incumbent in period 1. Consider how profits of the incumbent vary with n, the number of rivals in period 2.

Using the envelope theorem, we can write

$$\frac{d\Pi}{dn} = \left.\frac{\partial \Pi}{\partial n}\right|_{q_1} = q_1\frac{\partial p_1}{\partial n} + p_2\frac{\partial q_2}{\partial n} + q_2\frac{\partial p_2}{\partial n} - c\frac{\partial q_2}{\partial n} \tag{18}$$

Re-writing $\frac{\partial p_1}{\partial n}$ as $\frac{\partial p_1}{\partial p_2}\frac{\partial p_2}{\partial n}$ and applying the see-saw effect equation (7) with $\frac{dp_1}{dp_2} = \frac{\partial p_1}{\partial p_2}$ to (18), we obtain

$$\frac{d\Pi}{dn} = \underset{-}{\frac{\partial p_2}{\partial n}}\left(-nq_2 - n_A \underset{+}{\frac{dF}{dn_A}} \underset{-}{\frac{\partial n_A}{\partial p_2}}\right) + \underset{-}{\frac{\partial q_2}{\partial n}} \underbrace{(p_2 - c)}_{+} \tag{19}$$

Therefore, when the incumbent operates alone (n=0), the expression above is negative, and the incumbent has no incentive to face a tiny amount of competition; only if n is big enough, so that nq_2 is big enough, is it possible that the incumbent has an incentive to invite rivals. This is true for both homogeneous (dF/dn_A=0) and heterogeneous (dF/dn_A>0) buyers.

We can obtain further insight into the nature of the commitment problem as follows. Write profits of the incumbent (14) so that each period's quantity is shown as the product of the quantity per buyer and the number of buyers:

$$\Pr M_{n>0} = n_A\left[(P_1 - c)q_{A,1} + (P_2 - c)\frac{q_{A,2}}{1+n}\right] \tag{20}$$

We partially differentiate (20) with respect to n:

$$\frac{d\Pr M_{n>0}}{dn} = n_A\left\{\left[(P_1 - c)\frac{dq_{A,1}}{dP_1} + q_{A,1}\right]\underset{-}{\frac{dP_1}{dP_2}}\,\underset{-}{\frac{dP_2}{dn}} + \frac{1}{1+n}\left[(P_2 - c)\frac{dq_{A,2}}{dP_2} + q_{A,2}\right]\underset{-}{\frac{dP_2}{dn}} - \frac{(P_2 - c)q_{A,2}}{(1+n)^2}\right\} \tag{21}$$

Rearranging terms inside the square brackets, we have:

$$\frac{d\Pr M_{n>0}}{dn} = n_A\left\{\left[\frac{(P_1 - c)}{q_{A,1}}\frac{dq_{A,1}}{dP_1} + 1\right]q_{A,1}\underset{-}{\frac{dP_1}{dP_2}}\,\underset{-}{\frac{dP_2}{dn}} + \frac{1}{1+n}\left[\frac{(P_2 - c)}{q_{A,2}}\frac{dq_{A,2}}{dP_2} + 1\right]q_{A,2}\underset{-}{\frac{dP_2}{dn}} - \frac{(P_2 - c)q_{A,2}}{(1+n)^2}\right\}$$

Marginal cost is a constant and so plays a minor role in the analysis. Furthermore, we know from the analysis above that, when considering liner demand, we can set c=0 without loss of generality. We therefore set c=0. For simplicity, suppose buyers are homogenous, so that we can use (8) to substitute dP_1/dP_2:

$$\frac{d\Pr M_{n>0}}{dn}=n_A q_{A,2}\left\{\underbrace{\left\{-\left[\frac{P_1}{q_{A,1}}\frac{dq_{A,1}}{dP_1}+1\right]+\frac{1}{1+n}\left[\frac{P_2}{q_{A,2}}\frac{dq_{A,2}}{dP_2}+1\right]\right\}\frac{\partial P_2}{\partial n}}_{\textit{elasticity effect}}-\underbrace{\frac{P_2}{(1+n)^{2}}}_{\textit{loss of market share effect}}\right\} \tag{22}$$

The first two terms involving the square brackets make up the elasticity effects: given that an increase in period 2 competition lowers the second period price and enables M to raise the first period price, it benefits M if demand in the second period is elastic and demand in the first period inelastic. The third term, which falls with the square of the number of (1+n) firms in the second period, measures the reduction in market share which M suffers in period 2 as a result of increased competition. Thus, for the incumbent to want to invite rivals, the elasticity effect must dominate the loss of market share effect.

With linear demand, (19) looks as follows:

$$\frac{d\Pi}{dn}=n_A\frac{a^2}{\lambda(m+n+1)^3 m}\left\{2n-\frac{(m+n)(a-c)}{\lambda q_1}\right\} \tag{20}$$

which can be either positive or negative. In particular, at n=0, the expression is always negative, implying that, starting from a situation without rivals, the incumbent never wants to face an (infinitesimal) increase in competition. This confirms what we found in the general case of (19).

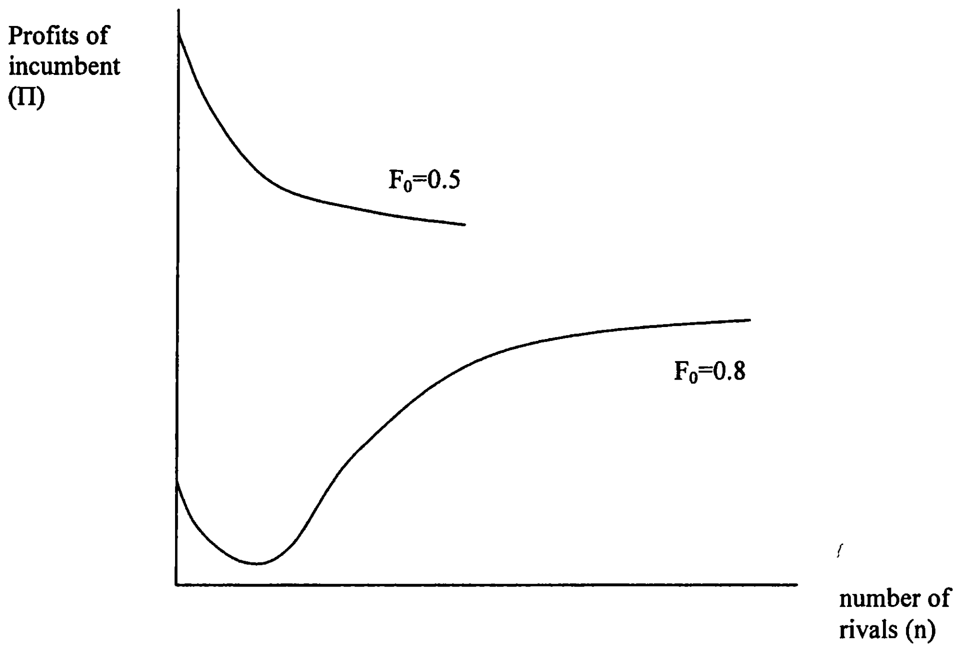
Unfortunately, with quadratic utility the algebra is complicated and we need to rely on simulations to perform some comparative statics. We do the analysis first for the simpler case of homogeneous buyers, and then for heterogeneous buyers.

With homogeneous buyers, so that F_1=0, all buyers face the same set-up cost $F=F_0$. In this case, there is in general no guarantee that buyers break even over the two periods; indeed, if the fixed cost F_0 is low enough, then CS_2-F>0, which implies $CS_1+CS_2-F>0$, and the model reduces to two ordinary Nash-Cournot games played over two periods with no link between them. However, if the fixed cost is high enough, then CS_2-F<0, and the link is retained, namely $CS_1+CS_2-F=0$. If the incumbent finds it profitable to sell to one buyer, it will find it profitable to sell to all of them, given that they are identical, in particular with respect to their fixed cost of entry. Assume that F_0 is high enough for the link to be retained. Set m=1, so that there is only 1 incumbent in period 1. Then our simulations suggest the following results.

Proposition 1: When F_0 is big enough, the incumbent invites rivals, provided enough potential rivals are available.

As an illustration[6], consider the example below with a=1 and λ=1:

Figure 2.1: How the set-up cost of buyers affects the relationship between profits of the incumbent and the number of rivals in period 2



[6] This diagram and those that follow in this and the next chapter involve continues lines/curves. Since the number of rivals is an integer, these lines represent sequences of dots. We use continues lines because sometimes a profit profile may exhibit a change of direction when the number of rivals is very large, so that it would be inconvenient (and sometimes impossible) to draw the exact number of dots.

The incumbent invites rivals when F_0 is big enough. This is because the incumbent in period 1 is effectively sustaining buyers into entering through a low enough p_1, and the higher F_0 is, the lower p_1 needs to be. If the incumbent could commit, it would be optimal for him to spread F_0 over both periods by equally lowering p_1 and p_2 (via higher output). But he can't. Inviting rivals in period 2 makes p_2 fall, allowing the incumbent to enjoy a higher p_1, and takes some of the burden of sustaining buyers off his shoulders. This result is consistent with the findings of Farrell and Gallini (1988). In fact, their analysis, which is for Bertrand competition, corresponds to the special case of $n=\infty$ in the diagram above. Our analysis extends theirs by showing the profile of profits for all finite values of n.

The simulations indicate that, with linear demand at least, one cannot have an internal optimum for the number of rivals. Instead, the simulations show that, for a sufficiently high F_0, the profit profile is U-shaped, and the optimum occurs at $n=\infty$. This implies that the standard competition effect dominates when the number of rivals is small, but the commitment mechanism dominates when the number of rivals is large. Also notice that the profit profile for $F_0=0.5$ is shorter than that for $F_0=1$. This is because at n=6 we find that $CS_1+CS_2-F>0$, so that there is a 'regime shift'. We come back to this case below.

Dick(1992) criticizes the commitment problem literature for not being able to explain why there are so many second-source producers. His criticism is based on the implicit assumption of Bertrand competition, which is found in the pioneering work of Farrell and Gallini (1988). In the model of this chapter, in which we assume Cournot competition, it is quite possible for the number of second-source producers to be very high.

Let's move on to some more comparative statics; in particular, we consider the impact of the demand parameters a and λ on the profit profile. For this purpose, it is useful to quantify the size of the commitment problem. If the incumbent could commit, he would maximize revenues by ensuring that $q_{A,1}=q_{A,2}$, thus completely smoothing the consumption path of buyers. However, the incumbent cannot commit, and so $q_{A,1}>q_{A,2}$, reflecting the incumbent's opportunistic behaviour in period 2. Thus the size of the commitment problem can be measured by ratio $q_{A,1}/q_{A,2}$:

$$\textit{Size of commitment problem} = \frac{q_{A,1}}{q_{A,2}} = \sqrt{\frac{2F_0\lambda}{a^2}\left(\frac{n+2}{n+1}\right)^2} - 1 \qquad (21)$$

Thus the commitment problem rises with F_0, as we have already seen, and with λ, essentially because a higher λ reduces $q_{A,2}$ (and also $q_{A,1}$, but by less, because q_1 is 'anchored' to F_0). In other words, more elastic demand tends to weaken the result that a greater number of rivals increases profits of the incumbent. This is because the more elastic demand is, the smaller the period 2 price reduction when the number of rivals increases, and hence the smaller the period 1 price increment that is transmitted via the see-saw effect. The commitment problem falls with the parameter a, essentially because a higher a increases $q_{A,2}$ (and $q_{A,1}$, but by less, again because q_1 is 'anchored' to F_0). Of course, the commitment problem falls with n, which is the whole point of inviting rivals.

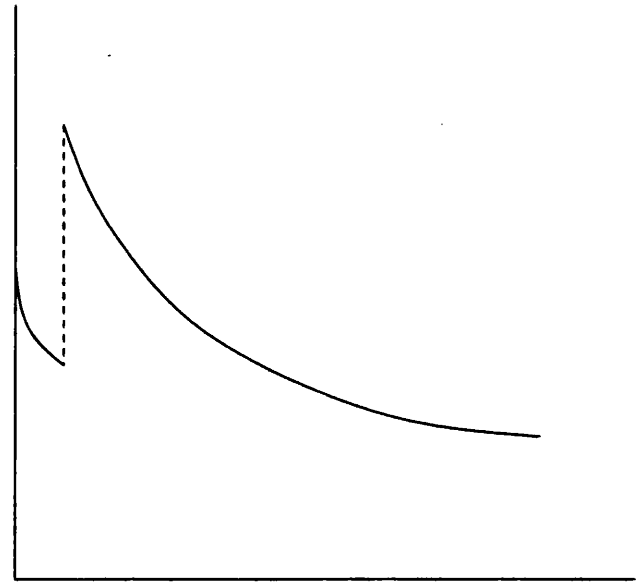
The commitment problem would not arise if the demand of each buyer was inelastic, since, in that case, buyers could be compensated for a high P_2 by a lower P_1 without any adverse effect (from the incumbent's point of view) for the average price. The commitment problem would only arise in the extreme sense considered by Farrell and Gallini (1988), whereby if the fixed cost F faced by buyers is very high, then there is no market at all, since buyers would demand a negative period 1 price, which is ruled out by assumption.

We mentioned above that increasing n can result in CS_2-F>0, so that it is no longer the case that CS_1+CS_2-F=0. This is true for a small enough F_0. This raises the following possibility.

Proposition 2: The incumbent may invite rivals so as stop sustaining buyers altogether. In this case, the optimal number of rivals is finite.

As an example, consider the example below with a=0.5, F_0=0.05 and λ=0.65:

Figure 2.2: The extreme case of an incumbent inviting rivals so as to enable all buyers to make a positive net surplus



For n large enough, we go from CS_2-$F \leq 0$ to CS_2-$F>0$, leading to a 'regime shift'. In this particular example, inviting 1 rival is enough to cause this regime shift ($n>0.04$ will do it, in fact) and results in higher profits than if the incumbent operates alone. Clearly inviting any more rivals must cause profits to decline (via the standard competition effect), so the incumbent invites just 1 rival. There is an internal optimum.

Intuitively, without competition in period 2, the incumbent has to charge a very low p_1 to make it possible for buyers to cover their fixed cost of entry F_0. Essentially, as we have seen, the incumbent pays for the fixed cost of buyers. Up to now, we have considered only parameter values for which CS_2-$F \leq 0$, in which case inviting rivals helps the incumbent by reducing the burden of sustaining buyers. In the example of the diagram above, inviting 1 rival is enough to cause CS_2-$F>0$, and the incumbent benefits from the fact that it no longer has to sustain buyers at all, i.e. the burden is completely lifted.

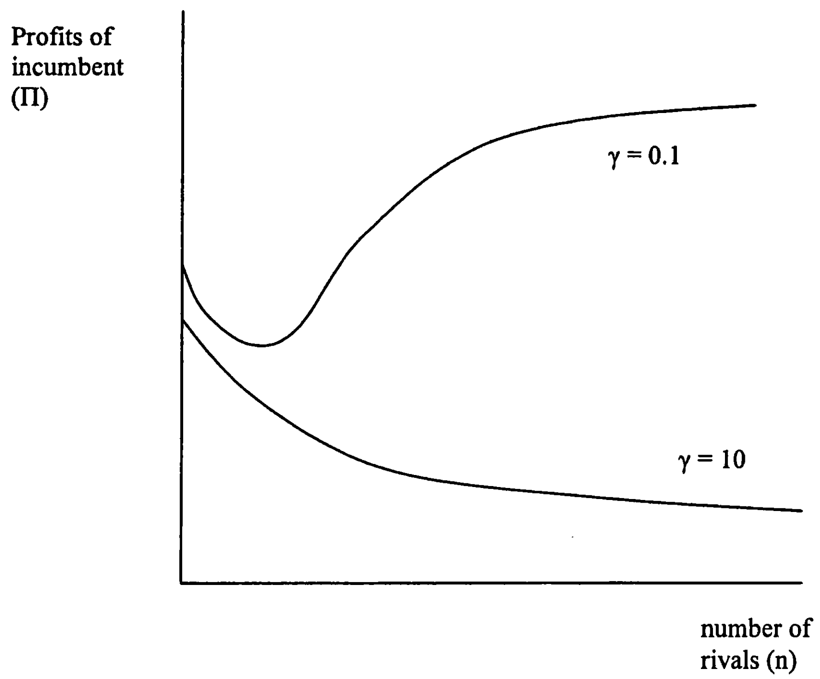
With heterogeneous buyers, we know that CS_1+CS_2-F=0 always. The comparative statics of F_0, λ and a are essentially the same as those of the homogeneous buyers case.

As for the effect of F_1, it is as one would expect: a higher F_1 lowers profits of the incumbent, but leaves the shape of the profit profile unaffected. As for the effect of γ, which increases with the degree of heterogeneity of buyers, we have the following, intuitive result.

Proposition 3: It is possible that an incumbent invites rivals only if buyers are sufficiently homogeneous in their fixed cost of entry.

A low enough γ makes profits of the incumbent rise with the number of rivals, and a high enough γ makes them fall. Intuitively, a high γ means that few extra buyers will enter when more rivals enter, so that overall profits of the incumbent fall; therefore the incumbent will not invite rivals when γ is high. Here is an example with a=1, F0=0.01, F1=0.001, λ=1:

Figure 2.3: How the relationship between profits of the incumbent and the number of rivals in period 2 is affected by the degree of buyer heterogeneity



It is worth mentioning that profits of rivals can rise with the number of rivals. This is the case, for example, with a=1, F_0=0.1, F_1=1, λ=0.1, γ=0.02, in the range n=[1,10]. This result requires buyers to be heterogeneous: if they were homogenous, the number of buyers could not rise with the number of rivals, and so there would be no possibility of profits of rivals increasing. Profits of rivals are more likely to rise with the number of entrants if F_0 is high, since in this case the commitment problem is severe; this implies that the presence of more firms and thus a lower price in period 2 boosts demand significantly, causing many buyers enter.

This result has a number of implications. Firstly, given that the incumbent benefits from entry of rivals, it implies that entry may be advantageous for all firms. Since, in addition, it is advantageous for buyers, it may be advantageous for all agents. Thus, at least in theory, increased competition can benefit all economic agents (without any redistribution among them). This cannot happen in a 1-period model, since there is no commitment problem.

Secondly, the result may help to explain the phenomenon of sub-licensing. Sub-licensing is observed in second-source agreements in the electronics industry, in particular in the semiconductor industry (Swann 1987), at very modest royalty fees. The commitment explanation of this paper may explain this phenomenon: the incumbent may allow sub-licensing so as to achieve the number of rivals it finds most profitable. The incumbent may allow sub-licensing instead of approaching sub-licensing firms directly because it may not have had contacts with the sub-licensing firms in the past, whilst a licensee might have; thus the transaction cost of licensing may fall via sub-licensing.

The model can give rise to multiple equilibria due to the self-fulfilling expectation of rivals. Suppose that rival firms expect that not enough buyers will enter in period 1. Then rival firms too will not enter in period 2, causing the period 2 price to stay high; as a consequence, buyers do not enter, validating the expectations of rival firms. A similar argument holds in the case where rivals have optimistic expectations. As multiple equilibria are easier to show in the case of the decreasing returns to scale problem of chapter 3 than in the case of the commitment problem of this chapter[7], we leave this discussion for Chapter 3.

[7] Fundamentally, this is because showing the presence of multiple equilibria in the case of the commitment problem requires that there be more than 1 potential rival entrant, raising issues of co-ordination among rivals.

4 Entry of rivals in period 1

Why should a monopolist incumbent give up its monopoly status in period 1? At first sight, we may think that this is illogical, since the problem of not being able to commit affects the second period, and thus it would seem sensible to require a remedy also in the second period, that is, entry of rivals in period 2.

However, this argument ignores the fact that the commitment problem, which arises in period 2, has repercussions in the first period. In particular, to induce a sufficiently large entry of customers, the incumbent may need to produce at price below marginal cost in period 1. If period 1 profits are sufficiently negative, then inviting rivals in period 1 enables the incumbent to 'share out' some of its losses. Remarkably, we can find parameter values such that the rival wants to enter in period 1, rather than wait until period 2, because doing so means that (plenty) more buyers enter in period 1.

We assume heterogeneous buyers, to ensure that (5) always holds.

Proposition 4: The incumbent may invite rivals to enter in period 1. It does so if it faces sufficiently large losses in period 1.

Let us begin by differentiating the incumbent's profits with respect to the number of period 1 firms m:

$$\frac{d\Pi}{dm} = p_1 \frac{dq_1}{dm} + q_1 \frac{dp_1}{dm} - c\frac{dq_1}{dm} + p_2 \frac{dq_2}{dm} + q_2 \frac{dp_2}{dm} - c\frac{dq_2}{dm} \tag{22}$$

Re-writing $\frac{dp_1}{dm}$ as $\frac{dp_1}{dp_2}\frac{dp_2}{dm}$ and applying the see-saw effect equation (7) to (22), we obtain

$$\frac{d\Pi}{dm}=\underset{-}{\frac{dp_2}{dm}}\left(-nq_2-n_A\underset{+}{\frac{dF}{dn_A}}\,\underset{-}{\frac{dn_A}{dp_2}}\right)+\underset{-}{\frac{dq_1}{dm}}\underbrace{(p_1-c)}_{+\,or\,-}+\underset{-}{\frac{dq_2}{dm}}\underbrace{(p_2-c)}_{+} \tag{23}$$

Therefore, starting from a situation without rivals (n=0), the incumbent has an incentive to invite rivals provided that $(p_1\text{-}c)<0$, i.e. if it is incurring losses in period 1, and if the period 1 losses are big enough. This would be the case if, for example, F_0 is big enough.

Additionally, if the incumbent incurs a loss in period 1, then he has a greater incentive to attract competition in period 1 than he does in period 2. This can be seen as follows.

$$\frac{d\Pi}{dm}=\left.\frac{\partial\Pi}{\partial m}\right|_{n_A}+\frac{\partial\Pi}{\partial n_A}\frac{dn_A}{dm} \tag{24}$$

Next, we write profits as follows:

$$\Pi=\Pi_1(q_1,n_A,m)+\Pi_2(q_2(n_A),n_A,m+n) \tag{25}$$

where Π_1 is profits in period 1 and Π_2 is profits in period 2. n_A is determined by (5), which can be written more extensively as $CS_1(q_{A,1}(q_1,n_A,m)) + CS_2(q_{A,2}(q_2,n_A,m+n))-F(n_A)=0$. Equation (25) is simply the general form of equation (17) for any concave utility function $U(q_{A,t})$. Notice that m affects profits in (25) via 2 more channels than n does: directly in Π_1, and indirectly by affecting CS_1 and therefore n_A in (5).

When the incumbent operates alone in period 1 (m=1), $\frac{\partial\Pi}{\partial n_A}=0$. This is because maximizing profits with respect to n_A is equivalent to maximizing profits with respect to q_1. This can been from (5): bearing in mind that $q_{A,1}=\frac{mq_1}{n_A}$ and that

$q_{A,2} = \dfrac{(m+n)q_2}{n_A}$, (5) shows that, after q_2 is determined in period 2 as a function of n_A, only q_1 and n_A are left to be determined; therefore, fixing q_1 also fixes n_A, and viceversa. Thus, for m=1, we can use (24) and (25) to write

$$\begin{aligned}\frac{d\Pi}{dm} = \left.\frac{\partial\Pi}{\partial m}\right|_{n_A} &= \left.\frac{\partial\Pi}{\partial n}\right|_{n_A} + \left.\frac{\partial(\Pi_1)}{\partial m}\right|_{n_A} \\ &= \frac{d\Pi}{dn} + \left.\frac{\partial(\Pi_1)}{\partial m}\right|_{n_A}\end{aligned} \tag{26}$$

where $\dfrac{d\Pi}{dn} = \left.\dfrac{\partial\Pi}{\partial n}\right|_{n_A}$ by virtue of the fact that $\dfrac{\partial\Pi}{\partial n_A} = 0$. If $\Pi_1<0$, then $\left.\dfrac{\partial\Pi_1}{\partial m}\right|_{n_A} > 0$, and $\dfrac{d\Pi}{dm} > \dfrac{d\Pi}{dn}$. In particular, if $\dfrac{d\Pi}{dn}$ is negative but small, $\dfrac{d\Pi}{dm}$ can still be positive. In this case we not only have a negation of what Farrell and Gallini suggested ("Obviously, second-sourcing that enables firms to compete with the innovator on equal terms immediately (and without royalty payments) is not profitable"), but a reversal; that is, the incumbent invites rivals only if they can enter without delay!

Consider the following example with a=0.15, λ=1, F_0=0.01, F_1=0.001 and γ=0.2:

Figure 2.4: Profits of the incumbent and the timing of entry of rivals

	Number of firms (= 1 + number of rivals)														
	1	2		3			4				5				
	m=1 n=0	m=1 n=1	m=2 n=0	m=1 n=2	m=2 n=1	m=3 n=0	m=1 n=3	m=2 n=2	m=3 n=1	m=4 n=0	m=1 n=4	m=2 n=3	m=3 n=2	m=4 n=1	m=5 n=0
Profits of m firms	46.9	29.6	49.4	35.2	40.0	46.6	44.5	38.8	41.0	43.6	54.6	39.8	38.2	39.0	40.1
Profits of n firms	-	49.2	-	33.2	49.9	-	26.4	39.9	46.0	-	22.0	33.4	38.6	41.5	-

The above example shows the best strategy of the incumbent given the number of potential rivals available. The underlined numbers indicate the best strategy of the incumbent given the number of potential rivals. Profits in the first period when the incumbent operates alone are –104.3, reflecting a strong incentive to invite rivals in the first period. If there is only 1 potential rival available (so that the number of firms is 2), then it's best to invite him to enter in period 1 (m=2, n=0). The rival also wants to enter in period 1 (because 49.4 > 49.2), so there is no problem of incentive compatibility.

If there are 2 potential rivals (so that the number of firms is 3), the incumbent still only wants to invite only 1 rival in the first period. The same strategy is optimal with 3 potential rivals. With 4 potential rivals, however, the optimal strategy is to invite all 4 to enter in the second period (m=1, n=4). This is because inviting a large number of rivals in period 2 alleviates the commitment problem significantly and either makes period 1 losses very small or turns them into gains (in the example, they turn into gains), thus removing the "loss-sharing" incentive to invite rivals in period 1. When the potential number of rivals is infinite, and it is profitable for the incumbent to invite them, he will invite them in period 2. This example makes clear that the optimal time of entry and the optimal number of rivals from the incumbent's point of view depend on the number of potential rivals.

What happens to the incumbent's optimal strategy as F_0 changes? For a high enough F_0, profits are negative for any number of firms, so there is no market (e.g. with F_0=0.1). For a small enough F_0, the incentive to invite rivals in period 1 disappears, as period 1 profits are only slightly negative or even positive, but the incentive to invite 4 rivals in period 2 remains (e.g. with F_0=0.001).

The model, by showing that it can be optimal for the incumbent to invite rivals in either period[8], may explain the second-sourcing phenomenon observed in the semiconductor industry whereby some original sources invite rivals to enter at the time of product launch (Dick(1992)), and others invite rivals to enter some time afterwards (Swann (1987))[9]. It is common to find that innovators incur a loss in the initial phases just after the product has been launched; whilst this may be a result of fixed cost such as R&D, it may also be a result of the commitment problem which we explore here, and which can be alleviated by inviting rivals to enter in period 1. The theory thus calls for an empirical study of how the timing of entry of rival firms is linked with the sign and size of profits at the time of product launch.

A related case is that of Apple Computers which, in 1994, chose to start licensing its proprietary operating system (Kende(1995a), who refers to Business Week, September 26, 1994). Many commentators argued that Apple ought to have taken this step earlier, to expand the client base more and at an earlier time, so as to encourage software producers to write software for Apple, and not just for Microsoft Windows. The results of our model may be interpreted to suggest that who was correct depends on the values of cost parameters; in particular, inviting rivals from the beginning would be optimal if the set-up cost of buyers was very large.

---

[8] Consistently throughout the chapter, indeed throughout the thesis, we disregard the possibility of royalties, which may affect the decision to invite firms in period 1 as opposed to period 2.
[9] It would be interesting to find out whether there are original sources which, for the same product, invite some rivals to enter in period 1, and some in period 2.

5 The social planner

In a first-best scenario, a social planner would maximize welfare by imposing the competitive outcome. The social planner's problem would be to maximize the sum of profits and consumer surplus net of buyers' fixed cost of entry over both periods, taking into account the demand functions of buyers. From the social planner's point of view, the number of firms would be irrelevant, since we are assuming equally efficient firms and constant returns to scale.

The fact that the number of firms would be irrelevant in a first-best scenario makes this analysis somewhat uninteresting. More interesting is to study a second-best scenario in which a social planner can control how many firms can enter (subject to the constraint that all firms make weakly positive profits), for example by forcing the incumbent to reveal its technological know-how and controlling access to the technology via licensing restrictions.

Two interesting results emerge. The first is that, with homogenous buyers, the number of rivals which a social planner finds optimal to invite may be finite. Suppose that entry is restricted to the second period, perhaps because entrant firms need some time to adjust to the new technology[10]. With homogeneous buyers, the surplus of buyers over 2 periods is always zero via the see-saw effect, no matter how much competition there is in period 2. This implies that welfare is simply the sum of industry profits. It follows that the usual deadweight argument whereby an increase in competition raises welfare does not apply. In particular, if industry profits decline as the number of rivals increases, then, because consumer surplus is unaffected, welfare falls. On the other hand, when industry profits rise, welfare too rises. In the examples below, we find that welfare rises with the number of entrants to begin with, but then declines.

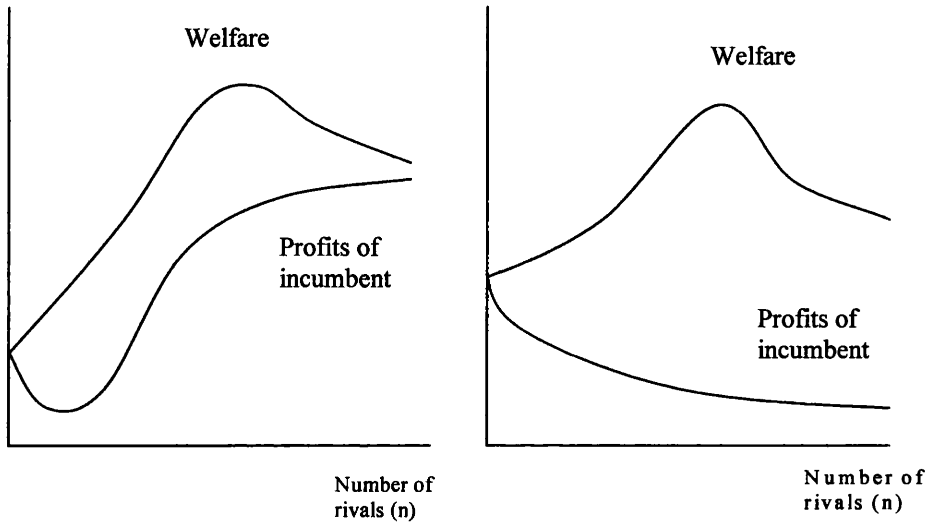
The second interesting result is captured in the following proposition.

[10] The model is not suitable for the study of entry in the first period in the presence of homogenous buyers, since there is no way of determining how much each firm produces in period 1, and therefore no Nash equilibrium.

Proposition 5: The incumbent may invite either too many or too few rivals from a social planner's point of view.

We illustrate both cases below. Let a=0.15, m=1, λ=0.6.

Figure 2.5: Comparing the incumbent's and the social planner's optimal number of rivals



(a) F_0=3 (b) F_0=2

In the graph on the right, the fixed cost $F=F_0$ is large enough to make the commitment problem severe, so that the incumbent invites rivals. The optimal number of rivals from the incumbent's point of view is infinitely large. However, the incumbent does not care that, by inviting more rivals, profits of existing rivals may fall. In the example, the fall of existing rivals' profits is so large that welfare eventually falls. Thus the socially optimal number of rival firms is less than the incumbent's optimal number of rival firms. In the diagram on the left[11], the fixed cost $F=F_0$ is smaller, the commitment problem is less severe, and the incumbent does not invite any rivals. From the social planner's point of view, however, the number of entrant rivals is positive, reflecting the fact that the incumbent is not concerned with enabling rivals firms to enjoy a profit.

[11] The vertical scales of the two diagrams are not comparable.

6 Conclusion

Fundamentally, the analysis points to a relationship of symbiosis between producers. This means that producers benefit from each other's presence without co-operating with each other.

This relationship is most evident in the result that an incumbent may benefit from the presence of rivals already from the start of the period in which the new product is launched. This is what we observe in many second-sourcing agreements in the semiconductor industry. We also know that this is done with the full knowledge of buyers, as the model requires.

The possibility that an incumbent may wish to invite rivals as soon as the product is launched is probably the most significant contribution of the model of this chapter to our understanding of the commitment mechanism by which an incumbent invites rivals. The other significant contribution is that the model identifies the optimal number of buyers from the incumbent's point of view, which is not possible in Farrell and Gallini (1988) because they assume Bertrand competition (with a homogenous good).

Indeed, the main model of section 2 of this chapter is the Cournot counterpart of the Bertrand model of Farrell and Gallini (1988), so the two papers serve as benchmarks. Since, with a homogeneous good, the Cournot case with an infinite number of firms and the Bertrand case with any number of rivals are equivalent, the case of Farrell and Gallini (1988) is a special case of our model.

There appear to be promising avenues of future research. The first is to study what the implications are of having producers of varying efficiency. Would an incumbent want to attract a low or a high efficiency producer? This is an important question, since there is evidence that original source producers are keen to attract certain rivals but not others. The analysis suggests that an incumbent may prefer to attract more efficient rivals, contrary to conventional thinking. A rigorous study of this question requires formal modelling.

The second possible area of further research is to allow the incumbent to charge royalties. Economides (1996a) shows that, when the direct network effect is large enough and thus the incentive to invite rivals is sufficiently strong, royalties can be negative. In our model, it would be useful to study how royalties affect the decision to invite rivals in one period rather than another.

The third avenue of research is that of R&D. The huge literature on R&D and competition shows that the effect of the latter on the former can be either positive or negative, depending on the extent to which an innovator can appropriate the returns on its innovation (Reinganum (1989)). Many results in the literature may well be reversed in the setting of our model in which more competition generates higher profits. This analysis would be useful in view of the increasingly important role which joint R&D exchange programmes, in which firms co-operate at the research stage but subsequently compete in the product market, have played in second-sourcing agreements from the mid 1980's onwards.

Chapter 3

Why firms invite rivals: the capacity constraint problem

1 Introduction

We saw in the previous chapter that an incumbent may invite rivals to overcome a commitment problem. In this chapter, we explore a different explanation as to why an incumbent may want to invite rivals, namely to alleviate a capacity constraint.

By inviting rivals, who are similar to the incumbent in that they too also face a capacity constraint, the incumbent can relax the capacity constraint at the industry level. This enables consumers in the second period to enjoy a lower price than that which the incumbent would have found profitable to achieve on its own. The incumbent gains by having a lower period 2 price via the see-saw effect explored in the previous chapter: consumers are willing to tolerate a higher price in period 1, and this positive effect more than outweighs the negative effects of a lower price and reduced market share in the second period, so that profits over both periods increase.

There is an additional motivation to explore the capacity constraint incentive to invite rivals. In the case of the commitment problem of the previous chapter, we saw that, with linear demand, an incumbent invites as many rivals as possible (except in the special case where there is a "regime shift"), ideally an infinite number. But in practise the number of rivals can be just a few. Swann (1987) states that, paradoxically, it may be in the original producer's interest to attract some second source producers, but only some, for beyond a certain point the original firm finds itself against "unwanted competition". This can be simply because there are only a few rivals available. Alternatively, as the model of this paper shows, it could be an endogenous outcome, the result of the profit maximization problem of an incumbent facing a capacity constraint.

For mathematical ease, wee model capacity constraint as decreasing returns to scale, which are a smooth, mild form of capacity constraint (Tirole, 1988). The assumption of decreasing returns to scale is appropriate in the semiconductor industry. Laboratory space and equipment constitute a factor of production which is fixed in the short to medium term. Furthermore, the short life cycle of microchips may act as a disincentive to the increase in size of laboratory capacity, since the firm may not be able to earn returns on these investments for long enough to justify expenditure on them.

Economides (1996b) goes as far as to include limited capacity in his definition of second-sourcing: "'Second-sourcing' occurs when a firm with a unique product but with a limited manufacturing capacity allows other firms to produce its product under license" (Economides (1996)). Dick(1992) reports that "microprocessor producers often have found themselves with insufficient internal capacity to satisfy downstream demand". Second-sourcing may be particularly attractive to a multinational corporation, since it may enable the multinational to launch a product on a specific market without any initial foreign direct investment, leaving itself the option of investing in its own plant should demand for its product turn out to be sufficiently large (UN(1986)). According to Varian (2001), whilst there are capacity constraints in every production process, for information goods, and particularly for chips, this is "the baseline case".

The assumption of limited capacity is central to Dick (1992)'s analysis of second-sourcing, where firms must choose capacity before uncertainty is resolved. In his model, firms sign a contract: whichever firm successfully innovates, it second sources production at a pre-agreed price. This enables each firm to select an ex-ante level of capacity that is lower and more profitable than that without contracting. The model of this chapter does not require that firms be able to sign contracts, unlike Dick (1992).

One of the explanations that we find in the empirical literature on semiconductors as to why an incumbent invites rivals is that it wishes to increase the demand for 'supporting products' produced by the original source. Often, microchips form groups whose members are in some way technologically complementary, so purchasing one microchip reduces the investment cost of another microchip in the same group (UN (1986), Swann (1987)).

The explanation offered in this paper has similarities with this motivation for inviting rivals. Indeed, we may think of the two periods of our model as two distinct markets whose products are complementary. Such products include computer hardware and software, video game consoles and cartridges, printers and printer cartridges, etc.. When competition increases in the second period, this is like competition increasing in one of the two markets. The lower price in the second period market stimulates demand for the complementary product in the first period market.

The chapter is structured as follows.

In section 2, we set up the model, and consider the incumbent's decision to invite rivals to enter in period 2.

In section 3, we consider the incumbent's incentive of inviting rivals in period 2.

In section 4, we discuss the possibility of multiple equlibria due to the self-fulfilling expectations of rivals.

In section 5, we draw some conclusions.

2 Model

The demand side of the model, which gives rise to the see-saw effect, is the same as in the chapter about the commitment problem. The production side, however, is different. We assume that the incumbent can commit in period 2 to its quantity in period 1, so that there is no incentive to invite rivals to alleviate a commitment problem. Instead, the incentive is to alleviate a capacity constraint problem, which we introduce by assuming decreasing returns to scale, rather constant returns as in the chapter about the commitment problem.

For ease of exposition, we present below the demand side, which we already introduced in the chapter about the commitment problem.

Buyers are assumed to be price-takers. We assume buyers' utility to be $V = U(q_{A,t}) + y_{A,t}$, where $q_{A,t}$ denotes the quantity chosen by buyers in time period $t = \{1,2\}$ and $y_{A,t}$ is a numeraire good. Given income $I = p_t q_{A,t} + y_{A,t}$, where p_t is the price in period t, the maximization problem $\underset{q_{A,t}, y_{A,t}}{Max}\ U(q_{A,t}) + y_{A,t} - \Lambda(I - p_t q_{A,t} - y_{A,t})$ is equivalent to

$$\underset{q_{A,t}}{Max}\ U(q_{A,t}) - p_t q_{A,t} \qquad (1)$$

by virtue of the fact that Λ=1. The maximization problem is intuitively appealing in that it simply involves maximizing the gap between benefits and costs. Buyers can be either firms of consumers. If they are firms, then they buy an input in production, for example a computer hardware manufacturer may buy a semiconductor component.

The resulting demand function is

$$\frac{dU}{dq_{A,t}} = p_t \tag{2}$$

Critically, each buyer faces a fixed cost of entry. If buyers are firms and the good in question is an input in production, then this cost could be an adaptation cost of adjusting the production process to accommodate the new input. If the good in question is a consumer good, it could be the cost of learning how to use the new product, or of having to buy complementary equipment so as to use the new product.

Buyers can be either homogeneous or heterogeneous in their fixed cost of entry, F_i. Suppose buyers are heterogeneous. Arranging buyers in order of increasing fixed cost of entry, we assume

$$F_i = F_0 + F_1 r^\gamma \tag{3}$$

where F_i is the fixed cost of buyer i, and r is his rank number in the sequence of buyers. In equilibrium, the marginal buyer's fixed cost F is

$$F = F_0 + F_1 n_A{}^\gamma \tag{4}$$

where n_A is the number of buyers, which is determined in period 1. The marginal buyer just breaks even, i.e.:

$$CS_1 + CS_2 - F = U(q_{A,1}) - p_1 q_{A,1} + U(q_{A,2}) - p_2 q_{A,2} - F = 0 \tag{5}$$

where CS_t is consumer (buyer) surplus in period t.

If buyers are homogeneous ($F_1=0$), then (5) applies to all of them as they all face the same fixed cost. In this case, F_0 needs to be sufficiently large for (5) to hold[1]. As long as consumer surplus in both periods is weakly positive, then buying takes place in both periods. Now, because (5) says that buyers must purchase in both periods to break even, it follows that a buyer would incur a net loss if he purchased in only one period. Thus buyers must enter in period 1. This implies that no buyers enter in period 2, simplifying the analysis considerably.

Totally differentiating (5) enables us to identify the see-saw effect:

$$\frac{dU}{dq_{A,1}} dq_{A,1} - p_1 dq_{A,1} - q_{A,1} dp_1 + \frac{dU}{dq_{A,2}} dq_{A,2} - p_2 dq_{A,2} - q_{A,2} dp_2 - \frac{dF}{dn_A} dn_A = 0 \tag{6}$$

Applying the first order conditions (2) to (6), some terms cancel out and we have the central equation of the this chapter:

$$-q_{A,1} dp_1 - q_{A,2} dp_2 - \frac{dF}{dn_A} dn_A = 0 \tag{7}$$

The see-saw effect is most clearly seen when buyers are homogenous. In this case $dF/dn_A=0$, and therefore

[1] With logarithmic utility, (5) holds in the case of homogenous buyers even for a small F_0 because the elasticity of demand is constant at -1 and so the incumbent focuses on minimizing costs by reducing quantity to the bare minimum (this being given by (5)).

$$\frac{dp_1}{dp_2} = -\frac{q_{A,2}}{q_{A,1}} \tag{8}$$

Thus the relationship between prices in different period is negative.

In the specific examples which we discuss below, we wish to consider the classic case of linear demand. This implies that we assume the quadratic utility function

$$U = aq_{A,t} - \frac{\lambda}{2} q_{A,t}^{\ 2} \tag{9}$$

The use of a quadratic utility function makes the period 1 optimization algebraically cumbersome, as we shall see. We experimented with different variants of quadratic utility with the aim of making the algebra as simple as possible, but could not find a superior alternative to (9). Inserting (9) into (1), we obtain that buyers in each period t solve the problem

$$\underset{q_{A,t}}{Max} \quad CS_t = aq_{A,t} - \frac{\lambda}{2} q_{A,t}^{\ 2} - p_t q_{A,t} \tag{10}$$

The resulting inverse demand function is

$$p_t = a - \lambda q_{A,t} \tag{11}$$

A higher λ therefore implies a less elastic demand schedule.

In period 1, there is only 1 firm, the incumbent. This is because the incumbent has no incentive to invite rivals in period 1. This in turn is because it never incurs a loss in period 1, so the "loss-sharing" motivation to invite rivals to enter in period 1 which we found in the case of the commitment problem is absent. It can easily be shown that the incumbent never makes a loss in period 1: assuming it can commit in period 1 to its period 2 quantity, it makes as much profit in period 1 as it does in period 2, and because it makes positive profits over both periods, it must make positive profits in period 1. As a consequence, the incumbent invites rivals to enter only in period 2.

We assume that there are ways for the incumbents to commit in period 1 to what its period 2 quantity $q_{2,M}$ will be, for example by signing an enforceable contract with buyers. Thus, the incumbent is a Stackleberg leader who moves in period 1, and all entrants are followers who move in period 2. This removes the commitment incentive to invite rivals explored in the previous chapter. Note that $q_{2,M}$ will be different from that of the rivals, q_2, since the incumbent chooses $q_{2,M}$ in period 1 with a view to inducing buyers to enter.

The incentive to invite rivals is given by the assumption of decreasing returns to scale, which is a mild form of capacity constraint. All firms face decreasing returns to scale, which we capture by means of quadratic costs. With the incumbent's choice of output $q_{2,M}$ being given in period 2, rivals firms face the following problem:

$$\begin{aligned} &\underset{q_2}{Max} \quad p_2 q_2 - \frac{\alpha}{2} q_2^{\ 2} \\ &s.t. \quad (11),\, q_{A,2} = \frac{q_2 + s_2 + q_{2,M}}{n_A} \end{aligned} \tag{12}$$

where s_2 is the combined output of the remaining n-1 rivals, n is the total number of rivals and n_A is the number of buyers, which determined in period 1 and so is exogenous in period 2[2]. In equilibrium $s_2 = (n-1)q_2$, and so the equilibrium value of q_2 is

[2] Note that a firm in period 2 will not find it profitable to deviate from the Nash equilibrium in order to attract more buyers (by increasing quantity and thus make the price fall so much that new buyers enter with fixed costs larger than those of existing buyers). This is because the incentive which a firm in period 2 faces is smaller than the incentive which the incumbent in period 1 faces when deciding how many buyers to attract: the incumbent in period 1 sells to the same buyers over both periods, not just in the second.

$$q_2 = \frac{\frac{an_A}{\lambda} - q_{2,M}}{n+1+\frac{\alpha n_A}{\lambda}} \tag{13}$$

In period 1, the incumbent chooses the quantities of output in both periods and then buyers incur their fixed cost of entry and enter. The incumbent in period 1 faces the following problem:

$$\begin{aligned} &\underset{q_1, q_{2,M}}{Max} \quad \Pi = p_1 q_1 - \frac{\alpha}{2} q_1^{\ 2} + p_2 q_{2,M} - \frac{\alpha}{2} q_{2,M}^{\ \ 2} \\ &s.t. \quad (5), (11), (13), q_{A,1} = \frac{q_1}{n_A}, q_{A,2} = \frac{q_{2,M} + nq_2}{n_A} \end{aligned} \tag{14}$$

The solutions for linear demand are complicated and simulations are necessary. Specifically, $q_{2,M}$ can be shown to be a cubic in n_A; and n_A in turn is given by a polynomial of order greater than three in $q_{2,M}$. However, we can obtain some general results for any concave utility function $U(q_{A,t})$.

3 Inviting rivals to enter in period 2

Let us begin by not making any assumptions about the nature of returns to scale. Profits of the monopolist are

$$\Pi = p_1 q_1 - TC_1 + p_2 q_{2,M} - TC_2 \tag{15}$$

where TC_1 and TC_2 are total cost in periods 1 and 2, respectively. Next, noting that q_1 and $q_{2,M}$ are both control variables of the incumbent in period 1, we can use the envelope theorem to write

$$\frac{d\Pi}{dn} = \left.\frac{\partial\Pi}{\partial n}\right|_{q_1,q_{2,M}} = p_1\frac{\partial p_1}{\partial n} + q_{2,M}\frac{\partial p_2}{\partial n} \tag{16}$$

Recognizing that, when n changes, the effect on p_2 and p_1 is via the change in nq_2, we can re-write the above expression as

$$\frac{d\Pi}{dn} = p_1\frac{\partial p_1}{\partial nq_2}\frac{\partial nq_2}{\partial n} + q_{2,M}\frac{\partial p_2}{\partial nq_2}\frac{\partial nq_2}{\partial n} = \frac{\partial nq_2}{\partial n}\left[p_1\frac{\partial p_1}{\partial nq_2} + q_{2,M}\frac{\partial p_2}{\partial nq_2}\right] \tag{17}$$

Now consider the FOC of profits with respect to $q_{2,M}$:

$$\frac{\partial\Pi}{\partial q_{2,M}} = \frac{\partial p_1}{\partial q_{2,M}} + p_2 + \frac{\partial p_2}{\partial q_{2,M}}q_{2,M} - \frac{\partial TC_2}{\partial q_{2,M}} = 0 \tag{18}$$

Next, we recognize that

$$\frac{\partial p_2}{\partial q_{2,M}} = \frac{\partial p_2}{\partial nq_2} \quad and \quad \frac{\partial p_1}{\partial q_{2,M}} = \frac{\partial p_1}{\partial nq_2} \tag{19}$$

Both equalities are due to the fact that what really matters in determining prices is the total period 2 output $q_{2,M} + nq_2$; so it does not matter which of the two changes, $q_{2,M}$ or nq_2, as long as they change by the same amount the impact on prices will be the same (both directly and indirectly via the number of buyers n_A).

Combining (17), (18) and (19), we have

$$\frac{d\Pi}{dn} = \frac{\partial nq_2}{\partial n}\left[-p_2 + \frac{\partial TC_2}{\partial q_{2,M}}\right] \tag{20}$$

Recognizing that $\frac{\partial nq_2}{\partial n} > 0$, we have the following result.

Proposition 1: Profits of the incumbent rise with the number of rivals if in period 2 its marginal cost exceeds price.

The intuition behind this result is as follows. An incumbent can lower its period 2 output by 1 unit and invite enough rivals to ensure that total output in period 2 stays constant (when considering infinitesimal changes, $\partial q_{2,M}=\partial(nq_2)$). This implies that p_2 stays constant. In addition, because p_2 is constant, the number of buyers who enter in period 1 is unaffected. Since p_2 is unaffected, the incumbent loses exactly p_2 in that it sells 1 unit less in period 2. The incumbent gains from producing one unit less of $q_{2,M}$ in that costs decline by the size of its marginal cost, which we call MC_M (M stands for monopoly, the incumbent's status in period 1). Thus the incumbent's net gain is MC_M-p_2. The reason why this net gain is scaled down by the factor $\partial(nq_2)/\partial n$ in (20) is that we are interested in the effect of an increase in n, rather than in an increase in nq_2.

How is it possible for price to exceed marginal cost? The key to this puzzle lies in recognizing that it refers to the marginal cost of the incumbent, not of rivals. p_2 is always greater than the marginal cost of rivals: after all, rivals in period 2 must al least break even. But the marginal cost of the incumbent is greater than that of rivals. This is because the marginal benefit is also greater. The marginal benefit of $q_{2,M}$ includes the positive effect of enabling the incumbent to reduce q_1 while maintaining the number of buyers constant via the see-saw effect equation (5). That is, the incumbent gains in that the burden of sustaining buyers is reduced.

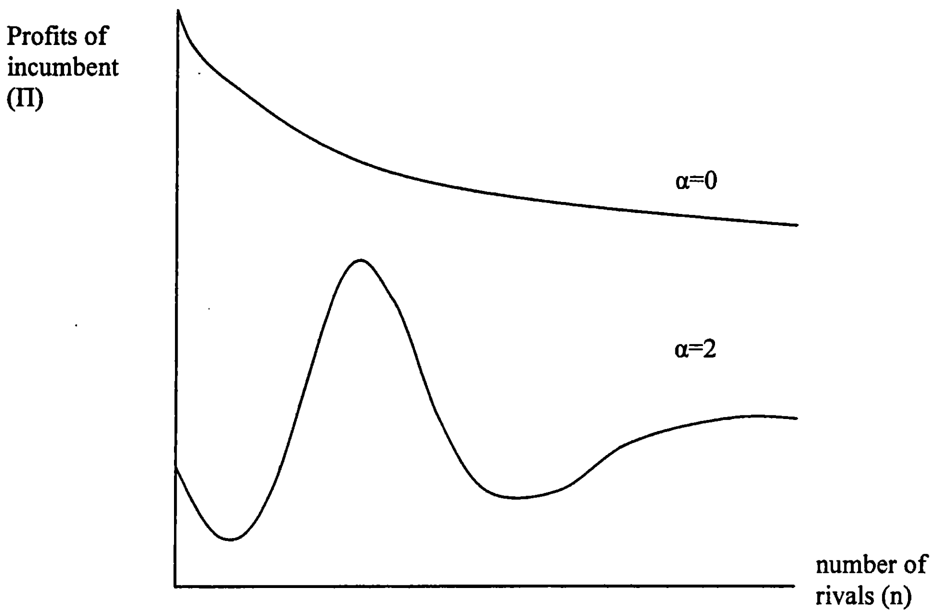
What would happen with constant returns to scale, and equally efficient firms? Now the marginal cost of the incumbent is equal to that of the rival(s), and because p_2 is always more than the MC of the rivals, it follows that p_2 is always more than the MC of the incumbent. Proposition 1 tells us that in this case the incumbent always loses if a rival enters, so it will never invite a rival. In particular, if α=0, MC_M is constant (and equal to zero), and it can immediately be seen from (20) that the incumbent never invites a rival. This discussion highlights the pivotal role that decreasing returns to scale play in motivating an incumbent to invite rivals. If rivals are more efficient than the incumbent, the result suggests that it might be profitable for the incumbent to invite rivals even under constant returns to scale; this is a promising line of further research.

Note that this result is more general than our model suggests. In particular, because the incumbent is assumed to be able to commit, the period 1 and period 2 markets could be interpreted as two markets for two complementary products in the same period, e.g. hardware and software. The question then arises, in which market would the incumbent wish to invite rivals? Is it the one in which MC_M-p_2 is greatest? Is it possible for the degree of complementarity between the two markets to be so high that an incumbent may invite rivals in both markets? Can entry in one market turn MC_M-p_2 in the other market from negative to positive? Again, we leave this as a topic for future research.

In the example below, we illustrate the point made above that the decreasing returns to scale (α>0) are necessary for the incumbent to invite rivals. Note that a and c always occur together as the term (a-c). So we can just set c=0 and focus exclusively on revenues. This is a consequence of assuming linear demand.

Consider the example below with a=2, λ=1.10, F_0=1, F_1=1, m=1, γ=0.2

<u>Figure 3.1: The need for decreasing returns to scale for profits of the incumbent to rise with the number of rivals</u>



As in the case of the commitment problem, sometimes the standard competition effect dominates, so that the profit profile is negatively sloped, and sometimes the decreasing returns to scale mechanism dominates, so that the profit profile is positively sloped. The above diagram indicates, furthermore, that an internal optimum for the number of rivals an incumbent wishes to invite is possible. Therefore, the model of this paper may account for Swann (1987)'s finding that that, paradoxically, it may be in the original producer's interest to attract some second source producers, but only some, for beyond a certain point the original firm finds itself against "unwanted competition".

We next address the question of how the degree of heterogeneity of buyers with respect to their fixed cost of entry affects the incumbent's incentive to invite rivals.

Proposition 2: With homogeneous buyers, profits of the incumbent are always weakly increasing in the number of rivals. With heterogeneous buyers, they can be increasing or decreasing.

This can be shown as follows. Using the envelope theorem, we can write

$$\frac{d\Pi}{dn} = \left.\frac{\partial \Pi}{\partial n}\right|_{q_1, q_2, M} = q_1 \frac{\partial p_1}{\partial n} + q_2 \frac{\partial p_2}{\partial n} \tag{21}$$

Re-writing $\frac{\partial p_1}{\partial n}$ as $\frac{\partial p_1}{\partial p_2}\frac{\partial p_2}{\partial n}$ and applying the see-saw effect equation (7) with $\frac{dp_1}{dp_2} = \frac{\partial p_1}{\partial p_2}$ to (21), we obtain

$$\frac{d\Pi}{dn} = \underset{-}{\frac{\partial p_2}{\partial n}}\left(-nq_2 - n_A \underset{+}{\frac{dF}{dn_A}} \underset{-}{\frac{\partial n_A}{\partial p_2}} \right) \tag{22}$$

With homogenous buyers, so that $\frac{dF}{dn_A} = 0$, we have $\frac{d\Pi}{dn} = 0$ for n=0,[3] and $\frac{d\Pi}{dn} > 0$ for n>0. With heterogeneous buyers, so that $\frac{dF}{dn_A} > 0$, the derivative can be either positive or negative[4]; in particular it is always negative at n=0.

This result has a potentially vast number of applications, as many economic activities, perhaps the majority of them, involve at least a small fixed cost faced by buyers, such as a learning cost. The other requirement for it to be applicable is that the economic activity in question should be characterized by decreasing returns to scale. Various authors have found evidence of decreasing returns to scale in manufacturing, e.g. Bils(1987) using U.S. data.

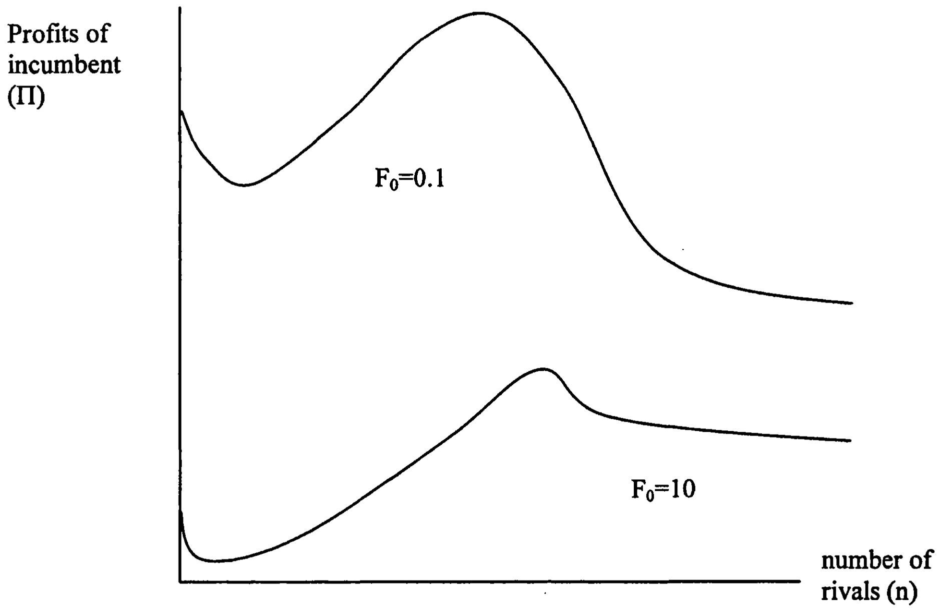
Let us now perform some comparative statics exercises.

Consider the parameter F_0. The simulations confirm the intuition that a higher F_0 increases the incentive for an incumbent to invite rivals so as to reduce the burden of sustaining buyers though a low p_1 and also a low p_2 (by means of a high q_1 and $q_{2,M}$): profits of the incumbent are higher at $n=\infty$ than they are at n=0 in the case of a high F_0, but they are lower in the case of a low F_0. E.g. with a=10, $\lambda=1$, $\alpha=2$, $F_1=1$, m=1, $\gamma=0.15$:

3 Notice that there is a mathematical subtlety here. So far we have not specified whether returns to scale are constant or decreasing. So the result appears to hold even with constant returns, which is intuitively impossible, since the incumbent would not have any incentive to invite rivals. As a matter of fact, we can show that returns to scale cannot be constant. Suppose they were. Then we know that the number of buyers chosen by the incumbent in period 1 is arbitrarily large, say $\overline{n_A}$, and independent of n. If n increases, and n_A does not, then, for given q_1 and $q_{2,M}$ as required in the application of the envelope theorem, $CS_1+CS_2-F>0$; so our model collapses. With decreasing returns to scale, n_A can increase to restore $CS_1+CS_2-F=0$.

4 We can show that, for heterogeneous buyers and $F_0=0$, with linear demand the slope of the profit profile is negative at $n=\infty$.

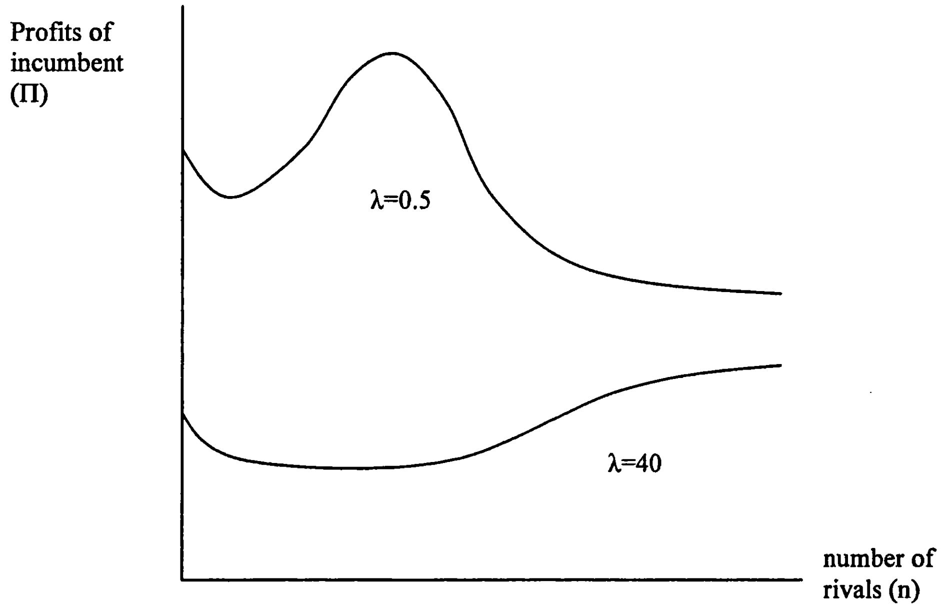
Figure 3.2: How the set-up cost of buyers affects the relationship between profits of the incumbent and the number of rivals



The results of the simulations for different parameter values of a, F_1 and γ are also as one might expect. A low a or a high F_1 make it more costly for the incumbent to sustain buyers through a low p_1 and p_2, and so the incentive to invite rivals is greater. A low γ allows the number of buyers to increase greatly when the number of rivals increases, making it profitable for the incumbent to invite them.

The results of the simulations for different parameter values of λ are less obvious. E.g. with a=10, α=2, F_0=1, F_1=1, m=1, γ=0.2:

Figure 3.3: How the demand parameter λ affects the relationship between profits of the incumbent and the number of rivals



In this example, an internal optimum is possible with a low λ, but not with a high λ. This is because when λ is low, the number of buyers grows with the number of rivals much faster than when λ is high. This in turn is because consumer surplus in period 2 expands faster: a lower λ implies that, for a given number of buyers, increased competition results in a greater increase in $q_{A,2}$. In addition, notice that profit at $n=\infty$ are greater than at $n=0$ for a high λ, but smaller for a low λ. This is because a high λ makes sustaining buyers more costly, and so inviting rivals is profitable for the incumbent (albeit only if there are enough of them available).

As in the case of the commitment problem of chapter 2, we find that profits of rivals can rise with the number of rivals. This is made possible by a sufficiently large increase in the number of buyers, which in turn is made possible by a low enough γ.

Again, as in the case of the commitment problem, we find that an incumbent may want to invite rivals to completely lift the burden of sustaining buyers, rather than just reduce it, which we referred to as a "regime shift". We also find that the number of rivals which the incumbent invites may be higher or lower than the socially optimal number, for reasons already described in the case of the commitment problem.

4 Multiple equilibria

In this section, we show that multiple equilibria are possible because of the self-fulfilling expectations of rivals. Consider the possibility of entry of rival firms in period 2. Suppose that rival firms expect that not enough buyers will enter in period 1. Then rival firms too will not enter in period 2, causing the period 2 price to stay high; as a consequence, buyers do not enter, validating the expectations of rival firms. A similar argument holds in the case where rivals have optimistic expectations.

We assume that all firms face a fixed cost of entry ε. We assume, furthermore, that the incumbent's technology is freely available to all, so that whether rival firms enter or not in period 2 is not something which the incumbent can control.

Suppose that potential entrant firms are optimistic and expect that there will be enough buyers around in period 2 for operational profits to cover the fixed cost ε. Suppose, furthermore, that at the start of period 1 they send a credible signal to the incumbent that they will enter in period 2, such as starting the necessary official paper work. Then the incumbent will select the number of buyers n_A knowing that rival firms will enter.

Suppose now that period 2 entrants are pessimistic and expect that there will not be enough buyers around for operational profits to cover the fixed cost ε. Thus they will not start any official paper work in period 1, the incumbent will receive a credible signal that potential rivals will not enter, which affects its choice of n_A. This n_A will be lower than that chosen by the incumbent had rival firms entered.

For multiple equilibria resulting from self-fulfilling expectations to exist, we require that ε be greater than profits of entrants were they to decide not to enter, and smaller than profits of entrants were they to decide to enter. More simply, we require that profits of entrants were they to decide to enter must be greater than profits of entrants were they to decide not to enter. In this case, we can find a fixed cost that lies between these two profit values which will deliver multiple equilibria.

Proposition 3: Provided that entrant firms face a fixed cost of entry that is neither too big nor too small, multiple equilibria due to the self-fulfilling expectations of rivals may exist.

Here is an example with a=10, α=0.01, F_0=1, F_1=1, λ=0.1, γ=1:

Figure 3.4: Multiple equilibria

	Number of rivals	n=0	n=1
Rivals are optimistic	Profits of incumbent	6,415	6,167
	Profits of each rival	n/a	2,651
	Number of buyers	37.61	48.11
Rivals are pessimistic	Profits of incumbent	6,415	5,415
	Profits of each rival	n/a	1,915
	Number of buyers	37.61	37.61

In this example[5], we have multiple equilibria if the fixed cost faced by entrant rivals is between 1,915 and 2,651.

The result that multiple equilibria are possible has important implications in reality. It may explain why some products become widely used (they become standards) and others do not: they do not when potential second-source producers do not believe that the product will attract enough buyers. In addition, there is a role for government to co-ordinate expectations.

[5] Notice that we do not require profits of the incumbent to rise with the number of rivals; this is because, in this section, we have assumed that the technology is available to all, so that the number of rivals is a not a control variable of the incumbent. Notice also that in reality the number of buyers must be an integer; imposing this restriction does not affect the results for this example.

5 Conclusion

This paper has made a number of contributions to our understanding of why some firms are observed to invite rivals. The most important are the following.

Firstly, this paper offers a novel explanation as to why an incumbent may invite rivals, namely to overcome a capacity constraint. By inviting rivals to enter in period 2, an incumbent achieves an increase in period 2 consumer surplus which would have been too costly to achieve on its own. This in turn enables buyers to tolerate a higher price in period 1, so that overall profits of the incumbent are higher than without rivals.

Secondly, we identified the second period condition which needs to be met for the incumbent to have an incentive to invite a rival: price must be less than its marginal cost. This is possible given that the incumbent chooses a very large period 2 quantity so as to induce entry of buyers in period 1.

Thirdly, we have addressed the question "how many rivals does an incumbent invite?" when it is in the incumbent's interest to invite rivals. This question is not addressed in the literature. For example, Farrell and Gallini (1988) assume that competition leads to equality between price and marginal cost regardless of how many firms there are, making their model incapable of addressing this question. The model of this paper shows that sometimes a finite, non-zero number of rivals can be optimal for an incumbent. We have therefore formally modelled Swann's intuition that an incumbent may want to invite only a limited number of rivals, for beyond that number it would find itself facing "unwanted competition".

Future avenues of research are as follows. Because the incumbent is assumed to be able to commit, the period 1 and period 2 markets could be interpreted as two markets for two complementary products in the same period, e.g. hardware and software. The question then arises, in which market would the incumbent wish to invite rivals? Is it the one in which MC_M-p_2 is greatest? Is it possible for the degree of complementarity between the two markets to be so high that an incumbent may invite rivals in both markets? Can entry in one market turn MC_M-p_2 in the other market from negative to positive? All these questions can be addressed with a minimal increase in the degree of complexity of the model, so that the benefits of extending the model appear to far outweigh the costs.

Chapter 4

An application to the labour market: competition for trained workers

1 Introduction

The objective of this paper is twofold. First, to present a theory of economic development which explains how an LDC may benefit when local firms poach workers trained by a multinational, the latter being the only source of foreign technology. Second, to study the implications of the *see-saw effect* - the inverse relationship between prices in different periods which is present throughout the thesis - when applied to the labour market. We believe both objectives to be worthy of consideration, in that they combine intellectual novelty and potential practical application.

Let us begin by giving an overview of the theory and of the intuition behind the main result. We consider a multinational[1] which decides to produce in an LDC, bringing with it new technology that is of use to local LDC firms. This technology is embodied in the workers by way of training in the first period, and is transferred to local firms when workers are poached in the second period. In the second period, the multinational and local firms compete for the labour of the trained workers. For simplicity, we assume that there is no product market competition between firms. The absence of product market competition between the multinational and local firms reflects the fact that multinationals often produce in LDC's to supply the markets of more developed nations, whilst LDC firms, particularly when in their infancy, tend to supply only the local market.

The most important result in this paper is that the multinational's profits rise when competition for its trained workers increases. The increased competition may be due either to training being of a more general nature (i.e. its usefulness to local firms is greater), or to there being more non-training firms demanding the services of trained workers. Consider an increase in the generality of training. When we say that training is

[1] There is nothing in the model which obliges us to interpret this firm as a multinational, such as the existence of a branch (headquarters or subsidiary) in another country. Thus the theory is more general than the example suggests.

more general, we mean either that the productivity of a poached worker in local firms is higher, or that there are more non-training firms which can benefit from poaching trained workers[2]. As a consequence, when training is more general, local firms attract more labour from the multinational. Contrary to standard intuition, the multinational's profits increase.

This result hinges on the multinational's ability to lower the first period wage to the lowest possible level consistent with workers leaving the agricultural sector in the first period to work in industry; in particular, whenever the second period wage increases, the multinational can recoup this cost by reducing the first period wage, maintaining the sum of wages over two periods which workers demand in industry at a level equal to the sum of wages in agriculture. We call the inverse relationship between wages in different periods the see-saw effect.

Consider now an increase in competition for trained workers in the second period due an exogenous increase in the productivity of workers in local firms, i.e. training that is more general. This results in an increase in the second period wage, enabling the multinational to lower the first period wage to the point where the sum of wages is back to its original level. It can be shown that the multinational gains from the increase in competition for workers in the second period. This is because in the first period it enjoys a fall in the wage when the quantity of labour it employs is high, whilst in the second period it incurs an increase in the wage (equal in absolute terms to the fall in the first period wage in the simplest model) when the quantity of labour it employs is low. Thus the reduction in the first period wage bill more than compensates for the increase in the second period wage bill. Furthermore, the net reduction in the wage bill exceeds the fall in period two revenues resulting from increased poaching; therefore profits of the multinational increase. Note the essential role which the see-saw effect plays in ensuring that profits of the multinational rise when training becomes more general.

In essence, the see-saw effect operates in this paper when the multinational faces delayed competition for inputs: when the latter increases, and workers take into consideration the sum of wages over the two periods, the multinational enjoys positive repercussions in the first period which outweigh the negative ones in the second.

---

[2] In the terminology of Stevens (1994), it is the degree of transferability that we are allowing to vary, general training being an extreme case of training that it is entirely transferable.

The assumption of decreasing returns to scale, or diminishing returns to labour given that we are considering a single input, is essential in generating the result that more competition increases profits of the incumbent because it guarantees, in the presence of wage-taking firms in period 2, that the second period wage rises with increased competition. Consider what would happen if, instead of diminishing returns, firms faced constant returns. In order to make the comparison between constant and diminishing returns sharp, suppose that firms are equally efficient. Then an increase in the number of rivals in period 2 would not affect the second period wage, and the see-saw effect would not operate[3].

The result that more general training results in higher profits of the incumbent does not require that in the first period there be only one incumbent. Indeed, there could be many incumbents, as we show later when we relax the main assumptions. With many firms present in both periods, the model presented here provides an extension of the benchmark, classic model of diminishing returns with only 1 period, in which it is never possible for profits of incumbent firms to increase when the number of rival firms increases, to a 2-period setting, in which it is possible.

What does training have to do with the see-saw effect? Local firms are assumed not to possess the multinational's training technology. Training provides workers with the skills that make workers in the multinational attractive to local firms, since local firms are assumed not to have the ability to train. Thus training is fundamental to the operation of the see-saw effect. We initially assume, for simplicity, that training is costless, and later consider costly training, which is more realistic. These costs may be 'direct', e.g. training courses, or 'indirect', e.g. a slower speed of production or a reduced quality of the product.

The theory is an application of the capacity constraint problem of chapter 3 to the context of the market of labour. It is not an application of the commitment problem of chapter 2: in the main model of this chapter on labour markets, we assume that workers supply their services inelastically, so they don't care what wage they obtain in each period, as long as the sum of wages is constant. Instead, in chapter 2 consumers' demand is elastic, so they do care about the price in each period; the incumbent, by inviting

---

[3] There is an unlikely scenario in which the multinational may gain from increased competition even under constant returns to scale: when local firms entering in period 2 are more efficient than the multinational and thus they poach all the multinational's workers. Then, if entry of these firms results in the second period wage being higher by virtue of their higher efficiency, the multinational could still gain via the see-saw effect.

rivals, achieves a mix of prices across periods that is superior from the consumer's point of view, and thus reduces the multinational's burden of sustaining buyers through a low first period price.

To the best of our knowledge, no other theoretical mechanism has been proposed hypothesizing that a training firm benefits when an increase in competition for its trained workers takes place. The literature on training has largely ignored the question of how more competition for trained workers affects training firms and thus the quantity of training, focusing instead on the question of whether firms will pay for training that is of a general nature. There are, to the best of our knowledge, only 2 papers in the training literature that deal with an increase in competition for trained workers. The first is Autor (2000) and the second is Stevens (1994).

In Autor (2000), so-called temporary help firms offer general training to good and bad workers, but only good workers accept it since it is more valuable to them by virtue of their higher productivity. Thus temporary help firms are valued by their customers for separating good from bad workers. An increase in market competition is here modelled as a discontinuous shift from a given number of firms environment to one with free-entry. More competition is assumed to lower profits to zero, in sharp contrast with the main result of this chapter. Training is increased as firms must, in a competitive environment, ensure that workers' utility is maximized, and increased training achieves this. In our paper, too, training increases with increased competition, though for different reasons.

In Stevens (1994), a random shock process determines the productivity of firms after training has been carried out and each firm pays a given worker only what its closest rival would pay him, so that the wage is less than the marginal product. This implies that there is imperfect competition in the labour market. More firms competing for a worker in her model results in lower expected profits for the training firm by virtue of the probability of poaching being greater. Specific (though not entirely specific) training becomes less valuable, as the worker is more likely to leave. Thus more competition in Stevens (1994) is associated with lower profits, as in Autor(2000) but in contrast with our paper, and less training, unlike in both Autor(2000) and here.

Theoretical papers on training and turnover have tended to focus on how much training an individual worker receives, which we call vertical training, as opposed to how many workers firms train, which we call horizontal training[4]. The framework suggested in this paper enables us to discuss simultaneously vertical and horizontal training, and in particular to see what happens to these quantities as the degree of generality of training varies. The decision as to whether to subsidize one form of training over the other is an important one for policy makers in LDC's[5]. For a discussion of the literature on training by multinationals in LDC's, see chapter1.

Finding empirical evidence to test the model has proven a problem. Whilst the model's predictions are largely testable[6], finding data is no easy task. We need to find data collected at the time when a multinational - or a group of multinationals, since the results do not depend on there being one firm in period 1- begins operating in the rural area of an LDC. However, this is precisely the time when data is scarcest, as the authorities' priority is to ensure that the project kicks off successfully rather than to organize precise statistical monitoring of the project and its effects on society. Our efforts to find Mexican data - Mexico is one of the best sources of LDC data, both in terms of quantity and quality - have proved futile. However, we may be able to find data elsewhere, for example in developed nations in the context of industry which is set up in rural areas.

The paper proceeds as follows. Section 2 presents the main model and the results for the market economy. In this section, we obtain the main result of the paper, namely that profits of the multinational are higher when training is more general or when the number of rival firms increases.

In section 3, we relax some of the assumptions and find that in some cases the key result is weakened, in others it is strengthened. Firstly, we show that having many firms in period 1 as opposed to one firm only does not affect the main result that more

[4] Furthermore, as Arvin (1993) points out, most of the literature that followed Becker's pioneering work in the '60's and '70's concentrated on specific training, in particular on the optimal way for workers and firms to share the costs and benefits of specific training. Current research, on the other hand, has put general training centre stage, in part as a result of the key role that human capital plays in modern growth theory. The emphasis of this paper on the role of human capital in the development process places it in the 'modern' camp. However, its simultaneous treatment of general and specific training puts the paper somewhere between these poles.

[5] This is confirmed in my discussions with Mexican government officials, trade unionists and industrialists. For example, training policy in the states of Tlaxcala and Veracruz includes subsidies to both horizontal and vertical training. This paper suggests that vertical training should be either taxed or subsidized depending on a particular condition that relates to training costs. It would be interesting to see which Mexican states, if any, tax vertical training.

[6] A problem is that the degree of generality of training is not directly observable, though it can be proxied.

competition increases profits of period 1 incumbents. Secondly, we show that, if workers face a cost of entering industry (e.g. moving to a different location), then the main result will continue to hold for 1 incumbent in period 1, but will collapse with many firms in period 1 if workers are sufficiently heterogeneous in their fixed cost of entry. Thirdly, we drop the assumption that each worker works a fixed number of hours, and allows individuals' supply of work to be elastic. We show that the main result is unaffected in the case of workers being homogeneous in their fixed cost of entry, but may collapse if they are sufficiently heterogeneous. Finally, allowing the multinational to hire workers in period 2 leaves the main result intact.

In section 4, we introduce training per worker. We distinguish between the amount of training per worker, called *vertical training*, and the number of workers trained, called *horizontal training*.

In section 5, we find that, compared to the social optimum, horizontal training is always underprovided, whilst vertical training may be either underprovided or overprovided, depending on cost of training parameters.

In section 6, we show that, when firms face a fixed cost of entry, multiple equilibria due to self-fulfilling expectations of local firms are possible, whereby if period 2 entrant firms believe in period 1 that there will not be enough workers available in period 2, then they will not enter, the multinational firm in period 1 will train less workers, and the expectations of the entrants will be confirmed.

The conclusion follows in section 7.

2 Main model

A multinational M requires skilled workers to produce. By skilled, we mean having the skills needed to generate industrial output. There are no skilled workers in the LDC at the start of period 1, since all workers are employed in agriculture. Thus M attracts unskilled labour from the agricultural sector in the first period and trains it, i.e. it turns unskilled into skilled labour. In the second period, a number n of local (LDC) firms L's are born which also utilize skilled labour as their only input. L's obtain skilled labour by poaching it from M, as M is the only source of skilled labour and L's are assumed not to be able to train workers[7]. Both types of firms face diminishing returns to labour, accounting for the 'spread' of skilled labour to L's in the second period and generating positive profits in the presence of barriers to entry[8]. We initially assume a perfectly elastic supply of unskilled labour for M's to draw, reflecting the presence of a large agricultural sector. The absence of product market competition between M's and L's reflects the fact that M's often produce in LDC's to supply the markets of more developed nations, whilst LDC firms, particularly when in their infancy, tend to supply only the local market. Since we disregard product market competition, we assume the product price to be constant and, for simplicity, equal to 1 for both M and L's.

Although we can shown that profits of the multinational increase with the productivity of poached workers in local firms for any production function, it is nevertheless useful to construct a specific model in order (a) to show explicitly the interplay of the various variables in determining profits of M, (b) to study how profits of local firms vary with specific parameters, such as productivity of workers in the multinational in period 2, and (c) to study how the result that more competition can increase an incumbent firm's profits is affected when we relax some assumptions.

M is active in both periods, whilst L's are active only in the second period. Their respective production function, which display decreasing returns to scale (or diminishing returns to labour, given that labour is the only input), are

$$Q_M = (\rho_t \; S_M)^{\alpha} \quad 0 < \alpha < 1 \quad \rho_1 = 1, \rho_2 = \rho > 1$$

and

$$Q_L = (\beta \; S_L)^{\alpha} \tag{1}$$

where S_M is skilled labour hired from the agricultural sector and trained by M, which we call horizontal training (in section 4 we introduce training per worker, which we call

[7] One possible rational for this assumption is that training programmes can only be set up once firms have established themselves, that is, after a number of periods of production.
[8] Barriers to entry are considered by many commentators to be higher in LDC's than in more developed nations, suggesting that non-zero profits is an attractive feature of a model of development.

vertical training, θ). We assume $\rho_1 = 1$ and $\rho_2 = \rho > 1$, reflecting the fact that training is an on-going process lasting one period (had it been instantaneous and therefore complete at the start of period 1, then we would have had $\rho_1 = \rho_2 = 1$). There are n identical local firms.

In practice, it is likely that $\beta \leq \rho$: given the same number of workers in the multinational and in each local firm, workers who stay in the multinational in period 2 are at least as productive as workers who move to local firms. However, it need not necessarily be the case: L's are at times capable of combining their old technology with that of M to achieve a more efficient technology than that of M.

The supply of unskilled labour coming from agriculture is horizontal and determined by the opportunity cost of not working in agriculture, which exhibits constant returns to labour and where the marginal product per period is 1[9]. This is in contrast with supply in the second period, which is vertical as a result of being determined in period 1, as we shall see shortly. The model thus displays the curious feature that supply in period 2 is determined by (the incumbent's) demand in period 1. Later on, we allow the incumbent to hire from agriculture also in period 2; in doing so, the result that an incumbent gains from the presence of more local firms continues to hold, but we lose the aforementioned feature, showing that it is not a necessary feature of the model for the key result to hold.

Since the sum of wages over two periods in industry must be at least as high as that in agriculture, in equilibrium we have

$$w_1 + w_2 = 2 \qquad (2)$$

(2) assumes for simplicity that each worker supplies his labour inelastically, i.e. that substitution and the income effect exactly match each other. Equation (2) embodies the see-saw effect, the inverse relationship between wages in different periods, on which the key result of the paper hinges, namely that more competition may result in higher profits for firms present in the market already from period 1.

[9] Later on we allow workers to be heterogeneous in the cost of moving from agriculture to industry. Another interpretation of such a set-up is that there are decreasing returns in agriculture, with the first workers to leave agriculture needing less to be compensated than subsequent workers.

The fact than demand for trained workers in period 2 exceeds that in period 1 ensures that $w_2 > w_1$. This implies that $w_2 > 1$, so that industrial workers in the second period stay in industry. We assume, for simplicity, that M, which is assumed to be the only firm capable of training, does not train workers from the agricultural sector in period 2 to compensate for those it loses to L's. This assumption is relaxed in section 3.

In period 2, both L's and M take w_2 as given[10]. We defining local firms' demand for labour as S_L and the multinational's demand for labour as $S_{M,2}$. The problem of L's is

$$\underset{S_L}{Max}\ \pi_{L,2} = (\beta\ S_L)^{\alpha} - w_2 S_L$$

which yields the inverse demand function

$$w_2 = \alpha\ \beta^{\alpha}(S_L)^{\alpha-1} \tag{3}$$

M's problem is

$$\underset{S_{M,2}}{Max}\ \pi_{M,2} = (\rho\ S_{M,2})^{\alpha} - w_2 S_{M,2}$$

which yields the inverse demand function

$$w_2 = \alpha\ \rho^{\alpha}(S_{M,2})^{\alpha-1} \tag{4}$$

[10] The fact that M is a wage-taker is justified if ρ is not too high relative to β, the productivity parameters of M and L's in period 2 respectively, and if the number n of L's is high, which is reasonable given that we are considering wage-taking L's.

In equilibrium, demand for workers equals supply, S_M, which is given (i.e. vertical) since it is determined by M in period 1:

$$nS_L + S_{M,2} = S_M \tag{5}$$

Combining (3), (4) and (5), we obtain that S_L is proportional to S_M:

$$S_L = B\, S_M \qquad where \quad B = \frac{1}{n + \left[\left(\frac{\beta}{\rho}\right)^{\alpha}\right]^{\frac{1}{\alpha-1}}} \tag{6}$$

This is an important sub-result which simplifies the subsequent analysis. Note that if $\beta=\rho$, then each firm simply obtains an equal share $1/(n+1)$ of supply.

In period 1, M's problem is to maximize profits over both periods by choice of S_M:

$$\underset{S_M}{Max}\, \pi_M = (S_M)^{\alpha} - w_1 S_M + (\rho\, S_{M,2})^{\alpha} - w_2\, S_{M,2} \tag{7}$$

subject to (2), (3), (4), (5) and (6).

M's profits can be written as

$$\underset{S_M}{Max}\, \pi_M = D(S_M)^{\alpha} - 2S_{M,2} \tag{8}$$

where $D = 1 + \rho^{\alpha}(1 - nB)^{\alpha} + \alpha n \beta^{\alpha} B^{\alpha}$

The optimal S_M is thus

$$S_M = \left[\frac{2}{\alpha D}\right]^{\frac{1}{\alpha-1}} \tag{9}$$

Le us proceed to the main result of this chapter.

Proposition 1: If the productivity parameter β of trained workers in local firms is higher, or if the number n of local firms is higher, profits of the multinational increase.

Proof: Either by showing that $dD/d\beta>0$ and that $dD/dn>0$, or more generally as follows. Let the production function of M be $f(S_i)$, where $S_i=S_M$ in period 1 and $S_i=S_{M,2}$ in period 2.

$$\underset{S_M}{Max}\, \pi_M = f(S_M) - w_1 S_M + f(\rho\ S_{M,2}) - w_2\ S_{M,2}$$

We wish to show that an increase in w_2, be it as a result of either β or n increasing, causes M's profits to increase. Using the envelope theorem,

$$\frac{d\pi_M}{dw_2} = \frac{\partial \pi_M}{\partial w_2}$$

Using Hotelling's lemma, which recognizes that the second period wage is equal to the marginal product, we have

$$\frac{\partial \pi_M}{\partial w_2} = -\frac{dw_1}{dw_2} S_M - S_{M,2}$$

Using condition (2), which says that $w_1+w_2=2$, we have that

$$\frac{dw_1}{dw_2} = -1 \qquad (10)$$

Combining the last two equations, we have

$$\frac{d\pi_M}{dw_2} = S_M - S_{M,2} > 0$$

where the inequality sign is due to the fact that the multinational faces competition in the second period, but not in the first.

The intuition behind the result for β is as follows. An increase in the degree of generality of training is equivalent in this model to an exogenous increase in productivity of workers in L's. Thus L's' demand for labour rises and therefore so does w_2. Given that workers care about the sum of wages over two periods, market clearing in period 1 implies that w_1 falls by the same absolute amount as w_2 goes up by. Since the amount of labour employed by M in the first period is greater than that in the second, the overall wage bill falls. Furthermore, the overall wage bill falls by more than second period revenues, so that M's overall profits increase. The fact that profits rise with β suggests that M would find it profitable to increase β if it was within its power to do so, provided that such an action is not too costly.

The intuition for an increase in the number n of rivals is very similar to that for an increase in β[11]. A greater number of L's increases competition for labour in period 2, raising w_2. A higher w_2 is beneficial to M via the see-saw effect. Note that, given our assumption of atomistic, wage-taking firms in period 2, the number n of local firms must always be sufficiently large.

[11] According to Stevens (1994), an increase in n is an increase in the generality of training, since it implies that training is of use to a greater number of firms. Using this interpretation of an increase in n, it is not surprising that both a higher n and a higher β cause the multinational's profits to be higher.

In essence, the see-saw effect operates in this paper when a firm faces delayed competition for inputs: when the latter increases, and workers take into consideration the sum of wages over the two periods rather than each period's wage individually, the firm enjoys positive repercussions in the first period which outweigh the negative ones in the second.

The multinational is ultimately a school of training, which charges workers, its 'pupils', the fall in w_1 when they become more employable. The pupils then charge local firms (as well as the multinational in period 2) the increase in w_2. Thus, the multinational effectively sells training skills to local firms, its ultimate customers. Indeed, when β rises to equal infinity, the multinational loses all its workers and thus sells nothing in period 2, yet it gains from the increase in beta, highlighting the intuition that the multinational gains by being a training firm.

Clearly, an increase in ρ, the multinational labour productivity parameter in period 2, increases profits of M, again by virtue of the fact that the second period wage increases.

The definition of 'general training' that we have been using, namely that training is more general if it increases the productivity β of trained workers in firms which poach workers, is quiet broad. It is similar to that used by authors such as Jones (1988), Ritzen(1991) and Hyman(1992), who define general training as training that is of use to some other firms, and Schackelton(1992), who defines it as training that is of use to at least one other firm.

One could, however, consider a more narrow definition. For example, training could be regarded as more general if it leads to an increase in productivity in poaching firms relative to that in the training firm, i.e. an increase in β/ρ. If we were to adopt this definition, we could no longer make the general claim that more general training increases profits if the training firm, because a higher ρ, which raises profits of the training firm, implies a reduction in β/ρ.

This discussion suggests that the definition of the term general as applied to training is important and affects the interpretation of the results of the models below. This point is also made by Stevens(1994), who warns that comparing models of different authors with different definitions of 'general training' is dangerous. Our choice of definition reflects, we believe, a commonly held notion of what general training means.

How does the presence of rivals help M? When rivals enter, the constraint of decreasing returns to scale (equivalently, diminishing returns to labour, as there is only one input) at the industry level is relaxed, enabling period 2 consumer surplus to rise and inducing further entry of workers in period 1 via the see-saw effect. In other words, more competition tomorrow results in a relaxation of what is effectively a form of capacity constraint. For this reason, M cannot simply achieve on its own the profits which it achieves in the presence of L's.

The importance of diminishing returns for Proposition 1 to hold is fundamental. If there were constant returns to scale, so that α=1, M's profits (8) would be

$$\max_{S_M} \pi_M = S_M(\rho - 1) \tag{11}$$

so that the presence of rivals would not affect profits of M. This is because, as the diminishing returns parameter α approaches 1, where $\alpha = 1$ represents constant returns, the equality between the wage and the marginal product is preserved, so that $w_2 = \rho$. This fixes the period 1 wage at 2-ρ, so that the see-saw effect is absent. Furthermore, profits in period 2 are zero, so that the market share of M in period 2 (undetermined with constant returns to scale) is irrelevant in determining profits of M.

We now show that the commitment problem of chapter 2 is not at work here. Suppose the multinational can commit. Then it would maximize

$$\max_{S_M, S_{M,2}} \pi_M = (S_M)^\alpha - w_1 S_M + (\rho\, S_{M,2})^\alpha - w_2\, S_{M,2}$$

s.t. $w_1+w_2=2$, $S_M=S_{M,2}$

where the second constraint comes from the fact that the multinational operates alone in both periods, so the workers available to it in the second period are those it hired and trained in the first. The maximization problem with commitment then reduces to

$$\underset{S_M}{Max}\ \pi_M = (1+\rho^{\alpha})(S_M)^{\alpha} - 2\ S_M$$

Now suppose that the multinational cannot commit. In period 2 supply is horizontal at $w_2=1$, which implies via the see-saw effect that $w_1=1$. More importantly, because the multinational operates alone in both periods, we have $S_M=S_{M,2}$, as with commitment. Then the maximization problem without the ability to commit turns out to be the same as the one with the ability to commit. We deduce that there is no commitment problem.

It is noteworthy that increased poaching is not the cause of increased profits of M, but rather both are a consequence of increased competition. At no point does the multinational benefit from poaching per se, but rather by the presence of a compensation mechanism, this being the see-saw effect. As the compensation received by the multinational for its poached workers exceeds the cost, the main result follows that more competition results in higher profits of the multinational.

The theory assumes that workers can sustain a wage cut in the first period. At first sight, it would appear that the theory is vulnerable to the criticism that, if the wage cut is sufficiently large, it may not be possible for the worker to survive. If a worker lacks savings, or cannot resort to (sufficiently cheap) borrowing, or if he does not have a family that will support him[12], then he will not be able to sustain the wage cut. This implies that, when competition increases in period 2, or when training becomes more general, so that the period 2 wage increases, the multinational in period 1 will not be able to pay a lower wage. This would then destroy the see-saw mechanism[13], so that an increase in competition in period 2 will result in the multinational incurring losses over both periods. However, this is only true in our simple model in which workers do not require a wage premium over what they earn in agriculture, i.e. they do not face any cost of moving from agriculture to industry. When they do face such a cost, the wage in period 1 need not fall below the level which is required by workers to survive, so that the above analysis goes through. What matters is not the level at which the sum of wages over the two periods is fixed, but simply that it is fixed. By the same argument, the first period wage need not fall below the wage earned in agriculture, thus fitting the facts.

[12] In addition, if workers and firms cannot sign contracts whereby workers make delayed payments to firms to cover the costs of training, possibly via deductions from workers' wages over time, then firms may not be able to forego wage reductions while imparting training. This is the problem of firms not being able to obtain property rights over the human capital sold to workers (Katz and Ziderman (1999)).

[13] Another reason why the see-saw mechanism may be destroyed is that there may be minimum wage legislation in place, making it illegal for the period 1 wage to fall below a certain level.

The theory assumes, furthermore, that workers are forward looking, so that they tolerate a lower wage today to enjoy a higher wage tomorrow. Given that we are considering a developing nation in which a forward looking mentality may be less pronounced than in a so-called developed nation, the assumption might be a demanding one.

Next, we consider how profits of L's vary with the number n of local firms and the productivity parameter ρ of M in period 2.

Profits of L's can be written as

$$\pi_L = (1-\alpha)(\beta B S_M)^{\alpha} \tag{12}$$

Note, first of all, that an increase in β, the productivity parameter of local firms, results in an increase in B, which measures the fraction of the total labour supply going to each local firm, and an increase in S_M. Therefore profits of L's rise with β, as one would expect.

Substituting B of (6) and S_M of (9) into (12), and differentiating with respect to n, we can show that profits of L always fall with n. This result is not obvious because an increase in n has two opposing effects: on the one hand B falls, reflecting the fact that each firm employs a lower fraction of the work force in period 2; on the other hand, it induces M to increase S_M, because the higher second period wage reduces the overall wage bill to fall via the see-saw effect.

Profits of L's, however, may either rise or fall with ρ, the productivity parameter of M in period 2. This is because, although B falls, S_M rises. In particular, substituting B of (6) and S_M of (9) into (12), and differentiating with respect to ρ, we have that

$$\frac{d\pi_L}{d\rho} = -\beta + \frac{n}{\left(n + \left(\frac{\rho}{\beta}\right)^{\frac{\alpha}{1-\alpha}}\right)^{\alpha}} \tag{13}$$

which can be either positive or negative. The $-\beta$ term reflects the reduction in B, the share of total labour supply, hired by each L in period 2 following an increase in ρ, whilst the positive term on the right reflects the increase in S_M induced by the increase in ρ. Intuitively, whilst local firms in period 2 tend to lose from an increase in the multinational's productivity parameter ρ and the resulting higher wage, they tend to gain from the fact that the multinational has a greater incentive to hire and train workers in period 2, resulting in a greater supply of workers in period 2 and thus a lower wage. Either of these two forces can dominate, generating the ambiguity in the movement of profits of local firms when ρ increases.

3 Relaxing the assumptions

3.1 Allowing for many firms in period 1

Suppose that there are many firms in period 1. For simplicity, let all firms be identical. The question is whether or not Proposition 1, which says that profits of a period 1 incumbent increase with the number of entrants in period 2, will be unaffected. The answer is yes, and the proof is as follows.

Let S be the total quantity of labour employed by firms, which is finite as a result of diminishing returns. Let there be m firms in period 1, and m + n firms in period 2. As in the proof for proposition1, $f(\cdot)$ represents any production function displaying decreasing returns. Then profits of period 1 incumbent firms are

$$\pi_m = f\left(\frac{S}{m}\right) - w_1 \frac{S}{m} + f\left(\frac{S}{m+n}\right) - w_2 \frac{S}{m+n}$$

Using Hotelling's lemma, which recognizes that the wage is equal to the marginal product, we have

$$\frac{d\pi_m}{dw_2} = -\frac{dw_1}{dw_2}\frac{S}{m} - \frac{S}{m+n}$$

Using condition (2), namely $w_1 + w_2 = 2$, which says that workers break even over two periods, we have

$$\frac{d\pi_m}{dw_2} = \frac{S}{m} - \frac{S}{m+n} > 0 \tag{14}$$

This result for each of m firms in period 1 is analogous to that for the monopolist in period 1 (see the expression below equation (10)), as can be seen by setting m=1 in (14). As in the case of a monopsony, the result is driven by the fact that the quantity of labour hired in period 1 exceeds that hired in period 2, and by the fact that wages move in opposite directions in such a way that the change in the period 1 wage is equal in size to the change in the second period wage.

This model is the 2-period extension of the classic, 1-period model with diminishing returns and price taking firms. It may be rather surprising to find that profits of period 1 incumbents increase with the number of the period 2 entrants.

3.2 Allowing for heterogeneous workers

We assume that workers face a fixed cost of entering industry, such as having to move home. We allow workers to be heterogeneous by assuming that they face different costs of moving from agriculture to industry. Another interpretation of this set-up is that there are decreasing returns to labour in agriculture, with the first workers to leave agriculture needing less to be compensated than subsequent workers (i.e. they have lower fixed cost of moving in the original interpretation). Let workers be ranked in order of increasing cost of entry. In this case, only the marginal worker breaks even, and the see-saw equation (2) is altered as follows:

$$w_1 + w_2 - F(S) = 2 \tag{15}$$

where S is the total number of workers, and F(S) is an increasing function of S.

3.2.1 M in period 1

Proposition 1 is unaffected because, even though the see-saw effect equation now contains a term F(S) that is increasing in the number of workers S, we have that

$$\frac{\partial w_1}{\partial w_2} = -1$$

which replaces the previous total derivative (10), namely $dw_1/dw_2=-1$, where the partial sign refers to the fact that S is kept constant.

Intuitively, M is in control of the number of workers who enter industry. Therefore, when the second period wage increases as a result of increased competition, the monopolist could choose to prevent new workers from entering, and reduce the first period wage by as much as the second period wage rises. This preserves the result of Proposition 1.

Clearly, with price-taking firms, no firm is in control of how many workers enter in period 1. This implies that, when the second period wage rises, so that the number of workers entering industry increases, the first period wage may, according to the see-saw effect equation (15), fall too little to cause profits of period 1 incumbents to increase. We explore this possibility below.

3.2.2 m wage-taking firms in period 2

As in the case without worker heterogeneity, we have

$$\frac{d\pi_m}{dw_2} = -\frac{dw_1}{dw_2}\frac{S}{m} - \frac{S}{m+n}$$

Using the condition $w_1 + w_2 - F(S) = 2$, which says that workers break even over two periods, we have

$$\frac{dw_1}{dw_2} = \frac{\partial w_1}{\partial w_2} + \frac{dF(S)}{dS}\frac{dS}{dw_2} = -1 + \frac{dF(S)}{dS}\frac{dS}{dw_2} \tag{16}$$

where the second term on the left-hand side is positive because both its constituent terms are positive.

$$\frac{d\pi_m}{dw_2} = \frac{S}{m}\left[1 - \frac{dF(S)}{dS}\frac{dS}{dw_2}\right] - \frac{S}{m+n} \tag{17}$$

Thus, profits of m's may fall with the second period wage provided that dF(S)/dS is sufficiently large. In the Appendix, we show that this conclusion is confirmed when we extend the basic model of section 2. We also show that profits of m's always fall with m. Similarly, we show that profits of n's increase with m's provided that worker heterogeneity is sufficiently low, and that profits of L's always fall with n.

In other words, when we extend the model of section 2, we find that there is a certain symmetry. With wage-taking firms in both periods, a firm which enters in one period always loses if another firm enters in that same period, and gains if another firm enter in the other period provided that worker heterogeneity is sufficiently low.

The proof is given in Appendix 7.1. The intuition is as follows.

If an extra firm enters in period 2, it causes the second period wage to rise, reducing profits of firms which also enter in period 2. However, the period 1 wage falls, because more workers enter. With the fixed cost incurred by workers rising sufficiently slowly with the number of workers, i.e. with sufficiently low worker heterogeneity, the first period wage will fall enough for profits of firms which enter in period 1 to increase.

If an extra firm enters in period 1, it causes the first period wage to rise, reducing profits of firms which also enter in period 1. However, the period 2 wage may fall. It does so provided that enough workers enter in period 1. This in turns takes place provided that the fixed cost incurred by workers is increasing sufficiently slowly with the number of workers. i.e. that worker heterogeneity is sufficiently low.

We summarize the results for the case of m wage-taking firms in period 1:

Figure 4.1: Can profits of firms which enter in period 1 (m's) and profits of firms which enter in period 2 (n's) increase with the number of entrants? Results for the case of inelastic supply of each worker

	Profits of m's (period 1 entrants)	Profits of n's (period 2 entrants)
Higher m	No	Yes if worker heterogeneity low
Higher n	Yes if worker heterogeneity low	No

3.3 Allowing each worker's supply to be elastic

Does Proposition1 continue to hold true when each individual worker's supply is elastic?

We show that, for any production function exhibiting diminishing returns, allowing each workers to supply his labour elastically does not affect Proposition 1 as long as worker are homogeneous in their fixed cost of entry. This is true both for a monopsony in period 1 and price-takers in period 1. If, however, workers are heterogeneous, we cannot come to any conclusion about how Proposition1 is affected. To do so, we would need to extend the model of section 2 to allow both the supply of labour of individual workers to be elastic and for workers to be heterogeneous in their fixed cost . Since that model of this does not extend easily, we are not in apposition to study how Proposition 1 is affected by this particular combination of altered assumptions.

We assume that a higher wage induces every worker to supply more work, so that the substitution effect dominates the income effect, and the supply curve slopes upward at all times.

In each period t, workers choose how many hours of work to supply taking the wage as given:

$$\underset{s_{A,t}}{Max}\; w_t s_{A,t} - c(s_{A,t})$$

where $s_{A,t}$ is the number of hours worked by each worker in period t, and $c(s_{A,t})$ is an increasing, convex function. The inverse demand function is thus

$$w_t = \frac{dc(s_{A,t})}{ds_{A,t}} \tag{18}$$

The condition for the marginal worker to break even is

$$w_1 s_{A,1} - c(s_{A,1}) + w_2 s_{A,2} - c(s_{A,2}) - F(S) = 2 \tag{19}$$

Equation (19) is the elastic labour supply version of equation (2), and embodies the see-saw effect, by which w_1 and w_2 move in opposite directions when demand for labour in period 2 increases exogenously via β or n.

Using the first order conditions (18), the marginal worker condition (19) can be seen to link the variables w_1, w_2 and S_M. In particular, suppose that w_2 rises as a result of more competition for workers in period 2. If only M is present in period 1, M has a choice of either reducing w_1, or increasing S_M, or both. If m wage-taking firms are present, then the increase in w_2 will certainly draw more workers S into the industry.

Consider now how the marginal worker condition is affected by small changes in the second period wage (which increases when n or β increase):

$$\frac{dw_1}{dw_2}s_{A,1}+\frac{ds_{A,1}}{dw_1}\frac{dw_1}{dw_2}w_1-\frac{dc(s_{A,1})}{dw_1}\frac{dw_1}{dw_2}+s_{A,2}+\frac{ds_{A,2}}{dw_2}w_2-\frac{dc(s_{A,2})}{dw_1}-\frac{dF(S)}{dS}\frac{dS}{dw_2}=0 \tag{20}$$

Using the first-order conditions (18), the above expression simplifies to

$$\frac{dw_1}{dw_2}s_{A,1}+s_{A,2}-\frac{dF(S)}{dS}\frac{dS}{dw_2}=0 \tag{21}$$

For simplicity, we assume that firms are identical. Firms maximize profits by choice of total number of hours worked. More specifically, in period 2 each firm chooses the product $S_{i,2}s_{A,2}$ which maximizes profits taking w_2 as given, where $S_{i,2}$ is the number of workers each firm i hires in period 2. In other words, $S_{i,2}$ and $s_{A,2}$ are perfect substitutes in the production function of firms.

The problem faced by firms in period 2 is thus:

$$\underset{S_{i,2}s_{A,2}}{Max}\ \pi_{i,2}=f(S_{i,2}s_{A,2})-w_2S_{i,2}s_{A,2}$$

which generates the inverse demand function

$$w_2 = \frac{df(S_{i,2}s_{A,2})}{d(S_{i,2}s_{A,2})} \tag{22}$$

There are 2 cases for the period 1 market structure, that of the monopsonist M and that of m wage-taking firms. As the proof is rather laborious, we relegate it to Appendix 7.2. Here, we restrict ourselves to giving the intuition behind the main result that, with homogenous workers, profits of period 1 incumbents, regardless of whether there is a monopsony or many wage-taking firms, increase with the number of period 2 entrants.

The key behind this result is the same as that with individually inelastic supply of labour, namely equation (21). This equation states shows how the marginal worker condition varies with the period 2 wage.

A monopsony, unlike wage-taking firms, is in control of the number of workers who enter industry. Therefore, when the second period wage increases as a result of increased competition, the monopolist could choose to prevent new workers from entering, and reduce the first period wage sufficiently to preserves the result of Proposition 1.

Clearly, with wage-taking firms, no firm is in control of how many workers enter in period 1. This implies that, when the second period wage rises, so that the number of workers entering industry increases, the first period wage may, according to the see-saw effect equation (19), fall too little to cause profits of period 1 incumbents to increase. Therefore, we cannot tell whether profits of incumbents rise or fall with the second period wage, except when workers are homogeneous, in which case dF(S)/dS = 0 in equation (21), so that profits increase.

3.4 Allowing hiring of workers from agriculture in period 2

Suppose now that firms which have entered in period 1 (of which there are m) have the ability to hire again from the agricultural sector in period 2 (and to train these new workers). Firms which enter in period 2 continue to rely on poaching to obtain skilled workers because they do not have training facilities.

Let workers who entered in period 1 be labelled 'old', and those who enter in period 2 be called 'new. Define $w_{2,N}$ and $s_{A,2,N}$ as the wage received by new workers and the hours worked by new workers, respectively. The period 2 problem of firms which have entered in period 1 is to then choose the total quantity $S_m s_{A,2}$ of old labour hours and the total quantity $S_N s_{A,2,N}$ of new labour hours.

$$\underset{S_m s_{A,2},\, S_N s_{A,2,N}}{Max} \quad \pi_{m,2} = \left(\rho S_m s_{A,2} + S_N s_{A,2,N}\right)^{\alpha} - w_2 S_m s_{A,2} - w_N S_N s_{A,2,N}$$

where we have assumed that new workers in period 2 are as productive as old workers in period 1 (i.e. their ρ takes a value of 1).

Since $S_m s_{A,2}$ and $S_N s_{A,2,N}$ are perfect substitutes in production (once we adjust for the parameter ρ), firms will not demand both new and old workers except in the knife-edge case where $w_{2,N}=w_2/\rho$, in which they would be indifferent between the two. Since we wish to consider a situation in which firms which enter in period 1 hire new workers in period 2, it must be the case that in period 2 they hire only new workers[14]. This is true provided that

$$w_{2,N} \leq \frac{w_2}{\rho}$$

Since firms which enter in period 1 demand only new workers in period 2, only firms which enter in period 2 are left to hire old workers. Therefore, when competition for old

[14] Introducing training costs can change this rather drastic outcome so that some old workers are retained by firms which enter in period 1.

workers increases in period 2, raising w_2, w_1 falls via the see-saw effect as before, but now firms which enter in period 1 do not suffer from poaching since they do not employ any old workers in period 2.

The original trade-off found in the case where training firms cannot hire again in period 2, which involved training firms losing the services of some workers in period 2 and thus producing less, is now absent: only the gain of a lower w_1 via the see-saw effect remains. Therefore, allowing firms which enter in period 1 to hire again from agriculture in period 2 strengthens the see-saw effect, so that firms always gain when there is more competition for trained workers.

Interestingly, we can show that an equilibrium in which hiring from agriculture takes place in period 2 is only possible if workers are homogeneous in their cost of entry. In other words, if workers are heterogeneous, no further hiring from agriculture in period 2 will take place. We prove this in the Appendix 7.3. Intuitively, if fixed costs of workers are increasing in the number of workers, the second period wage of newly hired workers would have to be correspondingly higher, and would result in workers, who would otherwise enter in period 1, entering in period 2 instead. Clearly this would not be an equilibrium, since no workers would enter in period 1.

We summarize our conclusions when we relax some important assumptions.

Proposition 2: In the case of monopsony in period 1, assuming that each worker's supply is inelastic, we find that introducing worker heterogeneity does not affect Proposition 1, namely that increased competition for workers in period 2 results in increased profits of the monopsony. Assuming each worker's supply to be elastic, we find that Proposition 1 is unaffected in the case where workers are homogeneous.

In the case of wage-taking firms in period 1, assuming that each worker's supply is inelastic, we find that Proposition 1, namely that increased competition for workers in period 2 results in increased profits of period 1 incumbents, holds provided that worker heterogeneity is sufficiently low. Assuming each worker's supply to be elastic, we find that Proposition 1 is unaffected in the case where workers are homogeneous.

Regardless of whether in period 1 there is a monopsony or a number of wage-taking firms, allowing new workers to be hired from agriculture in period 2 leaves Proposition 1 intact.

Proof: by referring to previous proofs involving different assumptions.

Therefore, our results suggest that Proposition 1 is more robust in the case of a monopsony in period 1 than in the case of wage-taking firms in period 2. Fundamentally, this is due to the monopolist's ability to control the number of workers who enter industry in period 1.

Below is a table summarizing the results of the extensions.

Figure 4.2: Do profits of period 1 incumbents increase with the number of rivals in period 2? Results under different assumptions

		Homogeneous workers		Heterogeneous Workers	
		Each worker's supply inelastic	Each worker's supply elastic	Each worker's supply inelastic	Each worker's Supply elastic
Mono-psony in period 1	No hiring in period 2	Yes	Yes	Yes	Result not available
	Hiring in period 2	Yes	Yes	Scenario not possible	Scenario not possible
m firms in period 1	No hiring in period 2	Yes	Yes	Yes if worker heterogeneity low, no if high	Result not available
	Hiring in period 2	Yes	Yes	Scenario not possible	Scenario not possible

Where 'hiring in period 2' refers to firms which enter in period 1 being able to hire labour from agriculture in period 2.

4 Training per worker and the number of workers trained

In this section, we wish to introduce training costs explicitly, and to consider a training firm's trade-off between the number of workers is trains, and the amount of training per worker.

For simplicity, we revert to the simpler, main model o section 2 in which there is a monopsony in period 1, workers do not face fixed costs of entry into industry, and workers supply their work inelastically[15]. We introduce a new control variable for M, namely vertical training, θ. Vertical training is the amount of training which M imparts to each worker in period 1, as opposed to horizontal training, which is simply the number of workers M hires (and trains) in period 1. The total cost of training S_M workers an amount θ each is $c\theta^{\kappa}S_M^{\iota}$.

Training per worker θ is determined by M in period 1. Therefore, in period 2, L's and M take both w_2 and θ as given in maximizing profits. The problem of L's is

$$\underset{S_L}{Max}\, \pi_{L,2} = (\beta\theta\, S_L)^{\alpha} - w_2 S_L$$

which yields the inverse demand function

$$w_2 = \alpha\, \beta^{\alpha}\theta^{\alpha}(S_L)^{\alpha-1} \tag{23}$$

M's problem is

$$\underset{S_{M,2}}{Max}\, \pi_{M,2} = (\rho\, \theta\, S_{M,2})^{\alpha} - w_2 S_{M,2}$$

[15] Assuming that each worker supplies his work elastically leaves the results of this section unaffected, but makes the algebra more complicated.

which yields the inverse demand function

$$w_2 = \alpha\ \rho^{\alpha}\theta^{\alpha}(S_{M,2})^{\alpha-1} \tag{24}$$

As before, in equilibrium demand for workers equals supply, S_M, which is given since it is determined by M in period 1:

$$nS_L + S_{M,2} = S_M \tag{25}$$

Combining (23), (24) and (25), we obtain the familiar relationship found in section 2 whereby S_L is proportional to S_M:

$$S_L = B\ S_M \qquad where \quad B = \frac{1}{n + \left[\left(\frac{\beta}{\rho}\right)^{\alpha}\right]^{\frac{1}{\alpha-1}}} \tag{26}$$

In period 1, M's problem is to maximize profits over both periods by choice of S_M and θ:

$$\underset{\theta, S_M}{Max}\ \pi_M = (\theta\ S_M)^{\alpha} - w_1 S_M + (\rho\theta\, S_{M,2})^{\alpha} - w_2 S_{M,2} - c\theta^k S_M{}^j$$

subject to (2), (24), (25) and (26) (27)

where c, j, k are all positive parameters. c is the unit training cost. k may be interpreted as the degree of returns to vertical training, so that k<1 implies increasing returns, k=1 implies constant returns, and k>1 implies decreasing returns. So, for example, k>1 could reflect diminishing returns to learning. A similar reasoning applies for the parameter j of horizontal training.

Partial differentiation and substitution yields the following equilibrium values for horizontal and vertical training:

$$S_M = \left\{ \frac{ck}{\alpha D} \left[\frac{2}{c(k-j)} \right]^{\frac{k-\alpha}{k}} \right\}^{\frac{k}{(\alpha-k)(1-j)+k(\alpha-j)}} \tag{28}$$

$$\theta = \left\{ \frac{2}{c(k-j)} S_M^{\;1-j} \right\}^{\frac{1}{k}} \tag{29}$$

where $D = 1 + \rho^{\alpha}(1 - nB)^{\alpha} + \alpha n \beta^{\alpha} B^{\alpha}$

Note that D is as in the main model of the paper, so that the presence of vertical training θ does not affect it.

Note furthermore that $k > j$ is needed for S_M to be positive. This reflects the fact that S_M involves a cost which θ does not, namely wages. The full explanation is as follows. Re-writing profits of M (27) using D of (28) and (29), M's problem can be restated as

$$\underset{\theta, S_M}{Max} \pi_M = (\theta S_M)^{\alpha} D - 2S_M - c\theta^k S_M^{\;j} \tag{30}$$

The expression above indicates that S_M and θ generate the same benefits and that therefore M has no way of differentiating between S_M and θ when looking purely at the benefits which they generate. These benefits consist of (a) revenue from production and (b) reduced wages over the two periods via the see-saw effect, and are captured by the term D. However, M can and does differentiate between S_M and θ by virtue of the fact that their costs are different: the cost of S_M includes both wages and the cost of training as captured by j, whilst the cost of θ includes only the cost of training as captured by k.

If $k \leq j$, there would be no incentive to carry out horizontal training, since this involves paying wages, whilst vertical training does not. Thus we require $k > j$.

Second order sufficiency conditions for a maximum are satisfied at the optimal choice of (S_M,θ) provided that all of the following are met:

$$j > 1 \quad or \quad j < 1 \quad and \quad (1-\alpha)\, k > j\,(1-j) \tag{31}$$

$$k > 1 \quad or \quad k < 1 \quad and \quad \alpha < k \tag{32}$$

$$\frac{(k-1)(1-\alpha)}{j} - \frac{(j-1)\alpha}{k} > 3 - 2\alpha \quad \forall\, j,k \tag{33}$$

In addition, the condition that profits of the multinational cannot fall below zero is assumed to be respected. At the optimal choice of (S_M,θ), this condition simplifies to

$$k(\frac{1}{\alpha} - 1) + j \geq 1 \tag{34}$$

The condition states that α has to be sufficiently low in relation to j and k, that is, the marginal productivity of trained workers must be sufficiently high in relation to the marginal cost of horizontal and vertical training. Note that ρ does not appear in (34); this is because even if ρ=0, M makes non-negative profits by virtue of the fact that first period revenues are positive and that the see-saw effect operates. Note also that n and β do not appear in (34); this is because the multinational can operate and make non-negative profits regardless of whether or not there are L's, or, equivalently, regardless of how productive L's are.

Since the term D of this section is the same as that of section 2, Proposition 1 continued to hold. In particular, if we set k = 0 and fix θ at some given value, so that vertical training is excluded, we have shown that Proposition 1 holds when we introduce training costs explicitly. The new result is as follows:

Proposition 3: If training by multinationals becomes more general, or if the number of rival firms increases, horizontal training always rises; vertical training rises if $j<1$ and falls if $j>1$.

Proof: Noting that $\partial D/\partial\beta > 0$ ($\partial D/\partial n>0$), differentiate (28) and (29) with respect to D. Use (34) to establish that the term $(\alpha-k)(1-j)+k(\alpha-j)$ in expressions (28) and (29) is negative.

The explanation of Proposition 3 hinges on the structure of costs, in particular the parameter j associated with horizontal training. When training becomes more general, or when the number n of rivals in period 2 is greater, the productivities of both vertical and horizontal training increase, as reflected in an increase in D. Upon considering profits of M as expressed in equation (30), we see that an increase in D increases the productivities of θ and S_M equally, in the sense that they rise by the same proportion. In deciding which of the two to increase, if not both, M considers costs. Since the vertical training cost parameter k exceeds the horizontal training cost parameter j, horizontal training always rises. To see whether or not vertical training rise, re-write costs as

$$Costs = S_M\left(2+\theta^k S_M{}^{j-1}\right) \qquad (35)$$

Thus, when S_M rises, the weight attached to θ^k, S_M^{j-1}, falls if $j<1$ and rises if $j>1$. Thus, when $j<1$, an increase in θ is consistent with minimizing costs in the multinational's problem (27).

Intuitively, if, for example, $j<1$, so that there are 'increasing returns to horizontal training', the increase in horizontal training is greater compared with that in the case of $j>1$; the large increase in horizontal training tends to increase the productivity of vertical training, and therefore the quantity of vertical training undertaken. According to Katz and Ziderman (1999), who write on training in developing nations, "it seems likely that [...] training [...] will be subject to economies of scale, in the sense that the training of more workers in a given specialization will be less costly per worker'. In the terminology of our model, this corresponds to the case of $j<1$.

5 The social planner

We next consider a social planner wishing to maximize welfare, which we define here as the sum of consumer surplus and profits. We are interested in finding the socially optimal values for horizontal and vertical training. Since wages received by workers are paid by firms, wages do not appear in the objective function of the social planner, which is given by revenues of local firms and of M. The social planner's problem is therefore:

$$\underset{\theta, S_M, S_L}{Max}\, n[\beta\theta S_L]^{\alpha} + [\theta S_M]^{\alpha} + [\rho\theta(S_M - nS_L)]^{\alpha} - c\theta^k S_M{}^j \tag{36}$$

where the multinational's period 2 profits are written taking into account that workers not poached by L's are employed by M.

We find that S_L* is proportional to S_M* in the same way as in the market outcome,

$$S_L^* = BS_M^* \qquad where \quad B^* = \frac{1}{n + \left(\frac{\beta}{\rho}\right)^{\frac{\alpha}{\alpha-1}}} \tag{37}$$

This result is not surprising since the period 2 market allocation is socially optimal by virtue of firms being price-takers.

Using (36) and (37), we can re-write the social planner's maximization problem as

$$\underset{\theta, S_M}{Max}\pi_M = [\theta S_M]^{\gamma} E - 2S_M - c\theta^k S_M{}^j \tag{38}$$

where $E = 1 + \rho^{\alpha}(1 - nB)^{\alpha} + n\beta^{\alpha}B^{\alpha}$

The solutions are

$$S_M * = \left\{ \frac{ck}{\alpha E}\left[\frac{2}{c(k-j)}\right]^{\frac{k-\alpha}{k}} \right\}^{\frac{k}{(\alpha-k)(1-j)+k(\alpha-j)}} \tag{39}$$

$$\theta * = \left\{ \frac{2}{c(k-j)} S_M^{*\,1-j} \right\}^{\frac{1}{k}} \tag{40}$$

The second order conditions are the same as those of the market equilibrium.

Notice the close resemblance between social planner's E of (38) and its market counterpart D of (29). For clarity of exposition, we re-write them: $E = 1 + \rho^{\alpha}(1 - nB)^{\alpha} + n\beta^{\alpha}B^{\alpha}$, and $D = 1 + \rho^{\alpha}(1 - nB)^{\alpha} + \alpha n\beta^{\alpha}B^{\alpha}$. D and E are the same except for the last of the three constituent terms, which is multiplied by α in D, but not in E. Given that D rises with n and β, and that $0<\alpha<1$, so does E. Hence, by symmetry with Proposition 3, we have the result that if training by multinationals becomes more general, or if the number of entrant firms in period 2 increases, the socially optimal level of horizontal training always rises; the socially optimal level of vertical training rises if $j<1$ and falls if $j>1$. The explanation is analogous to that of Proposition 3.

We can now compare the market allocation with that of the social planner.

Proposition 4: The market always underprovides horizontal training; it underprovides vertical training when $j<1$, and overprovides it when $j>1$.

Proof: by recognizing that the term E of the social planner in (38) is greater than the term D of the market equilibrium in (29), and then directly comparing the socially optimal and market amounts of horizontal and vertical training ((39) and (40), and (28) and (29) respectively).

The market underprovides horizontal training because (a) the multinational is not concerned with taking into consideration the n rivals' profits, and (b) the multinational seeks to exploit the see-saw effect fully by reducing horizontal training in period 1, raising the second period wage and thus lowering the first period wage. It can do so by virtue of its monopsony power in period 1.

Given that horizontal training is underprovided, the horizontal training parameter j determines if vertical training is underprovided or overprovided. If $j<1$, then the weight cS_M^j associated with vertical training θ in training costs $cS_M^j\theta^k$ in the social planners' problem is small, so that an increase in vertical training above the market level is socially desirable. If, on the other hand, $j>1$, the weight cS_M^j associated with vertical training θ in training costs $cS_M^j\theta^k$ in the social planners' problem is large, so that an increase in vertical training below the market level is socially desirable.

6 Multiple equilibria

In this section, we explore the possibility that there may be multiple equilibria due to the expectations of L's being self-fulfilling.

In order to find necessary conditions for multiple equilibria to exist, we would need to consider Nash Cournot firms, since we must concern ourselves with the market when there are only a few firms (indeed, only 2). To avoid presenting a new version of the model of section 2 with Nash firms, we continue to assume wage-taking firms, and look instead for sufficient conditions for multiple equilibria to exist.

Assume firms face a fixed cost of entry C. We will find a sufficient condition involving a range of C's for which the possibility of self-fulfilling equilibria exists.

Say there is a period zero, in which L's can take a (costless) action that commits them either to enter or not enter. This action can be the signing of a legal document. M moves in period 1, having observed whether L's have committed themselves to enter or not. L's can only enter in period 2 (perhaps because it takes a while for them to set up the necessary facilities).

If L's believe in period 1 that there won't be enough trained workers in period 2 to enable them to tolerate a given fixed cost C of entry, then they will choose not to enter, and will not sign the legal paperwork. We know that, via the see-saw effect, M will have a lower incentive to train and hire workers in period 1 than if L's did enter, and the expectation of L's will be confirmed. If, on the other hand, L's believe that there will be enough trained workers in period 2 to enable them to tolerate a given fixed cost of entry C, then they will choose to enter, and will sign the necessary paperwork. Via the see-saw effect, M will have a greater incentive to train and hire workers in period 1 than if L's did not enter, and the expectation of L's will be confirmed.

We use the basic model of section 2 with M in period 1 and worker homogeneity. If firms enter in period 2, we assume for simplicity that they and the ex-monopolist are identical, so that $\beta = \rho = 1$.

Let us begin by considering what the highest possible value of C is for a firm's expectation that it is not profitable to enter to be confirmed.

If L's commit not to enter, then M will operate as a monopolist in the 2 periods. In period 2, M will set the lowest possible wage, $w_2 = 1$. Since $w_1+w_2=2$, $w_1=1$. Thus the monopolist's period 1 problem is

$$\underset{S_M}{Max} \quad 2\left[(S_M)^{\alpha} - S_M\right] \tag{41}$$

The optimal S_M is

$$S_M = \left(\frac{1}{\alpha}\right)^{\frac{1}{\alpha-1}} \tag{42}$$

Now suppose that 1 L was to enter. The level of profit that it could make would be less than half the profits of a monopolist in period 2. These can be shown to be

$$\pi_{L,L\text{'s commit not to enter}} \leq \frac{1}{2}\left(\frac{1}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}(1-\alpha) \tag{43}$$

Therefore, a fixed cost of entry C higher than the level of profits given by (43) guarantees that it is not profitable for L to enter given that it has committed not to enter, so that its expectations are confirmed valid.

Now, let us consider what the lowest possible level of C is for a firm's expectation that it is profitable to enter to be confirmed. This is given by the lowest possible level of profits of L's, which, given that profits of L's are falling with the number n of L's, is found at n=∞. At n=∞, the negative effect of reduced market share is exactly balanced by the increase in the supply of workers S_M, so that profits of L's are constant. Using (12) with $n = \infty$,

$$\pi_{L,L\text{'s commit to enter}} = (1-\alpha)\left(\frac{\alpha^2}{2}\right)^{\frac{\alpha}{1-\alpha}} \tag{44}$$

Therefore, a fixed cost of entry C (weakly) lower than the level of profits given by (44) guarantees that it is profitable for L to enter given that it has committed to do so, so that its expectations are confirmed valid.

For multiple equilibria to be possible, it must be true that π_L given that L's commit not to enter be less than the fixed cost C of entry, and that π_L given that L's commit to enter be (weakly) greater than the fixed cost C of entry. In other words, if π_L given that L's commit not to enter are (weakly) less than π_L given that L's commit to enter, it is possible to find a fixed cost of entry C such that multiple equilibria are possible. Using (43) and (44), the condition for π_L given that L's commit not to enter to be (weakly) less than π_L given that L's commit to enter can be shown to be

$$\alpha^{\alpha} - 2^{2\alpha - 1} > 0 \qquad (45)$$

This expression can be shown to fall as α rises (for 0<α<1). In other words, returns to labour must be diminishing sufficiently steeply. This ensures that M benefits sufficiently from the presence of L's to make its choice of S_M be very large compared with that chosen when it faces no competition in period 2. When S_M is large, then profits of L's given that L's commit to enter will also be large, and therefore it will be possible for these to exceed profits of L's given that they commit not to enter.

Proposition 5: Fixed costs can always be found such that multiple equilibria are possible provided that (45) is satisfied.

Proof: see the derivation above.

Proposition 5 has implications for the development process. It suggests that, in the presence of diminishing returns to labour, whether or not a developing nation may experience industrialization in a particular rural area depends on the expectations of local firms which may or may not enter following the beginning of the industrialization process. It also suggests that there is a government role for co-ordinating the expectations of local firms.

7 Conclusion

The aim of this paper is to study the early stages of industrialization in LDC's where a multinational trains workers who are subsequently poached by local firms. In the process, we have shown that the see-saw effect may operate in the labour market. We have also shown the key role played by decreasing returns to scale in generating the main result that a multinational may enjoy higher profits when there is more competition for the workers which have undergone training by the multinational.

In essence, this paper offers an example of symbiosis between a multinational and local firms: it is possible for both parties to benefit from each other's presence without them co-operating with each other. This is reflected in the key results that the multinational's profits increase when local firms become better at utilizing workers trained by the multinational, and when the number of local firms rises. In both cases, according to the see-saw effect, the second period wage rises, enabling the multinational to offer workers a lower first period wage, which more than compensates the multinational for the higher second period wage. Local firms benefit from the presence of the multinational since the latter trains workers which local firms are assumed not to be capable of doing.

There are a number of avenues for further research.

First of all, it would be useful to explore how the elasticity of each worker's supply in the presence of worker heterogeneity affects the main result that more competition increases profits of the multinational. Secondly, if some workers who leave the multinational have entrepreneurial capabilities and set up their own firms, then the multinational, by determining how many workers leave its premises, would be able to control the number of local firms. It would be useful to explore how the multinational could use this extra control variable to take advantage of the see-saw effect. Thirdly, it would be instructive to consider the presence of two inputs in production, capital and labour, and to study how their interaction would affect our main result that increased competition for trained workers may increase profits of the multinational.

8 Appendix

7.1 Extending the main model of section 2 to allow for m price-taking firms in period 2 and for worker heterogeneity

For simplicity, we let firms be identical, so that $\beta = \rho = 1$.

The second period wage (see equation (3)) is then

$$w_2 = \alpha \ (S_2)^{\alpha-1}$$

where S_2 is the quantity of labour hired by each firm in period 2.

Total demand for workers in period 2 must equal supply:

$$(m+n)S_2 = S$$

Combing the two equations above we have

$$w_2 = \alpha\left(\frac{S}{m+n}\right)^{\alpha-1}$$

Similarly, in period 1 the wage is

$$w_1 = \alpha\ (S_1)^{\alpha-1}$$

Total demand for workers in period 2 must equal supply:

$$mS_1 = S$$

Combining the two equations above we have

$$w_1 = \alpha\left(\frac{S}{m}\right)^{\alpha-1}$$

To find the equilibrium number of workers S, we need to consider the condition whereby the marginal worker breaks even. Worker heterogeneity is assumed to take the form

$$F(S) = S^{\gamma}$$

Consequently, the condition stating that the marginal worker breaks even is

$$w_1 + w_2 - S^{\gamma} = 2$$

For algebraic tractability, we suppress the 2, implying that the benefit of working in agriculture is zero. The condition is thus

$$w_1 + w_2 - S^{\gamma} = 0$$

Substituting $w_1(S)$and $w_2(S)$ found above into the condition for the marginal worker to break even yields the equilibrium number of workers S:

$$S = \left\{\alpha\left[m^{\alpha} + (m+n)^{1-\alpha}\right]\right\}^{\frac{1}{\gamma+1-\alpha}}$$

Profits of m's are

$$\Pr_m = (1-\alpha)S^{\alpha}\left[\frac{1}{m^{\alpha}} + \frac{1}{(m+n)^{\alpha}}\right]$$

$$= (1-\alpha)^{\frac{\alpha}{\gamma+1-\alpha}}\left\{m^{\alpha} + (m+n)^{1-\alpha}\right\}^{\frac{\alpha}{\gamma+1-\alpha}}\left[\frac{1}{m^{\alpha}} + \frac{1}{(m+n)^{\alpha}}\right]$$

$$\frac{d\Pr_m}{dn} = \alpha(1-\alpha)^{\frac{\alpha}{\gamma+1-\alpha}}\frac{(m+n)^{\alpha+1}}{\left[m^{-\alpha} + (m+n)^{-\alpha}\right]^{\gamma-1}}\left\{\frac{(1-\alpha)(m+n)\left[m^{-\alpha} + (m+n)^{-\alpha}\right]}{\gamma+1-\alpha} - \left[m^{1-\alpha} + (m+n)^{1-\alpha}\right]\right\}$$

For sufficiently low worker heterogeneity as captured by a low γ, profits of m's rise with n. Conversely, they fall for sufficiently high worker heterogeneity. Profits of m's fall with m always.

Profits of L's are

$$\Pr_L = (1-\alpha)S^{\alpha}\left[\frac{1}{(m+n)^{\alpha}}\right]$$

$$= (1-\alpha)^{\frac{\alpha}{\gamma+1-\alpha}}\left\{m^{\alpha} + (m+n)^{1-\alpha}\right\}^{\frac{\alpha}{\gamma+1-\alpha}}\left[\frac{1}{(m+n)^{\alpha}}\right]$$

$$\frac{d\Pr_L}{dm} = \alpha(1-\alpha)^{\frac{\alpha}{\gamma+1-\alpha}}\frac{(m+n)^{\alpha+1}}{\left[m^{-\alpha} + (m+n)^{-\alpha}\right]^{\gamma-1}}\left\{\frac{(1-\alpha)(m+n)\left[m^{-\alpha} + (m+n)^{-\alpha}\right]}{\gamma+1-\alpha} - \left[m^{1-\alpha} + (m+n)^{1-\alpha}\right]\right\}$$

For low worker heterogeneity as captured by a low γ, profits of L's rise with m. Conversely, they fall for sufficiently high worker heterogeneity. Profits of L's fall with n always.

Remarkably, there is complete symmetry in that

$$\frac{d\,\mathrm{Pr}_m}{dn} = \frac{d\,\mathrm{Pr}_L}{dm}$$

The equality can be explained as follows.

First, consider the derivative on the left-hand side. 1 more firm in period 2 implies only a small increase in competition over the two periods for each firm present from period 1, but it also generates a small increase in the number of workers. So, there are two small effects working against each other, resulting in a net effect of a certain size.

Second, consider the derivative on the right-hand side. 1 more firm in period 1 implies a large increase in competition over a single period (the second) for each firm in period 2, but it also generates a large increase in the number of workers. So, there are two large effects working against each other, resulting in a net effect of the same size as the one described above.

7.2 Extending the main model of section 2 to allow individually elastic supply of labour, worker heterogeneity and either M or for m price-taking firms in period 1

7.2.1 M in period 1

In period 1, M maximizes profits by choosing the total number of hours $S_M s_{A,1}$, where we have replaced S in the notation in the main text with S_M to show that there is a monopsony in period 1. However, the marginal worker constraint (19) implies that $s_{A,1}$ is a function of S_M, so that M is only free to choose S_M. Since firms are identical, each firm in period 2 will employ $S_M/(1+n)$ of the labour force, since there are n+1 firms in period 2. M's period 1 problem is then

$$\underset{S_M}{Max}\ \pi_M = f(S_M s_{A,1}) - w_1 S_M s_{A,1} + f\left(\frac{S_M}{1+n} s_{A,2}\right) - w_2 \frac{S_M}{1+n} s_{A,2}$$

subject to (18), (19) and (22)

Next, recognizing from our previous discussion of the marginal worker condition that w_1 is a function of w_2 and of S, we can write profits of M to show that all variables are ultimately determined by the control variable S_M:

$$\pi_M = f(S_M s_{A,1}(w_1(w_2(S_M),F(S_M)))) - w_1(w_2(S_M),F(S_M))S_M s_{A,1}(w_1(w_2(S_M),F(S_M)))$$
$$+ f\left(\frac{S}{1+n}s_{A,2}(w_2(S_M))\right) - w_2(S_M)\frac{S_M}{1+n}s_{A,2}(w_2(S_M))$$

Upon differentiating the above expression with respect to S_M, and applying Hotelling's lemma in period 2, we find that

$$\frac{d\pi_M}{dS_M} = \frac{\partial f}{\partial s_{A,1}}\frac{ds_{A,1}}{dw_1}\frac{dw_1}{dS_M} + \frac{\partial f}{\partial S} - \left[s_{A,1}S\frac{dw_1}{dS_M} + w_1 S_M \frac{ds_{A,1}}{dw_1}\frac{dw_1}{dS_M} + w_1 s_{A,1}\right]$$
$$- s_{A,2}\frac{S}{1+n}\frac{dw_2}{dS_M} = 0$$

$$where \quad \frac{dw_1}{dS_M} = \frac{\partial w_1}{\partial w_2}\frac{dw_2}{dS_M} + \frac{\partial w_1}{\partial F}\frac{dF}{dS_M} \qquad \text{(A7.2.1.1)}$$

The last expression comes from the marginal worker condition, and shows the link between w_1, w_2 and S_M discussed above.

We now show that the marginal product of S_M is greater than w_1, a consequence of M being a monopsony. Rearranging the terms in (A7.2.1.1), we have

$$
\left[\frac{\partial f}{\partial s_{A,1}} - w_1 S_M\right]\frac{ds_{A,1}}{dw_1}\frac{dw_1}{dS_M}\left[\frac{\partial w_1}{\partial w_2}\frac{dw_2}{dS_M} + \frac{\partial w_1}{\partial F}\frac{dF}{dS}\right] + \left[\frac{df}{dS} - w_1 s_{A,1}\right] =
$$

$$
-\frac{\partial w_2}{\partial S_M}\left[\frac{\partial w_1}{\partial w_2}s_{A,1}S_M + s_{A,2}\frac{S_M}{1+n}\right] + \frac{\partial w_1}{\partial F}\frac{dF}{dS}s_{A,1}S_M
$$

(A7.2.1.2)

Using (21) to recognize that

$$
\frac{\partial w_1}{\partial w_2} = -\frac{s_{A,2}}{s_{A,1}},
$$

where the partial sign refers to the fact that that we are keeping S_M constant (this is useful in view of our use of the envelope theorem later on).

(A7.2.1.2) simplifies to

$$
\left[\frac{\partial f}{\partial s_{A,1}} - w_1 S_M\right]\underset{+}{\frac{ds_{A,1}}{dw_1}}\left[\underset{-}{\frac{\partial w_1}{\partial w_2}}\underset{-}{\frac{dw_2}{dS_M}} + \underset{+}{\frac{\partial \psi}{\partial F}}\underset{+}{\frac{dF}{dS_M}}\right] + \left[\frac{\partial f}{\partial S_M} - w_1 s_{A,1}\right] =
$$

$$
-\underset{-}{\frac{\partial w_2}{\partial S_M}}\left[s_{A,2}\left(-S_M + \frac{S_M}{1+n}\right)\right] + \underset{+}{\frac{\partial w_1}{\partial F}}\underset{+}{\frac{dF}{dS}}s_{A,1}S_M
$$

(A7.2.1.3)

Next, we recognize that, as a result of $s_{A,1}$ and S_M being perfect substitutes in production, we can write

$$
\frac{\partial f}{\partial s_{A,1}} = \frac{\partial f}{\partial S_M}\frac{S_M}{s_{A,1}} \qquad \text{(A7.2.1.4)}
$$

Using (A7.2.1.4), we can see that the sign of $\partial f/\partial S_M$ - w_1S_M in the first set of square brackets must be the same as that of $\partial f/\partial S_M - w_1s_{A,1}$ in the third set of square brackets. Since the right-hand side of (A7.2.1.3) is positive, it follows that these two terms in square brackets are also positive, so that

$$\frac{\partial f}{\partial s_{A,1}} - w_1 S_M > 0 \quad and \quad \frac{\partial f}{\partial S_M} - w_1 s_{A,1} > 0 \qquad \text{(A7.2.1.5)}$$

As usual, we wish to consider how an increase in w_2, which in turn is the result of increased competition for workers in period 2, affects M's profits. Applying the envelope theorem and Hotelling's lemma,

$$\frac{d\pi_M}{dw_2} = \frac{\partial w_1}{\partial w_2}\left\{\frac{ds_{A,1}}{dw_1}\left[\frac{\partial f}{\partial s_{A,1}} - w_1 S_M\right]\right\} - s_{A,2}\frac{S_M}{1+n} = 0 \qquad \text{(A7.2.1.6)}$$

Next, we take (A7.2.1.2), and multiply both sides of by the inverse of dw_2/dS_M. Using the terms in this modified version of (A7.2.1.2) to substitute for various terms in (A7.2.1.6), we have that

$$\frac{d\pi_M}{dw_2} = -\frac{1}{\frac{dw_2}{dS_M}}\left\{\frac{ds_{A,1}}{dw_1}\frac{\partial w_1}{\partial F}\frac{dF}{dS_M}\left[\frac{\partial f}{\partial s_{A,1}} - w_1 S_M\right] + \left[\frac{\partial f}{\partial S} - w_1 s_{A,1}\right] - s_{A,1}S_M\frac{\partial w_1}{\partial F}\frac{dF}{dS_M}\right\}$$

Since dw_2/dS_M is negative,

$$Sign\left\{\frac{d\pi_M}{dw_2}\right\} = Sign\left\{\underset{+}{\frac{\partial w_1}{\partial F}}\underset{+}{\frac{dF}{dS_M}}\left[\underset{+}{\frac{ds_{A,1}}{dw_1}}\underbrace{\left(\frac{\partial f}{\partial s_{A,1}} - w_1 S_M\right)}_{+} - s_{A,1}S_M\right] + \underbrace{\left(\frac{\partial f}{\partial S} - w_1 s_{A,1}\right)}_{+}\right\}$$

(A7.2.1.7)

Note that, if workers are homogeneous in their cost of entry into industry, $\partial F/\partial S_M$ is zero, and the expression is positive. Therefore, if the second period wage increases as a result of increased competition in period 2, M's profits always increase. If, on the other hand, dF/dS_M is positive and sufficiently large, the expression may be negative, because a rise in the second period wage will be accompanied an insufficiently large reduction in the first period wage. Whether or not the expression can be negative, however, remains to be shown using a specific model, which we are not in a position to do since the model of section 2 does not extend easily to include the assumptions of both worker heterogeneity and individually elastic supply of labour.

7.2.2 m wage-taking firms in period 1

The analysis for m wage-taking firms in period 1 is simpler than that with a monopsony. The problem of firms in period 1 is analogous to that in period 2:

$$\underset{S_{i,1}s_{A,1}}{Max}\ \pi_{i,1} = f(S_{i,1}s_{A,1}) - w_1 S_{i,1}s_{A,1}$$ (A7.2.1.8)

which generates the inverse demand function

$$w_1 = \frac{df(S_{i,1}s_{A,1})}{d(S_{i,1}s_{A,1})}$$ (A7.2.1.9)

Consider how profits of m's change when the second period changes as a result of an increase in competition:

$$\frac{d\pi_m}{dw_2} = -\frac{dw_1}{dw_2}\frac{S}{m}s_{A,1} + \frac{S}{m+n}s_{A,2} \qquad \text{(A7.2.1.10)}$$

where S is the total number of workers in the industry.

Using (21), which describes how the marginal worker condition varies with the second period wage, we have that

$$\frac{d\pi_m}{dw_2} = S\,s_{A,2}\left(\frac{1}{m} - \frac{1}{m+n}\right) - \frac{S}{m}\frac{dF}{dS}\frac{dS}{dw_2} \qquad \text{(A7.2.1.11)}$$

Therefore, when workers are homogeneous, so that dF/dS=0, profits of the m firms in period 1 always increase with the second period wage resulting from an increase in competition in period 2. When workers are not homogenous, so that dF/dS>0, establishing whether or not profits of the m firms can fall with the second period wage requires that we set up a specific model. Since the model of section 2 does not extend easily to include the assumptions of both worker heterogeneity and individually elastic supply of labour, we cannot carry out this analysis.

<u>7.3 Proof that an equilibrium in which hiring from agriculture takes place in period 2 is only possible if workers are homogeneous in their cost of entry</u>

For the sake of increased generality, in particular to show that this result is not dependent on the time discount factor being zero, we show that it holds for any time discount factor $\delta = (0,1)$, and for any wage in agriculture. We assume that the marginal product of agriculture, MPA, is constant no matter how many workers are drawn away from it into the industrial sector; with perfect competition in agriculture, the wage is also constant and equal to MPA. We also allow individual workers to supply their work elastically, again for greater generality.

We label workers hired from agriculture in period 1 S_1, and workers hired from agriculture in period 2 S_2.

For an equilibrium in which workers are hired in period 2, a number of conditions need to be met. We list three which, together, are sufficient to demonstrate that, for an equilibrium in which workers are hired in period 2, workers must be homogeneous in their fixed cost of entry.

(a) The marginal worker who enters in period 1 does so because in period 1 he finds it (weakly) more beneficial than entering in period 2:

$$w_1 s_{A,1} + \delta\, w_2 s_{A,2} - F(S_1) \geq MPA + \delta\left(w_{2,N} s_{A,2,N} - F(S_1)\right)$$

where $w_{2,N}$ and $s_{A,2,N}$ are the wage enjoyed by newly hired workers in period 2 and the hours worked by newly hired workers, respectively. We retain previously used notation for variables relating to workers hired already from the start of period 1.

(b) The marginal worker who enters in period 1 and works in industry in both periods is in equilibrium indifferent between doing this and working in agriculture in both periods:

$$w_1 s_{A,1} + \delta\, w_2 s_{A,2} - F(S_1) = (1+\delta) MPA$$

(c) The marginal worker who enters in period 2 is indifferent between doing this and working in agriculture:

$$w_{2,N} s_{A,2,N} - F(S_1 + S_2) = MPA$$

Combining conditions (a), (b) and (c), we find the following reduced condition:

$$F(S_1) + F(S_1 + S_2) \geq 0$$

which can only be satisfied if workers are homogeneous, i.e. if dF/dS=0.

Thus, for an equilibrium with hiring of new workers in period 2 to be possible, workers must be homogeneous in their fixed cost of moving from agriculture to industry.

Chapter 5

Conclusion

This thesis has made a number of contributions to our understanding of how more competition may increase profits of an incumbent. These are the most significant:

1. In chapter 2, we have extended the commitment model of Farrell and Gallini (1988), which assumes Bertrand competition, to Cournot competition. This has enabled us to identify the optimal number of rivals from the incumbent's point of view. It has also enabled us to show that an incumbent may invite rivals to enter from the time the product is launched, extending the commitment explanation of why an incumbent may invite rivals to second-source agreements in which rivals enter without delay. An incumbent has an incentive to invite a rival to enter without delay to 'share' the losses incurred in the initial period of production. Only in the model of chapter 2 is it possible for an incumbent to gain when the number of rivals increases in period 1. This suggests that the commitment mechanism is in a sense more powerful than the decreasing returns to scale mechanism of chapters 3 and 4.

2. Still in chapter 2, we have shown that whether or not profits of the incumbent increase with the number of rivals in period 2 can depend on how many buyers enter as a consequence of more competition lowering the second period price. How many buyers enter in turn depends on the degree of heterogeneity of buyers: if the set-up cost of buyers rises sufficiently slowly with the number of buyers, i.e. if buyers are sufficiently homogeneous, then many buyers will enter, and the incumbent will gain from increased competition. We obtain similar results in chapters 3 and 4.

3. In chapter 2, as in chapter 3, it is possible for all agents, i.e. all firms and buyers, to benefit from increased competition. The incumbent's gain is reflected in the fact that it invites entry; consumers gain from lower prices and higher quantities; entrant firms gain from the entry of other entrant firms by virtue of the increased number of buyers.

4. In chapter 3, we have suggested a new explanation of why an incumbent may want to invite rivals, namely to relax the capacity constraint at the industry level in the second period. This in turn raises consumer surplus in period 2, so that consumers are willing to tolerate a higher price in period 1. As in chapter 2, we have identified the optimal number of rivals from the incumbent's point of view.

5. In chapter 3, we have shown that the incumbent invite rivals if the simple condition is satisfied that in period 2 price is less than its marginal cost.

6. In chapter 4, we have applied the theoretical principle in chapter 3 to the labour market. A multinational which trains workers may enjoy higher profits when more of its workers are poached by rival firms if it is compensated by a sufficiently lower wage during training. Specifically, when demand for trained workers increases, workers considering training predict that the wage after training will be higher, and are willing to tolerate a lower wage during training. Demand for trained workers increases if there are more non-training firms or if training is of a more general nature. If there are lots of incumbents in the first period, the model is a straight-forward extension of the classic, decreasing returns to scale model with competitive firms from a one period setting to a two period setting. It may be surprising that in such a simple model more competition may increase profits of an incumbent.

7. Still in chapter 4, we have shown that increased competition for trained workers, whilst it always leads to an increase in the number of trained workers in our model, the quantity of training per worker may rise or fall, depending on training cost parameters.

8. In all chapters, we have found that multiple equilibria are possible due to the self-fulfilling expectations of rival entrants.

In summary, the thesis' contribution consists in having expanded upon an existing explanation of why incumbents invite rivals, developed a new explanation, and applied that new explanation to a novel setting, that of the labour market. Essentially, firms may benefit from each other's presence in a non-cooperative environment, in a process of symbiosis.

What possible avenues for further research are there? Testing the theory is probably the next logical step. Our attempts to find Mexican data to test the theory of chapter 4 have been unsuccessful. However, since the theory is more general than the specific example discussed in chapter 4 suggests, it may be possible to test it using data from developed nations. There is already some evidence from a CBI survey in 1989 indicating that many employers are not deterred by the prospect of losing their trained workers to rivals when deciding whether or not to train them.

There are two important theoretical extensions to explore in further research:

1. The result in chapter 3 that an incumbent invites rivals if in period 2 price (p_2) is less than its marginal cost (MC_M) is more general than our model suggests. In particular, because the incumbent is assumed to be able to commit, the period 1 and period 2 markets could be interpreted as two markets for two complementary products in the same period, e.g. hardware and software. The question then arises, in which market would the incumbent wish to invite rivals? Is it the one in which MC_M-p_2 is greatest? Or does the incumbent invite rivals in both markets, provided MC_M-p_2>0?

2. It would be interesting to explore how R&D and the see-saw effect interact. Since the literature indicates that more competition may either increase or decrease R&D, depending on the innovator's ability to appropriate the returns from R&D, it is unclear whether the presence of the see-saw makes more competition result in higher R&D or less, and might therefore be worthy of investigation. This extension could then be followed by the study of joint R&D between subsequent rivals, which is valuable considering that in the last decade there has been a shift in the semiconductor industry - our source of evidence for chapters 2 and 3 - away from simple second-sourcing to joint R&D between subsequent rivals.

References

1 References relating to chapters 2 and 3 on industrial organization

van Ackere,A and Reyniers, DJ, 1995, 'Trade-ins and introductory offers in a monopoly, Rand Journal of Economics, Vol 26, No 1, pp58-74, Spring

Anton,J and Yao,D, 1987, 'Second sourcing and the experience curve: price competition in defense procurement', Rand Journal of Economics, Vol 18, No 1, Spring, pp 57-76

Axelford,R, Mitchell,W, Thomas,RE, Scott Bennett,D and Bruderer,E, 1995, 'Coalition formation in standard-setting alliances', Management Science, Vol 41, No 9, September

Begg,A and Klemperer,P, 1992, 'Multi-period competition with switching costs', Econometrica, Vol 60, No 3, pp651-666, May

Bernhrdt,D and Joshua,S, 1998, 'Pirated for profit', Canadian Journal of Economics, Vol 31, No 4, October

Borenstein,S, MacKie-Mason,JK and Netz,JS, 1995, 'Antitrust policy in aftermarkets', Antitrust Law Journal, 63, pp455-482

Borenstein,S, MacKie-Mason,JK and Netz,JS, 1996, 'Exercising market power in proprietary aftermarkets', University of California, Berkeley, Working Paper

Business Week, 1989, 'Clonemakers don't scare Sun – It's sending them engraved invitations', July 24

Chen,Y, 1997, 'Paying customers to switch', Journal of Economic and Management Strategy, Vol 6, No 4, pp877-897, Winter

Church, J and Gandal,N, 1992, 'Network effects, software provision, and standardization', Journal of Industrial Economics, Vol 40, No 1, pp85-103

Conner,K, 1995, 'Obtaining strategic advantage from being imitated: when can encouraging "clones" pay?', Management Science, Vol 41, No 2, February, pp209-225

Conner,K and Rumlet,R, 1991, 'Software piracy': an analysis of protection strategies', Management Science, Vol 37(2), February, pp125-139

Costa,AC and Dierickx,I, 2002, 'Licensing and bundling', International Journal of Industrial Organization, 20, pp251-267

Crampes,C and Hollander,A, 1993, 'Umbrella pricing to attract early entry', Economica, 60, pp465-474

Dick,A, 1992, 'An efficiency explanation for why firms second source', Economic Enquiry, No 2, pp332-354

Dutta,J and John,G, 1995, 'Combining lab experiments and industry data in transaction cost - analysis - the case of competition as a safeguard', Journal of Law Economics and Organization, Vol 11, No 1, pp87-111

Economides,E, 1996a, 'Network externalities, complementarities, and invitations to enter', European Journal of Political Economy, Vol 12, pp211-233

Economides, N, 1996b, 'The economics of networks', International Journal of Industrial Organization, 14, pp673-699

Farrell,J and Gallini,N, 1988, 'Second sourcing as a commitment: monopoly incentives to attract competition', Quarterly Journal of Economics, 103, pp383-406

Farrell,J and Shapiro,J, 1987, 'Optimal contracts with lock-in', American Economic Review, 79(1), pp51-68

Farrell,J and Shapiro,J, 1987, 'Optimal contracts with lock-in', Rand Journal of Economics, 19, pp123-137

Gallini,N, 1984, 'Deterrence by market sharing: a strategic incentive for licensing', The American Economic Review, Vol 74, No 5, pp931-941, December

Gallini,N and Karp,L, 1989, 'Sales and consumer lock-in', Economica, Vol 56, pp279-294

Grindley,PC and Teece,DJ, 1997, 'Managing intellectual capital: licensing and cross-licensing in semiconductors and electronics', California Management Review, 39(2), Winter

Hall, BH and Ham, RM, 1999, 'The patents paradox revisited: determinants of patenting in the US semiconductor industry, 1980-94', National Bureau of Economic Research Working Paper No. 7062, March

Holmes,T, 1990, 'Consumer investment in product-specific capital - The monopoly case', Quarterly Journal of Economics, Vol 105, No 3, pp789-801

Katz,ML and Shapiro,C, 1992, 'Product introduction with network externalities', Journal of Industrial Economics, 40(1), pp55-84

Katz,L and Shapiro,C, 1986, 'Technology adoption in the presence of network externalities', Journal of Political Economy, Vol 94, No 4

Kende,M 1998, 'Profitability under an open versus a closed system', Journal of Economics and Management Strategy, Vol 7, No 2, pp307-326

Kende,M, 1995a, 'Licensing, software synergy, and indirect network externalities', INSEAD WP, 95/46/EPS

Kende,M, 1995b, 'Licensing and the battle between the standards, INSEAD WP, 95/47/EPS

Kende,M, 1992, 'Strategic standardization in trade with network externalities', INSEAD WP, 92/50/EP

Klemperer,P, 1987, 'Markets with consumer switching costs', Quarterly Journal of Economics, May

Klemperer,P, 1988, 'Welfare effects of entry into markets with switching costs', Journal of Industrial Economics, Vol 37, No 2, pp159-165, December

Klemperer,P, 1995, 'Competition when consumers have switching costs: an overview with applications to industrial organization, macroeconomics, and international trade', Review of Economic Studies, Vol 62, pp515-539

Lovell, EB, 1968, 'Domestic licensing practises', The Conferences Board, Experiences in Marketing Management, 18

McAfee,R and Schwartz,M , 1994, 'Opportunism in multilateral vertical contracting - nondiscrimination, exclusivity , and uniformity', American Economic Review, Vol 84, No 1, pp210-230

Reinganum,JF, 1989, The timing of innovation: research, development, and diffusion', in Handbook of Industrial Organization, edited by Schmalensee,R and Willig,RD, Vol 1

Riordan,M and Sappington,D, 1988, 'Commitment in procurement contracting', Scandinavian Journal of Economics, 90(3), pp357-372

Riordan,M and Sappington,D, 1989, 'Second sourcing', Rand Journal of Economics, Vol 20, No 1, Spring, pp41-58

Riordan,M and Sappington,D, 1988, 'Commitment in procurement contracting', Scandinavian Journal of Economics, 90(3), pp357-372

Rosenblaum,J, 1985, 'Licensing second-source suppliers', Les Nouvelles, June, pp93-94

Rosenthal,R 1980, 'A model in which an increase in the number of sellers leads to a higher price', Econometrica, Vol 48, No 6, September

Rossell,M and Walker,C, 1999, 'Entry solicitation and deterrence in R&D competition with spillovers', The Japanese Economic Review, Vol 50, No1, march, pp96-103

Segal,IR, and Whinston,MD, 2000, 'Exclusive contracts and protection of investments', Vol 31, No 4, pp603-633

Shapiro,C and Teece,DJ, 1994, 'Systems competition and aftermarkets: an economic analysis of Kodak', The Antitrust Bulletin, Spring, pp135-162

Shapiro,C, 1995, 'Aftermarkets and consumer welfare: making sense of Kodak', Antitrust Law Journal, 63, pp483-511

Shephard,A, 1987, 'Licensing to enhance demand for new technologies', Rand Journal of Economics, Autumn, pp360-368

Sirbu,M and Kent,H, 1986, 'Standardization of local area networks', Carnegie Mellon University Working Paper

Swann,GMP, 1985, 'Product competition in microprocessors', Journal of Industrial Economics, 34(1), September, pp33-53

Swann,GMP, 1987, 'Industry standard microprocessors and the strategy of second-source production' in 'Product standardization and competitive strategy, edited by Gabel,H, Amsterdam: North Holland, pp239-262

Taylor,R, 1984, 'Licensing in theory and practice: licensor-licensee relationships', Antitrust Law Journal, 53(3), pp561-609

Teece,D, 1986, 'Profiting from technological innovation: implications for integration, collaboration, licensing and public policy', Research Policy, 15, pp285-305

Tirole,J, 1988, 'The theory of industrial organization', MIT Press

United Nations Centre on Transnational Corporations, 1986, 'Transnational corporations in the international semiconductor industry', New York: United Nations, especially pp133-141

Varian,HR, 2001, 'Economics of information technology', http://www.sims.berkeley.edu/~hal/papers/mattioli.html

2 References relating to chapter 4 on labour economics

Autor, DH, 2000, 'Why do temporary help firms provide free general skills training?', NBER Working Paper n.7637

Acemoglu, D and Pischke,J, 1999, 'Beyond Becker: training in imperfect labour markets', The Economic Journal, n.109 (February)

Acemoglu, D and Pischke,J, 1998, 'The structure of wages and investment in general training', NBER Working Paper n.6357

Acemoglu,D, 1996a, 'Credit constraints, investment externalities and growth, in Booth,AL and Snower,DJ, 'Acquiring skills: market failures, their symptoms and policy responses', Cambridge University Press.

Acemoglu, D and Pischke,J, 1996b, 'Why do firms train? Theory and evidence', NBER Working Paper #5605, June

Arvin,B, 1993 'A model of general training, banking, and labour turnover', Bulletin of Economic Research, 45:3, p197-214

Autor,DH, 2000, 'Why do temporary help firms provide free general training?', NBER Working Paper No 7637, April

Von Bardeleben,R, Beicht,U and Fehér,K, 'Betriebliche Kosten und Nutzen der Ausbildung. Repräsentative Ergebnisse aus Industrie, Handel und Handwerk', Berichte zur beruflichen Bildung Heft 187. Bielefend: Bertelsmann

Becker,G, 1964, 'Human capital', New York, NBER, First edition.

Blömstrom,M, 1991, 'Host country benefits and foreign investment', in McFetridge,DG, 'Foreign investment, technology and economic growth', Toronto University Press.

Blömstrom,M and Kokko,A (1998), 'Multinational corporations and spillovers', Journal of Economic Surveys, 12(3), pp247-278

Blomstrom,M and Wang,JY, 1992, 'Foreign investment and technology transfer, International Economic Review?

CBI (Confederation of British Industry), 1989, 'Towards a Skills Revolution', London

Chang,C and Wang, 1995, 'A framework for understanding differences in labour turnover and human capital investment', Journal of Economic Behaviour and Organization, Vol 28, p91-105.

Chen,EKY, 1983, 'Multinational corporations, technology and employment', Macmillan, London

Das,S, 1987, 'Externalities, and technology transfer through multinational corporations', Journal of International Economics, 22, p171-182

Daesden,L, Machin,S, Reed,H and Wilkinson,D, 1997, 'Labour turnover and work-related training', May

Dunning,JH, 1992, 'Multinational enterprises and the global economy', Addison-Wesley Publishing Company.

Ethier,WJ and Markusen,JR, 1996, 'Multinational firms, technology diffusion and trade', Journal of International Economics, 41, pp1-28

Fairburn,JA and Malcomson,JM, 1994, 'Rewarding performance by promotion to a different job', European Economic Review, 38, p683-690

Feuer,M, Glick,H and Desai,A, 1987, 'Firm-financed education and specific human capital: a test of the insurance hypothesis', in 'Market failures in training?', Stern,D and Ritzen,J editors, Springer-Verlag

Feuer,M, Glick,H and Desai,A, 1991, 'Is firm-sponsored education viable?', Journal of Economic Behaviour and Organization, Vol 8, p121-136.

Fosfuri,A, Motta,M and Rønde,T, 2001, 'Foreign direct investment and spillovers through workers' mobility', Journal of International Economics, 53, pp205-222

Fujita,M and Thisse, JF, 1996, 'Economics of agglomeration', Journal of the Japanese and International Economies, 10, pp339-378

Gershenberg,I, 1990, 'The training and dissemination of managerial know-how in LDC's: Jamaica', Social and Economic Studies, 39:2

Gershenberg,I, 1987: 'The training and spread of managerial know-how, a comparative analysis of multinational and other firms in Kenya', World Development, 15:7, p931-939.

Glass,AJ and Saggi,K, 1999, 'Multinational firms, technology transfer and welfare', mimeo

Haacker,M, 1998, 'Spillovers from foreign investment through labour turnover: the supply of 'management skills', LSE, mimeo

Hamilton,J, Thisse,J and Zenou,Y, 2000, 'Wage competition with heterogeneous workers and firms', Journal of Labour Economics, Vol 18, No 3

Hyman,J, 1992, 'Training at work', Routledge

Jenkins,R, 1988, 'Comparing foreign subsidiaries and local firms in LDC's: theoretical issues and empirical evidence', ? [School of Development Studies, University of East Anglia].

Jones,I, 1988, 'An evaluation of YTS', Oxford Review of Economic Policy, 4, pp54-71

Katz,E and Ziderman,A, 1990, 'Investment in general training: the role of information and labour mobility', The Economic Journal, 100, December, p1147-1158.

Katz,E and Ziderman,A, 1999, '' in 'Regulation Strategies and Economic Policies – Essays in honour of Bernard Corry and Maurice Peston', Vol 3, edited by Daniel,S, Arestis,P and Grahl,J, Blackwell

Kokko,A, Tansini,R and Zejan,C, 1996, 'Local technological capability and productivity spillovers from FDI in the Uruguayan manufacturing sector', Journal of Development Studies, 32.4, April, p602-611.

Lilliard,R and Tan,H , 1986, 'Private sector training', Rand Corporation, R-3331.

Lynch,L, 1988, 'private sector training and the impact', Draft, MIT, October.

Markusen,JR and Venables,AJ, 1999, 'Foreign direct investment as a catalyst for industrial development', European Economic Review, 43, pp335-356

Mincer,J, 1994, 'Human capital - A review', in 'Labour Economics and Industrial Relations', Kerr,C and Staudohar,P editors, Harvard University Press

Motala,E, 1995, 'Comprehensive science and technology, education and training and labour market policies for manufacturing growth', Development Southern Africa, Vol 12, No 1, February

Noorbakhsh,F, Paloni,A and Youssef,A, 2001, 'Human capital and FDI inflows to developing nations: new empirical evidence', World Development, Vol 29(9), pp1593-1610

Niosi,J, Hanel,P and Fiset,L, 1995, 'Technology transfer to developing countries through engineering firms: the Canadian experience', World Development, 23, Vol 10, p1815-1824

Pack,H 1993, 'Exports and externalities: the sources of Taiwanese growth', mimeo

Pack,H and Saggi,H, 1997, 'Inflows of foreign technology and indigenous technological development', Review of Development Economics, 1(1), pp81-98

Ritzen,J, 1991, 'Market failure for general training, and remedies', in Market Failure in Training, Stern,D and Ritzen,J editors, Berlin Heidelberg

Rotenberg,JJ and Saloner,G, 2000, 'Competition and human capital accumulation: a theory of interregional specialization and trade, Regional Science and Urban Economics, 30, pp373-404

Shackleton,JR, 1992, 'training too much? A sceptical at the economics of skill provision in the UK', Hobart Paper 118, Institute of Economic Affairs, London

Schutz,N and Stahl,K, 1996, 'Do consumers search for the highest price? Equilibrium and monopolistic optimum in differentiated products markets', Rand Journal of Economics, 27, pp542-652

Stevens,M, 1994, 'A theoretical model of on-the job training with imperfect competition', Oxford Economic Papers, 46, pp537-562

Stevens,M, 1999, 'Human capital theory and UK vocational policy', Oxford Review of Economic Policy, Vol.15, n.1

Venables,AJ, 1995, 'International Location and Economic Efficiency - Economic Integration and the Location of Firms', American Economic Review, Vol 85, No 2

Venables, AJ, 1996, 'Equlibrium locations of vertically linked industries', International Economic Review, Vol 37, No 2, May

Walkenhorst,P, 2000, 'Foreign direct investment, technological spillovers and the agricultural transition in central Europe', Post-Communist Economies, Vol 12(1)

Walsh,J, 1993, 'Internalization v. Decentralization: an analysis of recent developments in pay bargaining', British Journal of Industrial Relations, 31:3, September1

United Nations Conference on Trade and Development, 1998, 'Foreign direct investment and development'