Strategic complementarities and network effects

Filomena Garcia

Ph.D. Thesis

Université Catholique de Louvain
Centre for Operations Research and Econometrics - CORE

January 2006

List of Figures

Contents

To Othman, with love

Acknowledgments

Many times before I have thought about what I would write in the acknowledgements of my thesis, but now, that the time is due, I feel overwhelmed by the number of people whose support along these years must be accounted for.

When I first left to Barcelona, I had in my mind that I would do a PhD. I had set up a goal, but I must confess that I was totally unprepared for the challenge that I was about to face. Since that moment I have been through many changes, I moved country three times and in each place, I had to recreate my life. I have been fortunate to find people along the way that supported and that taught me a great deal of things. I am thankful to all of them, especially the great friends I have left from my time in Barcelona and York, and that now are spread around the world.

In the changes that I have been through during the writing of my thesis I had always a fixed point, an anchor that held the boat in the hardest times. This fixed point was my family and friends back home. They were always there for me. Each time that I returned, they welcomed me as if I had never left. And when going back abroad was too difficult they and gave me the strength to keep on. For all their love and patience I am grateful. *Agradeço aos meus pais e padrinhos, a verdadeira prova de amor que me deram durante estes anos, apoiando-me e dando-me força para continuar em todas as circunstâncias.*

Even though once and a while I had to return to Portugal for reassurance and balance, I must say that I have found in Louvain-la-Neuve a true home and my years spent there have been very enriching. I thank my flatmates for their infinite patience and care and all the friends that have surrounded me with almost familiar attention and love. Writing a thesis is not always cheerful, but I had always the chance to find a helping hand to put me back on my feet when things went wrong and to share the joys that this work also brings. I like to think that I had an angel with me all along the way, and I am grateful for this.

I am indebted to Paolo for the support that he has given me in all these years. Without him, I wouldn't have had the courage to overcome all the adversities that I have endured.

I have not been alone in the writing of this thesis. The environment at CORE is unique and it provided me with the intellectual stimulus that I needed to advance in my research. From my short experience, I learned that people at CORE are nicer than in any other research centre, and the quality of the work produced there is a reflection of this warm environment. Besides sympathy I also benefited from the academic excellence that characterizes researchers at CORE. I am grateful to Rabah Amir, for having believed in my capacities when I transferred from York and for the countless number of discussions we had along these years.

As for Jean J. Gabszewicz, words wouldn't be enough to express the joy that it was to work with him. He has been a teacher and a friend, giving me the support I needed in the last year of my thesis. I can openly say that without him, I wouldn't have finished so soon. Working side by side with a person whose love for life and for research is so contagious has been a gift that I will always treasure. I just hope I can live up to his expectations.

I thank all my co-authors for their patience and for sharing with me many entertaining hours while working on our research.

I am also grateful to all the members of the jury, Paul Belleflamme, Luca Lambertini and Jacques Thisse, for their insightful and extremely useful comments.

A thesis also needs financial support, and I am indebted to the Fundação para a Ciência e Tecnologia, who provided me with a scholarship for my first years of PhD, and to Fundação Gulbenkian and CORE through the ARC project who co-financed my last year of the PhD.

I dedicate this thesis to all the people that have believed in me and that have supported me through these years, though, in particular I dedicate it to a very special person. Othman has been the first one to encourage me to write a thesis and go abroad.

He has always believed in my capacities, while being my teacher, my friend, and as he liked to say, my grandfather "honoris causa". Unfortunately he passed away before seeing this project accomplished, but I like to believe that he is here today, sharing my same happiness.

1

INTRODUCTION

The primary aim of this thesis is to contribute to the analysis of strategic complementarities in industrial organization problems. The thesis can be divided in two parts. The first part (chapter 2) deals with the interplay between strategic complementarities and endogenous heterogeneity. The second part (chapters 3 to 6) concentrates on strategic complementarities under the form of network effects in dynamic setups. While strategic complementarities arise as the *fil rouge* of the thesis, part 1 and 2 constitute two separate lines of research and should be motivated separately. This prologue serves the purpose of 1) introducing the reader to the concept of strategic complementarities, 2) relating strategic complementarities to network effects, 3) motivating the following chapters and finally 4) highlighting the main contributions of the thesis. I will thus start by presenting the context in which the study of strategic complementarities has been developed and then proceed in explaining the relationship between the first and the second part of the thesis and briefly sketching each chapter.

Recent developments in economic theory have focused on the so-called theory of supermodular games, or games with strategic complements. The reason is that complementarities arise naturally in many social and economics situations, ranging from competition issues in industrial organization to coordination failures in macroeconomics and finance and to equilibrium selection problems in game theory. The notion of complementarity is due to Edgeworth (1925, in *Papers Relating to Political Economy*, pp. 117-118) and refers to the case in which the marginal value of an action is increasing in the level of another action or variable. Bulow et al. (1985, in *Journal of Political*

Economy pp. 488-511) extended this definition to include strategic behavior: we say that the actions of two agents are strategic complements if the marginal profitability of an agent increases with the actions of the rival. Examples of this kind of interaction abound. In industrial organization firms selling complementary products would like to increase their prices if the rivals do so; in R&D investment, firms are more prone to invest if opponents invest more as well, in order to profit from the same probability of innovation; in fashion, consumers tend to follow others in their choices of outfits, cars or even vacation sites. Bank runs, currency attacks, Internet, technology adoption are all examples of human interaction in which strategic complementarities arise.

The interest in this particular kind of strategic interaction has grown considerably with the development of the appropriate modeling tools for incorporating complementarities. In fact, many of the problems could not be dealt with standard optimization techniques due to the non-smooth payoffs and multiple equilibria that usually arise in the framework of complementarities. The correct methodology to deal with strategic complementarities (and also strategic substitutability) is the theory of supermodular games. The monotonicity of the optimal strategies of agents with respect to other agents' actions is the cornerstone of this theory. Existence of pure strategy Nash equilibria in two-player games is guaranteed as long as players' reaction functions are increasing and there is a partial order in the strategy space, regardless of the non-smoothness of the payoff functions. Also, the problem of multiple equilibria is addressed by establishing an ordering of the equilibria and providing comparative statics with respect to the parameters of interest and the equilibria. When dealing with general action spaces, the theory of supermodular games is also suitable to study complementarities arising in dynamic settings or incomplete information settings.

This thesis pushes further the applicability of the theory of supermodular games to the analysis of symmetric games in which equilibria are asymmetric only. Although most economic thinking leans on the idea of symmetry among interacting agents there is no reason –besides simplicity- to disregard the real world heterogeneity. Moreover, heterogeneity is even a necessary condition for many economic activities to take place. Inspired by this finding, researchers from different economic and social fields have attempted to explain the emergence of asymmetries. We distinguish three different approaches to endogenous heterogeneity: the first one is based on coordination failures; the second on the idea that under certain conditions, the only stable equilibria are asymmetric and thus, symmetric equilibria can be disregarded; the third one -bred by the literature on business strategy- postulates that heterogeneity derives from the complex environment in which firms and individuals operate thus disabling them from correctly formulating strategies to their behavior.

Chapter 2 attempts to contribute to the debate generated by the different approaches along the lines of modern game theory. We analyze a priori symmetric two-player games that posses two properties: players' actions form strategic substitutes, and each player's payoff, though continuous, admits a nonconcavity along the diagonal in the action space. Strategic substitutability in this setting guarantees the existence of pure strategy Nash equilibrium due to the monotonicity of the optimal strategies mentioned above. Furthermore, the nonconcavity of the payoff function precludes symmetric equilibria. Hence, at any of the possibly multiple equilibria, a priori identical agents will necessarily choose different equilibrium actions. In our approach, endogenous heterogeneity arises from the interaction of fully informed rational agents, as a result of the nonsmoothness of the payoff function. We also consider two other related classes of games that always possess asymmetric, but never symmetric, pure-strategy equilibria although they are not of strategic substitutes. This suggests that the essence of the monotonicity is not as critical as the diagonal nonconcavity property in generating exclusively asymmetric outcomes. A second motivation for Chapter 2 is to generalize a class of two-player games that constitutes applications of our main propositions. These include two-stage business models with long-term investment decisions in the first stage and product market competition in the second stage.

The second part of the thesis also deals with strategic complementarities, however now assuming the specific form of network externalities. Payoffs are said to exhibit network externalities if they increase in the total number of users of the same commodity. When an individual chooses to join a network, the utility for others to join the same network also increases. Hence, the actions of joining a network are strategic complements among agents. The reason for this -apart from the possibility of directly interacting with other agents- is the fact that widely used commodities are more likely to be improved upon, to have better technical support and to be provided with more components. Take two of the most common examples: the satisfaction that a consumer derives from the usage of a telephone is higher, the higher is the number of other consumers who have access to telephone; in the same way, a user of a certain operative system obtains more returns the larger the number of users because of the gains in mobility, diversity of available software and so on.

As we can see from these examples, the network property is essentially dynamic in the sense that networks with consistently large number of consumers through time become more desirable and eventually might dominate the market in which they compete. Dynamics is therefore a crucial element in the study of market interactions in the presence of network effects. Chapters 3 to 6 concentrate on the study of dynamic aspects of network externalities.

In general, even in dynamic models of technology adoption, the formulation of the network effect most commonly used is the one where the utility of agents is increasing in the number of past adopters. From the moment in which a consumer joins a network, he adds to the utility of future users, without actually enjoying any more benefits from the network. In chapter 3 we contribute to the literature on technology adoption by considering agents who live for more than one period and receive payoffs from joining the network they have chosen in all periods of their lives. In this case, the action of each player is complementary to the actions of both past and future agents.

As stressed before, complementarities give rise to multiple equilibria, since it is in the best interest of the agents to coordinate. Nevertheless, in chapter 3, we show that a unique equilibrium exists in technology adoption, when agents have a forward-looking behavior and the value of the technology is stochastic. The setup is an overlapping generations game in which agents choose between two competing technologies. Each technology provides the consumers with an intrinsic benefit (so-called stand-alone value) and with a network value than can be divided in direct and indirect network effect. The indirect network effect is the benefit obtained from using a technology that has been historically chosen more often: it incorporates the improvements in technology due to the existence of initial users. The direct network effect results from the direct interaction between agents: for instance the possibility of communicating in the same language or the possibility of exchanging information with the same standard. We assume that the stand-alone value is stochastic, such that agents must consider the expected behavior of their future counterparts. Strategic complementarities allow us to develop the setup of chapter 3 within the class of monotone Bayesian games defined by Van Zandt and Vives (2003). Hence, it is possible to identify a unique equilibrium of the switching form in technology adoption. This equilibrium is characterized by the absence of lock in. The previous literature predicts that the dynamic adoption of technologies exhibiting network effects would end up by favoring one of the technologies that would dominate the market as time goes by. We show that this is not the case if the stand-alone values for the technology are stochastic; nevertheless, the equilibrium adoption path exhibits hystheresis due to strategic complementarities.

Chapter 3 does not address two important issues that arise naturally in dynamic models of network effects. The first one relates to the optimal pricing of technologies, and the second one is the introduction of new qualities in the market when it is already dominated by a quality with network externalities. Chapter 4 and 5 deal with both these issues, whereas chapter 6 concentrates on the optimal path of prices for a monopolist operating in a network industry.

Whenever a good possesses the property of providing its consumers with network effects, it creates an effect of 'brand loyalty'. Consumers must consider the benefits of joining the network, weighting them with the benefits of 'standing alone'. Often new qualities face difficulty to break into a market where strong network connections are present. Chapter 4 analyzes the incentive of a monopolist to introduce a new, higher quality in the market when he already produces a lower quality good and has constituted already a network of consumers. Chapter 5 analyses the same problem but from the point of view of an incumbent that is threatened by the entry of a higher quality brand in the market. In these chapters we assume that the network takes one period to be constituted. Thus, we rule out communication networks, like telephones or Internet, and consider only technologies whose networks are delayed, like the ones that rely on the word of mouth or that require learning. In this category we find the experience goods, i.e., goods whose value increases in the number of people who have tested them. We can also think about goods that can be improved through time thanks to the information conveyed by initial users. And examples abound: later versions of software are usually more reliable than earlier versions, vacation sites tend to offer better services the more people visit them, credit cards are accepted more often the higher the number of users of the card, and so on. This kind of externalities highlights more clearly the trade off between higher quality and network effects, which is the central theme of chapters 4 and 5. The main result is that quality improvement occurs if and only if the preference of the consumers for the network is not too high relative to the gain in quality between the old and the new version of the product. Moreover, quality improvement occurs more often under oligopoly than under monopoly. The reasons for this are 1) the cannibalization effect that the introduction of a new variety has in the monopolist's profit, and 2) the fact that preventing entry in the oligopoly is costly, whereas in monopoly it is costless. At the same time, chapters 4 and 5 discuss the problem of optimal pricing strategies in a dynamic setting with network effects. We conclude that in most cases, the optimal solution is to set very low initial prices in order to constitute the largest network in initial periods, and then to set a positive surplus extracting price in subsequent periods. This finding is compatible with the idea of penetrating pricing in business strategy and is also consistent with several empirical results: for instance, the free demo-versions of software, which are offered in the market, before the full-version is sold.

The question of introductory pricing is further analyzed in chapter 6, in which a monopolist sets the optimal pricing strategy for a finite number of periods during which he expects to be operating in the market. Besides the simplicity of the model, the contribution is that the path of optimal prices is increasing though time. In case of high network effects (or short horizon), the initial price -and only the initial one- is set to zero. All subsequent prices are chosen such that they extract the entire consumer surplus generated by the network (the so called bargain-then-rip-off strategy). The introductory price phenomenon is frequently observed (banks offer low rates for new users, phone companies set low rates for new adherents) and within this model we identify the precise condition under which this is the monopolist's optimal strategy. Our condition is related both to the intensity of the preferences for the network and the time horizon of the monopolist.

The thesis is organized as follows: Chapter 2 (*Endogenous asymmetry in strategic models*) is the result of joint work with Rabah Amir and Malgorzata Knauff; Chapter 3 (*Technology adoption with forward-looking agents*) is co-authored with Paolo Colla and Chapter 5 and 6 (respectively *Quality improvement and network externalities and A note on expanding demand and monopoly pricing*) are the product of cooperation with Jean J. Gabszewicz. Finally, Chapter 4 (*When does a monopoly improve its quality in a network industry?*) is a solo paper.

2

ENDOGENOUS HETEROGENEITY IN STRATEGIC MODELS : SYMMETRY-BREAKING VIA STRATEGIC SUBSTITUTES AND NONCONCAVITIES

2.1 Introduction

One of the most pervasive presumptions in modern economic analysis is the symmetric nature of interacting agents. While often intended solely as a simplifying assumption on *a priori* grounds, this presumption has also permeated economic thinking for a variety of other reasons. When considering noncooperative games, analysts often restrict attention to symmetric equilibrium points even when other asymmetric outcomes exist and may reflect more rationalizable or more pertinent behavior. In mechanism design or policy games, the social planner typically assumes identical treatment of identical agents, although global optimality might dictate otherwise. The design of various forms of joint ventures is also subject to a similar observation.

In most cases, the only justification beyond simplicity is what Schelling (1960) convincingly termed the focal nature of symmetric equilibrium outcomes. Indeed, it is widely recognized that inter-agent heterogeneity is often a critical dimension of several economic and social phenomena. From a positive perspective, heterogeneity is simply a necessary postulate to account for the simple fact that in the real world, one seldom observes identical agents, be it individuals, firms, industries or countries. In a similar vein, even from a normative standpoint, differences across interacting agents often constitute a necessary condition for many important economic activities such as trade or risk-sharing.

Understanding the origins and/or evolution of diversity across economic agents or disparities in economic performance across regions is increasingly perceived as a central goal of economic and social research in a number of different areas (see e.g. Geroski et al. (2003) and Ghemawat (1986)). Macroeconomists attempt to explain the causes of booms and recessions. Development economists wish to understand the forces behind poor and strong economic performances. Labor economists attempt to get a handle on discriminatory treatment of some groups of workers. Business strategists and industrial economists devote a lot of attention to the sources and sustainability of inter-firm heterogeneity within and across industries. Overall, much effort has been expanded with a view to explain "the diversity across space, time and groups" (Matsuyama, 2002).

In view of the diversity of economic research areas involved in this effort, it is not surprising that various conceptual and methodological approaches have been developed in connection with this complex task. While often tailored to a specific area, each of these approaches is broad in explanatory scope and has wide potential applicability. We now briefly review three of these general paradigms that share some relationship to the present paper.

The dominant approach, based on coordination failures, postulates a game with strategic complementarities and multiple Pareto-ranked pure-strategy Nash equilibrium points. Diversity is then synonymous with making different equilibrium selections, with the high-performing entity picking the Pareto-dominant equilibrium and the low-performing entity failing to do so. This argument is thus generally predicated on the presence of two identical and non-interacting economies, each operating under a different equilibrium out of the same equilibrium set. It may also be invoked to explain diversity across time within the same economy, with booms and recessions corresponding to operation under the Pareto dominant and inferior equilibria respectively. This literature includes as key studies Cooper and John (1988), Murphy, Schleifer and Vishny (1989), and is surveyed by Cooper (1999).

The coordination failures approach has been criticized for failing to offer any compelling argument for the diversity in equilibrium selections in the two-economy model or for the regime switch in the one-economy model. Matsuyama (2002) proposes a modification of the former model by creating an interactive link between the two sub-economies and allowing the two players to take two decisions, one in each sub-economy. Under some conditions on the larger game with *a priori* identical players, namely that the players' actions are pairwise strategic complements and each player's actions are substitutes due to a fixed total resource constraint, multiple equilibria arise with the property that the symmetric equilibrium is unstable while the two asymmetric equilibria are stable, both in the sense of Cournot dynamics. Endogenous heterogeneity in this approach is then predicated on the central postulate that only Cournot-stable equilibria are observable outcomes of this complex game, any of which would involve each agent taking different actions in the two *ex ante* identical sub-economies. Matsuyama (2000, 2002, 2004) coined the term "symmetry-breaking" to refer to this heterogeneity-generating process.

The third approach originates in the business strategy literature, and is often presented as part of a general critique of economic theory. With their traditional emphasis on investigating the workings of firms as complex organizations, strategy scholars have been particularly concerned with understanding the sources and dynamics of inter-firm heterogeneity along various functions and characteristics. In the dominant view, as articulated by Nelson and Winter (1982), firms operate in such highly complex and ever-changing environments that they entertain no hope of ever accumulating enough knowledge about their world to view it as a strategic game or formulate a precise game-theoretic strategy to guide their overall behavior. Rather, firms grope for economic performance via a heuristic learning process of trial and error and the continual updating of routines and rules of thumb eschewing optimization. In this evolutionary vision, heterogeneity is simply an inevitable outcome of this groping behavior, with firms ending up with different heuristic strategies and core capabilities to implement them. These "discretionary" differences can then be sustained over extended periods of time due to the presence of barriers to successful imitation generated by the differences in core competencies, and also by forces of path dependence in the evolution of firms' choices. This literature often criticizes economic theory for not adequately accounting for inter-firm differences, other than postulating them either as reflecting variations in initial conditions, or as exogenous consequences of the luck of a draw in stochastic models. This failure is attributed to the fact that economic theorists persist, as part of their excessive reliance on complete rationality, in "taking a firm's choice sets as obvious to it and the best choice similarly clear and obvious" (Nelson, 1991).[1]

[1] An interesting development over the last two decades is a strand of literature straddling the traditional boundaries between industrial organization and business strategy and addressing issues of interest to both fields, making them increasingly related: See Shapiro (1989), Rumelt, Schendel and Teece (1991), Roller and Sinclair-Desgagne (1996).

The present paper constitutes an attempt to contribute to this rich debate along standard lines of argument in applied game theory and industrial organization. Consider a two-player symmetric normal-form game characterized by two key properties: (a) actions form strategic substitutes, and (b) each player's payoff, though continuous, admits a key nonconcavity along the diagonal in action space, which results in a jump of the reaction correspondence across the 45^o line. Such a game always admits pure-strategy Nash equilibrium points due simply to the property of strategic substitutes. Furthermore, due to property (b), no such equilibrium could ever be symmetric. At any of the possibly multiple equilibria, which obviously occur in pairs due to the symmetry of the game, otherwise identical agents will necessarily take different equilibrium actions. While this description exactly fits the main result of the paper, we consider two other related classes of games that always possess asymmetric, but never symmetric, pure-strategy equilibria although they are, strictly speaking, not of strategic substitutes. This suggests that the latter property is not as critical as the diagonal nonconcavity property in generating exclusively asymmetric outcomes.

Since payoffs are continuous in actions in all three classes of games under consideration, these games will typically admit a symmetric mixed-strategy Nash equilibrium (Dasgupta and Maskin, 1986). As this would be the only focal equilibrium in the sense of Schelling (1960), it may reasonably be advanced as a plausible outcome of such a game. Nevertheless, in the actual realization of the equilibrium randomizations, the players will still end up playing different actions with high, if not full, probability. Hence, given a focus on explaining observed heterogeneity, this approach need not rule out mixed strategies *a priori*.

Towards the goal of generating endogenous heterogeneity, this approach is obviously closest in spirit to Matsuyama's symmetry-breaking explanation. By allowing for suitable discontinuities in the players' reaction curves, it dispenses with the need to interconnect two separate games in the somewhat complex (and subtle) manner proposed by Matsuyama. More importantly, it also provides a framework that is independent of the controversial argument of outright rejecting Cournot-unstable equilibria. Indeed, even when one ignores the focal nature derived from their symmetry, it is worthwhile to observe that these equilibria cannot be ruled out on account of any of the standard Nash equilibrium refinements, such as normal-form perfection or strategic stability

(Kohlberg and Mertens, 1986)[2]. Furthermore, in an experimental setting involving a symmetric two-player game with one unstable symmetric equilibrium and a pair of asymmetric equilibria, Cox and Walker (1998) found little support in the data for *any of the three equilibria.* This provocative finding suggests that while a Cournot-unstable equilibrium of a given game may be justifiably regarded as unobservable, it does not thereby follow that some Cournot-stable equilibrium of the same game will necessarily prevail and thus be observable. Rather, the presence of both Cournot-stable and unstable equilibria may engender a high level of indeterminacy, which may critically reduce the predictive power of the game.

Our findings may also be advanced as a rebuttal to the aforementioned criticism from the business strategy literature. Indeed, while sharing their motivation for understanding intra-industry heterogeneity, this approach underlies a general methodology for generating inter-firm differences out of a strategic game with fully rational and completely informed players. The contrast with the evolutionary explanation is rather striking. Instead of discretionary differences that inevitably arise out of the idiosyncratic heuristic response that each firm develops in isolation from other firms as a result of its multi-faceted operation in an extremely complex environment, we uncover strategic differences that arise out of a fully-fledged game-theoretic interaction amongst firms in a simple and completely known environment. We will return to this contrast in the specific context of an R&D game in a subsequent section.

The present paper may also be motivated in relation to various broad strands of literature in industrial economics dealing in some way with strategic endogenous heterogeneity along lines similar to ours here. The first literature that comes to mind is concerned with product differentiation. In a myriad of two-stage games where each firm chooses a quality level or a horizontal characteristic in the first stage, and then a price for its product in the second stage, endogenous heterogeneity naturally emerges out of the firms' perception that identical choices in the first stage will lead to zero profits in the second stage Bertrand competition due to the resulting homogeneity of the products. See in particular Gabszewics and Thisse (1979), Shaked and Sutton (1982) for vertical product differentiation and Tabuchi and Thisse (1995) for horizontal product differentiation with a non uniform density. The present paper will direcly generalize Motta (1993) and Aoki and Prusa (1996).

[2] Indeed, they are typically strict Nash equilibria (in the sense that a unilateral deviation will lead to a strict loss for the deviator), and thus would survive any of the well-known refinements.

The second, extensive literature deals with infinite-horizon industry dynamics allowing for entry and exit. One class of models, exemplified by Jovanovic (1982), postulates perfectly competitive firms for which differences emerge due to *exogenous* idiosyncratic technology shocks. Another class is formed by studies that do generate endogenous heterogeneity in long run dynamics by considering firms that invest in capacity expansion (e.g. Besanko and Doraszelski, 2002) or R&D (e.g. Doraszelski and Satterthwaite, 2004). Simpler two-stage models with similar flavor but without entry and exit also generate endogenous differences amongst competing firms: Reynolds and Wilson (2000) and Maggi (1996) for capacity expansion and Amir and Wooders (1999, 2000) for R&D.

There are several other studies in various areas of industrial organization where endogenous heterogeneity emerges in a strategic setting. Hermalin (1994) deals with a two-stage game where firms' choices of managerial structures take place before market competition. Mills (1996) and Amir (2000) deal with R&D games giving rise to equilibrium outcomes with maximal heterogeneity only, i.e. full R&D by one firm and no R&D by the rival.[3] In public economics Mintz and Tulkens (1986) exhibit asymmetric tax rates for identical member states.

As a second motivation, the present paper is an attempt to develop a unifying approach to understanding symmetry-breaking mechanisms in general classes of two-player games, encompassing many of the cited studies. These two-stage models share two key features that are critical for the symmetry-breaking arguments they present. The first is a fundamental nonconcavity in the payoffs, which may be confined to the diagonal in action space or hold globally, and the second is some form of strategic substitutes in first-period actions, possibly of an abstract sort (more on this in Section 2.4). While there is quite some variation in the precise manner versions of these two features are present and interact across all the models, we will be able to capture most of them in three separate general results, which though quite distinct at first sight, nonetheless bear some definite relationship at an abstract level.

The paper is organized as follows. Section 2.2 contains the overall set-up. Section 2.3 provides the results on the exclusive existence of asymmetric equilibria for submodular payoff functions. Section 2.4 deals with nonsubmodular payoff functions and Section 2.5 presents the results for games with convex payoffs. Section 2.6 discusses an extension

[3] Another strand of literature, not directly related to our setting, deals with endogenous heterogeneity arising out of hybrid models of joint ventures where firms make a cooperative decision in the first stage followed by product market competition in the second stage: See e.g. Salant and Shaffer (1998,1999) and Long and Soubeyran (2001).

to ordinal complementarity and substitutability conditions. Each section provides a summary of the relevant applications the results pertain to. The appendix provides a brief overview of the supermodularity notions and results.

2.2 Setup

This section lays out the general notation for use throughout the paper. The noonco-operative game described below may be a simple one-shot game or it may represent the payoffs of a two-stage game as a function of the first period actions, where the unique second stage pure-strategy equilibrium has been substituted in. In the latter case, which actually covers most of the applications of this paper, we obviously restrict consideration to subgame-perfect equilibria and analyze the resulting one-shot game.

Consider a two-player normal form game Γ given by the tuple (X, Y, F, G). Let X and Y be the action sets of player 1 and 2 respectively, such that $X = Y = [0, c] \subset R$. The maps F and $G : X \times Y \rightarrow R$ are the payoff functions of players 1 and 2 respectively and F can be expressed as:

$$F(x,y) = \begin{cases} U(x,y), x \geq y \\ L(x,y), x < y \end{cases} \tag{2.1}$$

By symmetry of the game Γ, G can be expressed as:

$$G(x,y) = \begin{cases} L(y,x), x \geq y \\ U(y,x), x < y \end{cases} \tag{2.2}$$

Observe that, somewhat contrary to standard practice, the first argument of F is the action of player 1 while the first argument of G is the action of player 2. It is useful to define the following sets:

$$\Delta_U = \{(x,y) \in R^2 : x \geq y\} \text{ and } \Delta_L = \{(x,y) \in R^2 : x \leq y\}.$$

It will be assumed throughout the paper that U, L, F and G are jointly continuous functions of the two actions. Define the best response correspondences (reaction curves) for players 1 and 2 respectively as $r_1(y) = \arg\max\{F(x,y) : x \in [0,c]\}$ and $r_2(x) = \arg\max\{G(y,x) : y \in [0,c]\}$.

As usual, a pure strategy Nash equilibrium (or PSNE for short), $(x^*, y^*) \in [0, c]^2$ is said to be symmetric if $x^* = y^*$, and asymmetric otherwise. It follows from the symmetry of the game that if (x^*, y^*) is a PSNE, (y^*, x^*) is also a PSNE.

Each of the next three sections investigates a separate class of normal-form symmetric games that always possess asymmetric Nash equilibria and no symmetric Nash equilibria. For each of the three classes, we provide a general result establishing both the existence and the inexistence conclusions and an illustration based on previous studies where a special case of the result was derived in a specific setting.

The definitions and main results from the theory of supermodular games used in this paper are reviewed in the appendix in a very simple way, which is sufficient for the purposes of this paper.

2.3 Endogenous heterogeneity with strategic substitutes

In this section, we consider a two-player symmetric normal-form game characterized by two key properties. The first is that actions form strategic substitutes. This means that an increase in one player's strategy lowers the other player's marginal returns to increasing his own strategy. As a result, players respond optimally to an increase of the opponent's choice with a decrease of their own variable. In other words, the best reply correspondences are downward-sloping and a pure-strategy Nash equilibrium exists (see, Vives, 1990 and Milgrom and Roberts, 1990).

The second key property is that each player's payoff, though jointly continuous in the two actions, admits a fundamental nonconcavity along the 45^o line, giving rise to a canyon shape along the diagonal. A key consequence of this feature is that a player would never optimally respond to an action of the rival by playing that same action.

Taken together, these two properties imply that each best reply is a decreasing correspondence with a (downward) jump over the 45^o line.[4] Hence, no PSNE could ever be symmetric. At any of the possibly multiple equilibria, which obviously occur in pairs due to the symmetry of the game, *ex ante* identical agents will necessarily take different equilibrium actions.

[4]In this paper, we will say that a function $f : R \to R$ is increasing (strictly increasing) if $x' > x$ implies $f(x') \geq (>) f(x)$. A correspondence is increasing if its maximal and minimal selections are increasing functions (as in the conclusion of Topkis's Monotonicity Theorem).

While all three results presented in this paper share this same flavor, the main result in terms of the generality of the assumptions and thus of the scope of applicability is this section's.

2.3.1 The results

Different subsets of the following assumptions will be needed for our conclusions below. The notation is as laid out in Section 2.2.[5] A full discussion of the assumptions and results is presented at the end of the section. Most of the proofs can be found in the appendix.

A1 U, L are submodular

A2 $U_1(x,x) > L_1(x,x), \forall x \in (0,c)$

A3 $U_1(0,0) > 0, L_1(c,c) < 0$

A1 says that on either side of the diagonal, but not necessarily globally, each player's marginal returns to increasing his action decrease with the rival's action. A2 holds that each player's payoff, though globally continuous in the two actions, has a kink along the diagonal in the shape of a "valley". The role of A3 is simply to rule out PSNEs at $(0,0)$ or (c,c).

These assumptions form a sufficiently general framework to encompass many of the studies mentioned in Section 5.1 as illustrated below. Furthermore, all three assumptions are easy to check in a particular model.

Theorem 1 *Assume that $A1-A3$ hold. Then the game Γ is of strategic substitutes, has at least one pair of asymmetric PSNEs and no symmetric PSNEs.*

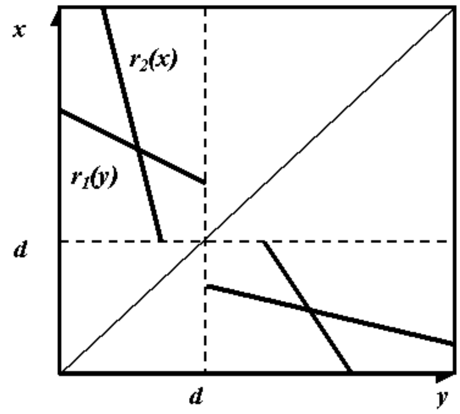
The idea of the proof is that overall submodularity of the payoff function is inherited from the submodularity of its components U and L in the presence of assumption $A2$. We know from Topkis's Monotonicity Theorem that global submodularity of the payoff function implies globally decreasing best replies. Assumptions $A2-A3$ imply that the best replies have a downward jump that crosses over the diagonal. This situation is depicted in figure 2.1.[6] (We caution the reader that the continuity of the reaction

[5] In addition, throughout the paper, partial derivatives are denoted by a subindex corresponding to the relevant variable, i.e. $U_1(x,y) = \frac{\partial U(x,y)}{\partial x}$ and $U_2(x,y) = \frac{\partial U(x,y)}{\partial y}$.

[6] Notice that unusually x is the variable in the vertical axis. This corresponds to analyzing the game from the point of view of player 1 that chooses x as a response to y. We maintain this convention throughout.

curves in each triangle over and below the diagonal is only there for the sake of a clearer figure. It needs not hold under assumptions A1- A3). The first result gives us existence of equilibrium via the strategic substitutes property, and the second precludes symmetric equilibria. The complete proof can be found in the appendix.

FIGURE 2.1. Decreasing reaction curves have a jump along the diagonal, and there is no symmetric equilibrium.



Theorem 1 does not rule out the existence of multiple pairs of PSNEs. Indeed, the two reaction curves may intersect several times above and below the diagonal. In case of multiple pairs of PSNEs, there will typically be co-existence of pairs of Cournot-stable and pairs of Cournot-unstable PSNEs. Nevertheless, Theorem 1 does imply that all of these PSNEs are asymmetric. Hence symmetry-breaking in this context does not rely on the rejection of Cournot-unstable symmetric PSNEs. In the same vein, this type of symmetry-breaking is not at odds with Schelling's (1960) notion of focalness of PSNEs.

The next result adds further restrictions of a general nature on the payoff components of our game that lead to a unique pair of PSNEs, which are then necessarily Cournot-stable.[7] In this case, symmetry-breaking is coupled with more predictive power of the game, although the selection of one PSNE from the pair still remains indeterminate, as is standard in symmetric settings.[8]

[7] It is worthwhile to point out here that our results indicate an even total number of PSNEs, in apparent conflict with the well-known odd number results. The explanation is that the latter results are based on degree theory and require continuity of the best-response form. Given our systematic and robust jump across the diagonal, our findings are actually consistent with the odd number result in a generic sense.

[8] As mentioned in the Section 5.1, assumptions $A1-A3$ imply that our game admits a *symmetric* mixed-strategy Nash equilibrium. However, in the actual realization of such an equilibrium, the two players will be heterogeneous with high, if not full, probability.

Theorem 2 *Assume that U and L are twice continuously differentiable and that the following holds:*

$$U_{11}(r(z), z) - U_{12}(r(z), z) \geq 0 \tag{2.3}$$

$$L_{11}(r(z), z) - L_{12}(r(z), z) \geq 0 \tag{2.4}$$

then there is exactly one pair of PSNEs.

The next result is devoted to comparing the two asymmetric PSNEs from any given pair from the point of view of the players' welfare. Given any pair of asymmetric PSNEs, it is often of interest to determine circumstances where a given equilibrium secures better payoffs for a player. In other words, under what conditions would each player prefer the PSNE where he is the high or the low-activity player? To this end, we need to impose a condition of monotonicity on the payoff function of the player in question along his opponent's best reply as stated in the following result, which lays out conditions for player 1 (say) to prefer the PSNE where he is the low-activity player.[9]

Theorem 3 *Let (x^*, y^*) and (y^*, x^*) be equilibria in Δ_U and Δ_L, respectively. Without loss of generality, let $x^* > y^*$. If $A1 - A3$ hold and moreover $U(r_1(y), y)$ and $L(r_1(y), y)$ are increasing in $y \in [0, c]$ then $F(x^*, y^*) \geq F(y^*, x^*)$.*

There is a dual statement giving conditions under which each player would prefer the high-activity equilibrium, given any pair of PSNEs. Being obvious from Theorem 3, it is omitted for the sake of brevity.

2.3.2 Applications

In this section we present examples of economic models that constitute special cases of the general framework developed above. While the assumptions validating Theorem 1 might at first appear somewhat special, they are satisfied in several *a priori* unrelated studies that have established endogenous heterogeneity in strategic settings. There are also some studies where asymmetric equilibria are produced via a mechanism similar to our Theorem 1, without being a special case in a formal sense. Going over some of these examples illustrates the unifying character of our results and allows us to

[9] This monotonicity assumption is clearly more general than assuming that each player's payoff is increasing in the rival's action. For an example illustrating this point, see von Stengel (2003).

provide some contextual interpretations of endogenous heterogeneity, or our version of symmetry-breaking.

2.3.2.1 R&D investment

The first example is based on the model by Amir and Wooders (2000). Two *a priori* identical firms with initial unit cost c are engaged in a two stage game of R&D investment and production. In the first stage, autonomous cost reductions x and y for firms 1 and 2, respectively, are chosen. The novel feature of this study is that spillovers are postulated to flow only from the more R&D active firm to the rival, but not vice versa. The effective (post-spillover) cost reductions X and Y when $x \geq y$ are given by:

$$X = x \text{ and } Y = \begin{cases} x \text{ with probability } \beta \\ y \text{ with probability } 1-\beta \end{cases} \tag{2.5}$$

Second stage product market competition, be it Cournot or Bertrand, is assumed to have a unique PSNE with equilibrium payoffs given by $\Pi : [0,c]^2 \rightarrow R$.[10] $\Pi(x,y)$ is the payoff of the firm whose unit cost is the first argument. $f : [0,c] \rightarrow R$ is a known R&D cost schedule. Amir and Wooders (2000) assume the following:

C1 Π and f are twice continuously differentiable

C2 Π is strictly submodular and $\Pi_1(x,y) < 0$ and $\Pi_2(x,y) > 0$

C3 $\Pi(x,x) < \Pi(y,y)$ if $x > y$

C4 $|\Pi_1(x,x)| > |\Pi_2(x,x)|, \forall x \in [0,c]$

C5 $f'(x) \geq 0$ and $f(0) \geq 0$

C6 $f\prime(0) < -\beta\Pi_2(c,c) - \Pi 1(c,c)$ and $f\prime(c) > -(1-\beta)\Pi_1(0,0)$

The overall payoff of firm $1, F(x,y)$, defined as in (2.1), is given by the difference between its second stage profit and first stage R&D cost. The payoff of firm 2 is $G(y,x)$ by symmetry.

$$U(x,y) = \beta\Pi(c-x, c-x) + (1-\beta)\Pi(c-x, c-y) - f(x) \tag{2.6}$$

$$L(x,y) = \beta\Pi(c-y, c-y) + (1-\beta)\Pi(c-x, c-y) - f(x) \tag{2.7}$$

[10] This is a standard assumption in the literature.

We can easily check that assumptions $A1, A2$ and $A3$ indeed hold in order to apply Theorem 1. U and L are continuous and differentiable because they result from the sum of continuous and differentiable functions. $U(x,x) = L(x,x), \forall x \in [0,c]$, so F and G are continuous. $A1$ can be checked by using the cross-partial test and the fact that $\Pi(x,y)$ is submodular (Assumption $C1$). Also, Using $C2$, and

$$U_1(x,x) = -[\Pi_1(c-x,c-x) + \beta\Pi_2(c-x,c-x)] - f'(x) \tag{2.8}$$

$$L_1(x,x) = -(1-\beta)\Pi_1(c-x,c-x) - f'(x) \tag{2.9}$$

we obtain that $U_1(x,x) > L_1(x,x)$, therefore $A2$ is verified.

From Theorem 1, we can conclude that payoff functions for both players are submodular and thus reaction curves are downward sloping and there exists at least one PSNE. Moreover the reaction curves do not intersect the diagonal and by $C6$, $U_1(0,0) > 0$ and $L_1(c,c) < 0$, so there is no symmetric equilibrium in $x \in [0,c]$. The uniqueness of a pair of asymmetric PSNEs is shown by imposing conditions that secure that the reaction curves are contractions, so that Theorem 2 can be applied. Similarly, extra conditions are needed to apply Theorem 3 and Amir and Wooders conclude that player 1 prefers the equilibrium in Δ_U (for details, see Amir and Wooders, 2000).

This model provides a good opportunity for a typical economic interpretation of strategic endogenous heterogeneity in a context that is of particular interest to business strategy scholars. Indeed that field typically attaches a great deal of importance to the innovation process and to its central role in dynamic competition. The key driving force behind asymmetric equilibrium outcomes here is the one-way nature of the spillover process. A firm will always react by performing either less R&D than its rival knowing that it may free ride on the difference in R&D levels, or, in case the rival's R&D is simply too low, by overtaking it. In this vision, firms will endogenously settle into R&D innovator and imitator roles simply as a reflection of the nature of the R&D spillover process. This critical difference arises as a consequence of strategic thinking in a fully interactive setting: though facing equal and known opportunities in all respects, firms emerge as fundamentally different in all possible equilibria of a robust and general model. In strategic settings, there may simply be no single "best choice" (to paraphrase Nelson, 1991) for all *ex ante* identical firms facing the same available choices, simply because one firm's choice has a direct influence on what becomes best for its rivals.

This difference in one key component of firms' overall strategies will then be a causal factor, through natural complementarity-reinforcing developments, for heterogeneity in other aspects of firms' strategies, including in particular firm size and organization (see Amir and Wooders, 1999 for details). This perspective stands in sharp contrast to the explanation for inter-firm differences characterized by idiosyncratic groping behavior on the part of firms and weak interaction amongst them in a world of high uncertainty and complexity, as often envisioned in the strategy literature.

2.3.2.2 Provision of Information

The second example deals with the provision of information in Bertrand oligopoly, see Ireland (1993). Two *a priori* symmetric firms produce a homogeneous product and play a two stage game. In the first stage, each firm sets the level of its product information and in the second stage they compete in prices. Information regards only the existence of the product. Consumers may obtain costless information about prices of products that they know to exist. The number of consumers is normalized to 1. The variables x and y are the proportions of consumers who know about product 1 and 2 respectively. Each consumer is not willing to pay a unit price higher than 1. Firm 1's sales are given by:

$$Q_1 = \begin{cases} x \text{ if } p_1 < p_2 \\ x - \frac{xy}{2} \text{ if } p_1 = p_2 \\ x - xy \text{ if } p_1 > p_2 \end{cases} \tag{2.10}$$

For $x = y = 1$ the Bertrand oligopoly has a pure strategy Nash equilibrium at $p_1 = p_2 = 0$. If information is not full (x or y or both are less than 1), no pure strategy Nash equilibrium exists. There exists a mixed strategy Nash equilibrium given by the distribution function $G_i\left(p_i\right)$ that has the following form respectively for firm 1 and 2 .

$$G_1(p) = \begin{cases} \frac{1-(1-x)/p}{x} \text{ if } 1-x \leq p \leq 1 \\ 0 \text{ if } 0 \leq p \leq 1-x \end{cases} \qquad G_2(p) = \begin{cases} \frac{1-(1-x)/p}{y} \text{ if } 1-x \leq p \leq 1 \\ 0 \text{ if } 0 \leq p \leq 1-x \end{cases} \tag{2.11}$$

The overall payoff for (say) firm 1 in the game, upon substituting the second stage equilibrium payoffs, is given by $F(x,y) = E_p\left(p_1 Q_1\right)$ or

$$F\left(x,y\right) = \begin{cases} U(x,y) = x\left(1-y\right) \text{ if } x \geq y \\ L\left(x,y\right) = x\left(1-x\right), \text{ if } x \leq y \end{cases} \tag{2.12}$$

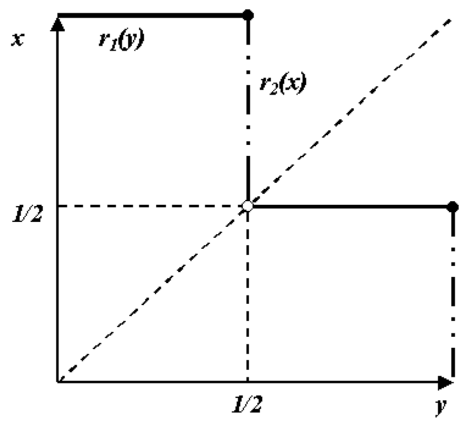
It is trivial to show that assumptions $A1, A2$ and $A3$ are verified in this example. There exists an equilibrium and this equilibrium cannot be symmetric for $p \in (0,1)$. Moreover $U_1(0,0) > 0$ and $L_1(1,1) < 0$. So we can conclude from Theorem 1 that no symmetric equilibria exist for $p \in [0,1]$.

The number of equilibria may be obtained by finding the explicit form of the reaction curves.

$$r_1(y) = \begin{cases} 1 \text{ if } y \leq 1/2 \\ 1/2 \text{ if } y > 1/2 \end{cases} \qquad r_2(x) = \begin{cases} 1 \text{ if } x \leq 1/2 \\ 1/2 \text{ if } x > 1/2 \end{cases} \tag{2.13}$$

Figure 2.2 illustrates that there exists exactly one pair of pure strategy Nash equilibria, namely $(1, \frac{1}{2})$ and $(\frac{1}{2}, 1)$.

FIGURE 2.2. Reaction curves are constant except for a jump down, which precludes symmetric equilibrium.



We can compare equilibria from the point of view of player 1, using the dual to the Theorem 3. The payoff function of player 1 when $x \leq y$, is constant along his best reply, i.e. $L_2(r_1(y), y) = 0$. Therefore we conclude that player 1 prefers the equilibrium where he is more active.

2.4 Endogenous heterogeneity without monotonic best replies

In the previous section we discussed symmetry breaking via strategic substitutes and nonconcavity of the payoff function. In this section we extend the analysis from the previous section to encompass other forms of strategic interaction. The main difference is that agents' strategies form (partially) strategic complements. This occurs when a more aggressive strategy from one agent rises the other player's marginal returns to increasing his own strategy. Consequently, an increase of the opponent's choice is responded to by an increase of own choice variable. This property implies that best-replies are increasing in own action. As stated above, strategic complementarities are partial in the sense that they are not observed overall. Reaction curves are piecewise increasing, however due to a symmetry breaking nonconcavity in the payoff function, they possess a jump down across the diagonal. The common aspect to the whole analysis is the canyon shape of the agents' payoff functions along the 45^o line. Once again, a player would never optimally respond to an action of the rival by playing that same action. Whereas in the previous section the submodularity of the payoff function (or alternatively the strategic substitutability) was a global feature, in this section we cannot state global supermodularity (or strategic complementarity). The main consequence is that we cannot guarantee without further assumptions the existence of a PSNE. Nevertheless, when it exists, it will never be symmetric.

When strategies are strategic complements, there is no need to assume the quasiconcavity of the payoff function in order to guarantee the existence of a PSNE. The reason is that existence might be based on the fact that reaction curves are increasing and continuity plays no role. Likewise, when payoff functions are quasiconcave supermodularity is not crucial for arguing the existence of a PSNE. When player's payoff function is partially quasiconcave and possesses a nonconcavity along the diagonal, we obtain the same type of symmetry-breaking already discussed. In this case, reaction curves are continuous (not necessarily increasing), except for the jump across the diagonal. Existence of PSNE can be assured (as before) via added conditions, however it is clear that no PSNE can be symmetric.

We provide now the main results of this section, first for the case in which strategic complementarities are present and then for the case of partially quasiconcave payoff functions.

2.4.1 The results

In this section we analyze the case in which the components of the payoff function, U and L, are not submodular. We first consider the case in which U and L are supermodular and then the case where U and L do not have this property, which is replaced by quasiconcavity.

Consider the following assumptions:

B1 U and L are supermodular.

B2 U and L are differentiable and $U_1(x,x) > L_1(x,x), \forall x \in (0,c)$

Define $\overline{r}_1(y) = \arg\max\{U(x,y) : (x,y) \in \Delta_U\}$ and $\underline{r}_1(y) = \arg\max\{L(x,y) : (x,y) \in \Delta_L\}$

B3 $U_2(\overline{r}_1(y), y) < L_2(\underline{r}_1(y), y), \forall y \in Y$

B4 $U_1(0,0) > 0, L_1(c,c) < 0$

B5 $U_1(x,0) > 0$, $\forall x < d$ and $L_1(x,c) < 0$, $\forall x > d$, for $d : r_1(d-\varepsilon) > d > r_1(d+\varepsilon)$

$B1$ says that on either side of the diagonal, but not necessarily globally, each player's marginal returns to increasing his action increase with the rival's action. $B2$ holds that each player's payoff, though globally continuous in the two actions, has a valley-like shape along the diagonal. $B3$ is responsible for the uniqueness of the jump in the reaction curves. $B4$ excludes the existence of equilibria in $(0,0)$ and (c,c). Finally, $B5$ restricts the reaction curves to a compact subset of the action space which enables us to prove the existence of an asymmetric PSNE.

These assumptions form a sufficiently general framework to encompass many of the studies mentioned in Section 5.1 as illustrated below. Furthermore, all assumptions are possible to check in a particular model.

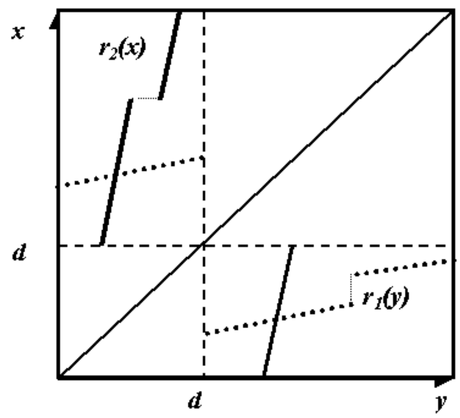
Lemma 1 *If $B1-B4$ hold, then there exists exactly one $d \in [0,c]$ such that $r_1(d-\varepsilon) > d > r_1(d+\varepsilon)$, $\forall \varepsilon > 0$.*

This Lemma guarantees that the reaction curves possess a jump and that this jump is unique. The flavor of the proof is the following: given supermodularity of U and L, we know that reaction curves are increasing in Δ_U and in Δ_L. Assumption $B2$ implies that there is no interior symmetric equilibrium due to the presence of a canyon along the diagonal. Furthermore, assumption $B4$ rules out symmetric equilibrium on the boundary. This is equivalent to saying that the reaction curves do not intersect the diagonal on the whole strategy space. Hence, there must be a jump across the diagonal. Finally assumption $B3$, that entails an idea of monotonicity of maxima along y, guarantees that in case a jump occurs, it is unique. From this Lemma we can conclude that the reaction curves possess a jump across the diagonal in a point d and that once it occurs, the reaction curves never jump back again. The point d is useful to define subsets of the strategy space, which are necessary in the next theorem.

The following theorem implies that no symmetric PSNE exists for a game where assumptions $B1-B5$ hold and that a PSNE exists. From these premises we can conclude that there are only asymmetric PSNEs.

Theorem 4 *Assume that $B1-B5$ hold, then there exist at least one pair of asymmetric PSNEs and no symmetric one.*

FIGURE 2.3. Reaction curves are partially increasing, but posses an unique jump down, which precludes any symmetric equilibrium.



The intuition behind this result is the following: from Lemma 1 we have that the reaction curves are partially increasing (as they increase in each side of the diagonal) and possess a unique downward jump at point d. Since reaction curves are not overall increasing we cannot guarantee without further assumptions the existence of a PSNE. Introducing assumption $B5$ guarantees that the reaction functions are well defined in $R = \{(x, y) : d \leq x \leq c, 0 \leq y \leq d\}, R \subset \Delta_U$ and $R' = \{(x, y) : 0 \leq x \leq d, d \leq y \leq c\}, R' \subset \Delta_L$, in the sense that they are completely contained in these compact subsets of the strategy space. We obtain, hence, increasing maps in compact sets and Tarski's Fixed Point Theorem can be applied to show that within these sets a PSNE exists.

Note that we can reorder one player 's action set in a nonstandard way to obtain a submodular game. Consider action space of (say) player 1 as $(d, 0] \cup [c, d)$, ordered from left to right. The other player's action space order remains without change. It can be verified, that the game satisfying $B1-B5$ then becomes a game of strategic substitutes.

Now consider another property of U and L.

B1' U and L are quasi-concave.

The assumption $B1$ can be replaced by the assumption $B1'$ which guarantees that player's best replies are continuous (not overall) and still the result holds. In other words, the supermodularity of U and L is not necessary (even though often observed in applications) for the existence of only asymmetric PSNE in this framework. In particular, assuming that U and L are quasiconcave implies that the reaction curves are partially continuous (even though not monotone) and thus, as long as the unique jump across the diagonal exists and the reaction curves are completely contained in compact subsets of the domain, we can show the existence of asymmetric PSNE. Since monotonicity cannot be guaranteed anymore Tarski's Fixed Point Theorem must be replaced by Brouwer's Fixed Point Theorem in showing the existence of PSNE.

Lemma 2 *If $B1'$ and $B2-B4$ hold, then there exists exactly one $d \in [0,c]$ such that $\forall \varepsilon > 0$: $r_1(d-\varepsilon) > d > r_1(d+\varepsilon)$,*

Theorem 5 *Assume that $B1'$ and $B2-B5$ hold then there exist at least one pair of asymmetric PSNEs and no symmetric one.*

The proofs of these results follow the same reasoning as the proofs of the precedent ones, however, existence of PSNE is now guaranteed through the application of Brouwer's Fixed Point Theorem (which is suitable given that payoff functions are quasiconcave).

Uniqueness of a pair of equilibria can be shown if reaction functions are contractions by using Banach's Fixed Point Theorem as in the previous section.

A comparison of equilibria when payoff functions are partially supermodular (or quasiconcave) can be done in the same spirit of Theorem 3. Instead of assuming $A1-A4$ we assume $B1-B5$ ($B1'-B5$). In the proof, the second inequality follows now from the fact that $x^* \geq d$ and Assumption $B3$.[11]

Adopting the results of Echenique (2004), wherein the order on the action spaces is not exogenously given would enlarge the scope of supermodular games. In particular,

[11] It is easy to see that Assumption $B3$ means that for $z > d$, $U(r_1(z), z) < L(r_1(z), z)$.

since our game here has at least two PSNEs, one can always find a partial order such that it becomes a supermodular game (Echenique, 2004, Theorem 5).

2.4.2 Applications: Quality Investment

We illustrate the results of this section with two papers dealing with quality investment problems. The first paper we analyze is Aoki and Prusa (1996). In this paper, two identical firms produce products differentiated by quality, in a two-stage setting. In the first stage firms 1 and 2 decide the level of quality investment $x \in [0, c]$ and $y \in [0, c]$ respectively, and in the second stage they simultaneously announce prices.[12]

Consumers are diversified in their willingness to pay for quality. Production cost is assumed to be 0 and firm 1 (firm 2) incurs a cost of quality investment $f(x) = kx^2$ $\left[f(y) = ky^2\right]$, $k > 0$. A detailed study of the second stage equilibrium of this game can be found in Gabszewicz and Thisse (1979) and Shaked and Sutton (1982). The subgame perfect equilibrium of the whole game can be obtained by backward induction and first stage overall payoffs for firm 1 are given as follows:

$$F(x, y) = \begin{cases} U(x, y) = \frac{4x^2(x-y)}{(4x-y)^2} - kx^2 \text{ if } x \geq y \\ L(x, y) = \frac{yx(y-x)}{(4y-x)^2} - kx^2 \text{ if } x < y \end{cases} \tag{2.14}$$

Now we can check whether assumptions of Theorem 4 are fulfilled. Given that U and L are differentiable we can check supermodularity (Assumption $B1$) by recurring to Topkis's Characterization Theorem. Consider that $U_1(x, x) = \frac{4}{9} - 2x$ and $L_1(x, x) = -\frac{1}{9} - 2x$, so $U_1(x, x) > L_1(x, x)$ which verifies $B2$. To check $B3$ we compute:

$$U_2(\overline{r}_1(y), y) = -4\overline{r}_1^2(y)\frac{2\overline{r}_1(y)+y}{(4\overline{r}_1(y)-y)^3} \quad L_2(\underline{r}_1(y), y) = -\underline{r}_1^2(y)\frac{2y+\underline{r}_1(y)}{(-4y+\underline{r}_1(y))^3} \tag{2.15}$$

Since $\overline{r}_1(y) > y > \underline{r}_1(y)$ the condition $B3$ holds. We know, hence, that the reaction curves are upward sloping except for a downward jump at a point d. Since $U(0, 0)$ is not defined we cannot check the first part of $B4$ directly, we must compute $\lim_{x\to 0} U_1(x, x)$. We obtain $\lim_{x\to 0} U_1(x, x) = \frac{4}{9} > 0$ which means that increasing quality investment at point $(0, 0)$ is profitable for both firms. Point $(0, 0)$ is thus ruled out as a pure strategy Nash equilibrium of the game. As for the second part of $B4$, it is easily verified by

[12] Aoki and Prusa (1996) consider unlimited quality investments. We impose the upper limit c, arbitrarily big, such that the strategy spaces are compact.

the fact that $L_1(c,c) = -\left(\frac{1}{9}+2kc\right) < 0$. From Lemma 1 we have that if there is an equilibrium, it cannot be symmetric in $[0,c]^2$.

To apply Theorem 4 we must also check, whether $B5$ is verified. To this end, we must find the point d where the reaction curve has a jump. If the reaction curve is given implicitly, there is no algorithm to find this point d, however it is possible to find it through numerical methods. In the case of the model by Aoki and Prusa (1996), the point d is equal to $\frac{1}{12k}$. It is possible to compute that $U_1(x,0) = \frac{1}{4} - 2kx \geq 0$ for $x > d$ and that for sufficiently big c, $L_1(x,c) = \lim_{y\to\infty} L_1(x,y) = -2kx + \frac{1}{16} < 0$ for $x > d$. So the assumptions of Theorem 4 are satisfied and we can conclude that there exist only asymmetric equilibria in the game. Furthermore, reaction curves in this model are contractions and so a unique pair of equilibria exists.

The second paper that can be analyzed under our framework is Motta (1993). He develops two versions of a vertical product differentiation model, one with fixed and the other with variable cost of investment in quality. In each of them, he compares price versus quality competition. The model with fixed cost and price competition can illustrate the results of this section. Considering the same notation as in Aoki and Prusa (1996), the payoff function is defined as follows:

$$F(x,y) = \begin{cases} \frac{4v^2x^2(x-y)}{(4x-y)^2} - \frac{x^2}{2} \text{ if } x \geq y \\ \frac{xyv^2(y-x)}{(4y-x)^2} - \frac{x^2}{2} \text{ if } x \leq y \end{cases} \tag{2.16}$$

Where v is the upper limit of the set of consumer's taste parameters. This model is analogous to Aoki and Prusa (1996) if we let $v = 1$ and $k = \frac{1}{2}$. All the results follow directly from the above discussion.

2.5 Convex Payoffs

Certain features of the production technologies or consumer preferences may lead to a situation where payoff functions are convex. In particular, the presence of highly convex demand functions or strongly concave costs translating intensely decreasing elasticity of demand or decreasing marginal costs might have this effect. This property of the payoff functions implies that agents prefer corner solutions. Moreover certain additional conditions on the payoffs might generate asymmetric equilibria and even rule out the symmetric ones. In this section we analyze a class of games in which players have convex payoff functions and only asymmetric PSNEs arise.

Let $S = [0, c]$ be the strategy space of a player in a two-player game, Γ. Consider $F : S \times S \to R$ as the payoff function of player 1. $G(y, x) = F(x, y)$ is the payoff of player 2 because the game is *a priori* symmetric. Then we can apply the following theorem to conclude about the properties of the equilibria of Γ.

Theorem 6 *Assume that the following assumptions hold:*

1. *F is strictly quasi-convex in own strategy*
2. *$F(c, 0) > F(0, 0)$ and $F(0, c) > F(c, c)$*

Then, the game has no symmetric equilibrium and it has exactly one pair of asymmetric equilibria given by $(0, c)$ and $(c, 0)$.

Proof. From the definition of strict quasi-convexity, we know that:

$$F(\lambda z_1 + (1 - \lambda) z_2) < \max\{F(z_1), F(z_2)\}, \forall z_1, z_2 \in S \times S$$

Let $z_1 = (0, y)$ and $z_2 = (c, y)$, then we know that any $x \in (0, c)$ yields a lower payoff than $x = 0$ or $x = c$, $\forall y \in [0, c]$. This means that $\forall y \in [0, c]$, $r_1(y) = 0$ or $r_1(y) = c$. Analogously, for player 2 we have the same result, due to symmetry.

Finally, from Assumption 2 we have that $(c, 0)$ and $(0, c)$ are the only PSNEs of the game. ■

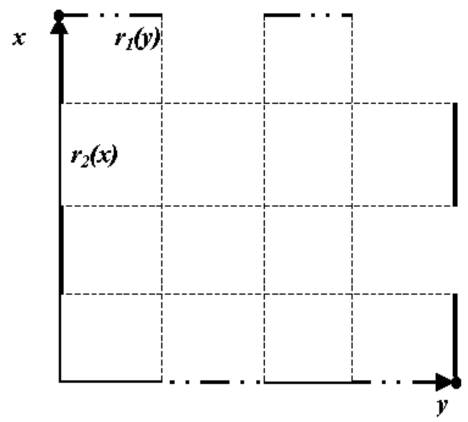
Figure 2.4 illustrates the results of this section. Notice that we depicted reaction curves which are not continuous, specifically, they possess an odd number of jumps. This is a direct consequence of the Assumptions 1 and 2 that imply either that the reaction curves are continuous, or that, if they jump, they must jump an odd number of times.

2.5.1 Applications

This theorem generalizes the results of Amir (2000) and Mills and Smith (1996), whose models are usually presented within the literature about endogenous heterogeneity of firms.[13] In these papers it is considered a two-stage duopoly game. In the first stage, firms make long term investment decisions that affect the production costs. In the

[13] Another example where the game can be reduced to a two person normal form game is presented in Boyer and Moreaux (1997). Conditions for the non-existence of symmetric equilibria are the same, even if the payoff function is not convex.

FIGURE 2.4. Quasi-convexity of the payoffs implies that players prefer corner solutions.



second stage, firms compete *à la* Cournot. Both firms face a linear demand function and for some parameterization of the cost functions, profits are convex in own quantity. In both papers it is then concluded that if some conditions on the parameters hold, asymmetric equilibria might arise. It is easily shown that the conditions presented in the papers can be deduced from the assumptions of Theorem 6.

2.6 Extensions

Most of the results of the previous sections depend on some form of monotonicity of the reaction curves. We used the cardinal notions of complementarity and substitutability to obtain this property mainly due to its convenient characterization through the cross partial derivatives of the payoff functions. Supermodularity and submodularity are cardinal notions that are not preserved by monotone transformations of the objective function. The usefulness of ordinal properties under which comparative statics results are invariant is clear. Milgrom and Shannon (1994) proved that the result of Topkis holds when single crossing property substitutes supermodularity of the objective function. In this section we show that our results can be generalized to the ordinal definitions of complementarity.

Let F be defined as in (2.1) and consider the following assumptions:

A1′ U and L have dual single crossing property

A4 $U_2(\overline{r}_1(y), y) < L_2(\underline{r}_1(y), y), \forall y \in [0, c]$

The Assumption $A1'$ is alternative to Assumption $A1$ defined in Section 2.3 and provides the ordinal condition for the reaction curves to be decreasing. Assumption $A4$ expresses monotonicity of U and L with respect to the opponent's action along the best replies.

We now show the main result of this section in the following steps: first we conclude about decreasing best-replies; then we observe that there is a downward jump in the reaction function at point d and that this jump is unique. Finally we use Topkis fixed point theorem to conclude of the existence of equilibria in the subsets of the strategy space defined using point d.

Finally we treat the setting of Section 2.4, extending its results to the ordinal definitions of complementarity.

Theorem 7 *If $A1', A2, A3$ and $A4$ hold, then the game Γ has at least one pair of asymmetric PSNEs and no symmetric one.*

The idea of the proof is that within Δ_U and Δ_L the dual single crossing property of the payoff function F holds from assumption $A1'$. As Milgrom and Shannon (1994) showed, the dual single crossing property allows us to draw the same conclusions as the submodularity of the payoff function in terms of monotonicity of the reaction curves. Furthermore the dual single crossing property has the advantage of being more general and preserved by monotonic transformations. Then assumption $A2$ implies that there exists a jump down in the reaction curves at a certain point d and $A4$ implies that this jump is unique. With point d we define compact subsets of the strategy space where reaction curves are decreasing and thus we can guarantee the existence of a PSNE by Tarski's Fixed Point Theorem.

Some results from Section 2.4 can also be extended into an ordinal version.

Let F be defined as in (2.1) and consider the following assumption:

$B1''$ U and L have the single crossing property

Theorem 8 *If the assumptions $B1''$ and $B2 - B5$ hold, then there exist at least one pair of asymmetric PSNEs and no symmetric one.*

From Milgrom and Shannon's Theorem Assumption $B1''$ implies that reaction curves are partially increasing in Δ_U and Δ_L. From this fact and as long as they are well defined in a compact subset of the strategy space, we can obtain existence of PSNEs.

The valley-shape of the profit precludes symmetric equilibria in the same spirit as Theorem 4.

2.7 Conclusion

Our theorems assert that, under specific conditions, heterogeneity in agent's behavior might arise even when they are *a priori* identical. This paper constitutes, hence, a contribution to the discussion about the sources of diversity across economic agents and disparities in economic performances. While previous literature stands on arguments related to multiplicity of equilibria and strategic complementarities (Cooper, 1999) or on strategic substitutability and stability of equilibria (Matsuyama, 2002), our approach stands on the existence of a fundamental nonconcavity of the payoff function and on some form of strategic substitutability. It is, thus, similar in spirit to Matsuyama's work. However, we show that endogenous heterogeneity does not rely on the idea that only stable equilibria are observable as in Matsuyama. With respect to Cooper's approach, where agents can still choose symmetrically, our results guarantee that symmetric equilibria can never arise in a two player setting. Even though we have, in our model some form of strategic substitutability (notice that when talking about two-player games strategic substitutability can be converted into complementarity through a simple inversion of one agent's strategy space), the critical assumption for the inexistence of symmetric equilibria is the nonconcavity of the payoffs. In fact, we show that strategic substitutability can be replaced by partial quasiconcavity and still the results follow.

An alternative explanation for endogenous heterogeneity can be found in business strategy literature. Our results can also be considered as a response to its critique as asymmetries arise here in a completely deterministic and rational setup. We should thus expect that heterogeneity generated in such a framework can be the origin of long-term diversity.

2.A Appendix

2.A.1 Summary of supermodular/submodular games

We give an overview of the main definitions and results in the theory of supermodular games that are used in the paper, in a simplified setting that is sufficient for our purposes. Details may be found in Topkis (1978).[14]

Let I_1 and I_2 be compact real intervals and $F : I_1 \times I_2 \to R$. F is (strictly) supermodular if $\forall x_1, x_2 \in I_1, x_2 > x_1$ and $\forall y_1, y_2 \in I_2, y_2 > y_1$ we have $F(x_2, y_2) - F(x_2, y_1)(>) \geq F(x_1, y_2) - F(x_1, y_1)$. F is (strictly) submodular if $-F$ is (strictly) supermodular.

Theorem 9 (Topkis's Characterization Theorem) *Let F be twice continuously differentiable. Then*

(i) $F_{12} = \frac{\partial^2 F}{\partial x \partial y} \geq 0$ $[\leq 0]$ for all $x, y \Leftrightarrow F$ is supermodular [submodular].

(ii) $F_{12} = \frac{\partial^2 F}{\partial x \partial y} > 0$ $[< 0]$ for all $x, y \Rightarrow F$ is strictly supermodular [submodular].

The supermodularity property is not preserved by monotonic transformations of the function F. An alternative notion (ordinal) is the single crossing property defined as follows: F has single crossing property [dual single crossing property] in (x, y) if $\forall x_1, x_2 \in I_1, x_2 > x_1$ and $\forall y_1, y_2 \in I_2, y_2 > y_1$ we have

$$F(x_1, y_2) - F(x_1, y_1) \geq 0\, [\leq 0] \Rightarrow F(x_2, y_2) - F(x_2, y_1) \geq 0\, [\leq 0]\,.$$

The single crossing property does not have a correspondent differential characterization and thus it is often more difficult to check. Now we present the main monotonicity theorems.

Theorem 10 (Topkis's Monotonicity Theorem) *If F is continuous in y and (strictly) supermodular [submodular] in (x, y), then $\arg\max_{y \in I_2} F(x, y)$ has (all of its) maximal and minimal selections increasing [decreasing] in $x \in I_1$.*

Theorem 11 (Milgrom and Shannon) *The conclusion of Topkis's Theorem continues to hold when supermodularity [submodularity] is replaced by the [dual] single crossing property.*

[14] Other aspects of the theory may be found in Topkis (1979), Milgrom and Roberts (1990) and Vives (1990).

We can introduce now the notion of supermodular game and of its properties. A two player game is supermodular (submodular) if both payoff functions are continuous, supermodular (submodular) and both action spaces are compact real intervals. [15] The fixed point theorems associated with this framework are due to Tarski (1955).

Theorem 12 (Tarski's Fixed Point Theorem) *Let $f : I_1 \times I_2 \to I_1 \times I_2$ be an increasing function, then f has a fixed point.*

Theorem 13 *A two player supermodular (submodular) game has a pure strategy Nash equilibrium.*

In general, this theory dispenses with assumptions of concavity or differentiability of payoff functions, making it an extremely general framework to study the properties of equilibria.

2.A.2 Proofs of Section 2.3

The proof of Theorem 1 is organized as follows: we begin with proving four preliminary lemmas, and then present the main proof in two steps: first we show existence of PSNE and afterwards that all PSNEs must be asymmetric.

The first lemma states that for a small enough square of points on the diagonal we have submodularity.

Lemma 3 *Consider the following points as depicted in figure 2.5. If $A1-A2$ hold, then for small enough $\alpha > 0 : U(x, x-\alpha) - U(x-\alpha, x-\alpha) \geq L(x, x) - L(x-\alpha, x), \forall x \in [0, c]$.*

Proof. (Lemma 3) Take any point (x, x) on the diagonal, belonging to the domain of F. For $\alpha > 0$ small enough

$$U_1(x, x) \simeq U(x+\alpha, x) - U(x, x) \text{ and } L_1(x, x) \simeq (x, x) - L(x-\alpha, x). \tag{2.17}$$

Hence, from $A2$:

$$U(x+\alpha, x) - U(x, x) > L(x, x) - L(x-\alpha, x). \tag{2.18}$$

[15] Compactness is not necessary, it is required in order to use a simplified version of Tarski's Fixed Point Theorem, without referring to lattices.

From $A1$ we know that

$$U(x+\alpha, x-\alpha) - U(x, x-\alpha) \geq U(x+\alpha, x) - U(x, x). \tag{2.19}$$

Take $\widehat{\varepsilon} > 0$ such that

$$U(x+\alpha, x) - U(x, x) \geq L(x, x) - L(x-\alpha, x) + \widehat{\varepsilon}. \tag{2.20}$$

From continuity of $U(x, y)$ in x (for fixed $y = x - \alpha$), it follows that

$$\forall 0 < \varepsilon \leq \frac{\widehat{\varepsilon}}{2},\ \exists \delta > 0 \text{ such that}$$
$$|x - (x-\alpha)| = \alpha \leq \delta \text{ and } |U(x, x-\alpha) - U(x-\alpha, x-\alpha)| \leq \varepsilon \leq \frac{\widehat{\varepsilon}}{2}.$$

Moreover,

$$\forall 0 < \varepsilon \leq \frac{\widehat{\varepsilon}}{2},\ \exists \delta > 0 \text{ such that}$$
$$|x + \alpha - x| = \alpha \leq \delta \text{ and } |U(x+\alpha, x-\alpha) - U(x, x-\alpha)| \leq \varepsilon \leq \frac{\widehat{\varepsilon}}{2}.$$

This allows us to establish that:

$$(U(x+\alpha, x-\alpha) - U(x, x-\alpha)) - (U(x, x-\alpha) - U(x-\alpha, x-\alpha)) \leq 2\varepsilon \leq \widehat{\varepsilon} \tag{2.21}$$

since, for $|A| < \varepsilon$ and $|B| < \varepsilon$, then $A - B < 2\varepsilon$.

Rewriting (2.21),

$$U(x+\alpha, x-\alpha) - U(x, x-\alpha) \leq U(x, x-\alpha) - U(x-\alpha, x-\alpha) + \widehat{\varepsilon}. \tag{2.22}$$

Finally, summarizing, we have:

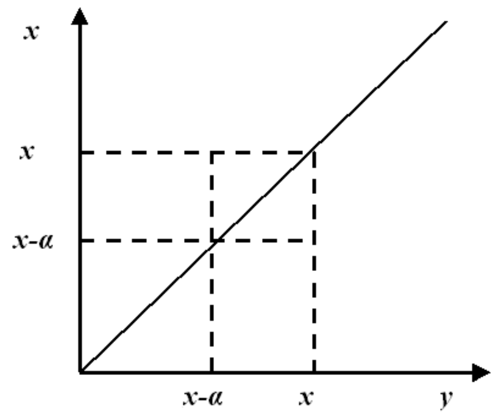
$$\begin{gathered}
L(x, x) - L(x-\alpha, x) + \widehat{\varepsilon} \leq \\
\leq U(x+\alpha, x) - U(x, x) \leq \\
\leq U(x+\alpha, x-\alpha) - U(x, x-\alpha) \leq \\
\leq U(x, x-\alpha) - U(x-\alpha, x-\alpha) + \widehat{\varepsilon}
\end{gathered}$$

where, the first inequality comes from (2.20), the second one from (2.19) and the last one from (2.22). So we obtain,

$$U(x, x-\alpha) - U(x-\alpha, x-\alpha) + \widehat{\varepsilon} \geq (x, x) - L(x-\alpha, x) + \widehat{\varepsilon}. \tag{2.23}$$

Subtracting $\widehat{\varepsilon}$ from both sides ends the proof. ■

FIGURE 2.5. Square of the length α.



The next lemma extends the property of submodularity of F form the small square of length α to any square with two vertices on the diagonal.

Lemma 4 *If Lemma 3 holds, then for any square with points in the diagonal, such as depicted in figure 2.7, we have,*

$$F(z, x) - F(z, z) \geq F(x, x) - F(x, z).$$

Proof. (Lemma 4) Consider the square formed by the four points defined in the lemma. Divide this square into rectangles such that their height is equal to the original height of the square and its length is not bigger than α, as defined in Lemma 3. We will now show that for points in the vertices of such rectangles the thesis holds and *a fortiori* it is possible to obtain the conclusion for the whole square. Figure 2.6 illustrates the proof.

Let $x - z = k\alpha$, where $k \in R$, and $\alpha > 0$ is small enough. Now consider the rectangle defined by the following points $(x, x), (x, x-\alpha)$, (z, x) and $(z, x-\alpha)$. From Lemma 3 we know that

$$F(x, x-\alpha) - F(x-\alpha, x-\alpha) \geq F(x, x) - F(x-\alpha, x)$$

Also, from $A1$ we know that

$$F(x-\alpha, x-\alpha) - F(z, x-\alpha) \geq F(x-\alpha, x) - F(z, x)$$

Adding these two inequalities we obtain that

$$F(x, x-\alpha) - F(z, x-\alpha) \geq F(x, x) - F(z, x) \tag{2.24}$$

Repeating the procedure, consider the rectangle defined by: $(x, x-\alpha)$, $(x, x-2\alpha)$, $(z, x-\alpha)$, $(z, x-2\alpha)$. From $A1$ we know that:

$$F(x, x-2\alpha) - F(x-\alpha, x-2\alpha) \geq F(x, x-\alpha) - F(x-\alpha, x-\alpha)$$

and also

$$F(x-2\alpha, x-2\alpha) - F(z, x-2\alpha) \geq F(x-2\alpha, x-\alpha) - F(z, x-\alpha).$$

Using Lemma 3 we know that

$$F(x-\alpha, x-2\alpha) - F(x-2\alpha, x-2\alpha) \geq F(x-\alpha, x-\alpha) - F(x-2\alpha, x-\alpha)$$

Adding the three inequalities we obtain:

$$F(x, x-2\alpha) - F(z, x-2\alpha) \geq F(x, x-\alpha) - F(z, x-\alpha)$$

From (2.24) we obtain

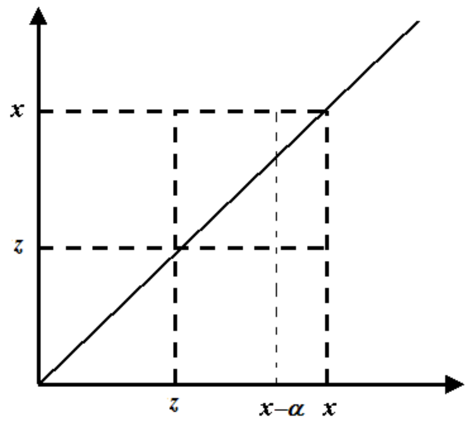
$$F(x, x-2\alpha) - F(z, x-2\alpha) \geq F(x, x) - F(z, x)$$

We can repeat this argument k times until getting a rectangle whose length is not bigger than α. Once again we apply assumption $A1$ and Lemma 3 to show that submodularity holds for this rectangle as well and we can conclude that,

$$F(x,z) - F(z,z) \geq F(x,x) - F(x,z).$$

Hence submodularity is satisfied for any square with vertices coinciding with the diagonal. ■

FIGURE 2.6. Partition of the square whose vertices coincide with the diagonal.



The following Lemma establishes that the analysis of submodularity of any rectangle formed by four points of the domain $[0,c]^2$ can be reduced to the analysis of submodularity for points placed in such a way that they form a square with vertices coinciding with the diagonal.

Lemma 5 *If $A1$ and $A2$ hold, then F and G are submodular on* $[0,c]^2$.

Proof. (Lemma 5) Due to the kink along the diagonal, one cannot invoke Topkis's simple cross-partial test (Topkis's Characterization Theorem in the Appendix) to verify submodularity of F and G. Instead, we use definition of submodularity for any configuration of four points in the domain, constituting a rectangle. If the rectangle is completely contained in either Δ_U or Δ_L, the submodularity condition follows from $A1$. Every other situation can be reduced by adding subrectangles, each of which lying fully in either Δ_U or Δ_L, to the situation depicted in figure 2.5 as we now show, say for F. Consider the case of figure 2.7 with the four points (x,z), (z,z), (x,y), (z,y) as shown. With $z < x < y$, we know from $A1$ that, since $F = U$ on Δ_U,

$$F(x,y) - F(x,x) \geq F(z,y) - F(z,x).$$

From Lemma 4, submodularity holds for the configuration of the square (x,x), (z,x), (z,z), (z,x), hence we have

$$F(x,x) - F(x,z) \geq F(z,x) - F(z,z).$$

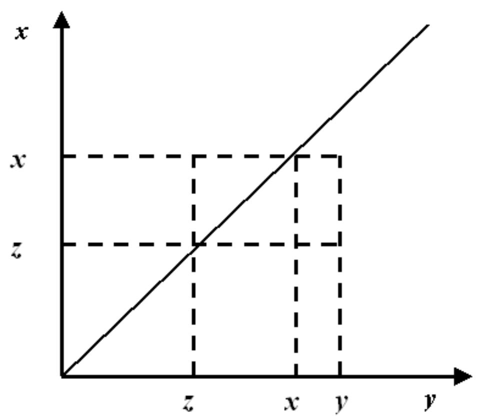
Adding the two inequalities yields

$$F(x,y) - F(x,z) \geq F(z,y) - F(z,z),$$

which is just the definition of submodularity for the original points (x,z), (z,z), (x,y) and (z,y).

It can be shown via analogous steps that the submodularity of F for any other configuration of points can be reduced to showing submodularity for squares with two vertices on the diagonal. The details are left out. ■

FIGURE 2.7. If F satisfies submodularity on the square on the diagonal, this implies it satisfies submodularity of the rectangle.



The next result allows us to conclude that the two reaction curves always admit a discontinuity that skips over the diagonal, a key step for our endogenous heterogeneity result.

Lemma 6 *Given $A1-A3$, there exists exactly one point $d \in (0, c)$, such that $r_i(d-\varepsilon) > d > r_i(d+\varepsilon), i = 1, 2$.*

Proof. (Lemma 6) From Topkis's Monotonicity Theorem and Lemma 5 , all the selections from the best reply correspondences are downward sloping. Hence, both $r_i(d-\varepsilon)$ and $r_i(d+\varepsilon)$ exist for any selection of r_i and are independent of the selection.

From assumption $A3$, we know that $(0,0) \notin Graph\ r_i$ and $(c,c) \notin Graph\ r_i$ (i.e. r_i does not go through $(0,0)$ or (c,c)). These two properties imply that r_i cannot be identically 0 or c.

We next show that the reaction correspondence r_1 (say) cannot ever cross the 45^o line at an interior point, i.e. in $(0,c)$. The generalized first order condition for a maximum of F (say) to occur at a point (x,x) with $x \in (0,c)$, which applies even in the absence of differentiability, is that $U_1(x,x) \leq L_1(x,x)$. Assumption $A2$ rules out this case . Hence no $x \in (0,c)$ can ever be a best reply to itself, meaning that the reaction curves do not cross the 45^o line at any interior point.

Since r_1 starts strictly above 0 (for $y = 0$) and ends strictly below c (for $y = c$), the above properties of r_1 imply that there exists exactly one $d \in (0,c)$ such that $r_1(d-\varepsilon) > d > r_1(d+\varepsilon)$. In words, there must exist a jump in the best reply function past the diagonal as in figure 2.1. ■

Using the Lemmas 5 and 6 we can now prove Theorem 1.

Proof. (Theorem 1) From Lemma 5 we have overall submodularity of the payoff function. This guarantees that a PSNE exists.

Consider now the behavior of the reaction curves in the area Δ_U. The same conclusion follows for Δ_L by symmetry. Define the following restricted reaction curves: $r_1|_{\Delta_U}(y) : [0,d] \to [d,c]$ and $r_2|_{\Delta_U}(x) : [d,c] \to [0,d]$ both decreasing as implied by Lemma 5. Define the mapping $B : [d,c] \to [d,c]$, $B(x) = r_1|_{\Delta_U} \circ r_2|_{\Delta_U}(x)$, which is increasing given that each of $r_1|_{\Delta_U}$ and $r_2|_{\Delta_U}$ is decreasing. From Tarski's Fixed Point Theorem, we know that there exists $\bar{x}$ such that $B(\bar{x}) = r_1|_{\Delta_U} \circ r_2|_{\Delta_U}(\bar{x})$, therefore $(\bar{x}, r_2(\bar{x})|_{\Delta_U})$ is a PSNE. From Lemma 6, there is no symmetric PSNE in $[0,c]$. Hence, there must exist at least one pair of asymmetric PSNEs. ■

Theorem 2 rules out the existence of multiple pairs of asymmetric equilibria.

Proof. (Theorem 2) Once again we concentrate on the area Δ_U. Conclusions follow for the area Δ_L by symmetry. Whenever r_1 $[r_2]$ is interior, first order condition $U_1(r_1(y), y) = 0$ $[L_1(r_2(x), x) = 0]$, together with the implicit function theorem and the assumptions (2.3) and (2.4), implies that r_1 $[r_2]$ is differentiable in Δ_U and that $r_1'(y) = -\frac{U_{12}(r_1(y),y)}{U_{11}(r_1(y),y)} \geq -1$, also $r_2'(x) = -\frac{L_{12}(r_2(x),x)}{L_{11}(r_2(x),x)} \geq -1$. Hence, $r_1(y)|_{\Delta_U}$ and $r_2(x)|_{\Delta_U}$ are contractions. Using Banach's fixed point theorem we can conclude that there exists exactly one pure strategy Nash equilibrium in Δ_U.[16] In the same way there exists exactly one pure strategy Nash equilibrium in Δ_L. Concluding, we have exactly one pair of pure strategy Nash equilibrium. ■

Finally we provide a proof of Theorem 3.

Proof. (Theorem 3) Since $x^* > d > y^*$, we have $F(x^*, y^*) = U(x^*, y^*)$ and $F(y^*, x^*) = L(y^*, x^*)$. Also $U(r_1(d), d) = L(r_1(d), d)$ if d denotes the unique point of jump of reaction curve between Δ_U and Δ_L, as defined in Lemma 6. Then

$$\begin{aligned} F(x^*, y^*) &= U(x^*, y^*) = \\ &= U(r_1(y^*), y^*) \leq U(r_1(d), d) = \\ &= L(r_1(d), d) \leq L(r_1(x^*), x^*) = L(y^*, x^*) = F(y^*, x^*) \end{aligned}$$

where both inequalities follow from the monotonicity of $U(r_1(y), y)$ and $L(r_1(y), y)$. ■

2.A.3 Proofs of Section 2.4

First we prove Lemma 1 and Lemma 2 since these proofs are similar. We then move to proving Theorems 4 and 5.

Proof. (Lemma 1) We consider the area Δ_U. From $B1$ and Topkis's Monotonicity Theorem, we know that the reaction curves are increasing. The generalized first order condition for a maximum to occur in the point (x, x) in the absence of differentiability of F (and G, by symmetry) is that $U_1(x, x) \leq L_1(x, x)$. Assumption $B2$ rules out this possibility so we have that no y (nor x), belonging to $(0, c)$, can be best reply to itself, meaning that the reaction curves do not cross the 45^o line. Assumption $B4$ excludes that 0 can be a best reply to 0 and that c can be a best reply to c. Hence, there must exist a $d \in [0, c]$, such that $r_1(d - \varepsilon) > d > r_1(d + \varepsilon)$.

To exclude the possibility of another jump we use assumption $B3$. Consider $\underline{r}_1(y)$ and $\overline{r}_1(y)$ defined as in Section 2.4. Denote $W(y) = L(\underline{r}_1(y), y)$ and $V(y) = U(\overline{r}_1(y), y)$. From the Envelope Theorem, $\frac{\partial W(y)}{y} = L_2(\underline{r}_1(y), y)$, and $\frac{\partial V(y)}{y} = U_2(\overline{r}_1(y), y)$. Hence, when $B3$ holds, we know that W increases in y quicker then V does. It means that

[16](Banach's Fixed Point Theorem): Let $S \subset \mathcal{R}^n$ be closed and $f : S \to S$ be a contraction mapping, then there exists $x \in S : f(x) = x$.

when the overall reaction curve jumps down along the diagonal, it never jumps up again. ■

Proof. (Lemma 2) If $B1'$ holds reaction curves are continuous in Δ_U and in Δ_L. $B2$ and $B4$ rules out the possibility that F (and G, by symmetry) has a maximum in a point $[x, x]$, that secures that the reaction curve must have a jump down in a point $d \in [0, c]$. As in the proof of Lemma 1, $B3$ rules out other possible upward jumps between Δ_L and Δ_U. ■

We may now show that only asymmetric pure strategy Nash equilibria exist.

Proof. (Theorem 4) Consider $R \subset \Delta_U$ as defined in Section 2.4. Define as before the restricted reaction curves as $r_1(y)|_{\Delta_U}$ and $r_2(x)|_{\Delta_U}$. From assumption $B5$, the best reply of player 1 to $y = 0$ cannot be less than d as U is decreasing in x for $y = 0$ when $x < d$. For $y > 0$, and given assumption $B1$ (supermodularity), Topkis's Monotonicity Theorem allows us to conclude that the best reply of 1 is increasing. Hence $r_1(y)|_{\Delta_U} \in R$. Seemingly the best reply of player 2 for $x = c$ cannot exceed d as L is decreasing in y when $x = c$. Also by Topkis's, the reaction curve is an increasing map. Therefore $r_2(x)|_{\Delta_U} \in R$. Consider the mapping, $B : R \to R$ such that $B(x, y) = (r_1(y), r_2(x))$. $B(x, y)$ is an increasing correspondence, given that both its components are increasing. R is a compact set and hence we may use Tarski's Fixed Point Theorem to conclude that there exists a pair $(\bar{x}, \bar{y})$ such that $B(\bar{x}, \bar{y}) = (r_1(\bar{y}), r_2(\bar{x}))$. $(\bar{x}, \bar{y})$ is a pure strategy Nash equilibrium.

Finally, by Lemma 1 we know that no equilibrium can be symmetric. ■

Proof. (Theorem 5) Define R, $r_1(y)|_{\Delta_U}$ and $r_2(x)|_{\Delta_U}$ as before. From assumption $B5$, $r_1(y)|_{\Delta_U} \in R$ and $r_2(x)|_{\Delta_U} \in R$ and from assumption $B1'$ they are continuous. Consider the mapping $B : R \to R$ such that $B(x, y) = (r_1(y), r_2(x))$. B is a continuous correspondence, given that both its components are continuous, R is a compact set and hence we may use Brouwer's Fixed Point Theorem to conclude that there exists a pair $(\bar{x}, \bar{y})$ such that $B(\bar{x}, \bar{y}) = (r_1(\bar{y}), r_2(\bar{x}))$. $(\bar{x}, \bar{y})$ is a pure strategy Nash equilibrium.

By Lemma 2 we know that no equilibrium can be symmetric. ■

2.A.4 Proofs of Section 2.6

To prove the theorem we first formulate a useful lemma.

Lemma 7 *If $A1\prime, A2$ and $A4$ hold, then there exist exactly one point $d \in [0, c]$ such that $r_1(d - \varepsilon) > d > r_1(d + \varepsilon), \varepsilon > 0$.*

Proof. (Lemma 7) Consider Δ_U. From $A1'$ and Topkis's Monotonicity Theorem we know that reaction curve is decreasing in this area. The generalized first order condition for a maximum to occur in (x, x) is (in the absence of differentiability in this point) $U_1(x, x) \leq L_1(x, x)$. Assumption $A2$ rules out this possibility, thus no $x \in (0, c)$ can be a best response to itself. Moreover, neither $(0, 0)$ nor (c, c) can be an equilibrium, since from $A3$ follows, that for player 1 it is always profitable to deviate from any of these points. Hence, the reaction curve does not cross the 45^o line, and there must exist a point, call it $d \in [0, c]$ such that $r_1(d - \varepsilon) > d > r_1(d + \varepsilon)$.

Now, we prove uniqueness of this point. Consider $\underline{r}_1(y)$ and $\overline{r}_1(y)$ defined as in Section 2.4. Denote $W(y) = L(\underline{r}_1(y), y)$ and $V(y) = U(\overline{r}_1(y), y)$. From the Envelope Theorem, $\frac{\partial W(y)}{y} = L_2(\underline{r}_1(y), y)$, and $\frac{\partial V(y)}{y} = U_2(\overline{r}_1(y), y)$. Hence, if $A3$ holds, we know that W increases in y quicker then V does. It means that when the overall reaction curve jumps down along the diagonal, it never jumps up again. ■

Proof. (Theorem 7) Consider restricted reaction curves $r_1|_{\Delta_U}(y)$ and $r_2|_{\Delta_U}(x)$, both decreasing as implied by $A1'$ and Topkis's Monotonicity Theorem. From Lemma 7 and the monotonicity it follows that $c \geq r_1|_{\Delta_U}(0) \geq r_1|_{\Delta_U}(d) > d$ and $0 < r_2|_{\Delta_U}(0) \leq r_2|_{\Delta_U}(d) \leq d$, therefore $r_1|_{\Delta_U}(y) : [0, d] \to [d, c]$ and $r_2|_{\Delta_U}(x) : [d, c] \to [0, d]$ are well defined. Define the mapping $B : [d, c] \to [d, c]$, $B(x) = r_1|_{\Delta_U} \circ r_2|_{\Delta_U}(x)$, which is increasing given that each of $r_1|_{\Delta_U}$ and $r_2|_{\Delta_U}$ is decreasing. From Tarski's Fixed Point Theorem, we know that there exists $\bar{x}$ such that $B(\bar{x}) = r_1|_{\Delta_U} \circ r_2|_{\Delta_U}(\bar{x})$, therefore $(\bar{x}, r_2(\bar{x})|_{\Delta_U})$ is a PSNE.

From Lemma 7, there is no symmetric PSNE in $[0, c]$. Hence, there must exist at least one pair of PSNEs. ■

References

Amir, R. (1996), Cournot oligopoly and the theory of supermodular games, *Games and Economic Behavior*, 15, 132-148.

Amir, R. (2000), R&D returns, market structure and research joint ventures, *Journal of Institutional and Theoretical Economics*, 156(4), 583-598.

Amir, R. and J. Wooders (1999), Effects of one-way spillovers on market shares, industry prices, welfare and R&D cooperation, *Journal of Economics and Management Strategy*, 8(2), 223-249.

Amir, R. and J. Wooders (2000), One-way spillovers, endogenous innovator/imitator roles, and Research Joint Ventures, *Games and Economic Behavior*, 31, 1-25.

Aoki, R and T. Prusa (1996), Sequential versus simultaneous choice with endogenous quality, *International Journal of Industrial Organization*, 15, 103-121.

Besanko, D. and U. Doraszelski (2002), Capacity dynamics and endogenous asymmetry in firm size, *RAND Journal of Economics*, 35, 23-49.

Boyer, M and M. Moreaux (1997), Capacity commitment versus flexibility, *Journal of Economics and Management Strategy*, 6(2), 347-376.

Cooper, R. (1999), *Coordination games: complementarities and macroeconomics*, Cambridge, UK: Cambridge University Press.

Cooper, R. and A. John. (1988), Coordinating coordination failures in Keynesian models, *Quarterly Journal of Economics*, 103, 441-463.

Cox, J. and M. Walker (1998), Learning to play Cournot duopoly strategies, *Journal of Economic Behavior and Organization*, 36, 141-161.

Dasgupta, P. and E. Maskin (1986), The existence of equilibrium in discontinuous economic games, I: theory, *Review of Economic Studies*, 53, 1–26.

Doraszelski, U. and M. Satterthwaite (2004), Foundations of Markov-perfect industry dynamics: existence, purification, and multiplicity, Mimeo.

Echenique, F. (2004), A characterization of strategic complementarities, *Games and Economic Behavior*, 46(2), 325-347.

Echenique, F. and A. Edlin (2004), Mixed equilibria in games of strategic complements are unstable, *Journal of Economic Theory*, 118(1), 61-79.

Gabszewicz J.J. and J.-F. Thisse (1979), Price competition, quality and income disparities, *Journal of Economic Theory*, 20, 340-359.

Geroski, P. A., Lazarova S., Urga G. and C. F. Walters (2003), Are differences in firm size transitory or permanent?, *Journal of Applied Econometrics*, 18(1), 47-59.

Ghemawat, P. (1986), Sustainable advantage, *Harvard Business Review*, 64, 53-58.

Hermalin, B.E. (1994), Heterogeneity in organizational form: why otherwise identical firms choose different incentives for their managers, *RAND Journal of Economics*, 25(4), 518-537.

Ireland, J.N. (1993), The provision of information in a Bertrand oligopoly, *Journal of Industrial Economics*, 41(1), 61-76.

Jovanovic (1982), Selection and the evolution of industry, *Econometrica*, 50, 649-70.

Kohlberg, E. and J.-F. Mertens (1986), Strategic stability of equilibria, *Econometrica*, 54, 1003-37.

Long, N. V. and A. Soubeyran (2001), Cost manipulation games in oligopoly, with costs of manipulating, *International Economic Review*, 42(2), 505-533.

Maggi G. (1996), Strategic trade policies with endogenous mode of competition, *American Economic Review*, 86(1), 237-258.

Matsuyama, K (2000), Endogenous inequality, *Review of Economic Studies*, 67, 743-759.

Matsuyama, K. (2002), Explaining diversity: symmetry-breaking in complementarity games, *American Economic Review*, 92, 241-246.

Matsuyama, K. (2004), Financial market globalization, symmetry-breaking, and endogenous inequality of nations, *Econometrica*, 72, 853-884.

Milgrom, P. and J. Roberts (1990), Rationalizability, learning and equilibrium in games with strategic complementarities, *Econometrica*, 58(6), 1255-1277.

Milgrom, P. and C. Shannon (1994), Monotone comparative statics, *Econometrica*, 62: 157-180.

Mills, D. E. and W. Smith (1996), It pays to be different: endogenous heterogeneity of firms in an oligopoly, *International Journal of Industrial Organization*, 14, 317-329.

Mintz, J. and H. Tulkens (1986), Commodity tax competition between member states of a federation: equilibrium and efficiency, *Journal of Public Economics*, 29, 133-172.

Motta, M. (1993), Endogenous quality choice: price vs. quality competition, *Journal of Industrial Economics*, 41(1), 113-131.

Murphy, K., A. Schleifer and R. Vishny (1989), Industrialization and the big push, *Journal of Political Economy*, 97, 1008-1026.

Nelson, R. and R. Winter (1982), *An evolutionary theory of economic change*, Harvard University Press, Cambridge, MA.

Nelson, R. (1991), Why do firms differ and why does it matter? *Strategic Management Journal*, 12, 61-74.

Reynolds, S. and B. Wilson (2000), Bertrand-Edgeworth competition, demand uncertainty and asymmetric outcomes, *Journal of Economic Theory*, 92, 122-141.

Röller L-H and B. Sinclair-Desgagné (1996), On the heterogeneity of firms, *European Economic Review*, 77, 402-420.

Rumelt, R., Schendel, D. and D. Teece (1991), Strategic management and economics, *Strategic Management Journal*, 12, 5-29.

Salant, S.W. and G. Schaffer (1998), Optimal asymmetric strategies in research joint ventures, *International Journal of Industrial Organization*, 16, 195-208.

Salant, S.W. and G. Shaffer (1999), Unequal treatment of identical agents in Cournot equilibrium, *American Economic Review*, 83, 585-604.

Schelling, T. (1960), *The strategy of conflict*, Harvard University Press, Cambridge.

Shaked, A. and J. Sutton (1982), Relaxing price competition through product differentiation, *Review of Economic Studies*, 49(1), 3-13.

Shapiro, C. (1989), The theory of business strategy, *RAND Journal of Economics*, 20, 125-137.

Stengel, von B. (2003), Follower payoffs in symmetric duopoly games, *CDAM Research Report LSE-CDAM-2003-13*.

Tabuchi, T. and J.-F. Thisse (1995), Asymmetric equilibria in spatial competition, *International Journal of Industrial Organization*, 13, 213-227

Tarski, A. (1955), A lattice-theoretic fixed point theorem and its applications, *Pacific Journal of Mathematics*, 5, 285-309.

Topkis, D. (1978), Minimizing a submodular function on a lattice, *Operations Research*, 26, 305-321.

Topkis, D. (1979), Equilibrium points in nonzero sum n-person submodular games, *SIAM Journal of Control and Optimization*, 17, 773-787.

Vives, X. (1990), Nash equilibrium with strategic complementarities, *Journal of Mathematical Economics*, 19, 305-321.

Vives, X. (1999), *Oligopoly pricing: old ideas and new tools*, The MIT Press, Cambridge, Massachussets.

3

TECHNOLOGY ADOPTION WITH FORWARD LOOKING AGENTS

3.1 Introduction

We characterize the unique equilibrium in a game of technology adoption where consumers are forward-looking. Our analysis focuses on technologies exhibiting network externalities, in that they become more valuable for each individual the more consumers adopt them (see Rolfhs (1974) for a seminal contribution). The literature has treated the problem of both static and dynamic decision-making in the presence of network externalities (Katz and Shapiro (1986), Shy (1996)). However, the possibility for agents to consider the behavior of their future counterparts –obviously ruled out within a static setup– has not received much attention within dynamic frameworks. In fact, most of the papers assume that agents realize their payoffs at the time they purchase a given technology. From then onwards they belong to the network, thus increasing future adopters' benefits but not receiving any further payoff. As a consequence, agents regard predecessors' adoptions only, that is they consider the installed base when opting for one technology. In this paper we deal with consumers receiving payoffs in all periods of their permanence in the network, in which case expectations over future agents' choices need to be factored in.

We consider an overlapping generations setup in which consumers live for two periods and opt for one of two technologies. We focus on stochastic technology values, i.e. consumers are not aware of the technology value in successive periods. This feature is in line with the analysis in Adner (2003): the marginal utility of technological improvements for consumers is not constant in every step of the technology evolution.

In fact, it assumes the form of an S-curve: agents have a very high willingness to pay for the first improvements, but this willingness decreases after a certain threshold. If we think about two technologies whose values evolve through time, it is possible that at a certain point consumers have higher willingness to pay for one technology, and at another point they prefer the other technology.

Consumers obtain utility from the value of the technology, the so called *stand alone value*, and from interaction with other consumers, the *network value* as in Arthur (1989), Choi (1994), Farrell and Saloner (1992) and Shy (1996). We distinguish between direct and indirect network externalities. Direct effects are related to the increase in the quality of the product due to an increment in the number of contemporaneous users. Indirect effects entail broader benefits stemming from past adoptions. Even though direct interaction with remote consumers is not possible, the fact that a technology has had many previous users increases its value for the consumer: there is higher likelihood that the technology has less flaws, technical assistance might be prompter and more and better components might be available (see Liebowitz and Margolis (1994) for a complete characterization of indirect network externalities).

Before making their choice, consumers observe technology values and predecessors' actions, and form expectations about future behavior. The main result of our paper is that a unique equilibrium exists in which individuals adopt technologies via switching strategies that depend on predecessors' choices. The presence of strategic complementarities due to network externalities induces multiple equilibria in adoption. If individuals expect the adoption of a certain technology, they have an incentive to choose this same technology and enjoy the network benefit. As Levin (2001) points out, 'because the expectation of a certain behavior tends to be self-fulfilling, there is the strong possibility of multiple expectations-driven equilibria' (Levin (2001), p. 1). The main problem is therefore to select among these equilibria the one that will actually be played. Given stochastic technology values and network effects, our game is a (dynamic) Bayesian game of strategic complementarities. By Van Zandt and Vives (2003) a static Bayesian game of strategic complementarities admits a greatest and a least Bayes-Nash equilibrium in monotone strategies. We show that in our dynamic framework this result still holds, and moreover that the greatest and least equilibria coincide. Hence the game is dominance solvable and has a unique equilibrium.

Our uniqueness result relates to the literature on so-called global games (see Morris and Shin (2003) for a general overview).[1] The main idea is that if 1) each player observes a noisy signal of the payoffs and 2) the payoffs' space includes values that make each action strictly dominant, then iterative dominance leads to a unique equilibrium as the noise becomes small. In our setup, stochastic technology values serve the purpose of a noisy signal and strategic complementarities arise due to network externalities.

A second result of our analysis is that technologies cannot *lock in* by historical events. Instead they will alternatively be chosen through time. Arthur (1989) presents a dynamic model in which agents choose between two competing technologies based on their personal preferences (stand alone value) as well as the network benefit, with the latter described by a function of the total number of previous adopters. His main finding is that whenever a given technology achieves a sufficient mass of adopters, it locks in attracting all future consumers. He considers as well the possibility that consumers take into account the future base. However he restricts his analysis to the case of lock-in by one of the technologies, neglecting in the first place how forward-looking behavior interplays with the existence of lock-in. In the present paper we address this issue, allowing agents to obtain payoffs every period of their lives and thus to form expectations about the future base. A similar standpoint is taken in Ochs and Park (2004), that highlight the importance of forward looking behavior in network formation.[2]

The absence of lock in is in line with Liebowitz and Margolis (1994; 1995), that question the empirical relevance of lock-in patterns in technology adoption, arguing that lock-ins are extremely unlikely to occur. We provide a different rationale for the absence of lock-ins. Agents incorporate successors' choices and face technologies with stochastic stand alone values. Given that these values are independent of the number of adopters, network externalities are not the only driving force behind technology adoption. A lead in terms of installed base is not enough to attract all consumers. This

[1] More recently, Burdzy, Frankel and Pauzner (2001), Frankel and Pauzner (2000) and Frankel (2001) study the conditions under which an equilibrium is selected for dynamic games with binary actions and strategic complementarities. Giannitsarou and Toxvaerd (2003) provide similar results for recursive games and Levin (2001) extends the main findings to overlapping generation games. For an application of Levin's result to the technology adoption game described in the present article see Colla and Garcia (2004).

[2] Their model differs from ours in that individual types (or technology values as in our setup) are uncorrelated, thus the global games equilibrium selection approach cannot be applied. In Ochs and Park (2004) it is possible to identify a unique symmetric perfect bayesian equilibrium since agents choose both when and whether to join a network.

is because agents are concerned with the stand alone value granted by the technology upon purchase, as well as with the value provided in subsequent periods.

Even though lock-in does not emerge in our setup, path dependence in technology choices is still present. Once the model parameters are laid out, the technology adopted by the ancestor affects the current user's decision. When this occurs the equilibrium path exhibits *hysteresis*. Moreover, the likelihood of observing hysteresis increases in the network benefit: the larger the latter, the higher the importance of predecessors' choices.

One empirical prediction of our model is that competing technologies coexist in the market. Several cases within the computer industry are compatible with our results. In fact, the computer industry does not comply entirely with early leading technologies becoming dominant (see Gandal, Greenstein and Salant (1999) for further details). In the early 80's, CP/M, a highly adopted operating system, was orphaned both by users and developers at the advantage of DOS, an operating system that had been recently developed.[3] In this case the technology that had initially gained some market share lost its relevance over time. One reason for this is the heterogeneity of adopters: early buyers are usually highly skilled, while late buyers are less skilled and might go for a more user friendly platform. Another example that contradicts the idea of lock-in is the persistence in time of the two computer platforms Macintosh and PC (IBM-PC). These two platforms became important at the time the 16-bit microprocessor was invented, and have shared the market since then. Passing through several waves of technological advances none of the two managed to become dominant in the market (see Bresnahan and Greenstein (1999)). An explanation for this evidence is the idea that technologies' relative values oscillate through time. Improvements in microprocessors' speed stand as a clear example of this. In the beginning microprocessors were very slow and any marginal increase in the speed was extremely valued by the consumers. However, as the microprocessors proved faster, speed increases became less important. In the Macintosh/PC example, the permanence in the market of the two standards is justified by the oscillating preferences of the consumers for the technological improvements in the two platforms.

We characterize as well the unique equilibrium in technology adoption when a converting device is available and allows agents from different networks to interact. The

[3] 'Orphaning occurs when late adopters choose a technology incompatible with the technology adopted by early users and suppliers of supporting services' (Gandal, Greenstein and Salant (1999), p. 88)

conclusion on the absence of lock in (that is a direct consequence of the stochastic pattern of the technology value), does not change when converters are allowed. However converters contribute to mitigate hysteresis: agents have weaker incentives to coordinate their choices, and as a consequence technology adoption depends less on predecessors' decisions.

Since technologies do not lock in, a different measure of dominance should be studied. The first measure we propose is the limiting probability of technology adoption. This describes the likelihood that each technology is chosen in the long run, and provides a rough estimate of its expected demand schedule over long horizons. We show that technologies are more likely to be adopted in the long run if they 1) provide higher stand alone values, or 2) agents prefer bigger networks, or 3) they embed a converter device. In short run equilibrium a technology will be adopted for a certain number of consecutive periods, and then replaced. From the producer's point of view, it is therefore important to determine what is the expected time of adoption, our second measure for dominance. We find that the expected time of adoption decreases in the presence of converters due to the fact that compatible technologies reduce path dependence and one observes switching between technologies more often.

The remainder of the paper is organized as follows. The sequential move game with incompatible technologies is described in section 3.2. Section 3.3 characterizes the unique equilibrium and analyses the impact of the underlying parameters on the equilibrium outcome. Partial converters are introduced in section 3.4. The long run behavior of our adoption game is described in section 3.5. Section 3.6 focuses on the robustness of the equilibrium outcome to alternative specifications. Finally section 3.7 concludes.

3.2 Theoretical model

Within a discrete time and infinite horizon setting we consider a sequence of users who are planning to adopt a technology. At each time $n \in N$, player n enters the game and chooses among two competing technologies A and B. Each player lives for two periods. We borrow the terminology from the OLG literature and refer to the young (resp. old) generation at time n as the n-th player (resp. the $(n-1)$-th player). In the first period, user n buys a single unit of one of the two technologies and commits to his choice in period $n+1$. Choice irreversibility can be introduced by assuming that it is prohibitely costly to switch, say, to technology A at time $n+1$ after having adopted technology B at time n. This assumption is quite standard in (dynamic) network adoption games.[4] We denote player n's action set as $A_n = \{0, 1\}$ where $a_n = 1$ (resp. $a_n = 0$) corresponds to technology A (resp. B). We restrict our analysis to unsponsored technologies resulting in their supply at a price equal to zero.[5] Agents discount future payoffs via the factor $\beta \in (0, 1)$.

3.2.1 Technology value

The technology value at time n is given by x_n, and can be thought of as player n's relative preference between the two technologies when no other agent shares the same base. We model the evolution of technology values through time as a random walk with barriers $-\alpha$ and α, where $\alpha > 0$, i.e.

$$x_n = x_{n-1} + \sigma\varepsilon_n \tag{3.1}$$

where $\sigma > 0$. The innovation ε is characterized by the density $g(\varepsilon)$:

$$g(\varepsilon) = \begin{cases} -1 & \text{prob. } q = 1 - p \\ 1 & \text{prob. } p \end{cases} \tag{3.2}$$

where $p \in (0, 1)$. The boundaries α and $-\alpha$ are partially reflecting. In other words, if the random walk is in state α at time n, the following period it can either go down to $\alpha - \sigma$ with probability q or stay at α with probability $p = 1 - q$.[6]

3.2.2 Individual preferences

Player's valuation of a technology reflects two components: the stand alone value and the network value. The former captures the utility the user derives if no other player adopts the same technology, while the latter is the benefit from interacting with other users.

[4] To our knowledge Farrell and Saloner (1985) are the only ones that obtain -rather than imposing- choice irreversibility. However they are admittedly unable to explain why this result obtains (see Farrell and Saloner (1985), footnote 9 as well as Malin (2003))

[5] The term unsponsored technologies was first used by Arthur (1989) to refer to technologies that are non-appropriable. The absence of property rights leads to entry in the market until marginal cost pricing condition is met, i.e. in our case until prices are zero.

[6] Similarly when $x_n = -\alpha$, the technology value can either jump to $-(\alpha - \sigma)$ (with probability p) or stay at $-\alpha$ (with probability q) at time $n + 1$.

3.2.2.1 Stand alone value

The $n-$th period stand alone value depends on the contemporaneous technology value x_n. After opting for one technology, generation n receives the stand alone value at period n. The time n stand alone value is therefore a function of the current technology value x_n and player n's chosen technology a_n according to

$$\pi_S = \pi_S(a_n, x_n) = \begin{cases} \alpha + x_n & \text{if } a_n = 1 \\ \alpha - x_n & \text{if } a_n = 0 \end{cases} \tag{3.3}$$

Given (3.3), players' relative preference for technology A over B is increasing in the technology value through an affine function.[7]

3.2.2.2 Network Value

Network externalities assume two different aspects. On one hand, consumers may interact between each other because they use the same technology (direct network value). On the other, consumers benefit from the technologies installed by previous generations, even without direct interaction (indirect network value). Each user lives for two periods and at each time n there are two generations active in the market. Thus player n's direct network value depends on the action chosen by his immediate predecessor $n-1$ and by his immediate successor $n+1$.

We consider the direct network value to be linear in the number of users of the same technology. Further we assume that the two technologies are identical in terms of the direct network benefit they provide per unit of member in the base, $\nu_D > 0$. If player n and $n-1$ adopt the same technology, they receive a direct network benefit of ν_D. The same payoff is received by user n and $n+1$ if they coordinate on the same technology. Generation n's per-period direct network payoff π_D is given by:

$$\pi_D = \pi_D(a_n, a_{-n}) = \begin{cases} \nu_D a_{-n} & \text{if } a_n = 1 \\ \nu_D(1 - a_{-n}) & \text{if } a_n = 0 \end{cases} \tag{3.4}$$

where a_{-n} denotes the technology purchased by the other generation active in the market, i.e. $a_{-n} = a_{n-1}$ at time n and $a_{-n} = a_{n+1}$ at time $n+1$.

[7] Our choice for the barriers in the stochastic process(3.1) together with the payoff in (3.3) imply that the stand alone value belongs to the interval $[0, 2\alpha] \in \Re^+$. This normalization provides individuals with a (weakly) positive utility from purchasing one of the two techologies in the absence of network benefits.

As for the indirect network value, we assume that the technologies chosen by previous generations affect player n's choice through a summary statistic $h_n = h(a_{n-M}, ..., a_{n-2})$, $M \geq 2$, which describes the past behavior, or equivalently the history of technology adoption. We require the function $h(\cdot)$ to be real-valued, bounded, separable and increasing in each of its arguments such that the technology chosen by each generation from $n-M$ to $n-2$ affects player n's utility.[8] Without loss of generality we let $h_n : \{0,1\}^{M-1} \rightarrow [0, H]$, where H is the value taken by the statistic h_n when all predecessors chose technology A, i.e. $H = h(1, ..., 1)$. The indirect network value is then given by

$$\pi_I = \pi_I (a_{n-M}, ..., a_{n-2}) = \begin{cases} \nu_I h_n & \text{if } a_n = 1 \\ \nu_I (H - h_n) & \text{if } a_n = 0 \end{cases} \tag{3.5}$$

where $\nu_I > 0$ describes the value attached to technology history.

3.2.2.3 Total Payoffs

Adding up the stand-alone component (3.3) and the network components (3.4-3.5) gives player n's overall utility $u_n = u(a_n, a_{n-1}, a_{n+1}, h_n, x_n)$:

$$u_n = \begin{cases} \alpha + x_n + \nu_D (a_{n-1} + \beta a_{n+1}) + \nu_I h_n & \text{if } a_n = 1 \\ \alpha - x_n - \nu_D [(1 - a_{n-1}) + \beta (1 - a_{n+1})] + \nu_I (H - h_n) & \text{if } a_n = 0 \end{cases}$$

For the remainder of our analysis it is useful to rewrite player n's payoff in the following compact way:

$$u_n = \pi_S (a_n, x_n) + \pi_D (a_n, a_{n-1}) + \beta \pi_D (a_n, a_{n+1}) + \pi_I (a_n, h_n) \tag{3.6}$$

The terms on the RHS in eq. (3.6) capture respectively the stand alone value, the direct network component -i.e. the sum of the payoff related to the existing base and to the (discounted) future base- and the indirect network component.[9]

[8] The assumption that h_n does not depend on the immediate predecessor action, a_{n-1}, is made without loss of generality. Suppose that the indirect network value depends on $(a_{n-M}, ..., a_{n-2})$ and on a_{n-1} as well through some statistic $h(\cdot)$. This parametrization would be equivalent to another one in which the direct network value attached to a_{n-1} is different from the one attached to a_{n+1}. This would not alter qualitatively our main results. See also section 3.2.4.

[9] According to (3.6), despite player n lives for two periods, he receives the technology stand alone value only once. This assumption is made for simplicity, and we could easily encompass a more articulated specification in which the utility depends on both x_n and x_{n+1} via, say, $\pi_S (a_n, x_n) + \beta \pi_S (a_n, x_{n+1})$. The main results would be unchanged, and the interested reader is referred to the working paper version of our work (see Colla and Garcia (2004)). A similar argument applies to the introduction of the (discounted) indirect network value, $\beta \pi_I (a_n, h_{n+1})$.

3.2.3 Timing and Strategies

At time n, agent n is aware of the current technology valuation x_n and the predecessors' choices $a_{n-M}, ..., a_{n-1}$. A strategy for player n is, thus, a function $s_n(a_{n-1}, h_n, x_n) : \{0,1\} \times [0, H] \times [-\alpha, \alpha] \to \{0,1\}$. Player n chooses his action to maximize the expected payoff in (3.6) . Letting $s_{n+1} = s(a_n, h_{n+1}, x_{n+1})$ denote the strategy of player $(n+1)$, the $n-$th player solves the following:

$$\max_{a_n \in A_n} E\left[u\left(a_n, a_{n-1}, s_{n+1}, h_n, x_n\right) | a_{n-1}, h_n, x_n\right] \tag{3.7}$$

From (3.7) it emerges that, when choosing which technology to purchase, player n takes into account the effect of his action in determining the future size of the network, which in turn affects player's $n+1$ compatibility payoff. It is worth noting that the process (3.1) allows the technological valuation to be correlated through time, which turns out to be crucial in enabling player n to forecast the next generation's strategy after observing x_n. We allow player n to receive network benefits from the (expected) future base via (3.7), and thus we explicitly bring in a role for predicting future technology values.

3.2.4 Discussion

We now briefly discuss the modelization assumptions related to timing, technology values and players' payoffs.

- *overlapping generations.* The overlapping generations structure has been chosen for two reasons. First, OLG models are appropriate to study the choice of buying a durable good and holding it for several periods. Second, OLG models allow agents to form expectations over their finite lifetime. Shy (1996) considers an OLG setup as well, but like Arthur (1989) the future behavior of agents is not regarded as a determinant of consumers' choices.

- *stochastic technology values.* We consider stochastic technology values (as in Choi (1994)) in order to capture the idea that the stand alone value at time $n+1$ will be revealed to players at the beginning of that period, such that x_{n+1} is a random variable from player n's standpoint at time n. Thus, we view the technology value as inherent to the technology itself, rather than being agent-specific. The step size $\sigma > 0$ in (3.1) can be thought of as a measure of (intertemporal) heterogeneity in technology valuation, and drives the randomness in players' payoff. As it has been mentioned in the Introduction, the literature on strategy and management provides an intuitive explanation for the stochastic technology values' assumption, known as the demand S-curve for technology improvements (see Adner (2003)). Alternatively, if we regard technologies as evolving through time, x_n measures the relative value of technologies' developments at time n.

- *correlated technology values.* The correlation through time of the technological valuation is crucial in enabling player n to forecast the next generation's strategy after observing x_n. Technology values are assumed to be independent through time in Arthur (1989), while the technology innovation -and thus the stand alone value- is given by a deterministic -rather than stochastic- process in Shy (1996). In the sequential move games of Arthur (1989) and Shy (1996) the intertemporal pattern of technology values does not play a relevant role, since users choose their action based on the installed base only. Unlike these works, we allow player n to receive network benefits from the (expected) future base via (3.7), and thus we explicitly bring in a role for predicting future technology values. Our specification (3.1) is closely related to Oyama (2003), which employs a similar process to describe the pattern of fundamentals within an OLG speculative attack model. One can also explain the correlation through time in technology values resorting to taste shocks in users' preferences like in Macskasi (2002). In this case each generation tastes' are evolving, as to say that the relative preferences for the young generation at time n might differ from the tastes of the same generation when old at time $n+1$. Note however that at each point in time both the young and the old generations agree on the technology value. This argument leads again to think of technology values as time specific.

- *affine stand alone value.* It is assumed that the stand alone depends on a technology value through an affine function. This formulation is fairly standard in the literature. Arthur (1989) considers two (classes of) agents with the former (resp. the latter) displaying a natural preference for technology A (resp. technology B). His formalization for the stand alone value can be obtained from ours restricting x to take only two values, say $x = \{-1, 1\}$, with equal probabilities. The specification for π_S in (3.3) is analogous to Farrell and Saloner (1992) with the only difference that technology values belong to the interval $[-\alpha, \alpha]$ in our model rather than $[0,1]$. Shy (1996) considers a richer specification for the stand alone value, without restricting it to affine functions of x_n.

- *direct network value.* In line with much of the previous literature we consider the direct network value to be linear in the number of users of the same technology. Further we assume that the two technologies provide the same direct network benefit per unit of member in the base, $\nu_D > 0$. This assumption is quite standard in the literature (see Choi (1994) and Farrell and Saloner (1992)). Allowing for asymmetric network benefits as in Arthur (1989) can be easily incorporated into our framework without changing the main conclusions.

- *indirect network value.* The specification (3.5) serves as a reduced form for more complex relationships between the installed base, i.e. the actions $a_{n-M}, ..., a_{n-2}$, and individual payoffs. Such indirect network externalities usually stem from the increase in the variety of complementary products available for each technology (see for example Clements and Ohashi). The requirements imposed on the function h allow for several specifications for such indirect network externalities. [10]

3.3 Solving for an equilibrium

3.3.1 The benchmark model: no network externalities

If there are no network externalities, utility is simply given by the stand alone value (3.3). The maximization problem (3.7) therefore resumes to:

$$\max_{a_n \in A_n} \pi_S(a_n, x_n)$$

[10] For example one might take the sum of predecessors' actions as summarizing past behaviour, and consequently set $h_n = \sum_{j=2}^{M} a_{n-j}$ and $H = M - 1$. In this case player n puts a weight equal to ν_I to each action chosen by generations from $n-M$ to $n-2$. Alternatively, a weighted average of past actions can be employed to capture the benefit from indirect interaction with predecessors, i.e. $h_n = \sum_{j=2}^{M} \omega_j a_{n-j}$ with $\sum_{j=2}^{M} \omega_j = 1$, $\omega_j > 0$ for all $j = 2, ..., M$ and $H = 1$. One can then choose decaying weights, i.e. $\omega_j > \omega_{j+1}$, in order to capture the idea the technologies more recently chosen have a larger impact on generation n payoff.

In the absence of network benefits, i.e. $\nu_D = \nu_I = 0$, agent n is indifferent between the two technologies when the technology value $x_n = \hat{x}$ solves:

$$\pi_S(1, \hat{x}) = \pi_S(0, \hat{x})$$

We refer to the technology value $\hat{x}$ as the pivotal point, that is the threshold above which the agent has a strict preference for technology A. Plugging (3.3) in the above equation gives $\hat{x} = 0$. Without network externalities the pivotal point is zero because, in equilibrium, actions are entirely determined by the technology value. If player n prefers technology A to B –this would occur when $x_n > 0$– then he opts for A.

3.3.2 The model with network externalities

We now analyze equilibrium strategies when both ν_D and ν_I are strictly positive. In order to solve for the equilibrium in the sequential move game outlined in section 3.2 we first prove the following:

Lemma 1 *Let the model parameters be such that*

$$2\alpha > \nu_D(1+\beta) + \nu_I H \ . \tag{3.8}$$

Then the space of technology values contains a region $[\bar{x}, \alpha]$ where technology A is dominant, and a region $[-\alpha, \underline{x}]$ where technology B is dominant, i.e.

$$a_n(x_n) = \begin{cases} 1 & \text{if } x_n \geq \bar{x} = \frac{\nu_D(1+\beta)+\nu_I H}{2} \\ 0 & \text{if } x_n < \underline{x} = -\bar{x} \end{cases}$$

As for the meaning of (3.8), recall from (3.3) that the maximal stand alone value is given by 2α, which would occur whenever the stocastic process in (3.1) reaches one barrier. From (3.4) the direct per-period network benefit is ν_D and from (3.5) the maximal indirect network benefit is $\nu_I H$. According to Lemma 1 there exist technology values for which the maximal stand alone valuation offsets the (discounted) benefits from coordinating on a network, i.e. adopting the installed technology ($\nu_D + \nu_I H$) given that it will be chosen by the immediate successor ($\beta\nu_D$). Lemma 1 ensures that some individuals would choose a technology regardless of the network size: technology A is dominant whenever x_n is above the critical value $\bar{x}$, whereas technology B is dominant if the technology value falls below $\underline{x}$. Condition (3.8) guarantees that the values $\bar{x}$ and

$\underline{x}$ lie in the state space for the technology process X. In general, multiple equilibria would occur within the interval $[\underline{x}, \bar{x}]$.[11] Lemma 1 plays a key role in applying an iterated elimination argument, and thus solving for the unique equilibrium. Proposition 1 constitutes our main result.

Proposition 1 *Under (3.8) the game has a unique equilibrium, in which for all n*

$$s\left(a_{n-1}, h_n, x_n\right) = \begin{cases} 1 \text{ } \mathit{if} \text{ } x_n \geq x^*\left(a_{n-1}, h_n\right) \\ 0 \text{ } \mathit{if} \text{ } x_n < x^*\left(a_{n-1}, h_n\right) \end{cases}, \tag{3.9}$$

where $x^*\left(a_{n-1}, h_n\right)$ *is a decreasing function of both* a_{n-1} *and* h_n*. Moreover, let the step size* σ *be sufficiently small such that:*

$$\sigma < \bar{\sigma} = \nu_D\left(1 - \frac{\beta}{2}\right) . \tag{3.10}$$

Then the cut-off points are given by:

$$x^*\left(0, h_n\right) = \frac{\nu_D\left(1 - \beta p\right) + \nu_I\left(H - 2h_n\right)}{2} \tag{3.11}$$

$$x^*\left(1, h_n\right) = \frac{-\nu_D\left(1 - \beta q\right) + \nu_I\left(H - 2h_n\right)}{2} \tag{3.12}$$

Proof: See the Appendix.

According to Proposition 1, switching strategies are played at equilibrium. Within the global game literature this is a common finding due to strategic complementarities in (dynamic) games. Player n has an incentive to move to higher actions as soon as the successor raises his strategy from s_{n+1} to $s'_{n+1} > s_{n+1}$. The cut-off points $x^*\left(a_{n-1}, h_n\right)$ specify the technology values at which user n is indifferent between the two technologies. These cut-offs depend on the immediate predecessor's observed action, on the technology history and on the expected behavior of the immediate successor, and due to condition (3.8) they belong to the region $[\underline{x}, \bar{x}]$. Proposition 1 yields $x^*\left(0, h_n\right) > x^*\left(1, h_n\right)$, so that when technology A is highly valuable relative to B –this occurs when a player observes a relatively high value for x_n– it is going to be adopted regardless of the predecessors' choices. On the other hand when technology B is more valuable,

[11] This region is symmetric around zero, i.e. the pivotal point for dominant actions in the absence of network benefits. Note that the gap $\bar{x} - \underline{x}$ depends positively on the network benefits. The fact that $\bar{x} - \underline{x}$ increases in both ν_D and $\nu_I H$ stems from the fact that when individuals attach a large positive value to joining a network, a high stand alone component is needed in order to adopt technologies regardless of other users' choices. As a result both $\bar{x}$ and $\underline{x}$ would move away from zero the larger are ν_D, ν_I and H.

player n is more likely to purchase technology A only if he observes his predecessor choosing A. Unlike condition (3.8), the inequality (3.10) does not play any role for the equilibrium uniqueness result. Thanks to condition (3.10), there exists at least one value of the technology process X within the two cut-off points. This rules out the uninteresting situation in which the technology values are such that the agents jump from one equilibrium to the other, outside the region $[x^*(1,h_n), x^*(0,h_n)]$, every few periods.

Consider the time during which all types fall into one of the dominance regions, say $x_n > \bar{x}$. In this case technology network benefits are not strong enough to observe A-lover individuals choosing the competing technology. When technology values fall into $[\bar{x}, \underline{x}]$ results from supermodular games allow to determine the unique equilibrium path. More specifically, when x_n is above $x^*(0,h_n)$ technology A is chosen regardless of the predecessors' actions (similarly, technology B is adopted for x_n below $x^*(1,h_n)$). Model uncertainty –captured by the technology value heterogeneity σ– plays a role in determining the width of the region $[x^*(1,h_n), x^*(0,h_n)]$ but not $[\bar{x}, \underline{x}]$. When the uncertainty about future types is very large, the interval $[x^*(1,h_n), x^*(0,h_n)]$ is less effective in refining the dominance regions, i.e. $\lim_{\sigma\to\infty}[\bar{x} - x^*(0,h_n)] = \lim_{\sigma\to\infty}[x^*(1,h_n) - \bar{x}] = 0$.[12]

The interval $[x^*(1,h_n), x^*(0,h_n)]$ generates hysteresis, since player n's choice depends on the predecessors' actions (as well as the expectation of the successor's action) whenever x_n falls into the hysteresis band, i.e. the gap $x^*(0,h_n) - x^*(1,h_n) = \bar{\sigma}$. In other words, when $x_n \in [x^*(1,h_n), x^*(0,h_n)]$ individual n's choice is determined by his predecessors' actions, and equilibrium adoption is path dependent. When $\bar{\sigma}$ is small the likelihood of hysteresis is reduced, i.e. the probability of simultaneously observing a technology value falling into the interval $[x^*(1,h_n), x^*(0,h_n)]$ and a user choosing based on the past history is small.

We can perform the following comparative statics over the hysteresis band: 1) it narrows with the discount rate β, 2) it widens with the direct network benefit ν_D and 3) it is unaffected by the indirect network benefit ν_I. High values for β mean that the importance of the future base is high relative to the installed base. In such

[12] Other things equal, this refining ability vanishes with the future base importance. When individual n neglects the impact of his successor's action, i.e. for values of β close to zero, the equilibrium selection criterion has no power in eliminating the multiplicity of equilibria within $[\bar{x}, \underline{x}]$.

a situation individuals tend to disregard the predecessor's action and the switching points get closer to each other. On the other hand an increase in the direct network externality ν_D would increase the importance of the network component relative to the stand alone value in the individuals' expected utility. In order to explain the third finding, suppose that h_n is the arithmetic mean and at time n one has $h_n > H/2$, i.e. more than a half of the predecessors have chosen A. Then an increase in ν_I would move both cut-offs downwards by the same amount. This reflects the fact that technology A is more attractive since it takes a lower technology value x_n to have individuals choosing A. These effects are summarized in figure 3.1-Panel A. Figure 3.1 shows a sample path for the random walk (3.1) with $N = 100$, $p = 1/2$, $x_0 = 0$, $\sigma = 0.5$, $\alpha = 4$ and $\nu_I = 0$; individual preferences are captured by $\beta = 0.5$ and $\nu_D = 5$. The hysteresis band narrows when the discount factor increases to $\beta = 0.95$ (see Panel A.2), and widens when the direct network benefit increases to $\nu_D = 6$ (see Panel A.3).

Note from (3.11,3.12) that when the technology value is expected to be constant through time, i.e. $p = 1/2$, the cut-off points are symmetric around $\nu_I (H - 2h_n)/2$, i.e. the pivotal point in the absence of direct network externalities. Further the cut-offs collapse to zero as the network externalities die out; in this case technologies are chosen in equilibrium depending on their stand alone value only. Similarly, when the technology value is expected to move away from its current value –because either $p > 1/2$ or $p < 1/2$–, the cut-off points converge to the pivotal point $\nu_I (H - 2h_n)/2$ as ν_D goes to zero, i.e. $\lim_{\nu_D \to 0} x(0) = \lim_{\nu_D \to 0} x(1) = \nu_I (H - 2h_n)/2$.

Finally, if p increases, the process for technology values is expected to move upwards with a higher probability as time goes by, so that there is a higher probability that the successor turns out to adopt technology A. As discussed in the previous section, this lowers the minimum level of stand alone value required for the current player to choose technology A. An increase in p thus decreases both the cut-off points by the same amount (similarly for a decrease in p).

3.4 Introducing Converters

3.4.1 Individual payoffs

We now consider the effect of converters enabling imperfect compatibility between technologies A and B. As in section 3.2, we assume perfect competition, implying a null price for the converter. Let $r \in (0,1)$ denote the compatibility of technology A with B (s is defined similarly). The stand alone component is not influenced by the existence of converters. Converters have an effect on the utility derived from networks, in that they allow agents to profit from the network even if no one else has chosen the same technology. Consider for example player n choosing technology A while both the previous and the next generations opt for technology B. Compatibility results in a payoff of $\nu_D r(1+\beta)$ contrasting with a null network benefit in the absence of converters. Similarly player n receives $\nu_D s(1+\beta)$ if he adopts technology B and both agent $n-1$ and $n+1$ choose A. Let $u_n^c = u^c(a_n, a_{n-1}, h_n, a_{n+1}, x_n)$ denote the individual payoff with compatibility.

Using the payoffs defined for the incompatibility case (see (3.3), (3.4) and (3.5)) gives:

$$u_n^c = \begin{cases} u(1, a_{n-1}, a_{n+1}, h_n, x_n) + \nu_D r((1-a_{n-1}) + \beta(1-a_{n+1})) & \text{if } a_n = 1 \\ u(0, a_{n-1}, a_{n+1}, h_n, x_n) + \nu_D s(a_{n-1} + \beta a_{n+1}) & \text{if } a_n = 0 \end{cases}$$

or equivalently:

$$u_n^c = u_n + \nu_D r a_n (1+\beta) + \nu_D (a_{n-1} + \beta a_{n+1})(s - a_n(r+s)) \tag{3.13}$$

Note that when both the compatibility levels are equal to zero one gets the $u_n^c = u_n$ and the previous results with incompatible technologies follow.

3.4.2 Equilibrium and interpretation

The equilibrium in the absence of external benefits ($\nu_D = \nu_I = 0$, see subsection 3.3.1) does not change with the introduction of converters as they act as network enhancers only. On the contrary, the equilibrium with $\nu_D > 0$ is different from the incompatibility case and is characterized in the following:

Proposition 2 *Let the model parameters satisfy the restriction (3.8). Then the technology adoption game displays dominance regions. Technology A is dominant if* $x_n \geq \bar{x}_c = \bar{x} - r\nu_D(1+\beta)/2$ *and technology B is dominant if* $x_n < \underline{x}_c = \underline{x} + s\nu_D(1+\beta)/2$.

The game admits a unique equilibrium in which for all n:

$$s^c(a_{n-1},h_n,x_n)=\begin{cases}1 & \text{if } x_n\geq x_c^*(a_{n-1},h_n)\\ 0 & \text{if } x_n< x_c^*(a_{n-1},h_n)\end{cases}$$

where the cut-off $x_c^*(a_{n-1},h_n)$ is decreasing in both a_{n-1} and h_n. Furthermore let the step size σ be sufficiently small according to:

$$\sigma<\bar{\sigma}_c=\bar{\sigma}-\frac{\nu_D\left[r(1-\beta q)+s(1-\beta p)\right]}{2} \tag{3.14}$$

Then the cut-off points are given by:

$$x_c^*(0,h_n) = x^*(0,h_n)-\frac{\nu_D(r-\beta ps)}{2} \tag{3.15}$$

$$x_c^*(1,h_n) = x^*(1,h_n)+\frac{\nu_D(s-\beta qr)}{2} \tag{3.16}$$

Proof. (Proposition 2) See the appendix. ■

In the presence of converters the dominance regions are affected by the probability of upward movements p and the network benefits ν_D and ν_I along the same lines as the incompatibility case. An increase in either ν_D or ν_I would move $\underline{x}_c$ and $\bar{x}_c$ towards the barriers α and $-\alpha$, thus shrinking the dominance regions. When technologies provide substantial network benefits, it takes higher stand alone values (and therefore higher technology values) in order to make individual opt without considering network externalities. Note that converters act in the opposite direction, that is an increase in either r or s (or both) would reduce the dominance regions. High values for r provide an individual opting for A with network payoffs even if other players choose B, thus making technology A more appealing and widening the region $[\bar{x}_c,\alpha]$. As it emerges comparing the dominance regions under compatibility and incompatibility (see Lemma 1), the presence of converters shrinks the interval in which no action is dominant, i.e. $[\underline{x}_c,\bar{x}_c]\subset[\underline{x},\bar{x}]$. The reason behind this is that when technologies are compatible the gains from coordinating (i.e. three generations choosing the same technology) are reduced, since player n profits from some network externalities even if he chooses a different technology relative to players $n-1$ and $n+1$. The same argument implies that the hysteresis band narrows with compatible technologies. In fact from eq. (3.14) one has $\bar{\sigma}_c<\bar{\sigma}$. Given the cut-off points in (3.15,3.16), we get the following comparative statics results:

1. when direct network payoffs are negligible, individuals switch around the pivotal point, namely $\nu_I (H - 2h_n)/2$.

2. an increase in the discount rate β reduces the hysteresis band along the same lines as in section 3.3.2. However, in the presence of converters this gap reduces less than in the case without converters.[13] This is because agents derive utility from the installed base even if they buy a technology that has not been chosen by the previous generation;

3. converters affect cut-off points in an asymmetric fashion: higher values of s increase both $x_c^* (0, h_n)$ and $x_c^* (1, h_n)$ while an increase in r would lower both the switching points. This finding has the following interpretation. Assume that, given predecessors' choices, player n is indifferent between the two technologies. Other things equal an increase in s makes technology B more attractive since user n achieves higher gains from compatibility with the competing technology A. As a consequence it would take a higher x_n, i.e. an individual that is relatively more prone to purchase technology A, to restore indifference. A similar reasoning would hold with respect to an increase in r. Despite this asymmetry note that an increase in either of the compatibility parameters would narrow the hysteresis band $\bar{\sigma}_c$. Partial converters bring in network benefits that can be reaped when other players purchase the competing technology. A higher level of the one-way compatibility s has a higher impact in the utility function when the predecessor chose A than when he chose B. In order to restore indifference, the cut-off for $a_{n-1} = 1$ increase more than the cut-off for $a_{n-1} = 0$ for a given technology history h_n, thus narrowing the hysteresis band. A similar argument applies to changes in r.

The main point here is that, other things equal, converters decrease direct network benefits from coordination on the same technology and dominance regions widen. Moreover, with symmetric converters ($s = r$) the cutoff points $x_c^* (0, h_n)$ and $x_c^* (1, h_n)$ are symmetric around $(H - 2h_n)/2$ as in the incompatibility case, and $[x_c^* (0, h_n), x_c^* (1, h_n)] \subset [x^* (0, h_n), x^* (1, h_n)]$. As a result hysteresis is less likely to occur with partial converters (contrast Panel A.1 with Panel B.1 in figure 3.1). On the other hand asymmetric

[13] From (3.14) one has $\partial \bar{\sigma}_c / \partial \beta = (rq + sp - 1) \nu_D / 2$, which is negative since $(p, r, s) \in (0, 1)^3$.

converters would drive individual choices towards the more compatible technology. In order to see this, compare Panel B.2 and B.3 in figure 3.1. Given a sample path for (3.1) and other exogenous parameters, the case of symmetric converters $r = s = 0.2$ is considered in Panel B.1, while asymmetric converters $r = 0.6$ and $s = 0.2$ (resp. $r = 0.2$ and $s = 0.6$) appear in Panel B.2 (resp. Panel B.3). Note that the level of overall compatibility $r + s$ is the same for both Panel B.2 and B.3, such that the difference $x_c^*(0, h_n) - x_c^*(1, h_n)$ does not change from one case to the other. Recall that $r > s$ corresponds to higher compatibility gains for technology A and note from Panel B.1 and B.2 that more users adopt technology A in the asymmetric converters case (similarly more users opt for technology B in Panel B.3 relative to Panel B.1).

3.5 Long run behavior and the expected time of adoption

3.5.1 Technology lock-in and limiting behavior

We investigate lock-in effects in our OLG setup where individuals explicitly take into account the actions of future generations when choosing between two competing technologies. As mentioned in the Introduction, in Arthur (1989) one technology emerges as dominant as time goes by. In other words there exists a time after which *all* players opt for the same technology. In our setup, due to the stochastic nature of the individual types, the emergence of a technology as dominant is related to: 1) the likelihood of the stochastic process for x_n hitting the barriers $\underline{x}$, $\bar{x}$, $x(0, h_n)$ and $x(1, h_n)$ (as well as their counterpart with converters) and 2) the impact of the underlying parameters on the mentioned barriers. As for the latter point, we have provided several comparative statics results in sections 3.3 and 3.4. The long run characterization of our adoption game is thus captured by the limiting behavior of the technology process (3.1-3.2) like in Kandori, Mailath and Rob (1993). Technology A locks in if and only if x_n is always above $x(0, h_n)$ for large n (similarly technology B locks in if and only if x_n is below $x(1, h_n)$). More formally, let the adoption probabilities of the two technologies be defined as $\pi_A = \lim_{n\to\infty} \Pr(x_n \geq x(0, 0))$ and $\pi_B = \lim_{n\to\infty} \Pr(x_n \leq x(1, H))$. Adoption probabilities with compatible technologies are defined similarly, i.e. with respect to the relevant cut-offs $x_c^*(0, 0)$ and $x_c^*(1, H)$, and denoted by $\pi_{A,c}$ and $\pi_{B,c}$. Then technologies lock in if and only if either $\pi_A = 1$ or $\pi_B = 1$ (lock-ins with compatible technologies are defined similarly).

Proposition 3 *No technology lock-in occurs in the long run, regardless of the presence of converters.*

Proof. (Proposition 3) See the appendix. ■

The idea behind Proposition 3 is that the process (3.1) hits any barrier with positive probability, regardless of the uncertainty about future values σ. As a consequence, no technology can emerge as dominant in the long run, and lock-ins can occur only temporarily.

We now consider the interplay between the parameters in our game and the long run probabilities of adopting each technology. This point is clearly related to the impact of initial parameters on the cut-offs (see sections 3.3 and 3.4), since changes in the cut-offs affect the probabilities of adoption π_A and π_B.

Corollary 1 *i) π_A and π_B (and their counterparts with compatible technologies) increase with β and decrease with both ν_D and ν_I; ii) π_A and $\pi_{A,c}$ increase with p (resp. π_B and $\pi_{B,c}$ decrease with p). iii) $\pi_{A,c}$ increases in r and decreases in s (resp. $\pi_{B,c}$ increases in s and decreases in r); iv) $\pi_{A,c} > \pi_A$ and $\pi_{B,c} > \pi_B$ if and only if $r/s \in \left(\beta p, \dfrac{1}{\beta(1-p)}\right)$. Moreover for $p = 1/2$, v) $\pi_A = \pi_B = \pi$ and $\pi_{A,c} = \pi_{B,c} = \pi_c$ if and only if $(s = r)$*

Proof. (Corollary 1) See the appendix. ■

According to property *ii*) an increase in p increases the likelihood of adopting technology A. This is intuitive since higher values for p mean that technology A is more valuable, in that it provides higher stand alone values. Similarly, an increase in compatibility level r makes technology A more valuable, and thus increases the probability it becomes dominant in the long run (see property *iii*). Note from *iv*) that converters increase the adoption probabilities if and only if the ratio r/s belongs to the above interval. At a first glance this might seem to be in sharp contrast with findings in Arthur (1989) , where converters always increase adoption probabilities. However note that in Arthur (1989) consumers receive payoffs (both stand alone and network benefits) only upon purchase. This would correspond $\beta = 0$ in our setup, such that property *v*) implies that converters increase the probability of adoption, alike Arthur (1989). Finally for $p = 1/2$ the long run probabilities for the process (3.1) are all equal across states and as a consequence the adoption probabilities coincide.

3.5.2 Expected time of adoption

We now consider the length of time the random walk (3.1) takes to move from one state to the other. The aim here is to determine the expected time to observe individuals switching from one technology to the other. Consider incompatible technologies and assume that at time n the technology value is immediately below $x^*(1, h_n)$. We know from Proposition 1 that player n would adopt technology B. The following period, player $n+1$ makes his choice comparing x_{n+1} with $x^*(0, h_{n+1})$. Since $h_{n+1} \leq h_n$ for $a_n = 0$, it follows that $x^*(0, h_{n+1})$ is always above x_{n+1}. Therefore technology B is chosen by player $n+1$, as well as the following generations until the process for technology values passes through $x^*(0, 0)$. From now onwards it is technology A to be chosen until X crosses $x^*(1, H)$ and so on. Given that x_n is immediately below $x^*(1, h_n)$ we define by $m_B = m_B(h_n)$ the expected number of adopters of technology B. m_B is therefore related to the average number of periods the random walk (3.1) takes to move from $x^*(1, h_n)$ to $x^*(0, 0)$. Similarly $m_A = m_A(h_n)$ is the expected number of agents choosing technology A given that the process x is immediately above $x^*(0, h_n)$, and corresponds to the average time the random process (3.1) takes to exit the band $[x^*(1, H), x^*(0, h_n)]$ after entering from $x^*(0, h_n)$. Knowledge of m_A and m_B is useful to determine how often we are likely to observe adopters switching from one technology to the other for a given technology history. With convertible technologies m_A^* and m_B^* are defined along the same lines and have a similar interpretation. Formulas to compute m_A and m_B (and their counterpart with compatibility) are given in the Appendix. One might be interested in determining how compatible technologies affect the average time of adoption. For a given set of parameters $(p, \sigma, \alpha, \beta, \nu_D, \nu_I, H)$ and convertibility values (r, s) we say that compatible technologies decrease the likelihood of switching -or equivalently increase path dependence- whenever $m_A^* > m_A$ and $m_B^* > m_B$.

Proposition 4 *The introduction of symmetric converters reduces path dependence whenever* $r/s \in \left(\beta p, \dfrac{1}{\beta(1-p)}\right)$.

Proof. (Proposition 4) See the appendix. ■

This finding follows from the fact that converters reduce the hysteresis band, which in turn implies that the expected adoption time for both technologies cannot increase. This will be relevant in an extended version of our model with sponsored technologies and firms setting technology prices based on the expected demand schedule.

3.6 Discussion

We now briefly discuss the impact of alternative assumptions on our equilibrium outcomes. First of all equilibrium uniqueness is preserved under a generalization of the expected payoffs in (3.7) . More specifically, linearity –albeit convenient– is not needed in order to preserve equilibrium existence, which is a consequence of expected payoffs exhibiting increasing differences (see Van Zandt and Vives (2003) and the proof of Proposition 1). A different specification of the payoffs would obviously imply different equilibrium cut-off points.

Second, the random walk specification (3.1-3.2) is not necessary to select a unique equilibrium. We could have used a different cumulative distribution function for the innovation ε –including for instance a continuous distribution with bounded support– as well as a random walk without barriers. We choose a binary distribution for the technology innovation for its simplicity and impose elastic barriers for the technology valuation in order to have bounded stand alone values like in the previous literature. What one needs is technology values to exhibit first order stochastic dominance, i.e. higher types of player n believe that successors are more likely to be of higher types as well. Third, an important feature that must be imposed on the game is that it displays dominance regions, such that one can apply an iterative dominance argument and select a unique equilibrium. This means that payoffs should be specified in such a way that for some technology values the actions chosen by other players (via the installed and future base), play no role in determining player n's choice.[14]

3.7 Conclusion

We have analyzed the technology adoption choices of agents characterized by a forward looking behavior. The existing literature does not encompass users getting utility from the purchased technology over their whole life time. Due to this assumption, agents take into consideration the installed base only, but do not form expectations of the future base. The OLG model allows us to consider agents that: 1) receive benefits in all periods of their permanence in a network and 2) take them into account when choosing the technology in the first period. Agents must therefore form expectations about future

[14] The model in Arthur (1989) does not belong to this class: it is not true for all n that one action is optimal no matter the technology value and the history. This happens because the stand alone value is bounded but the network value is unbounded and increasing in the actions of all the predecessors.

behavior. This feature, together with the strategic complementarities arising from the network effects, yields multiple equilibria in technology adoption. However, due to stochastic technology values, this game belongs to the class of Bayesian monotone games. We can, thus, build upon the Van Zandt and Vives (2003) which demonstrate that a greatest and least equilibrium exists for this particular class of games, and show that the equilibrium is unique. In essence, even within our dynamic setting the existence result of Van Zandt and Vives (2003) for static games still holds. Therefore, the main result is that an equilibrium exists and it is unique: agents, based on past observations, choose the technologies through a particular switching strategy. A second result is that lock in cannot occur in our setup. The intuition is that, given the thresholds of the equilibrium strategy, it is always possible to find a technology value for which in any point in time agents switch from the most used strategy to the least used one.

Further investigation could be done including agents living for more than two periods. In this case we would not expect the qualitative conclusions of our model to change. However this step could make the setup more realistic, in that at each point in time more than two generations are active in the market. Another promising direction is to introduce sponsored technologies produced by competing firms.

3.A Appendix

Proof. (Lemma 1) Let $\bar{\Delta}_n$ and $\underline{\Delta}_n$ denote respectively the value of Δ_n in (3.19) when both the users $n-1$ and $n+1$ choose the low and high action respectively, i.e. $\bar{\Delta}_n = \Delta(0,0,h_n,x_n)$ and $\underline{\Delta}_n = \Delta(1,1,h_n,x_n)$. From (3.19) one has:

$$\begin{aligned}\bar{\Delta}_n &= 2x_n - \nu_D(1+\beta) - \nu_I H \\ \underline{\Delta}_n &= 2x_n + \nu_D(1+\beta) + \nu_I H\end{aligned}$$

We now solve for the technology values making individuals indifferent between the two technologies. Let $\bar{x}$ (resp. $\underline{x}$) be the type that makes the individual n indifferent between choosing one of the two technologies when both the predecessor and the successor coordinate on technology A (resp. B), i.e. $\Delta(0,0,0,\bar{x})$ (resp. $\Delta(1,1,H,\underline{x})$). Using the above expressions for $\bar{\Delta}_n$ and $\underline{\Delta}_n$ one gets $\bar{x} = \frac{\nu_D(1+\beta)+\nu_I H}{2}$ and $\underline{x} = -\bar{x}$. As is clear, $\bar{x} > \underline{x}$ and one needs to check that $\alpha > \bar{x}$ for both $\bar{x}$ and $\underline{x}$ to belong to $(-\alpha,\alpha)$. This is equivalent to require that condition (3.8) holds true. ■

Proof. (Proposition 1) With finite action space, monotone strategies can be represented by the cut-off technology values $x(a_{n-1},h_n)$ where players switch from $a_n=0$ to $a_n=1$; moreover $x(.)$ is decreasing in a_{n-1} and h_n as in (3.9).

Step 1: existence of a greatest and least equilibrium in monotone strategies

Van Zandt and Vives (2003) (Theorem 1) set out minimal assumptions on agents' payoff function (u_n supermodular in a_n and with increasing differences in any pair of variables) as well as on beliefs (first order stochastic dominance) guaranteeing that there exist a greatest and least Bayes-Nash equilibrium in monotone strategies. As is clear, the assumption about player types is fullfilled by the random walk (3.1-3.2) that we use for the technology process. We now show that the remaining conditions as layed out by Van Zandt and Vives are fulfilled in our model. Let $U_n = U(a_n,a_{n-1},h_n,x_n;s_{n+1})$ be the expected utility (see (3.7)) of player n, when $n+1$ plays the monotone strategy s_{n+1} defined in (3.9). One can rewrite U_n as

$$\begin{aligned}U &= \Pr(x_{n+1} \geq x^*(a_n,h_{n+1})\,|x_n)\,u(a_n,a_{n-1},h_n,1,x_n) + \\ &\quad \Pr(x_{n+1} \leq x^*(a_n,h_{n+1})\,|x_n)\,u(a_n,a_{n-1},h_n,0,x_n)\end{aligned}$$

We now show that U_n has increasing differences in $(a_n, a_{n-1}), (a_n, h_n), (a_n, x_n)$ and (a_n, a_{n+1}).[15] To this end, we first show that u_n has the same properties.

Let $\Delta_n = \Delta(a_{n-1}, a_{n+1}, h_n, x_n)$ denote the difference in player n's payoff when he switches between technology A and B and $\Lambda_n = \Lambda(a_n, a_{n-1}, h_n, x_n)$ denote the difference in player n's payoff when player $n+1$ switches between technology A and B:

$$\Delta_n = u(1, a_{n-1}, a_{n+1}, h_n, x_n) - u(0, a_{n-1}, a_{n+1}, h_n, x_n) \tag{3.17}$$

$$\Lambda_n = u(a_n, a_{n-1}, 1, h_n, x_n) - u(a_n, a_{n-1}, 0, h_n, x_n) \tag{3.18}$$

Using the payoffs in (3.6) we have

$$\Delta_n = -\nu_D(1+\beta) - \nu_I + 2x_n + 2\nu_I h_n + 2\nu_D(a_{n-1} + \beta a_{n+1}) \tag{3.19}$$

$$\Lambda_n = \nu_D \beta (2a_n - 1) \tag{3.20}$$

Inspection of (3.19-3.20) reveals that our payoff structure displays the following properties:

i) Δ_n is increasing in a_{n-1}, a_{n+1}, h_n and x_n

ii) Λ_n depends only on player's n action, i.e. $\Lambda(a_n, a_{n-1}, h_n, x_n) = \Lambda(a_n)$, and $\Lambda(1) > 0 > \Lambda(0)$.

While property *i)* is in fact equivalent to u_n having increasing differences in the relevant pair of variables (see Van Zandt and Vives (2003), pg. 21), property *ii)* is essential to establish increasing differences of the expected utility, U_n, within a dynamic monotone Bayesian game like the one we consider.[16] This property implies that there is strong complementarity between player n's action and his successor's, i.e. not coordinating brings disutility. With increasing differences only, coordinating increases the payoff of the players. However, not coordinating might not bring about losses. Thus, by means of property *ii)*, player n has a strict incentive to coordinate with player $n+1$.

[15] In fact, increasing differences among any pair of arguments of U_n is equivalent to supermodularity in each argument since U_n is defined on the product of linearly ordered sets. Moreover it is straightforward to see that U_n is supermodular in a_n, which is another requirement for the monotonicity of the argmax of U_n.

[16] In fact, from property *i)* one only has that $\Lambda(1, a_{n-1}, h_n, x_n) > \Lambda(0, a_{n-1}, h_n, x_n)$.

Let $D_n = D(a_{n-1}, h_n, x_n; s_{n+1})$ be the difference in player n's expected payoff from taking the high and the low action when agent $n+1$ plays according to (3.9):

$$\begin{aligned} D(a_{n-1}, h_n, x_n; s_{n+1}) &= U(1, a_{n-1}, h_n, x_n; s_{n+1}) - U(0, a_{n-1}, h_n, x_n; s_{n+1}) \\ &= \Pr\left(x_{n+1} \geq x_{n+1}\left(0, h_{n+1}\right) | x_n\right) \times \Delta\left(a_{n-1}, 1, h_n, x_n\right) + \\ &\quad \Pr\left(x_{n+1} < x_{n+1}\left(1, h_{n+1}\right) | x_n\right) \times \Delta\left(a_{n-1}, 0, h_n, x_n\right) + \\ &\quad \Pr\left(x_{n+1}\left(0, h_{n+1}\right) > x_{n+1} \geq x_{n+1}\left(1, h_{n+1}\right) | x_n\right) \times \\ &\quad \left[u\left(1, 1, h_n, x_n\right) - u(0, 0, h_n, x_n)\right] \end{aligned}$$

Therefore, showing that U_n has increasing differences in any pair of relevant arguments is equivalent to showing that $D\left(a_{n-1}, h_n, x_n; s_{n+1}\right)$ is increasing in a_{n-1}, h_n, x_n and a_{n+1}.

Letting G be the cumulative distribution function of the increment ε in the stochastic process for the technology values (see (3.2)), one can rewrite D_n as

$$\begin{aligned} D_n &= \Delta\left(a_{n-1}, 1, h_n, x_n\right) + \Lambda\left(0\right) G\left(\frac{x_{n+1}\left(0, h_{n+1}\right) - x_n}{\sigma}\right) \\ &\quad -\Lambda\left(1\right) G\left(\frac{x_{n+1}\left(1, h_{n+1}\right) - x_n}{\sigma}\right) \end{aligned} \tag{3.21}$$

As is clear, D_n is independent of a_{n+1}. Moreover:

- D_n is increasing in x_n from properties $i)$ and $ii)$, and G decreasing in x_n
- D_n is increasing in a_{n-1}, from properties $i)$ and $ii)$. Furthermore $h_{n+1} = h\left(a_{n+1-M}, ..., a_{n-1}\right)$ is increasing in a_{n-1} and $x_{n+1}\left(a_n, h_{n+1}\right)$ is decreasing in h_{n+1}, such that G is decreasing in a_{n-1}.
- D_n is increasing in h_n from properties $i)$ and $ii)$ and from h_{n+1} increasing in h_n.[17]

Following the result of Van Zandt and Vives (2003), Theorem 1, The existence of a greatest and least equilibrium in monotone strategies follows from the supermodularity of the expected utility function and the existence of dominance regions. In fact,

[17] In fact, suppose that h_n increases because any of the actions $a_{n+1-M}, ..., a_{n-2}$ increase. Then, h_{n+1} increases as well. If the increase in h_n is due to a higher value for a_{n-M}, then h_{n+1} is not affected.

$D(a_{n-1}, h_n, x_n; s_{n+1}) < 0$ if $x_n < \underline{x}$ and $D(a_{n-1}, h_n, x_n; s_{n+1}) > 0$ if $x_n > \underline{x}$, which provides us with the starting point for the greatest and least best reply.[18]

Step 2: equilibrium uniqueness

The final step of the proof consists of showing that the greatest and least equilibrium of this game coincide. Another alternative would be to show that the best-reply mapping is a contraction as in Levin (2001). Define $\overline{x}^{\infty}(a_{n-1}, h_n)$ as the limit of the sequence of iterated deletion of dominated strategies when player n observes a_{n-1} and h_n. The threshold $\overline{x}^{\infty}(a_{n-1}, h_n)$ defines the infimum technology value above which player n chooses technology A:

$$D(a_{n-1}, h_n, \overline{x}^{\infty}(a_{n-1}, h_n); s_{n+1}) \leq 0 \tag{3.22}$$

Similarly, let $\underline{x}^{\infty}(a_{n-1}, h_n)$ be the supremum of x_n below which player n chooses technology B, i.e.

$$D(a_{n-1}, h_n, \underline{x}^{\infty}(a_{n-1}, h_n); s_{n+1}) \geq 0 \tag{3.23}$$

Thus all strategies for player n have been deleted but the ones such that

$$s(a_{n-1}, h_n, x_n) = \begin{cases} 1 \text{ if } x_n \geq \overline{x}^{\infty}(a_{n-1}, h_n) \\ 0 \text{ if } x_n < \underline{x}^{\infty}(a_{n-1}, h_n) \end{cases}$$

where $\underline{x}^{\infty}(a_{n-1}, h_n) \leq \overline{x}^{\infty}(a_{n-1}, h_n)$. For a unique equilibrium to exist we must have that $\underline{x}^{\infty}(a_{n-1}, h_n) = \overline{x}^{\infty}(a_{n-1}, h_n)$.

Suppose player n observes $a_{n-1} = 1$ and h_n. Then (3.22) becomes

$$\pi_s(1, \overline{x}^{\infty}(1, h_n)) + \pi_D(1,1) + \pi_I(1, h_n) + \beta E[\pi_D(1, a_{n+1}) | \overline{x}^{\infty}(1, h_n)] \leq$$
$$\pi_s(0, \overline{x}^{\infty}(1, h_n)) + \pi_I(0, h_n) + \beta E[\pi_D(0, a_{n+1}) | \overline{x}^{\infty}(1, h_n)] \ ,$$

that is

$$\alpha + \overline{x}^{\infty}(1, h_n) + \nu_D + \nu_I h_n + \beta p \nu_D \leq \alpha - \overline{x}^{\infty}(1, h_n) + \nu_I(H - h_n) + \beta \nu_D \ .$$

[18] For a graphical representation of the iterated elimination of dominated strategies see Frankel, Morris and Pauzner (2001).

Solving the latter inequality for $\overline{x}^{\infty}(1,h_n)$ gives

$$\overline{x}^{\infty}(1,h_n) \leq \frac{-\nu_D(1-\beta p)+\nu_I(H-2h_n)}{2}.$$

The inequality (3.23) yields

$$\alpha+\underline{x}^{\infty}(1,h_n)+\nu_D+\nu_I h_n+\beta p\nu_D \geq \alpha-\underline{x}^{\infty}(1,h_n)+\nu_I(H-h_n)+\beta\nu_D$$

such that

$$\underline{x}^{\infty}(1,h_n) \geq \frac{-\nu_D(1-\beta p)+\nu_I(H-2h_n)}{2}$$

Since $\underline{x}^{\infty}(1,h_n) \leq \overline{x}^{\infty}(1,h_n)$ it follows that

$$\underline{x}^{\infty}(1,h_n)=\overline{x}^{\infty}(1,h_n)=x^*(1,h_n)$$

Following the same procedure for $a_{n-1}=0$, we get

$$\underline{x}^{\infty}(0,h_n)=\overline{x}^{\infty}(0,h_n)=x^*(0,h_n)$$

where $x^*(0,h_n)$ is given in the main text. Thus a unique equilibrium exists. ■

Proof. (Proposition 2). The incremental payoffs for the compatibility case $\Delta_{n,c}$ and $\Lambda_{n,c}$ are defined along the same lines of (3.17,3.18). Using (3.13) they can be written in terms of their counterpart under incompatibility:

$$\Delta_{n,c} = \Delta_n+\nu_D r(1+\beta)-\nu_D(r+s)(a_{n-1}+\beta a_{n+1}) \tag{3.24}$$

$$\Lambda_{n,c} = \Lambda_n+\beta\nu_D[s-(r+s)a_n] \tag{3.25}$$

The dominance regions can be determined as before since:

$$\begin{aligned}\bar{\Delta}_{n,c} &= \Delta_c(0,0,0,x_n)=\bar{\Delta}_n+r\nu_D(1+\beta)\\ \underline{\Delta}_{n,c} &= \Delta_c(1,1,H,x_n)=\underline{\Delta}_n-s\nu_D(1+\beta)\end{aligned}$$

which yield $\bar{x}_c$ and $\underline{x}_c$ as in the main text. Since $\bar{x}_c<\bar{x}$ (and similarly $\underline{x}_c>\underline{x}$) it follows that condition (3.8) is sufficient to guarantee that the adoption game displays dominance regions with compatible technologies.

From (3.24-3.25) one has that $\Delta_{n,c}$ is increasing in a_{n-1}, a_{n+1}, h_n and x_n due to property $i)$ for Δ_n and $r+s<2$. Moreover

$$\Lambda_c\left(a_n,1,h_n,x_n\right)=\beta\nu_D\left(1-r\right)>0>-\beta\nu_D\left(1-s\right)=\Lambda_c\left(a_n,0,h_n,x_n\right)$$

so that property $ii)$ holds. In order to show that the expected payoff U_c has increasing differences in any pair of variables, we have from (3.21) that

$$\begin{aligned} D_{n,c} &= \Delta_{n,c}\left(a_{n-1},1,h_n,x_n\right)+\Lambda_c\left(0\right)G\left(\frac{x_{n+1}\left(0,h_{n+1}\right)-x_n}{\sigma}\right) \\ &\quad -\Lambda_c\left(1\right)G\left(\frac{x_{n+1}\left(1,h_{n+1}\right)-x_n}{\sigma}\right) \end{aligned} \tag{3.26}$$

It is easy to see that the $D_{n,c}$ is increasing in x_n, a_{n-1}, h_n and independent of a_{n+1} from the analysis for D_n and the expressions (3.24-3.25). So there exists a greatest and least equilibrium in monotone strategies.

Uniqueness follows closely from the proof of Proposition 1.

Suppose that player n observes $a_{n-1}=1$ and h_n. The infimum technology value above which he chooses technology A is given by

$$D_c(a_{n-1},h_n,\overline{x}_c^{\infty}\left(1,h_n\right);s_{n+1})\leq 0$$

that is

$$\overline{x}_c^{\infty}\left(1,h_n\right)\leq\frac{-\nu_D\left[1-s-\beta q\left(1-r\right)\right]+\nu_I\left(H-2h_n\right)}{2}$$

The supremum of x_n below which player n chooses technology B is given by

$$D_c(a_{n-1},h_n,\underline{x}_c^{\infty}\left(1,h_n\right);s_{n+1})\geq 0,$$

yielding

$$\underline{x}_c^{\infty}\left(1,h_n\right)\geq\frac{-\nu_D\left[1-s-\beta q\left(1-r\right)\right]+\nu_I\left(H-2h_n\right)}{2}$$

and uniqueness follows from $\underline{x}_c^{\infty}\left(1,h_n\right)\leq\bar{x}_c^{\infty}\left(1,h_n\right)$. Applying the same reasoning for the case $a_{n-1}=0$ gives the cutoff $x_c^*\left(0,h_n\right)$ in the main text. ■

Proof. (Proposition 3). Let S denote the state space for the process x. Without loss of generality we assume $\{0,\alpha\}\in S$ and let $\bar{n}$ denote the number of steps the process takes to move from 0 to α, i.e. $\bar{n}=\alpha/\sigma$. Therefore the (finite) state space S comprises

$2\bar{n}+1$ elements and is given by $S = \{-\alpha, -(\alpha-\sigma), ..., -\sigma, 0, \sigma, ..., \alpha\}$. In what follows $s_i \in S$ denotes the i-th state, where the subscript i is an integer between 0 and $2\bar{n}$, i.e. $s_0 = -\alpha, ..., s_{\bar{n}} = 0, ..., s_{2\bar{n}} = \alpha$. We apply results for Markov chains to the random walk (3.1). The one-step transition matrix is given by:

$$P = \begin{bmatrix} p & q & 0 & \dots & 0 & 0 & 0 \\ q & 0 & p & & 0 & 0 & 0 \\ & \vdots & & \ddots & & \vdots & \\ 0 & 0 & 0 & & q & 0 & p \\ 0 & 0 & 0 & \dots & 0 & q & p \end{bmatrix} \tag{3.27}$$

with $p + q = 1$. Given that at time n the Markov chain is in state s_i, P_{ij} gives the probability that the Markov chain moves to state s_j next period, i.e. $P_{ij} = \Pr(x_{n+1} = s_j | x_n = s_i)$ for $n \in N$. The Markov chain described by (3.27) is irreducible, aperiodic and regular. It follows that the n-step ahead matrix P^n converges as $n \to \infty$ to a positive matrix $\Pi = \mathbf{1}^\top \boldsymbol{\pi}$ (Çinlar (1975), Corollary 2.11).[19] The probability vector $\boldsymbol{\pi}$ is the unique solution to:

$$\boldsymbol{\pi} P = \boldsymbol{\pi} \tag{3.28a}$$

$$\boldsymbol{\pi}\mathbf{1}^\top = 1 \tag{3.28b}$$

$$\pi_j > 0, \; j = 0, ..., 2\bar{n} \tag{3.28c}$$

Since $\pi_j > 0, \forall j$, each element in the limiting matrix Π is strictly positive. This means that every state can occur with a positive probability as $n \to \infty$ and as a consequence technology lock-ins are ruled out in our game. This result is not affected by the compatibility between the two technologies, since the stationary distribution $\boldsymbol{\pi}$ is driven by the random process (3.1) only. ■

[19] In what follows boldface characters denote row vectors; for example $\mathbf{1}$ is the $1 \times (2\bar{n}+1)$ unity vector. π_j is the probability to reach state s_j as time goes to infinite, i.e. $\pi_j = \lim_{n\to\infty} P_{ij}^{(n)}$.

Proof. (Corollary 1). First of all we proceed in determining the probability vector $\boldsymbol{\pi}$. Using the transition matrix (3.27) the constraints in $(3.28a, 3.28b)$ become:[20]

$$\begin{gathered}
\pi_0 q + \pi_1 q = \pi_0 \\
\pi_0 p + \pi_2 q = \pi_1 \\
\pi_{i-1} p + \pi_{i+1} q = \pi_i, i = 2, ..., 2\bar{n} - 2 \\
\pi_{2\bar{n}-2} p + \pi_{2\bar{n}} p = \pi_{2\bar{n}} \\
\textstyle\sum_{i=0}^{2\bar{n}} \pi_i = 1
\end{gathered}$$

Letting $\rho = \frac{p}{q}$ the above system may be rewritten as:

$$\pi_1 = \rho \pi_0 \tag{3.29a}$$

$$\pi_i = \pi_{i-1} \rho, i = 2, ..., 2\bar{n} - 1 \tag{3.29b}$$

$$\pi_{2\bar{n}} = \pi_{2\bar{n}-1} \rho \tag{3.29c}$$

$$\textstyle\sum_{i=0}^{2\bar{n}} \pi_i = 1 \tag{3.29d}$$

By recursive substitution[21] one has $\pi_i = \pi_1 \rho^{i-1} = \pi_0 \rho^i$ for $i = 2, ..., 2\bar{n} - 1$, and $\pi_{2\bar{n}} = \pi_1 \rho^{2\bar{n}-1} = \pi_0 \rho^{2\bar{n}}$ such that $\sum_{i=0}^{2\bar{n}} \pi_i = \pi_0 \sum_{i=0}^{2\bar{n}} \rho^i$. Thus:

$$\begin{array}{ll}
\pi_0 = (2\bar{n} + 1)^{-1} & \text{if } p = \frac{1}{2} \\
\pi_0 = \left(\frac{1 - \rho^{2\bar{n}+1}}{1 - \rho} \right)^{-1} & \text{if } p \neq \frac{1}{2}
\end{array} \tag{3.30}$$

and the stationary distribution $\boldsymbol{\pi}$ is given by (3.30,3.29a-3.29c). Recall that by construction, for a given step size σ, the barriers $(\alpha, -\alpha)$ coincide with admissible states for the random walk process $(s_{2\bar{n}}, s_0)$. On the other hand this might not occur for the cut-off points $x(0,0)$ and $x(1,H)$. Before determining the adoption probabilities, one has to determine the states for the random walk process in (3.1) corresponding to the cut-offs. We therefore consider s_{n_0} as the nearest state to $x(0,0)$, i.e. $n_0 = \{\min n \in (0, 2\bar{n}) : s_{n_0+1} > x(0,0)\}$. Similarly s_{n_1} is defined by $n_1 = \{\max n \in (0, 2\bar{n}) : s_{n_1-1} < x(1,H)\}$ As is clear from $x^*(0,0) > x^*(1,H)$ it follows that $n_1 < n_0$. It follows that the adoption probabilities are $\pi_A = \sum_{k=n_0}^{2\bar{n}} \pi_k$ and $\pi_B = \sum_{k=0}^{n_1} \pi_k$, or equivalently (adoption

[20] For the vector $\boldsymbol{\pi}$ to be uniquely determined as a solution for (3.28a,3.28b) one equation in the system (3.28a) is redundant (or equivalently, the determinant of $(I - P)$ needs to be null). Thus we drop the $(2\bar{n} - 1)$-th equation in (3.28a).

[21] When $\bar{n} = 1$, i.e. the state space for x is $S = \{-\alpha, 0, \alpha\}$, equations (3.29b) are not defined. The stationary distribution is defined by equations (3.29a,3.29c,3.29d) only.

probabilities with compatible technologies are defined similarly, considering $n_{0,c}$ and $n_{1,c}$ instead of n_0 and n_1):

$$\begin{array}{c} \pi_A = \pi_0 \sum_{k=n_0}^{2\bar{n}} \rho^k \\ \pi_B = \pi_0 \sum_{k=0}^{n_1} \rho^k \end{array} \tag{3.31}$$

As it is obvious from (3.31), the probability of adopting technology A in the long run is decreasing in n_0, and similarly the probability of adopting technology B is increasing in n_1. For $p = 1/2$ one has $\rho = 1$ and $\pi_0 = (2\bar{n}+1)^{-1}$ from (3.30). The adoption probabilities (3.31) become $\pi_A = \dfrac{2\bar{n} - n_0 + 1}{2\bar{n}+1}$ and $\pi_B = \dfrac{n_1+1}{2\bar{n}+1}$. On the other hand for $p \neq 1/2$ equations (3.30,3.31) give:

$$\begin{aligned} \pi_A &= \left(1 - \rho^{2\bar{n}+1}\right)^{-1} \left(\rho^{n_0} - \rho^{2\bar{n}+1}\right) \\ \pi_B &= \left(1 - \rho^{2\bar{n}+1}\right)^{-1} \left(1 - \rho^{n_1+1}\right) \end{aligned}$$

From the expression for the cut-offs in (3.11,3.12,3.15,3.16) one has

$$\begin{aligned} x^*(0,0) &= \frac{\nu_D(1-\beta p) + \nu_I H}{2} \; ; \; x^*(1,H) = -\frac{\nu_D(1-\beta q) + \nu_I H}{2} \; ; \\ x^*_c(0,0) &= \frac{\nu_D(1-r-\beta p(1-s)) + \nu_I H}{2} \; ; \; x^*_c(1,H) = -\frac{\nu_D(1-s-\beta q(1-r)) + \nu_I H}{2} \end{aligned}$$

Therefore properties (*i-iii*) easily follows taking the appropriate derivatives in the above expressions. Finally for $r/s \in \left(\beta p, \dfrac{1}{\beta(1-p)}\right)$, the cut-offs with compatible technologies are within $[x^*(1,H), x^*(0,0)]$, and v) obtains. Note that the above interval is always non-empty since $\dfrac{1}{\beta(1-p)}$ is strictly bigger than βp. For $p = 1/2$ the cut-offs $x^*(0,0)$ and $x^*(1,H)$ are symmetric around the origin, i.e. $x^*(0,0) = -x^*(1,H)$, it follows that $n_0 = 2\bar{n} - n_1$ which gives the first claim in property (v). Similarly $x^*_c(0,0) = -x^*_c(1,H)$ if and only if $s = r$. ■

Proof. (Proposition 4). For each state s_j let τ_j be the (function giving the) number of times that the process is in state s_j. The expected first passage time from state s_i to state s_j is given by $m_{ij} = E_i(\tau_j)$. For the Markov chain (3.27) the matrix $M = \langle m_{ij} > 0 \rangle$ is given by (see Kemeny and Snell (1976), chapter VII):

$$m_{ij} = \begin{cases} \frac{1}{\pi_i} & i = j \\ (4\bar{n}+1-i)\,i - (4\bar{n}+1-j)\,j & i > j \\ j(j+1) - i(i+1) & i < j \end{cases} \quad \text{if } p = \frac{1}{2} \tag{3.32}$$

$$m_{ij} = \begin{cases} \frac{1}{\pi_i} & i = j \\ \frac{1}{2p-1}\left(\frac{\rho^{2\bar{n}+1}\left(\rho^{-j}-\rho^{-i}\right)}{\rho-1} + j - i\right) & i > j \quad \text{if } p \neq \dfrac{1}{2} \\ \frac{1}{2p-1}\left(j - i - \frac{\rho^j-\rho^i}{(\rho-1)\rho^{i+j}}\right) & i < j \end{cases} \tag{3.33}$$

Fix a pair $(i,j) \in [1,2\bar{n}] \times [1,2\bar{n}]$ with $i > j$, such that m_{ij} is below the diagonal of M. Take another pair (k,l) with $k > l$, $k \geq i$ and $l \leq j$. Using (3.32,3.33) the difference $m_{kl} - m_{ij}$ becomes:

$$m_{kl} - m_{ij} = \begin{cases} (4\bar{n}+1)(k-i+j-l) - \left(k^2-l^2\right) + \left(i^2-j^2\right) & \text{if } p = 1/2 \\ \frac{1}{2p-1}\left[\frac{\rho^{2\bar{n}+1}\left(\rho^{-i}-\rho^{-k}+\rho^{-l}-\rho^{-j}\right)}{\rho-1} - (k-i+j-l)\right] & \text{if } p \neq 1/2 \end{cases}$$

Taking $k = i$ and $j > l$ gives:

$$m_{il} - m_{ij} = \begin{cases} [4\bar{n}+1-(j+l)]\,(j-l) & \text{if } p = 1/2 \\ \frac{1}{2p-1}\left[\frac{\rho^{2\bar{n}+1}\left(\rho^{-l}-\rho^{-j}\right)}{\rho-1} + l - j\right] & \text{if } p \neq 1/2 \end{cases}$$

Taking $k > i$ and $j = l$ gives:

$$m_{kj} - m_{ij} = \begin{cases} [4\bar{n}+1-(k+i)]\,(k-i) & \text{if } p = 1/2 \\ \frac{1}{2p-1}\left[\frac{\rho^{2\bar{n}+1}\left(\rho^{-i}-\rho^{-k}\right)}{\rho-1} + i - k\right] & \text{if } p \neq 1/2 \end{cases}$$

It follows that $m_{il} > m_{ij}$ for $j > l$ and $m_{kj} > m_{ij}$ for $k > i$. Therefore m_{ij} decreases with j and increases with i below the main diagonal.

Now consider: 1) the pair $(i,j) \in [1,2\bar{n}] \times [1,2\bar{n}]$ with $i < j$, i.e. m_{ij} is above the diagonal of M, and 2) the pair (k,l), $k < l$ with $k \leq i$ and $l \geq j$. From (3.32,3.33) one has:

$$m_{kl} - m_{ij} = \begin{cases} (l-j+i-k) + \left(l^2-k^2\right) - \left(j^2-i^2\right) & \text{if } p = 1/2 \\ \frac{1}{2p-1}\left[(l-j+i-k) - \frac{\rho^{j+l}\left(\rho^i-\rho^k\right)+\rho^{i+k}\left(\rho^l-\rho^j\right)}{(\rho-1)\rho^{i+j+k+l}}\right] & \text{if } p \neq 1/2 \end{cases}$$

Proceeding as before consider $k = i$ and $j < l$:

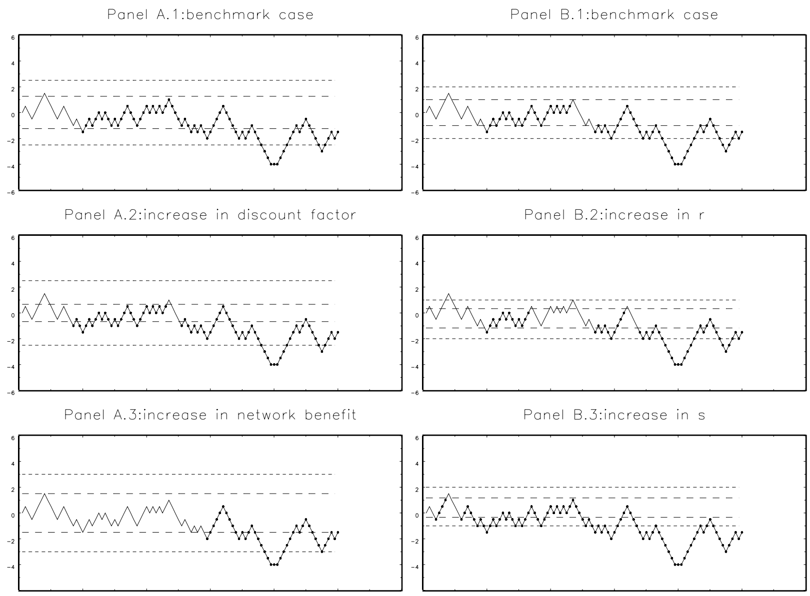
$$m_{il} - m_{ij} = \begin{cases} (l-j)\,(1+j+l) & \text{if } p = 1/2 \\ \frac{1}{2p-1}\left(l - j - \frac{\rho^l-\rho^j}{(\rho-1)\rho^{j+l}}\right) & \text{if } p \neq 1/2 \end{cases}$$

When $k < i$ and $j = l$ one has:

$$m_{kj} - m_{ij} = \begin{cases} (i-k)(1+i+k) & \text{if } p = 1/2 \\ \frac{1}{2p-1}\left(i - k - \frac{\rho^i - \rho^k}{(\rho-1)\rho^{i+k}}\right) & \text{if } p \neq 1/2 \end{cases}$$

Therefore m_{ij} increases in j and decreases in i above the main diagonal. For a given model parametrization, the cut-off points are obtained using (3.11,3.12). Let these states be denoted by s_u and s_l respectively for the upper $(x^*(0,0))$ and the lower $(x^*(1,H))$ cut-off. Note that $s_l < s_u$ since $l < u$. Similarly let s_u^* and s_l^* be the corresponding states in the presence of convertible technologies obtained via (3.15,3.16). It follows that the expected number of adopters is $m_A = m_{ul}$ and $m_B = m_{lu}$. We know that converters reduce the hysteresis band, i.e. $s_u^* - s_l^* < s_u - s_l$. Moreover it is easy to show that $[x_c^*(1,H), x_c^*(0,0)] \subset [x^*(1,H), x^*(0,0)]$ whenever $r/s \in \left(\beta p, \dfrac{1}{\beta(1-p)}\right)$. It follows that it cannot occur simultaneously that: 1) m_{lu}^* is above and to the right of m_{lu} and 2) m_{ul}^* is below and to the left of m_{ul}. Since m_{lu} and m_{ul} are respectively above and below the main diagonal, it follows that it cannot be that $m_A^* > m_A$ and $m_B^* > m_B$ simultaneously. ■

FIGURE 3.1. Technology adoption pattern.



Several technology adoption patterns are depicted. Panels A.1,A.2 and A.3 refer to incompatible technologies, while Panels B.1,B.2 and B.3 consider compatible technologies. In each Panel the solid line corresponds to one sample path for the process x_n. Black circles indicate individuals choosing technology B. The dashed lines display the cut-off points and the small-dashed lines the dominance regions limits. The following parameterization has been used: $N = 100$, $p = 1/2$, $x_0 = 0$, $\sigma = 0.5$ and $\alpha = 4$ (technology value process); $\beta = 0.5$ and $\nu = 5$ (individual preferences); $r = s = 0.2$ (compatibility values). Panel A.1 (resp. B.1) is the benchmark case with incompatible technologies (resp. compatible technologies). Panels A.2,A.3,B.2 and B.3 carry out some comparative statics exercises. Panel A.2 considers an increase in the individual discount factor ($\beta = 0.95$); Panel A.3 displays the effects of an increase in the network benefit ($\nu = 6$). Panels B.2 and B.3 consider asymmetric converters ($r = 0.6$ and $s = 0.6$ respectively).

References

Adner, R. (2003), A demand based view of technology life-cycles, mimeo, Insead

Arthur, W.B. (1989), Competing technologies, increasing returns, and lock-in by historical events, *The Economic Journal*, 99, 116-131

Athey, S. (2001), Single crossing properties and the existence of pure strategy equlibria in games of incomplete information, *Econometrica*, 69, 861-890.

Bresnahan, T. and S. Greenstein (1999) Technological competition and the structure of the computer industry", *Journal of Industrial Economics*, 68, 1-40

Burdzy, K., Frankel, D. and A. Pauzner (2001), Fast equilibrium selection by rational players living in a changing world, *Econometrica*, 68, 163-190

Choi, J.P. (1994), Irreversible choice of uncertain technologies with network externalities, *RAND Journal of Economics*, 25(3), 382-401

Çinlar, E. (1975), Introduction to stochastic processes, Prentice-Hall, New Jersey

Clements, M. and H. Ohashi, Indirect network effects and the product cycle: U.S. video games, 1994 - 2002, *The Journal of Industrial Economics*, forthcoming

Colla, P. and F. Garcia (2004), Technology adoption with forward-looking agents, *CORE Discussion paper,* 43/2004

Farrell, J. and G. Saloner (1992), Converters, compatibility, and the control of interfaces, *Journal of Industrial Economics*, 40(1), 9-35

Farrell, J. and G. Saloner (1985), Standardization, compatibility and innovation, *RAND Journal of Economics*, 16(1), 70-83

Frankel, D. and A. Pauzner (2000), Resolving indeterminacy in dynamic settings, the role of shocks, *Quarterly Journal of Economics*, 115, 283-304

Frankel, D. (2001), Payoff shocks and equilibrium selection with endogenous flexibility, mimeo, Tel Aviv University

Gandal, N., Greenstein, S. and D. Salant (1999) Adoptions and Orphans in the early microcomputer market, *The Journal of Industrial Economics*, 47, 87-105

Giannitsarou, C. and F. Toxvaerd (2003), Recursive global games, mimeo, Universidade Nova de Lisboa

Kandori, M., Mailath, G. J. and R. Rob (1993), Learning, mutation and long run equilibria in games, *Econometrica*, 61(1), 29-56

Katz, M. and C. Shapiro (1986), Technology adoption in the presence of network externalities, *The Journal of Political Economy*, 94, 4, 822-841

Kemeny, J.G. and J.L. Snell (1976), Finite Markov chains, Springer-Verlag, New York

Levin, J. (2001), A note on global equilibrium selection in overlapping generations games, mimeo, Stanford University

Liebowitz, S. and S.E. Margolis (1994), Network externality, an uncommon tragedy, *Journal of Economic Perspectives*, 8, 133-150

Liebowitz, S. and S.E. Margolis (1995), Path dependence, lock-in and history, *Journal of Law, Economics and Organization*, 11(1), 205-226

Macskasi, Z. (2002), Customer lock-in with long term contracts, mimeo, Northwestern University

Malin, E. (2003), Endogenous timing and intensity of inertia effect in a bandwagon equilibrium, mimeo, Université de Rennes 1

Morris, S. and H. S. Shin (2003), Global games, theory and applications, in *Advances in Economics and Econometrics (Proceedings of the Eight World Congress of the Econometric Society)* (M. Dewatripont, L. Hansen and S. Turnovsky Eds.)

Ochs, J. and I.-U. Park (2004), Overcoming the coordination problem, dynamic formation of networks, mimeo, University of Pittsburgh

Oyama, D. (2003), Booms and slumps in a game of sequential investment with the changing fundamentals, *Japanese Economic Review*,55 (3), 311-320

Rohlfs, J. (1974), A theory of interdependent demand for a communication service, *Bell Journal of Economics*, 5, 16-37

Shy, O. (1996), Technology revolutions in the presence of network externalities, *International Journal of Industrial Organization*, 14, 785-800

Van Zandt, T. and X. Vives (2003), Monotone equilibria in bayesian games of strategic complementarities, *CEPR Discussion Papers* 4103

4

WHEN DOES A MONOPOLY IMPROVE ITS QUALITY IN A NETWORK INDUSTRY?

4.1 Introduction

The objective of this paper is to analyze the incentive of a monopolist to introduce a higher quality brand in the market, when he sells already a low quality one and network effects are present. Frequently firms obtain the necessary technological developments in order to offer a higher quality good to the consumers. However, whether to introduce this new quality in the market is a strategic decision, especially if network externalities on consumption are present.

We consider a monopolist that produces a low quality good in two periods, benefiting from network externalities from first period consumption. In the second period he has to decide whether or not to introduce a high quality good in the market. In this context, the monopolist is able to create through first period pricing an installed base of consumers that secures him a high demand in the second period. He then sets the monopoly price over this higher demand. If a high quality product comes to compete with the low quality one in the second period, it could cannibalize the benefits resulting from the installed base for the low quality product. However, in that situation, the monopolist is able to offer a higher quality and charge a higher price. The optimal solution for the monopolist in terms of introducing or not the high quality depends on how important are the network effects relative to the quality differential. In this article I uncover the necessary and sufficient conditions under which a monopolist 1) introduces the high quality good in the market, 2) stops selling the low quality, 3) sells both qualities, 4) disregards the quality improvement. While doing so, I discuss the role of introductory pricing in markets when networks are present. The main result is that quality improvement takes place if and only if the intensity of the consumer' preference for the network is smaller than the quality differential of the goods.

This analysis is fit to study goods with delayed network externalities, i.e. goods whose network takes at least one period of time to become effective.[1] Many network industries exhibit a delay in the creation of the network. Early buyers do not benefit from the network as it takes some period of time in order to be effective. As Bensaid and Lesne (1996) argue, the delay might reflect the existence of a *learning by doing* externality or a *word of mouth* process. In the first case, late consumers benefit from improved versions of the initial product, which become available given that some consumers initially used the good and tested it. In the second case, consumers are not completely informed of the qualities of the product and the existence of an installed base reduces the search cost of obtaining this information. Vettas (1997), also presents a model where agents learn from other consumers the quality level of the product, such that the network effect is lagged in time. The software industry can be used as an example of an industry where the network effects are delayed. First, the software industry is characterized by the presence of delayed network externalities: early buyers do not benefit from the network both because technical assistance is less expanded in the beginning and initial versions of the product are more subject to bugs and defaults. Late buyers benefit from both these externalities. Second, the software industry is characterized by frequent innovations and quality improvements, not always compatible with the earlier versions of the product.

The problem of the trade off between technology improvement and network effect has been treated previously by several authors. Farrell and Saloner (1986), Katz and Shapiro (1992), Shy (1996) and Fudenberg and Tirole (2000) are all concerned with the problem of quality improvement. However, they do not use the vertical differentiation model to represent how the trade off between technology improvement and network effect operates.[2] The above authors all assume that consumers have the same preferences.[3] This implies that as soon as the new quality becomes available all consumers

[1] Bensaid and Lesne (1996) have the same modelization for the network effect, while Cabral et al. (2000) focus on immediate (or communication) networks.

[2] Baake and Boom (1999) and Bental and Spiegel (1995) have also combined the use of a vertical product differentiation model with network externalities. However, Baake and Boom (1999) focus on compatibility decisions, without addressing the quality improvement problem. Similarly Bental and Spiegel (1995) analyse the market coverage issue when network effects are present in a static framework.

[3] In fact, Fudenberg and Tirole (2000) assume the existence of two types of consumers; however, they are homogeneous in their preferences for the technological improvement and differ only in their stand-alone values.

(or none) adopt it, depending on the magnitude of the network effect at the time of innovation. By contrast, in my approach it is assumed that consumers are heterogeneous. Furthermore, these authors concentrate on the problem of a duopolistic market structure, whereas I analyze a monopolistic structure in which the innovation is available for sale by the monopolist and not by a potential entrant. In this respect, this paper relates to the literature on multimarket monopolist and the trade off between cannibalization and surplus extraction.[4]

When a good generates network effects, the consumers' willingness to pay is increasing with the number of users. The consumers that buy in initial periods do not only have an effect on immediate profit, but they also enhance firm's future profit. Setting initial low prices, i.e. practising an introductory pricing strategy, can be an optimal in order to obtain increased profits in subsequent periods. Bensaid and Lesne (1996) address this issue in a model where they test the so-called Coase conjecture on durable goods pricing in a model with network externalities. They find that a monopolist has incentive to pursue an increasing-through-time path of prices, contrary to what Coase conjectured.[5]

Cabral et al.(2000) also address the issue of introductory pricing in a model where network externalities are a necessary and sufficient condition for this pricing strategy to be optimizing. In the present article, I conclude that introductory pricing always takes place. However I distinguish between the introductory pricing where the seller maximizes the network constituted in the first period (by setting null price) and the situation where the path of prices is increasing but the initial price is positive. I show that if the network effect is mild, the loss from setting the initial price to zero is not compensated by the increase in the willingness to pay of consumers in the second period. The firm might just prefer to choose the monopoly price in both periods rather than concentrating profit in second period. Nevertheless the path of prices is increasing because the existence of some installed base (even if not maximal) increases the second period demand. I also show that for high network effects, setting initial prices to zero is always the optimal strategy, regardless of the introduction or not of the high quality good.[6]

[4] There is an extensive literature on cannibalization in business strategy.

[5] The present article contributes to this discussion allowing for the presence of the high quality good in the second period.

[6] This result closely relates to the findings on penetrating prices in marketing science.

This paper also contributes to the literature on vertical product differentiation with network externalities. Lambertini and Orsini (2003) have studied the optimal level of quality chosen by a monopolist when network effects are present and the fixed costs are increasing and convex in the quality. Lambertini and Orsini (2004) analyze the incentive to perform costly R&D activities to achieve quality improvement in a vertically differentiated setting with network externalities. This paper adds to this literature by extending the monopolist's choice set to include the possibility of becoming multimarket. This allows us to study the issue of multiproduct pricing and intertemporal pricing in the presence of network externalities and vertical product differentiation.

In vertical product differentiation models (Gabszewicz and Thisse, 1978, 1980), a monopolist always chooses to innovate removing the low quality from the market. I show that this result does not hold anymore in the presence of significant network externalities in consumption. Moreover in the presence of network externalities, the issue of market coverage when more than one quality may be offered, becomes relevant. In the traditional model, with consumers' valuations distributed in the interval $[0, 1]$, the market is never covered at positive prices. The reason for this is that the willingness to pay for the lowest type is zero. When network effects are present, not only the unanimity of choice among consumers disappears, but also, the lowest consumer type has a positive valuation even at positive prices. It is not clear whether the optimal choice of the monopolist in this context is to cover or uncover the market. In this paper I allow for both possibilities and conclude that for higher values of the network effect, the covering of the market yields higher profits, whereas for low values, it is more profitable to uncover the market. [7] The results of this article are closely related to Gabszewicz and Garcia (2005) who present a model where the incumbent, selling a low quality good with network effects, faces potential competition from an entrant with a high quality good. The main result is that the incumbent prefers to block the entry of the high quality good when the network effect more than doubles the quality differential of the goods. As it will be shown in the following sections, quality improvement is more likely to occur in a duopolistic market structure (an incumbent and an entrant) than in a monopolistic one.

The paper is structured as follows: the model is presented in the following section as well as the results for each possible *regime* that the monopolist may choose; sec-

[7] Constantatos and Perrakis (1997) address the issue of market coverage for a multiproduct monopolist threatned by entry, in the absence of network externalities.

tion 4.3 investigates the social optimality of the monopolist's behavior and section 4.4 concludes. Proofs of the propositions can be found in the appendix.

4.2 The model

Consider a monopolist who sells a good which exhibits delayed network externalities, say good 1, at constant average cost normalized to 0. The presence of network externalities implies that the utility each consumer obtains from buying the good is increasing on the number of other consumers who also acquire it. Once the good is sold, it takes one period of time for network externalities to become effective. The model has two periods. In each period, there is a continuum of consumers indexed by a parameter v uniformly distributed on $[0, 1]$. Consumers buy at most one unit of the good and cannot delay consumption. The consumer obtains utility from using the good independently from the network, the so-called *stand-alone value*. The parameter v differentiates the consumers in terms of their stand-alone value: a consumer with a low v does not attribute a very high stand-alone value to the good and derives most of his utility from the network effect. In all other aspects, consumers are identical, in particular they value the network effect in the same way through a parameter α. The utility of the consumers has therefore two components: the stand-alone value and the network value. Since the network takes one period of time to become effective, consumers who acquire the good in the first period do not benefit from the network externality. On the contrary, consumers who buy in the second period enjoy both the intrinsic quality of the good and its network value. Their utility is thus increasing in the size of first period sales. More precisely, the utility for consumer v in period 1 for buying product 1 is given by $\beta_1 v - p_{11}$,with $\beta_1 > 0$ measuring the quality of product 1, and p_{11} the price set by the monopolist in period 1. Define the demand of product 1 in period 1 as $D_{11}(p_{11})$. The utility of buying product 1 in period 2 is given by: $\beta_1 v + \alpha D_{11}(p_{11}) - p_{12}$.

In the second period, the monopolist is able to sell a new good with higher quality, say good 2.[8] I also assume a zero average cost for selling good 2 and that there are no additional costs related to the introduction of the new quality in the market. The quality of good 2 is given by β_2. In the first period, the monopolist sells only good 1. A consumer v will buy if $\beta_1 v \geq p_{11}$. The demand is given by:

[8] We can think that some costless R&D activity has been undertaken and it resulted in quality improvement.

$$D_{11}(p_{11}) = 1 - \frac{p_{11}}{\beta_1}. \tag{4.1}$$

The monopolist serves the market in the interval $\left[\frac{p_{11}}{\beta_1}, 1\right]$. The profit $\pi_1(p_{11})$, in the first period is given by,

$$\pi_1(p_{11}) = p_{11}\left(1 - \frac{p_{11}}{\beta_1}\right). \tag{4.2}$$

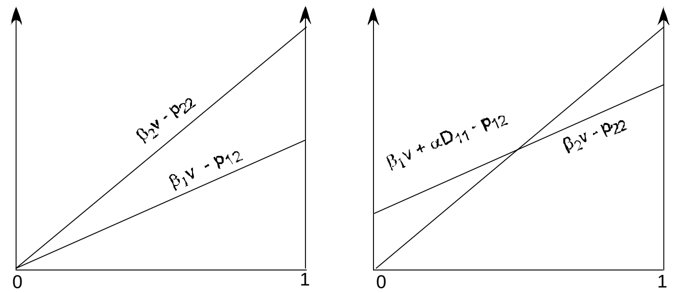
Now consider period 2. Technological innovation allows the monopolist to sell a higher quality good, good 2, whose quality is represented by the parameter β_2. There are two sources of quality differentiation between the two goods: one exogenous given the the quality differential $(\beta_2 - \beta_1)$, and one endogenous, given by the network effect generated in period 1. The monopolist is thus able to manipulate, in the first period, the endogenous component of good 1's quality through its pricing choices. Then, he has to decide whether to sell good 1 and 2, or only good 1, or only good 2 and to choose its prices, respectively p_{12} and p_{22}, in each of these alternatives.

Several elements can influence consumers' choice in the second period: the quality differential between the two variants, $\beta_2 - \beta_1$, the network effect α, the pricing policy p_{11} of the monopolist in the first period and, finally, the prices p_{12} and p_{22} for both qualities chosen in the second period. Traditional models of vertical differentiation are characterized by the fact that at identical prices consumers rank unanimously the different qualities. This property is no longer satisfied in the presence of network externalities. The reason is that the network externality brings a different source of quality differentiation that is independent of consumer types. Low types which were refraining from consuming good 1 in period 1, are now induced to buy product 1 due to its network effect. Figure 4.1 illustrates this point. The first graph corresponds to the second period market configuration, in the absence of network effects and when prices are zero. The second graph corresponds to the second period market configuration at zero prices and when network effects are positive.

I characterize now the behavior of the consumers in the second period when both qualities coexist. Denote $v_1(p_{12}, p_{22})$ the consumer who is indifferent between buying good 1 or good 2 in the second period. The value $v_1(p_{12}, p_{22})$ is the solution to the equation

$$\beta_1 v + \alpha D_{11} - p_{12} = \beta_2 v - p_{22},$$

FIGURE 4.1. The effect of the network on the unanimity of choice.



which implies that $v_1(p_{12}, p_{22}) = \frac{p_{22}-p_{12}+\alpha D_{11}}{\beta_2-\beta_1}$. It is easily seen that if α is taken to be 0, the value $v_1(p_{12}, p_{22})$ corresponds to the indifferent consumer in the traditional model of vertical differentiation. In the present case, the marginal consumer is shifted by $\frac{\alpha D_{11}}{\beta_2-\beta_1}$,a measure of the relative weight of the two sources of quality differentiation. The consumer indifferent between buying 1 or not buying at all is given by $v_2(p_{12})$,which is the solution to

$$\beta_1 v + \alpha D_{11} - p_{12} = 0.$$

namely, $v_2(p_{12}) = \frac{p_{12}-\alpha D_{11}}{\beta_1}$. For both demands to be positive and smaller than 1 at strictly positive prices, I must have that $0 < v_1 < 1$, which is equivalent to

$$\begin{aligned} p_{22} - p_{12} + \alpha D_{11} - (\beta_2 - \beta_1) &< 0 \\ p_{12}\beta_2 - p_{22}\beta_1 - \alpha D_{11}\beta_2 &< 0 \end{aligned}$$

Within the area of prices defined by the previous inequalities, I have that the demand for good 1, $D_{12}(p_{12}, p_{22})$ is given by:

$$D_{12}(p_{12}, p_{22}) = \frac{p_{22} - p_{12} + \alpha D_{11}}{\beta_2 - \beta_1} - \max\left\{0, \frac{p_{12} - \alpha D_{11}}{\beta_1}\right\}. \quad (4.3)$$

The demand for good 2 $D_{22}(p_{12}, p_{22})$ obtains as

$$D_{22}(p_{12}, p_{22}) = 1 - \frac{p_{22} - p_{12} + \alpha D_{11}}{\beta_2 - \beta_1}. \quad (4.4)$$

When $\frac{p_{12}-\alpha D_{11}}{\beta_1} > 0$, the market is incompletely covered. In the traditional model of vertical differentiation, with consumers' types distributed in the interval $[0, 1]$, a multiproduct monopolist never covers the market at positive prices. This is so because the lowest-type consumer only buys if the price is set to be zero: his valuation is indeed equal to zero. When network effects are present, the consumer type zero buys even at positive prices, in order to benefit from the network. It is therefore necessary to distinguish whether, after entry, demands correspond to that of a covered, or uncovered, market. The multimarket profit of the monopolist in period 2, $\pi_2^m(p_{12}, p_{22})$ is defined as

$$\pi_2^m(p_{12}, p_{22}) = p_{12}D_{12} + p_{22}D_{22}.$$

The monopolist then maximizes profits by selecting one of the following *regimes* in the second period: sell both goods, the "multimarket solution", sell only good 1 and disregarding the potential product improvement, the "no improvement solution", and sell only good 2 and stopping the sale of good 1 altogether, the "strict improvement solution". In the second option the demand for good 1 in the second period is given by

$$D_{12}(p_{12}, p_{11}) = 1 - \frac{1}{\beta_1}\left[p_{12} - \alpha\left(1 - \frac{p_{11}}{\beta_1}\right)\right] \tag{4.5}$$

while in the third option, the demand the monopolist faces in period 2 is independent of the first period choices, and is given by

$$D_{22}(p_{22}) = 1 - \frac{p_{22}}{\beta_2}. \tag{4.6}$$

I analyze now these regimes for the monopolist in more detail.

4.2.1 Multiproduct monopolist

We first analyze the optimal behavior of the monopolist in the regime where he sells both products and the market remains uncovered.

4.2.1.1 The monopolist does not cover the market

I shall start by analyzing the second period problem of the monopolist, given the choice of first period price, p_{11}. Consider that the monopolist chooses to sell goods 1 and 2 in the second period, without covering the market. Since $\beta_2 > \beta_1$, the market is not covered if the indifferent consumer between buying product 1 or not buying is strictly larger than 0. That is, I must have $v_2(p_{12}) > 0$, which is equivalent to $p_{12} > \alpha\left(1 - \frac{p_{11}}{\beta_1}\right)$.

Definition 1 *A [strictly] uncovered market is a situation where both products have positive demand and the sum of the demands is [strictly] less than 1 at positive prices.*

Proposition 1 *When the monopolist decides to sell both products in period 2, there exists a positive value of α, say $\overline{\alpha}$, such that the market is uncovered whenever $\alpha < \bar{\alpha}$.*[9]

This proposition states that for α sufficiently small, it may happen that the monopolist sells both goods in the second period, leaving the market uncovered. However, this case does not exhaust completely the situations in which it is optimal for the monopolist to sell both products in period 2 but, in such cases, he will be seen to cover the market. When $\alpha < \bar{\alpha}$, the monopolist chooses prices given by

$$p_{11}^{*} = \frac{\left(2\beta_1^2\left(\beta_2-\beta_1\right)-\alpha^2\beta_2\right)\beta_1}{4\beta_1^2\left(\beta_2-\beta_1\right)-\alpha^2\beta_2};$$

$$p_{12}^{*} = \frac{\beta_1}{2} + \frac{2\alpha\beta_1^2\left(\beta_2-\beta_1\right)}{2\left(4\beta_1^2\left(\beta_2-\beta_1\right)-\alpha^2\beta_2\right)} \text{ and } p_{22}^{*} = \frac{1}{2}\beta_2,$$

with corresponding total profits

$$\Pi^U = \frac{1}{4}\frac{4\beta_1^2\left(\beta_2^2-\beta_1^2\right)-\alpha^2\beta_2^2}{4\beta_1^2\left(\beta_2-\beta_1\right)-\alpha^2\beta_2}.$$

Notice that $p_{11}^{*} < p_{12}^{*}$, and thus, if we define introductory pricing by requiring a decreasing path of prices over periods, introductory pricing is indeed a candidate equilibrium. However, $p_{11}^{*} > 0$ and consequently, the installed base is not maximal in period 1. Intuitively, as α is low, the consumers place not enough importance on the network to induce the monopolist to decrease the initial price up to zero and thereby foregoing completely profits in period 1.

4.2.1.2 The monopolist covers the whole market

I consider now the second period problem when the monopolist decides to cover the market with product 1 and 2. Given that $\beta_2 > \beta_1$, the market is covered when $p_{12} \leq \alpha\left(1-\frac{p_{11}}{\beta_1}\right)$.

Definition 2 *A covered market is a situation where both products have positive demand and the sum of the demands is exactly equal to 1 at positive prices.*

Proposition 2 *When the monopolist decides to sell both products in period 2, there exists an interval of α values, namely $\left(\overline{\overline{\alpha}}, \beta_2-\beta_1\right)$ such that the market is covered.*[10]

[9] The value of $\overline{\alpha}$ is given by $\overline{\alpha} \equiv \frac{\beta_1}{\beta_2}\left(\sqrt{\left(5\beta_2-\beta_1\right)\left(\beta_2-\beta_1\right)}-\left(\beta_2-\beta_1\right)\right)$

[10] Where $\overline{\overline{\alpha}}$ is defined as $\overline{\overline{\alpha}} \equiv \frac{1}{2}\left(\sqrt{\left(\beta_2-\beta_1\right)\left(7\beta_1+\beta_2\right)}-\left(\beta_2-\beta_1\right)\right)$.

This proposition states that for some values of the quality differential and the network effect intensity, selling both products and covering the market is a candidate equilibrium. The corresponding prices are given by

$$p_{11} = 0,\ p_{12} = \alpha \text{ and } p_{22} = \frac{\alpha + \beta_2 - \beta_1}{2};$$

and total profit by

$$\Pi^C = \frac{(\alpha + \beta_2 - \beta_1)^2}{4(\beta_2 - \beta_1)}.$$

In this equilibrium, the monopolist chooses the maximal size of the installed base by setting initial price to zero and then extracting all the surplus generated by the network in the next period.

Corollary 2 *When, given α, the monopoly decides to sell both goods in the second period, he must either cover the market, or uncover it: both possibilities are not available simultaneously.*

Proof. (Corollary 2) The uncovered market solution occurs when $\alpha \in (0, \bar{\alpha})$ and the covered market solution occurs when $\alpha \in \left(\overline{\overline{\alpha}}, \beta_2 - \beta_1\right)$. One can show that for $\beta_2 > 2\beta_1$, we get $\overline{\alpha} < \overline{\overline{\alpha}}$ (see appendix). Thus the intervals of α-values for which the covered market solution and uncovered market solution are indeed solutions to the monopolist optimization problem are necessarily disjoint. ■

4.2.2 Strict product improvement

Now consider that the monopolist sells good 1 in the first period while he decides to substitute good 1 for good 2 in the second. The monopolist avoids cannibalizing sales of product 1 in the second period by introducing a high quality good. Given that network effects are delayed by one period, he disregards the impact of first period sales in period 2 and sets monopoly prices in both periods. The equilibrium is thus given by: $p_{11}^* = \frac{1}{2}\beta_1$ and $p_{22}^* = \frac{1}{2}\beta_2$ and total profits are $\pi(p_{11}^*) + \pi(p_{22}^*) = \frac{1}{4}(\beta_1 + \beta_2)$.

4.2.3 No product improvement

Finally assume that the monopolist decides to sell only good 1 in the two periods, therefore disregarding completely quality improvement.

Proposition 3 *When the monopolist decides to disregard the quality improvement, there is a treshold value for* α*, namely,* $\widetilde{\alpha} = \beta_1$*, such that the optimal monopolist solution is given by* $p_{11}^* = \frac{(\alpha-\beta_1)\beta_1}{(\alpha-2\beta_1)} > 0$ *and* $p_{12}^* = \frac{\beta_1^2}{(2\beta_1-\alpha)}$ *when* $\alpha \in (0, \widetilde{\alpha})$, *and by* $p_{11}^* = 0$ *and* $p_{12}^* = \alpha$ *when* $\alpha \in (\widetilde{\alpha}, \infty)$.

From this proposition we know that, in an equilibrium at which the monopolist disregards the quality improvement by selling only good 1, the network is only maximal if α is not too large with respect to the quality of good 1. Admittedly, the monopolist also chooses introductory pricing when α is larger than $\widetilde{\alpha}$, since in that case also p_{11}^* $> p_{12}^*$, however, the network constituted in the initial period is not maximal.

4.2.4 Comparison between alternative regimes

As seen in the previous section, the monopolist adopts different optimizing behavior depending on the value of the network effect intensity α.

FIGURE 4.2. The monopolist's choices as a function of α.

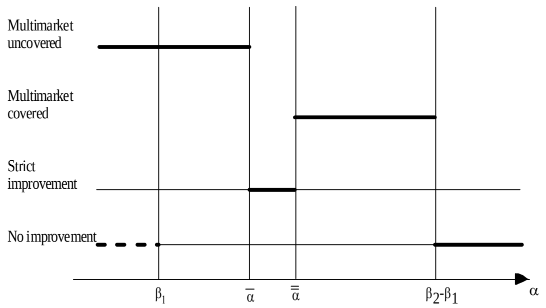


Figure 4.2 represents the regimes that the monopolist may choose as a function of the parameter α and β_1 and β_2. Stressed in bold are the optimal regimes for each partition of the space of the parameter α. Given these alternatives, I state the following proposition.[11]

Proposition 4 *Quality improvement in the second period occurs if and only if* $\alpha < \beta_2 - \beta_1$.

[11] From the start we have already assumed that $\beta_2 > 2\beta_1$. In order to allow a simple comparison among the alternatives opened to the monopolist, we assume here that $\beta_2 > 5\beta_1$. Without this assumption we would have to subdivide further the interval $(\overline{\alpha}, \overline{\overline{\alpha}})$, to distinguish between those values of α, for which strict improvement would dominate the no improvement solution, and *vice versa*. Thus, although simplifying, this assumption does not restrict substantially our analysis.

1. *When* $\alpha \in \left(\overline{\overline{\alpha}}, \beta_2 - \beta_1\right)$*, the monopolist chooses to sell both goods while covering the market and creates the maximal network in period 1.*

2. *When* $\alpha \in \left(\overline{\alpha}, \overline{\overline{\alpha}}\right)$ *the monopolist sells the high quality good only in the second period.*

3. *When* $\alpha \in (0, \overline{\alpha})$ *the monopolist chooses to sell both products without covering the market, and chooses a first period price smaller than the price of period 2, but strictly larger than zero.*

When the network intensity is extremely low, the monopolist does not have any interest in lowering the first period price up to zero: the increase in the installed base does not compensate for the loss of profits in period 1 resulting from a zero price. Moreover, given that α is not very high, the quality of good 1 in the second period as resulting from its stand alone value and the network effect, is not large enough to justify disregarding the quality improvement. Hence, it is profitable to sell also good 2 so as to extract the surplus from the high valuation types. As α increases ($\overline{\alpha} > \alpha > \overline{\overline{\alpha}}$), it becomes optimal to choose the solution in which quality 1 is sold in period 1, but withdrawn from the market in period 2 to the benefit of quality 2. In both periods, the monopolist quotes the monopoly price. The difference with the previous situation is that selling simultaneously both products in the second period becomes less attractive: qualities are closer to each other and the cannibalization effect is stronger. The monopolist finds it profitable to stop selling good 1 altogether. However, if α is higher, in particular when $\alpha > \overline{\overline{\alpha}}$, stopping to sell good 1 in period 2 is not profitable anymore. The monopolist can sell good 1 at period 1 at a zero price and thereby maximize the second period network, then, in period 2 he succeeds in covering the market by selling both qualities. Finally, if α continues to increase, the monopolist loses interest in the new high quality good. This is so because the quality differential between the products does not compensate anymore the size of the network effect obtained in period 1. It is profitable for the monopolist to disregard quality improvement and to sell only good 1 in both periods.

4.3 Welfare considerations

In this section we study the optimal welfare solutions, as a function of the parameters of the model. The welfare function is given by

$$W = CS_1 + CS_2 + \pi_1 + \pi_2, \tag{4.7}$$

where, CS_i and π_i denote respectively, period i's consumer surplus and profit. Let W^U[resp. W^C] denote the welfare of the solution when both qualities are sold at strictly positive prices in both periods and the market remains uncovered [resp. covered]. Similarly, W^{SI} [resp. W^{NI}] stands for the welfare of the solution in which only the high [resp. low] quality is sold in the second period. We must distinguish the value of welfare in the "no improvement" solution when $\alpha \in [0, \beta_1]$ and when $\alpha \in [\beta_1, \infty]$ since the equilibrium prices differ (see Proposition 3).We denote this value respectively as $W^{NI}_{[0,\beta_1]}$ and $W^{NI}_{[\beta_1,\infty]}$. The values of the the different components of W for the different alternatives faced by the monopolist are presented in appendix.

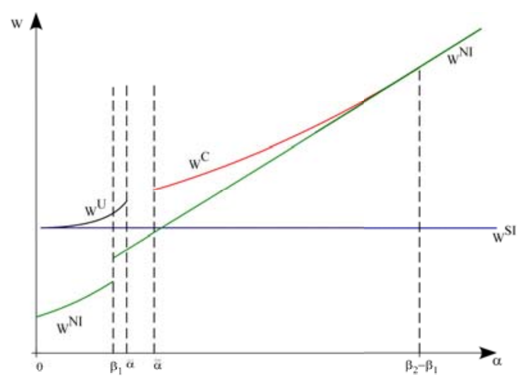
Comparing these values with the welfare realized at the monopoly solution for each value of α, and assuming that the quality differential is sufficiently large (in particular we must guarantee that $\beta_2 > 5.72\beta_1$), we obtain the following.

Proposition 5 *When the optimal monopolist solution consists in selecting a particular regime, this regime also corresponds to the one that maximizes welfare at monopoly prices for sufficiently large quality differential.*

Proof. See the appendix. ■

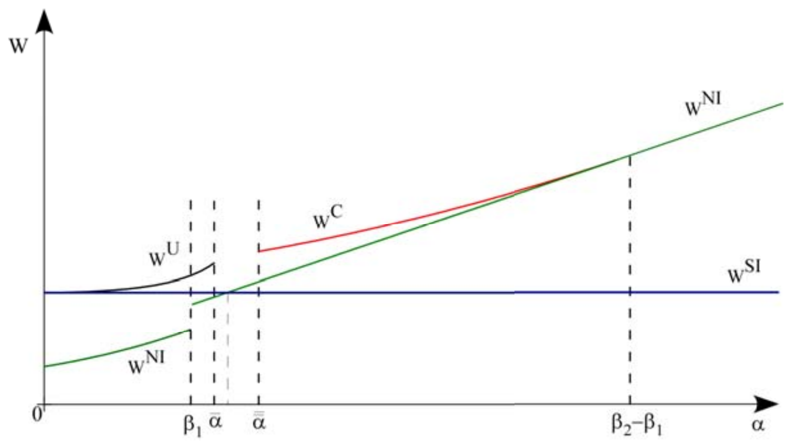
Figure 4.3 illustrates the proposition.

FIGURE 4.3. Welfare as a function of the network effect, α. β_1=1, β_2=6.



It is important to highlight that for smaller values of β_2, a comparative analysis could still be performed. However, in that case the monopolist's choice of regime would not correspond entirely to the on that maximizes the welfare. Namely, there would exist a α, say $\widetilde{\alpha} \in \left(\overline{\alpha}, \overline{\overline{\alpha}}\right)$, for which if $\overline{\alpha} < \alpha < \widetilde{\alpha}$ strictly improvement dominates the no improvement solution in terms of welfare and if $\widetilde{\alpha} < \alpha < \overline{\overline{\alpha}}$, the contrary holds. Instead, from the point of view of the monopolist, in this space of α, the strict improvement solution always dominates the no improvement one (see Figure 4.4).

FIGURE 4.4. Welfare as a function of the network effect, α. β_1=1, β_2=5.



4.4 Conclusion

In this paper I have analyzed the necessary and sufficient conditions under which a monopolist, operating in a market with network externalities has incentive to improve its quality. Although the model is rather simple, the existence of many possible regimes available for the monopolist renders the analysis complex and extensive. In fact, we must consider that the monopolist might decide to operate as a multimarket monopolist or to sell only one quality (the high or the low). In case he decides to sell both qualities he still faces the choice of covering the market or not. The sequential structure of the game also contributes to the complexity of its analysis, given that the situation of the market in the second period is determined by the first period choices. Nevertheless, it is possible to identify the precise conditions under which each of the possible regimes is optimal. The main conclusion is that quality innovation occurs if and if the network intensity is smaller than the quality differential. This result closely relates to Gabszewicz and Garcia (2005). In their paper, the monopolist accommodates the entrance of a higher quality good sold by a rival if and only if the network intensity is smaller than twice the quality differential. In fact, comparing these two results, we can confirm the intuition that quality improvement occurs with higher likelihood in a duopolistic market structure than in a monopolistic one.

4.A Appendix

Proof. (Proposition 1) When the market is uncovered, the monopolist solves the following optimization problem:

$$\max_{p_{12},p_{22}} p_{12}\left(\frac{p_{22}-p_{12}+\alpha D_{11}}{\beta_2-\beta_1}-\frac{p_{12}-\alpha D_{11}}{\beta_1}\right)+p_{22}\left(1-\frac{p_{22}-p_{12}+\alpha D_{11}}{\beta_2-\beta_1}\right)$$

The first order conditions are:

$$\begin{aligned}\frac{(2p_{22}-p_{12}+\alpha D_{11})}{\beta_2-\beta_1}+\frac{(\alpha D_{11}-p_{12})}{\beta_1}-\frac{p_{12}\beta_2}{\beta_1(\beta_2-\beta_1)} &= 0\\ \frac{p_{12}-p_{22}}{\beta_2-\beta_1}+\frac{(p_{12}-p_{22}-\alpha D_{11})}{\beta_2-\beta_1}+1 &= 0\end{aligned}$$

The optimal solution is:

$$\begin{aligned}p_{12} &= \frac{\beta_1+\alpha D_{11}}{2}\\ p_{22} &= \frac{1}{2}\beta_2\end{aligned}$$

and the equilibrium demands are:

$$\begin{aligned}D_{12}^{*} &= \frac{\beta_2 D_{11}\alpha}{2(\beta_2-\beta_1)\beta_1}\\ D_{22}^{*} &= \frac{1}{2}-\frac{\alpha D_{11}}{2(\beta_2-\beta_1)}.\end{aligned}$$

I must guarantee that $D_{12}^{*}+D_{22}^{*}<1$, that is $\frac{(\beta_1+\alpha D_{11})}{2\beta_1}<1$, which occurs when

$$\alpha D_{11}<\beta_1. \tag{4.8}$$

Condition [4.8] depends on D_{11} and consequently on p_{11}. Its fulfillment can only be assessed after disclosing the optimal choice of p_{11}. The monopolist chooses p_{11} to maximize overall profit

$$\Pi(p_{11})=\frac{\left(\beta_1(\beta_2-\beta_1)+\alpha^2\left(1-\frac{p_{11}}{\beta_1}\right)^2\right)\beta_2}{4(\beta_2-\beta_1)\beta_1}+p_{11}\left(1-\frac{p_{11}}{\beta_1}\right).$$

This function is concave when $\alpha^2\beta_2 - 4\beta_1^2(\beta_2-\beta_1) < 0$ and convex otherwise. [12] When the function is concave, the optimum is attained at the root of the first derivative of $\Pi(p_{11})$, namely,

$$p_{11}^* = \frac{\left(\alpha^2\beta_2 - 2\beta_1^2(\beta_2-\beta_1)\right)\beta_1}{\alpha^2\beta_2 - 4\beta_1^2(\beta_2-\beta_1)}.$$

This price p_{11} belongs to the admissible interval $(0,\beta_1)$ if and only if $\alpha^2\beta_2 - 2\beta_1^2(\beta_2-\beta_1) < 0$.[13] If $\alpha^2\beta_2 - 2\beta_1^2(\beta_2-\beta_1) > 0$ and $\alpha^2\beta_2 - 4\beta_1^2(\beta_2-\beta_1) < 0$ the function is concave, however $p_{11}^* = 0$. Finally, when $\alpha^2\beta_2 - 4\beta_1^2(\beta_2-\beta_1) > 0$ the function is convex and $p_{11}^* = 0$. I must now verify whether condition [4.8] is verified both when $p_{11}^* = 0$ and $p_{11}^* > 0$.

If $p_{11}^* = 0$ the condition [4.8], for the existence of an uncovered market optimal solution, is $\alpha < \beta_1$. However, since I assume $\beta_2 > 2\beta_1$, whenever the optimal solution is $p_{11} = 0$ the condition in never met.[14] If $p_{11}^* = \frac{\left(2\beta_1^2(\beta_2-\beta_1)-\alpha^2\beta_2\right)\beta_1}{4\beta_1^2(\beta_2-\beta_1)-\alpha^2\beta_2} > 0$, the condition [4.8] can be rewritten as

$$\alpha\left(1 - \frac{1}{\beta_1}\frac{\left(2\beta_1^2(\beta_2-\beta_1) - \alpha^2\beta_2\right)\beta_1}{4\beta_1^2(\beta_2-\beta_1) - \alpha^2\beta_2}\right) < \beta_1$$

that is satisfied when $\alpha < \overline{\alpha}$. It is easy to show that $\bar{\alpha} < \sqrt{2}\beta_1\sqrt{\frac{\beta_2-\beta_1}{\beta_2}}$. Hence an uncovered market solution exists in the domain $(0,\bar{\alpha})$ and it is equal to: $p_{11}^* = \frac{\left(2\beta_1^2(\beta_2-\beta_1)-\alpha^2\beta_2\right)\beta_1}{4\beta_1^2(\beta_2-\beta_1)-\alpha^2\beta_2}$, $p_{12} = \frac{\beta_1}{2} + \frac{2\alpha\beta_1^2(\beta_2-\beta_1)}{2\left(4\beta_1^2(\beta_2-\beta_1)-\alpha^2\beta_2\right)}$, $p_{22} = \frac{1}{2}\beta_2$. ■

Proof. (Proposition 2) The monopolist solves the following problem.

$$\max_{p_{12},p_{22}} \; p_{12}\left(\frac{p_{22} - p_{12} + \alpha\left(1 - \frac{p_{11}}{\beta_1}\right)}{\beta_2-\beta_1}\right) + p_{22}\left(1 - \frac{p_{22} + \alpha\left(1-\frac{p_{11}}{\beta_1}\right) - p_{12}}{\beta_2-\beta_1}\right)$$

$$s.t.\ p_{12} \le \alpha\left(1 - \frac{p_{11}}{\beta_1}\right)$$

Writing the Lagrangian, I have:

$$L = \left(p_{12}\left(\frac{p_{22} + \alpha D - p_{12}}{\beta_2-\beta_1}\right) + p_{22}\left(1 - \frac{p_{22} + \alpha D - p_{12}}{\beta_2-\beta_1}\right)\right) - l\left(p_{12} - \alpha D\right)$$

[12] Solving the inequality for α we have that whenever $\alpha < 2\beta_1\sqrt{\frac{\beta_2-\beta_1}{\beta_2}} \equiv \hat{\alpha}$ the profit function Π is concave.

[13] This condition can be written as $\alpha < \sqrt{2}\beta_1\sqrt{\frac{\beta_2-\beta_1}{\beta_2}}$.

[14] p_{11}^* is zero when $\alpha > \sqrt{2}\beta_1\sqrt{\frac{\beta_2-\beta_1}{\beta_2}}$. There exists an uncovered market solution when $p_{11}^* = 0$ if $\alpha < \beta_1$. However $\beta_1 < \sqrt{2}\beta_1\sqrt{\frac{\beta_2-\beta_1}{\beta_2}}$, so, no uncovered market solution exists with $p_{11} = 0$.

and the Kuhn Tucker conditions are

$$\begin{aligned}
\frac{p_{12}}{\beta_2-\beta_1}-\frac{p_{22}}{\beta_2-\beta_1}+\frac{1}{\beta_2-\beta_1}\left(p_{12}-D\alpha-p_{22}\right)+1 &= 0\\
\frac{p_{22}}{\beta_2-\beta_1}-\frac{p_{12}}{\beta_2-\beta_1}-l+\frac{1}{\beta_2-\beta_1}\left(D\alpha-p_{12}+p_{22}\right) &= 0\\
l\left(p_{12}-\alpha D\right) &= 0
\end{aligned}$$

The unique finite solution for this problem occurs when $p_{12}=\alpha D$, $p_{22}=\left(\frac{\alpha D+\beta_2-\beta_1}{2}\right)$ and $l=1$. I must check whether the demands are positive and smaller than 1 at this solution. The equilibrium demand of good 1 is $D_{12}=\frac{\alpha D_{11}+\beta_2-\beta_1}{2(\beta_2-\beta_1)}$, and I must guarantee that,

$$\alpha D_{11}<\left(\beta_2-\beta_1\right) \tag{4.9}$$

In the first stage the monopolist chooses p_{11} to maximize overall profit given by

$$\Pi\left(p_{11}\right)=p_{11}\left(1-\frac{p_{11}}{\beta_1}\right)+\frac{\left(\alpha\left(1-\frac{p_{11}}{\beta_1}\right)+\beta_2-\beta_1\right)^2}{4\left(\beta_2-\beta_1\right)}$$

This function is concave when $\alpha^2-4\beta_1\beta_2+4\beta_1^2<0$, or equivalently when $\alpha<2\sqrt{\beta_1\left(\beta_2-\beta_1\right)}$ and convex otherwise. Whenever the function is concave, the function attains its maximum at $p_{11}^*=\frac{\beta_1\left((\alpha-2\beta_1)(\beta_2-\beta_1)+\alpha^2\right)}{\alpha^2-4\beta_1(\beta_2-\beta_1)}$. This value of p_{11}^* is strictly lower than β_1, however, if $\left(\left(\alpha-2\beta_1\right)\left(\beta_2-\beta_1\right)+\alpha^2\right)>0$ or equivalently, $\alpha>\overline{\overline{\alpha}}$, it can assume negative values, in which case the optimum is attained in the boundary of the admissible set, namely at $p_{11}^*=0$. When the global profit is convex, it is easily seen that the optimum is attained at $p_{11}^*=0$. In summary, for $\alpha\geq 2\sqrt{\beta_1\left(\beta_2-\beta_1\right)}$ the function is convex and $p_{11}^*=0$, for $\overline{\overline{\alpha}}<\alpha<2\sqrt{\beta_1\left(\beta_2-\beta_1\right)}$, the profit function is concave but the optimum is attained at $p_{11}^*=0$, for $\alpha\leq\overline{\overline{\alpha}}$, the profit function is concave and the optimal p_{11}^* is positive.[15]

I must now check whether the solutions $p_{11}^*=0$, and $p_{11}^*>0$, allow to maintain consistency with the initial condition for the existence of the second period covered market equilibrium, (4.9). When $p_{11}^*=0$, the solution is valid iff $\alpha<\left(\beta_2-\beta_1\right)$. As $p_{11}^*=0$ is a solution if and only if $\alpha\in\left(\overline{\overline{\alpha}},\infty\right)$, the first and second period solution are valid in the domain $\alpha\in\left(\overline{\overline{\alpha}},\beta_2-\beta_1\right)$ that can be shown to be non empty (for $\beta_2>2\beta_1$).

[15]It is easily shown that $\overline{\overline{\alpha}}<2\sqrt{\beta_1\left(\beta_2-\beta_1\right)}$.

When $p_{11}^{*} > 0$ condition [4.9] implies that

$$\alpha \left(1 - \frac{\left(\left(\alpha - 2\beta_1\right)\left(\beta_2 - \beta_1\right) + \alpha^2\right)}{\alpha^2 - 4\beta_1\left(\beta_2 - \beta_1\right)}\right) > \beta_2 - \beta_1$$

which is equivalent to $\alpha > 2\left(\beta_2 - \beta_1\right)$. The solution $p_{11}^{*} > 0$, is valid for $\alpha < \overline{\overline{\alpha}}$, which is smaller than $2\left(\beta_2 - \beta_1\right)$, so, I cannot have $p_{11} > 0$, inducing a covered market solution in the second stage. ■

Proof. (Corollary 2) I will now show that $\overline{\alpha} < \overline{\overline{\alpha}}$.

$$\begin{aligned}
& 2\frac{\beta_1}{\beta_2}\left(\sqrt{\left(5\beta_2 - \beta_1\right)\left(\beta_2 - \beta_1\right)} - \left(\beta_2 - \beta_1\right)\right) \\
< \; & \sqrt{\left(\beta_2 - \beta_1\right)\left(7\beta_1 + \beta_2\right)} - \left(\beta_2 - \beta_1\right) \Leftrightarrow \\
& \frac{1}{\beta_2}\left(\beta_2 - 2\beta_1\right)\left(\beta_2 - \beta_1\right) \\
< \; & \sqrt{\left(\beta_2 - \beta_1\right)\left(7\beta_1 + \beta_2\right)} - 2\frac{\beta_1}{\beta_2}\sqrt{\left(5\beta_2 - \beta_1\right)\left(\beta_2 - \beta_1\right)}
\end{aligned}$$

The RHS is positive for $\beta_2 > 2\beta_1$ (it is increasing in β_2, for $\beta_2 > 2\beta_1$ and 0 for $\beta_2 = 2\beta_1$). Taking squares,

$$\frac{1}{\beta_2^2}\left(\beta_2 - 2\beta_1\right)^2\left(\beta_2 - \beta_1\right)^2 <$$

$$\left(\left(7\beta_1 + \beta_2\right) + 4\frac{\beta_1^2}{\beta_2^2}\left(5\beta_2 - \beta_1\right) - 4\frac{\beta_1}{\beta_2}\sqrt{\left(7\beta_1 + \beta_2\right)\left(5\beta_2 - \beta_1\right)}\right)\left(\beta_2 - \beta_1\right) \Leftrightarrow$$

$$\left(\frac{1}{\beta_2^2}\left(\beta_2 - 2\beta_1\right)^2\left(\beta_2 - \beta_1\right) - \left(7\beta_1 + \beta_2\right) - 4\frac{\beta_1^2}{\beta_2^2}\left(5\beta_2 - \beta_1\right)\right)$$

$$< -4\frac{\beta_1}{\beta_2}\sqrt{\left(7\beta_1 + \beta_2\right)\left(5\beta_2 - \beta_1\right)} \Leftrightarrow$$

$$-12\left(\beta_1 + \beta_2\right)\frac{\beta_1}{\beta_2} < -4\frac{\beta_1}{\beta_2}\sqrt{\left(7\beta_1 + \beta_2\right)\left(5\beta_2 - \beta_1\right)}$$

$$3\left(\beta_1 + \beta_2\right) > \sqrt{\left(7\beta_1 + \beta_2\right)\left(5\beta_2 - \beta_1\right)}$$

$$9\left(\beta_1 + \beta_2\right)^2 > \left(7\beta_1 + \beta_2\right)\left(5\beta_2 - \beta_1\right)$$

$$9\left(\beta_1 + \beta_2\right)^2 - \left(7\beta_1 + \beta_2\right)\left(5\beta_2 - \beta_1\right) > 0$$

$$4\left(\beta_2 - 2\beta_1\right)^2 > 0$$

■

Proof. (Proposition 3) Firm 1 is a monopolist so its demand is given by the mass of consumers whose valuation is above price, i.e.

$$D(p_{11}) = \int_{p_1}^{1} \beta_1 v(r) df(r) = 1 - \frac{p_{11}}{\beta_1} \tag{4.10}$$

The demand at period 2 is given by:

$$D_{12}(p_{12}) = \int_{p_{12}-\alpha(1-\frac{p_{11}}{\beta_1})}^{1} \beta_1 v(r) df(r) = 1 - \frac{1}{\beta_1}\left[p_{12} - \alpha\left(1 - \frac{p_{11}}{\beta_1}\right)\right] \tag{4.11}$$

In the second period, the monopolist solves the following problem:

$$\begin{aligned} &\max_{p_{12}} p_{12}\left(1 - \frac{1}{\beta_1}\left(p_{12} - \alpha\left(1 - \frac{p_{11}}{\beta_1}\right)\right)\right) \\ &s.t.\ 1 - \frac{1}{\beta_1}\left(p_{12} - \alpha\left(1 - \frac{p_{11}}{\beta_1}\right)\right) \leq 1 \end{aligned}$$

The Lagrangian of the problem is the following:

$$L = p_{12}\left(1 - \frac{1}{\beta_1}(p_{12} - \alpha D)\right) + \lambda\left(1 - \left(1 - \frac{1}{\beta_1}(p_{12} - \alpha D)\right)\right)$$

and the Kuhn Tucker conditions are:

$$\begin{aligned} \frac{\lambda}{\beta_1} + \frac{(D\alpha - 2p_{12})}{\beta_1} + 1 &= 0 \\ \lambda\left(1 - \left(1 - \frac{1}{\beta_1}(p_{12} - \alpha D)\right)\right) &= 0 \end{aligned}$$

This system has two solutions. I consider first the case where the restriction is not binding, i. e. $\lambda = 0$ and $p_{12} = \frac{1}{2}(D\alpha + \beta_1)$. The demand must be positive and less than 1, and this occurs iff $D_{12} = \frac{1}{2}\frac{(\alpha D_{11} + \beta_1)}{\beta_1} < 1$. The condition that must be satisfied is:

$$\alpha D_{11} < \beta_1 \tag{4.12}$$

If the solution is valid, the monopolist chooses p_{11}, to maximize:

$$\Pi(p_{11}) = p_{11}\left(1 - \frac{p_{11}}{\beta_1}\right) + \frac{1}{4\beta_1}\left(\left(1 - \frac{p_{11}}{\beta_1}\right)\alpha + \beta_1\right)^2$$

This function is concave when $\alpha < 2\beta_1$, in which case the profit is maximized for $p_{11}^* = \frac{(\beta_1 - \alpha)\beta_1}{(2\beta_1 - \alpha)}$. So I have that when $\alpha \geq 2\beta_1$, the profit function is convex and the optimum is obtained at $p_{11}^* = 0$. When $\beta_1 < \alpha < 2\beta_1$, the function is concave but the optimum is attained in the lower boundary of the admissible set for p_{11}, namely $[0, \beta_1]$. When $\alpha \leq \beta_1$, the function is concave and $p_{11}^* > 0$. I must now check whether $p_{11}^* = 0$,or $p_{11}^* > 0$ are consistent with the condition [4.12]. For $p_{11}^* = 0$, I must have $\alpha < \beta_1$, however the solution $p_{11}^* = 0$ occurs only when $\alpha > \beta_1$, and thus this solution is not consistent with the second period optimal solution given by $p_{12}^* = \alpha D_{11} + \beta_1$.As for $p_{11}^* = \frac{(\alpha-\beta_1)\beta_1}{(\alpha-2\beta_1)}$, the condition [4.12] boils down to $-2\frac{(\beta_1-\alpha)\beta_1}{(2\beta_1-\alpha)} < 0$, which is true whenever $\alpha < \beta_1$, situation where $p_{11}^* > 0$ is effectively optimal.

I analyze now the possibility that the restriction is binding. In that case $p_{12}^* = \alpha D_{11}$ and $\lambda = \alpha D_{11} - \beta_1$. The condition for this solution to be valid is:

$$\alpha D_{11} > \beta_1 \tag{4.13}$$

If this solution is valid, the monopolist maximizes in the first period, the following overall profit function:

$$\Pi(p_{11}) = (p_{11} + \alpha)\left(1 - \frac{p_{11}}{\beta_1}\right)$$

This function is overall concave and the optimum is attained at: $p_{11}^* = \frac{(\beta_1-\alpha)}{2}$. If $\alpha > \beta_1$, $p_{11}^* = 0$, if $\alpha < \beta_1$, $p_{11}^* > 0$. I must now check whether the condition [4.13] works. Consider first $p_{11}^* = 0$, then I must have $\alpha > \beta_1$, which coincides with the area of parameters where $p_{11}^* = 0$ is effectively optimal. Now, if $p_{11}^* > 0$, the condition is $\alpha\left(1 - \frac{(\beta_1-\alpha)}{2\beta_1}\right) - \beta_1 = \frac{1}{2}\frac{(\alpha+2\beta_1)(\alpha-\beta_1)}{\beta_1} > 0 \Leftrightarrow \alpha > \beta_1$. However for $\alpha > \beta_1$, $p_{11}^* > 0$ is not optimal. The condition is not met. ■

Proof. (Proposition 4)

I start by identifying the monopolist's options as follows:

a) Selling both goods, leaving a part of the market unserved and not practising introductory pricing. The profit is $\Pi^U = \frac{1}{4}\frac{4\beta_1^2(\beta_2^2-\beta_1^2)-\alpha^2\beta_2^2}{4\beta_1^2(\beta_2-\beta_1)-\alpha^2\beta_2}$.

b) Selling only good 1 in the second period and not practising introductory pricing (no improvement option). Profit is $\Pi^{NI}_{[0,\beta_1]} = \frac{\beta_1^2}{(2\beta_1-\alpha)}$.

c) Selling only good 2 in the second period and obtaining profit $\Pi^{SI} = \frac{\beta_2+\beta_1}{4}$.

d) Selling only good 1 in the second period and practising introductory pricing (no improvement option). Profit $\Pi^{NI}_{[\beta_1,\infty]} = \alpha$

e) Selling both goods, covering the market unserved and practising introductory pricing. The profit is $\Pi^C = \frac{(\alpha+\beta_2-\beta_1)^2}{4(\beta_2-\beta_1)}$

Lemma A.1 *Whenever $\beta_2 > 2\beta_1$, $\beta_1 < \overline{\alpha} < \overline{\overline{\alpha}} < \beta_2 - \beta_1$.*

Proof. (Lemma A.1) First I show that $\beta_1 < \overline{\alpha}$.

$$\begin{gathered}
\beta_1 < \overline{\alpha} \Leftrightarrow \\
\beta_1 < \frac{\beta_1}{\beta_2}\left(\sqrt{(5\beta_2 - \beta_1)(\beta_2 - \beta_1)} - (\beta_2 - \beta_1)\right) \Leftrightarrow \\
(2\beta_2 - \beta_1) < \sqrt{(5\beta_2 - \beta_1)(\beta_2 - \beta_1)} \Leftrightarrow \\
(2\beta_2 - \beta_1)^2 < (5\beta_2 - \beta_1)(\beta_2 - \beta_1) \Leftrightarrow \\
(2\beta_2 - \beta_1)^2 - (5\beta_2 - \beta_1)(\beta_2 - \beta_1) < 0 \Leftrightarrow \\
\beta_2 (2\beta_1 - \beta_2) < 0
\end{gathered}$$

Whenever $\beta_2 > 2\beta_1$ I obtain $\beta_1 < \overline{\alpha}$. From Corollary 2 it follows that $\overline{\alpha} < \overline{\overline{\alpha}}$. I must just show that $\overline{\overline{\alpha}} < \beta_2 - \beta_1$.

$$\begin{aligned}
\overline{\overline{\alpha}} &< \beta_2 - \beta_1 \Leftrightarrow \\
\left(\frac{1}{2}\sqrt{(\beta_2 - \beta_1)(7\beta_1 + \beta_2)} - \frac{1}{2}(\beta_2 - \beta_1)\right) &< \beta_2 - \beta_1 \Leftrightarrow \\
\sqrt{(\beta_2 - \beta_1)(7\beta_1 + \beta_2)} &< 3(\beta_2 - \beta_1) \Leftrightarrow \\
(7\beta_1 + \beta_2) &< 9(\beta_2 - \beta_1) \Leftrightarrow \\
(7\beta_1 + \beta_2) - 9(\beta_2 - \beta_1) &< 0 \\
8(2\beta_1 - \beta_2) &< 0
\end{aligned}$$

Whenever $\beta_2 > 2\beta_1$ I obtain $\overline{\overline{\alpha}} < \beta_2 - \beta_1$.

■

The monopolist faces the following options for the different subsets of the parameter space, α.

Case 1: $\alpha \in (0, \beta_1)$, then the monopolist chooses between options a), b) and c).

Case 2: $\alpha \in (\beta_1, \overline{\alpha})$, the monopolist may chooses between options a), c) and d).

Case 3: $\alpha \in \left(\overline{\alpha}, \overline{\overline{\alpha}}\right)$, the monopolist chooses between options c) and d).

Case 4: $\alpha \in \left(\overline{\overline{\alpha}}, \beta_2 - \beta_1\right)$, the monopolist chooses between c), d) and e).

Case 5: $\alpha \in (\beta_2 - \beta_1, \infty)$, the monopolist chooses between options c) and d).

I will now analyze each case separately:

Case 1: $\alpha \in (0, \beta_1)$ I show first that a) dominates b) and then that a) dominates c).

$$\begin{aligned}
\Pi^U &> \Pi^{NI}_{[0,\beta_1]} \\
\frac{1}{4}\frac{4\beta_1^2\left(\beta_2^2-\beta_1^2\right)-\alpha^2\beta_2^2}{4\beta_1^2\left(\beta_2-\beta_1\right)-\alpha^2\beta_2} &> \frac{\beta_1^2}{\left(2\beta_1-\alpha\right)} \Leftrightarrow \\
\left(\alpha+2\beta_1\right)\left(\alpha\beta_2-2\beta_1\beta_2+2\beta_1^2\right)^2 &> 0
\end{aligned}$$

Also,

$$\begin{aligned}
\Pi^U &> \Pi^{SI} \\
\frac{1}{4}\frac{4\beta_1^2\left(\beta_2^2-\beta_1^2\right)-\alpha^2\beta_2^2}{4\beta_1^2\left(\beta_2-\beta_1\right)-\alpha^2\beta_2} &> \frac{\beta_2+\beta_1}{4} \Leftrightarrow \\
\beta_2\beta_1\alpha^2 &> 0
\end{aligned}$$

The optimal choice of the monopolist is to sell both products, uncovering the market and not practising introductory pricing.

Case 2: $\alpha \in (\beta_1, \overline{\alpha})$, the monopolist may choose between options a), c) and d). As shown in Case 1, $\Pi^U > \Pi^{SI}$ (the proof was independent of the interval of values for the parameter α). Now I show that $\Pi^U > \Pi^{NI}_{[\beta_1,\infty]}$. The strategy that I will use is the following: First I show that $\Pi^{NI}_{[0,\beta_1]} > \Pi^{NI}_{[\beta_1,\infty]}$ whenever $\alpha \in (\beta_1, \overline{\alpha})$ and then I use the result in Case 1 to conclude that $\Pi^U > \Pi^{NI}_{[0,\beta_1]} > \Pi^{NI}_{[\beta_1,\infty]}$.

$$\Pi^{NI}_{[0,\beta_1]} > \Pi^{NI}_{[\beta_1,\infty]} \Leftrightarrow \frac{\beta_1^2}{\left(2\beta_1-\alpha\right)} > \alpha$$

I can easily show that $2\beta_1 > \alpha$ whenever $\alpha \in (\beta_1, \overline{\alpha})$. Therefore the above inequality can be rewritten as:

$$\begin{aligned}
\beta_1^2 &> \alpha\left(2\beta_1-\alpha\right) \Leftrightarrow \beta_1^2-\alpha\left(2\beta_1-\alpha\right) > 0 \\
\left(\beta_1-\alpha\right)^2 &> 0
\end{aligned}$$

Then $\Pi^U > \Pi^{NI}_{[\beta_1,\infty]}$.

Case 4: $\alpha \in \left(\overline{\overline{\alpha}}, \beta_2 - \beta_1\right)$, the monopolist chooses between c), d) and e). We have that $\Pi^C > \Pi^{NI}_{[0,\beta_1]}$

$$\Pi^C > \Pi^{NI}_{[0,\beta_1]} \Leftrightarrow \frac{(\alpha + \beta_2 - \beta_1)^2}{4(\beta_2 - \beta_1)} > \alpha$$
$$\frac{(\alpha + \beta_1 - \beta_2)^2}{4(\beta_2 - \beta_1)} > 0$$

and $\Pi^C > \Pi^{SI}$.

$$\Pi^C > \Pi^{SI} \Leftrightarrow \frac{(\alpha + \beta_2 - \beta_1)^2}{4(\beta_2 - \beta_1)} > \frac{\beta_2 + \beta_1}{4}$$
$$(\alpha + \beta_2 - \beta_1)^2 4 - (\beta_2 + \beta_1) 4 (\beta_2 - \beta_1) > 0$$
$$4\left(2(\alpha - \beta_1)(\beta_2 - \beta_1) + \alpha^2\right) > 0$$

Then the optimal solution is to cover the market selling both products.

Case 5: $\alpha \in (\beta_2 - \beta_1, \infty)$, the monopolist chooses between options c) and d).

$$\Pi^{NI}_{[\beta_1,\infty]} > \Pi^{SI} \Leftrightarrow \alpha > \frac{\beta_2 + \beta_1}{4}$$

If this inequality holds for the smallest value of α, it holds everywhere. For $\alpha = \beta_2 - \beta_1$ we have:

$$4(\beta_2 - \beta_1) - (\beta_2 + \beta_1) = (3\beta_2 - 5\beta_1) > 0$$

And so I can conclude that $\Pi^{NI}_{[0,\beta_1]} > \Pi^{SI}$ holds.

Finally, for case 3, I must check whether $\Pi^{NI}_{[\beta_1,\infty]} \gtreqless \Pi^{SI}$. If $\frac{\beta_1+\beta_2}{4} > \overline{\overline{\alpha}}$, I know that in the interval $\alpha \in \left(\overline{\alpha}, \overline{\overline{\alpha}}\right)$, $\alpha < \frac{\beta_1+\beta_2}{4}$ and thus $\Pi^{NI}_{[0,\beta_1]} < \Pi^{SI}$. The monopolist would then choose to sell only good 2 in period 2. However I can guarantee that $\frac{\beta_1+\beta_2}{4} > \overline{\overline{\alpha}}$ if and only if $\beta_2 > 4.7889\beta_1$ (this is the reason why we have assumed that $\beta_2 > 5\beta_1$). Likewise, if $\frac{\beta_1+\beta_2}{4} < \overline{\alpha}$, in the interval $\alpha \in \left(\overline{\alpha}, \overline{\overline{\alpha}}\right)$, $\alpha > \frac{\beta_1+\beta_2}{4}$. The monopolist would then prefer to sell only good 1 in both periods. I can guarantee that $\frac{\beta_1+\beta_2}{4} < \overline{\alpha}$ if and only if, $2\beta_1 < \beta_2 < 3.4798\beta_1$. The decision of the monopolist between stopping to sell good 1 or good 2 depends on the quality differential of these two goods. If good 2 has a very high quality with respect to good 1, then it is profitable to stop selling good 1 and just selling good 2. If good 2 has a higher quality but the differential is not very high, then it is more profitable to disregard quality improvement and sell only good 1.

■

Proof. (Proposition 5) First we compute the welfare of each alternative available for the monopolist.

In the uncovered market case, it is easy to obtain,

$$CS^U = \frac{1}{8}\frac{\alpha^4\beta_2^3 - 4\beta_1^2\beta_2\left(\beta_1-\beta_2\right)\left(\beta_1-2\beta_2\right)\alpha^2 + 16\beta_1^4\left(\beta_1+\beta_2\right)\left(\beta_1-\beta_2\right)^2}{\left(4\beta_1^2\left(\beta_2-\beta_1\right)-\alpha^2\beta_2\right)^2} \tag{4.14}$$

and

$$W^U = \frac{1}{8}\frac{3\alpha^4\beta_2^3 - 4\beta_1^2\beta_2\left(6\beta_2+\beta_1\right)\left(\beta_2-\beta_1\right)\alpha^2 + 48\beta_1^4\left(\beta_1+\beta_2\right)\left(\beta_2-\beta_1\right)^2}{\left(4\beta_1^2\left(\beta_2-\beta_1\right)-\alpha^2\beta_2\right)^2} \tag{4.15}$$

For the covered market solution,

$$CS^C = \frac{1}{8}\frac{\beta_2^2 + 6\beta_1\beta_2 - 2\alpha\beta_2 + 2\beta_1\alpha - 7\beta_1^2 + \alpha^2}{\beta_2-\beta_1} \tag{4.16}$$

$$W^C = \frac{1}{8}\frac{3\alpha^2 + 2\left(\beta_2-\beta_1\right)\alpha + \left(3\beta_2+5\beta_1\right)\left(\beta_2-\beta_1\right)}{\beta_2-\beta_1} \tag{4.17}$$

Also, if strict improvement is the optimal solution the consumer surplus obtains as

$$CS^{SI} = \frac{\beta_1+\beta_2}{8}. \tag{4.18}$$

And the overall welfare is

$$W^{SI} = \frac{3}{8}\left(\beta_1+\beta_2\right). \tag{4.19}$$

Finally, for the no improvement solution, the consumer surplus is given by

$$CS^{NI}_{[0,\beta_1]} = \frac{1}{2}\beta_1\frac{2\beta_1^2-\alpha^2}{\left(2\beta_1-\alpha\right)^2} \tag{4.20}$$

and

$$CS^{NI}_{[\beta_1,\infty]} = \beta_1. \tag{4.21}$$

Total welfare is

$$W^{NI}_{[0,\beta_1]} = \frac{1}{2}\beta_1\frac{6\beta_1^2-2\beta_1\alpha-\alpha^2}{\left(2\beta_1-\alpha\right)^2} \tag{4.22}$$

$$W^{NI}_{[\beta_1,\infty]} = \alpha+\beta_1 \tag{4.23}$$

Now we perform the comparison of the welfare in the different subsets of parameter's α space.

Case 1: $\alpha \in (0, \beta_1)$, we have seen that $\Pi^U > \Pi^{SI}$. Now we show that also $CS^U > CS^{SI}$. It is easy to see that CS^U can be written as $\frac{K}{8} + \frac{k}{8}(\beta_1 + \beta_2)$,where K is positive and $k > 1$, such that

$$\frac{K}{8} + \frac{k}{8}(\beta_1 + \beta_2) > \frac{1}{8}(\beta_1 + \beta_2)$$

is always verified. We can thus conclude that $W^U > W^{SI}$. Also, $\Pi^U > \Pi^{NI}$ and $CS^{SI} > CS^{NI} \Rightarrow CS^U > CS^{NI}$, so,$W^U > W^{NI}$.

Case 2: $\alpha \in (\beta_1, \overline{\alpha})$. As before, $W^U > W^{SI}$. And now, also $W^U > W^{NI}_{[\beta_1,\infty]}$, because $CS^{NI}_{[\beta_1,\infty]} < CS^{NI}_{[0,\beta_1]} < CS^U$and $\Pi^U > \Pi^{NI}_{[0,\beta_1]} > \Pi^{NI}_{[\beta_1,\infty]}$.

Case 3: $\alpha \in \left(\overline{\alpha}, \overline{\overline{\alpha}}\right)$, whenever $\beta_2 > 5.72\beta_1$, $W^{SI} > \Pi^{NI}_{[\beta_1,\infty]}$.If $3\beta_1 > \beta_2 > 5.72\beta_1$, there exists a value of α, say $\alpha^* \in \left(\overline{\alpha}, \overline{\overline{\alpha}}\right)$, such that: for $\alpha \in (\overline{\alpha}, \alpha^*)$, the solution of strict improvement dominates the no improvement and for $\alpha \in \left(\alpha^*, \overline{\overline{\alpha}}\right)$, the solution of no improvement dominates the strict improvement.

Case 4: $\alpha \in \left(\overline{\overline{\alpha}}, \beta_2 - \beta_1\right)$, first we show that $W^C > W^{SI}$

$$\begin{aligned} \frac{1}{8}\frac{3\alpha^2 + 2(\beta_2 - \beta_1)\alpha + (3\beta_2 + 5\beta_1)(\beta_2 - \beta_1)}{\beta_2 - \beta_1} - \frac{3}{8}(\beta_2 + \beta_1) &= \\ \frac{\left(2\alpha(\beta_2 - \beta_1) + 2\beta_1(\beta_2 - \beta_1) + 3\alpha^2\right)}{8(\beta_2 - \beta_1)} &> 0 \end{aligned}$$

now, we show that $W^C > W^{NI}_{[\beta_1,\infty]}$

$$W^C - W^{NI}_{[\beta_1,\infty]} = \frac{3}{8}\frac{(\alpha + \beta_1 - \beta_2)^2}{(\beta_2 - \beta_1)} > 0$$

Case 5: $\alpha \in (\beta_2 - \beta_1, \infty)$. We know that for $\alpha > \frac{1}{8}(3\beta_2 - 5\beta_1)$, $W^{NI}_{[\beta_1,\infty]} > W^{SI}$. Given that $\frac{1}{8}(3\beta_2 - 5\beta_1) < \beta_2 - \beta_1$, whenever $\alpha \in (\beta_2 - \beta_1, \infty)$ the result holds.

Therefore, as long as $\beta_2 \geq 5,72\beta_1$, the solutions chosen through welfare maximization coincide with the solutions chosen by the monopolist. ∎

References

Bensaid, Bernard and Jean-Philippe Lesne (1996), Dynamic Monopoly Pricing with Network Externalities, *International Journal of Industrial Organization*, 14, 837-855

Cabral, L. , D. Salant and G. Woroch (1999), Monopoly Pricing with Network Externalities, *International Journal of Industrial Organization*, 17, 199-214

Constantatos, C., and S. Perrakis (1997), Vertical Differentiation: Entry and Market Coverage with Multiproduct Firms, *International Journal of Industrial Organization*, 16, 81-103

Farrell, J. and G. Saloner (1986), Installed base and compatibility: innovation, product preannouncements and predation, *American Economic Review*, 7(5), 940-955

Fudenberg, D., and J. Tirole (2000), Pricing a network good to deter entry, *Journal of Industrial Economics*, 48, 373-390

Gabszewicz, J. and F. Garcia (2005), Quality improvement and network externalities, mimeo, Université Catholique de Louvain

Gabszewicz, J. and J. Thisse (1978), Price competition, quality and income disparities, *Journal of Economic Theory*, 20 (3), 340-359

Gabszewicz, J. and J. Thisse (1980), Entry (and exit) in a differentiated industry, *Journal of Economic Theory*, 22 (2), 327-338

Katz, M. and C. Shapiro (1992), Product introduction with network externalities, *Journal of Industrial Economics*, 40, 55-83

Lambertini, L. and R. Orsini (2003), Monopoly, Quality, and Network Externalities, *Keio Economic Studies*, 40(2), 1-16

Lambertini, L. and R. Orsini (2004), R&D for Quality Improvement and Network Externalities, WP 516, Dipartimento di Scienze Economiche, Università degli Studi di Bologna.

Rohlfs, J. (1974), A theory of interdependent demand for a communication service, *Bell Journal of Economics*, 5, 16-37

Shy, O. (1996), Technology Revolutions in the Presence of Network Externalities, *International Journal of Industrial Organization*, 14(6), 785-800

5

QUALITY IMPROVEMENT AND NETWORK EXTERNALITIES

5.1 Introduction

In this paper, we analyze a network industry in which an incumbent faces potential competition by a high quality brand. The incumbent benefits from the size of its installed base and decides whether or not to bar the entry of the higher quality brand and consequently, whether or not quality improvement takes place. The aim of this paper is to identify a necessary and sufficient condition under which quality improvements are spontaneously introduced, in spite of the existence of network effects.

It is frequently argued that the presence of network effects in an industry slows down technological progress. This occurs because consumers tend to prefer a good that is already in the market with an installed base rather than a new product without installed base, even if this product results from a less expensive technology. In this case, as argued by Shy (1996), the speed of technology adoption differs from the speed of technology innovation. We examine an analogous question for the case of quality innovation: in the presence of network effects, does the adoption path of higher quality goods differ from their innovation path? Clearly network effects should dampen the willingness of consumers to consume a new, higher quality product, simply because the existing one, even when of smaller quality, does benefit from the installed base. The entrant should provide a new brand whose quality sufficiently exceeds the quality of the old one in order to compensate this dampening effect, and induce consumers to switch to the better variant.

In order to understand the trade off between higher quality and network effect, we develop a two-stage business model in which an incumbent sells in the first period a certain good whose characteristics include the fact that consumers buying it in the second period may enjoy network externalities. We introduce heterogeneity in consumer's valuation for the goods in a vertical differentiation model and analyze the optimal behavior of an incumbent when a new product of higher quality becomes available in the second period. The main difference from traditional models of vertical product differentiation is the presence of network effects. The consumers are not unanimous in their preferences for the high quality good, even when prices are equal. Due to the network externality, at equal prices, some agents may prefer to exploit the network effect by buying a low quality good, rather than enjoying the benefits of a higher quality one[1]. We identify a sufficient and necessary condition under which the incumbent prefers to accommodate the higher quality brand in equilibrium rather than barring its entry. This condition says that the intensity of network effects on consumers' preferences should not exceed twice the differential of intrinsic qualities existing between the two variants.

It is important to highlight the kind of goods that our analysis comprehends. As mentioned above, we consider a two stage model where the network takes one period of time to become effective. We consider therefore a *delayed* network effect.[2] Also we assume that agents buy only in one period and cannot delay consumption. Our analysis is, in particular, appropriate for reputation goods, or goods which require a learning period. For instance, consider the market for a reputation good when reputation increases with the size of its market. The initial buyers of such a good spread information about how they enjoy its consumption, and often create a core of fans who praise its quality and virtues. The set of consumers who buy the good can be viewed as a network, improving the utility of its potential consumers and thereby the potential profits' opportunities of the firm selling the good. This is the case of several products resulting from recent innovation, and whose characteristics can only be revealed by direct consumption. The larger the set of initial buyers, the larger the utility subsequent buyers expect to obtain. Also for many goods, the size of the initial network of buyers determines the availability of complementary services the seller would be willing to develop around the

[1] The idea of analysing network effects in the framework of a vertically differentiated industry is not new (see, in particular, Baake and Boome (1999) and Bental and Spiegel (1995).

[2] The study of network externalities has been introduced for the first time in Rohlfs (1974).

variant which is initially supplied. Examples are: a new durable good and the number of shops which increases with the number of its consumers; a new machine and the number of its components when they are available to subsequent buyers, like in the case of a computer whose number of compatible programs increases with the number of its initial buyers; or, more generally, any good whose quality increases through time when the number of its buyers increases. In all such cases, the utility of the subsequent buyers for the good increases with the network size of its initial buyers. Notice however that such product improvements or reputation effects need some period of time to be finalized, creating a time lag between the creation of the network, and its effect on the utility of subsequent buyers.

The problem of the trade off between technology improvement and network effect has been treated previously by several authors. Farrell and Saloner (1986), Katz and Shapiro (1992), Shy (1996) and Fudenberg and Tirole (2000) are all concerned with the problem of quality improvement. However, they do not use the vertical differentiation model to represent how the trade off between technology improvement and network effect operates.[3] The above authors all assume that consumers have the same preferences.[4] This implies that as soon as the new quality becomes available all consumers (or none) adopt it, depending on the magnitude of the network effect at the time of innovation. By contrast, we assume in our approach that consumers are heterogeneous. Specifically, this heterogeneity allows that, at identical prices, some consumers prefer the old technology with an installed base and lower quality, while others prefer the new technology without installed base. The individual preferences of consumers in the vertical differentiation model rely on a small number of parameters capturing the essentials of the trade off between network effects and quality. This is the reason why we are able to identify precisely how the optimal strategies of the incumbent are depending on these parameters and, beyond that, when it is optimal (for him) to accommodate entry and open the door for quality improvement. There are several other significant differences between these models and ours. All these papers consider an infinite horizon model in which a new technology becomes available at a certain defined period, while we assume

[3]Baake and Boom (1999) and Bental and Spiegel (1995) have also combined the use of a vertical product differentiation model with network externalities. However, Baake and Boom (1999) focus on compatibility decisions, without addressing the quality improvement problem. Similarly Bental and Spiegel (1995) analyse the market coverage issue when network effects are present in a static framework. The closest approach to ours is Gabszewicz, Pepall and Thisse (1992), who examine the entry problem when there are switching costs in a vertical product differentiation context.

[4]In fact, Fudenberg and Tirole (2000) assume the existence of two types of consumers; however, they are homogeneous in their preferences for the technological improvement and differ only in their stand-alone values.

initially a two-period model and show that the conclusions obtained can be extended to an infinite horizon. Shy (1996) models the result of technology improvement as an increase in the stand-alone value. With this respect, Shy's approach is close to ours even though he does not take into account the consequences of this improvement on prices and, accordingly, the market participation constraint is not active.[5] Furthermore, as stated above, consumers are assumed to be homogeneous in their tastes for quality improvement. Katz and Shapiro (1992) model technological improvement through a decrease in production costs while we assume a higher stand-alone value for the new product; they analyze the pricing strategies and the timing of product introduction. Their findings are the following: when technologies are compatible, pricing is the result of a Bertrand game with differentiated costs since technological progress reduces the entrant's cost. When technologies are incompatible the incumbent sets his price so as to make consumers indifferent between buying immediately upon arrival in the market, or delaying their purchase until the entry of the new technology. The incumbent is never able to deter entry if the opponent sells a compatible product. On the contrary, if the new product is incompatible, entry deterrence is possible. Therefore, the incumbent can deter entry with certainty by denying the entrant the possibility to sell a compatible product. In our model, entry deterrence is obtained through the level of prices set in initial periods, rather than through the compatibility decision. Farrell and Saloner (1986) analyze the possibility that the incumbent prevents entry by the rival through predatory pricing; however they do not discuss the conditions under which it arises. Fudenberg and Tirole (2000) develop a model of limit pricing based on the idea that the installed base of a network good can fill a preemptive role similar to that of investment in physical capacity. In their model a sequence of potential entrants threatens the incumbent monopoly power and leads to possible entry deterrence strategy. If entry occurs, all consumers prefer the improved good and this drives the price of the incumbent to the point where he is indifferent between deterring entry or not. In our case technological progress is modeled by the creation of a new product with higher quality but a certain degree of substitutability. Price competition does not drive the incumbent's profit to the point where he is indifferent between deterring entry or not, because goods are vertically differentiated. Also, in Fudenberg and Tirole (2000) entry

[5] Shy (1996) considers technologies to be supplied at competitive prices and accordingly he does not analyse price decisions. He also treats the possibility that a monopolist produces the technology and its improvements, however he does not consider the possibility of a threatening opponent.

deterrence is always possible due to the existence of an entry cost. In our model, there is no entry cost and thus entry deterrence is necessarily the result of the pricing policy pursued during initial periods.

We describe the model in section 5.2. Section 5.3 analyses when this deterrence of the quality improvement does not take place. Some final remarks and avenues for further research are gathered in the conclusion.

5.2 The model

Consider the market for a good with positive network externalities, say good 1, sold in period 1 by a monopolist at constant average cost normalized to 0. The monopolist is assumed to be protected against entry by a patent during period 1, but free entry is assumed to be authorized in the next period. Moreover, we assume that the life cycle of product 1 does not extend beyond the end of period 2. In each period there is a cohort of consumers, indexed by a parameter v uniformly distributed on $[0,1]$, entering the market. In period 1, these consumers, who are assumed to buy at most one unit of the good, obtain utility only from its *stand-alone value:* the network effect generated by good 1 takes one period of time to become effective. Thus consumers who acquire the good in the first period do not benefit from the network externality due to the delay in the network effect. Consequently, the utility for consumer v in period 1 for buying good 1 is simply given by $\beta_1 v - p_{11}$, with $\beta_1 > 0$ measuring the quality of product 1, and p_{11} the price set by the incumbent in period 1.

In the second period, a new, higher intrinsic quality good, say good 2, becomes available for sale by an entrant, firm 2. We shall also assume a zero average cost for producing good 2. The stand alone value of good 2 for consumer v is given by $\beta_2 v$. We assume throughout that $\beta_2 \geq 4\beta_1$. The utility of type v consumer in period 2 for buying product 2 at price p_{22} is given by $\beta_2 v - p_{22}$. Notice that even if we conduct the analysis as if product 2 does not generate network effects in the subsequent period, we shall later explain that this assumption does not affect the conclusion of our analysis.

Now, the cohort of consumers entering the market in period 2 and buying good 1 not only enjoy its stand alone value, but also its *network value*, generated in the previous period. We assume that all consumers value the network effect in the same way through a parameter α. Consequently, the utility of buying product 1 in period 2 for consumer v writes as $\beta_1 v + \alpha D_{11}(p_{11}) - p_{12}$, with p_{12} denoting the price of good 1 in period 2 and $D_{11}(p_{11})$ the size of first period sales. Thus, in period 2, competition takes place between two firms: the incumbent, firm 1, selling product 1 and enjoying its network effects resulting from its first period sales and the entrant, firm 2, selling a higher intrinsic quality product. Define the *quality index* of product i as the utility obtained by consumer of type $v = 1$ at zero prices when he consumes product i. Accordingly, the quality index of good 1 in period 2 is given by $\beta_1 + \alpha D_{11}(p_{11})$ while the quality index of good 2 is equal to β_2.

In the first period, the incumbent is alone in the market and sells only good 1. A consumer v will buy if $\beta_1 v \geq p_{11}$. The demand is given by:

$$D_{11}(p_{11}) = 1 - \frac{p_{11}}{\beta_1}. \tag{5.1}$$

The monopolist serves the market in the interval $[\frac{p_{11}}{\beta_1}, 1]$. The profit of the incumbent in the first period is given by

$$\pi_{11}(p_{11}) = p_{11}\left(1 - \frac{p_{11}}{\beta_1}\right). \tag{5.2}$$

The incumbent anticipates the entry by firm 2. The goods sold by firm 1 and 2 are exogenously differentiated in quality, but also endogenously differentiated, due to the first mover advantage of firm 1. This advantage allows the incumbent to manipulate the quality of its product sold in period 2 through the choice of its first period price. In the first period, the incumbent sets price p_{11} and, in the second period, he competes in prices with the entrant. In the second period, we assume that these prices are determined at a price equilibrium of the corresponding game. However, firm 1, anticipating entry by firm 2, is able to control, through the first period price p_{11}, the competitive conditions in the market in period 2: depending on whether p_{11} is high or low, few or many consumers buy the good in period 1, and the resulting network created in the first period is small or large.

Several elements can influence consumers' choice in the second period: the quality differential between the two variants, $\beta_2 - \beta_1$, the intensity of the network effect α, the pricing policy p_{11} of the incumbent in the first period and, finally, the prices p_{12} and $_{22}$ chosen by both firms in the second period. Traditional models of vertical differentiation are characterized by the fact that at identical prices all consumers rank in the same way the different qualities. This property is no longer satisfied in the presence of network externalities. This occurs because the network externality brings a different source of quality differentiation that is independent of the type of consumer. Low types are induced to buy product 1.

Denote by $v_1(p_{12}, p_{22}; p_{11})$ the consumer who is indifferent between buying good 1 or good 2 in the second period. The value $v_1(p_{12}, p_{22}; p_{11})$ is the solution to the equation

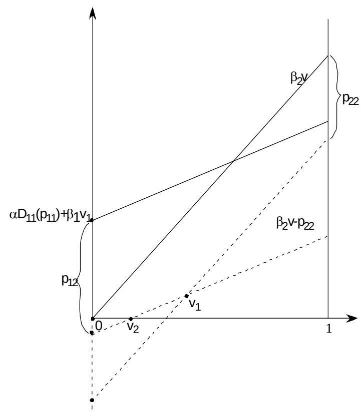
$$\beta_1 v + \alpha D_{11}(p_{11}) - p_{12} = \beta_2 v - p_{22},$$

which implies that $v_1(p_{12}, p_{22}; p_{11}) = \frac{p_{22}-p_{12}+\alpha D_{11}(p_{11})}{\beta_2-\beta_1}$. If α is taken to be 0, the value $v_1(p_{12}, p_{22}; p_{11})$ corresponds to the indifferent consumer in the traditional model of vertical differentiation. In the present case, the indifferent consumer is shifted by $\frac{\alpha D_{11}(p_{11})}{\beta_2-\beta_1}$, a measure of the relative weight of the two sources of quality differentiation. The consumer indifferent between buying good 1 or not buying at all is given by $v_2(p_{12})$,which is the solution to

$$\beta_1 v + \alpha D_{11}(p_{11}) - p_{12} = 0.$$

namely, $v_2(p_{12}; p_{11}) = \frac{p_{12}-\alpha D_{11}(p_{11})}{\beta_1}$.

FIGURE 5.1. The marginal consumers.



For both demands to be strictly positive and smaller than 1 we must have that $0 < v_2 < v_1 < 1$, which is equivalent to

$$p_{22} - p_{12} + \alpha D_{11}(p_{11}) - (\beta_2 - \beta_1) < 0 \tag{5.3}$$

$$p_{12}\beta_2 - p_{22}\beta_1 - \alpha D_{11}(p_{11})\beta_2 < 0 \tag{5.4}$$

$$p_{12} - \alpha D_{11}(p_{11}) > 0 \tag{5.5}$$

Within the area of prices (p_{12}, p_{22}) defined by the previous inequalities, we have that the demand in period 2 for good 1 is given by

$$D_{12}(p_{12}, p_{22}; p_{11}) = \frac{p_{22} - p_{12} + \alpha D_{11}(p_{11})}{\beta_2 - \beta_1} - \frac{p_{12} - \alpha D_{11}(p_{11})}{\beta_1}, \tag{5.6}$$

while the demand for good 2 obtains as

$$D_{22}(p_{12}, p_{22}; p_{11}) = 1 - \frac{p_{22} - p_{12} + \alpha D_{11}(p_{11})}{\beta_2 - \beta_1}. \tag{5.7}$$

Condition 5.5 guarantees that there exist some consumers who buy neither of the two products in the market: the market is *uncovered*. When the reverse inequality holds, all consumers buy one of the products: the market is *covered*. In that case, demands are given by

$$D_{12}(p_{12}, p_{22}; p_{11}) = \frac{p_{22} - p_{12} + \alpha D_{11}(p_{11})}{\beta_2 - \beta_1}, \tag{5.8}$$

$$D_{22}(p_{12}, p_{22}; p_{11}) = 1 - \frac{p_{22} - p_{12} + \alpha D_{11}(p_{11})}{\beta_2 - \beta_1}. \tag{5.9}$$

Notice that in the traditional model of vertical differentiation, with consumer types distributed over the interval $[0, 1]$, the market would never be covered at positive prices in a duopoly industry. This is so because the lowest-type consumer only buys if the price is set to be zero: his valuation is indeed equal to zero. When network effects are present, the consumer type zero buys even at positive prices, in order to benefit from the network. It is therefore necessary to distinguish whether after entry, demands correspond to that of a covered, or uncovered, market. Reminding that $D_{11}(p_{11}) = \left(1 - \frac{p_{11}}{\beta_1}\right)$, the corresponding second period profits for both firms obtain as

$$\begin{aligned}
\pi_{12}(p_{12}, p_{22}) &= p_{12}\left[\frac{p_{22} + \alpha\left(1 - \frac{p_{11}}{\beta_1}\right) - p_{12}}{\beta_2 - \beta_1} - \max\left\{\frac{p_{12} - \alpha\left(1 - \frac{p_{11}}{\beta_1}\right)}{\beta_1}, 0\right\}\right] \\
\pi_{22}(p_{12}, p_{22}) &= p_{22}\left[1 - \frac{p_{22} + \alpha\left(1 - \frac{p_{11}}{\beta_1}\right) - p_{12}}{\beta_2 - \beta_1}\right].
\end{aligned}$$

For the quality improvement resulting from the introduction of good 2 to take place, it is needed that the marginal consumer v_1 corresponding to the optimal price policy of the incumbent, be strictly smaller than 1, i.e. $p_{22}^* - p_{12}^* + \alpha D_{11}(p_{11}^*) < \beta_2 - \beta_1$, with the pair (p_{12}^*, p_{22}^*) denoting equilibrium values of prices in the second period and p_{11}^* the optimal price in period 1.

The subsequent analysis aims at identifying a necessary and sufficient condition to allow product improvement.

5.3 Equilibrium analysis

In this section we analyze the general case in order to identify the conditions under which quality improvement does actually take place. As mentioned above, the incumbent is able to drive the second period market equilibrium through its choice of first period price. In other words, the size of the network that he chooses to constitute is determinant in the market conditions that the entrant faces. Moreover, the incumbent might be interested in practising an aggressive strategy to limit entry by the opponent (the so-called *limit pricing strategy*). Starting with the second period market subgame, two possibilities might be pursued by the incumbent: either to allow entry, or to deter it.[6]

5.3.1 Limit pricing

We examine first the entry deterrence outcome. In order to deter entry, firm 1 must quote the price

$$p_{12}(p_{22}) = \alpha\left(1 - \frac{p_{11}}{\beta_1}\right) - (\beta_2 - \beta_1) + p_{22}$$

which cancels demand of firm 2 when it quotes a price equal to p_{22}. Entry deterrence occurs when $D_{22}(p_{12}, p_{22}) = 0$ even if $p_{22} = 0$. Then, the candidate equilibrium is the pair of prices $\left(p_{12}^L, p_{22}^L\right)$ defined by

$$\begin{aligned} p_{12}^L &= \alpha\left(1 - \frac{p_{11}}{\beta_1}\right) - (\beta_2 - \beta_1) \\ p_{22}^L &= 0. \end{aligned}$$

We immediately notice that, whatever the price p_{11} in $[0, \beta_1]$ selected in the first period, p_{12}^L is smaller than 0 when $\alpha < \beta_2 - \beta_1$. Therefore no limit pricing strategy exists for

[6] The reverse possibility where the entrant forces exit of the incumbent cannot arise. To force the exit of the incumbent the entrant should quote a price such that, even when $p_{12} = 0$, demand of firm 1 could not be positive. However, even quoting $p_{22} = 0$, $D_{12} > 0$.

the incumbent in this area of the parameters: when α is small relative to the difference in the intrinsic qualities, the cost of entry deterrence is prohibitive for the incumbent. Consequently, when $\alpha < \beta_2 - \beta_1$, *whatever the price p_{11} selected by the incumbent in period 1, he is constrained to accommodate entry in period 2.*

On the contrary, when $\alpha \geq \beta_2 - \beta_1$, it is possible to practise limit pricing, which entails a second period profit of $\alpha\left(1 - \frac{p_{11}}{\beta_1}\right) - (\beta_2 - \beta_1)$. The optimal choice of the incumbent for first period price is then given by the solution to the following optimization problem:

$$\max_{p_{11}} \Pi_{total}\left(p_{11}\right) = p_{11}\left(1 - \frac{p_{11}}{\beta_1}\right) + \alpha\left(1 - \frac{p_{11}}{\beta_1}\right) - (\beta_2 - \beta_1)$$

It is easily shown that the optimal admissible value for p_{11} is zero. [7] Thus, at the optimal first-period strategy, the incumbent obtains a total profit equal to $\alpha - (\beta_2 - \beta_1)$ when using his limit pricing strategy in period 2. We now compute the incumbent's optimal price profile for the situation in which both firms are active in the market.

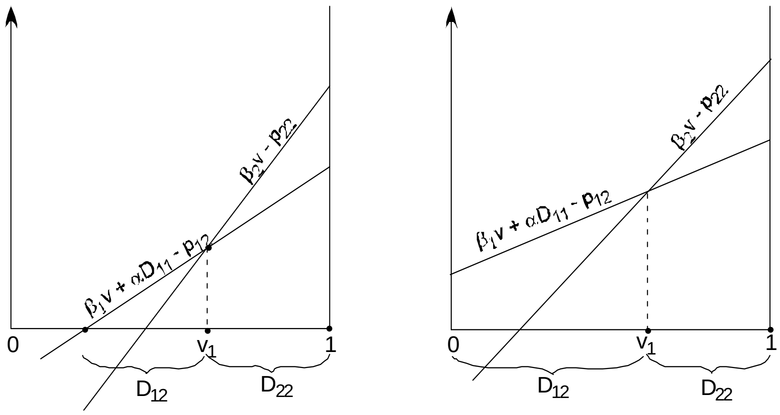
5.3.2 Entry accommodation

In traditional models of vertical product differentiation, with types distributed over the $[0, 1]$ interval, the market is always uncovered. In the presence of network externalities it is however necessary to discuss whether the optimal solution is to cover the market or not, since the existence of the network gives a positive utility to the lowest type even at positive prices. The distinction between a covered or uncovered market is important since demands are defined differently in one case or the other. This can be observed in Figure 5.2.

Consider first a pair of prices such that both firms are active in the market in period 2 but the market is *uncovered*. In this alternative, profits are given by:

[7] The total profit function is concave and attains its maximum at $p_{11} = \frac{1}{2}\left(\beta_1 - \alpha\right)$. If $\beta_2 > 4\beta_1$, as assumed in the model, the above value for p_{11} is negative and therefore, when $\beta_2 - \beta_1 < \alpha$, the optimal value obtains at $p_{11} = 0$.

FIGURE 5.2. Covered and uncovered market.



$$
\begin{aligned}
\pi_{12}\left(p_{12}, p_{22}\right) &= p_{12}\left(\frac{p_{22}+\alpha\left(1-\frac{p_{11}}{\beta_1}\right)-p_{12}}{\beta_2-\beta_1}-\frac{p_{12}-\alpha\left(1-\frac{p_{11}}{\beta_1}\right)}{\beta_1}\right) \\
\pi_{22}\left(p_{12}, p_{22}\right) &= p_{22}\left(1-\frac{p_{22}+\alpha\left(1-\frac{p_{11}}{\beta_1}\right)-p_{12}}{\beta_2-\beta_1}\right).
\end{aligned}
$$

The following necessary and sufficient conditions must be satisfied at an interior equilibrium,

$$
\begin{aligned}
\frac{\partial \pi_{12}}{\partial p_{12}} &= \frac{\beta_1 p_{22}-2p_{12}\beta_2+\alpha\left(1-\frac{p_{11}}{\beta_1}\right)\beta_2}{\beta_1\left(\beta_2-\beta_1\right)}=0 \\
\frac{\partial \pi_{22}}{\partial p_{22}} &= \frac{\beta_2-\beta_1-2p_{22}+p_{12}-\alpha\left(1-\frac{p_{11}}{\beta_1}\right)}{\beta_2-\beta_1}=0,
\end{aligned}
$$

with resulting candidate equilibrium prices

$$
\begin{aligned}
p_{12}^{*} &= \frac{\beta_1\left(\beta_2-\beta_1\right)+\alpha\left(1-\frac{p_{11}}{\beta_1}\right)\left(2\beta_2-\beta_1\right)}{4\beta_2-\beta_1}, \\
p_{22}^{*} &= \beta_2\frac{2\left(\beta_2-\beta_1\right)-\alpha\left(1-\frac{p_{11}}{\beta_1}\right)}{4\beta_2-\beta_1}.
\end{aligned}
$$

The price p_{12} is always positive while a sufficient condition for the price p_{22} to be positive is $\alpha \leq 2(\beta_2 - \beta_1)$. The demands at the interior candidate equilibrium are

$$\begin{aligned} D_{12}(p_{12}^*, p_{22}^*) &= \beta_2 \frac{(2\beta_2 - \beta_1)\alpha\left(1 - \frac{p_{11}}{\beta_1}\right) + (\beta_2 - \beta_1)\beta_1}{\beta_1(4\beta_2 - \beta_1)(\beta_2 - \beta_1)} \\ D_{22}(p_{12}^*, p_{22}^*) &= \beta_2 \frac{2(\beta_2 - \beta_1) - \alpha\left(1 - \frac{p_{11}}{\beta_1}\right)}{(4\beta_2 - \beta_1)(\beta_2 - \beta_1)}. \end{aligned}$$

The demand $D_{12}(p_{12}^*, p_{22}^*)$ is always positive while a sufficient condition for $D_{22}(p_{12}^*, p_{22}^*)$ to be positive is again $\alpha \leq 2(\beta_2 - \beta_1)$. Furthermore we must check that $D_{12}(p_{12}^*, p_{22}^*) + D_{22}(p_{12}^*, p_{22}^*) < 1$ which holds if and only if

$$\beta_2 \frac{3\beta_1 + 2\alpha\left(1 - \frac{p_{11}}{\beta_1}\right)}{(4\beta_2 - \beta_1)\beta_1} < 1. \tag{5.10}$$

To these prices and demands correspond the interior candidate equilibrium profits,

$$\begin{aligned} \pi_{12}(p_{12}^*, p_{22}^*) &= \frac{\beta_2\left(\beta_1(\beta_2 - \beta_1) + \alpha\left(1 - \frac{p_{11}}{\beta_1}\right)(2\beta_2 - \beta_1)\right)^2}{(4\beta_2 - \beta_1)^2(\beta_2 - \beta_1)\beta_1} \\ \pi_{22}(p_{12}^*, p_{22}^*) &= \beta_2^2 \frac{\left(2(\beta_2 - \beta_1) - \alpha\left(1 - \frac{p_{11}}{\beta_1}\right)\right)^2}{(4\beta_2 - \beta_1)^2(\beta_2 - \beta_1)}. \end{aligned}$$

We observe that the interior candidate price equilibrium is a function of the price p_{11} selected by the incumbent in period 1, which determines the network size in period 2.

Now assume that interior candidate equilibrium prices lead to a *covered* market. The profits are then given by:

$$\begin{aligned} \pi_{12}(p_{12}, p_{22}; p_{11}) &= p_{12}\left(\frac{p_{22} + \alpha\left(1 - \frac{p_{11}}{\beta_1}\right) - p_{12}}{\beta_2 - \beta_1}\right) \\ \pi_{22}(p_{12}, p_{22}; p_{11}) &= p_{22}\left(1 - \frac{p_{22} + \alpha\left(1 - \frac{p_{11}}{\beta_1}\right) - p_{12}}{\beta_2 - \beta_1}\right). \end{aligned}$$

The first order conditions imply that at an interior candidate price equilibrium:

$$\begin{aligned} \frac{p_{22} - 2p_{12} + \alpha\left(1 - \frac{p_{11}}{\beta_1}\right)}{(\beta_2 - \beta_1)} &= 0 \\ \frac{(\beta_2 - \beta_1) + (p_{12} - 2p_{22}) - \alpha\left(1 - \frac{p_{11}}{\beta_1}\right)}{(\beta_2 - \beta_1)} &= 0, \end{aligned}$$

with corresponding prices,

$$p_{22} = \frac{1}{3}\left(2\left(\beta_2 - \beta_1\right) - \alpha\left(1 - \frac{p_{11}}{\beta_1}\right)\right),$$
$$p_{12} = \frac{1}{3}\left(\alpha\left(1 - \frac{p_{11}}{\beta_1}\right) + \beta_2 - \beta_1\right)$$

and equilibrium profits given by

$$\pi_{22}\left(p_{11}\right) = \frac{1}{9}\frac{\left(2(\beta_2-\beta_1)-\alpha\left(1-\frac{p_{11}}{\beta_1}\right)\right)^2}{\beta_2-\beta_1},$$
$$\pi_{12}\left(p_{11}\right) = \frac{1}{9}\frac{\left(\alpha\left(1-\frac{p_{11}}{\beta_1}\right)+\beta_2-\beta_1\right)^2}{\beta_2-\beta_1}.$$

In order to identify the optimal profile of prices for the incumbent, we have now to analyze his first period optimal strategy both for the case of covered and uncovered second-period market. We have assumed in the model that $\beta_2 > 4\beta_1$. As it is shown in the proof of the following lemma, this assumption gives a sufficiently high intrinsic quality differential between the two goods in order to induce the incumbent to create the largest network in the first period by quoting a zero price.

Lemma 1 *When $\beta_2 - \beta_1 < \alpha$ and the incumbent accommodates entry of firm 2, its optimal first period price is given by $p_{11}^* = 0$.*

Proof: See the appendix.

From Lemma 4.1, it follows that, under entry accommodation, the incumbent always chooses to maximize his installed base in the first period by quoting price zero, both in the case of covered and uncovered market. It is easy to see however, that in the same domain, the condition (5.10) guaranteeing the existence of an uncovered market at equilibrium prices in period 2, is never met when $p_{11} = 0$. Thus, we can rule out the possibility that the incumbent would drive the equilibrium path to an equilibrium with an uncovered market in stage 2. Thus, under entry accommodation, only the total profit corresponding to a zero price in period 1 and the market coverage solution in period 2 has to be taken into account by the incumbent when deciding whether or not to deter entry, namely

$$\Pi_{total}\left(0\right) = \frac{1}{9}\frac{\left(\alpha + \beta_2 - \beta_1\right)^2}{\beta_2 - \beta_1}.$$

5.3.3 Total profit comparisons

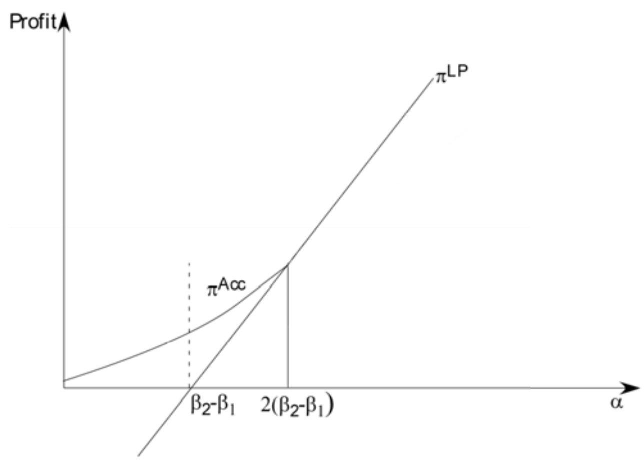
Now we can compare the total profits under limit pricing and entry accommodation, always having in mind to identify the optimal pricing profile of the incumbent. This comparison has to be performed only when $\alpha \in [\beta_2 - \beta_1, 2(\beta_2 - \beta_1)]$. Indeed, when $\alpha < \beta_2 - \beta_1$, no limit pricing is possible while, when $\alpha \geq 2(\beta_2 - \beta_1)$, no accommodation is possible at equilibrium since demand of firm 2 is negative at equilibrium prices. It is easy to see that when $\alpha \in [\beta_2 - \beta_1, 2(\beta_2 - \beta_1)]$,

$$\Pi_{total}(0) = \frac{1}{9} \frac{(\alpha + \beta_2 - \beta_1)^2}{\beta_2 - \beta_1} > \alpha - (\beta_2 - \beta_1),$$

where the last term of this inequality represents the total profit for the entry deterrence strategy: in the domain $\alpha \in [\beta_2 - \beta_1, 2(\beta_2 - \beta_1)]$, the incumbent gets a higher total profit by accommodating entry than by practising a limit pricing strategy (see Figure 5.3).

Consequently, for all $\alpha \in [0, 2(\beta_2 - \beta_1)]$, the optimal strategy for the incumbent consists in accommodating entry. Thus we conclude:

FIGURE 5.3. Profits under entry deterrence and entry accommodation.



Proposition 1 *A necessary and sufficient condition for quality improvement is*

$$\alpha \leq 2(\beta_2 - \beta_1).$$

The intensity of the network effect in the preferences of the consumers should not be too large compared with the intrinsic quality differential between the two variants: otherwise entry deterrence would necessarily take place. In fact, when $\alpha \geq 2(\beta_2 - \beta_1)$, the difference between the network effect intensity α and the stand-alone values differential $\beta_2 - \beta_1$, makes it more advantageous to the incumbent to prevent both products to compete with each other in the second-period market. The size of the network effect amplifies the intrinsic quality of the standard variant sold in period 1 to such an extent that it becomes the product with the highest quality index in period 2. It is easy to anticipate that lowering the amplitude of the network effect below the differential of the intrinsic quality parameters can make the use of the entry deterrence strategy less attractive, and give room for quality improvement adoption.

This result relates directly to the main findings in Garcia (2005). When a monopolist sells a good with network externalities and has the possibility of introducing a higher quality brand in the market, Garcia (2005) shows that quality improvement takes place if and only if the network intensity, α, is smaller than the quality differential, $\beta_2 - \beta_1$. The threshold is thus higher when the new quality is sold by a rival than if it is sold by the monopolist himself. Quality improvement is more likely under competition than under monopoly and the reason is that, for some values of the network effect, it is too costly to deter the entry of a rival. On the contrary, under monopoly, it is costless to deter the introduction of the high quality.

The above analysis has been conducted assuming that the life cycle of good 1 has two periods, and that good 2 would not generate network externalities in subsequent periods. The first assumption guarantees that the horizon of the incumbent does not extend beyond period 2. Otherwise he would have to take into account all subsequent periods when defining his optimal pricing strategy. The resulting increase in complexity would not be compensated by the development of new insights in the problem. We will now argue that the second assumption does not play any role in our main conclusion. When $\alpha > 2(\beta_2 - \beta_1)$, the incumbent practises limit pricing, which entails a zero demand for firm 2 and, accordingly, no network can be constituted having effect in subsequent periods. If $\alpha \leq 2(\beta_2 - \beta_1)$, the incumbent is willing to accommodate entry when the rival sells a good without network effect in the subsequent periods. If there would exist network effects related to this second good, the rival would certainly quote in the subsequent periods a lower price in period 2 than the equilibrium price without network effects which we have identified above. In particular he could possibly set a price equal to zero in order to create the largest installed base for period 3. Since the incumbent accommodates entry in the latter situation, he will *a fortiori* accommodate entry when his opponent can constitute a network. Thus, the condition $\alpha \leq 2(\beta_2 - \beta_1)$ remains necessary and sufficient for quality improvement. Notice that if the condition $\alpha > 2(\beta_2 - \beta_1)$ is satisfied, the entry of product 2 is barred and firm 1 remains a monopolist over periods 1 and 2, disappearing at the end of the latter. Assuming that the barred product can no longer enter the market, product 3 is the only candidate for sale in period 3. When the condition $\alpha > 2(\beta_4 - \beta_3)$ also holds, firm 3 remains a monopolist over periods 3 and 4, and disappears at the end of period 3. This reasoning can be pursued for periods beyond period 4, as long as the condition for entry deterrence is satisfied. The market structure is a sequence of monopolies with effective quality improvement arising every two periods due to the bounded life cycle of the variants.

5.4 Welfare Comparisons

It is interesting to identify the circumstances under which the use of an accommodating strategy pareto-dominates the practise of limit pricing. This comparison is only meaningful in the domain of parameters in which the incumbent disposes of both possibilities, namely when $\alpha \in (\beta_2 - \beta_1, 2(\beta_2 - \beta_1))$. Remind that the optimal first period strategy of the incumbent always consists in quoting $p_{11} = 0$. Accordingly, under limit pricing, the total surplus W^{LP} over the two periods writes as

$$\begin{aligned} W^{LP} &= CS_1^{LP} + CS_2^{LP} + \Pi_1^{LP} + \Pi_2^{LP} = \\ &= \int_0^1 (\beta_1 v)\, dv + \int_0^1 (\beta_1 v + \alpha - p_{12}^*)\, dv + p_{12}^* = \\ &= \alpha + \beta_1, \end{aligned}$$

with CS_i^{LP} denoting the consumers' surplus of period i and Π_i^{LP} the profit of period i under limit pricing. Likewise, the total surplus of the accommodating strategy obtains as

$$\begin{aligned} W^{Acc} &= CS_1^{Acc} + CS_2^{Acc} + \Pi_1^{Acc} + \Pi_2^{Acc} \\ &= \int_0^1 (\beta_1 v)\, dv \\ &\quad + \int_{v_1(p_{12}^*, p_{22}^*)}^1 (\beta_2 v - p_{22}^*)\, dv + \int_0^{v_1(p_{12}^*, p_{22}^*)} (\beta_1 v + \alpha - p_{12}^*)\, dv \end{aligned}$$

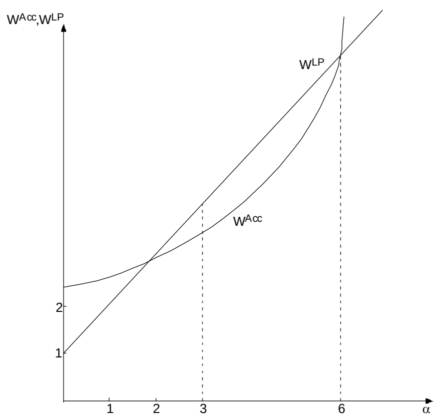
$$+p_{12}^{*}\left(\frac{p_{22}^{*}-p_{12}^{*}+\alpha}{\beta_2-\beta_1}\right)+p_{22}^{*}\left(1-\frac{p_{22}^{*}-p_{12}^{*}+\alpha}{\beta_2-\beta_1}\right)=$$
$$=\frac{1}{18}\frac{5\alpha^2+4\alpha\left(\beta_2-\beta_1\right)+2\left(5\beta_1+4\beta_2\right)\left(\beta_2-\beta_1\right)}{\beta_2-\beta_1},$$

with CS_i^{Acc} denoting the consumers' surplus of period i and Π_i^{Acc} the profit of period i under accommodation. The difference between the total surpluses obtains as:

$$W^{Acc}-W^{LP}=\frac{1}{18}\frac{8\left(\beta_2-\beta_1\right)^2+5\alpha^2-14\alpha\left(\beta_2-\beta_1\right)}{\beta_2-\beta_1},$$

which is negative whenever $\alpha\in\left(\frac{4}{5}\left(\beta_2-\beta_1\right),2\left(\beta_2-\beta_1\right)\right)$ and positive otherwise. As an illustration, Figure 5.4 depicts the total surplus under limit pricing and accommodation as a function of the network intensity parameter α for $\beta_2=4$ and $\beta_1=1$.

FIGURE 5.4. Welfare.



When $\alpha\in\left(\frac{4}{5}\left(\beta_2-\beta_1\right),2\left(\beta_2-\beta_1\right)\right)$, it would be better from the welfare point of view that only product 1 is available to consumers. However, at the equilibrium solution of period 2, the incumbent chooses to accommodate entry when $\alpha\in\left(\left(\beta_2-\beta_1\right),2\left(\beta_2-\beta_1\right)\right)$, thus making the suboptimal decision concerning its welfare consequences. In the referred domain, if total welfare is decomposed into its consumers' and firms' components, the presence of both products in the market leads to higher prices than in the limit pricing solution, which tends to reduce consumers' surplus. The positive difference in profits between accommodating and deterring does not compensate for the loss in consumers' surplus.

5.5 Conclusion

We have raised the question whether the existence of network effects could prevent the quality improvements resulting from the increase in the stand alone value. Even if this issue has been treated previously in the literature, it has never been examined in the context of a vertical product differentiation model. This exploration enables us to identify a very simple necessary and sufficient condition under which intrinsic quality improvements are effectively realized at equilibrium. This condition says that the intensity of the network effect should be smaller than twice the differential between the intrinsic qualities of the entrant's and incumbent's variants. We have also shown that for some combination of parameters, the monopolist's decision concerning quality improvement is suboptimal.

Several questions naturally arise from our work. One possibility would consist in keeping the model considered in this paper and analyzing the dynamics of entry assuming that the condition for quality improvement is always fulfilled. In this case, two products can simultaneously survive in the market in each period, giving rise to a sequence of duopolistic market structures.

5.A Appendix

Proof. (Lemma 1): Assume first that the market is uncovered; the profit function is then given by

$$\pi_{tot}(p_{11}) = \frac{\beta_2\left(\beta_1(\beta_2-\beta_1)+\alpha\left(1-\frac{p_{11}}{\beta_1}\right)(2\beta_2-\beta_1)\right)^2}{(4\beta_2-\beta_1)^2(\beta_2-\beta_1)\beta_1} + p_{11}\left(1-\frac{p_{11}}{\beta_1}\right).$$

This function is quadratic in p_{11} and concave (resp. convex) when α belongs to the domain $(0,\hat{\alpha})$ with $\hat{\alpha} = \frac{\beta_1}{\beta_2}\sqrt{\beta_2(\beta_2-\beta_1)}\frac{4\beta_2-\beta_1}{2\beta_2-\beta_1}$(resp. $(\hat{\alpha},\infty)$). Whenever $4\beta_2 - 13\beta_1 > 0$, $\hat{\alpha} < \beta_2 - \beta_1$ and thus in the relevant domain the function $\pi_{tot}(p_{11})$ is convex. The maximum must obtain at one of the boundaries of the admissible domain $[0,\beta_1]$ for p_{11}. Comparing these profits at the two extreme values of the domain, it is easily seen that $\pi_{tot}(p_{11})$ is always maximal at $p_{11}^* = 0$.Now assume that p_{11} entails a covered market in period 2. Then the incumbent chooses p_{11} that maximizes:

$$\Pi_{total}(p_{11}) = \frac{1}{9}\frac{\left(\alpha\left(1-\frac{p_{11}}{\beta_1}\right)+\beta_2-\beta_1\right)^2}{\beta_2-\beta_1} + p_{11}\left(1-\frac{p_{11}}{\beta_1}\right).$$

This function is convex whenever $\alpha > 3\sqrt{\beta_1(\beta_2-\beta_1)}$. In that case is easily shown that the maximum is attained at $p_{11} = 0$. Assume first that $3\sqrt{\beta_1(\beta_2-\beta_1)} < \beta_2 - \beta_1$, in which case $\beta_2 > 10\beta_1$. Then for $\alpha \in [\beta_2-\beta_1, 2(\beta_2-\beta_1)]$, the profit function is convex and $p_{11} = 0$. Now assume that $\beta_2 - \beta_1 < 3\sqrt{\beta_1(\beta_2-\beta_1)} < 2(\beta_2-\beta_1)$, in which case $\frac{13}{4}\beta_1 < \beta_2 < 10\beta_1$. Then for $\alpha \in \left[3\sqrt{\beta_1(\beta_2-\beta_1)}, 2(\beta_2-\beta_1)\right]$ the function is convex and $p_{11} = 0$. For $\alpha \in \left[\beta_2-\beta_1, 3\sqrt{\beta_1(\beta_2-\beta_1)}\right]$the profit function is concave and its first derivative vanishes at $p_{11} = \frac{1}{2}\beta_1\frac{2\alpha(\alpha+(\beta_2-\beta_1))-9\beta_1(\beta_2-\beta_1)}{\alpha^2-9\beta_1(\beta_2-\beta_1)}$, which is negative in the relevant domain of parameters. Accordingly, the optimal value of p_{11} is once again 0. QED ■

References

Baake, P. and A. Boom (1999), Vertical Product Differentiation, Network Externalities, and Compatibility Decisions, Mimeo, Humboldt University.

Bental, B. and M. Spiegel (1995), Network competition, product quality, and market coverage in the presence of network externalities, *Journal of Industrial Economics*, 43, 197-208

Farrell, J. and G. Saloner (1986), Installed base and compatibility: innovation, product preannouncements and predation, *American Economic Review*, 7(5), 940-955

Fudenberg, D. and J. Tirole (2000), Pricing a Network Good to Deter Entry, *Journal of Industrial Economics*, 48, 373-390

Gabszewicz, J.J., Pepall L. and J.-F. Thisse (1992), Sequential Entry with Brand Loyalty Caused by Consumer Learning-by-Using, *Journal of Industrial Economics*, 40, 397-416

Garcia, F. (2005), When does a monopoly improve its quality?, Mimeo, CORE.

Katz, M. and C. Shapiro (1992), Product introduction with network externalities, *Journal of Industrial Economics*, 40, 55-83

Rohlfs, J. (1974), A theory of interdependent demand for a communication service, *Bell Journal of Economics*, 5, 16-37

Shy, O. (1996), Technology Revolutions in the Presence of Network Externalities, *International Journal of Industrial Organization*, 14(6), 785-800

6

A NOTE ON EXPANDING DEMAND AND MONOPOLY PRICING

6.1 Introduction

In this paper we analyze the pricing decision of a monopolist who sells a good subject to network externalities for a finite number of periods. We consider that the network takes one period of time to be constituted so that consumers derive utility in each period from the volume of past sales (see Rohlfs (1974), Katz and Shapiro (1985, 1994) and Farrell and Saloner (1985)). This type of network externalities can be found when the quality of the good depends on some "word of mouth" or "learning by doing" process, according to the terminology in Bensaid and Lesne (1996). Early consumers allow firms to improve their product, or to offer better services and assistance in later periods, which increases the benefit of consuming the good for later users. Our analysis is, in particular, appropriate for reputation goods (see Rogerson (1982)), or goods which require a learning period. For instance, consider the market for a reputation good when reputation increases with the size of its market. The initial buyers of such a good spread information about how to use it and enjoy its consumption. The group of consumers who buy the good initially can be viewed as a network, improving the utility of its potential consumers and thereby the potential profits opportunities of the firm. This is the case of several products resulting from recent innovation, and whose characteristics can only be revealed by direct consumption. Also for many goods, the size of the initial network of buyers determines the availability of complementary services the seller would be willing to develop around the variant which is initially supplied. Examples are: a bank and the number of branches which increases with the number of its customers, a new machine and the number of its components when they are available to subsequent buyers, like in the case of a computer whose number of compatible programs increases with the number of its initial buyers. Notice however that such product improvements or reputation effects need some period of time to be finalized, creating a time lag between the creation of the network, and its effect on the utility of subsequent buyers. This structure of preferences implies that consumers always consider past consumers' decisions before deciding to acquire the good or not. Consequently, the monopolist's pricing decision must consist in an intertemporal strategy: the price chosen in early periods influences future sales and profits.

In this note we uncover the monopolist's optimal strategy and study how it depends on the intensity of the preference for the network and on the horizon's length. In initial periods, the network effect is mild and thus the surplus that the monopolist could extract from the consumers by setting the instantaneous monopoly price is low. On the contrary, setting a lower price attracts consumers and increases the demand in future periods. The incentive to set low prices is, nevertheless decreasing in time because the closer the horizon, the weaker the interest of increasing further future demand. As a consequence, we first observe that the path of optimal prices is increasing through time. We also observe that this path of prices approaches the sequence of instantaneous monopoly prices corresponding to the sequence of accumulated demands, while being exactly equal to this monopoly price in the last period.

Furthermore,for large values of network intensity, we show that the monopolist chooses initial prices to be zero, indicating that he would even set a negative price if this would be admissible. We also show that, in this context, the monopolist selects a path of prices which increases through time, and that this path is steeper than when the network intensity is weaker.

The initial zero price phenomenon is frequently observed. Banks usually offer low rates for new users, software is usually offered at very low price (or, even, given for free) in initial periods (whereas updates are expansive); similarly, phone companies or network providers have low price deals for new adherents (See Ackere and Reyniers (1995)). In this note, we provide a necessary and sufficient condition under which the initial price is zero, implying that the seller has interest in offering free sample of the good in the first period. To the best of our knowledge, this condition has not been discussed so far in the literature.

Our note relates to previous literature on the pricing of durable goods and experience goods. In particular, Cabral *et al.* (1999) study monopoly pricing in a network industry and conclude that introductory pricing occurs if consumers are price takers, if there is incomplete information about demand or asymmetric information about cost. We show that introductory pricing occurs in a setup where information is complete and symmetric. Clarke *et al.* (1982) discuss, within a continuous time framework, the consequences on monopoly pricing of experience effects in demand. They are also led to the conclusion that the path of prices should be increasing through time. Dhebar and Oren (1985,1986), analyzing linear and nonlinear pricing strategies for the monopolist, also conclude that his profit margins should be increasing. Our note contributes to this discussion by providing explicitly the optimal path of prices. We also obtain comparative statics results related to the length of the horizon and to the intensity of the network.

6.2 The model and results

Consider a profit maximizing monopolist that sells a durable good at constant average cost set to zero. The monopolist will be active in the market for T periods.

At each period $t, t = 1, ..., T$,a new cohort of consumers enters the market. Consumers value the good for its intrinsic utility and for the network benefits that it entails. The more consumers buy the good the higher is its value for each individual. Consumers are differentiated according to their *stand alone value.* They are distributed uniformly in $[0, 1]$ by increasing order of intrinsic valuation for the good.

The demand for the good at time t, is denoted by $D_t(p_t, p_{t-1}, ..., p_1)$. The utility of agent v at time t is given by:

$$u_t\left(v; p_{t-1}, ..., p_1\right) = v + \alpha \sum_{j=1}^{t} D_{t-j}\left(p_{t-j}, ..., p_1\right); \tag{6.1}$$

The parameter α, $\alpha > 0$, measures the intensity of the network effect.

First, we derive the demand function for the product in each period t, $t = 1, ..., T$. Since the consumer, in her decision problem at period t, only considers the network constituted in the past periods, the functional form of demand in each period is independent of the subsequent periods and depends only on prices p_τ, $\tau = 1, ...t$. Let us start with $T = 2$. The demand at period 1 is given by:

$$D_1(p_1) = 1 - p_1, 1 \geq p_1 \geq 0. \tag{6.2}$$

In period 2, taking into account the network effect created in period 1, we identify the consumer $\widehat{v}$ who is indifferent between buying the product and not buying in period 2 by the condition

$$v + \alpha(1 - p_1) - p_2 = 0,$$

namely, $\widehat{v} = p_2 - \alpha(1 - p_1)$. Thus, demand in period 2 is given by:

$$D_2(p_2, p_1) = 1 - p_2 + \alpha(1 - p_1). \tag{6.3}$$

Likewise, in period 3 and taking into account the network effect in periods 1 and 2, the consumer who is indifferent between buying the product and not buying in period 3, denote it by $\widehat{\widehat{v}}$,is given by the condition

$$v + \alpha(1 - p_1 + 1 - p_2 + \alpha(1 - p_1)) - p_3 = 0,$$

namely, $\widehat{\widehat{v}} = p_3 - \alpha(1 - p_1 + 1 - p_2 + \alpha(1 - p_1))$. Thus, the demand in period 3 obtains as

$$D_3(p_3, p_2, p_1) = 1 - p_3 + \alpha(1 - p_1 + 1 - p_2 + \alpha(1 - p_1)).$$

Proceeding by induction, it is easy to see that the demand in period $t \leq T$ obtains as

$$D_t(p_t, ..., p_1) = (1 + \alpha)^{t-1} - \alpha \sum_{j=2}^{t} (1 + \alpha)^{t-j} p_{j-1} - p_t.$$

Notice that, when $\alpha > 1$, the demand might react with more intensity to previous prices, p_2 and p_1, than to current price. This is so because, under this condition, lower initial prices increase the value of the good for subsequent users. On the contrary, for $\alpha < 1$, the effect of initially low prices tends to vanish with time. An alternative expression for the demand obtains as

$$D_t(p_t, ..., p_1) = 1 - p_t + \alpha \sum_{j=1}^{t} D_{t-j}.$$

The demand in period t is a linear decreasing function of p_t with the constant term increasing in t. All the above expressions for demands at period t must of course satisfy the fact that demand can never exceed 1.

The *total* profit function with horizon T writes as:

$$\begin{aligned}\Pi(T; p_T, ..., p_1) &= \sum_{k=1}^{T} p_k D_k \\ &= \sum_{k=1}^{T} p_k \left((1+\alpha)^{k-1} - \alpha \sum_{j=2}^{k} (1+\alpha)^{k-j} p_{j-1} - p_k \right).\end{aligned}$$

6.3 "Weak" network intensity

Consider first that the condition

$$\alpha < \frac{1}{T-1} \qquad \text{(C1)}$$

is satisfied by the parameters of the model.

Proposition 1 *Under C.1, the set of values given by*

$$p_t(T) = \frac{1-(T-t)\alpha}{2-(T-1)\alpha}, \ t = 1, ..., T \qquad (6.4)$$

constitutes the unique solution to the monopolist's profit maximization problem.

Proof. We differentiate $\Pi(T; p_T, ..., p_1)$ with respect to $p_1, ..., p_T$ and, by the first order necessary conditions for an interior maximum, we get a system of T equations defined by

$$\frac{\partial \Pi(T; p_T, ..., p_1)}{\partial p_s} = 0, \ s = 1, ..., T \qquad (6.5)$$

or, equivalently

$$(1+\alpha)^{s-1} - \alpha \sum_{j=2}^{s} (1+\alpha)^{s-j} p_{j-1} - 2p_s - \alpha \sum_{i=1}^{T-s} p_{s+i} (1+\alpha)^{i-1} = 0, s = 1, ...T.$$

It is easy to see that, the vector of prices (6.4) is the unique solution to the above system whenever condition C.1 holds (notice that, $p_t(T)$ increases with t; accordingly, $p_1(T) = \frac{1-(T-1)\alpha}{2-(T-1)\alpha}$ is positive if, and only if, C.1 is satisfied). Now, it remains to check when the first order necessary conditions identified above are also sufficient for (6.4) being a solution to the maximization problem of the monopolist. The Hessian matrix associated with the maximization problem writes as

$$A(T) = -\begin{bmatrix} 2 & \alpha & \alpha(1+\alpha) & \ldots & \alpha(\alpha+1)^{T-2} \\ \alpha & 2 & \alpha & \ldots & \alpha(\alpha+1)^{T-3} \\ \alpha(1+\alpha) & \alpha & 2 & \ldots & \alpha(\alpha+1)^{T-4} \\ \ldots & \ldots & \ldots & 2 & \alpha \\ \alpha(\alpha+1)^{T-2} & \alpha(\alpha+1)^{T-3} & \alpha(\alpha+1)^{T-4} & \alpha & 2 \end{bmatrix}$$

The total profit function $\Pi(T; p_T, ..., p_1)$ is strictly concave iff,

$$\begin{aligned} \det(A(T)) &= (-1)^T((T-1)\alpha - 2)(\alpha+2)^{T-1} < 0, T \text{ even} \\ & \quad \textit{or}, \\ \det(A(T)) &= (-1)^T((T-1)\alpha - 2)(\alpha+2)^{T-1} > 0, T \text{ odd} \end{aligned}$$

or, equivalently

$$\alpha < \frac{2}{T-1}. \tag{C.2}$$

This condition is trivially satisfied when α satisfies C.1 and it also guarantees the uniqueness of the solution to the maximization problem. ■

To illustrate the above, the following table represents the sequence of optimal prices corresponding to various values of the horizon T.

$T=2$	$\frac{1-\alpha}{2-\alpha}$	$\frac{1}{2-\alpha}$			
$T=3$	$\frac{1-2\alpha}{2-2\alpha}$	$\frac{1-\alpha}{2-2\alpha}$	$\frac{1}{2-2\alpha}$		
$T=4$	$\frac{1-3\alpha}{2-3\alpha}$	$\frac{1-2\alpha}{2-3\alpha}$	$\frac{1-\alpha}{2-3\alpha}$	$\frac{1}{2-3\alpha}$	
...	...	...	...	...	
T	$\frac{1-(T-1)\alpha}{2-(T-1)\alpha}$	$\frac{1-(T-2)\alpha}{2-(T-1)\alpha}$	$\frac{1-(T-3)\alpha}{2-(T-1)\alpha}$	...	$\frac{1}{2-(T-1)\alpha}$

It is interesting to discuss briefly how this solution depends on the main ingredients of the model, namely, the time horizon T and the network intensity parameter α. In order to perform the desired comparative statics we easily compute the value of demand D_t, $t = 1, ..., T$, at optimal prices $p_t(T), ..., p_1(T)$, namely

$$D_t(p_t(T), ..., p_1(T)) = \frac{1}{2-(T-1)\alpha}.$$

First notice that the value of each demand at period t evaluated at optimal prices is smaller than 1. Furthermore, we find that this value is constant, given the horizon T. Finally, it is increasing with the horizon T and decreasing with α.

The period t-profits at the optimal solution $\pi_t(T)$ obtain as

$$\begin{aligned}\pi_t(T) &= p_t(T) \cdot D_t(p_t(T), ..., p_1(T)) \\ &= \frac{1-(T-t)\alpha}{(2-(T-1)\alpha)^2}.\end{aligned}$$

Therefore, the per-period profits $\pi_t(T)$ are increasing over time and also increasing with the network intensity α (under C.1).

Total profits evaluated at optimal prices, given the horizon T, obtain as

$$\Pi(T, p_t(T), ..., p_1(T)) = \sum_{t=1}^{T} \pi_t(T) = \frac{T}{2(2-(T-1)\alpha)}.$$

An immediate conclusion is that total profits are an increasing function of the horizon, as well as of the network intensity α.

6.4 "Strong" network intensity

By "strong network intensity", we mean that condition C.1 is violated. First, let us assume that

$$\frac{1}{T-1} \leq \alpha \leq \frac{2}{T-1}. \tag{C.3}$$

As shown above, total profits are also strictly concave in this domain, but some, or all, components of the vector $\left\{p_t(T) = \frac{1-(T-t)\alpha}{2-(T-1)\alpha}\right\}_{t=1}^{T}$ must be equal to 0: all prices are to be nonnegative. When α increases beyond the value $\frac{1}{T-1}$,the first component of this vector to become smaller or equal to 0 at (6.4)is $p_1(T)$. Then, how to characterize the optimal path of prices for $p_2(T), p_3(T), ..., p_T(T)$?

Proposition 2 *Under C.3, the set of values given by*

$$\begin{aligned}p_1(T) &= 0 \\ p_t(T) &= (t-1)\alpha, \ t = 2, ..., T\end{aligned} \tag{6.6}$$

constitutes the unique solution to the monopolist's profit maximization problem.

Proof. First, notice that the domain of values for which $p_1(T) = \frac{1-(T-1)\alpha}{2-(T-1)\alpha} \leq 0$, namely, $\left\{\alpha : \frac{1}{T-1} < \alpha < \frac{2}{T-1}\right\}$, strictly includes the domain for which $p_1(T+1) \leq 0$. Consequently, the strongest condition for having $p_1(T) = 0$ obtains for the horizon $T = 2$ so that, whatever the value of the horizon $T \geq 2$, we must have $p_1(T) = 0$. We shall now prove the proposition by induction. Consider first $T = 2$. Then, the second and third period demands are equal to:

$$\begin{aligned} D_2(p_2, p_1) &= 1 - p_2 + \alpha, \\ D_3(p_3, p_2, p_1) &= 1 - p_3 + \alpha(1 - p_2 + \alpha). \end{aligned}$$

We show that, under condition C.3, the optimal second period price $p_2(2)$ is equal to α. Total profit $\Pi(2, p_2, 0)$ is then equal to $1 + p_2(1 - p_2 + \alpha)$. The first-order condition writes as

$$1 - 2p_2 + \alpha = 0,$$

which is satisfied for $p_2 = \frac{1+\alpha}{2}$. Substituting this value in $D_2(p_2, p_1)$,we see that $D_2 = \frac{1+\alpha}{2}$, which exceeds 1, given that, by condition C.3, we have $\alpha > 1$. Accordingly, this value should be discarded as a solution to the maximization problem, and the optimal solution obtains at the highest value for p_2 keeping the demand smaller than 1, namely, $p_2 = p_2(2) = \alpha$.

Finally we must show that, if the path of prices $(0, \alpha, 2\alpha, ...(T-2)\alpha)$ is optimal up to $T-1$, then the optimal price path up to T is given by the sequence $(0, \alpha, 2\alpha, ...(T-1)\alpha)$. So, assume that $(0, \alpha, 2\alpha, ...(T-2)\alpha)$ is the optimal sequence of prices when the horizon is $T-1$. The demand function D_T then writes as $1 - p_T + \alpha(T-1)$, with corresponding total profits equal to $p_T(1 - p_T + \alpha(T-1))$.The first order condition writes as

$$1 - 2p_T + \alpha(T-1) = 0,$$

which is satisfied for $p_T = \frac{1+\alpha(T-1)}{2}$. Substituting this value in D_T, we see that, by condition C.3 ($\alpha \geq \frac{1}{T-1}$), the corresponding demand exceeds the value 1,which is not admissible. Then, $p_T(T) = \alpha(T-1)$, which is the price that corresponds to demand $D_T = 1$. ■

Now, consider the case in which condition

$$\alpha > \frac{2}{T-1} \tag{C.4}$$

holds. Then, the total profit function is no longer concave, but convex. Accordingly, the solution to the monopolist's problem occurs at the boundary of the admissible domain of prices, namely $\{(p_1, ..., p_T) : 0 \leq D_t(p_t, ..., p_1) \leq 1, t = 1, ..., T\}$. In what follows we show:

Proposition 3 *Under C.4, the set of values given by*

$$\begin{aligned} p_1(T) &= 0 \\ p_t(T) &= (t-1)\alpha, \;\; t = 2, ..., T \end{aligned} \tag{6.7}$$

constitutes the unique solution to the monopolist's profit maximization problem.

Proof. Suppose first that $T = 2$ and that C.4 is satisfied, namely, $\alpha \geq \frac{1}{2}$. Then, the total profit function of the monopolist writes as:

$$\Pi(2, p_2, p_1) = p_1(1 - p_1) + p_2(1 - p_2 + \alpha(1 - p_1))$$

which is a convex function in p_1 and p_2. The optimal solution must lie on the boundary, namely, $p_1 = 0$ and $p_2 = \alpha$, the prices for which both demands are equal to 1. Then proceeding as in the proof of Proposition 2, it is easy to show that the sequence of prices $0, \alpha, ..., (T-1)\alpha$ constitutes the optimal monopoly solution when C.4 holds. ■

6.5 Conclusion

Our analysis is developed along the values of two major parameters, namely, the network intensity and the horizon length, which are related to each other by conditions C.1 to C.4. When the network intensity and the horizon length meet condition C.1, the market is never saturated: given the weakness of network intensity, the horizon is not sufficiently far for the monopolist to cover the market and keep it covered for the remaining periods. When conditions C.2 and C.3 are fulfilled, namely, when network intensity is sufficiently strong, the monopolist is induced to saturate the market from the very beginning, and keep it covered for all subsequent periods. Notice that when the horizon length tends to infinity, the domain of network intensity values to which corresponds an interior solution, i.e. in which zero pricing is not practised, tends to a set of measure zero. Thus we conclude that, when the horizon is sufficiently remote, zero pricing in the starting period should be the rule.

The vector of prices identified in the optimal policies above corresponds to a solution in which the monopolist is not allowed to revise his strategy. However, we have checked that the solution would not be altered if this revision would be allowed. This property follows from the fact that consumers arrive in the market by successive waves and cannot delay consumption, which implies that they cannot arbitrate between periods when making their purchase decisions.

References

Ackere, A. and D. Reyniers (1995), Trade-ins and introductory offers in a monopoly, *RAND Journal of Economics*, 26 (1), 58-74

Bensaid, B. and J.-P. Lesne (1996), Dynamic Monopoly Pricing with Network Externalities, *International Journal of Industrial Organization*, 14, 837-855

Cabral, L. , D. Salant and G. Woroch (1999), Monopoly Pricing with Network Externalities, *International Journal of Industrial Organization*, 17, 199-214

Clarke, F. and Darrough, M. and J. Heineke (1982), Optimal pricing policy in the presence of experience effects, *The Journal of Business*, 55(4), 517-530

Dhebar, A. and S. Oren (1985), Dynamic nonlinear pricing in networks with interdependent demand, *Operations Research*, 34 (3), 384-394

Dhebar, A. and S. Oren (1985), Optimal dynamic pricing for expanding networks, *Marketing Science*, 4 (4), 336-351

Farrell, J. and G. Saloner (1985), Standardization, compatibility,and innovation, *Rand Journal of Economics*, 16, 70-83.

Katz, M. and C. Shapiro (1985), Network effect, competition, and compatibility, *American Economic Review*, 75(3), 424-440

Katz, M. and C. Shapiro (1994), System competition and network effects, *Journal of Economic Perspectives*, 8(2), 93-115

Rogerson, W. (1983), Reputation and product quality, *The Bell Journal of Economics*, 14 (2), 508-516

Rohlfs, J. (1974), A theory of interdependent demand for a communication service, *Bell Journal of Economics*, 5, 16-37