Technology, Industrial Structure, Financial Institutions and Economic Growth

by

Jian Tong

London School of Economics and Political Science

A Thesis Submitted for the Degree of Doctor of Philosophy

of the

University of London

September 2001

Abstract

This thesis studies the relationship between technology and industrial structure in the context of a growing market economy.

Chapters 2 and 5 develop some general equilibrium models which permit a study of the relationship between quality competition, market structure and growth. Both market structure and the rate of growth are determined endogenously as functions of underlying parameters describing the pattern of technology and tastes, and the institutional environment. It is argued that quality competition constitutes an economic mechanism of primary importance, which provides essential incentives for innovation at the industry level, while also contributing to aggregate technological progress by way of R&D spillover effects. A related theme of the thesis is that constraints on quality competition are detrimental to growth.

Chapter 3 presents a theoretical model which explains certain statistical regularities regarding cohort survival patterns, the persistence of firm turnover, and the appearance of shakeouts during an industry life cycle. By treating the market as comprising a number of strategically independent submarkets, this analysis separates the strategic interaction effects which occur at the submarket level, from the independence effects which operate across submarkets.

Chapter 4 studies competition between two cohorts of radically different but substitutable technologies. By analyzing the entry of new-technology-based firms, the exit of incumbents and subsequent quality competition, this chapter explores the impact of a radical innovation on market structure and on the turnover of firms. Two critical levels of the parameter which measures the efficiency of the new technology are identified: the first must be attained for 'creative destruction' to take place, while the second must be attained for this 'creative destruction' process to take a 'drastic' form which involves the complete replacement of currently active firms by a wave of new entrants.

Acknowledgments

My deepest thanks go to my supervisor, Professor John Sutton, whom I have been incredibly fortunate to meet at LSE and to whom I am intellectually indebted. This is the case not least because John's work on technology and market structure serves as a basis for much of what appears in this thesis and I have benefited enormously from his guidance throughout the preparation of this thesis, moreover, his intellectual integrity has set a worthy standard to which to aspire, not only during this work, but, I believe, for the many years to come in my research career.

I am also grateful to Chenggang Xu, who has provided me invaluable encouragement and advice at all levels through the years of my PhD at LSE with great kindness. Through the many helpful comments on my research, through the many discussions about economics in general, and especially through our joint work on Endogenous Financial Institutions, R&D Selection and Growth, Chenggang has influenced my understanding of economics with an emphasis on the roles of institutions in economic growth and development.

Many other members of the LSE have provided extremely helpful comments and suggestions on some part of the thesis, amongst whom I must mention Leonardo Felli, Nobu Kiyotaki and Max Steuer, who gave useful comments on Chapter 3; Alan Manning, Michele Piccione, Max Steuer, Mark Schankerman, Danny Quah and David Webb, whose comments and suggestions have helped to improve the contents or presentation of Chapter 2.

I would like to thank all my friends and colleagues in the Economics of Industry group and fellow PhD students at LSE. I have benefited particularly from discussions with Volker Nocke, Sonie Reiche, Silvia Sonderegger, Maria-Luisa Mancusi, Leopoldo Yanes, Thomas Buettner and Konstantinos Koufopoulos.

My sincere thanks also go to the seminar and conference participants at LSE, Warwick, Queen's University (Canada), Southampton, the Econometric Society European Meeting (2001) and the European Association of Researchers in Industrial Economics (EARIE, 2001) for stimulating discussions and helpful comments.

This work was mostly carried out at STICERD with its excellent research facilities; my special thanks go to Nora Lippincott, who proofread Chapters 2 and 3 for the STICERD: Economics of Industry Discussion Paper Series. Financial supports from the K. C. Wong Education Foundation and the Economics Department of LSE are gratefully acknowledged. My participation at conferences was made possible with the financial supports from the Review of Economic Studies and the Royal Economic Society.

Finally I am deeply grateful for the support, understanding and encouragement that I have received from my friends and family; particularly my wife, Yuanzhen, who has been through all the ups and downs with me, and my parents in law.

I dedicate this thesis to my Parents.

Contents

1 Introduction **8**

2 Quality Competition, Market Structure and Endogenous Growth **15**

2.1 Introduction 15

2.2 The Basic Model: Understanding 'Competition-Push' 17

2.2.1 The Setup 17

2.2.2 The Game of Quality Competition with Free Entry 21

2.3 Equilibrium 23

2.3.1 Market Structure and R&D in Each Industry 23

2.3.2 Balanced Growth Equilibrium 25

2.3.3 Discussion 27

2.4 The Interplay of 'Competition-push' and 'Rent-pull' 29

2.4.1 A Game of Quality Competition with Restricted Entry 29

2.4.2 Complementarity between 'Competition-Push' and 'Rent-Pull' 31

2.4.3 Institutional Barriers to Entry and Growth 33

2.4.4 Romer (1990) Revisited 36

2.5 Extending the Model: The Impact of Credit Constraint 38

2.5.1 Corner Solutions and Fragmented Market Structure 39

2.5.2 Credit Constrained Growth 40

2.5.3 A Poverty Trap 42

2.6 Conclusion 43

3 Submarkets, Shakeouts and Industry Life Cycle **44**

3.1 Introduction 44

3.2 The Model 48

3.2.1 'Independent' Submarkets 48

3.2.2 Industrial Growth via the Emergence of Submarkets 49

3.2.3 Modelling Competition within a Submarket 52

3.2.4 The Games of Quality Competition with and without Credit Constraints 53

3.3 Equilibrium 55

3.3.1 Market Structure in the Credit-un-constrained Phase 55

3.3.2 Market Structure in the Credit-constrained Phase 56

3.3.3 Miniature Shakeouts and Cohort Survival Pattern 57

3.3.4 Aggregate Shakeouts 59

3.4 Simulations of Industry Life-Cycle 59

3.4.1 Case 1: With Aggregate Shakeouts 59

3.4.2 Case 2: Without Aggregate Shakeouts 62

3.5 Concluding Remarks 65

4 Radical Innovation, Selection of Technology and Market Structure **66**

4.1 Introduction 66

4.2 Basic Model 68

4.2.1 Consumer Choice over Quality 68

4.2.2 A Reduced-form Profit Function 69

4.2.3 The Game 69

4.2.4 Equilibrium 71

4.2.5 Endogenizing the Number of Entrants 76

4.3 Anticipated Future Technological Breakthroughs 83

4.4 Conclusion 89

5 Quality Competition, Market Structure and Technological Progress with Cohort Replacements **91**

5.1 Introduction . . . 91

5.2 The Model . . . 93

5.2.1 The Basic Setup . . . 93

5.2.2 The Game of Quality Competition with Free Entry . . . 95

5.3 Equilibrium . . . 97

5.3.1 Market Structure and R&D in Each Industry . . . 97

5.3.2 A Balanced Growth Equilibrium . . . 99

5.4 Constraints on Quality Competition . . . 103

5.5 Discussion: The Immunity Period . . . 106

5.5.1 Market Structure and the Duration of Immunity . . . 106

5.5.2 Constraints on Quality Revisited . . . 109

5.6 Concluding Remarks . . . 111

6 Conclusions **112**

6.1 Three General Ideas . . . 112

6.2 Three Major Themes . . . 114

A For Chapter 2 **116**

B For Chapter 3 **120**

C For Chapter 4 **123**

D For Chapter 5 **126**

Chapter 1

Introduction

This thesis is concerned with the way in which imperfectly competitive markets guide the allocation of resources toward innovation, and with analyzing how the structure of each industry, over its life cycle, develops in tandem with the process of economy wide growth.

One central theme that runs through the thesis relates to the joint determination of market structure and innovative efforts by firms. This theme emerges clearly in the context of quality competition among firms. It is argued that quality competition is a primary economic mechanism that provides essential incentives for innovation at the industry level. In Chapters 2 and 5, it is further argued that quality competition also contributes to aggregate technological progress through R&D spillover effects. Chapters 2 and 3 in particular looks at how the roles of institutions, particularly financial institutions, shape the market environments in which quality competition occurs. The dynamic general equilibrium framework developed in the Chapters 2 and 5 allows us to draw implications for economic growth, the determinants of which are traced to underlying parameters describing the technology of the industries and institutional environment within which it works. A further theme explored by Chapters 3, 4 and 5 addresses the pattern of, and economic forces underpinning, industry life cycles.

In the existing R&D-based growth literature (Romer 1990, Grossman and Helpman 1992, Aghion and Howitt 1992), monopoly rents are treated as the driving force of innovation and the determinants of rents and R&D efforts are usually traced to some exogenous market structure. Therefore the chain of causation runs from market structure through the 'rent pull' to R&D and growth. This is in contrast with the recent literature in the Industrial Organization which aims to link market structure and R&D intensity. This literature has developed at some length the view that market structure and R&D intensity should be seen as jointly determined within an equilibrium system (Phillips (1971), Dasgupta and Stiglitz (1980), Sutton (1998, Chapter 1)), and it has identified a general economic mechanism, the escalation mechanism (Sutton, 1998), which jointly determines R&D intensity and market structure. The macro-economic implications of this mechanism have not, however, been explored in the literature to date.

In Chapter 2 I develop a growth model in which I explore this topic. The basic idea in Chapter 2 is that "quality competition" plus "R&D spillover effects" constitute a growth engine distinct from the "rent-pull growth mechanism" that has been the focus of attention in the recent 'endogenous growth' literature. We look at R&D activities which are driven by quality competition between rival firms, (as opposed to a 'patent race', which leads to the "winner-takes-all" scenario.) The term 'quality competition', in this thesis, means rivalry in vertical product differentiation or cost-reducing process innovation, (for the equivalence of these two forms of rivalry, see Sutton (1998), pp24-25 and Appendix 15.1.) 'Quality' is a special good, which has nonrival and partially excludible features. While firms compete in 'quality' the total benefits created by their R&D efforts go beyond their own industries, due to 'spillover effects', and these spillovers enhance the degree to which industrial R&D contributes to the aggregate technological progress.

Based on the above idea, Chapter 2 first develops a basic growth model, which extends the quality competition model developed by Sutton (1998), adopting the framework originally used by Romer (1990). This model demonstrates the working of a "competition-push" engine of growth. The stage-game setting embedded in this model then allows us to investigate how some underlying institutional factors can affect both market structure and the rate of growth, through affecting the market environment for quality competition. Against the benchmark set by the basic model, we parameterize, in one extension, the institutional barriers to entry; in another, the factors that restrict quality competition, e.g., credit constraints on R&D investment due to imperfection in capital markets. The conclusion is that the role of monopoly rents should not be over-stated, and that constraints on quality competition, e.g., credit constraints, are bad for growth.

Over the past decade the study of industry life-cycles in the growth-of-firms literature has devoted most of its attention to characterizing and explaining the 'shakeout' process (Klepper (1990), Jovanovic and MacDonald (1994), Klepper (1996, 1999)), i.e., the statistical regularity whereby the number of producers tends first to rise to a peak and later to fall to some lower level in many industries. Recent empirical findings by Horvath et al. (1997) and Klepper (1999) shed light on two other statistical regularities regarding industry life-cycles: (1) the 'turbulence' (firm turnover), i.e., a statistical regularity whereby the entry-exit process persists throughout an industry life-cycle and the observation that the gross entry rate and the gross exit rate are positively correlated across industries; and (2) the cohort survival pattern, i.e., all entry cohorts share a qualitatively similar survival pattern, which displays a significantly higher exit hazard rate at an early age than at subsequent ages.

When the 'turbulence' and the cohort survival pattern are examined jointly, a surprising pattern appears. As first noted by Horvath et al. (1997), despite the fluctuation in entry rates, the timing of exits for different cohorts of entrants is remarkably similar over time: the exit hazard rate peaks at a very early age of every cohort's life and drops dramatically to low levels at subsequent ages. In this sense, a typical industry life-cycle can be roughly re-described as follows: a miniature shakeout (i.e., an excess entry followed by dramatic early-age exit then followed by gradual subsequent exits) happens in the life of each cohort of entrants in a similar way throughout the whole industry life-cycle, and the 'shakeout' originally described by Klepper is an aggregation of these overlapping-cohort miniature shakeouts, together with the fact that the early-stage and late-stage cohorts have small numbers of entrants and the interim cohorts have large numbers of entrants.

Chapter 3 presents a theoretical model which explains these regularities and the links between them. Based on the insight described by Sutton (1997a 1997b, 1998), I begin the analysis by treating the market as comprising a number of strategically independent submarkets, so that I can separate the strategic interaction effect at the submarket level and the independence effects which operate across these independent submarkets at the aggregate level. This chapter then proposes an unusual but plausible extension of the game-theoretic quality competition literature, with an emphasis on independent submarkets, to explain these aforementioned statistical regularities in a typical industry life-cycle. The major scenario to be described is as follows: The uncertainty and informational problems surrounding the beginning of a new submarket tend to impose credit constraints upon fixed expenditures by producers. This in turn restricts the pressure of quality competition and leads to the viability of an excessive number of entrants in the early period of the life of the submarket. The miniature shakeout takes place later when the initial credit constraints are removed as the initial uncertainty is resolved and quality competition escalates, leading to the non-viability of a large fraction of existing producers in the submarket. Since an industry usually contains many independent submarkets which emerge at different times, the above scenario is repeated over time, so we observe continuing 'firm turnover' in the industry. When the emergence of submarkets slows down, the gross exit of producers will eventually dominate gross entry, and an "aggregate shakeout" takes place.

A frequently asked question about a new technology is whether it is commercializable? One factor that may deny a new technology commercial value is that the industry currently operates using an established substitute technology; the new technology may make the existing technology obsolete and replace it. Chapter 4 studies competition between two cohorts of radically different but substitutable technologies, and aims to gain some insights on how markets settle such a contest.

A contest between alternative technologies is sometimes represented as competition between firms advancing along alternative technological trajectories. One suitable measure of the prevalence of a technological trajectory is the market share captured by the firms who go along that trajectory. The technological contest is deeply influenced by the fact that the progress along a trajectory requires fixed outlays in R&D, which may be trajectory-specific and therefore involve sunk costs, in the sense that most of the value is not transferable across trajectories. As a result, switching trajectories will not involve a smooth transition for any incumbent firm which has committed sunk costs to the old trajectory; an opportunity may emerge which allows profitable entry by new entrants who have no vested interests in the old technology. We ask: under what circumstances does this happen, and how are market shares allocated between incumbents and new entrants at equilibrium?

While the incumbents employing the existent technology have a first mover advantage, potential new entrants have the advantage of jumping to the more efficient new technology. By conducting equilibrium analysis on the entry of new firms, the exit of incumbents and subsequent quality competition, I examine the impact of a radical innovation on market structure and on the turnover of firms. The analysis fully characterizes the conditions for three possible outcomes: either (1) the radical new technology is blocked by the existent technology, or (2) the two technologies coexist in the marketplace, or (3) the new technology replaces the existent technology. These conditions imply two bounds to the efficiency of the new technology, the first of which must be reached for 'creative destruction' to take place, and the second of which suffices for its effects to be 'drastic'.

Some of the insight gained in the partial equilibrium analysis of Chapter 5 carries on to the general equilibrium analysis of Chapter 5, where I study the role of quality competition in endogenous technological progress with cohort replacement. In contrast to the Schumpeterian 'creative destruction' type models in the R&D-based growth literature (Grossman and Helpman 1992, Aghion and Howitt 1992), which identify aggregate technological progress as a sequence of monopolizable achievements along a 'quality ladder', the present analysis relates the technological cohort replacement and the associated firm turnover in an industry with some radical change in R&D technology, and the new round of quality competition induced by this change. We model the radical change in R&D technology as emerging 'exogenously', i.e., outside the industry to which it is applied, but this process is still endogenous from an economy-wide point of view, while the pace of technological change depends on the aggregate spillover effects from the R&D efforts in all industries.

One insight arising in the Schumpeterian 'creative destruction' models is that the prospect of future creative destruction has a negative impact on the effort put into the current generation of technology. This remains true in the present study, though an additional insight emerges: a cohort of established technologies can enjoy a period of immunity when is creative-destruction-free. This scenario emerges as a result of a finding in Chapter 4, that stronger (but not 'sufficiently stronger') new technologies may be blocked by established technologies; therefore before 'sufficiently stronger' new technologies emerge, (and this takes time), established technologies are immune to 'creative-destruction'.

Creative destruction and firm turnover, in the setting of Chapter 5, emerge as particular outcomes of the more fundamental economic force of quality competition in association with radical technological changes. This strengthens the central claim of Chapter 2, that quality competition provides an essential incentive for innovation at industry level and, through R&D spillover effects, it also underlies aggregate technical progress. The present analysis shows that this mechanism is very robust, in that it does not rely on whether the aggregate technological progress proceeds through variety expansions a la Romer (1990) (as assumed in Chapter 2) or through cohort replacements a la Aghion-Howitt (1992). This also echoes another claim made in Chapter 2: that constraints on quality consumption reduce R&D intensity and lower the rate of growth.

In a short concluding chapter, an overall view of technology and growth is set out, which integrates some of the ideas of Chapters 2-5, and places them in a broader context.

Chapter 2

Quality Competition, Market Structure and Endogenous Growth

2.1 Introduction

One important theme of the recent endogenous growth literature relates to the connection between market structure (concentration) and the rate of growth (See for example Aghion and Howitt (1998, Chapter 7)). Discussions of this connection have been conducted in settings where concentration is taken as an exogenous variable, so that the "chain of causation" runs from concentration to growth. This mirrors an old debate in the related I.O. literature which aimed to relate concentration to the rate of innovation (or R&D intensity). That early debate was marked by the difference of views regarding the "direction of causality", some writers seeing the direction as running from concentration to R&D intensity, others seeing R&D intensity as affecting concentration. This debate had been resolved in the I.O. literature by 1980, by which time it being widely accepted that concentration and R&D intensity should be seen as jointly determined within an equilibrium system. (Phillips (1971), Dasgupta and Stiglitz (1980), Sutton (1998, Chapter 1)).

In this chapter I re-examine the more complex relationship between concentration and the rate of growth from the same standpoint. In the model developed in the chapter, we trace the determinants of market concentration and growth to primitive parameters describing the technologies (notably, a parameter that captures the "effectiveness of R&D"). This framework allows a re-explanation of the connection between such parameters, and the level of concentration and the rate of growth.

At the level of this analysis lie two economic mechanisms:

- a 'rent pull' mechanism of a familiar kind, reminiscent of the "Schumpeterian" argument, by which monopolist rent offers an incentive to innovate.
- a 'competition effect', where the push of competition by rivals creates an incentive to innovate.

It is the interplay of the two mechanisms which is central to what follows.

The novel feature of this analysis, relating to the existing models in the R&D-based growth literature, is that it builds upon the 'quality competition' models[1] of the related I.O. literature (See for example Sutton (1991, 1998), Dasgupta and Stiglitz (1980, section II)). In contrast to "patent race"[2], or "quality ladder"[3] models, these models allow equilibrium outcomes in which several rival firms simultaneously offer substitutable 'quality goods' or compete in a homogeneous good market via cost-reducing process innovations[4]. Thus we can have a range of market structures

[1]In this paper the term 'quality competition' is used in a broad sense, as to include both vertical product differentiation and cost-reducing process innovation, emphasizing the contrast between 'innovation as a device of competition' and pure price competition.

[2]See for example Dasgupta and Stiglitz (1980, section IV). For a comprehensive survey of the related literature, see Reinganum (1989).

[3]See for example Grossman and Helpman (1991) and Aghion and Howitt (1992).

[4]The possibility of coexistence of directly competing firms is neither ruled out by the possibility of "natural monopoly" nor ruled out by patent protection. Related to the former point, Dasgupta and Stiglitz (1980) and Sutton (1991, 1998) show that, with 'quality choice', an equilibrium outcome supporting multiple firms is feasible. Related to the latter point, Levin et al. (1987), provided empirical evidence, suggesting that the effectiveness of patents to prevent competition is, in general, very limited.

that run from monopoly through to a 'fragmented' (competitive) structure.

Once we are exposed to how some underlying parameters affect both market structure and the rate of growth through affecting the game environment, we can distinguish the natural tendency of market competition and the outcomes marked by the intervention of some institutional variables. Examples of intervention of such kind are that "monopolist rent" can be affected by the underlying parameter regarding institutional barriers to entry[5], that R&D intensity can be dependent on the parameter which describes the imperfection of the capital market e.g., credit constraint on R&D investment.

2.2 The Basic Model: Understanding 'Competition-Push'

2.2.1 The Setup

Two sectors

The model represents an economy which consists of two sectors: a final good sector and an intermediate good sector with multiple industries producing different varieties. For convenience, in what follows, the term 'industry' is reserved for an industry in the intermediate good sector. The final good can be used in consumption and can also be used as input to produce intermediate goods. The variety of intermediate goods can grow as a result of aggregate technological progress. In the final good sector there are constant returns to scale and perfect competition. In each industry firms compete in 'quality' as well as in price. Due to the non-rival and excludable features of 'quality' and the consequent non-convexity, the firms are not price takers and the market structure is endogenous.

[5]By 'institutional barriers to entry', we mean measures which affect the effectiveness of patents as means to exclude competition, or the lack or ineffectiveness of antitrust law and policy.

Industry Level Technology, Non-rival and Excludable 'qualities'

The technology in an industry can be captured by an endogenous cost structure. The variable cost is all capital cost of final good and has constant returns to scale. The 'quality' represents either consumers' 'perceived quality' measure by their willingness to pay, or productivity level. Mathematically these are isomorphous, and both can be measured by the amount of effective units produced by unit input of final good. So in the rest of chapter, quality and productivity are regarded as interchangeable terms. Quality can be improved by R&D which requires a fixed cost. The R&D cost is all the cost of hiring R&D personnel, and is a convex increasing function of quality index. An intermediate good producer can trade off between R&D cost and variable cost. There are overall increasing returns to scale due to the fact that the quality is non-rival, and can be applied to all units of output with zero marginal cost. The quality is excludable through exclusion devices such as patent or secrecy, so it is only accessible by the firm which invests in it.

Aggregate Level Technology and Externalities

The emergence of new varieties of intermediated goods and therefore new markets in each period, $\dot{A}$, is a function of innovative efforts made by firms in industries, measured by L_2, and the spillover from knowledge embedded in existing intermediated goods, A, i.e., $\dot{A} = \delta L_2 A$. The spillover from A to $\dot{A}$ has been well recognized by the R&D-based growth literature. The R&D spillover effects[6] from quality creation by industries to aggregate growth in knowledge need a few more comments. In this model, the R&D personnel (scientists and engineers) L_2 are hired by firms in industries to create qualities. In doing so, they also contribute to generic progress in knowledge and generate prototype designs of new varieties of intermediated goods as by-products. So this kind of spillover is flowing from one type of innovation,

[6]The empirical evidence on the R&D spillover effects, as surveyed by Nadiri (1993), suggests that they are sizable, and may be even larger than the effects of own R&D at industry level.

i.e., quality improvement, towards another type of innovation, i.e., new prototype creation. The emergence of the prototype design of a new variety predicts a new emerging market which can soon be entered and competed in.

3-stage Game in Each Industry

Whenever a new industry emerges, a three-stage game is played in the emerging industry. In the first stage, potential entrants simultaneously choose whether to enter the new industry. In the second stage, those firms which have entered the market choose their 'quality' levels (u_i), which will be achieved by R&D and incurring costs of hiring R&D personnel. In the third stage, firms with their achieved 'quality' level compete in quantity in a Cournot manner[7].

Consumption

There are constant L identical consumers who live infinitely and each has one unit of labor per period. The labor can be supplied either as simple labor used in final good sector, or innovative labor used in industries to do R&D. Each consumer has a discounted, constant elasticity preference:

$$\int_0^\infty \frac{c^{1-\varepsilon}}{1-\varepsilon} e^{-\rho t} dt \text{ for } \varepsilon > 0,$$

where c is consumption of final good, ρ is time preference and ε is the constant relative risk averse coefficient. The Euler equation for intertemporal optimization for given interest rate r is that $\frac{\dot{c}}{c} = \frac{r-\rho}{\varepsilon}$.

[7] The form of price competition throughout this thesis is Cournot quantity competition. Tougher price competition as captured by a Bertrand price competition, or softer price competition which may involve some form of collusion are beyond the scope of this thesis.

Final Good Sector

The final good can be treated as the numeraire. Its production function is

$$Y = L_1^{1-\gamma} \int_0^A \left(\sum_{j=1}^{N_i} u_{ij} \tilde{x}_{ij} \right)^{\gamma} di,$$

where Y is the total output of final good, L_1 is labor input, $\tilde{x}_{ij}$ is the input of the intermediate good provided by firm j in industry i, whose corresponding firm-specific 'quality index' is u_{ij}, N_i is the number of firms in industry i, and A is the number of intermediate good industries, $0 < \gamma < 1$.

To simplify the formulation, we define the quality-adjusted quantity as:

$$x_{ij} \equiv u_{ij} \tilde{x}_{ij},$$

and define the 'quality-adjusted price' of each unit of quality-adjusted quantity as:

$$p_{ij} \equiv \frac{\tilde{p}_{ij}}{u_{ij}}.$$

Since every unit of quality-adjusted quantity in industry i is homogeneous, the following 'one price principle' (or non-arbitrage condition) must apply:

$$p_{ij} = p_i \text{ for } j = 1, 2, \cdots, N_i; i = 1, 2, \cdots, A.$$

After the transformation, the original production function can be replaced by:

$$Y = L_1^{1-\gamma} \int_0^A x_i^{\gamma} di, \tag{2.1}$$

where $x_i \equiv \sum_{j=1}^{N_i} x_{ij} = \sum_{j=1}^{N_i} u_{ij} \tilde{x}_{ij}$. Now u_{ij} can be re-interpreted as a 'productivity index' measured in units of quality-adjusted quantity. Mathematically this new formulation is isomorphous to the one which may describe the setting where in each industry firms produce a homogeneous good with firm-specific productivity.

There is perfect competition in the final good market due to constant returns to scale, so equating the marginal product of labor and the wage rate gives us the inverse labor demand function

$$w = \frac{(1-\gamma) Y}{L_1}. \tag{2.2}$$

Similarly by equating the marginal product of input i and its price, we have the inverse demand function for intermediate good i,

$$p_i = \gamma L_1^{1-\gamma} x_i^{\gamma-1}. \tag{2.3}$$

2.2.2 The Game of Quality Competition with Free Entry

In each emerging industry a 3-stage game is played. The solution concept adopted in this chapter is subgame perfect Nash equilibrium, and it can be solved by backward induction. The third stage subgame is a Cournot game of quantity competition, a Cournot-Nash equilibrium exists and is unique for a given set of quality levels of all firms and the demand condition of the economy. The subgame Nash equilibrium determines for each firm a reduced-form profit function as follows

$$\pi_i\left(u_i \mid u_{-i}\right) = \underbrace{S}_{\text{market size}} \underbrace{\frac{1}{1-\gamma}\left(1 - \frac{N-1+\gamma}{\sum_{j=1}^{N}\frac{u_i}{u_j}}\right)}_{\text{market share}} \underbrace{\left(1 - \frac{N-1+\gamma}{\sum_{j=1}^{N}\frac{u_i}{u_j}}\right)}_{\text{mark-up ratio}}, \tag{*}$$

where $S \equiv \gamma^{\frac{1}{1-\gamma}} L_1 \left(\frac{N-1+\gamma}{\sum_{j=1}^{N}\frac{1}{u_j}}\right)^{\frac{\gamma}{1-\gamma}}$, u_i is the quality level of firm of i, u_{-i} is a $(N-1)$-tuple of quality levels of other firm except firm i, for $i = 1, 2, \cdots, N$; N is the number of firms in the industry and L_1 is the amount of labor employed in the final sector.

Interested readers can find the derivation of this reduced-form profit function[8] in Appendix A. Its key feature is that a firm's profit increases with its relative quality level against its rivals', i.e., $\sum_{j=1}^{N}\frac{u_i}{u_j}$. This function tells how the strategic environment responds to vertical differentiations of firms. It captures the Darwinian selection pressure embedded in the environment constituted by customers and rivals. This environment provides firms with incentives to outperform their rivals in R&D and quality. When they vie in quality, each of them has to bear the burden of R&D

[8] This formulation is extended from the one developed in Appendix 15.1 of Sutton (1998). It allows a general treatment of $\gamma \in [0,1)$ while the Sutton (1998) formulation only deals with a special case, equivalent to: $\gamma = 0$.

cost, which is a function of quality level as shown below:

$$F\left(u_i; w, A\right) = w\frac{\mu}{A}u_i^{\beta}, \tag{**}$$

where $\frac{\mu}{A}u_i^{\beta}$ ($\beta > 1$) is the input of R&D personnel to generate quality level u_i, w is the market wage rate and μ is a constant. The R&D cost is therefore all labor cost. It is worth emphasizing that the increase in A, i.e., the aggregate level stock of knowledge, can shift the fixed (R&D) cost curve downward as shown in Figure 2.1. The intuition is that knowledge spillover from aggregate technical progress enhances the efficiency of R&D by each new firm. This positive trend may be offset by the negative trend caused by an increase in wage rate, also shown in Figure 2.1.

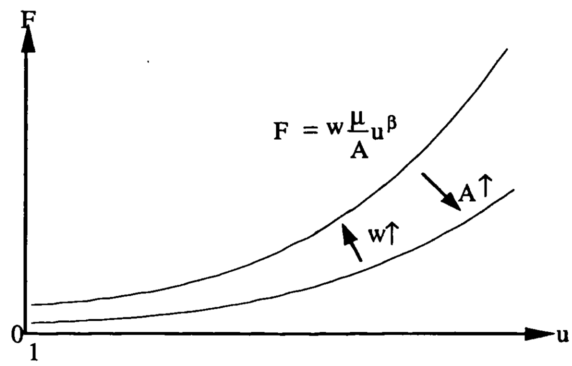


Figure 2.1: Fixed (R&D) cost function

Therefore each firm's objective in the second stage subgame is to maximize its net profit, in the present value sense, by choosing its own quality level given others' quality levels,

$$\max_{u_i}\left\{\frac{S}{r\left(1-\gamma\right)}\left(1-\frac{N-1+\gamma}{\sum_{j=1}^{N}\frac{u_i}{u_j}}\right)^2 - \frac{w\mu}{A}u_i^{\beta}\right\}, \tag{2.4}$$

where r is the interest rate. If we take the reduced form profit function and the R&D cost function as given, then the game is in essence a game of quality competition with free entry.

2.3 Equilibrium

2.3.1 Market Structure and R&D in Each Industry

Equilibrium analysis is focused on a symmetric subgame perfect Nash equilibrium, where each firm equates the marginal benefit of increasing quality and the marginal R&D cost. The implied best-reply function for each firm in a symmetric outcome is that

$$F = \frac{w\mu}{A} u^{\beta} = \frac{S}{r}\frac{1}{\beta}\left(\frac{\gamma}{N^3} + 2\left(N-1\right)\frac{N-1+\gamma}{N^3}\right), \tag{2.5}$$

where $S = \gamma^{\frac{1}{1-\gamma}} L_1 \left(\frac{N-1+\gamma}{N} u\right)^{\frac{\gamma}{1-\gamma}}$ is the industry's total revenue per period.

It is assumed that any intermediate good industry has free-entry. This results in the following zero-profit condition, which equates the R&D cost of the firm $\frac{w\mu}{A}u^{\beta}$ and the present value of profit flows $\frac{S}{r}\frac{1-\gamma}{N^2}$,

$$F = \frac{w\mu}{A} u^{\beta} = \frac{S}{r}\frac{1-\gamma}{N^2}. \tag{2.6}$$

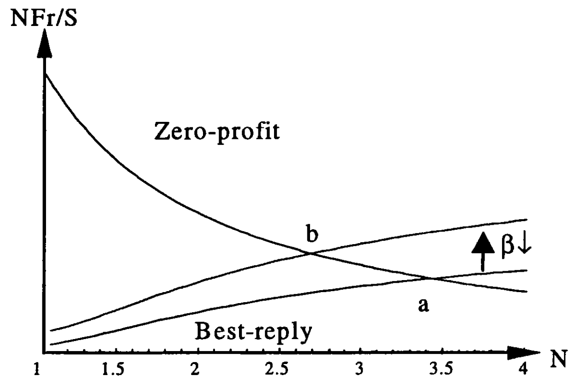


Figure 2.2: Market structure and industry R&D intensity

Figure 2.2 presents the above best-reply condition and zero-profit condition as two curves. The vertical axis is the industry R&D intensity and the horizontal axis is the number of firms in each industry. The graph shows that the Best-reply curve is upward sloping and the zero profit curve is downward slopping and they cross once so that it implies that a unique class of symmetric Nash equilibria exist, where R&D level and market structure are simultaneously determined. This suggests that there is a two-way causality between market concentration and R&D intensity. This result defies the view which portrays a simple causal relationship between the two variables. The graph also shows that when the parameter β decreases, it shifts the Best-reply curve upward therefore increases both market concentration $1/N$ and industry R&D intensity[9] (from **a** to **b**). The intuition for this result is as follows. A lower β means a higher degree of overall increasing returns to scale at firm level[10]. The market equilibrium finds a way to exploit this by having more concentrated market structure in each intermediate good industry. Sutton (1998) dubs this kind of economic mechanism as an 'escalation mechanism'. These findings are formally stated in the following

Proposition 1 [11] *The market structure and the R&D intensity of each intermediate*

[9] For a casual observer who could not see the underlying shift in β, there would be a temptation for him to claim that concentration causes high R&D intensity. But for a careful econometrician who would manage to control β, such a postulated positive causal relation could be dismissed. This scenario briefly reflects the evidence in the empirical literature. According to a comprehensive survey by Cohen and Levin (1989) on the empirical evidence, most studies on the relationship between market concentration and R&D intensity report a positive relation, though some find a negative one, and others argue for a non-monotonic relationship, e.g., Scherer (1967) suggested a non-linear, 'inverted U' relationship between the two variables. Furthermore, Scott (1984) and Levin et al. (1985) provide evidence that when some industry-specific effects are controlled for, concentration is found to contribute very little to the explanation of the variance in R&D intensity, and its effect becomes statistically less significant or even insignificant.

[10] The overall production function of each intermediate good producer is $x_i = u_i y_i = \left(\frac{A}{\mu} L_{2i}\right)^{\frac{1}{\beta}} y_i$, which is homogeneous of degree $1 + \frac{1}{\beta}$ in (L_{2i}, y_i). To see this in detail, note that L_{2i} is the input of innovative labor to achieve u_i, which is the productivity level, $L_{2i} = \frac{\mu}{A} u_i^{\beta} \Leftrightarrow u_i = \left(\frac{A}{\mu} L_{2i}\right)^{\frac{1}{\beta}}$, and y_i is the final good used as input to produce the intermediate good.

[11] The results here are consistent with those found in Dasgupta and Stiglitz (1980) and Sutton

good industry are simultaneously determined in the symmetric equilibrium. The equilibrium number of firms in each industry is $N_f = n(\beta, \gamma) \equiv \frac{n_0 + \sqrt{n_0^2 - 2(2-\gamma)}}{2}$*, where* $n_0 \equiv \frac{(2+\beta)(1-\gamma)}{2} + 1$*. Furthermore* N_f *is increasing in* β*.*

Proof. See Appendix A.

2.3.2 Balanced Growth Equilibrium

Following the conventional strategy to characterize a dynamic general equilibrium model, the balanced growth equilibrium is considered, where $\frac{\dot{Y}}{Y} = \frac{\dot{c}}{c} = \frac{\dot{A}}{A} = \frac{\dot{w}}{w} = g$ is satisfied and L_1, L_2 are constant over time. Immediate implications of these include that $g = \frac{\dot{c}}{c} = \frac{r-\rho}{\varepsilon}$, that $g = \frac{\dot{A}}{A} = \delta L_2$ and that $\frac{Y}{A}$ and $\frac{w}{A}$ are constant over time. A further implication is that the demand condition: $p = \gamma L_1^{1-\gamma} x^{\gamma-1}$ and the R&D cost function: $F(u; w, A) = w\frac{\mu}{A}u^{\beta}$, therefore the 3-stage game played in each industry are identical over time. Then it is reasonable to focus attention only on symmetric outcomes.

In the symmetric balanced growth equilibrium, the total revenue in each industry per period is identically

$$S = px = \frac{\gamma Y}{A}, \tag{2.7}$$

where γ is the cost share of all intermediate goods in the final good sector. Therefore γY is the total revenue of all industries per period and $\frac{\gamma Y}{A}$ is the total revenue of each industry. Given $\frac{Y}{A}$ is constant, S should also be constant.

The symmetric balanced growth equilibrium also implies that u^{β} is identical for all industries. Then the demand for innovative labor L_2 can be calculated as follows,

$$L_2 = \dot{A} N_f \frac{\mu}{A} u^{\beta}, \tag{2.8}$$

(1991). The setup of the game in Dasgupta and Stiglitz (1980) is a simultaneous game, unlike the sequential game setting in this chapter. Sutton (1991) differs slightly in that γ is fixed to 0. With a general γ in the current setting, it is found that N_f is decreasing in γ.

where $\dot{A}$ is number of emerging industries per period, N_f is the number of firms in each industry and $\frac{\mu}{A}u^\beta$ is the demand for innovative labor by each new firm. This condition combined with the aggregate growth equation

$$\dot{A} = \delta L_2 A. \tag{2.9}$$

pins down the quality level in each industry as follows

$$u^\beta = \frac{1}{\mu N_f \delta}. \tag{2.10}$$

To close the model, it remains to add the labor market clearing condition $L = L_1 + L_2$. Given the zero-profit condition $\frac{w\mu}{A}u^\beta = \frac{S}{r}\frac{1-\gamma}{N^2}$, a few manipulations with eq. (2.2), (2.10) and (2.7) to eliminate w, u^β, S put in place the final brick of the model as

$$L_1 = \frac{rN_f}{\gamma\delta}. \tag{2.11}$$

The system is finally boiled down to four equations: $\begin{cases} L = L_1 + L_2 \\ L_1 = \frac{rN_f}{\gamma\delta} \\ L_2 = \frac{g}{\delta} \\ g = \frac{r-\rho}{\varepsilon} \end{cases}$, which imply[12] the balanced growth rate:

$$g = \frac{\gamma\delta L - \rho N_f}{\gamma + N_f\varepsilon}. \tag{2.12}$$

The expression of the growth rate has the following features captured by

Proposition 2 *The growth rate g is decreasing in the equilibrium number of firms in each industry, N_f, and it is decreasing in β.*

Proof. $\frac{\partial g}{\partial N_f} = -\frac{\gamma(\rho+\varepsilon\delta L)}{(\gamma+N_f\varepsilon)^2} < 0$. Given $\frac{\partial N_f}{\partial \beta} > 0$, $\frac{\partial g}{\partial \beta} = \frac{\partial g}{\partial N_f}\frac{\partial N_f}{\partial \beta} < 0$. ■

The results should be interpreted with Diagram 1 in mind, knowing the causal relationship between variables. We see in eq. (2.12) that growth rate is linked with

[12] Other results are $L_1 = N_f\frac{\varepsilon\delta L+\rho}{\delta(\alpha+N_f\varepsilon)}$, $r = \alpha\frac{\varepsilon\delta L+\rho}{\alpha+N_f\varepsilon}$, $L_2 = \frac{\alpha\delta L-\rho N_f}{\delta(\alpha+N_f\varepsilon)}$.

market structure (fragmentation, measured by N_f, is negatively correlated with g). But it would be a misunderstanding if one inferred that growth could be enhanced by making industrial structures more concentrated, because the market structure is not an exogenous variable which can be controlled in the case of free entry. The deep reason for this correlation to exist is that market structure (n) and R&D effort (F) are jointly determined in equilibrium, both dependent on the industry level technological parameter β. Through the causal linkage between R&D and growth, β also affects g. Figure 2.3 shows that because β simultaneously has a positive effect on n and a negative effect on g, n and g must be negatively correlated, though not causally.

The demonstrated feasibility of positive rate of R&D-based growth in the case of free entry is very informative. It shows that even without positive net profit (rent), under certain micro-level and macro-level technological conditions, quality competition still provides a 'competition-push' engine of growth which works.

$$\beta \rightarrow \left.\begin{matrix}(n,F)\\ \delta\end{matrix}\right\} \rightarrow g$$

Diagram 1. The causal relationship between β, n and g

2.3.3 Discussion

Free Competition without Price-taking

When quality competition prevails and non-convexity results from the non-rival feature of 'quality', the endogenous R&D cost will form a natural barrier to entry. Consequently, though without any regulation on entry, profitability will dictate free but limited entry. In this environment price-taking behavior cannot be sustained, but the market should nevertheless be regarded very competitive due to

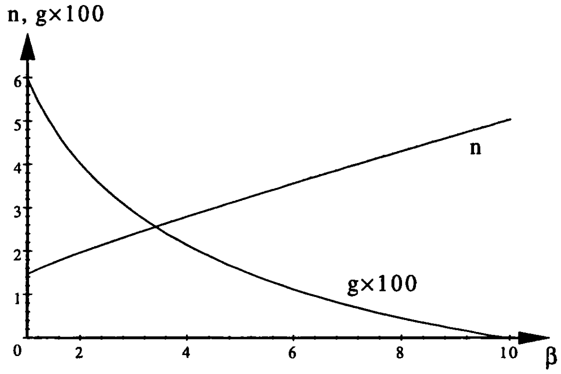


Figure 2.3: β jointly determines n and g

free entry. In this sense, a competitive market without perfect (price) competition is conceivable and non-convexity is compatible with (quality-)competitive markets. By characterizing the quality-based increasing returns to scale and the natural rule of market competition, this quality competition model stands in between the two extreme cases most studied in the economic literature, i.e., monopoly and perfect competition, presenting a picture of natural oligopolies. It suggests that the essence of a free market system does not necessarily lie in price-taking behavior, rather the 'competition-push' provides sufficient Darwinian pressure which contributes to the evolution of technologies.

Variety Expansion or Cohort Replacement

The above analysis of quality competition-based growth is presented in the context of a variety-expansion economy. It is assumed that the emergence of new grounds of quality competition relies on the expansion of varieties in a Romer (1990) manner. The major findings from the analysis, however, are more general than this particular setting. In Chapter 5, which presents a different version of the quality competition-based growth model, the same key results have been achieved from the context of a cohort-replacement economy. There the emergence of new grounds of quality competition is assumed to follow the replacement of earlier cohorts of technologies by successive ones in a manner akin to the Aghion-Howitt (1992) specification of 'creative destruction'. The departure from the Aghion-Howitt (1992) model is that the model in Chapter 5 allows multiple firms of the same cohort to compete in quality within each single market.

Given that variety expansion and cohort replacement are the two major modes of fundamental innovations, showing that quality competition-pushed innovation can complement either variety-expansion or cohort-replacement in generating endogenous growth assures the robustness of the analysis.

2.4 The Interplay of 'Competition-push' and 'Rent-pull'

2.4.1 A Game of Quality Competition with Restricted Entry

Having seen how quality competition jointly determines market structure and R&D intensity in the setting of the basic model, we proceed in the next step to analyze the interplay of the 'competition-push' and the 'rent-pull' familiar in the existing of R&D-based growth models. While free entry dries rents up, institutional barriers to entry tend to generate positive rents. In this section it is assumed that institutional barriers to entry impose a constraint on the number of potential entrants. If the number of potential entrants is smaller than N_f i.e., the number of firms determined in the free-entry equilibrium, then this constraint is binding and the zero-profit condition (2.6) does not hold. The actual number of firms in each industry, N, satisfies the condition $1 \leq N \leq N_f$. This modifies the game played in each industry to a 'game of quality competition with restricted entry'. The equilibrium of the changed game then is characterized by the restricted N, which has deviated from N_f, and the best reply condition (2.5).

Rents vs. R&D Costs

Figure 2.4 shows that when N deviates from N_f $(N < N_f)$ the new equilibrium is no longer at point **a**, which is the cross of the zero-profit curve and Best-reply curve. Instead, it shifts to point **b** (we will discuss the reason later) and the extra market power generates positive net profits, i.e., rents, residual after R&D costs are covered.

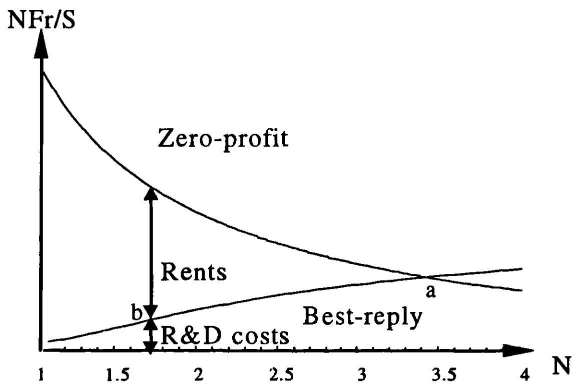


Figure 2.4: Market structure and spilt of profit

The present value of the rents therefore is a function of β, γ, N, S and r as follows,

$$V(\beta, \gamma, N, S, r) = \frac{S}{r}\frac{1-\gamma}{N^2} - F, \tag{2.13}$$

where $\frac{S}{r}\frac{1-\gamma}{N^2}$ is the present value of future profits flows and
$F = \frac{S}{r}\frac{1}{\beta}\left(\frac{\gamma}{N^3} + 2(N-1)\frac{N-1+\gamma}{N^3}\right)$ is the R&D cost by eq. (2.5).

To simplify discussion, rent and R&D cost can be denominated by the present value of industry total revenue S/r. Define $\Pi(\beta, \gamma, N) \equiv \frac{V(\beta,\gamma,N,S,r)}{S/r} = \frac{(1-\gamma)\beta N - \gamma - 2(N-1)(N-1+\gamma)}{\beta N^3}$ and $F_0(\beta, \gamma, N) \equiv \frac{F}{S/r} = \frac{1}{\beta}\left(\frac{\gamma}{N^3} + 2(N-1)\frac{N-1+\gamma}{N^3}\right)$. It is easy to see that $\Pi(\beta, \gamma, N) > 0$ for $N < N_f$ and that $\Pi(\beta, \gamma, N) = 0$ for $N = N_f$.

Market Structure, Profitability and Split of Profits

Figure 2.4 also shows that when $1 < N \leq N_f$, increasing extra market power (decreasing N) will increase the share of rents and decrease the share of R&D costs. The intuition is that when the restriction on free entry is tightened the split of profits will shift in favor of rents against R&D costs. This result is formally stated by the following

Proposition 3 *If* $1 < N \leq N_f$ *then* $\frac{\partial(N\Pi(\beta,\gamma,N)+NF_0)}{\partial N} < 0$, $\frac{\partial(N\Pi(\beta,\gamma,N))}{\partial N} < 0$ *and* $\frac{\partial(NF_0)}{\partial N} > 0$.

Proof. See appendix A.

2.4.2 Complementarity between 'Competition-Push' and 'Rent-Pull'

The aggregate growth equation (2.9) used in the basic model is too simple to capture the familiar rent-pull engine of growth. We extend it as follows:

$$\dot{A} = \delta \min\left(\frac{L_2}{\theta}, \frac{L_3}{1-\theta}\right) A, \text{ for } 0 \leq \theta \leq 1, \tag{2.14}$$

where L_2 is the hired R&D personnel, L_3 is the labor of entrepreneurs, involved in searching, creating and preempting new grounds of quality competition, providing leadership and gaining extra market power. In the context of a variety expansion economy, entrepreneurial efforts may be needed for the creation of new industries. The entrepreneurs are the founders of industrial firms and their efforts are compensated by the residual incomes of the firms, i.e., the rents, instead of wages. The parameter θ determines to what extent L_2 or L_3 is a scarce resource for the purpose of economic growth. For example, if $\theta = 1$, then the equation is reduced to $\dot{A} = \delta L_2 A$, which implies that L_3 is not scarce at all, as in our basic model. On the other hand, the familiar models of rent-pull usually assume that $\theta = 0$, which implies that L_2 is not scarce at all, and $\dot{A} = \delta L_3 A$. In between these two extreme cases, $\dot{A} = \delta\frac{L_2}{\theta}A$ applies if $\frac{L_2}{\theta} < \frac{L_3}{1-\theta}$ holds in equilibrium and $\dot{A} = \delta\frac{L_3}{1-\theta}A$ applies if $\frac{L_2}{\theta} \geq \frac{L_3}{1-\theta}$ holds in equilibrium. Given this specific functional form, it is immediately clear that for a given value of θ, the growth-maximizing ratio between L_2 and L_3 is $\frac{L_2}{L_3} = \frac{\theta}{1-\theta}$.

Due to perfect mobility of labor between L_2 and L_3, each entrepreneur's income should be equal to the income of a hired scientist or engineer in equilibrium. Since L_2's total income equals to total R&D costs and L_3's total income equals to total rents, the 'entrepreneurial arbitrage condition' requires that the ratio between L_2 and L_3 should be equal to the R&D costs-rents ratio, i.e.,

$$\frac{L_2}{L_3} = \frac{F_0}{\Pi}. \tag{2.15}$$

Therefore there exists a growth-maximizing split of profit such that $\frac{F_0}{\Pi} = \frac{\theta}{1-\theta}$ or equivalently, $\frac{NF_0}{N(F_0+\Pi)} = \theta$, which imply[13] $N = N^* \equiv \frac{n_1+\sqrt{n_1^2-2(2-\gamma)}}{2}$, where $n_1 \equiv \frac{(\beta\theta+2)(1-\gamma)}{2} + 1$.

Figure 2.5 shows that the growth-maximizing split of gross profit: $\frac{NF_0}{N(F_0+\Pi)} = \theta$ as the dashed curve for a given value of θ. The Best-reply curve between **a** and **b** is the actual split of profit. It can be seen in the graph that point **e** represents the unique growth-maximizing market structure. For a point between **a** and **e**, such as **d**, the share of rents is too low for the purpose of growth. On the other hand, for a point between **b** and **e**, such as **c**, the share of rents is too high and the market is too concentrated. So two points can be made from the graph: (1) both the competition-push engine and the rent-pull engine may be needed for the purpose of growth. (2) more extra market concentration does not always mean a higher rate of growth. These points are made even clearer when the balanced growth rate is solved explicitly.

[13]They imply the following quadratic equation $N^2 - \left(\frac{(\beta\theta+2)(1-\gamma)}{2} + 1\right)N + \frac{2-\gamma}{2} = 0$.

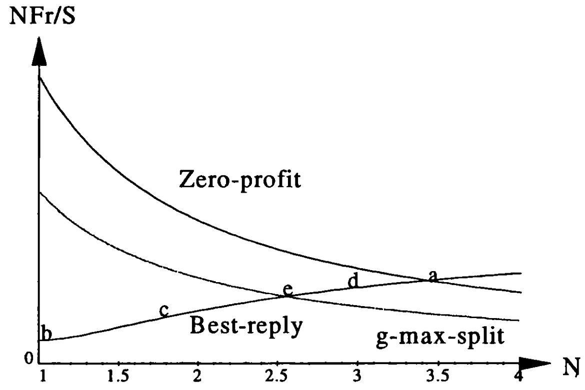


Figure 2.5: Growth-maximizing split of profit

2.4.3 Institutional Barriers to Entry and Growth

To obtain the closed-form growth function within this extended framework, one can refer to the derivations within the basic model, but should note that the zero profit condition (eq. (2.6)) is replaced by the entrepreneurial arbitrage condition (eq. (2.15)), that eq. (2.14) replaces eq. (2.9) and that the new labor market clearing condition is changed to $L = L_1 + L_2 + L_3$.

Then the implied balanced growth rate is

$$g = \frac{\frac{\gamma\delta N}{1-\gamma}\min\left(\frac{F_0}{\theta}, \frac{\Pi}{1-\theta}\right) L - \rho}{\varepsilon + \frac{\gamma}{N}}, \tag{2.16}$$

which, for a given θ, is a single-peaked function of N. The location of the peak is at the previously solved growth-maximizing market structure N^*, which depends crucially on parameter θ. The results are summarized by the following

Lemma 4 *N^* is the growth-maximizing market structure. N^* is increasing in θ. When $\theta = 1$, $N^* = N_f$. Furthermore when $N < N^*$, $\frac{\partial g}{\partial N} > 0$; when $N \geq N^*$, $\frac{\partial g}{\partial N} < 0$.*

Proof. See Appendix A.

Now we have to discuss how the actual market structure N is jointly determined by underlying technological parameters and a parameter measuring the effect of institutional barriers to entry. The crucial assumption we make is that the interplay of the technological parameters and the institutional barrier parameter is such that,

- the technological parameters determine a benchmark market structure, such as in a free-entry equilibrium ($N_f = n(\beta, \gamma)$).
- the institutional barrier parameter then determines how far the restricted number of potential entrants can deviate from the benchmark market structure determined above.

Denoting the institutional barrier parameter by ϕ, the interplay of the above two effects is captured by that

$$N = \phi + (1 - \phi) n(\beta, \gamma), \text{ for } \phi \in [0, 1].$$

Note that ϕ is the weight of a convex combination of the monopolist market structure and the free-entry market structure, $\phi = \frac{N_f - N}{N_f - 1}$, and that when $\phi = 0$, $N = N_f$ (no deviation); when $\phi = 1$, $N = 1$ (full deviation). Define $\phi^* = \frac{N^* - N}{N^* - 1}$, then by Lemma 4 it is straightforward to see

Proposition 5 *ϕ^* is growth-maximizing. ϕ^* is increasing in θ. When $\theta = 1$, $\phi^* = 0$. Furthermore when $\phi < \phi^*$, g is increasing in ϕ; when $\phi > \phi^*$, g is decreasing in ϕ.*

Results obtained in this section should be interpreted with Diagram 2 in mind, being aware of the directions of the "chains of causation". In this appended setting, the influence of the industry level technological parameter β feeds into the 'natural tendency' of market structure, captured by the free-entry equilibrium market structure n. This natural tendency then interplays with the 'institutional barriers to entry' parameter ϕ to jointly determine the actual market structure N, the gross profits, and the split of the gross profits into R&D costs (F_0) and rents (Π). For a given β, ϕ is a controllable parameter which determines (N, F_0, Π) jointly. According to another underlying parameter θ, which measures the relative scarcity of hired scientists and engineers vs. entrepreneurs, (N, F_0, Π) feed into the determination of the rate of growth. So for a fixed β, the most interesting parameters are θ and ϕ, and we visualize their influences on the rate of growth g in figure 2.6.

$$\left.\begin{array}{l} \left.\begin{array}{r} \beta \to n \\ \phi \end{array}\right\} \to (N, F_0, \Pi) \\ \hspace{6em}\theta \\ \hspace{6em}\delta \end{array}\right\} \to g$$

Diagram 2. The causal relationship between β, n, N and g

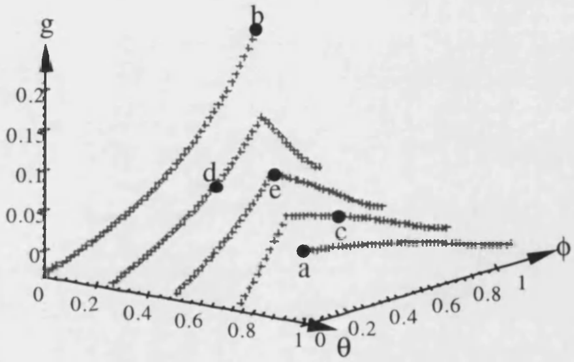


Figure 2.6: The relationship between ϕ and g for given θ

The graph shows that when $\theta = 1$, which describes the case of a single competition-push at work, then an institution which guarantees free entry (point **a**: $\theta = 1$, $\phi = 0$) is growth-maximizing. On the other hand, when $\theta = 0$, which describes the case of a single rent-pull engine at work, unsurprisingly, the institution which generates monopolist structure (point **b**: $\theta = 0$, $\phi = 1$) is growth maximizing. In between the two extreme cases lie three curves which describe cases where the twin engines are both at work. Then the relationship between ϕ and g is non-monotone, or more precisely of a "inverted U" shape. As for whether a given ϕ is too big (as at point **c**), too small (as at point **d**) or just right (as at point **e**) for the purpose of growth, it all subtly depends on the parameter θ. When θ decreases (i.e., the rent-pull becomes more important), relevant institutional barriers to entry become more desirable for the purpose of growth. However, it is also quite safe to claim that monopolization is not in general growth-maximizing, given that θ should not be too close to 0 generically. This final point is formally stated by the following

Proposition 6 *When $\theta > \frac{\gamma}{\beta(1-\gamma)}$, $N^* > 1$, i.e., monopolization is not growth-maximizing.*

Proof. $\theta > \frac{\gamma}{\beta(1-\gamma)} \Rightarrow n_1 > 2 - \frac{1}{2}\gamma \Rightarrow N^* > 1$. ■

Consider a numerical example: $\gamma = 1/3$, $\beta = 5$, $N_f = 3.0611$, $\frac{\gamma}{\beta(1-\gamma)} = 0.1$. It is quite reasonable to assume that $\theta > 0.1$, therefore monopolization should not be growth-maximizing in this case. The policy implication is stark: to keep the essential competition-push engine to work for the purpose of growth, monopolization of broad markets should not be encouraged.

2.4.4 Romer (1990) Revisited

Having fully characterized a model of the twin engines of growth, it is then interesting to see whether this framework can help to improve our understanding of some of the major existing R&D-based growth models, from a different perspective. Taking as an example the textbook version[14] of Romer (1990) model, two interesting questions can be asked:

1. can the result in the Romer model being replicated from this framework by taking a limiting case?

2. if 'yes' to the above question, what then can be said about the 'close neighborhood' of the Romer model within the scope of this framework?

Indeed the answer to the first question is 'yes', as elaborated by the following

[14]See for example Aghion and Howitt (1998) provides a textbook version of Romer (1990) model.

Proposition 7 [15]*If* $\theta = 0$, $\phi = 1$, $\beta \to \infty$, *then* $g \to \frac{\gamma\delta L-\rho}{\varepsilon+\gamma}$ *(simplified Romer result).*

Proof. $\theta = 0$, $\phi = 1 \Rightarrow g = \frac{\gamma\delta L-\rho}{\varepsilon+\gamma} - \frac{\gamma^2\delta L}{\beta(\varepsilon+\gamma)(1-\gamma)}$, $\Rightarrow \lim_{\beta\to\infty} g = \frac{\gamma\delta L-\rho}{\varepsilon+\gamma}$. ∎

This result suggests that if (1) there is no R&D spillover effects from the post-entry R&D, (2) the post-entry R&D is also a very ineffective means of competition, and (3) the institutional barriers to entry guarantee monopolization, then the outcome in the Romer (1990) model can be derived from this framework.

In the "close neighborhood" of the Romer (1990) model, i.e., setting $\theta = 0$ and $\phi = 1$, the closed-form growth rate then is reduced to:

$$g = \frac{\gamma\delta L - \rho}{\varepsilon + \gamma} - \frac{\gamma^2\delta L}{\beta\,(\varepsilon+\gamma)\,(1-\gamma)}, \tag{2.17}$$

which unambiguously increases in β.

Proposition 8 *If* $\theta = 0$, $\phi = 1$, $\frac{\partial g}{\partial \beta} > 0$, *i.e., in the above setting of 'monopolistic competition with (post-entry) quality choice' the more effective the post-entry R&D is, the lower the rate of growth.*

The intuition is straightforward: the more effective the post-entry R&D, monopolists will be more involved in cross-industry quality competition, which will only hurt rents and growth.

This analysis sheds new light on what a monopolistic competition setting tends to imply about the relation between competition and growth. Although in the setting of the Romer (1990) model there is no appropriate way[16] to relate competition and growth, the analysis of a setting in the 'close neighborhood' does reveal useful information on its 'logical implication'.

[15]Note, in a simplified Romer model (see for example Aghion and Howitt (1998, pp35-39)) the closed-form growth rate is equivalent to $g = \frac{\gamma\delta L-\rho}{\varepsilon+\gamma}$.

[16]Note that in the simplified Romer (1990) model the parameter γ which relates to demand elasticity, is also related to the share of capital in income distribution in the final sector. The mixed meaning makes it not an appropriate measure of competition.

2.5 Extending the Model: The Impact of Credit Constraint

The previous section examined how quality competition can be affected by institutional barriers to entry. This section looks at an extension of the basic model by asking how quality competition may be restricted in another possible way, viz. through being constrained by credit limits. This line of analysis leads to one possible explanation of the observed link between financial development and economic growth. The idea explored here is that quality competition-induced investment in R&D is usually large in absolute terms and takes an illiquid form, an so it requires external finance. This results in the capital market playing an important role in endogenous growth through the financing of R&D. In the basic model, it has been implicitly assumed that the capital market is perfect, so these issues don't arise. In practice, however, this assumption is very restrictive, and needs to be relaxed.

The reasons why the capital market may not be perfect lie beyond the scope of this thesis[17]. Here we simply explore the consequences that follow if firms cannot finance investment up to their preferred levels. The simplest way to capture this problem is to assume that there are exogenously imposed credit limits on all firms' R&D expenditures. Since in the symmetric balanced growth equilibrium there is a one-to-one mapping between R&D cost F and quality level u, credit constraints impose constraints on quality levels. The analysis can be simplified, therefore, by simply assuming that there are exogenously imposed constraints on quality levels. Consequently, the unconstrained maximization problem (2.4) is replaced by the

[17]For example, Kiyotaki and Moore (1997) attribute the reason why creditors may set credit limits to the fact that some debtors cannot credibly commit *not* to strategically default on debt repayments.

following constrained maximization problem

$$\max_{u_i} \frac{S}{r(1-\gamma)} \left(1 - \frac{N_c - 1 + \gamma}{\sum_{j=1}^{N_c} \frac{u_i}{u_j}}\right)^2 - \frac{w\mu}{A} u_i^\beta,$$

$$s.t. \ : \ u_i^\beta \leq \eta u^\beta \tag{2.18}$$

where $u^\beta = \frac{1}{\mu N_f \delta}$ is the unconstrained maximum point in the balanced growth equilibrium (See eq. (2.10)) and η is the index of credit constraint, $0 < \eta < 1$.

2.5.1 Corner Solutions and Fragmented Market Structure

The definition of η implies that when $\eta < 1$ the unconstrained equilibrium quality level is no longer feasible, therefore the credit constraint must be binding and it results in the following corner solution to the maximization problem,

$$u_i^\beta = u_c^\beta = \eta \frac{1}{\mu N_f \delta}, \tag{2.19}$$

where u_c is the constrained maximum point for each firm.

Similar to eq. (2.10), in the balanced growth equilibrium each firm's quality level is pinned down by

$$u_c^\beta = \frac{1}{\mu N_c \delta},$$

where N_c is the number of firms in each industry. Comparing the above two equations immediately reveals the relationship between N_c and N_f as follows,

$$N_c = \frac{N_f}{\eta}, \text{ for } 0 < \eta < 1, \tag{2.20}$$

which then implies

Proposition 9 $\frac{\partial N_c}{\partial \eta} < 0$, *i.e., tighter credit constraint leads to more fragmented market structure.*

Figure 2.7 shows that when there is a credit constraint, the Best-reply condition no longer holds. Instead of point **a**, point **b** becomes the equilibrium, where the industry R&D cost level is lower and the market structure is more fragmented.

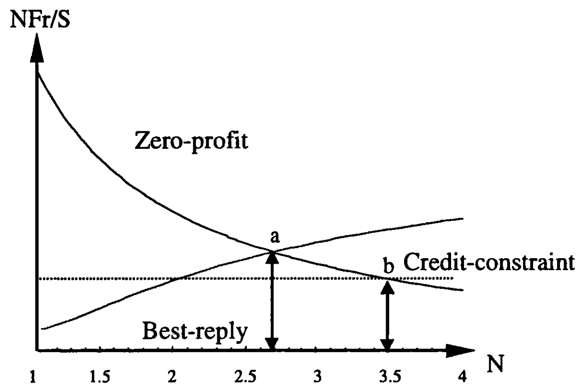


Figure 2.7: Credit constraint and market structure

2.5.2 Credit Constrained Growth

In this extension of the model, the growth function (2.9) is replaced by

$$g = \frac{\gamma\delta L - \rho N_c}{\gamma + N_c\varepsilon} = \frac{\gamma\delta L - \rho\frac{n(\beta,\gamma)}{\eta}}{\gamma + \frac{n(\beta,\gamma)}{\eta}\varepsilon}. \tag{2.21}$$

Simple comparative statics shows that

$$\frac{\partial g}{\partial \eta} > 0, \tag{2.22}$$

which means

Proposition 10 *If there is a binding credit constraint in the capital market, i.e., $\eta < 1$, then g is increasing in η, i.e., ceteris paribus, the less credit-constrained economy has higher growth rate.*

Diagram 3 reveals the causal relationship between variables. For a given β, the binding credit constraint η ($\eta < 1$) causes a deviation of the actual market structure (N_c) from its natural tendency (n) while imposing its influence on R&D expenditures (F). Through the causal relation between R&D and the rate of growth, tighter credit constraint (smaller η) eventually has a negative impact on g, as shown in Figure 2.8 and eq. (2.22). Since tightening credit constraint (decreasing η) has a negative

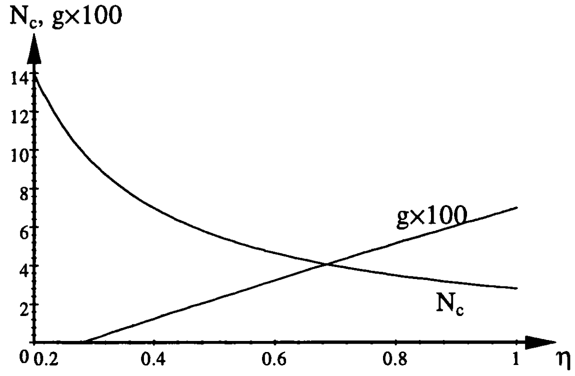


Figure 2.8: η jointly determines N_c and g

impact on g, and a positive effect on N_c, the correlation between N_c and g then must be negative.

$$\left.\begin{array}{c} \left.\begin{array}{l} \beta \to n \\ \quad\;\; \eta \end{array}\right\} \to (N_c, F) \\ \delta \end{array}\right\} \to g$$

Diagram 3. The causal relationship between β, n, N_c and g

These stark results are consistent with the empirical finding that there is a positive correlation between financial development and growth performance across countries. (King and Levine, 1993, Rajan and Zingales, 1996). The new insight gained from this analysis suggests a new channel through which finance can affect growth. This analysis emphasizes that finance can affect the competitive pattern. When firms are financially constrained in competing in quality, they are forced to compete in price in more fragmented markets. But the work of the 'competition-push' engine relies on quality competition rather than price competition, it therefore slows down. So this analysis predicts that a low growth rate is likely to be associated with market fragmentation, conditional on there not being institutional barriers to entry. This prediction is consistent[18] with the evidence provided by Kumar, Rajan and Zingales

[18]This claim relies on two maintained hypotheses: (1) economic growth rate is positively corre-

(1999) showing that firm size in external-finance-dependent industries is positively correlated with financial development.

Furthermore, although this analysis only focuses on the impact of credit constraints on quality competition and growth, the logic of the analysis can well be extended to studying the impact of constraints in the human capital market, such as employees' adoption and investment in learning and improving new technologies.[19]

2.5.3 A Poverty Trap

Eq. (2.21) also implies that there is a threshold of η, below which the sustained positive balanced growth rate is impossible, i.e.,

$$\eta \leq \frac{\rho n(\beta,\gamma)}{\gamma\delta L} \Rightarrow g \leq 0, \tag{2.23}$$

which suggests

Proposition 11 *There exists a lower bound of η, $\eta_{low} \equiv \frac{\rho n(\beta,\gamma)}{\gamma\delta L}$, such that if $\eta \leq \eta_{low}$, then $g = 0$, i.e., growth can not be sustained in an economy which has too tight credit constraints in capital market.*[20]

The implication is that the competition-push engine of growth cannot function properly due to imperfection in the capital market. Severe financial underdevelopment may cause the engine to stop working at all.

It is worth mentioning that the credit constraint discussed above is not necessarily due to unavailability of funds within an economy. It may be mainly due to the fact that the infrastructure, and institutions in the economy, are incapable of solving various informational problems or contract enforcement problems. It may be the

lated with financial development, and (2) firm size is positively correlated with market concentration.

[19]Related to this view, Chandler (1990) has already addressed the possibility that institutional constraints in the human capital market may restrict the exploitation of scale economies.

[20]Due to irreversibility of innovation, g is bounded from below by 0 in our model, but zero growth is possible.

case that this is a major contributory factor underlying poor economic performance. If so, poverty is not primarily driven by a lack of accumulation of physical capital. The ultimate cause lies in the idea that the institutions are incapable of supporting growth through their failure to allocate capital efficiently.

2.6 Conclusion

Understanding how quality competition jointly determines market structure and the rate of growth is the key to understanding why 'competition-push' is an essential engine of growth. The analysis needs to be carried out by beginning with the set of underlying parameters that determine both variables jointly. At this level, we can re-open the question regarding the relation between market structure and growth. It is suggested here that the interplay of the 'twin engines of growth', namely 'competition-push' and the more familiar 'rent-pull', complicates the link between market concentration and the rate of growth, leading to an inverted U shape relationship. It follows that, the need for 'rent-pull' notwithstanding, monopolization is, in general, not good for growth.

Chapter 3

Submarkets, Shakeouts and Industry Life Cycle

3.1 Introduction

In the past decade the study of industry life-cycles in the firm-growth literature has devoted most of its attention to characterizing and explaining the process of 'shakeout'[1], i.e., the statistical regularity according to which the number of producers tends to first rise to a peak and later falls to some lower level. This process has been observed in a large number of industries. Recent empirical findings by Horvath et al. (1997) and Klepper (1999) shed light on two other statistical regularities regarding the industry life-cycle: (1) 'turbulence' (firm turnover), i.e., the statistical regularity that the entry-exit process persists throughout an industry life-cycle and gross entry and gross exit rates are positively correlated; and (2) the cohort survival pattern, i.e., all entry cohorts share a qualitatively similar survival pattern, which displays a significantly higher exit hazard rate at early age than subsequent ages. These patterns may be closely associated with the process of 'shakeout'. This chapter presents a theoretical model which offers an explanation of these regularities, and

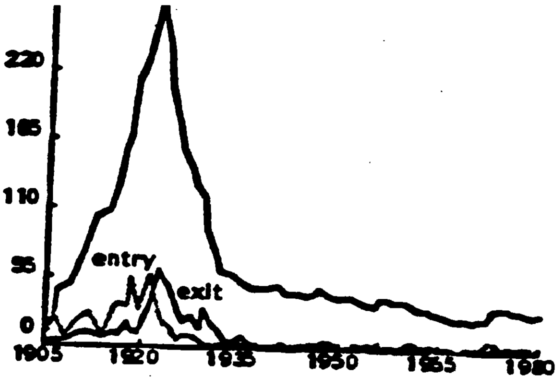
[1] See for example, Klepper (1990), Jovanovic and MacDonald (1994), Klepper (1996, 1999).

of the links between them.

Panel A of Figure 3.1 shows a striking example of 'shakeout' and the associated pattern of firm turnover ('turbulence'). Panel B of Figure 1 shows the cohort survival pattern. When the 'turbulence' and the cohort survival pattern are jointly examined, a surprising pattern emerges. As first noted by Horvath et al. (1997), despite the fluctuation in entry rates, the timing of exits for different cohorts of entrants is remarkably similar over time: the exit hazard rate is peaked at a very early age of every cohort's life and drops dramatically to low levels for subsequent ages[2]. In this sense, a typical industry life-cycle can be roughly re-described as follows: a miniature shakeout (i.e., an excess entry followed by dramatic early-age exit then followed by gradual subsequent exits) actually happens in the life of each cohort of entrants in a similar way throughout the whole industry life-cycle, and the industry level shakeout that has been widely observed emerges as an aggregation of these overlapping-cohort miniature shakeout, in association with a gross entry pattern in which the early-stage and late-stage cohorts have small numbers of entrants and the interim cohorts have large numbers of entrants.[3]

[2]Note that the vertical axis of Panel B of Figure 1 is in $\log_{10}$ scale, which implies that a seemingly straight line in such a space would actually be as convex as a $\log_{10}$ function if the vertical axis were in a linear scale. Such a convex curve would mean that the earlier the age the higher the exit rate.

[3]It is worth mentioning that such a stylized 'industry life-cycle' is directly related to three of the four well documented statistical regularities, which are about (1) the size-survival-growth relationship and size distribution, (2) age-survival-growth relationship, (3) the shakeout, and (4) the turbulence (or firm turnover). For detailed description on these, see Sutton (1997b).



A.Number of Firms (Tires, US)

horizontal axis: year

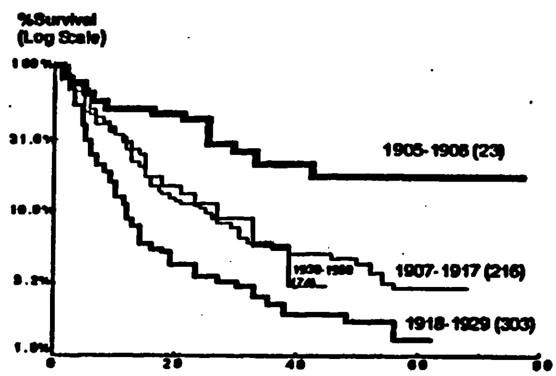


Figure 3.1: B. Cohort Survival Patern (Tires,US)

horizontal axis: year, vertical axis: survival rate (in log scale)

Source: Klepper 1999

The example of the tires industry shown in Figure 3.1, however, is an extreme case. Dramatic aggregate shakeout is not a universal phenomenon. According to research currently in progress by Steven Klepper, a good example of an exception to the 'shakeout' pattern occurs in the laser industry, which has experienced a rather long history of growth but so far has shown no sign of an aggregate shakeout. These differences in the dynamics of aggregate firm numbers across industries notwithstanding, one statistical regularity holds good across the general run of conventionally-defined (4-digit SIC) industries. That is, there are persistent waves of entries over time. The key message conveyed by this simple fact is that independent opportunities keep emerging in a conventionally-defined industry before it matures. This observation matches with the insight described by Sutton (1997a 1997b, 1998), that " most conventionally defined industries exhibit both some strategic interdependence, and some degree of independence across submarkets". If an industry comprises many independent submarkets, then it is natural to see independent opportunities emerge over time, which attract persistent waves of entries. When the notion of independent submarkets *a la* Sutton (1997a 1997b, 1998) is applied to the issue of industrial growth, the logic would suggest that both the pattern of industrial expansion through emergence of independent submarkets, and effect of strategic interaction within each submarket, should leave their fingerprints in the observed pattern of the industry life-cycle.

Suppose for example, that within each submarket the strategic interaction takes the form of price competition and quality choice by producers. (We may interpret quality competition as taking the form of vertical product differentiation, or cost-reducing process innovation, which requires endogenous R&D or advertising costs.) As we noted in Chapter 2, a concentrated market is the normal outcome where quality competition prevails.[4]

This chapter proposes an unusual but plausible extension of the game-theoretic quality competition literature, which emphasizes the role of independent submarkets, in order to explain the aforementioned statistical regularities noted above. The major scenario to be described is as follows. The uncertainty and informational problems surrounding the opening of a new submarket tend to impose credit constraints on fixed expenditures by producers. This in turn restricts the pressure of quality competition and leads to the viability of an excessive number of entrants in

[4]As noted in Chapter 2, this scenario is best formalized by Dasgupta and Stiglitz (1980), Sutton (1991, Ch. 3) and Sutton (1998, Ch. 15) in the game-theoretic literature. This theme is echoed by Klepper (1996, 1999) in the firm-growth literature.

the early life of a submarket. The miniature shakeout takes place later when the initial credit constraints are removed as initial uncertainty regarding the market is resolved, and this leads to the escalation of quality competition, which induces the exit of a large fraction of existing producers in the submarket. Since an industry usually contains many independent submarkets, which emerge and develop in a sequence over time, the above scenario is repeated over time, and this induces 'firm turnover' in the industry. When the emergence of submarkets slows down, gross exit of producers will eventually dominate gross entry, and an aggregate shakeout takes place.

3.2 The Model

3.2.1 'Independent' Submarkets

There are S identical consumers in the economy, each of whose utility function has the form of:

$$U = \frac{1}{\Psi(k)} \sum_{i=1}^{k} x_i^{\gamma(k)} + y,$$

where x_i is the consumption of variety i of the 'X' good, k is number of varieties of the 'X' good in the given period, y is the numeraire, standing for all other goods. The difference between this utility function and the usual formulation is that it allows the increase in k to bring some unconventional shocks to the utility function. This feature is embedded in $\Psi(k)$ and $\gamma(k)$ such that $0 < \gamma(k) < 1$, $\frac{\partial \gamma(k)}{\partial k} > 0$ and $\frac{\partial \Psi(k)}{\partial k} > 0$, therefore they capture the idea that the increase of the varieties has a business stealing effect on all existing varieties and it makes all varieties closer substitutes between themselves. The strength of the business stealing effect will depend on the specification[5] of $\Psi(k)$ and $\gamma(k)$. For the sake of simplicity

[5]For example, if it is specified that $\Psi(k) = k$, the utility function will become $U = \frac{1}{k}\sum_{i=1}^{k} x_i^{\gamma(k)} + y$, which implies a very strong business stealing effect.

and without loss of generality over the issues to be discussed, it is specified as $\Psi(k) = \gamma(k)$, hence the utility function is specified further to

$$U = \frac{1}{\gamma(k)} \sum_{i=1}^{k} x_i^{\gamma(k)} + y. \tag{3.1}$$

In any period each consumer maximizes U subject to the budget constraint: $\sum_{i=1}^{k} p_i x_i + y \leq I$, where I is the total consumption in the given period[6]. The first order condition of this maximization program is:

$$x_i^{\gamma-1} = p_i,$$

which implies that the demand function of variety i is

$$X_i = S\left(\frac{1}{p_i}\right)^{\frac{1}{1-\gamma}}, \tag{3.2}$$

where $X_i \equiv Sx_i$.

The price elasticity of demand is constant in each period as follows,

$$\zeta \equiv -\frac{\partial \ln X_i}{\partial \ln p_i} = \frac{1}{1-\gamma}, \tag{3.3}$$

for which $\frac{\partial \gamma(k)}{\partial k} > 0$ implies $\frac{\partial \zeta}{\partial k} > 0$ (elasticity enhancing effect).

The above specific utility function implies that the business stealing effect notwithstanding, the demand over each variety is only dependent on its own price in any given period. In other words, the submarkets of the 'X' industry are strategically independent of each other.

3.2.2 Industrial Growth via the Emergence of Submarkets

The above specific utility function also implies that the growth of the industry is through the emergence of new submarkets. It is further assumed in this chapter

[6]Each consumer's intertemporal optimization program is completely trivialized by the specification of a quasi-linear utility function and unity discount factor, and the implicit assumption that all 'X' goods are unstorable.

that the number of submarkets in period t, i.e., $k(t)$, follows a generalized Logistic diffusion curve:

$$\begin{cases} \dot{k} = ak^{\theta}(b-k)^{\lambda} \\ k(t) = 0 \text{ for } t < 0 \\ k(0) = 1 \end{cases}, \tag{3.4}$$

where b ($b \geq 1$) is the saturation number of independent submarkets within the industry. This law of motion captures the feature that growth of the number of submarkets is dependent on the existing number and the potential which hasn't been fulfilled. This feature can be demonstrated easily with the example: $\theta = \lambda = 1$, which simplifies the law of motion to:

$$\dot{k} = ak(b-k),$$

to which the closed-form solution is[7]:

$$k = \frac{be^{abt}}{c + e^{abt}},$$

where c is a constant. The implied rate of emergence of submarkets is:

$$\dot{k} = \frac{ab^2ce^{abt}}{(c + e^{abt})^2}.$$

Figures 3.2 and 3.3 show some general features of the generalized Logistic diffusion curve: (1) it is initially convex up to some point, then (2) it becomes concave, and finally (3) it becomes flat. Accordingly, the growth rate initially increases up to a peak, then declines, finally converges to zero.

The second order derivative of $k(t)$ is as follows,

$$\ddot{k} = a(\theta+\lambda)k^{\theta-1}(b-k)^{\lambda-1}\left(\frac{\theta b}{\theta+\lambda} - k\right) \gtreqless 0 \text{ when } k \lesseqgtr \frac{\theta b}{\theta+\lambda},$$

which reveals that the peak of growth rate is located at the point such that $k = \frac{\theta b}{\theta+\lambda}$.

[7]Note that $\frac{1}{k}\frac{dk}{dt} + \frac{1}{b-k}\frac{dk}{dt} = ab$ and $\frac{d}{dt}\ln\frac{k}{(b-k)} = ab$, which imply the displayed general solution.

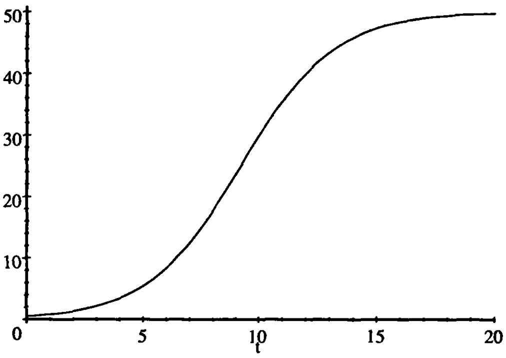


Figure 3.2: $k(t)$

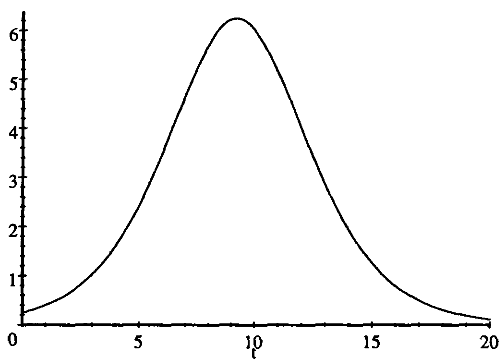


Figure 3.3: $\dot{k}(t)$

3.2.3 Modelling Competition within a Submarket

Whenever a new independent submarket emerges, the following game starts to be played.

Phase One (credit-constrained phase):

Stage one: entry decision when there is credit constraint $\rightarrow N_1$,

Stage two: quality (or productivity) choice,

Stage three: quantity competition (Cournot).

Phase Two (credit-unconstrained phase):

Stage one: entry decision when there is no credit constraint$\rightarrow N_2(t)$,

Stage two: quality (or productivity) choice,

Stage three: quantity competition (Cournot).

The Phase One game is played only in one period, and the Phase Two game is repeated in each subsequent period. The rationale of this game is that the first period, which is the beginning of the submarket, is marked by the uncertainty about each producer's capability of handling the technology of improving quality, therefore no external finance is involved in the investments in quality. As a result each producer's quality choice is restricted by a credit constraint. At the end of the first period, the uncertainty is resolved. Only those players who have proven high capabilities enter Phase Two, when credit constraints are removed. In other words, after one period in the history of each submarket, a wave of selective financial shocks affects the efficient players and transforms them into credit-un-constrained players. It is further assumed that only those players who have successful track records in the previous period are eligible for re-entry in any period of phase two.

The above description is the idea which motivates our modelling. However for the sake of simplicity, we will not explicitly model the efficiency-based financial selection process. Instead we will assume that some random selection takes place among the symmetric players, which selects a particular equilibrium among all the possible symmetric equilibrium outcomes.

3.2.4 The Games of Quality Competition with and without Credit Constraints

It is assumed that the quality achieved by each player in any period is not carried over to the next period due to depreciation. Under this assumption, except for a shift of parameter, each period of the repeated game almost becomes an isolated game, the equilibrium of which can be characterized without reference to what happens elsewhere. Therefore the subgame perfect Nash equilibrium is a sufficient solution concept for characterizing the game formulated in this chapter.

The three-stage game in each submarket during each period then can be solved by backward induction. The third stage subgame is always a Cournot game of quantity competition, for which a Cournot-Nash equilibrium exists and is unique up to a given set of quality levels of all incumbents. The subgame Nash equilibrium determines for each producer a reduced form profit function as follows[8]:

$$\pi_i \left(u_i \mid u_{-i}\right) = S \left(\frac{N-1+\gamma}{\sum_{j=1}^{N} \frac{1}{u_j}} \right)^{\frac{\gamma}{1-\gamma}} \frac{1}{1-\gamma} \left(1 - \frac{N-1+\gamma}{\sum_{j=1}^{N} \frac{u_i}{u_j}} \right)^2, \tag{3.5}$$

where u_i is the quality level of producer i, u_{-i} is a $(N-1)$-tuple of quality levels of other producers except producer i, for $i = 1, 2, \cdots, N$; N is the number of producers active in the submarket and S is the population of consumers in the economy.

Interested readers can find the derivation of this reduced form profit function in Appendix B. The key feature of it is that a producer's profit increases with its relative quality level against its rivals', i.e., $\sum_{j=1}^{N} \frac{u_i}{u_j}$. This function tells how the strategic environment responds to vertical differentiations of producers. It captures the Darwinian selection pressure embedded in the environment constituted

[8]This formulation is an extension of the one developed in Appendix 15.1 of Sutton (1998). It allows a general $\gamma \in [0, 1)$ while the Sutton (1998) formulation deals with a special case: $\gamma = 0$.

by customers and rivals. This environment provides producers with incentives to outperform their rivals in R&D and quality. When they vie in quality, each of them has to bear a fixed cost, which is a function of quality level as shown below:

$$F(u_i) = \begin{cases} \mu u_i^\beta & \text{in credit-un-constrained phase} \\ \begin{cases} \mu u_i^\beta & \text{if } \mu u_i^\beta \leq \delta \\ \infty & \text{if } \mu u_i^\beta > \delta \end{cases} & \text{in credit-constrained phase} \end{cases}, \tag{3.6}$$

where δ is the credit limit in the first period. The above fixed cost function is depicted in Figure 3.4, which shows that in the credit constrained phase, the effect of the credit constraint (δ) is equivalent to putting an upper bound to the quality level.

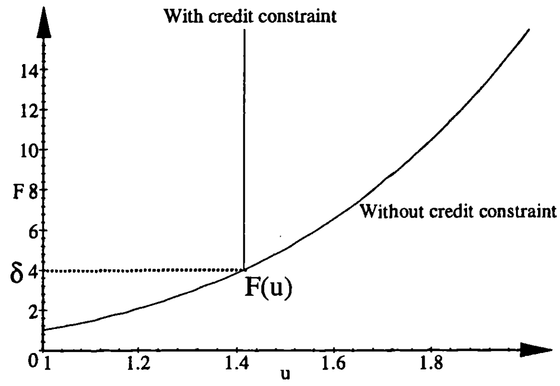


Figure 3.4: Fixed cost functions

Therefore each producer's objective in second stage subgame is to maximize its net profit by choosing own quality level given others' quality levels,

$$\max_{u_i} \left\{ S \left(\frac{N-1+\gamma}{\sum_{j=1}^{N} \frac{1}{u_j}} \right)^{\frac{\gamma}{1-\gamma}} \left(1 - \frac{N-1+\gamma}{\sum_{j=1}^{N} \frac{u_i}{u_j}} \right)^2 - F(u_i) \right\}. \tag{3.7}$$

If we take the above payoff function as given, then the game in each period is merely a simple game of quality competition with free entry.

3.3 Equilibrium

3.3.1 Market Structure in the Credit-un-constrained Phase

We proceed to characterize the equilibrium, starting with Phase Two, when credit constraints have been removed for remaining players. The focus is on a symmetric subgame perfect Nash equilibrium, where each producer equates the marginal benefit of increasing quality and the marginal cost. The implied best-reply function for each producer in a symmetric outcome is

$$F = \mu u^{\beta} = \frac{S}{\beta}\left(\left(1-\frac{1-\gamma}{N_2}\right)u\right)^{\frac{\gamma}{1-\gamma}}\left(2\left(N_2-1\right)\frac{N_2-1+\gamma}{N_2^3}+\frac{\gamma}{N_2^3}\right), \tag{3.8}$$

where N_2 is the number of producers in a Phase Two equilibrium.

Free entry into the submarket results in the following zero profit condition in a symmetric outcome:

$$F = \mu u^{\beta} = S\frac{1-\gamma}{N_2^2}\left(\left(1-\frac{1-\gamma}{N_2}\right)u\right)^{\frac{\gamma}{1-\gamma}}. \tag{3.9}$$

The above two conditions determine the equilibrium submarket structure and the equilibrium endogenous fixed cost level as shown in Figure 3.5. The graph also indicates that the equilibrium without credit constraints depends on parameter γ: when γ increases the submarket structure becomes more concentrated (point **B** vs. point **A**), i.e., N decreases.[9]

Proposition 12 [10]*In a symmetric Phase Two equilibrium, the number of producers is* $N_2 = n(\beta,\gamma) \equiv \frac{n_0+\sqrt{n_0^2-2(2-\gamma)}}{2}$, *where* $n_0 \equiv 1+\frac{(\beta+2)(1-\gamma)}{2}$. *Furthermore* N_2 *is decreasing in* γ.

Proof. See Appendix B.

[9] An increase in γ shifts the Zero-profit curve downward and to the left, but shifts the Best-reply curve upward and to the right. The net effect however, is dominated by the former. In fact, as illustrated in the graph the second effect is quantitatively very small compared to the first, so it can be ignored for the sake of simplicity.

[10] The result is consistent with that found in Dasgupta-Stiglitz (1980).

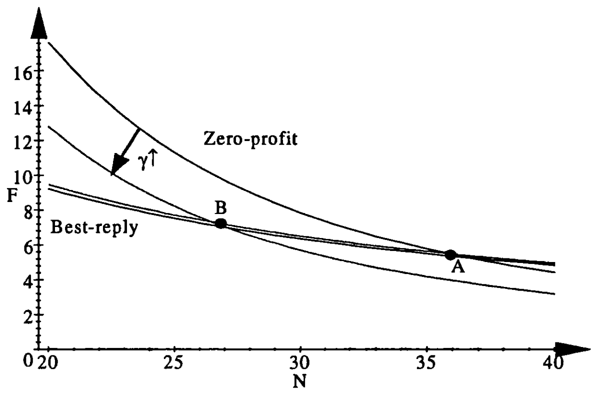


Figure 3.5: Equiliria without credit constraints

3.3.2 Market Structure in the Credit-constrained Phase

It is assumed that with the credit constraints, the credit-un-constrained equilibrium fixed cost level is not feasible, therefore the credit constraints must be binding. So the following binding credit constraint should replace the best reply condition in determining the equilibrium fixed cost level:

$$F_c = \mu u_c^\beta = \delta. \tag{3.10}$$

Accordingly, the zero profit condition should be modified to:

$$F_c = \mu u_c^\beta = S\frac{1-\gamma}{N_1^2}\left(\left(1-\frac{1-\gamma}{N_1}\right)u_c\right)^{\frac{\gamma}{1-\gamma}}, \tag{3.11}$$

where N_1 is the number of producers in the Phase One (credit-constrained) equilibrium.

The above two conditions determine the equilibrium submarket structure and the equilibrium endogenous fixed cost level as shown in Figure 3.6. It can be seen from the graph that the submarket is more fragmented when there is credit constraint (point **B**) than when there isn't (point **A**), yet the tighter the credit constraints the more fragmented the submarket (point **C** vs. point **B**).

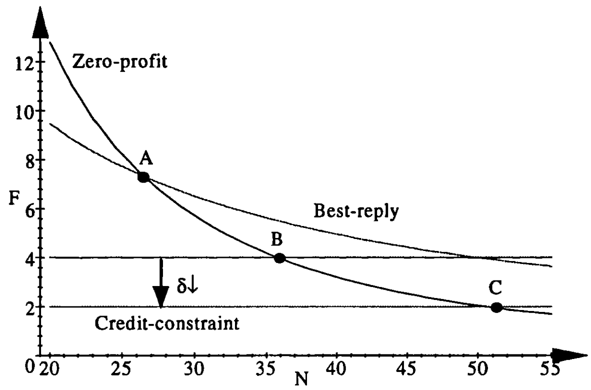


Figure 3.6: Equilibria with and without credit constraints

Proposition 13 *Under the assumption: $\beta > \frac{\gamma}{1-\gamma}$, the tighter the binding credit constraints are, the more fragmented the submarket is. Also, there is a lower bound to N_1 such that $N_1 > n(\beta, \gamma)$, i.e., when the credit constraints are strictly binding, the submarket is more fragmented than if there were no credit constraints.*

Proof. See Appendix B.

3.3.3 Miniature Shakeouts and Cohort Survival Pattern

Since all submarkets which emerge in the same period are similar, any cohort dynamics is merely a reflection of the dynamics of a representative submarket which belongs to that cohort, so a good point of departure in understanding cohort dynamics is to look at what kind of shocks a typical submarket experiences over time.

As the industry grows, it becomes 'tighter' in the sense that it is filled with more submarkets. Consequently, the price elasticity of demand in each submarket increases. This insight is captured by $\frac{\partial \gamma_t}{\partial k} > 0$, which implies:

$$\frac{d\gamma_t}{dt} = \frac{\partial \gamma_t}{\partial k} \frac{dk}{dt} > 0. \tag{3.12}$$

For a representative submarket, this means a series of shocks to the price elasticity of demand over time. By Proposition 12, the effects of these shocks should

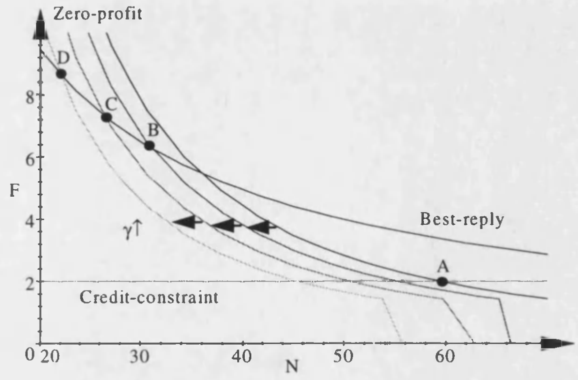


Figure 3.7: Evolution of submarket structure

make the submarket more and more concentrated, which implies a persistent exit of producers over time. In addition to this kind of exit pressure, the selective financial shocks which occur in the first period of the representative submarket may cause a more dramatic wave of exit of incumbents.

The evolution of the of structure of a typical submarket is illustrated by Figure 3.7, where point **A** stands for the initial credit-constrained equilibrium and points **B**, **C** and **D** stand for the subsequent credit-un-constrained equilibria. The jump from **A** to **B** is caused by two waves of shocks: the selective financial shocks and the shock to demand elasticity. Subsequent shifts: from **B** to **C**, from **C** to **D** and so on are due only to shocks to demand elasticity[11], and therefore are less dramatic. The typical cohort survival pattern is described by the following

Proposition 14 *Each cohort experiences a sequence of shocks such that after one period, there is a wave of selective financial shocks, and in all periods there are gamma shocks which shift the parameter γ upward. Consequently, each cohort experiences a miniature shakeout after one period, and continuous decline of number of producers in subsequent periods. If the binding credit constraints are sufficiently*

[11]The increases of γ should shift the best-reply curve upward and to the left. Since the effects of these changes are dominated by the effects of the changes to zero-profit curve, therefore are ignored in the graph, without causing non-trivial bias.

tight then each entry cohort has a higher first period hazard rate than subsequent periods.

3.3.4 Aggregate Shakeouts

We now examine the pattern that emerges at the aggregate (industry-wide) level. The aggregate pattern depends, for any given cohort survival pattern, on the industry-specific characteristics regarding growth, which can be captured by a set of parameters such as (a, b, θ, λ), embedded in the law of motion of industrial growth: $\frac{dk}{dt} = ak^{\theta}(b-k)^{\lambda}$.

For example, in an industry which consists of a large number of strategically independent submarkets where the links between submarkets are very weak, then b should be large, $\theta \to 0$ and $\lambda \to 0$. Consequently the observer would not see a dramatic aggregate shakeout at all. Instead, she would observe a very slow but steady build-up of number of producers during a long period in that industry. This pattern can arise due to the extreme specification of the generalized logistic curve:

$$\frac{dk}{dt} = ak^{\theta}(b-k)^{\lambda} \to a.$$

In less extreme settings, however, the observer would be able to see some kind of aggregate shakeout. Generally, the larger θ and λ in a industry, the more dramatic the aggregate shakeout. In the next section, we calibrate the parameters of the model and simulate the process of an industry life-cycle. Two different cases will be distinguished with respect to the aggregate pattern.

3.4 Simulations of Industry Life-Cycle

3.4.1 Case 1: With Aggregate Shakeouts

The specification of the k-γ relationship is as follows, $\gamma = 1 - \epsilon \exp(\omega(-k+1))$. The specification of the parameters are shown in the following table. (The range of parameter values has been chosen rather arbitrarily in order to illustrate the working of the mechanism, and is not justified by reference to specific empirical examples.)

Parameter:	a	b	θ	λ	β	δ	S	μ	ϵ	ω
Value:	0.0004	50	1.8	1.5	200	10	200000	1	0.9	0.05

The rough simulation indicates that the model can account for the typical patterns of aggregate shakeout, persistent turbulence and cohort dynamics. Figure 3.8 depicts the simulated aggregate shakeout and the correlated gross entry and gross exit over time. This graph resembles the real world picture presented in Panel A of Figure 3.1. The predicted correlation between gross entry and gross exit over time is 0.76, which is comparable with an estimated number from another source[12]. Figure 3.9 further breaks the aggregate shakeout pattern down to separated cohort dynamics. It suggests that the dramatic rise and fall of the total number of producers around the peak are largely due to impacts of a very few big cohorts (9, 10 and 11), which feature "mass entry waves followed by mass exit waves". The pattern of 'miniature shakeout' is demonstrated by Figure 3.10, which shows that the early-age exit hazard is significantly higher than subsequent ages in each cohort. The consequent typical cohort survival pattern then is presented in Figure 3.11, which shows the average survival rates over age. The simulated pattern qualitatively resembles the real life pattern presented in Panel B of Figure 3.1.

[12]Empirically, within any one country, a strong correlation is found to exist between entry and exit rates by industry. For example, Paul Geroski (1991) reports a correlation coefficient of 0.796 for a sample of 95 industries in the U.K. in 1987.

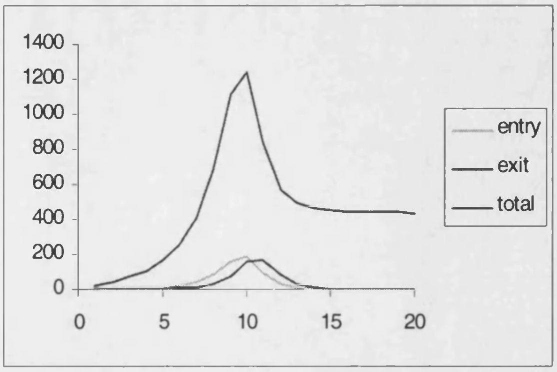


Figure 3.8: vertical axis: number of producers, horizontal axis: time (simulation 1)

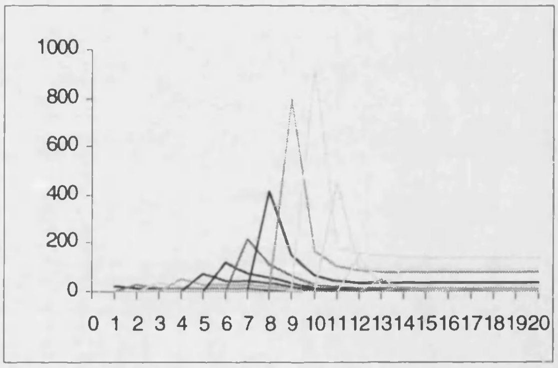


Figure 3.9: Cohort dynamics (simulation 1), vertical axis: number of producers, horizontal axis: time

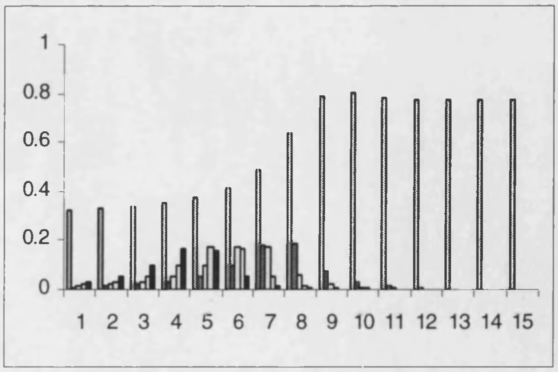


Figure 3.10: Cohort exit rates: age 2-5 (simulation 1), horizontal axis: cohort

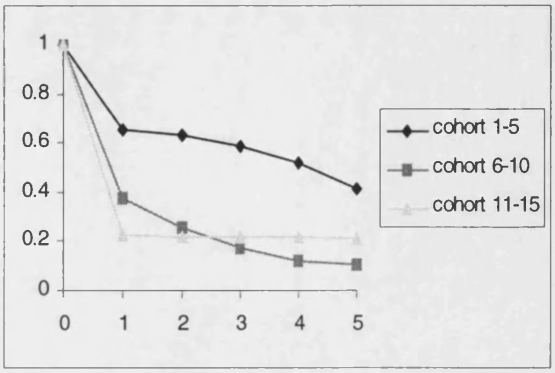


Figure 3.11: Average cohort survival pattern (simulation 1), horizontal axis: age

3.4.2 Case 2: Without Aggregate Shakeouts

Now we proceed to verify the point that if an industry consists of a large number of submarkets which have very weak links with the rest of the industry, i.e., b is sufficiently large and θ and λ are sufficiently small, then dramatic aggregate shakeout should not take place. The following specification of parameters embodies these features.

Parameter:	a	b	θ	λ	β	δ	S	μ	ϵ	ω
Value:	0.4	50	0.01	0.01	200	10	150000	1	0.9	0.0005

The simulation confirms the conjecture. Figure 3.12 shows a simulated industry life-cycle without an aggregate shakeout. It is worth mentioning that this result does not rely on any qualitative difference in cohort survival pattern. Figure 3.13-3.15 show that the simulated cohort survival pattern in this new specification is qualitatively similar to the previous one.

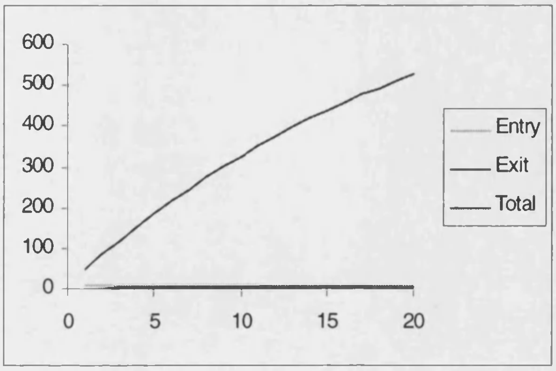


Figure 3.12: vertical axis: number of producers, horizontal axis: time (simulation 2)

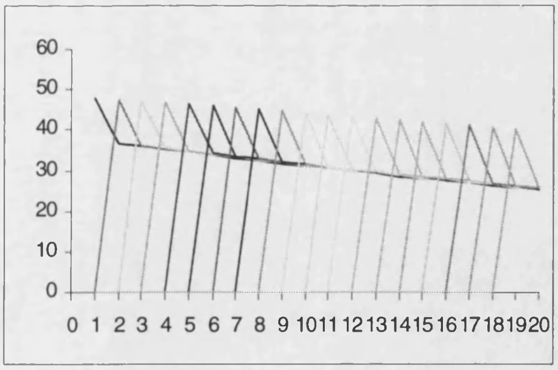


Figure 3.13: Cohort dynamics (simulation 2), vertical axis: number of producers, horizontal axis: time

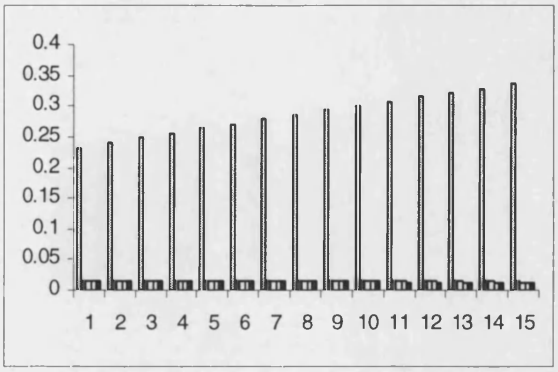


Figure 3.14: Cohort exit rates: age 2-5 (simulation 2), horizontal axis: cohort

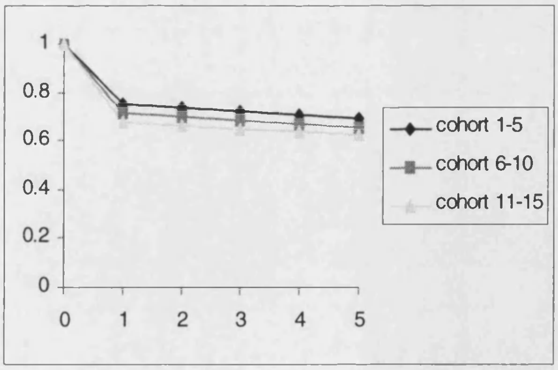


Figure 3.15: Average cohort survival pattern (simulation 2), horizontal axis: age

3.5 Concluding Remarks

This analysis uses the notion of 'independent submarkets' introduced by Sutton (1997a, 1997b), in which a game-theoretic model of the markets that incorporates strategic interactions within submarkets with (approximate) independence across submarkets. In this chapter, we have used this idea to explain some well-documented statistical regularities regarding the evolution of market structure. The results offer a simple candidate explanation for these regularities, while some extensions of the model might be needed in order to capture some secondary features of the process, this simple structure appears to be adequate as a representation of the main features of these processes on which empirical researchers have agreed.

Chapter 4

Radical Innovation, Selection of Technology and Market Structure

4.1 Introduction

It has been widely noted that the early history of many industries, such as aircraft, is marked by coexistence of rival technical trajectories (biplanes, monoplanes; wooden construction, metal construction, etc.). But in due course one trajectory emerges as the winner, or 'dominant trajectory' (Abernathy and Utterbach, 1978).

Even when an industry has settled down along some established trajectory, however, it is often the case that new discoveries make possible the exploration of a rival trajectory. This can have profound consequences for market structure.

For example, when valve technology was replaced by transistor technology, this was accompanied by the exit of all major US valve producers, and their replacement by a different population of firms, that developed expertise in the new silicon-based technologies (Sutton, 1998). A contemporary example, which is emerging at present, relates to traditional color film, which may be displaced by the new generation of digital cameras.

The relation between market competition and the rate and direction of technological change under such circumstances may take a complex form. On the one hand, competition may exist among firms all of whom follow the same "technology" of innovation, or 'technological trajectory'. On the other hand, different firms may follow different "technologies" of innovation, or 'technological trajectories', which are substitutable.

In this chapter we examine competition in an industry which faces the appearance of a new (R&D) trajectory (technology).[1]The new technology is superior to the old one in the sense that it costs less to achieve the same product quality level. Will the new technology replace the old one, will it fail to have any impact, or will both technologies coexist in the new equilibrium?

By characterizing equilibrium outcomes in terms of the arrival of new entrants, the subsequent form of quality competition and the exit of incumbent firms, we examine how the market resolves the competition between the rival technologies. We find that the outcome depends crucially on the relative attractiveness of the new technology as against the old one.

In one polar case, if the new trajectory is sufficiently superior to the old, then entrant firms following the new trajectory possess great advantages over incumbents in winning market share; their profits are large enough to cover their sunk costs and therefore justify their entry. On the other side, the incumbents' profitability is so badly damaged by the intensified competition that they are forced to exit the market. In the opposite polar case where the advantage of the new trajectory over the old one is relatively small, new entrants will not be able to win enough market share and earn sufficient profit to cover their initial fixed costs. Due to their first movers' advantage, the incumbents who suffer a slight technological disadvantage can successfully deter new entry. In cases intermediate between these extremes, we get the coexistence of incumbents and new entrants, and so the presence of both the

[1]There is of course an analogy between the competition among rival trajectories, and the competitive process analyzed by Schumpeter (1934), who was concerned with the process of 'creative destruction'.

old and the new technologies. An important conclusion in what follows is that these qualitative patterns are not affected by whether the innovation is unanticipated, or is anticipated, by incumbents.

The plan for the remainder of the chapter is as follows. Section 2 presents the basic model, starting with the simplest setting, where the technological breakthrough is unanticipated and takes place immediately after the incumbents have committed their sunk costs to the existing technology; the number of new entrants is assumed to be fixed to one at first, and is later endogenized. Section 3 looks at the case where a future technological breakthrough is anticipated, so that the incumbents take this into account when investing in R&D along the existing technology.

4.2 Basic Model

4.2.1 Consumer Choice over Quality

Following Sutton (1998), we consider the following utility function shared by all consumers:

$$U = \left(\sum_{i=1}^{N} u_i x_i \right)^{\alpha} z^{1-\alpha},$$

where x_i and u_i are the quantity and quality of 'quality good' provided by the firm i and z is the quantity of the 'outside good'. A consumer's decision problem is to maximize her utility subject to a budget constraint. This Cobb-Douglas form utility function has the following feature: The consumer spends fraction α of her income on the quality good, and fraction $(1 - \alpha)$ on the 'outside good'. The total expenditure on the quality good is therefore independent of qualities and prices and equals a fraction α of the total consumer income. We notate the total expenditure on quality good by S.

Since each consumer chooses a quality that maximizes u_i/p_i, it follows that all goods that command positive sales at equilibrium must have equal quality-adjusted price, i.e. $p_i/u_i = \lambda$, for all i. The constant λ, and so the vector of equilibrium prices, is determined by $\sum p_i x_i = \sum \lambda u_i x_i = S$, so that $\lambda = S/(\sum u_i x_i)$.

4.2.2 A Reduced-form Profit Function

We assume all the 'quality goods' are produced at some constant marginal cost $c > 0$. In the Cournot-Nash equilibrium in quantities, the profit flow earned by firm i among the N firms offering these goods is[2]

$$\Pi\left(u_i \mid (u_{-i})\right) = S\left(1 - \frac{N-1}{\sum_{j=1}^{N} \frac{u_i}{u_j}}\right)^2 .$$

A nice feature of the above reduced-form profit function is that the term $\left(1 - \frac{N-1}{\sum_{j=1}^{N} \frac{u_i}{u_j}}\right)$ represents firm i's market share as well as its mark-up ratio. Therefore firm i will only be active if $1 - \frac{N-1}{\sum_{j=1}^{N} \frac{u_i}{u_j}} > 0$, otherwise it exits the market.

4.2.3 The Game

The quality good industry is organized as a quality competition game which has two parts as follow.

Part one[3]

- *Stage 1*: A sufficiently large number of firms simultaneously choose whether to enter a market[4]. The number is large enough in the sense that some will choose not to enter in equilibrium.
- *Stage 2*: After observing the number of entry N, the firms which have already entered simultaneously choose their quality levels and implement the quality

[2]For the derivation of this reduced form profit function, see Appendix 15.1 of Sutton (1998).

[3]Part one of the model is exactly the Cournot model with quality choice by Sutton (1991). This part of game is played entirely independent of what will happen in the second part because they are unexpected at all.

[4]We assume that all entrants at this stage can commit to invest at stage 2.

by investing in R&D. The R&D costs are determined by the following fixed cost function:

$$F(u) = (\mu u)^{\beta},$$

where u is the quality level and μ is the parameter of the R&D technology. The overall payoff function for firm i then is:

$$Y(u_i \mid (u_{-i})) = \Pi(u_i \mid (u_{-i})) - F(u_i) = S\left(1 - \frac{N-1}{\sum_{j=1}^{N} \frac{u_i}{u_j}}\right)^2 - (\mu u_i)^{\beta}.$$

For the sake of simplicity we confine attention to the case where the outcome of Part one is a symmetric equilibrium such that all firms have equal quality level $\overline{u}$.

Part two

We assume that after the fixed and sunk costs have been incurred, an unexpected technological shock occurs.

- *Stage 3*: A new R&D technology indexed by μ' ($\mu' < \mu$) becomes available. One potential new entrant chooses whether to enter the market. If she enters then she also chooses her quality level, which without loss of generality can be notated by $k\overline{u}$. The new entrant's payoff function is:

$$Y(k\overline{u} \mid (\overline{u})) = S\left(1 - \frac{N}{Nk+1}\right)^2 - (\mu' k\overline{u})^{\beta}.$$

Here we have implicitly assumed that the incumbents are not allowed to respond to a new entry by changing their quality levels. Later in Appendix C we will show that relaxing this assumption will not change the result: to make no change in their quality levels after the new entry is the equilibrium strategy for the incumbents.

4.2.4 Equilibrium

The equilibrium concept used throughout this chapter is pure strategy sub-game perfect equilibrium. The analysis is made much easier when we only focus on the symmetric outcomes. We assume that in stage 2, all firms choose quality level $\overline{u}$ except firm i chooses quality level u. Then firm i's maximization problem is:

$$\max_{u} Y_i\left(u \mid \bar{u}\right) = \max_{u}\left(S\left(1 - \frac{N-1}{\frac{u}{\bar{u}}\left(N-1\right)+1}\right)^2 - u^{\beta}\right), \tag{4.1}$$

where we normalize $\mu = 1$ for the incumbents.

The first order condition for any symmetric outcome is:

$$\left.\frac{\partial Y_i}{\partial u}\right|_{u=\overline{u}} = 2S\left(1 - \frac{N-1}{N}\right)\frac{\left(N-1\right)^2}{N^2\overline{u}} - \beta\overline{u}^{\beta-1} = 0. \tag{4.2}$$

Since each firm has a non-entry option, the following non-loss-making condition should be satisfied:

$$\left.Y_i\right|_{u=\overline{u}} = S\left(1 - \frac{N-1}{N}\right)^2 - \overline{u}^{\beta} \geq 0. \tag{4.3}$$

Combining the above two conditions we can derive the following inequality which determines the number firms in Part one:

$$N + \frac{1}{N} - 2 \leq \frac{\beta}{2}. \tag{4.4}$$

At equilibrium N is the maximum integer[5] that satisfies the above inequality.

Remark 15 *The driving force behind market concentration in this model is the presence of increasing returns to scale, in the following sense: the specifications of the utility function: $U = \left(\sum_{i=1}^{N} u_i x_i\right)^{\alpha} z^{1-\alpha}$, and the fixed cost function: $h_i = u_i^{\beta}$, $(\beta > 0)$[6] and the assumption of constant marginal cost imply that the effective long-term production function is:*

$$y_i = u_i x_i = h_i^{\frac{1}{\beta}} f\left(k_i, l_i\right)$$

[5]In this study, we emphasize the integer effect, which allows positive net payoff to exist even when there is free entry.

[6]Here we notate the R&D effort by h.

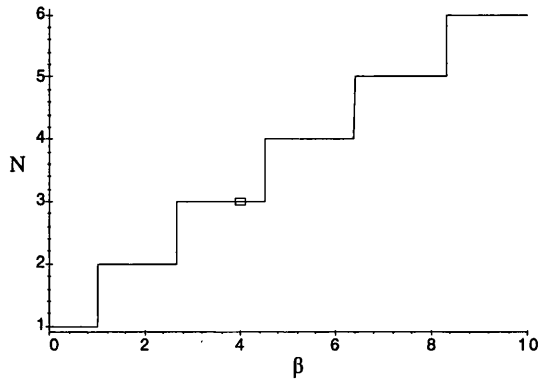


Figure 4.1: $N = n(\beta)$

where $f(\cdot,\cdot)$ is the short-term production function and is homogenous of degree 1 in (k_i, l_i) , so the output y_i is homogeneous of degree $(1 + 1/\beta)$ in overall input (h_i, k_i, l_i), which indicates increasing returns to scale. Furthermore, the value of parameter β determines the value of N in the following way:

$$\begin{array}{lllllllll} 0 < \beta < 1 & \Leftrightarrow & 2 < 1 + 1/\beta & \Leftrightarrow & N = 1 & \Leftrightarrow & \textit{natural monopoly} \\ 1 \leq \beta < 8/3 & \Leftrightarrow & 1 + 3/8 < 1 + 1/\beta \leq 2 & \Leftrightarrow & N = 2 & \Leftrightarrow & \textit{natural duopoly} \\ 8/3 \leq \beta & \Leftrightarrow & 1 + 1/\beta \leq 1 + 3/8 & \Leftrightarrow & N \geq 3 & \Leftrightarrow & \textit{natural oligopoly} \end{array}$$

At equilibrium N is a function of β, i.e., $N = n(\beta)$, and N is the largest integer such that

$$N \leq 1 + \frac{1}{4}\beta + \frac{1}{4}\sqrt{8\beta + \beta^2}. \tag{4.5}$$

The function $n(\beta)$ is not continuous, therefore it jumps at critical values of β. This can be visualized by Figure 4.1. The small box in the figure represents the specific example of $\beta = 4$. Since we are interested in the impacts of increases in firm numbers we will by-pass the monopoly and duopoly cases and focus on the case where that $\beta > \frac{8}{3}$ is satisfied and consequently $N \geq 3$.

Solving eq. (4.2) for $\bar{u}$ as a function of S, N and β:

$$\bar{u} = \left(\frac{2S\,(N-1)^2}{\beta N^3} \right)^{\frac{1}{\beta}} \equiv v\,(S, N, \beta)\,. \tag{4.6}$$

So far, what we have got are all standard outcomes from the Cournot model with 'perceived quality'. The novel content of the analysis begins from part two.

In stage 3, a single entrant maximizes her overall payoff:

$$Y \equiv S \left(1 - \frac{N}{kN+1} \right)^2 - (\mu' k \bar{u})^{\beta}\,, \tag{4.7}$$

the first order condition of which is:

$$\frac{\partial Y}{\partial k} = 2S \left(1 - \frac{N}{kN+1} \right) \frac{N^2}{(kN+1)^2} - \frac{\beta}{k} (\mu' k \bar{u})^{\beta} = 0. \tag{4.8}$$

There are two unknowns, k and μ', in equation (4.8). To solve the model, we need to introduce a second equation linking these parameters.

It is worth to mention that our main concern is under what circumstances a profit maximizing potential new entrant will enter in equilibrium. We can imagine that there exists some threshold value for the parameter μ' such that below which, an entry will occur and above which any potential entry will be deterred in equilibrium. If this is the case, then we want to pin down the threshold value of μ'. Apparently in such a threshold case, the new entrant should be indifferent between entering or not, i.e. the entrant is exactly at the break-even point. So the following break-even condition holds:

$$S \left(1 - \frac{N}{kN+1} \right)^2 = (\mu' k \bar{u})^{\beta}\,. \tag{4.9}$$

Combining equations (4.8) and (4.9) to eliminate $(\mu' k \bar{u})^{\beta}$ and S results in the following equation with k as the only unknown:

$$\beta N^2 k^2 - Nk\,(-2\beta + \beta N + 2N) + \beta - \beta N = 0. \tag{4.10}$$

The unique positive solution to the above equation is:

$$k = \frac{-2\beta N + \beta N^2 + 2N^2 + \left(-8\beta N^3 + \beta^2 N^4 + 4\beta N^4 + 4N^4\right)^{\frac{1}{2}}}{2\beta N^2} \equiv K(N, \beta). \tag{4.11}$$

Manipulating eq. (4.8) to solve for μ' leads to

$$\mu' = \frac{1}{k\bar{u}} \left(-\frac{2SN^3k^2 + 2SN^2k - 2SN^3k}{-\beta k^3 N^3 - 3\beta k^2 N^2 - 3\beta k N - \beta} \right)^{\frac{1}{\beta}}. \tag{4.12}$$

Inserting eq. (4.6) into eq. (4.12) and eliminating $\bar{u}$ and S, μ' turns out to be a function of k, N and β as follows

$$\begin{aligned} \mu' &= \frac{1}{k} \left(\frac{N^5 k (Nk + 1 - N)}{(k^3N^3 + 3k^2N^2 + 3Nk + 1)(N^2 - 2N + 1)} \right)^{\frac{1}{\beta}} \\ &\equiv \Xi(k, N, \beta). \end{aligned} \tag{4.13}$$

By $N = n(\beta)$ and eq. (4.11), eq. (4.13) turns out to be a function of β as follows

$$\begin{aligned} \mu' &= \Xi(K(N, \beta), N, \beta) \\ &= \Xi(K(n(\beta), \beta), n(\beta), \beta) \\ &\equiv \Phi(\beta). \end{aligned} \tag{4.14}$$

Figure 4.2 illustrates $\Phi(\beta)$. We see that it is not continuous: it jumps at critical values of β because N jumps. Again, the small box represents the specific example of $\beta = 4$.

In Appendix C we show that the sufficient condition for the maximum has been satisfied by checking on the related second order condition. In Appendix C we also show that even if the incumbents are allowed to respond to the entrant's quality choice, the equilibrium outcome is exactly the same as if the incumbents were not allowed to respond.

Proposition 16 *For the basic model, there exists a threshold value of μ', i.e. μ_{th}, such that:*

If $\mu' > \mu_{th}$, then no new entry occurs in equilibrium;

If $\mu' < \mu_{th}$, then new entry occurs in equilibrium.

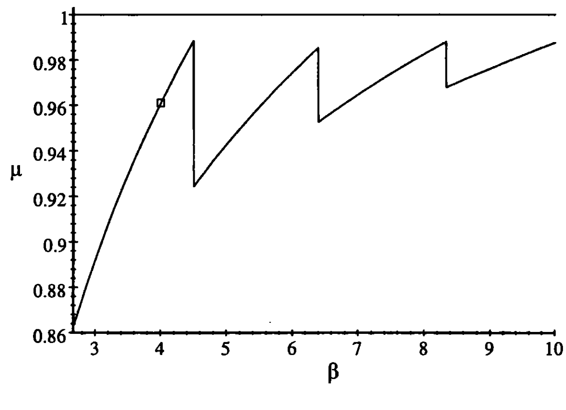


Figure 4.2: $\mu_{th} = \Phi(\beta)$

Example 17 *When* $\beta = 4$, $\overline{u}^4 = \frac{2}{27}S$ *(or* $\overline{u} = 0.5217S^{\frac{1}{4}}$*),*
$k = \frac{5}{12} + \frac{1}{12}\sqrt{57} = 1.0458$, $\mu_{th} = 0.9610$.

This proposition claims that the vulnerability of the existing market structure to a technological breakthrough is conditional on the relative merits of the new and old technologies. If the new technology doesn't enjoy a significant advantage over the old technology, the potential new entry will be deterred and the existing market structure will not change. On the other hand, if the new technology possesses a significant advantage over the old, it will be adopted by some potential entrant(s). The existing market structure will change as new entry occurs.

Corollary 18 *An anticipated technological breakthrough which will lead to* $\mu' > \mu_{th}$ *(*μ_{th} *as defined in Proposition 16) has no effect on the incumbents' competitive behavior.*

The intuition for the above corollary is as follows: consider the actions chosen by incumbents in the original game. For this configuration of parameter values, these actions constitute an optimal reply to the potential entrant's decision to not enter. Moreover, given these actions, it is optimal for the potential entrant not to enter.

4.2.5 Endogenizing the Number of Entrants

We now relax the assumption that the number of new entrants is fixed at unity. Part two of the game is now altered as follows:

Part two

- *Stage 3:* A new technology μ', $(\mu' < 1)$, becomes available. A (large) number of potential entrants choose whether to seize the opportunity to enter the market.

- *Stage 4:* After observing the number of new entry N_1, the new entrants simultaneously choose their quality levels and invest in R&D to implement the quality levels.

Again we will only focus on symmetric outcomes, i.e., we assume all the incumbents have the same quality level $\bar{u}$ and all the new entrants have the same quality level $\bar{k}\bar{u}$. Consequently, an incumbent's payoff is:

$$Y_i\left(\bar{u} \mid (\bar{u}), (\bar{k}\bar{u})\right) = S\left(1 - \frac{N + N_1 - 1}{N + N_1\frac{1}{\bar{k}}}\right)^2 - \bar{u}^\beta.$$

Conditional on all the others playing their equilibrium strategies in the symmetric equilibrium, the new entrant j whose quality level is $k\bar{u}$ faces the following maximization problem:

$$\max_k Y_j\left(k\bar{u} \mid (\bar{u}), (\bar{k}\bar{u})\right) \equiv \max_k \left(S\left(1 - \frac{N + N_1 - 1}{Nk + (N_1 - 1)\frac{k}{\bar{k}} + 1}\right)^2 - (\mu' k\bar{u})^\beta\right), \tag{4.15}$$

the first order condition of which for a symmetric outcome is:

$$\left.\frac{\partial Y_j}{\partial k}\right|_{k=\bar{k}} = 2S\left(1 - \frac{N + N_1 - 1}{\bar{k}N + N_1}\right)\frac{(N + N_1 - 1)\left(\bar{k}N + N_1 - 1\right)}{\left(\bar{k}N + N_1\right)^2\bar{k}} - \frac{\beta}{\bar{k}}\left(\mu'\bar{k}\bar{u}\right)^\beta = 0. \tag{4.16}$$

Again, the following non-loss-making condition must be satisfied:

$$Y_j \mid_{k=\bar{k}} = S\left(1 - \frac{N + N_1 - 1}{N\bar{k} + N_1}\right)^2 - \left(\mu'\bar{k}\bar{u}\right)^\beta \geq 0.$$

In a symmetric equilibrium, for a given μ', N_1 should be the maximum integer that satisfies the above conditions. Otherwise one more potential entrant could profitably deviate by entering.

In the threshold cases, the new entrants just break even, therefore the following break-even condition becomes useful:

$$Y_j \mid_{k=\bar{k}} = S\left(1 - \frac{N + N_1 - 1}{N\bar{k} + N_1}\right)^2 - \left(\mu'\bar{k}\bar{u}\right)^\beta = 0. \tag{4.17}$$

Combining equations (4.16) and (4.17) to eliminate $\left(\mu'\bar{k}\bar{u}\right)^\beta$ and S results in the following equation:

$$\beta N^2\bar{k}^2 - \left(N_1\left(2-\beta\right) + \left(N-1\right)\left(2+\beta\right)\right)N\bar{k} - \left(N_1\left(2N_1 - 4 + 2N + N\beta - \beta\right) + 2 - 2N\right) = 0 \tag{4.18}$$

the unique positive solution of which is

$$\bar{k} = \kappa_1\left(N_1, N, \beta\right), \tag{4.19}$$

where $\kappa_1\left(N_1, N, \beta\right)$
$\equiv \frac{(N_1(2-\beta)+(N-1)(2+\beta))N+\sqrt{((N_1(2-\beta)+(N-1)(2+\beta))N)^2+4\beta N^2(N_1(2N_1-4+2N+N\beta-\beta)+2-2N)}}{2\beta N^2}$

Knowing that $N = n\left(\beta\right)$, equation (4.19) implies:

$$\bar{k} = \kappa_1\left(N_1, n\left(\beta\right), \beta\right) \equiv \kappa\left(N_1, \beta\right). \tag{4.20}$$

Equations (4.6) and (4.17) jointly imply that threshold value of μ' is

$$\mu_{th} = \chi_1\left(N_1, N, \bar{k}, \beta\right),$$

where $\chi_1\left(N_1, N, \bar{k}, \beta\right) \equiv \left(\left(1 - \frac{N+N_1-1}{N\bar{k}+N_1}\right)^2 \frac{\beta N^3}{2N^2-4N+2}\right)^{\frac{1}{\beta}} \frac{1}{\bar{k}}$.

Given that $N = n\left(\beta\right)$, μ_{th} then is a known function of N_1 and β, i.e.,

$$\mu_{th} = \chi_1\left(N_1, n\left(\beta\right), \kappa\left(N_1, \beta\right), \beta\right) \equiv \chi\left(N_1, \beta\right). \tag{4.21}$$

A useful way to gain more insight into the relation between the technological parameter μ', the market structure N_1 and the quality level of the new entrant(s) k is to work out a concrete example.

Example 19 *Set* $\beta = 4$, *then* $N = 3$, $\bar{k} = \frac{-N_1+6+\sqrt{9N_1^2+28N_1+20}}{12}$ *and* $\mu_{th} = \left(1 - \frac{2+N_1}{3\bar{k}+N_1}\right)^{\frac{1}{2}} \left(\frac{27}{2}\right)^{\frac{1}{4}} \frac{1}{\bar{k}}$. *Consequently,*
when $N_1 = 1$, $\bar{k} = 1.0458$ *and* $\mu_{th} = 0.9610$*;*
when $N_1 = 2$, $\bar{k} = 1.2153$ *and* $\mu_{th} = 0.8516$*;*
when $N_1 = 3$, $\bar{k} = 1.3835$ *and* $\mu_{th} = 0.7598$*.*

In the above example, $\mu_{th} = 0.9610$, 0.8516 and 0.7598 are the three threshold values of μ' at which the market structure changes. For a given value of μ' we can calculate the corresponding market structure by looking at which interval μ' falls into.

After determining the equilibrium market structure, we then can calculate the associated quality level of the new entrant(s) $\bar{k}$ by manipulating equations (4.6) and (4.16), which determine $\bar{k}$ as a function of μ' via the following equation

$$\frac{N^3\left(\bar{k}N - N + 1\right)\left(N + N_1 - 1\right)\left(\bar{k}N + N_1 - 1\right)}{\left(N-1\right)^2\left(\bar{k}N + N_1\right)^3} = \left(\mu'\bar{k}\right)^\beta.$$

In the example: $\beta = 4$, $N = 3$, the above equation takes the form of

$$\frac{27}{4}\left(3\bar{k} - 2\right)\left(N_1 + 2\right)\frac{3\bar{k} + N_1 - 1}{\left(3\bar{k} + N_1\right)^3} = \mu'^4\bar{k}^4.$$

Figure 4.3 illustrates the relation between μ' and $\bar{k}$, conditional on N_1. We see in the graph that for a given pair of N and N_1, when μ' decreases, $\bar{k}$ increases, i.e., the relation is monotone. We can see that when μ' reaches the next threshold value, N_1 will increase by 1.

Our calculation leads to the following results:

- When $\mu' > 0.9610$, there is no new entry;

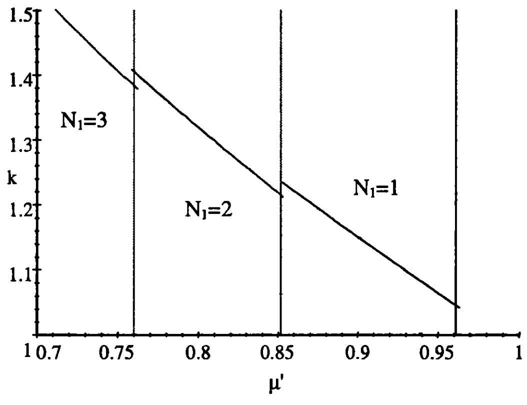


Figure 4.3:

- When $0.9610 \leq \mu' < 0.8516$, there is one new entrant, and as μ' decreases from 0.9610 towards 0.8516, $\bar{k}$ increases from 1.0458 towards 1.2365.
- When $0.8516 \leq \mu' < 0.7598$, there are two new entrants, and as μ' decreases from 0.8516 towards 0.7598, $\bar{k}$ increases from 1.2153 towards 1.4051.
- When $\mu' \geq 0.7598$, the number of new entrants becomes three, and this is the highest possible number of new entrants in any symmetric equilibria in this example.

It is worth to notice that when μ' is sufficiently small, the number of the new entrants can reach its highest possible level and the quality level of the new entrants can be so high that competition between the new entrants drives the quality-adjusted prices below the incumbents' marginal costs, so that the incumbents find it optimal to exit to avoid further losses. To characterize this situation, we first look at the threshold value of $\bar{k}$ for an incumbent to exit:

$$\left(1 - \frac{N + N_1 - 1}{N + \frac{N_1}{k}}\right) = 0,$$

where the left hand side is the market share as well as the mark-up ratio of an incumbent. The equation implies

$$\bar{k} = \frac{N_1}{N_1 - 1}. \tag{4.22}$$

In the current example, the largest possible number for N_1 is 3, so if $\bar{k} = \frac{3}{3-1} = 1.5$ then the incumbents' total market share becomes zero, and accordingly, $\mu_{th} = 0.7119$. When $\mu' \leq 0.7119$ and $\bar{k} \geq 1.5$, we know that all the incumbents will exit the market; but we want to know whether the incumbents have already had any impact on the competitive behavior of the new entrants.

The situation in which the incumbents all exit after the technological breakthrough and so make no impact on the competitive behavior of the new entrants will be referred to in what follows as a drastic innovation (or drastic replacement). In the case of a drastic innovation, a new entrant's maximization problem is replaced by the following problem:

$$\max_k Y_j\left(k\bar{u} \mid \left(\bar{k}\bar{u}\right)\right) \equiv \max_k \left(S\left(1 - \frac{N_1 - 1}{\left(N_1 - 1\right)\frac{k}{\bar{k}} + 1}\right)^2 - \left(\mu' k\bar{u}\right)^{\beta}\right),$$

implying that the incumbents can be ignored when a new entrant chooses her quality level. The symmetric first order condition is

$$2S\frac{\left(N_1 - 1\right)^2}{N_1^3} = \beta\left(\mu'\bar{k}\bar{u}\right)^{\beta} \tag{4.23}$$

and the non-loss-making condition is

$$\frac{S}{N_1^2} \geq \left(\mu' k\bar{u}\right)^{\beta}. \tag{4.24}$$

The above two conditions imply that

$$N_1 - 2 + \frac{1}{N_1} \leq \frac{\beta}{2},$$

which is the same as (4.4), therefore $N_1 = N = n(\beta)$.

Then (4.23) jointly with eq. (4.6) suggests

$$\left(\bar{k}\mu'\overline{u}\right)^{\beta} = 2S\frac{(N_1-1)^2}{\beta N_1^3} = 2S\frac{(N-1)^2}{\beta N^3} = (\overline{u})^{\beta},$$

which implies

$$\bar{k} = \frac{1}{\mu'}.$$

So the sufficient and necessary condition for a drastic replacement is

$$\begin{cases} N_1 = N = n(\beta) \\ \bar{k} = \frac{1}{\mu'} \\ \bar{k} \geq \frac{N_1}{N-1} \end{cases},$$

which is equivalent to $\mu' \leq \frac{N-1}{N}$. Then the corresponding threshold value of μ' is

$$\mu_{th} = \frac{N-1}{N}. \tag{4.25}$$

Proposition 20 *When $\mu' \leq \mu_{th} = \frac{N-1}{N}$, there are N new entrants, $\bar{k} \geq \frac{N}{N-1}$, all the incumbents exit and they have had no impact on the competitive behavior of the new entrants, i.e., the innovation is drastic.*

In the current example: $\beta = 4$, $N = 3$ then when $\mu' \leq \frac{3-1}{3} = \frac{2}{3} = 0.6667$, $N_1 = 3$, $\bar{k} \geq 1.5$ and all the incumbents exit the market. Our previous calculation has already shown that when $\mu' \leq 0.7119$, $N_1 = 3$, $\bar{k} \geq 1.5$ and all the incumbents will exit the market. So we want to know what is actually going on when $0.6667 < \mu' \leq 0.7119$. It turns out that

- when $0.6667 < \mu' \leq 0.7119$ there are three new entrants entering, choosing the quality level exactly such that $\bar{k} = 1.5$ and all the incumbents exit.

In the case stated above, although the incumbents exit the market, they have already had impacts on the quality level of the new entrants, which would be lower otherwise. Note that in this case, the first order condition (4.16) does not hold

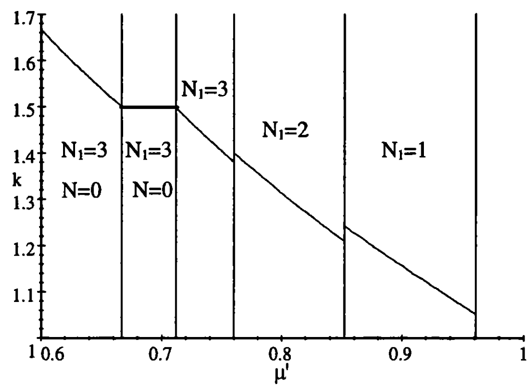


Figure 4.4:

because at $\bar{k} = \frac{N}{N-1}$ the reduced-form profit function is not smooth in k and its first order derivative is not defined; instead each new entrant has a corner solution to her maximization problem.

Figure 4.4 shows how the market structure and the new entrants's quality are related to the parameter of the new technology μ'. Table 4.1 summarizes all the threshold values of μ', corresponding market structure (N, N_1) and the relative quality level of the entrants $\bar{k}$ calculated for the example case of $\beta = 4$.

From the figure as well as the table we can see a weak monotone relationship between μ' and equilibrium number of new entrants N_1. The greater the gap between the new technology and the old, the closer we come to the outcome in which the market comes to be dominated by new entrants.

The last five columns of the table show the market shares and net payoffs of the entrants and the incumbents at each threshold value of μ'. From the table we can see that a radical technological breakthrough may reshape the market structure and cause reallocation of net payoffs, i.e. oligopoly rents. The size of the gap between the new technology and the old determines the extent of the 'creative destruction', notably, the there is a weak monotone relation between μ' and the combined market share of new entrants.

μ_{th}	N	N_1	$\overline{k}$	all ents' share	1 ent's share	1 ent's payoff	1 inc's share	1 inc's payoff
0.9610	3	0	N/A	0	0	0	0.3333	0.0370S
0.9610	3	1	1.0458	0.2749	0.2749	0	0.2417	-0.0157S
0.8516	3	1	1.2365	0.3630	0.3630	0.0407S	0.2123	-0.0290S
0.8516	3	2	1.2153	0.5830	0.2915	0	0.1390	-0.0548S
0.7598	3	2	1.4051	0.7128	0.3564	0.0308S	0.0957	-0.0649S
0.7598	3	3	1.3835	0.9021	0.3007	0	0.0326	-0.0730S
0.7119	3	3	1.5	1	0.3333	0.0148S	0	-0.0741S
0.7119	0	3	1.5	1	0.3333	0.0148S	0	-0.0741S
0.6667	0	3	1.5	1	0.3333	0.0370S	0	-0.0741S

Table 4.1: The symmetric equilibria after the unanticipated technological breakthrough

Through the above numerical example we have illustrated the relation between market structure, the new entrants' quality jump and the underlying parameter describing the new technology. As to more general claims, in addition to Propositions 16 and 20 we also have one further result:

Corollary 21 *There exist two bounds of μ', i.e., $\mu_n < \mu_1 < 1$, such that when $\mu' > \mu_1$ no new entrant will enter; and when $\mu' < \mu_n$, $n(\beta)$ new entrants will enter and all incumbents will exit.*

Figure 4.5 visualizes the existence of the two bounds of μ' in regard of market structure.

4.3 Anticipated Future Technological Breakthroughs

In this section we examine the case of an anticipated future technological breakthrough, with a view to analyzing how the expectation of this breakthrough by the

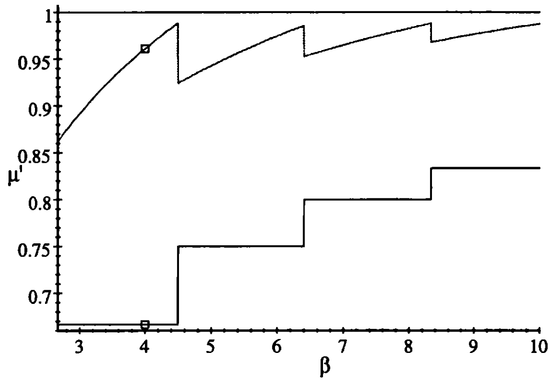


Figure 4.5: The two bounds

incumbents affects their behavior, and so impacts on outcomes before and after the technological breakthrough. For the sake of simplicity we study a two-period model, associated with the following 4-stage game:

Period one

- *Stage 1*: A sufficiently large number of firms simultaneously choose whether to enter a market. The market will exist for 2 periods. In each period, the market size is S and the discount rate is 1. It is anticipated that at the end of period one, there will be a technological breakthrough, i.e. a new technology will be available.

- *Stage 2*: After observing the number of entry N, the firms which have already entered simultaneously choose their quality levels and implement the quality levels by investing in R&D. The fixed cost function is indexed by μ ($\mu = 1$).

Period two

- *Stage 3*: A new technology indexed by μ' ($\mu' < \mu$) becomes available. A (large) number of potential entrants choose whether to enter the market.

- *Stage 4*: Having observed the number of new entrants N_1, each of the new entrants simultaneously chooses its quality level and incurs the corresponding cost of R&D. The fixed cost function is indexed by μ'.

To characterize the symmetric equilibrium, let us assume all incumbents choose quality level $\overline{u}$, except for incumbent i who chooses u. Then firm i maximizes its total payoff in the two periods:

$$\begin{aligned}
&\max_u Y_i \left(u \mid (\bar{u}), (\bar{k}\bar{u})\right) \\
\equiv\; &\max_u \left(S\left(1 - \frac{N-1}{(N-1)\frac{u}{\bar{u}}+1}\right)^2 + S\left(1 - \frac{N+N_1-1}{(N-1)\frac{u}{\bar{u}} + N_1\frac{u}{\bar{k}\bar{u}}+1}\right)^2 - u^\beta \right).
\end{aligned}$$

For symmetric equilibrium, the first order condition is:

$$2S\frac{(N-1)^2}{N^3} + 2S\frac{\left(\frac{N_1}{\bar{k}} - N_1 + 1\right)(N+N_1-1)\left(N-1+\frac{N_1}{\bar{k}}\right)}{\left(N+\frac{N_1}{\bar{k}}\right)^3} = \beta\bar{u}^\beta, \tag{4.26}$$

and the non-loss-making condition is:

$$\frac{S}{N^2} + S\left(1 - \frac{N+N_1-1}{N+\frac{N_1}{\bar{k}}}\right)^2 - \bar{u}^\beta \geq 0. \tag{4.27}$$

In the second period, we assume all the new entrants choose quality level $\bar{k}\overline{u}$, except for new entrant j who chooses $k\overline{u}$. Then firm j's maximization problem becomes:

$$\begin{aligned}
&\max_k Y_j \left(k\bar{u} \mid (\bar{u}), (\bar{k}\bar{u})\right) \\
\equiv\; &\max_u \left(S\left(1 - \frac{N+N_1-1}{Nk+(N_1-1)\frac{k}{\bar{k}}+1}\right)^2 - (\mu' k\bar{u})^\beta \right),
\end{aligned}$$

which is exactly the same as problem (4.15). Therefore equations (4.16) to (4.19) also apply to the characterization of the symmetric outcome in this setting, and we can determine $\bar{k}$ for the threshold cases.

Furthermore, eq. (4.17) and (4.26) together imply that

$$\mu_{th} = \left(\frac{\beta\left(1 - \frac{N+N_1-1}{N\bar{k}+N_1}\right)^2}{2\left(\frac{(N-1)^2}{N^3} + \frac{\left(\frac{N_1}{\bar{k}}-N_1+1\right)(N+N_1-1)\left(N-1+\frac{N_1}{\bar{k}}\right)}{\left(N+\frac{N_1}{\bar{k}}\right)^3}\right)} \right)^{\frac{1}{\beta}} \frac{1}{\bar{k}}, \tag{4.28}$$

which determines the threshold values of μ'.

Finally, eq. (4.26) implies

$$\bar{u} = \left(\frac{2}{\beta}\left(\frac{(N-1)^2}{N^3} + \frac{\left(\frac{N_1}{k} - N_1 + 1\right)(N + N_1 - 1)\left(N - 1 + \frac{N_1}{k}\right)}{\left(N + \frac{N_1}{k}\right)^3}\right)\right)^{\frac{1}{\beta}} S^{\frac{1}{\beta}}, \quad (4.29)$$

which determines the corresponding quality levels of the incumbents $\bar{u}$.

We consider a numerical example in order to provide some insight into the relation between the technological parameter μ', market structure (N, N_1) and the quality levels of the new entrant(s) k and of incumbents $\bar{u}$.

Example 22 *Set $\beta = 4$, then $N = 3$, $\bar{k} = \frac{-N_1+6+\sqrt{9N_1^2+28N_1+20}}{12}$,*

$$\bar{u} = \left(\frac{1}{2}\left(\frac{4}{27} + \frac{\left(\frac{N_1}{k}-N_1+1\right)(2+N_1)\left(2+\frac{N_1}{k}\right)}{\left(3+\frac{N_1}{k}\right)^3}\right)\right)^{\frac{1}{4}} S^{\frac{1}{4}} \text{ and } \mu_{th} = \left(\frac{2\left(1-\frac{3+N_1-1}{3k+N_1}\right)^2}{\frac{4}{27}+\frac{\left(\frac{N_1}{k}-N_1+1\right)(2+N_1)\left(2+\frac{N_1}{k}\right)}{\left(3+\frac{N_1}{k}\right)^3}}\right)^{\frac{1}{4}} \frac{1}{k}.$$

Consequently,
when $N_1 = 1$, $\bar{k} = \frac{5+\sqrt{57}}{12} = 1.0458$, $\bar{u} = 0.6145S^{\frac{1}{4}}$ and $\mu_{th} = 0.8159$;
when $N_1 = 2$, $\bar{k} = \frac{1+\sqrt{7}}{3} = 1.2153$, $\bar{u} = 0.5898S^{\frac{1}{4}}$ and $\mu_{th} = 0.7532$;
when $N_1 = 3$, $\bar{k} = \frac{3+\sqrt{185}}{12} = 1.3835$, $\bar{u} = 0.5428S^{\frac{1}{4}}$ and $\mu_{th} = 0.7303$.

As we can see, $\mu_{th} = 0.8159$, 0.7532 and 0.7303 are the three threshold values of μ' at which market structure changes. For a given value of μ' we can determine the corresponding market structure by finding which interval μ' lies in. Then we can calculate the quality level of the new entrant(s) $\bar{k}$ using equations (4.16) and (4.26), which determine $\bar{k}$ as a function of μ' via the following equation:

$$\frac{\left(\bar{k}N - N + 1\right)(N + N_1 - 1)\left(\bar{k}N + N_1 - 1\right)}{\left(\frac{(N-1)^2}{N^3} + \frac{\left(\frac{N_1}{k}-N_1+1\right)(N+N_1-1)\left(N-1+\frac{N_1}{k}\right)}{\left(N+\frac{N_1}{k}\right)^3}\right)\left(\bar{k}N + N_1\right)^3} = \left(\mu'\bar{k}\right)^{\beta}.$$

In the example: $\beta = 4$, $N = 3$, the above equation takes the form of

$$\frac{\left(3\bar{k} - 2\right)(2 + N_1)\left(3\bar{k} + N_1 - 1\right)}{\left(\frac{4}{27} + \frac{\left(\frac{N_1}{k}-N_1+1\right)(2+N_1)\left(2+\frac{N_1}{k}\right)}{\left(3+\frac{N_1}{k}\right)^3}\right)\left(3\bar{k} + N_1\right)^3} = \left(\mu'\bar{k}\right)^4.$$

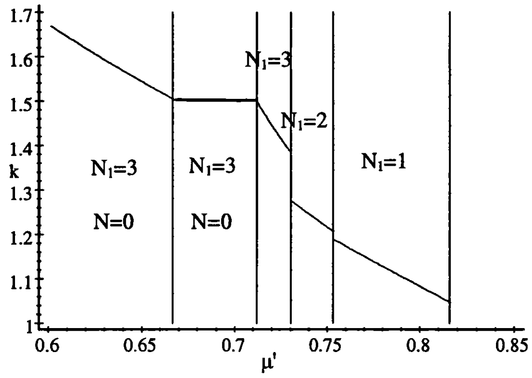


Figure 4.6:

In particular, when $N_1 = 3$, if $\mu_{th} = 0.7119$ then $\overline{k} = 1.5$ (the incumbents' exit threshold) and the incumbents' market share is zero.

When μ' is sufficiently small, the incumbents can foresee that they will exit in the second period. So their maximization problem becomes

$$\max_u Y_i\left(u \mid \bar{u}\right) = \max_u \left(S\left(1 - \frac{N-1}{\frac{u}{\bar{u}}\left(N-1\right)+1}\right)^2 - u^\beta \right)$$

which is the same as problem (4.1). As in problem (4.1), in this case $\left(\overline{u}\right)^\beta = 2S\frac{(N-1)^2}{\beta N^3}$ and $N = n\left(\beta\right)$. Moreover, all the arguments which lead to Proposition 20 still hold in this case, so that the claim in Proposition 20 is applicable, i.e. when $\mu' \leq \mu_{th} = \frac{N-1}{N}$, there are N new entrants, $\bar{k} \geq \frac{N}{N-1}$, all the incumbents exit and they have had no impact on the competitive behavior of the new entrants, i.e., the innovation is drastic. In the current example, the threshold value of μ' for a drastic innovation is $\mu_{th} = \frac{3-1}{3} = \frac{2}{3}$.

Figure 4.6 illustrates the following results:

- When $\mu' > 0.8195$, there is no new entry;
- When $0.7532 \leq \mu' < 0.8195$, there is one new entrant, and as μ' decreases from 0.8195 towards 0.7532, $\bar{k}$ increases from 1.0458 towards 1.1874.

μ_{th}	N	N_1	$\overline{u}$ ($\times S^{\frac{1}{4}}$)	$\overline{k}$	all ents' share	1 ent's share	1 ent's payoff	1 inc's share	1 inc's payoff
0.8159	3	0	0.6204	N/A	0	0	0	0.3333	0.0740S
0.8159	3	1	0.6145	1.0458	0.2749	0.2749	0	0.2417	0.0270S
0.7532	3	1	0.6088	1.1874	0.3424	0.3424	0.0294S	0.2192	0.0218S
0.7532	3	2	0.5898	1.2153	0.5830	0.2915	0	0.1390	0.0094S
0.7303	3	2	0.5836	1.2840	0.6330	0.3165	0.0105S	0.1224	0.0101S
0.7303	3	3	0.5428	1.3835	0.9021	0.3007	0	0.0326	0.0254S
0.7119	3	3	0.5217	1.5	1	0.3333	0.0148S	0	0.0370S
0.7119	0	3	0.5217	1.5	1	0.3333	0.0148S	0	0.0370S
0.6667	0	3	0.5217	1.5	1	0.3333	0.0370S	0	0.0370S

Table 4.2: The symmetric equilibria in the model with anticipated future technological breakthrough (All the figures are calculated only for the second period except that those in the last column are calculated for the sum of both periods.)

- When $0.7303 \leq \mu' < 0.7532$, there are two new entrants, and as μ' decreases from 0.7532 towards 0.7303, $\bar{k}$ increases from 1.2153 towards 1.2840.

- $0.7119 \leq \mu' < 0.7303$, there are three new entrants, and as μ' decreases from 0.7303 towards 0.71119, $\bar{k}$ increases from 1.3835 towards 1.5.

- when $0.6667 < \mu' \leq 0.7119$, there are three new entrants, choosing a quality level such that $\bar{k} = 1.5$ and all the incumbents exit.

- when $< \mu' \leq 0.6667$, the innovation is drastic and $\bar{k} \geq 1.5$.

Table 4.2 summarizes all the threshold values of μ', the corresponding market structure (N, N_1), the relative quality levels of the entrants $\overline{k}$ and the quality level of the incumbents $\bar{u}$ calculated for the example $\beta = 4$.

Comparing the results reported in Table 4.2 with those in Table 4.1 of the previous section, it is reassuring to see that the pattern is qualitatively similar except

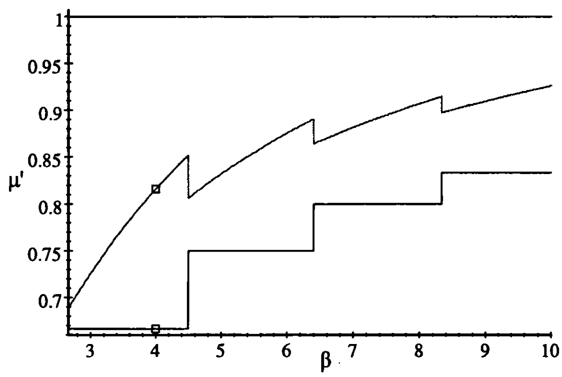


Figure 4.7: The two bounds

that the efficiency of the anticipated new technology impacts the quality level of the incumbents: the more efficient the new technology is, the less aggressive the incumbents are in investing in quality.

It can be shown, finally, that certain results obtain in this setting, for all values of β ($\beta > 8/3$). Here, we summarize these results in the form of two properties:

Property A: There exists a threshold value of μ', i.e. μ_{th}, such that:

If $\mu' > \mu_{th}$, then no new entry occurs in equilibrium;

If $\mu' < \mu_{th}$, then new entry occurs in equilibrium.

Property B: There exist two bounds of μ', i.e., $\mu_n < \mu_1 < 1$, such that when $\mu' > \mu_1$ no new entrant will enter; and when $\mu' < \mu_n$, $n(\beta)$ new entrants will enter and all incumbents will exit.

The existence of the two bounds is illustrated by Figure 4.7.

4.4 Conclusion

In this chapter we have studied the competition between two cohorts of radically different but substitutable technologies. While the incumbents employing the existent technology have a first mover advantage, potential entrants have the advantage of jumping to the more efficient new technology. By conducting equilibrium analysis on the entry of new firms, the exit of incumbents and subsequent quality competition, we have examined the impact of a radical innovation on market structure and on the turnover of firms. The analysis fully characterizes the conditions for three possible outcomes: either (1) the radically new technology is blocked by the existent technology, or (2) the two technologies coexist in the marketplace, or (3) the new technology replaces the existent technology. These conditions imply two bounds to the efficiency of the new technology, the first of which must be reached for 'creative destruction' to take place and the second of which suffices for its effects to be 'drastic'.

Chapter 5

Quality Competition, Market Structure and Technological Progress with Cohort Replacements

5.1 Introduction

In exploring the issues raised in Chapter 4, we have already introduced the idea that a drastic replacement of the population of firms may follow a technical shock. This idea, which I will refer to as 'cohort replacement' in what follows, has been recognized as an important aspect of technological progress since the contribution of Schumpeter (1942). The phenomenon has recently become the subject of formal economic modelling in the context of discussions of the Schumpeterian 'creative destruction' in the R&D-based growth literature (Grossman and Helpman 1992, Aghion and Howitt 1992). Differences in details notwithstanding, this strand of literature has a few common features: technological cohort replacement is discussed in the setting of a 'quality ladder'; and the advance of firms up the quality ladder is driven by the monopoly rent earned by patenting the next quality level. This chapter aims to explore technological cohort replacement from a somewhat different perspective.

The novel feature in what follows lies in relating technological cohort replacement to a radical change in R&D technology, which initiates a process of the follow-up innovative activities by rival firms. In this setting we develop an alternative explanation for the discontinuous feature often observed in quality improvements, and for the pattern of cohort replacement.

As in earlier chapters, we characterize an R&D technology by mapping from fixed outlays into product or process attributes. We saw in Chapter 4 that when radical change in technology induces cohort replacement, the outcome depends crucially on the gap between the new technology and the old. Cohort replacement does not necessarily take place even if the new technology is superior to the old; a (slightly) better new technology may be blocked by an established technology because of the presence of sunk costs invested in the old technology. Building upon this insight from Chapter 4, we consider in what follows the possibility that the established technology can enjoy a period of immunity vis-a-vis creative-destruction prior to the arrival of a strongly superior technology. The story developed here is consistent with the observation that technological progress may involve discrete jumps in market structure, which occur at well separated points in time.

As in earlier chapters, market structure and the level of R&D intensity are jointly determined within the model. In this setting, when a new round of quality competition is induced by the arrival of a more efficient R&D technology, a process of cohort replacement impacts on the equilibrium market structure and on the rate of firm turnover. The effects do not stop there. From a general equilibrium point of view, R&D spillovers induce further effects on aggregate technological progress, which in turn impacts upon the emergence of new R&D technologies. This chain of events can continue indefinitely, generating a continuing process of technological advance and economic growth.

The rest of the chapter is organized as follows. Section 2 presents the model. Section 3 characterizes the main mechanism of technological progress and economic growth. Section 4 analyzes the impacts of constraints on quality competition on market structure and growth. In Section 5 we discuss the robustness of the key results.

5.2 The Model

5.2.1 The Basic Setup

Two sectors

We consider a two sector model, comprising a final good sector and an intermediate good sector, with multiple industries producing different varieties. There is a continuum of intermediate good industries, the number of which is normalized to unity. The final good can be used in consumption and can also be used as input to produce intermediate goods. In the final good sector there are constant returns to scale and perfect competition. In each industry firms compete in 'quality' as well as in price. The general framework follows that which has been introduced in Chapter 2.

Aggregate Level Technology and Externalities

The advance of the frontier of the aggregate technology in each period, $\dot{A}_{\max}$, is a function of R&D efforts made by all firms in all industries, measured by L_2, and the spillover from existing knowledge stock $A_{\max}$, i.e., $\dot{A}_{\max} = \delta L_2 A_{\max}$. The idea of a spillover from $A_{\max}$ to $\dot{A}_{\max}$ is familiar within the R&D-based growth literature. The spillover from quality improvement by firms to aggregate advances in knowledge requires some comment. In this model, the R&D work force L_2 is hired by individual firms to improve the quality of the firms' products. In so doing, they also contribute to a generic progress in knowledge. In other words, we posit a spillover from industry-specific innovations, i.e., quality improvements, towards generic aggregate technical progress. Secondly, we further assume a reverse flow of information, in that an advance in the economy-wide level of technical expertise feeds into each individual industry.

Following Aghion-Howitt (1998) it is assumed that the stochastic structure of the aggregate technical progress is captured by a Poisson process with arrival rate λ, that the pace of technical progress by each aggregate innovation is γ, that the contributions from L_2 to aggregate innovations are independent so that the total arrival rate is $L_2\lambda$ and the rate of aggregate technical progress is $\dot{A}_{\max}/A_{\max} = L_2\lambda \ln \gamma = L_2\delta$, where $\delta = \lambda \ln \gamma$. The application of the frontier aggregate innovation to a particular industry a is given by a random draw from a uniform distribution between $[0,1]$. The success of an application depends on that the industry's current technology being of a vintage $\tau \geq T$; if $\tau < T$ the probability for the occurrence of a drastic cohort replacement is zero. The motivation for this assumption is that, for a particular industry to attract the attention of the outside scientific community, it is necessary that an opportunity exists for the introduction of a drastically superior new technology. As the pace of general scientific advance is slow and gradual, a certain length of time must elapse before the accumulated knowledge base affords the possibility for this kind of advance.

Modelling Cohort Replacement

We model cohort replacement as a three-stage game. In the first stage, potential entrants simultaneously choose whether to enter the industry. In the second stage, those firms which have already entered the market choose their 'quality' levels (u_i), which will be achieved by carrying out R&D and incurring the costs of hiring labor for the R&D function. In the third stage, firms' quality levels are given, and they compete in quantity (Cournot competition).

Consumption

There are L identical consumers who live infinitely, and each can supply one unit of labor per period. The labor can be supplied either as production labor used in final good sector, or used to do R&D. Each consumer has a discounted, linear utility function:

$$\int_0^\infty ce^{-\rho t}dt,$$

where c is consumption of final good, ρ is the discount rate. The Euler equation for intertemporal optimization is that $r = \rho$.

Final Good Sector

The final good can be treated as the numeraire. Its production function is

$$Y = L_1^{1-\alpha} \int_0^1 x_i^\alpha di, \tag{5.1}$$

where Y is the total output of final good, L_1 is labor input, x_i is the input of intermediate good i, $0 < \alpha < 1$.

There is perfect competition in final good market due to constant returns to scale, so equating the marginal product of labor to the wage rate yields the inverse labor demand function:

$$w = \frac{(1-\alpha)Y}{L_1}. \tag{5.2}$$

Similarly equating the marginal product of input i with its price leads to the inverse demand function for intermediate good i,

$$p_i = \alpha L_1^{1-\alpha} x_i^{\alpha-1} = \frac{\alpha Y}{x_i} \frac{x_i^\alpha}{\int_0^1 x_i^\alpha di}. \tag{5.3}$$

5.2.2 The Game of Quality Competition with Free Entry

In each industry experiencing cohort replacement a 3-stage game is played. As before, we characterize a subgame perfect Nash equilibrium. In the third stage, ('quantity competition') a Nash equilibrium in quantities exists and is unique, given the set of quality levels and the economy-wide demand conditions. The subgame Nash equilibrium determines for each firm a reduced form profit function, according to which the profit per period of firm k is

$$\pi_k (u_k \mid u_{-k}) = \frac{\alpha^{\frac{1}{1-\alpha}} L_1}{1-\alpha} \left(\frac{N-1+\alpha}{\sum_{j=1}^{N} \frac{1}{u_j}} \right)^{\frac{\alpha}{1-\alpha}} \left(1 - \frac{N-1+\alpha}{\sum_{j=1}^{N} \frac{u_k}{u_j}} \right)^2, \qquad (*)$$

where u_k is the quality level of firm k, u_{-k} is a $(N-1)$-tuple of quality levels of other firms except firm k, for $k = 1, 2, \cdots, N$; N is the number of firms in the industry and L_1 is the amount of labor employed in the final good sector.

The derivation of this reduced form profit function is give in Appendix[1] A. The key feature of the function is that a firm's profit increases with its relative quality level vis-a-vis its rivals', i.e., $\sum_{j=1}^{N} \frac{u_k}{u_j}$. Each firm incurs an R&D cost, in the form of a labor cost, as a function of its quality level as follows:

$$F(u_k; w, A_i) = w\mu \left(\frac{u_k}{A_i^{\frac{1-\alpha}{\alpha}}} \right)^{\beta}, \qquad (**)$$

where $\mu \left(\frac{u_k}{A_i^{\frac{1-\alpha}{\alpha}}} \right)^{\beta}$ $(\beta > 1)$ is the input of labor to generate quality level u_k, w is the market wage rate and μ is a constant. It is worth emphasizing that the increase in A_i, i.e., the aggregate level stock of knowledge, can shift the fixed cost curve downward as shown in Figure 5.1. The intuition here is that knowledge spillovers from aggregate technical progress enhances the efficiency of R&D by each new firm. This positive trend may be partially offset by the negative trend caused by an increase in the wage rate, also shown in Figure 5.1. The net effect is that the new cohort of firms have a technical advantage over old firms.

Each firm's objective in the second stage subgame is to maximize its net profit, in the present value sense, by choosing its own quality level given others' quality levels,

[1] A slight difference is that here we have replaced γ with α.

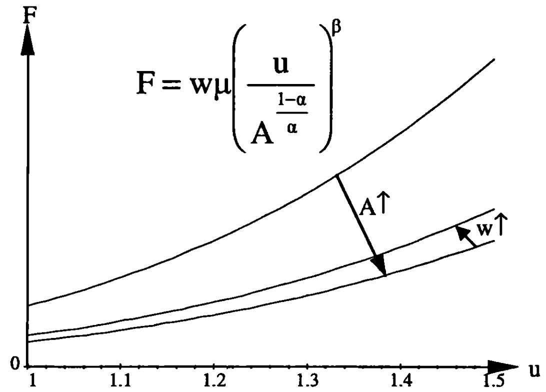


Figure 5.1: Fixed cost function

$$\max_{u_k}\left\{\frac{\alpha^{\frac{1}{1-\alpha}}L_1 v}{(1-\alpha)}\left(\frac{N-1+\alpha}{\sum_{j=1}^{N}\frac{1}{u_j}}\right)^{\frac{\alpha}{1-\alpha}}\left(1-\frac{N-1+\alpha}{\sum_{j=1}^{N}\frac{u_k}{u_j}}\right)^2 - w\mu\left(\frac{u_k}{A_i^{\frac{1-\alpha}{\alpha}}}\right)^{\beta}\right\}, \quad (5.4)$$

where $v \equiv \int_0^\infty \int_0^\tau e^{-rs} ds f(\tau)\, d\tau$ is the expected discounted present value of total revenue in a market of unit size, $f(\tau)$ is the probability density function of a cohort being replaced in the interval $(\tau, \tau + d\tau)$ after taking over the market. If we take the reduced form profit function and the fixed cost function as given, then the game is in essence a game of quality competition with free entry.

5.3 Equilibrium

5.3.1 Market Structure and R&D in Each Industry

Equilibrium analysis is focused on a symmetric subgame perfect Nash equilibrium, where each firm equates the marginal benefit of increasing quality and the marginal R&D cost. The best-reply function for each firm in a symmetric outcome is defined implicitly by the FOC for industry i:

$$S_i v\frac{1}{u}\left(\frac{\alpha}{N^3} + 2(N-1)\frac{N-1+\alpha}{N^3}\right) = \frac{\beta}{u}w\mu\left(\frac{u}{A^{\frac{1-\alpha}{\alpha}}}\right)^{\beta}$$

where $S_i \equiv \alpha^{\frac{1}{1-\alpha}} L_1 \left(\frac{N-1+\alpha}{N} u\right)^{\frac{\alpha}{1-\alpha}}$ is industry i's total revenue per period.

Now it is a feature of the model that the only industry that carries out quality innovation is the 'leading edge' industry, i.e., the industry with the highest value of A, i.e., $A = A_{\max}$. Denote the market size of this industry by $S_{\max}$ and its quality level by $u_{\max}$. We can now derive, from the FOC above, applied to the leading edge industry, the following equation:

$$F = w\mu \left(\frac{u_{\max}}{A_{\max}^{\frac{1-\alpha}{\alpha}}} \right)^{\beta} = S_{\max} v \frac{1}{\beta} \left(\frac{\alpha}{N^3} + 2(N-1) \frac{N-1+\alpha}{N^3} \right), \tag{5.5}$$

where $S_{\max} \equiv \alpha^{\frac{1}{1-\alpha}} L_1 \left(\frac{N-1+\alpha}{N} u\right)^{\frac{\alpha}{1-\alpha}}$.

It is assumed that any intermediate good industry is characterized by free-entry. Given our assumption on the technology of intermediate good production, this results in the following zero-profit condition, which equates the fixed cost of the firm with the present value of profit flows,

$$F = w\mu \left(\frac{u_{\max}}{A_{\max}^{\frac{1-\alpha}{\alpha}}} \right)^{\beta} = S_{\max} v \frac{1-\alpha}{N^2}. \tag{5.6}$$

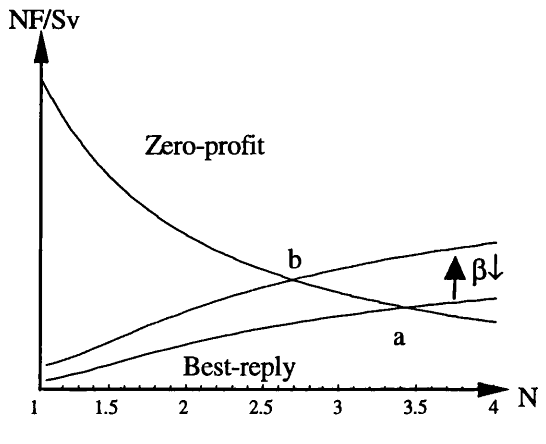


Figure 5.2: Market structure and industry R&D intensity

Figure 5.2 illustrates the above best-reply condition and zero-profit condition as two curves. The vertical axis shows the industry's R&D intensity and the horizontal axis shows the number of firms in each industry. The graph shows that the best-reply function is upward sloping and the zero profit schedule is downward sloping, so there is a unique symmetric Nash equilibrium, at which the level of R&D and market structure are simultaneously determined. The graph also shows that when the parameter β decreases, it shifts the best-reply function upwards and so increases both market concentration $1/N$ and industry R&D intensity (from **a** to **b**). The intuition for this result is as follows: A lower value of β means a higher degree of increasing returns to scale at firm level. The result is a more concentrated market structure in each intermediate good industry. These findings are summarized in the following

Proposition 23 *Market structure and the R&D intensity of each intermediate good industry are simultaneously determined in a symmetric equilibrium. The equilibrium number of firms in each industry is* $N = n(\beta, \alpha) \equiv \frac{n_0 + \sqrt{n_0^2 - 2(2-\alpha)}}{2}$, *where* $n_0 \equiv \frac{(2+\beta)(1-\alpha)}{2} + 1$. *Furthermore* N *is increasing in* β.

Proof. This follows[2] the proof of Proposition 1.

5.3.2 A Balanced Growth Equilibrium

Following the conventional strategy, we examine a balanced growth equilibrium, where $\frac{\dot{Y}}{Y} = \frac{\dot{c}}{c} = \frac{\dot{A}_{\max}}{A_{\max}} = \frac{\dot{w}}{w} = g$ and L_1, L_2 are constant over time. It is an immediate implication that $g = \frac{\dot{A}_{\max}}{A_{\max}} = \delta L_2 = \lambda L_2 \ln \gamma$, and that $\frac{Y}{A_{\max}}$ and $\frac{w}{A_{\max}}$ are constant over time.

By eq. (5.3) the total revenue in industry i per period is

$$S_i = p_i x_i = \alpha Y \frac{x_i^\alpha}{\int_0^1 x_i^\alpha di}. \tag{5.7}$$

The demand for R&D labor per period, L_2, is,

$$L_2 = (1 - \Omega) \lambda L_2 N \mu \left(\frac{u}{A^{\frac{1-\alpha}{\alpha}}} \right)^\beta, \tag{5.8}$$

[2]The slight difference is that here we have replaced γ with α.

where Ω is the proportion of industries which have vintage younger than T, therefore $(1-\Omega)$ is the proportion of industries which have vintage older than T, λL_2 is the arrival rate of aggregate innovation, and $N\mu\left(\frac{u}{A^{\frac{1-\alpha}{\alpha}}}\right)^{\beta}$ is the demand for innovative labor by each industry which experiences cohort replacement. This equation implies that

$$\left(\frac{u}{A^{\frac{1-\alpha}{\alpha}}}\right)^{\beta}=\frac{1}{(1-\Omega)\,\lambda\mu N}, \tag{5.9}$$

which further implies that, using the subscript max to denote the leading edge industry, as before:

$$\frac{u^i}{u_{\max}}=\left(\frac{A_i}{A_{\max}}\right)^{\frac{1-\alpha}{\alpha}}, \tag{5.10}$$

where u^i is the quality level of industry i, A_i is the aggregate technical level when industry i was last time at the technical frontier, and $u_{\max}$ is the quality level of the industry which is at the technological frontier for the time being.

Define

$$a_i \equiv \frac{A_i}{A_{\max}}, \tag{5.11}$$

which measures the relative technical level of a cohort against the frontier technical level. It follows that

$$\frac{u^i}{u_{\max}}=a_i^{\frac{1-\alpha}{\alpha}}.$$

By eq. (A.2) in Appendix A, it follows that $x_i=\alpha^{\frac{1}{1-\alpha}}L_1\left(\frac{N-1+\alpha}{N}u^i\right)^{\frac{1}{1-\alpha}}$, which further implies that

$$\frac{x_i}{x_{\max}}=\left(\frac{u^i}{u_{\max}}\right)^{\frac{1}{1-\alpha}}=a_i^{\frac{1}{\alpha}},$$

where $x_{\max}$ is the output of the frontier industry.

By eq. (5.7), it follows that

$$S_{\max}=\alpha Y\frac{x_{\max}^{\alpha}}{\int_0^1 x_i^{\alpha}di}=\alpha Y\frac{1}{\int_0^1\left(\frac{x_i}{x_{\max}}\right)^{\alpha}di}=\alpha Y\frac{1}{\int_0^1 a_i di}. \tag{5.12}$$

We rank a_i according to the vintage τ, so that

$$a(\tau) = e^{-\delta L_2 \tau}, \tag{5.13}$$

which means that an industry of vintage τ has relative technical level $e^{-\delta L_2 \tau}$.

The hazard rate of a cohort of vintage τ is

$$f(\tau) = \begin{cases} 0 & \text{if } \ 0 < \tau < T \\ L_2 \lambda e^{-L_2 \lambda (\tau - T)} & \text{if } \ \tau \geq T \end{cases}, \tag{5.14}$$

which is time-invariant in balanced growth equilibrium. This implies that in balanced growth equilibrium Ω is constant overtime therefore

$$\Omega \frac{1}{T} d\tau = (1 - \Omega) L_2 \lambda d\tau,$$

where $\frac{1}{T} d\tau$ is the proportion among those industries which are younger than T at the beginning of a very short period, which are aging to older than T during that period, and $L_2 \lambda d\tau$ is the proportion of those industries which are older than T at the beginning of the short period, and which undergo cohort replacement during that period. The equation implies that at any moment the transitions between the immune and exposed groups of industries are balanced, and the implied steady state proportion of the immune group is

$$\Omega = \frac{L_2 \lambda T}{1 + L_2 \lambda T}, \tag{5.15}$$

and consequently,

$$1 - \Omega = \frac{1}{1 + L_2 \lambda T}. \tag{5.16}$$

Substituting the above equation into eq. (5.9) results in

$$\left(\frac{u}{A^{\frac{1-\alpha}{\alpha}}} \right)^{\beta} = \frac{1 + L_2 \lambda T}{\lambda \mu N}. \tag{5.17}$$

In a steady state, the vintage distribution is time-invariant and so the vintage distribution density function coincides with the Poisson probability density function for $\tau \geq T$. This implies that the vintage distribution density function has the following form:

$$h(\tau) = \begin{cases} \frac{\Omega}{T} & \text{for} \quad 0 < \tau < T \\ (1-\Omega) L_2 \lambda e^{-L_2 \lambda (\tau - T)} & \text{for} \quad \tau \geq T \end{cases}. \tag{5.18}$$

At this point we make a critical simplifying assumption, which is motivated by the intuition that aggregate technological progress tends to advance in slow and gradual manner over time. In this spirit, we assume that technological progress at any period is made up of an infinite number of infinitesimal jumps in $A_{\max}$, and that the aggregate pace of progress is represented by a finite constant. Technically, this assumption implies that $\gamma \to 1$, $\lambda \to \infty$ and $\lambda \ln \gamma = \delta$ $(\delta > 0)$.

This in turn implies that $\Omega \to 1$, i.e., almost all industries are below the vintage T; once an industry reaches the vintage T, a cohort replacement immediately happens. Consequently, the following relations hold:

$$\int_0^1 a_i di = \int_0^T a(\tau) \frac{1}{T} d\tau = \frac{1}{T} \int_0^T e^{-\delta L_2 \tau} d\tau = \frac{\left(1 - e^{-\delta L_2 T}\right)}{\delta L_2 T}, \tag{5.19}$$

$$v = \int_0^\infty \int_0^\tau e^{-rs} ds f(\tau) d\tau = \int_0^T e^{-r\tau} d\tau = \frac{1 - e^{-rT}}{r} \tag{5.20}$$

and

$$\left(\frac{u}{A^{\frac{1-\alpha}{\alpha}}}\right)^\beta = \frac{L_2 T}{\mu N}. \tag{5.21}$$

Eq. (5.12) then implies that

$$S_{\max} = \frac{\alpha Y \delta L_2 T}{1 - e^{-\delta L_2 T}}. \tag{5.22}$$

To close the model, it remains to add the labor market clearing condition $L = L_1 + L_2$. Given the zero-profit condition $w\mu \left(\frac{u}{A^{\frac{1-\alpha}{\alpha}}}\right)^\beta = S_{\max} v \frac{1-\alpha}{N}$, a few manipulations to eq. (5.2), (5.21), (5.22) and (5.20) to eliminate w, $\left(\frac{u}{A^{\frac{1-\alpha}{\alpha}}}\right)^\beta$, $S_{\max}$ and v lead to the conclusion that

$$L_1 = \frac{\left(1 - e^{-\delta L_2 T}\right) rN}{\alpha \delta \left(1 - e^{-rT}\right)}. \tag{5.23}$$

The system reduces to four equations: $\begin{cases} r = \rho \\ L_1 = \frac{\left(1-e^{-\delta L_2 T}\right) rN}{\alpha\delta\left(1-e^{-rT}\right)} \\ g = \delta L_2 \\ L = L_1 + L_2 \end{cases}$, which imply that along a balanced growth path,

$$g = L\delta - \frac{\left(1 - e^{-gT}\right)\rho N}{\alpha\left(1 - e^{-\rho T}\right)}. \tag{5.24}$$

The properties of the implied growth rate are summarized in

Proposition 24 *For given T, the growth rate g is decreasing in the equilibrium number of firms in each industry, N, and it is decreasing in β.*

Proof. $\frac{\partial g}{\partial N} = -\frac{\rho\left(1-e^{-gT}\right)}{\alpha\left(1-e^{-\rho T}\right)+e^{-gT}N\rho T} < 0$. Given $\frac{\partial N}{\partial \beta} > 0$, then $\frac{\partial g}{\partial \beta} = \frac{\partial g}{\partial N}\frac{\partial N}{\partial \beta} < 0$. ■

The result provides some new insights about R&D based growth. The equilibrium market structure of each industry, represented by N, has a negative impact on growth rate. This does not imply, however, that growth could be enhanced by making market structure more concentrated, since the market structure is endogenously determined at equilibrium. A change in market structure occurs only in response to a change in some parameter underlying the quality competition game, such as β. For example, a decrease in β induces a more concentrated market structure, and an increase in the level of R&D conducted in each industry. Through spillovers from industrial R&D, the whole economy benefits from the shift, in the sense that it enjoys a higher growth rate.

5.4 Constraints on Quality Competition

In the last section we have demonstrated the feasibility of a growth mechanism which is driven by quality competition among rival firms. We might expect that constraints on quality competition would reduce the R&D intensity and the growth rate. To check this intuition, we characterize how constraints on quality competition affect the results.

With constraints on quality, the maximization problem (5.4) becomes:

$$\max_{u_i} \left\{ \frac{\alpha^{\frac{1}{1-\alpha}} L_1 v}{(1-\alpha)} \left(\frac{N-1+\alpha}{\sum_{j=1}^{N} \frac{1}{u_j}} \right)^{\frac{\alpha}{1-\alpha}} \left(1 - \frac{N-1+\alpha}{\sum_{j=1}^{N} \frac{u_i}{u_j}} \right)^2 - w\mu \left(\frac{u_i}{A_i^{\frac{1-\alpha}{\alpha}}} \right)^{\beta} \right\}$$

$$s.t. \quad : \quad \left(\frac{u_i}{A_i^{\frac{1-\alpha}{\alpha}}} \right)^{\beta} \leq \eta \left(\frac{u}{A^{\frac{1-\alpha}{\alpha}}} \right)^{\beta}, \quad 0 < \eta < 1,$$

where u is the quality level in the unconstrained equilibrium for a given level of A, and so by eq. (5.9) $\left(\frac{u}{A^{\frac{1-\alpha}{\alpha}}} \right)^{\beta} = \frac{L_2 T}{\mu N}$ holds; η is the quality constraint parameter. Given that $0 < \eta < 1$, the quality constraints are binding and this results in the following corner solution to the above optimization problem:

$$\left(\frac{u_c}{A_c^{\frac{1-\alpha}{\alpha}}} \right)^{\beta} = \eta \frac{L_2 T}{\mu N}. \tag{5.25}$$

In the new balanced growth equilibrium, equations (5.6) - (5.24) still hold, since the derivation of these equations does not depend on the best-reply condition (5.5), which is the only condition alters in the new setting. Thus eq. (5.9) implies that

$$\left(\frac{u_c}{A_c^{\frac{1-\alpha}{\alpha}}} \right)^{\beta} = \frac{L_{2c} T}{\mu N_c}, \tag{5.26}$$

which in connection with eq. (5.25) leads to

$$\frac{L_{2c}}{N_c} = \frac{\eta L_2}{N}. \tag{5.27}$$

Given that $g = \delta L_2$ and $g_c = \delta L_{2c}$, eq. (5.27) implies that,

$$\frac{g_c}{N_c} = \frac{\eta g}{N}, \tag{5.28}$$

or alternatively,

$$g_c(N, \eta) = \frac{\eta g(N)}{N} N_c(N, \eta). \tag{5.29}$$

Partially differentiating the both sides of eq. (5.29) with respect to η results in

$$\frac{\partial g_c(N,\eta)}{\partial \eta} = g(N)\frac{N_c(N,\eta) + \eta\frac{\partial N_c(N,\eta)}{\partial \eta}}{N}. \tag{5.30}$$

The application of eq. (5.24) in this new setting implies

$$g_c(N,\eta) = L\delta - \frac{\left(1 - e^{-g_c(N,\eta)T}\right)\rho N_c(N,\eta)}{\alpha\left(1 - e^{-\rho T}\right)} \tag{5.31}$$

Partially differentiating the both sides of the above equation with respect to η and solving for $\frac{\partial g_c(N,\eta)}{\partial \eta}$ results in

$$\frac{\partial g_c(N,\eta)}{\partial \eta} = \frac{\partial g_c}{\partial N_c}\frac{\partial N_c(N,\eta)}{\partial \eta} = -\frac{\rho\left(1 - e^{-g_cT}\right)}{\alpha\left(1 - e^{-\rho T}\right) + e^{-g_cT}N_c\rho T}\frac{\partial N_c(N,\eta)}{\partial \eta}. \tag{5.32}$$

We can combine eq. (5.30) and (5.32) to solve for $\frac{\partial N_c(N,\eta)}{\partial \eta}$ as follows

$$\frac{\partial N_c(N,\eta)}{\partial \eta} = -\frac{g(N)N_c(N,\eta)}{\frac{\rho(1-e^{-g_cT})N}{\alpha(1-e^{-\rho T})+e^{-g_cT}N_c\rho T} + g(N)\eta}, \tag{5.33}$$

which implies

$$\frac{\partial N_c(N,\eta)}{\partial \eta} < 0. \tag{5.34}$$

Substituting eq. (5.33) into (5.32) leads to

$$\frac{\partial g_c(N,\eta)}{\partial \eta} = \frac{\rho\left(1 - e^{-g_cT}\right)g(N)N_c(N,\eta)}{\rho\left(1 - e^{-g_cT}\right)N + \left(\alpha\left(1 - e^{-\rho T}\right) + e^{-g_cT}N_c\rho T\right)g(N)\eta},$$

which implies

$$\frac{\partial g_c(N,\eta)}{\partial \eta} > 0. \tag{5.35}$$

Proposition 25 *The effect of constraints on quality competition is to make market structure more fragmented, i.e.,* $\frac{\partial N_c(N,\eta)}{\partial \eta} < 0$, *while the rate of growth decreases, i.e.,* $\frac{\partial g_c(N,\eta)}{\partial \eta} > 0$.

5.5 Discussion: The Immunity Period

5.5.1 Market Structure and the Duration of Immunity

In the two preceding sections we have assumed that the duration of the immunity period is fixed, independently of market structure. This assumption greatly simplifies the analysis, but the question arises as to whether our main results are driven by this assumption. To check on this, we explore the relation between the duration of the immunity period T and market structure[3], measured by N.

We first consider the meaning of a 'drastic cohort replacement'. A drastic cohort replacement is one in which the incumbents' relative quality levels are so low relative to that of new arrivals that they make no impact on post-entry competition. This implies that at the time of the radical replacement, the incumbents' mark-up ratio is non-positive, i.e., they make losses if they produce. According to eq. (A.1) in Appendix A, the mark-up ratio of a firm takes the form of

$$\frac{p-\frac{1}{u_i}}{p}=\left(1-\frac{\tilde{N}-1+\alpha}{\sum_{j=1}^{\tilde{N}}\frac{u_i}{u_j}}\right).$$

At the time of a drastic replacement, the new entrants obtain quality level $u_{\max}$ while the incumbents have quality level $u_{\min}$. If an incumbent deviated from the equilibrium strategy, by continuing production, then the number of firms in this market will be $\tilde{N}=N+1$ and the incumbent's mark-up ratio would be

$$1-\frac{N+1-1+\alpha}{N\frac{u_{\min}}{u_{\max}}+1}\leq 0,$$

which implies

$$\frac{u_{\max}}{u_{\min}}\geq\frac{N}{N-1+\alpha}. \tag{5.36}$$

Recall eq. (5.10)

$$\frac{u^i}{u_{\max}}=\left(\frac{A_i}{A_{\max}}\right)^{\frac{1-\alpha}{\alpha}},$$

[3] We are mainly interested in the case of oligopoly where $N\geq 3$.

which has the implication that

$$\frac{\dot{u}_{\max}}{u_{\max}} = \frac{1-\alpha}{\alpha}\frac{\dot{A}_{\max}}{A_{\max}} = \frac{1-\alpha}{\alpha}g, \tag{5.37}$$

i.e., the growth rate of quality in the balanced growth equilibrium is $\frac{1-\alpha}{\alpha}g$. Therefore at the time of drastic cohort replacement, the quality ratio between the new entrants and the incumbents is

$$\frac{u_{\max}}{u_{\min}} = e^{\frac{1-\alpha}{\alpha}gT}. \tag{5.38}$$

We assume that a drastic cohort replacement takes place immediately once inequality (5.36) is satisfied. It follows that we can deduce from eq. (5.38) that

$$e^{\frac{1-\alpha}{\alpha}gT} = \frac{N}{N-1+\alpha}. \tag{5.39}$$

The implications of eq. (5.39) include the following four equivalent equations:

$$gT = \ln\left(\frac{N}{N-1+\alpha}\right)^{\frac{\alpha}{1-\alpha}},$$

$$g = \frac{\alpha}{(1-\alpha)T}\ln\frac{N}{N-1+\alpha} \tag{5.40}$$

$$T = \ln\left(\frac{N}{N-1+\alpha}\right)^{\frac{\alpha}{(1-\alpha)g}},$$

and

$$N = \frac{1-\alpha}{1-e^{-\frac{1-\alpha}{\alpha}gT}}.$$

Combining the above equations with eq. (5.24) leads to the following three equations, which indicate the relations between the pairs: (N,T), (N,g) and (T,g) respectively.

$$\frac{\alpha}{(1-\alpha)T}\ln\frac{N}{N-1+\alpha} = L\delta - \frac{\left(1-\left(\frac{N}{N-1+\alpha}\right)^{-\frac{\alpha}{1-\alpha}}\right)\rho N}{\alpha\left(1-e^{-\rho T}\right)} \tag{5.41}$$

$$g = L\delta - \frac{\left(1 - \left(\frac{N}{N-1+\alpha}\right)^{-\frac{\alpha}{1-\alpha}}\right)\rho N}{\alpha\left(1 - \left(\frac{N}{N-1+\alpha}\right)^{-\frac{\alpha\rho}{(1-\alpha)g}}\right)} \tag{5.42}$$

$$g = L\delta - \frac{(1-\alpha)\left(1 - e^{-gT}\right)\rho}{\alpha\left(1 - e^{-\rho T}\right)\left(1 - e^{-\frac{1-\alpha}{\alpha}gT}\right)} \tag{5.43}$$

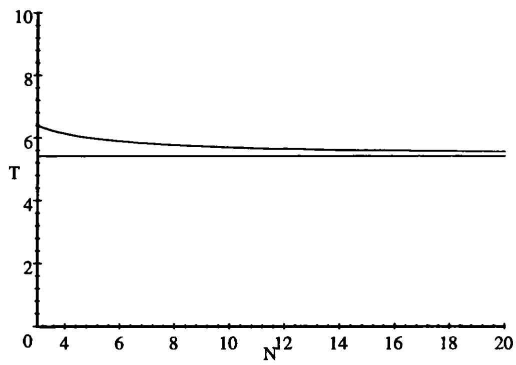
Characterizing these relations analytically appears to be difficult; in what follows we appeal to some numerical examples and graphical characterizations. We report one typical example below; further examples are reported in Appendix D.

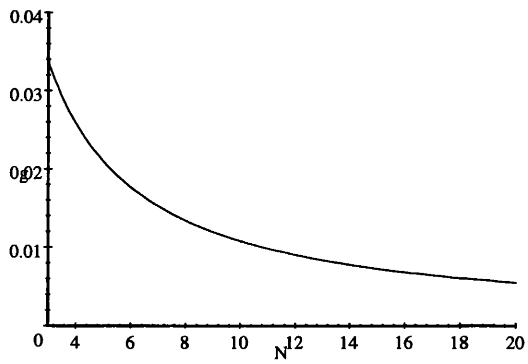
Example 26 *Set* $\alpha = 3/5, L = 1, \delta = 1/5, \rho = 3/100$, *then*

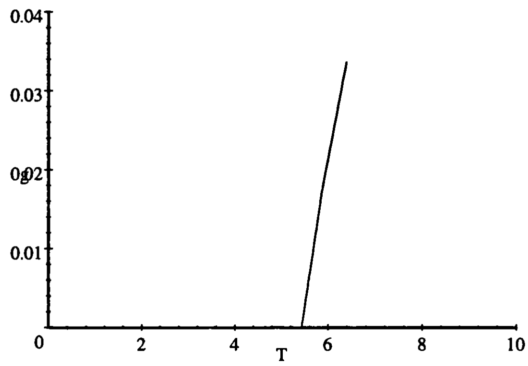
$\frac{3}{2T}\ln\frac{N}{N-\frac{2}{5}} = \frac{1}{5} - \frac{1}{20}\frac{\left(1-\left(\frac{N}{N-\frac{2}{5}}\right)^{-\frac{3}{2}}\right)N}{1-e^{-\frac{3}{100}T}}$,

$g = \frac{1}{5} - \frac{1}{20}\frac{\left(1-\left(\frac{N}{N-\frac{2}{5}}\right)^{-\frac{3}{2}}\right)N}{1-\left(\frac{N}{N-\frac{2}{5}}\right)^{-\frac{9}{200g}}}$ *and*

$g = \frac{1}{5} - \frac{1}{50}\frac{1-e^{-gT}}{\left(1-e^{-\frac{3}{100}T}\right)\left(1-e^{-\frac{2}{3}gT}\right)}$ *hold.*







Remark 27 *In the light of the above example, it is suggested that the possible response of T to the change in N is bounded both from below and from above, and it decreases in N. (See the upper left panel of the above figure, where the horizontal dashed line is the lower bound to T. As N increases, T decreases but converges to a lower bound which is bounded away from zero); The upper right panel of the above figure shows that g also decreases in N, consequently as shown in the lower panel of the above figure, T and g respond in the same direction when N changes. (These patterns are confirmed by the more numerical examples reported in Appendix D.) Given that the response of T to the change in N is bounded, the simplifying assumption that T is independent of N turns out not have biased the characterization of the relation between N and g qualitatively.*

In Appendix D we discuss how to pin down the lower bound and an upper bound to T in general, and provide some numerical examples to show that the upper bound and the lower bound to T can be quite close to each other therefore the range for T to change is quite limited.

5.5.2 Constraints on Quality Revisited

We now return to the setting where there are constraints on quality, and extend our consideration of the endogeneity of the immunity period to this setting. Here, eq.

(5.27) is modified to

$$\frac{L_{2c}T_c}{\mu N_c} = \eta \frac{L_2 T}{\mu N}. \tag{5.44}$$

Given that $g_c = \delta L_{2c}$ and $g = \delta L_2$, the above equation translates into

$$\frac{g_c T_c}{N_c} = \eta \frac{gT}{N}. \tag{5.45}$$

By the same token, we have from eq. (5.40),

$$g_c T_c = \frac{\alpha}{1-\alpha} \ln \frac{N_c}{N_c - 1 + \alpha}. \tag{5.46}$$

Combining eq. (5.40), (5.45) and eq. (5.46) to eliminate $g_c T_c$ and gT, we get the following equation which relates N_c to N and η,

$$\frac{\ln \frac{N_c}{N_c-1+\alpha}}{N_c} = \eta \frac{\ln \frac{N}{N-1+\alpha}}{N}, \tag{5.47}$$

which implies that N_c is function of N and η, i.e., $N_c = N_c(N, \eta)$. Differentiating both sides of the above equation with respect to η and solving for $\frac{\partial N_c(N,\eta)}{\partial \eta}$ leads to

$$\frac{\partial N_c(N,\eta)}{\partial \eta} = -\frac{N_c^2 \ln \frac{N}{N-1+\alpha}}{N\left(\frac{1-\alpha}{N_c-1+\alpha} + \ln \frac{N_c}{N_c-1+\alpha}\right)}, \tag{5.48}$$

which implies

$$\frac{\partial N_c(N,\eta)}{\partial \eta} < 0. \tag{5.49}$$

We may summarize the implications as follows:

1. Tightening the constraint on quality competition leads to a more fragmented market structure.

2. If $\frac{\partial g_c}{\partial N_c} < 0$ (as is the case in the examples considered here), then $\frac{\partial g_c(N,\eta)}{\partial \eta} = \frac{\partial g_c}{\partial N_c}\frac{\partial N_c(N,\eta)}{\partial \eta} > 0$.

The discussion in this section provides us with some insight which partially justifies the simplifying assumption that the duration of immunity period can be treated as fixed. First, in the examples we consider, the response of the duration of the immunity period is very limited, because it is bounded both from above and below. Second, endogenizing the duration of the immunity period does not qualitatively change the main results obtained under the simplifying assumption.

5.6 Concluding Remarks

Creative destruction and firm turnover, in the setting of this chapter, emerge as outcomes of quality competition, in association with radical technological changes. This strengthens the claim of Chapter 2, that quality competition constitutes a key vehicle in driving innovation and growth at industry level, and in combination with R&D spillover effects it also drives aggregate technical progress. The analysis shows that this mechanism is robust in the sense that it does not rely on whether aggregate technological progress proceeds through variety expansions a la Romer (1990) or through cohort replacements a la Aghion-Howitt (1992). This study also strengthens another claim developed in Chapter 2: that constraints on quality consumption reduce R&D intensity, and lower the rate of growth.

Chapter 6

Conclusions

Before summing up the major themes that run through these chapters, it is worth remarking on three underlying ideas that form a point of departure for the analysis.

6.1 Three General Ideas

1. Firm-specific Accumulation of Capability: (Partial) Non-rivalry and (Partial) Exclusion

The output of an innovation process takes the form of an increase in the stock of knowledge, skills or organizational capabilities of the firm. By committing resources to innovative activities firms can build up their capabilities, where we can interpret the firm's capability by (i) a shift parameter u ('quality') that moves the firm's demand schedule outwards, and/or (ii) a measure of the level of the marginal cost of production ('productivity'). In the preceding chapters, the analysis has been conducted in terms of a mapping $F(u)$ that links R&D spending to quality ('product innovation') or the unit cost of production ('process innovation').

Despite the proprietary nature of these capabilities, there remain a number of channels through which 'spillovers' of know-how occur, and this observation underlies one of the basic themes developed below.

2. R&D Competition between Rival Firms

Much of the R&D-based growth literature relies on representation of product market competition which, from the perspective of the industrial organization literature, appear rather special. The monopolistic competition framework excludes the kind of R&D escalation mechanism discussed in the foregoing chapters, while the ladder model of quality competition implies that each market as monopolized at each point in time. The framework used in the forgoing chapters allows us to examine the race of R&D competition, while permitting a realistic representation of market structure.

3. Increasing Returns to Scale

The use of the term 'increasing returns' has led to some confusion in the recent literature, in that different authors use the term to mean different things. It is worth distinguishing a number of uses, relating to the analysis of the preceding chapters:

(i) Technical 'increasing returns'. This relates to the shape of the $F(u)$ function. In the present thesis, this displays diminishing returns.

(ii) A second sense in which the item is used relates to the shape of the firm's average cost schedule. Since we assume fixed outlays in combination with a flat marginal cost schedule, the present analysis assumes increasing returns in this sense.

(iii) A third sense relates to the mapping from fixed outlays F to profit π. In the above chapters, the relation between profit and R&D inputs is 'S' shaped. In other words, increasing returns in this sense can hold over a certain limited range in the setting that features constant returns to scale in production, together with an R&D technology which displays decreasing technical returns in the sense of (i) above.

The degree of increasing returns to scale has a significance on market structure, and on resource allocation as has been demonstrated in all four previous chapters.

6.2 Three Major Themes

The central aim of the thesis has been to show how market structure shapes, and changes in line with the evolution of technology. This has been done by developing a class of quality competition models, within which three basic themes have been developed.

(a) Quality Competition: The Joint Determination of R&D Intensity and Market Structure

The first major theme relates to the joint determination of R&D intensity and market structure in the context of quality competition; the models developed in Chapters 2-5 all display this feature. This places in the forefront of the analysis the role of a general economic mechanism, the "escalation mechanism" (Sutton, 1998), in shaping equilibrium outcomes. The aim of this thesis has been to fit this economic mechanism into some broader frameworks and to explore its implications.

In Chapters 2 and 5, we show how the escalation mechanism can be incorporated into a general equilibrium framework, so that market structure, R&D intensity and the rate of economic growth all emerge as endogenous variables. In Chapter 3, we show how the escalation mechanism can be related to the process of shakeout of firms during a typical industry life cycle. In Chapter 4, we show how a radical change in the technology of R&D may affect the interaction between quality competition and market structure.

(b) Creative Destruction and the Birth of New Industries

While Chapter 2 captures the mechanisms associated with the birth and growth of new industries, Chapter 5 is concerned with the complementary process of 'creative destruction'. Taken together, these chapters paint a picture of the economic growth process which features a continuous turnover in the population of industries.

The mechanisms driving this process rest upon the process of spillovers of knowledge, and so on the limits to the proprietary nature of capabilities mentioned earlier. What is novel to the present literature is that these spillovers are modelled as occurring across industries, as opposed to over time, as in the existing growth literature.

(c) Financial Institutions and Market Environments

We have shown in Chapter 2, 3 and 5, that whether quality competition and the escalation mechanism can be effective in driving the process of innovation and growth depends on whether the institutional factors are present which support investments in quality building. If the only source of finance lies in retained earnings, then the uptake of profitable innovation opportunities may be constrained.

The analysis of Chapter 2 and 5 elaborates on the ways in which the absence of well functioning financial markets will impinge in damaging ways on the growth process.

Appendix A

For Chapter 2

Derivation of the reduced form profit function

To ease notation, the subscript which identifies industries is dropped. Then the inverse demand function (2.3) faced by a representative industry becomes

$$p = \gamma L_1^{1-\gamma} x^{\gamma-1},$$

where x is the industry's output, $x = \sum_{j=1}^{N} x^j$, x^j is the output of firm j and N is the number of firms. It can rewritten as

$$p = \left(\frac{\gamma^{\frac{1}{1-\gamma}} L_1}{\sum_{j=1}^{N} x^j} \right)^{1-\gamma}.$$

Consider the Cournot competition in quantity: for given (x^j), $j \neq i$, firm i chooses x^i to maximize its profit:

$$\max_{x^i} \left(\left(\frac{\gamma^{\frac{1}{1-\gamma}} L_1}{\sum_{j=1}^{N} x^j} \right)^{1-\gamma} - \frac{1}{u_i} \right) x^i,$$

where $\frac{1}{u_i}$ is the cost of producing one unit of quality-adjusted quantity. u_i is the 'quality' (or productivity) index. The FOC of the problem is,

$$-(1-\gamma)\frac{\left(\frac{\gamma^{\frac{1}{1-\gamma}}L_1}{\sum_{j=1}^N x^j}\right)^{1-\gamma}}{\sum_{j=1}^N x^j}x^i+\left(\frac{\gamma^{\frac{1}{1-\gamma}}L_1}{\sum_{j=1}^N x^j}\right)^{1-\gamma}-\frac{1}{u_i}=0.$$

Substituting $\left(\frac{\gamma^{\frac{1}{1-\gamma}}L_1}{\sum_{j=1}^N x^j}\right)^{1-\gamma}$ with p and solving for x^i, gives the following equation:

$$x^i=\frac{\left(p-\frac{1}{u_i}\right)x}{(1-\gamma)p}.$$

The market share of firm i is then

$$\frac{x^i}{x}=\frac{1}{1-\gamma}\frac{p-\frac{1}{u_i}}{p}.$$

Summing up the market shares of all firms to eliminate x^i and x: $\sum_{i=1}^N \frac{x^i}{x}=1=\frac{1}{1-\gamma}\frac{Np-\sum_{i=1}^N\frac{1}{u_i}}{p}$ and solving for p leads to

$$p=\frac{\sum_{j=1}^N\frac{1}{u_j}}{N-1+\gamma}.$$

By the above equations the market share of firm i can be solved as $\frac{x^i}{x}=\frac{1}{1-\gamma}\left(1-\frac{N-1+\gamma}{\sum_{j=1}^N\frac{u_i}{u_j}}\right)$.

The mark-up ratio of firm i is

$$\frac{p-\frac{1}{u_i}}{p}=\left(1-\frac{N-1+\gamma}{\sum_{j=1}^N\frac{u_i}{u_j}}\right). \tag{A.1}$$

The industry output in equilibrium then is

$$x=\gamma^{\frac{1}{1-\gamma}}L_1\left(\frac{N-1+\gamma}{\sum_{j=1}^N\frac{1}{u_j}}\right)^{\frac{1}{1-\gamma}}. \tag{A.2}$$

The total revenue of the industry turns out to be

$$S\equiv px=\gamma^{\frac{1}{1-\gamma}}L_1\left(\frac{N-1+\gamma}{\sum_{j=1}^N\frac{1}{u_j}}\right)^{\frac{\gamma}{1-\gamma}}.$$

Finally, the reduced form profit function is derived as

$$\pi_i = \frac{\gamma^{\frac{1}{1-\gamma}} L_1}{1-\gamma} \left(\frac{N-1+\gamma}{\sum_{j=1}^{N} \frac{1}{u_j}} \right)^{\frac{\gamma}{1-\gamma}} \left(1 - \frac{N-1+\gamma}{\sum_{j=1}^{N} \frac{u_i}{u_j}} \right)^2 \text{ or}$$

$$\pi_i \left(u_i \mid u_{-i} \right) = \underbrace{S}_{\text{market size}} \underbrace{\frac{1}{1-\gamma} \left(1 - \frac{N-1+\gamma}{\sum_{j=1}^{N} \frac{u_i}{u_j}} \right)}_{\text{market share}} \underbrace{\left(1 - \frac{N-1+\gamma}{\sum_{j=1}^{N} \frac{u_i}{u_j}} \right)}_{\text{mark-up ratio}}.$$

Proof of Proposition 1

By the best reply condition eq. (2.5) and the zero profit condition eq. (2.6), the following quadratic equation of N_f can be derived:

$$N_f^2 - N_f \left(\frac{(2+\beta)(1-\gamma)}{2} + 1 \right) + \frac{2-\gamma}{2} = 0,$$

the largest positive solution of which is

$$N_f = n(\beta, \gamma) \equiv \frac{n_0 + \sqrt{n_0^2 - 2(2-\gamma)}}{2},$$

where $n_0 \equiv \frac{(2+\beta)(1-\gamma)}{2} + 1$. It is easy to see that

$$\frac{\partial N_f}{\partial \beta} = \frac{\partial N_f}{\partial n_0} \frac{\partial n_0}{\partial \beta} > 0.$$

Proof of Proposition 3

By Proposition 1, $N_f = \frac{n_0 + \sqrt{n_0^2 - 2(2-\gamma)}}{2}$, where $n_0 \equiv \frac{(2+\beta)(1-\gamma)}{2} + 1$.
Then the following goes through:

$$\begin{aligned} N_f &> 1 \Rightarrow \sqrt{n_0^2 - 2(2-\gamma)} > 2 - n_0 \\ &\Rightarrow \begin{cases} n_0^2 - 2(2-\gamma) > (2-n_0)^2 & \text{if } n_0 \leq 2 \\ n_0^2 - 2(2-\gamma) > 0 & \text{if } n_0 > 2 \end{cases} \\ &\Rightarrow \begin{cases} n_0 > (2-\gamma) + \frac{1}{2}\gamma & \text{if } n_0 \leq 2 \\ n_0 > 2 & \text{otherwise} \end{cases} \\ &\Rightarrow n_0 > (2-\gamma) + \frac{1}{2}\gamma > 2 - \gamma. \end{aligned}$$

Then the following comparative statics can be derived:

$$\begin{aligned}
\frac{\partial\left(N\Pi\left(\beta,\gamma,N\right)\right)}{\partial N} &= \frac{-2\left(Nn_0-(2-\gamma)\right)}{\beta N^3} < \frac{-2\left((N-1)\left(2-\gamma\right)\right)}{\beta N^3} < 0 \text{ and} \\
\frac{\partial\left(NF_0\right)}{\partial N} &= \frac{2\left(N-1\right)\left(2-\gamma\right)}{\beta N^3} > 0 \text{ when } N > 1.
\end{aligned}$$

Proof of Lemma 4

Given $g = \frac{\frac{\gamma\delta N}{1-\gamma}\min\left(\frac{F_0}{\theta},\frac{\Pi}{1-\theta}\right)L-\rho}{\varepsilon+\frac{\gamma}{N}}$.

Note that if $N > N^*$ and $g > 0$, then $\frac{F_0}{\theta} > \frac{\Pi}{1-\theta}$; and if $g = \frac{\frac{\gamma\delta N}{1-\gamma}\frac{\Pi}{1-\theta}L-\rho}{\varepsilon+\frac{\gamma}{N}} > 0$, then $\frac{\gamma\delta N\Pi}{(1-\gamma)(1-\theta)}L > \rho$. Consequently,

$$\begin{aligned}
\frac{\partial g}{\partial N} &= \frac{\partial}{\partial N}\left(\frac{\frac{\gamma\delta LN^2\Pi}{(1-\gamma)(1-\theta)}-\rho N}{N\varepsilon+\gamma}\right) \\
&= \frac{\partial}{\partial N}\left(\frac{\gamma\delta L}{(1-\gamma)(1-\theta)}\frac{(1-\gamma)\beta N-\gamma-2(N-1)(N-1+\gamma)}{\beta N} - \rho N\right)\frac{1}{N\varepsilon+\gamma}+\frac{\partial}{\partial N}\left(\frac{1}{N\varepsilon+\gamma}\right)\left(\frac{\gamma\delta LN\Pi}{(1-\gamma)(1-\theta)} - \rho\right)N \\
&= -\underset{(+)}{\left(\gamma\delta L\frac{2\left(N^2-1\right)+\gamma}{(1-\gamma)(1-\theta)\beta N^2} + \rho\right)}\frac{1}{N\varepsilon+\gamma} - \frac{\varepsilon}{(\varepsilon N+\gamma)^2}\underset{(+)}{\left(\frac{\gamma\delta LN\Pi}{(1-\gamma)(1-\theta)} - \rho\right)}N < 0.
\end{aligned}$$

Also note that if $N \leq N^*$ and $g > 0$, then $\frac{F_0}{\theta} \leq \frac{\Pi}{1-\theta}$; and if $g = \frac{\frac{\gamma\delta N}{1-\gamma}\frac{F_0}{\theta}L-\rho}{\varepsilon+\frac{\gamma}{N}} > 0$, then $\frac{\gamma\delta LNF_0}{(1-\gamma)\theta} > \rho$. Consequently,

$$\begin{aligned}
\frac{\partial g}{\partial N} &= \frac{\partial}{\partial N}\left(\frac{\frac{\gamma\delta LN^2F_0}{(1-\gamma)\theta}-\rho N}{N\varepsilon+\gamma}\right) \\
&= \frac{\partial}{\partial N}\left(\frac{\gamma\delta L}{(1-\gamma)\theta}N^2F_0 - \rho N\right)\frac{1}{N\varepsilon+\gamma} - \frac{\gamma\delta LNF_0-\rho}{(1-\gamma)\theta}N\frac{\partial}{\partial N}\left(\frac{1}{N\varepsilon+\gamma}\right) \\
&= \left(\frac{\gamma\delta L}{(1-\gamma)\theta}\frac{\partial}{\partial N}\left(\frac{\gamma+2(N-1)(N-1+\gamma)}{\beta N}\right) - \rho\right)\frac{1}{N\varepsilon+\gamma} - \frac{\gamma\delta LNF_0-\rho}{(1-\gamma)\theta}N\frac{\partial}{\partial N}\left(\frac{1}{N\varepsilon+\gamma}\right) \\
&= \left(\frac{\gamma\delta L}{(1-\gamma)\theta}\frac{2N^2+\gamma-2}{\beta N^2} - \rho\right)\frac{1}{N\varepsilon+\gamma} + \frac{\gamma\delta LNF_0-\rho}{(1-\gamma)\theta}N\frac{\varepsilon}{(N\varepsilon+\gamma)^2} \\
&< \left(\frac{\gamma\delta L}{(1-\gamma)\theta}\frac{2N^2+\gamma-2}{\beta N^2} - \frac{\gamma\delta LNF_0}{(1-\gamma)\theta}\right)\frac{1}{N\varepsilon+\gamma} + \frac{\gamma\delta LNF_0-\rho}{(1-\gamma)\theta}N\frac{\varepsilon}{(N\varepsilon+\gamma)^2} \\
&= \left(\frac{\gamma\delta L}{(1-\gamma)\theta}\frac{2N^2+\gamma-2}{\beta N^2} - \frac{\gamma\delta L}{(1-\gamma)\theta}\frac{\gamma+2(N-1)(N-1+\gamma)}{\beta N^2}\right)\frac{1}{N\varepsilon+\gamma} + \frac{\gamma\delta LNF_0-\rho}{(1-\gamma)\theta}N\frac{\varepsilon}{(N\varepsilon+\gamma)^2} \\
&= \underset{(+)}{\frac{\gamma\delta L}{(1-\gamma)\theta}\frac{2(2-\gamma)(N-1)}{\beta N^2}}\frac{1}{N\varepsilon+\gamma} + \underset{(+)}{\frac{\gamma\delta LNF_0-\rho}{(1-\gamma)\theta}}N\frac{\varepsilon}{(N\varepsilon+\gamma)^2} > 0.
\end{aligned}$$

Appendix B

For Chapter 3

Derivation of the Reduced Form Profit Function

The subscript can be removed when a representative submarket in an arbitrary period after its emergence is addressed. So the demand function in eq. (3.2) becomes:

$$X = S\left(\frac{1}{p}\right)^{\frac{1}{1-\gamma}},$$

.where[1] X is the submarket's output, $X = \sum_{j=1}^{N} x^j$, x^j is the output of producer j and N is the number of producers. It can be rewritten as

$$p = \left(\frac{S}{\sum_{j=1}^{N} x^j}\right)^{1-\gamma}.$$

Consider the Cournot Competition in quantity: for given vector (x^j), $j \neq i$, producer i chooses x^i to maximize her profit:

$$\max_{x^i} \left(\left(\frac{S}{\sum_{j=1}^{N} x_j}\right)^{1-\gamma} - \frac{1}{u_i}\right) x^i,$$

where $\frac{1}{u_i}$ is the cost of producing one unit of effective quantity. u_i is the 'quality'

[1] If X is interpreted as effective quantity, then this demand function can be applied to both homogeneous good or vertically differentiated good.

(or productivity) index. The FOC of the program is:

$$-(1-\gamma)\frac{\left(\frac{S}{\sum_{j=1}^{N}x^j}\right)^{1-\gamma}}{\sum_{j=1}^{N}x^j}x^i+\left(\frac{S}{\sum_{j=1}^{N}x^j}\right)^{1-\gamma}-\frac{1}{u_i}=0.$$

Substituting $\left(\frac{S}{\sum_{j=1}^{N}x^j}\right)^{1-\gamma}$ with p and solving for x^i gives the following equation:

$$x^i=\frac{\left(p-\frac{1}{u_i}\right)X}{(1-\gamma)\,p}.$$

The submarket share of producer i then is

$$\frac{x^i}{X}=\frac{1}{1-\gamma}\frac{p-\frac{1}{u_i}}{p}.$$

Summing up the submarket shares of all producers to eliminate x^i and X: $1=\sum_{i=1}^{N}\frac{x^i}{X}=\frac{1}{1-\gamma}\frac{Np-\sum_{i=1}^{N}\frac{1}{u_i}}{p}$ and solving for p leads to

$$p=\frac{\sum_{j=1}^{N}\frac{1}{u_j}}{N-1+\gamma}.$$

By the above equations the submarket share of producer i can be solved as $\frac{x^i}{X}=\frac{1}{1-\gamma}\left(1-\frac{N-1+\gamma}{\sum_{j=1}^{N}\frac{u_i}{u_j}}\right)$.

The mark-up ratio of producer i is $\frac{p-\frac{1}{u_i}}{p}=\left(1-\frac{N-1+\gamma}{\sum_{j=1}^{N}\frac{u_i}{u_j}}\right)$.

The submarket total output in equilibrium then is

$$X=S\left(\frac{N-1+\gamma}{\sum_{j=1}^{N}\frac{1}{u_j}}\right)^{\frac{1}{1-\gamma}}.$$

The total revenue of the submarket turns out to be

$$pX=S\left(\frac{N-1+\gamma}{\sum_{j=1}^{N}\frac{1}{u_j}}\right)^{\frac{\gamma}{1-\gamma}}.$$

Finally the reduced form profit function of producer i is derived as

$$\begin{aligned}
\pi_i\left(u_i \mid u_{-i}\right) &= \left(p-\frac{1}{u_i}\right)x^i = S\left(1-\frac{N-1+\gamma}{\sum_{j=1}^N \frac{u_i}{u_j}}\right)^2 \frac{1}{1-\gamma}\left(\frac{N-1+\gamma}{\sum_{j=1}^N \frac{1}{u_j}}\right)^{\frac{\gamma}{1-\gamma}} \\
&= \underbrace{S\left(\frac{N-1+\gamma}{\sum_{j=1}^N \frac{1}{u_j}}\right)^{\frac{\gamma}{1-\gamma}} \frac{1}{1-\gamma}}_{\text{total sales}} \underbrace{\left(1-\frac{N-1+\gamma}{\sum_{j=1}^N \frac{u_i}{u_j}}\right)}_{\text{submarket share of producer } i} \underbrace{\left(1-\frac{N-1+\gamma}{\sum_{j=1}^N \frac{u_i}{u_j}}\right)}_{\text{mark-up ratio of producer } i}.
\end{aligned}$$

Proof of Proposition 12

By eq. (3.8) and eq. (3.9) the following quadratic equation can be derived: $N_2^2 - n_0 N_2 + \frac{2-\gamma}{2} = 0$, the largest positive solution of which is $N_2 = \frac{n_0+\sqrt{n_0^2-2(2-\gamma)}}{2}$. It is not difficult to verify that $\frac{\partial N_2}{\partial \gamma} < 0$.

Proof of Proposition 13

Combining eq. (3.10) and eq. (3.11) leads to the following equation:
$S\frac{1-\gamma}{N_1^2}\left(\left(1-\frac{1-\gamma}{N_1}\right)\left(\frac{\delta}{\mu}\right)^{\frac{1}{\beta}}\right)^{\frac{\gamma}{1-\gamma}} = \delta$. Taking the logarithm transformation of both sides of the above equation and manipulating it can lead to:
$\ln\left(S\left(1-\gamma\right)\right) - \frac{\gamma}{1-\gamma}\frac{1}{\beta}\ln\mu - 2\ln N_1 + \frac{\gamma}{1-\gamma}\ln\left(1-\frac{1-\gamma}{N_1}\right) = \left(1-\frac{\gamma}{1-\gamma}\frac{1}{\beta}\right)\ln\delta$. Differentiating both sides w.r.t. δ and reorganizing the equation reveals the following derivative: $\frac{\partial N_1}{\partial \delta} = -\frac{\left(1-\frac{\gamma}{1-\gamma}\frac{1}{\beta}\right)N_1(N_1-1+\gamma)}{2N_1-2+\gamma}$. It is easy to see that $\frac{\partial N_1}{\partial \delta} < 0$ when $\beta > \frac{\gamma}{1-\gamma}$.

Appendix C

For Chapter 4

The Second Order Condition

Here we check on the related second order condition:

$$\begin{aligned}
\frac{\partial^2 Y}{\partial k^2} &= 2S\left[\frac{N^4}{(kN+1)^4} - \left(1 - \frac{N}{kN+1}\right)\frac{2N^3}{(kN+1)^3}\right] - \frac{\beta(\beta-1)}{k^2}\left(\mu' k\overline{u}\right)^\beta \\
&= 2S\left[\frac{N^4}{(kN+1)^4} - \left(1 - \frac{N}{kN+1}\right)\frac{2N^3}{(kN+1)^3}\right] - \frac{\beta(\beta-1)}{k^2}S\left(1 - \frac{N}{kN+1}\right)^2 \\
&= -\frac{S}{k^2(kN+1)^4}\left[2N^3k^2(2kN+2-3N) + \beta(\beta-1)(kN+1-N)^2(kN+1)^2\right].
\end{aligned}$$

Figure C.1 shows that $\frac{\partial^2 Y}{\partial k^2} < 0$ for $\beta > 1$, which confirms that the sufficient condition for maximization is satisfied.

Possible Response by the Incumbents

If the incumbents are allowed to respond after they observe the new entry and the entrant's quality level $k\overline{u}$, a fourth stage should be added to the game as follows:

- *Stage 4*: After observing the new entrant's quality level, the incumbents simultaneously choose to adjust their quality levels.

We want to know if this makes the outcome different from the case where the incumbents are not allowed to change their quality levels. If this does change the

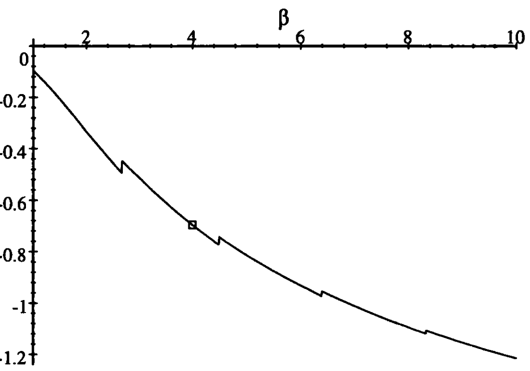


Figure C.1: $\frac{\partial^2 Y}{\partial k^2}\frac{1}{S}$

outcome, then we must be able to see an incumbent wants to deviate from the outcome of that case. Given other players' strategies the same as in the equilibrium in that case, any incumbent's payoff function is

$$y\left(k_1\overline{u} \mid (\overline{u}), k\overline{u}\right) = S\left(1 - \frac{N}{k_1(N-1) + \frac{k_1}{k} + 1}\right)^2 - (k_1\overline{u})^\beta + (\overline{u})^\beta, \ k_1 \geq 1.$$

The marginal payoff of increasing k_1 at the point $k_1 = 1$ is

$$\begin{aligned}
&\left.\frac{\partial}{\partial k_1} y\left(k_1\overline{u} \mid (\overline{u}), k\overline{u}\right)\right|_{k_1=1} \\
&\left.\frac{\partial}{\partial k_1}\left(S\left(1 - \frac{N}{k_1(N-1) + \frac{k_1}{k} + 1}\right)^2 - (k_1\overline{u})^\beta + (\overline{u})^\beta\right)\right|_{k_1=1} \\
= \ & 2S\left(\left(1 - \frac{N}{N + \frac{1}{k}}\right)\frac{N}{(N + \frac{1}{k})^2}\left(N - 1 + \frac{1}{k}\right) - \beta\overline{u}^\beta\right). \qquad \text{(C.1)}
\end{aligned}$$

Combining eq. (4.2) and (C.1) to eliminate $\beta\overline{u}^\beta$ results in

$$\begin{aligned}
&\left.\frac{\partial}{\partial k_1} y\left(k_1\overline{u} \mid (\overline{u}), k\overline{u}\right)\right|_{k_1=1} \\
= \ & 2S\left(\left(1 - \frac{N}{N + \frac{1}{k}}\right)\frac{N}{(N + \frac{1}{k})^2}\left(N - 1 + \frac{1}{k}\right) - \frac{N^2 - 2N + 1}{N^3}\right) \\
= \ & 2Sf(\beta)
\end{aligned}$$

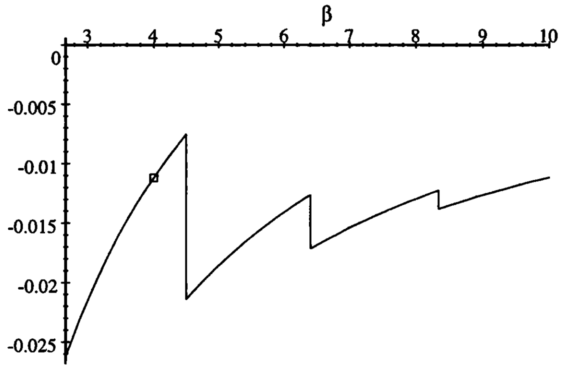


Figure C.2: $f(\beta)$

where $f(\beta) \equiv \left(1 - \frac{N}{N+\frac{1}{k}}\right) \frac{N}{\left(N+\frac{1}{k}\right)^2} \left(N - 1 + \frac{1}{k}\right) - \frac{N^2-2N+1}{N^3}\bigg|_{N=n(\beta),k=K(n(\beta),\beta)}$.

Figure C.2 shows that $f(\beta) < 0$ for $\beta > \frac{8}{3}$, which implies that

$$\frac{\partial}{\partial k_1} y\left(k_1\overline{u} \mid (\overline{u}), k\overline{u}\right)\bigg|_{k_1=1} < 0,$$

i.e., increasing quality is not an optimal response by any incumbent. Given that the incumbents' R&D costs have been sunk, they can not be saved by decreasing quality, i.e., the constraint $k_1 \geq 1$ must be satisfied, the optimal solution is the corner solution $k_1 = 1$, i.e., no response in quality.

Remark 28 *In the basic model, when $\beta > \frac{8}{3}$, even if the incumbents are allowed to respond to the entrant's quality choice, the equilibrium outcome is exactly the same as if the incumbents were not allowed to respond.*

Appendix D

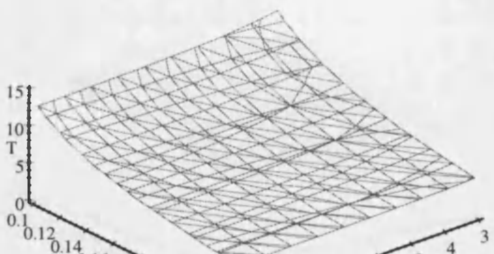
For Chapter 5

More Numerical Examples

Example 29 *Set $L = 1, \alpha = 2/5, \rho = 3/100$, then*

$$g = \delta - \frac{3}{40}\left(1 - \frac{1}{\left(\sqrt[3]{\left(\frac{N}{N-\frac{3}{5}}\right)}\right)^{2}}\right)\frac{N}{1-\left(\frac{N}{N-\frac{3}{5}}\right)^{-\frac{1}{50g}}},$$

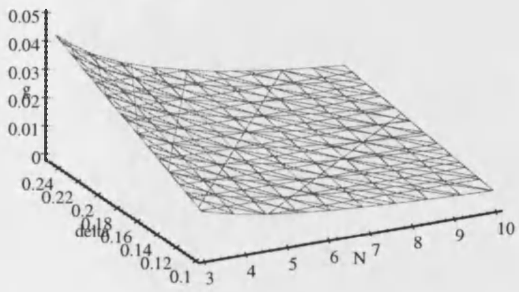
$$\frac{2}{3T}\ln\frac{N}{N-\frac{3}{5}} = \delta - \frac{3}{40}\left(1 - \frac{1}{\left(\sqrt[3]{\left(\frac{N}{N-\frac{3}{5}}\right)}\right)^{2}}\right)\frac{N}{1-\exp\left(-\frac{3}{100}T\right)} \quad \textit{hold.}$$

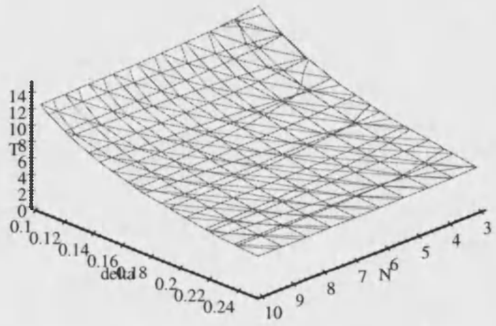


Example 30 *Set $L = 1, \alpha = 3/5, \rho = 3/100$, then*

$$g = \delta - \frac{1}{20}\left(1 - \frac{1}{\left(\sqrt{\left(\frac{N}{N-\frac{2}{5}}\right)}\right)^{3}}\right)\frac{N}{1-\left(\frac{N}{N-\frac{2}{5}}\right)^{-\frac{9}{200g}}},$$

$$\frac{3}{2T}\ln\frac{N}{N-\frac{2}{5}} = \delta - \frac{1}{20}\left(1-\frac{1}{\left(\sqrt{\left(\frac{N}{N-\frac{2}{5}}\right)}\right)^{3}}\right)\frac{N}{1-e^{-\frac{3}{100}T}} \textit{ hold.}$$

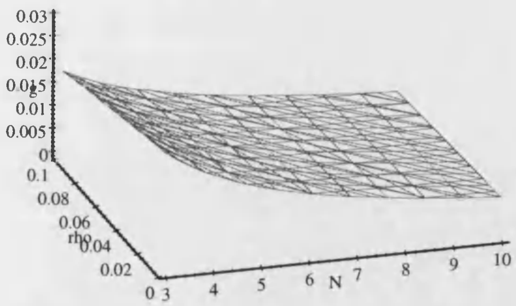


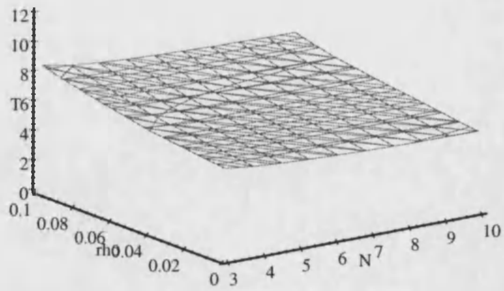


Example 31 *Set* $L=1, \alpha = 2/5, \delta = 1/5,$ *then*

$$g = \frac{1}{5} - \frac{5}{2}\left(1-\frac{1}{\left(\sqrt[3]{\left(\frac{N}{N-\frac{3}{5}}\right)}\right)^{2}}\right)\rho\frac{N}{1-\left(\frac{N}{N-\frac{3}{5}}\right)^{-\frac{2}{3}\frac{\rho}{g}}},$$

$$\frac{2}{3T}\ln\frac{N}{N-\frac{3}{5}} = \frac{1}{5} - \frac{5}{2}\left(1-\frac{1}{\left(\sqrt[3]{\left(\frac{N}{N-\frac{3}{5}}\right)}\right)^{2}}\right)\rho\frac{N}{1-e^{-\rho T}} \textit{ hold.}$$

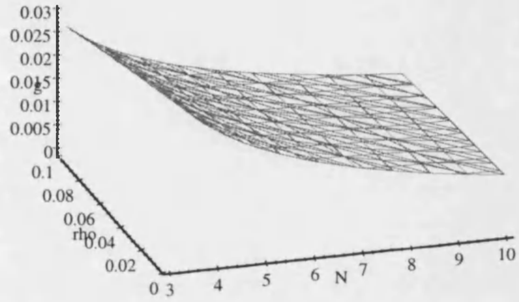


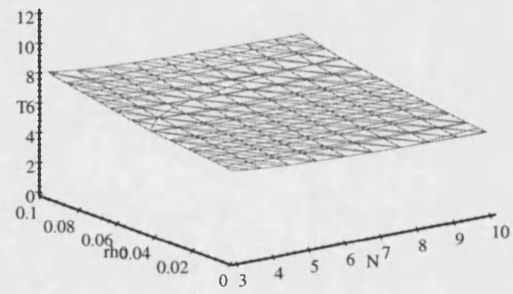


Example 32 *Set* $L=1, \alpha = 3/5, \delta = 1/5,$ *then*

$$g = \frac{1}{5} - \frac{5}{3}\left(1-\frac{1}{\left(\sqrt{\left(\frac{N}{N-\frac{2}{5}}\right)}\right)^{3}}\right)\rho\frac{N}{1-\left(\frac{N}{N-\frac{2}{5}}\right)^{-\frac{3}{2}\frac{\rho}{g}}},$$

$\frac{3}{2T}\ln\frac{N}{N-\frac{2}{5}}=\frac{1}{5}-\frac{5}{3}\left(1-\frac{1}{\left(\sqrt{\left(\frac{N}{N-\frac{2}{5}}\right)}\right)^{3}}\right)\rho\frac{N}{1-e^{-\rho T}}$ *hold.*





The Bounds to the Duration of the Immunity

The Lower Bound

$$\lim_{N\to\infty}\frac{\alpha}{(1-\alpha)\,T}\ln\frac{N}{N-1+\alpha}=\lim_{N\to\infty}\left(L\delta-\frac{\left(1-\left(\frac{N}{N-1+\alpha}\right)^{-\frac{\alpha}{1-\alpha}}\right)\rho N}{\alpha\left(1-e^{-\rho T}\right)}\right)$$

$$\lim_{N\to\infty}\frac{\alpha}{(1-\alpha)\,T}\ln\frac{N}{N-1+\alpha}=0$$

$$\lim_{N\to\infty}\left(L\delta-\frac{\left(1-\left(\frac{N}{N-1+\alpha}\right)^{-\frac{\alpha}{1-\alpha}}\right)\rho N}{\alpha\left(1-e^{-\rho T}\right)}\right)=\frac{L\delta e^{\rho T}-L\delta-\rho e^{\rho T}}{e^{\rho T}-1}$$

$$\frac{L\delta e^{\rho T}-L\delta-\rho e^{\rho T}}{e^{\rho T}-1}=0$$

$$\inf_{N}T=\frac{\ln\left(\frac{L\delta}{L\delta-\rho}\right)}{\rho}\tag{D.1}$$

An Upper Bound We consider the oligopoly case where $N \geq 3$. Therefore substituting $N = 3$ into eq. (5.41) gives us following equation which determines the upper bound to T,

$$3\left(1-\left(\frac{3}{2+\alpha}\right)^{-\frac{\alpha}{1-\alpha}}\right)\frac{\rho}{\alpha\left(1-e^{-\rho T}\right)} = L\delta - \frac{\alpha}{(1-\alpha)\,T}\ln\frac{3}{2+\alpha}, \tag{D.2}$$

which also implies

$$T = \frac{\frac{\alpha}{(1-\alpha)}\ln\frac{3}{2+\alpha}}{L\delta - 3\left(1-\left(\frac{3}{2+\alpha}\right)^{-\frac{\alpha}{1-\alpha}}\right)\frac{\rho}{\alpha(1-e^{-\rho T})}}. \tag{D.3}$$

Since $T \geq \inf_N T = \frac{\ln\left(\frac{L\delta}{L\delta-\rho}\right)}{\rho}$, the following inequality holds,

$$\frac{\alpha}{(1-\alpha)}\ln\frac{3}{2+\alpha} \geq \left(L\delta - 3\left(1-\left(\frac{3}{2+\alpha}\right)^{-\frac{\alpha}{1-\alpha}}\right)\frac{\rho}{\alpha\left(1-e^{-\rho T}\right)}\right)\frac{\ln\left(\frac{L\delta}{L\delta-\rho}\right)}{\rho},$$

which further implies that there exists an upper bound to T such that

$$T \leq \frac{\ln\left(\frac{\alpha\left(L\delta-\frac{\alpha\rho}{(1-\alpha)}\frac{\ln\frac{3}{2+\alpha}}{\ln\left(\frac{L\delta}{L\delta-\rho}\right)}\right)}{\alpha\left(L\delta-\frac{\alpha\rho}{(1-\alpha)}\frac{\ln\frac{3}{2+\alpha}}{\ln\left(\frac{L\delta}{L\delta-\rho}\right)}\right)-3\left(1-\left(\frac{3}{2+\alpha}\right)^{-\frac{\alpha}{1-\alpha}}\right)\rho}\right)}{\rho}.$$

Overall, the following holds

$$\frac{\ln\left(\frac{L\delta}{L\delta-\rho}\right)}{\rho} \leq T \leq \frac{\ln\left(\frac{\alpha\left(L\delta-\frac{\alpha\rho}{(1-\alpha)}\frac{\ln\frac{3}{2+\alpha}}{\ln\left(\frac{L\delta}{L\delta-\rho}\right)}\right)}{\alpha\left(L\delta-\frac{\alpha\rho}{(1-\alpha)}\frac{\ln\frac{3}{2+\alpha}}{\ln\left(\frac{L\delta}{L\delta-\rho}\right)}\right)-3\left(1-\left(\frac{3}{2+\alpha}\right)^{-\frac{\alpha}{1-\alpha}}\right)\rho}\right)}{\rho}.$$

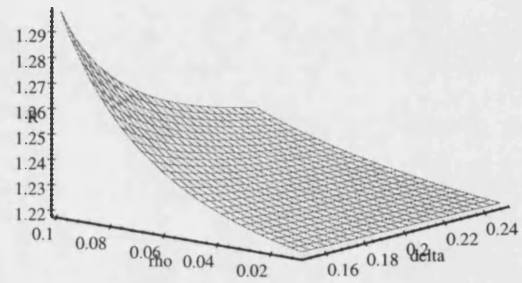
To get an idea of how close these two bounds can be, we look at the ratio of the upper bound to the lower bound:

$$R = \frac{\ln\left(\frac{\alpha\left(L\delta-\frac{\alpha\rho}{(1-\alpha)}\frac{\ln\frac{3}{2+\alpha}}{\ln\left(\frac{L\delta}{L\delta-\rho}\right)}\right)}{\alpha\left(L\delta-\frac{\alpha\rho}{(1-\alpha)}\frac{\ln\frac{3}{2+\alpha}}{\ln\left(\frac{L\delta}{L\delta-\rho}\right)}\right)-3\left(1-\left(\frac{3}{2+\alpha}\right)^{-\frac{\alpha}{1-\alpha}}\right)\rho}\right)}{\ln\left(\frac{L\delta}{L\delta-\rho}\right)},$$

and go through a few numerical examples.

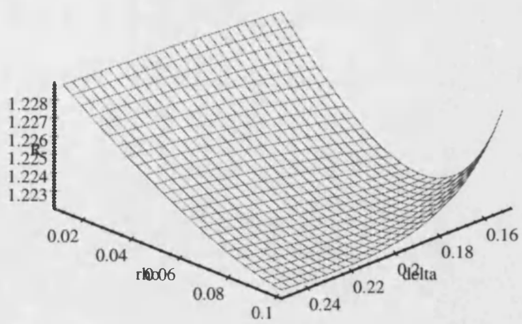
Example 33 *Set $L = 1, \alpha = 2/5$, then*

$$R = \frac{\ln \frac{2}{5} \frac{\delta - \frac{2}{3}\rho \frac{\ln \frac{5}{4}}{\ln \frac{\delta}{\delta-\rho}}}{\frac{2}{5}\delta - \frac{4}{15}\rho \frac{\ln \frac{5}{4}}{\ln \frac{\delta}{\delta-\rho}} - 3\left(1 - \frac{1}{5}\sqrt[3]{5}\left(\sqrt[3]{4}\right)^2\right)\rho}}{\ln \frac{\delta}{\delta-\rho}} \text{ holds.}$$



Example 34 *Set $L = 1, \alpha = 3/5$, then*

$$R = \frac{\ln \frac{3}{5} \frac{\delta - \frac{3}{2}\rho \frac{\ln \frac{15}{13}}{\ln \frac{\delta}{\delta-\rho}}}{\frac{3}{5}\delta - \frac{9}{10}\rho \frac{\ln \frac{15}{13}}{\ln \frac{\delta}{\delta-\rho}} - 3\left(1 - \frac{13}{225}\sqrt{15}\sqrt{13}\right)\rho}}{\ln \frac{\delta}{\delta-\rho}} \text{ holds.}$$



Remark 35 *The above numerical examples show that the upper bound and the lower bound to T can be quite close to each other therefore the range for T to change is quite limited.*

Bibliography

[1] Abernathy, William J. and James M. Utterbach (1978), "Patterns of Industrial Innovation", Technology Review 80: 40-47.

[2] Aghion, Philippe and Peter Howitt (1992), "A Model of Growth through Creative Destruction", Econometrica, 60, 2 (March), 323-351.

[3] Aghion, Phillipe and Peter Howitt (1998), Endogenous Growth Theory, Cambridge, Massachusetts: The MIT Press.

[4] Arthur, W. Brian (1994), Increasing Returns and Path Dependence in the Economy, Ann Arbor: The University of Michigan Press.

[5] Barro, Robert J. and Xavier Sala-I-Martin (1995), Economic Growth, New York: McGraw-Hill, Inc.

[6] Chandler, Alfred D. (1990), Scale and Scope, Cambridge, Massachusetts: The Belknap Press of Harvard University Press.

[7] Cohen, Wesley M. and Richard C. Levin (1989), "Empirical Studies of Innovation and Market Structure", Ch 18, Handbook of Industrial Organization, Volume II, Ed. by R. Schmalensee and R. D. Willig, Elsevier Science Publishers B. V..

[8] Dasgupta, P. and J. Stiglitz (1980), "Industrial Structure and the Nature of Innovation Activity". Economic Journal, 90: 266-293.

[9] Dixit, A. K. and J. Stiglitz (1977), "Monopolistic Competition and Optimum Product Diversity", The American Economic Review 67(3): 297-308.

[10] Dosi, Giovanni (1982) "Technological paradigms and technological trajectories", Research Policy, vol. 11, pp147-162.

[11] Dosi, Giovanni (1988), "Sources, Procedure, and Microeconomic Effects of Innovation" Journal of Economic Literature, vol 26, issue 3, Sept. pp 1120-1171.

[12] Grossman, Gene M. and Elhanan Helpman (1991), "Quality Ladders in the Theory of Economic Growth", Review of Economic Studies LVIII 43-61.

[13] Hellwig, Martin and Andreas Irmen (1999), "Endogenous Technical Change in a Competitive Economy", mimeo, University of Mannheim.

[14] Hopenhayn, Hugo A. (1993), "The Shakeout", Economics Working Paper 33, Universitat Pompeu Fabra.

[15] Horvath, Michael, Fabiano Schivardi and Michael Woywode (1997), "On Industry Life-Cycle: Delay, Entry, and shake-out in Beer Brewing", Mimeo, Stanford University.

[16] Jovanovic, Boyan (1982), "Selection and the Evolution of Industry", Econometrica, vol. 50, no. 3, pp. 327-38.

[17] Jovanovic, Boyan and Glenn M. MacDonald (1994), "The Life Cycle of a Competitive Industry", Journal of Political Economy, vol. 102, no. 2, 322-47.

[18] King, Robert G. and Ross Levine (1993a), "Finance and Growth: Schumpeter Might Be Right", Quarterly Journal of Economics CVIII 717-738.

[19] King, Robert G. and Ross Levine (1993b), "Finance, Entrepreneurship and Growth: Theory and Evidence", Journal of Monetary Economics 32: 513-542.

[20] Kiyotaki, Nobuhiro and John Moore (1997), "Credit Cycles", Journal of Political Economy vol.105, no.2: 211-248.

[21] Klepper, Steven and Elizabeth Graddy (1990), "The Evolution of New Industries and the Determinants of Market Structure", Rand Journal of Eco-

nomics, 21(1), pp. 27-44.

[22] Klepper, Steven (1996), "Entry, Exit, Growth, and Innovation Over the Product Life Cycle", American Economic Review, 86(3), pp. 562-83.

[23] Klepper, Steven (1999), "Firm Survival and the Evolution of Oligopoly", Mimeo, Carnegie Mellon University.

[24] Kuhn, Thomas (1962), The Structure of Scientific Revolutions, Chicago: Chicago University Press.

[25] Kumar, Krishna B., Raghuram G. Rajan and Luigi Zingales (1999), "What Determines Firm Size?", NBER Working Paper 7208, Cambridge, Massachusetts.

[26] Lakatos, I. (1980) The Methods of Scientific Research Programmes, Cambridge: Cambridge University Press.

[27] Levin, R. C., W. M. Cohen and D. C. Mowery (1985), "R&D Appropriability, Opportunity, and Market Structure: New Evidence on Some Schumpeterian Hypothesis", American Economic Review Proceedings, 75: 20-24.

[28] Levin, R. C., A. K. Klevonick, R. R. Nelson and S. Winter (1987), "Appropriating the Returns from Industrial R and D." Brookings Papers on Economic Activity, 3: 783-820.

[29] Mas-Colell, Andreu, Michael D. Whinston and Jerry R. Green (1995), Microeconomic Theory, Oxford: Oxford University Press.

[30] Milgrom, Paul and John Roberts (1992), Economics, Organization and Management, London: Prentice-Hall International, Inc.

[31] Nadiri, M. Ishaq (1993), "Innovations and Technological Spillovers", NBER Working Paper 4423, Cambridge, Massachusetts.

[32] Nelson, Richard and Sidney G. Winter (1982) An Evolutionary Theory of Economic Change. Cambridge, MA: Belknap Press of Harvard University Press.

[33] Nelson, Richard (1994), "Economic growth via the coevolution of technology and institutions", in Loet Leydesdortff and Peter Van Den Besselar, ed., Evolutionary Economics and Chaos Theory, London: Pinter Publishers.

[34] National Science Board (US), Science and Engineering Indicators: 2000.

[35] Phillips, A. (1971), Technology and Market Structure. Lexington, Mass: D. C. Health.

[36] Rajan, Raghuram G. and Luigi Zingales (1996), "Financial Dependence and Growth", NBER Working Paper 5758, Cambridge, Massachusetts.

[37] Rajan, Raghuram G. and Luigi Zingales (1999), "Financial Systems, Industrial Structure, and Growth", mimeo, University of Chicago.

[38] Reinganum, Jennifer F. (1989), "The Timing of Innovation: Research, Development, and Diffusion", Ch 14, Handbook of Industrial Organization, Volume I, Ed. by R. Schmalensee and R. D. Willig, Elsevier Science Publishers B. V..

[39] Romer, Paul (1990), "Endogenous Technological Change", Journal of Political Economy 98(5) part 2: 71-102.

[40] Scherer, F. M. (1967), "Market Structure and Employment of Scientists and Engineers", American Economic Review, 57: 524-531.

[41] Schumpeter, J. A. (1934), The Theory of Economic Development, Cambridge, Mass: Harvard University Press.

[42] Schumpeter, J. A. (1942), Capitalism, Socialism, and Democracy. New York: Harper.

[43] Scott, J. T. (1984) "Firm versus Industry Variability in R&D Intensity", in Z. Griliches, ed., R&D, Patents, and Productivity. Chicago: University of Chicago Press for the National Burea of Economic Research.

[44] Shaked, A. and John Sutton (1983), "Natural oligopolies". Econometrica, 1469-1484.

[45] Shaked, A. and John Sutton (1987), "Product differentiation and industrial structure". Journal of Industrial Economics, 131-146.

[46] Sutton, John (1991), Sunk Costs and Market Structure, Cambridge, Mass: The MIT Press.

[47] Sutton, John (1997a), "Game Theoretic Models of Market Structure." Advances in Economics and Econometrics, edited by D. Kreps and K. Wallis. (Proceedings of the World Congress of the Econometric Society, Tokyo, 1995.) Cambridge: Cambridge University Press, pp. 66-86.

[48] Sutton, John (1997b), "Gibrat's Legacy", Journal of Economic Literature, vol. XXXV (March), pp. 40-59.

[49] Sutton, John (1998), Technology and Market Structure, Cambridge, Mass: The MIT Press.

[50] Teece, David J. (1988) "The Nature and the Structure of Firms" in Dosi et al., Technical Change and Economic Theory. London: Francis Pinter; New York: Columbia University Press.

[51] Tirole, Jean (1988), The Theory of Industrial Organization, Cambridge, Mass: The MIT Press.