The Dynamics of Market Structure and R&D Competition

Ana Sofia Domingues Rodrigues

PhD Thesis

The University of York

Department of Economics and Related Studies

May 2007

Contents

Abstract 8

Acknowledgements 11

Declaration 13

1 Introduction 14

2 Understanding Gibrat's Law with a Markov-Perfect Dynamic Industry Model 23

2.1 Introduction 23

2.2 Gibrat's Law and Empirical Findings 25

2.3 The Model 28

2.3.1 The Industry 30

2.3.2 Evolution of Market Structure 33

2.3.3 Dynamic Equilibrium 38

2.4 Firm Size Distribution 41

2.4.1 Simulation 41

2.4.2 Distribution Results 44

2.4.3 Cross-Sectional Growth-Size Properties 47

2.5 Variation in the Firm Size Distribution 51

2.5.1 Fixed Costs 52
2.5.2 Sunk Costs 54
2.5.3 Rate of Cost Reduction 57
2.5.4 Technological Opportunity 59
2.5.5 Rate of Spillovers 61
2.6 Conclusions 63
2.7 Appendix 98

3 The Endogenous Determination of Market Structure and R&D Spillovers **104**

3.1 Introduction 104
3.2 The Model 109
3.2.1 The Spot Market 109
3.2.2 Dynamic Processes 111
3.2.3 R& D and Appropriability 112
3.2.4 Dynamic Equilibrium 117
3.3 Numerical Results 131
3.3.1 Methodology 131
3.3.2 Costless R&D Spillovers 132
3.3.3 Absorptive Capacity 140
3.4 Conclusions 144

4 Monopolistic Competition and Dynamic R&D **165**

4.1 Aims and Motivation 165
4.2 Background 167
4.3 The Model 171

4.3.1 The Spot Market 172

4.3.2 The dynamic environment 176

4.3.3 The R&D Choice 180

4.3.4 Endogenous Growth 182

4.4 Solution Methodology 186

4.5 Firm Size Distribution 188

4.6 Conclusions 192

4.7 Appendix 202

5 Final Remarks **203**

Bibliography **209**

List of Tables

2.1 Parameter values 65

2.2 Distribution Statistics for Baseline (Natural Logs) 66

2.3 Distribution Statistics for Baseline Excluding Firms with <1 Employee (Natural Logs) 67

2.4 Average Summary Statistics in Natural Logs 68

2.5 Regression Results 69

2.6 Tests of Standard Deviation of Growth Rates by Size Class 70

3.1 Parameter Values 147

4.1 Parameter Values 194

List of Figures

2.1 US Firm Size Distribution by Employment, 2002 71

2.2 Employment Distribution 72

2.3 Log of Employment Distribution 73

2.4 Values Distribution 74

2.5 Log of Values Distribution 75

2.6 Sales Distribution 76

2.7 Log Sales Distribution 77

2.8 Fixed Costs Effects (Levels) 78

2.9 Fixed Costs Effects (Logs) 79

2.10 Effects of Fixed Costs on Firm Size Distribution (Levels) 80

2.11 Effects of Fixed Costs on Firm Size Distribution (Logs) 81

2.12 Sunk Costs Effects (Levels) 82

2.13 Sunk Costs Effects (Logs) 83

2.14 Effects of Sunk Costs on Firm Size Distribution (Levels) 84

2.15 Effects of Sunk Costs on Firm Size Distribution (Logs) 85

2.16 Rate of Cost Reduction Effects (Levels) 86

2.17 Rate of Cost Reduction Effects (Logs) 87

2.18 Effects of the Rate of Cost Reduction on Firm Size Distribution (Levels) 88

2.19 Effects of the Rate of Cost Reduction on Firm Size Distribution (Logs) 89
2.20 Technological Opportunity Effects (Levels) 90
2.21 Technological Opportunity Effects (Logs) 91
2.22 Effects of the Productivity of Investment on Firm Size Distribution (Levels) 92
2.23 Effects of the Productivity of Investment on Firm Size Distribution (Logs) 93
2.24 The Rate of Technological Spillovers Effects (Levels) 94
2.25 The Rate of Technological Spillovers Effects (Logs) 95
2.26 Effects of the Rate of Spillovers on Firm Size Distribution (Levels) 96
2.27 Effects of the Rate of Spillovers on Firm Size Distribution (Logs) 97

3.1 The Impact of Spillovers on the Probability of Innovation Under the Different Appropriability Scenarios 146
3.2 The Impact of Costless Spillovers on the Optimal Investment Policy Function 148
3.3 The Impact of Costless Spillovers on the Optimal Probability of Innovation 149
3.4 The Impact of Costless Spillovers on the Optimal Firm's Value . . 150
3.5 The Value Function for b=0.0 151
3.6 The Impact of Costless Spillovers on the Ergodic Distribution of Market Structures 152
3.7 The Impact of Costless Spillovers on the Concentration Index C1 153
3.8 The Impact of Costless Spillovers on Investment, Probability of Success and Firm's Value 154

3.9 The Impact of Costless Spillovers on Consumers' Surplus 155

3.10 The Impact of Costless Spillovers on Producer Benefits 156

3.11 The Impact of Spillovers on the Optimal Investment Policy Function in the Case of Absorptive Capacity 157

3.12 The Impact of Spillovers on the Optimal Investment Probability of Success in the Case of Absorptive Capacity 158

3.13 The Impact of Spillovers on Firm's Value in the Case of Absorptive Capacity 159

3.14 The Impact of Spillovers on the Ergodic Distribution of Market Structures in the Case of Absorptive Capacity 160

3.15 The Impact of Spillovers on the Concentration Index C1 in the Case of Absorptive Capacity 161

3.16 The Impact of Spillovers on Investment, Probability of Success and Firm's Value in the Case of Absorptive Capacity 162

3.17 The Impact of Spillovers on Consumers' Surplus in the Case of Absorptive Capacity 163

3.18 The Impact of Spillovers on Producer Benefits in the Case of Absorptive Capacity 164

4.1 The Change in the Marginal Cost Curve with PS 195

4.2 Firm's Value as a function of the Efficiency Level 197

4.3 Investment as a function of the Efficiency Level 198

4.4 The Distribution of Firm Sizes for a Sample of British Firms . . . 199

4.5 The Distribution of Employment 200

4.6 The Distribution of Ln Employment 201

Abstract

This thesis investigates the articulation of the incentives to perform Research and Development of profit seeking firms. Throughout the thesis, the dynamic evolution of the distribution of these incentives across firms is the engine of industry transformation and growth. Thus, in order to assess the impact of different industry characteristics on the market structure, we need a faithful picture of the context where firms make their R&D choices.

Chapter one exposes more in detail the motivation to pursue the analysis developed in each chapter independently, and how they combine to build up the search for the understanding of the interactions between R&D, appropriability and market structure.

Chapter two presents a dynamic model of the firm size distribution. Empirical studies of the firm size distribution often compare its moments to those of a log-normal distribution, as implied by Gibrat's Law, and note important deviations. Thus, the first and basic questions addressed in the first chapter are how well does the dynamic industry model reproduce Gibrat's Law and how well does it match the deviations uncovered in the literature. We show that the model reproduces these results when testing the simulated output using the techniques of the empirical literature. We then use the model to study how structural parameters affect the firm size distribution. We find that, among other things, fixed

and sunk costs increase both the mean and variance of the firm size distribution while generally decreasing the skewness and kurtosis. The rate of growth in an industry also raises the mean and variance, but has non-monotonic effects on the higher moments.

In the third chapter we explore the implications of different degrees of R&D appropriability on market structure and welfare. We propose a framework to pursue this analysis by extending the Markov-Perfect dynamic industry model proposed by Ericson and Pakes (E-P, henceforth) (1995) through the introduction of a non-proprietary productivity component to R&D as part of a dynamic, stochastic process. We first assume that spillovers are costlessly absorbed and exploited by firms in the industry, and find that, in this case, a free rider problem arises, thereby decreasing the incentives for investment. This leads to a lower amount of innovations being developed in the industry, which in turn, implies lower consumer welfare while leaving the degree of concentration in the industry fairly unchanged. We then model a setting where it is assumed that in order to build its absorptive capacity the firm has to engage in some R&D of its own. In this case, we find that an increase in spillovers will enhance both consumer and producer welfare substantially, and increases the likelihood of neck-and-neck competition, therefore reducing the level of concentration in the industry. These results arise from the fact that having absorptive capacity being built as a by-product of R&D enhances the productivity of R&D investment, compensating for the free rider effect associated with the lack of appropriability.

The frameworks used in the two first chapters suffer from the "curse of dimensionality", such that the industries under analysis are limited in terms of the number of agents simultaneously active. In order to overcome this problem, in chapter four we move away from oligopolistic market structures and propose a model of monopolistic competition, where firms are sufficiently large to generate a firm size distribution with a certain degree of asymmetry, although each firm is too small to affect the industry's outcome. Furthermore, we account for industry growth by having the industry's output increasing over time as a result of knowledge externalities. The rich micro set-up of this model is analogous to that of E-P (1995), as it is composed by heterogeneous firms making their investment decisions in a world of uncertainty, but we abstract from entry and exit and instead of an oligopolistic market structure we model a monopolistic competition environment with many, heterogeneous firms. In this setting, firms are asymmetric in terms of the technology they use to produce a given commodity, and they are able to increase the likelihood of decreasing their marginal costs of production by investing in Cost Reducing R&D. In order to evaluate their future stream of profits and make their investment decisions firms only care about the evolution of their efficiency and the long-run efficiency index in the industry. Cutting down the oligopolistic interactions present in the E-P framework, and having firms looking at the long-run average industry state, allows us to overcome the curse of dimensionality usually associated with dynamic models with agent heterogeneity. Therefore, we are able to simulate the model with a large number of firms competing in the industry and we show that, contrary to most existing endogenous growth models, this model is able to deliver a firm size distribution with a substantial degree of heterogeneity.

Chapter 6 presents the final remarks to the investigation carried out in this thesis.

Acknowledgements

I am grateful to *Fundação para a Ciência e a Tecnologia*, from Portugal, for the financial support provided.

I am very grateful to Huw Dixon for his supervision throughout these years, insightful ideas and discussions, that helped me reason on matters and construct some economic thought of my own. No matter where he was, from York to Singapore, he has always been available. I would also like to thank Huw for his optimism which has been undoubtedly crucial in some more worrying moments and his good mood and joy during this time, that made the whole thing a lot more enjoyable.

A very special thanks to Christopher Laincz. He helped me overcome one of the biggest obstacles I faced: dealing with the *"monster"* of computing in industrial dynamics with multi-agent heterogeneity. Working with Chris was a great pleasure and a privelige. The time I spent doing research with Chris in Philadelphia was extremely important for my career, and an enriching experience in my life. For this, I would like to thank not only Chris but Vibhas Madan, the director of the LeBow College of Business in Drexel University, and all the staff, for having welcomed me there.

I would also like to thank my Thesis Advisory Group, John Hutton and Gabriel Talmain, for their insights and support.

I am grateful to Ariel Pakes for his comments on my first chapter, during a course in Industry Dynamics in Helsinki. I would also like to thank John Sutton and all other seminar participants at the RES conference 2006, for helpful suggestions.

There are many friends that deserve a thank you in these pages for having helped me in one way or the other throughout this time. Thank you Ana, Iris, Carolina, Fakhri, Pelim and Giorgious for all the time we spent together in York or Philly. Thanks to Filomena Garcia and Alex Kostakis for their comments on my work.

A very special thanks to Marisa. Unconditional friendship and complicity, together with lots of fun and many easy going discussions in economics, is something hard to find, I am happy to experience that with you.

I am always grateful to João Pedro for his incomparable friendship and for motivating me more than anybody else to study economics.

To Rui, thank you for your help, your love and your patiente. I thank you for making my goals your goals, you literally helped me always and everywhere during this time as, actually, during all others.

Finally, there are not enough thanks to my family for having given me everything. They built for me a path of fulfillment, joy, intellectual and emotional growth, in every sense possible. For their invaluable support, I thank in particular to my parents, Zé[1] and Bruno. Pai e Mãe, I can just say this wouldn't be without you.

I dedicate this thesis to my grandfather, a unique person whose dignity and wonderful personal qualities still surprise me today, although only in memories. I wish I could still make you proud.

[1]Including a special thanks for the many "e-mails" and visits received.

Declaration

Chapter 2 was developed in co-authorship with Christopher Laincz of Drexel University, Philadelphia. This is the paper 05/34 of the series of Discussion Papers of the Department of Economics of the University of York and is currently at the revise and resubmit status at the RAND Journal of Economics. The joint work that culminated in the paper was initiated during my stay as a visiting research scholar at LeBow College of Business, Drexel University. An earlier version of the paper was presented in November, 2004 at the conference "Innovation, Entrepreneurship and Growth" at the Royal Institute of Technology in Stockholm. In the same year, a poster of this chapter was presented at the EC^2 conference "Econometrics of Industrial Organisation", in Marseille. The current version of chapter 2 was presented at the Royal Economic Society conference at Nothingham, in April 2006, and at the workshop series of the Department of Economics of *Universidade do Minho,* Braga (Portugal), in June 2006.

Chapter 3 was presented at the "2005 ESSID European Summer School in Industrial Dynamics", in Cargese, Corsica.

An earlier version of Chapter 4 was presented at the PhD students workshops of the Department of Economics and Related Studies of the University of York in June 2004.

Chapter 1

Introduction

"In the years after the Second World War the economist's attitude gradually changed. The vast expenditures on Research and Development made it increasingly obvious that inventive activity was - or could be made to be - responsive to economic needs (or even to non-economic needs if such needs received sufficient financial support). Clearly much of the search activity of R and D was highly purposive: business firms were looking for new techniques in specific categories of products, they spent much money upon this search, and they were sometimes highly successful."

Nathan Rosenberg, 1974

This thesis explores the dynamics of organic growth through technological process as an output of R&D effort at the corporate level and the role of the expansion of scientific knowledge in the dynamics of market structure. Technological change has assumed a primordial role in the search for the engines of growth and development, and throughout the past decades Research and Development has predominantly become the province of firms rather than the outcome of government efforts or stand alone inventors. Thus, the inventive activity and the growth of scientific knowledge should be accounted for in the context of firms' investment choices, which arise as a response to the economic environment determining the incentives to perform R&D.

We aspire at providing some insights on the nature of R&D competition, and on how it shapes the market structure of an industry. We believe that the market environment, captured by the demand and cost conditions, the market structure, the productivity of innovations in enhancing the attractiveness of a firm's product or in improving the efficiency in which it combines inputs to produce its output, and the extent to which a firm is able to appropriate the outcome of its own investment govern the allocation of the innovative effort in the industry, and thus the pace of technological growth.

In our attempt at modelling firms' innovative activity, we allow for uncertainty in the outcome of R&D effort. By increasing the amount of resources allocated to the R&D activity, a firm enhances its likelihood of developing an innovation but it also faces the possibility of failure. The importance of accounting for uncertainty when modelling R&D competition was highlighted by Kamien and Schwartz (1971) and Dasgupta and Stiglitz (1980). In their approach to R&D competition, Dasgupta and Stiglitz (1980) also highlight the other dimension determining the pace of R&D: the market structure. Given the conditions for the productivity and appropriability of R&D in the industry, the relative position of the players in the market is what determines the magnitude and distribution of research effort across firms. Following this path of research implies moving away from the traditional approach (Schumpeter (1942)) of modelling the connection between the degree of concentration and R&D unilaterally, addressing not only the role of market structure in affecting the incentives to invest, but also the endogenous determination of the market structure itself with respect to the R&D

program.

In the approach presented here, product market competition generates incentives for profit maximising firms to engage in R&D competition, with the aim of potentially reducing the unit costs of production by improving its production process (process innovations). The uncertain outcome of the innovative R&D activity will, in turn, determine the innovation path of the firm and its profitability in the spot market. The returns to R&D depend not only on the firm's success in innovating, but also on the outcome of the rivals' R&D effort. A firm's competitiveness at a given moment in time depends on the path of both its own and its rivals' past successes in the R&D activity.

In the first chapter, we investigate the economic forces shaping the market structure in the industry. Given the dynamic character of R&D competition and the relevance of accounting for heterogeneity at the firm level and turbulence when modelling an industry environment, we depart from the E-P (1995) dynamic industry framework, which is able to replicate the variability of similar firms fates empirically observed. This framework also accounts for the endogeneity of market structure with respect to R&D competition, which is essential to the task of exploiting how industry conditions and R&D rivalry affect the distribution of the incentives to engage in R&D when its outcome is uncertain.

The distribution of market shares across the set of firms operating in an industry and the asymmetry in firm size have long been an issue of interest in economic theory. The first approach to this issue was conducted in 1931 by a French economist, Robert Gibrat, who proposes a simple model to explain the statistical properties of the distributions of a number of realities such as *"l'inégalites des richesses, la concentration des entreprises, la population des villes"*, among others. Gibrat presents intuitively the importance of his results: "Wouldn't it be astonishing the existence of a general tendency of an increase to 11 workers in firms with ten, and an increase to 1.100 in those with 1.000 workers?" . This intuitive approach was formalised as Gibrat's *law of proportionate effect,* according to which the relative effect of industrial fluctuations in the number of workers of a firm (Gibrat takes the number of worker as a proxy for firm size), or their absolute effect on the log of the size of the firm does not depend on the number of workers of the firm, such that the distribution of firm sizes can be well approximated by a lognormal distribution. This subject has attracted much attention since the late 50's, and has motivated many empiricists in testing the validity of Gibrat's statements.

The first chapter of this thesis, however, is not related with the testing of Gibrat's law, but rather with the search of the economic forces that shape the distribution of firm sizes and make the law of the proportionate effect a reasonable, yet imperfect approximation for the firm size distribution. What are the economic forces influencing the higher moments of firm size distribution? How do industry characteristics affect them? These research questions gain further interest given the results of empirical studies uncovering cross-industry variation in the higher moments of the firm size distribution such as Machado and Mata (2000), Lotti and Santerelli (2004) and Audretsch et al. (2004).

We tackle these research questions through the attempt of unraveling how industry characteristics, captured by a set of parameters in our E-P based model, affect the expected returns to R&D investment at the level of product market competition and R&D rivalry. In a first stage, we demonstrate that the model is able to generate a set of simulated data for firm sizes with similar statistical properties to those of the real data. We then vary the values of the parameters capturing the industry characteristics whose impact on the distribution of firm sizes we want to analyse. We find that the firm size distribution in industries with higher fixed or sunk costs are flatter and have larger mean size of firms. Fixed costs make it more difficult for firms to survive in the industry and act as a requirement of a minimum scale of operations to produce non-negative profits and survive, implying that the less efficient firms will be better off by exiting the industry and receiving the scrap value, while sunk costs act as entry barriers and discouraging potential players to enter the industry. As one would expect, both a higher technological progress and a higher productivity of R&D in the industry will lead to higher mean size and variance, but we found the impact of these industry characteristics on the higher moments of the firm size distribution to be non monotonic. Finally, we should expect that the higher the appropriability of R&D in the industry, the higher the mean and variance in the size of firms, but the effects on skewness and kurtosis are highly non-monotonic. These theory-driven results are in line with many of the findings of the empirical literature on cross industry variation in the firm size distribution, and yield a series of testable hypothesis on the nature and direction of the effects of key industry characteristics on the properties of the distribution of firm sizes.

In the third chapter of this thesis, we address the impact of cost-reducing R&D spillovers on the evolution of the ergodic distribution of market structures. Our aim is to unravel the dynamic forces driving the market structure changes brought about by the presence of externalities, and understand the relationship between the market structure and the trade off between the damage to the incentives for R&D effort and the reduction in wasteful duplication. The contribution of this piece of research lies on the analysis of the impact of externalities on market structure and welfare in a dynamic, stochastic context, given that the investigations carried out in the literature rest on static, symmetric models, or models that fail to account for the uncertainty in firms' R&D decisions.

In the existing literature concerning the effects of R&D externalities, it is usually assumed that the external effect of each firm's R&D is to lower the rival firms' unit cost of production (when modelling process innovations) or increase the quality of the competitors' product (product innovations). In chapter three, we propose a specification which incorporates uncertainty in the outcome of R&D such that the external effect of a firm's R&D is to enhance the likelihood of a rival firm experiencing a reduction in its marginal cost of production. Therefore, we eliminate the determinism in the relationship between R&D effort and marginal costs. Higher R&D does not ensure successful innovation, it rather delivers a more favourable innovation path over time.

We first introduce the conventional form of spillovers modelled in the literature- *Costless R&D Spillovers.* Under this setting, there are benefits of the R&D undertaken to firms other than the one who has bore its cost. A portion of the R&D effort by each firm flows to the public pool of knowledge, becoming readily available to the other firms. The proportion of the R&D effort leaking to the public pool of R&D is determined by the appropriability conditions in the industry such as patent policies, secrecy (particularly relevant for process innovations), lead-time, the extent of knowledge embodied in the output of the innovation process, the ease of imitation, worker mobility, etc. Under this scenario, firms substitute costly R&D effort by a free external source of R&D and cost optimization allows them to achieve higher values. However, the classical effect of R&D externalities of potentially improving the rival's state decreases the incentives to invest in R&D. Fewer innovations are developed in the industry and consumers will experience a welfare loss from reduced appropriability.

These implications of spillovers, however, are sensible to the assumption of

R&D spillovers being a pure public good. If one assumes, such as in Cohen and Levinthal (1989), that the absorption and exploitation of rivals' R&D spillovers is not a costless process, but that it rather depends on a crucial element, this being the firms' absorption capacity, built as a by product of R&D, the implications of R&D spillovers for market structure and welfare change dramatically. In this scenario, the productivity of R&D increases with an offsetting effect on the reduction of the incentives to perform R&D that arise from the free rider problem. Even with a reduction in the R&D investment undertaken in the industry, specially in the case of the leader firm (the firm with an advantage in the efficiency level), the presence of externalities improves the total amount of R&D devoted to the innovative process, delivering a more favourable innovation path which implies lower marginal costs in the industry. Consequently, consumer welfare will increase, but at the producer level, only the follower firm will experience a welfare gain. The improvement in the follower position relative to the leader reduces the asymmetry in the market structure.

Thus, we show that spillovers do not necessarily have a detrimental effect on R&D investment incentives and that the Schumpeterian results are weakened when imposing the need of an absorption capacity, built as a by product of R&D investment, in order for firms to absorb and exploit the spillovers. In fact, the literature on this subject has recurrently treated external information as costlessly absorbed, and with no other countervaling effects, the free riding problem dooms the lack of appropriability to be welfare reducing.

Dynamic industry models with agent heterogeneity can be very enlightening for the understanding of a number of important dimensions related to firm behaviour and market evolution, such as the ones addressed in chapters two and three. However, their usage is limited by the computational burden involved in computing rational expectations concerning the expected future states of the set of players in the industry. This *"curse of dimensionality"* associated with the dynamic framework used throughout the thesis prevents its use to study industries with more than a handful of firms. The dimensions of agent heterogeneity are also restricted by the computational burden.

There have been a series of attempts do overcome this limitation of the modelling of dynamic strategic interaction, namely a version of the Pakes and McGuire (2001) algorithm that computes expectations over the states of the ergodic set and ignoring the rest of all possible states and improves the computational efficiency. Dorazelski and Judd (2005) also propose a continuous time alternative which reduces to zero the probability that the state of two firms change simultaneously, improving the computational speed. In the fourth chapter of this thesis, we propose an alternative dynamic industry model where each firm is assumed too small to affect the industry outcome, but large enough so that its size is significant. As a result, each firm, individually, is too small to have a strategic effect on its rivals' decisions, and only the average industry state affects firms' decision. Furthermore, we simplify firms' decision process by having them treating the long run average of the industry as constant. Thus, in computing their perceptions of the evolution of the future payoffs, firms only care about the expected long-run average state of the industry and the dynamics of their own state, whose motion is given by the firm's own transition function.

Cutting down the oligopolistic interactions present in the original E-P (1995) framework allows to model industries with many heterogeneous agents. Firms will invest in R&D in order to improve their efficiency, and their decisions are a function of the long run average industry state and the probability dynamics of their own state. We propose an algorithm and the corresponding code to find the optimal R&D investment decisions. For doing this we have the firms forecasting the long run average industry state, which ultimately affects its profits. Given that expectation, we find the optimal R&D decision for each of the possible states each firm can find itself in and, using those, we simulate the industry market structure to obtain actual series of states for the firms in the industry. We perform this process until the expected long run industry average that firms assume as given for their optimization process is the actual long run average industry state that arises when simulating the industry's market structure using the optimal policy functions. We show that the equilibrium long run distribution of firm sizes entails a substantial degree of heterogeneity.

Chapter 2

Understanding Gibrat's Law with a Markov-Perfect Dynamic Industry Model

2.1 Introduction

Studies on the firm size distribution and Gibrat's Law to date have been the province of empiricists. We can write down various reduced form models, as in McCloughan (1995), to reproduce many of the statistical facts surrounding the firm size distribution and Gibrat's Law of Proportionate Effect, which states that the growth rate of a firm is independent of its size, and the well know deviations from this law found in the empirical literature. However, little of the empirical work has been guided by a formal structural model. In Caves' (1998) survey on the recent empirical findings in industrial organization, he states, "Although the importance of these facts for economic behaviour and performance is manifest, their development has not been theory-driven." This paper seeks to take a step towards filling this gap.

We employ an extension of the Ericson and Pakes (1995) model of a dynamic industry that allows for firm growth developed by Laincz (2004a). By varying key priors, the simulations demonstrate potential sources for the various, and sometimes conflicting, results on Gibrat's Law uncovered in the empirical literature. We demonstrate that the model matches empirical findings on Gibrat's Law.

A more recent literature uncovers significant cross-industry variation in the higher moments of the firm size distribution. Machado and Mata (2000) find that industry characteristics such as technological orientation and capital-intensity are significantly related to the skewness. Lotti and Santerelli (2004) show how the distribution of a new cohorts differs across different industries and over time. Audretsch et al. (2004) present evidence suggesting that the firm size distribution of the service industry differs from manufacturing. We use the model to develop theoretical reasoning for many of these findings, however, our analysis also emphasizes that some variables have strong non-monotonic effects on the moments of the firm size distribution suggesting caution in generalizing empirical results based on linear specifications.

After briefly reviewing the lengthy empirical literature on Gibrat's Law and its relationship to the firm size distribution in the next section, section 2.3 presents the basic model. In section 2.4 we compare the results of a baseline simulation to the empirical literature on the firm size distribution and Gibrat's Law. Section 2.5 then documents how varying key structural parameters alters the firm size distribution. Section 2.6 summarizes the results.

2.2 Gibrat's Law and Empirical Findings

Following the seminal works of Hart and Prais (1956) and Ijiri and Simon (1964), the industrial organization literature devoted much energy into exploring the statistical regularity known as Gibrat's Law as it applies to the firm size distribution. Figure 2.1 shows the size distribution of enterprises for the U.S. in 2001. Notably, the distribution is significantly skewed to the right with the large peak for the smallest size class with non-zero employment. The following simple statistical process generates almost the same distribution. Let x_i be the size of firm i, then growth from one period to the next is represented as:

$$x_i(t) = x_i^{\beta}(t-1)\exp\left[u_i(t)\right], \beta > 0 \tag{2.1}$$

where $u_i(t) \backsim iid\ N(\mu, \sigma^2)$. Defining $y_i(t) = \ln x_i(t)$, then:

$$y_i(t) = \beta y_i(t-1) + u_i(t). \tag{2.2}$$

When $\beta = 1$ we have Gibrat's Law wherein the growth rate of a firm is independent of its size and the process yields a log-normal distribution of firm sizes. Empirical work on the firm size distribution finds that this characterization is a close, but imperfect proxy for the data. The earliest work on Gibrat's Law only had data available for large firms. Hart and Prais (1956), for example, included only firms listed on the London Stock Exchange between 1885 and 1950. They found that Gibrat's Law provided a good statistical approximation for the distribution. Simon and Bonini (1958) found similar results for large US firms.

More recently, Hart and Oulton (1996) compare the implications of (2.1) to a large sample of firms measured by employees, net sales, and net assets. They find that the distribution has a long right tail, with skewness coefficient estimates ranging from 0.19 to 0.75, and leptokurtic with values from 4.58 to 6.20. However, they argue the deviations should not be compared with the extreme of matching the log-normal distribution exactly and that the close approximation justifies the use of Gibrat's Law in empirical work.

Our task is rather different. We are specifically interested in the deviations themselves. We want to construct a sensible model of optimizing firm behaviour that can both approximate the distribution and provide us with a tool to understand the deviations and, moreover, cross-industry differences. Before turning to the model, we look at the literature that explicitly rejects the strong form of Gibrat's Law where β is exactly one.

Mansfield (1962) was perhaps the first to explicitly deal with the problems that entry and exit present for the interpretation of Gibrat's Law. Specifically, since exiting firms effectively have a growth rate of -100%, does Gibrat's Law hold for all firms, only the survivors, or for firms exceeding a size threshold such as minimum efficient scale? Of the three, he found that the latter interpretation fit his data the best using a χ^2 test on the lognormality of the distributions for each of his industries in each time period. In growth size regressions, Mansfield found that in the entire sample of survivors, firms grow less than proportionally, i.e. $\beta < 1$. However, analyzing large firms only, he found that the mean growth rate is independent of size, i.e. $\beta = 1$. He still concluded that Gibrat's law does not hold for any of the versions considered due to the fact that, even for the case of larger firms only, the variance of growth rates decreases with size.

Subsequent empirical analysis largely confirmed Mansfield's initial foray into the subject. Using more advanced econometric techniques to deal with heteroscedasticity and sample selection bias, Hall (1987) and Evans (1987) found that Gibrat's Law generally holds for large firms, but not for the entire population. They uncover a negative relationship between size and growth. Dunne and Hughes (1994) also find that while size evolves proportionally for medium and large firms, small firms' growth rates have higher variance and tend to decrease with size.

Another set of growth regression studies focused on the persistence of deviations of firm size from the mean, which would imply biased estimates for β. Singh and Whittington (1975) and Kumar (1985) found evidence for serial correlation in the growth rates of firms supporting the variant of Gibrat's Law proposed in Ijiri and Simon (1964). Kumar (1985) confirms the previous findings rejecting the strong form of Gibrat's Law, by showing that the earlier conclusions were robust to correcting for autocorrelation in the growth rates.

One of the problems that has plagued this literature, particularly the early work by Hart and Prais (1956) and Simon and Bonini (1958), has been data without a balanced representation of small firms. Dunne and Hughes (1994) and Hart and Oulton's (1996) work tries to address the problem by using a database with broad representation of small firms. They use this database to test for differences in growth rates among firms of different size classes and find the differences to be significant in contrast to Gibrat's Law. In the analysis of our model, we find the same differences and we also note that how small firms are counted matters when analyzing the firm size distribution itself.

A newer literature focuses on cross-industry variation. Santarelli and Lotti (2004) look at the evolution of the size distribution of new firms in four industries. Over a period of five years most of the distributions approach the log normal distribution, however, they find that the more technologically oriented industries achieved the lognormal faster. Audretsch et al. (2004) find evidence that services may exhibit different distributional properties than manufacturing, the main focus of the empirical literature to date. Looking at the Dutch hospitality sector they find that growth is independent of size, whereas the majority of studies focusing on manufacturing find the negative growth-size relation discussed above. Machado and Mata (2000) use quantile regressions to examine the effect of industry characteristics on different portions of the distribution for Portuguese data. While their results are mixed for some characteristics and distribution measures, they find that the impact of industry characteristics on skewness is the most stable over time. Both technology measures and the rate of growth in an industry reduce the skewness of the distribution, while turbulence increases it.

However, all of the results of the previous paragraph lack a solid theoretical base for their findings. It is this gap in the literature we seek to fill by proposing a fully dynamic model of optimizing firms that generates the distributional characteristics found in the empirical literature.

2.3 The Model

To capture the forces that affect firm size distribution in a structural model, we apply a variant of the Ericson and Pakes (1995) model described in Laincz (2004a). The modification allows for continually falling marginal costs through process R&D such that we can discuss both firm and industry growth rates. That enables us to perform analogous growth-size tests on the resulting simulated data.

We specify an industry with a finite set of imperfect substitutes such that one of the common drawbacks of the Ericson-Pakes framework does not apply. Because the state space for a single industry can be very large, it limits the total number of firms that the computational algorithm can handle, often to no more than about 10 firms. In order to generate a cross-sectional distribution with a reasonable number of observations, our industry is characterized by a finite set of imperfect substitutes, but each good is produced by a Cournot oligopoly. We solve for the dynamics associated with each substitute separately treating them as highly disaggregated goods and then aggregate across the varieties. We think of each good as being defined at a 7-digit level in the SIC or NAIC codes, for example, and the aggregation taking place at a less detailed industry level such as 4 or 5 digits.[1]

The older literature presumed that a market contained a series of isolated opportunities and assigned exogenous probabilities that these opportunities would be undertaken by either incumbents or new entrants. By specifying independent products within the same broader market, we follow Sutton (1998) in bridging the literature between the earlier stochastic models and the more recent literature devoted to strategic interaction, by using a distinction between the market as a whole, and a number of more or less independent submarkets within it. As Sutton (1997) states, the assumption is "crude," however, ". . . most conventionally defined industries exhibit both some strategic interdependence within submarkets, and some degree of independence across submarkets."[2]. Our characterization allows for strategic interdependence within each product market, but independence across products within the industry.

[1]The approach is similar to Sutton (1998), pgs. 19-20. However, our use of the term "submarkets" differs from his and accords more with his notion of "subindustry" (see pages 297-298).

[2]Sutton (1997), p. 49.

2.3.1 The Industry

We characterize the industry as producing intermediate goods sold into a perfectly competitive final goods sector. Firms producing the intermediate goods choose quantity produced, investment in R&D, and whether to exit or not, if they are currently active in the product market, or enter if they are not currently active. The dynamic equilibrium is a Markov Perfect Nash Equilibrium which imposes that decisions are functions only of the current state which is the current market structure. The basic timing of the model begins with incumbent firms first choosing whether to exit or not. The remaining incumbent firms then compete in a Cournot fashion in the product market and determine their optimal levels of investment in R&D to lower future costs which follow a stochastic process. Potential entrants then compare their opportunity cost of remaining outside the industry to the expected value of entering in the next period. These potential entrants draw on a public stock of knowledge which increases overtime through spillovers according to another stochastic process. At the end of the period, R&D outcomes and the public stock of knowledge for the next period are determined by the results of the stochastic processes.

The Product Markets

Demand for intermediate goods comes from a perfectly competitive final goods sector with a CRS production function.[3] Output in period t, Y_t, of the final goods sector is given by the production function:

$$Y_t = k_{1t}^{\theta 1} \cdot k_{2t}^{\theta 2} \cdots k_{Mt}^{\theta M}, \text{ where } \sum_{m=1}^{M} \theta_m = 1. \tag{2.3}$$

[3]We could analagously think of (2.3) as the utility function for a consumer and apply the framework to imperfectly competitive final goods producers.

Each k_{mt} is the input from subsector m, where m denotes the products within the industry. Within each subsector, multiple firms engage in Cournot competition providing a homogenous good to gain market share. The demand for each intermediate good k_{mt} is given by:

$$k_{mt} = \frac{\theta_m P_y Y_t}{P_{mt}}. \tag{2.4}$$

Firms producing intermediate goods at any given time have a technology for production of intermediate goods where the marginal costs are constant although they vary across firms. All firms are assumed to face constant fixed costs which do not vary either with time or across firms. Each intermediate goods firm, $n \in [1, N_m]$, in industry m faces the following optimization problem for choosing quantity:

$$\max_{q_{jm}} \pi_{jm} = P_m \left(\theta_m P_y Y, \sum_{n=1}^{N_m} q_{nm} \right) q_{jm} - MC_{jm} q_{jm} - f \tag{2.5}$$

where market size, $\theta_m Y$, and total quantity, $\sum_{n=1}^{N_m} q_{nm}$, determine the price of the intermediate good, P_m. q_{jm} is the quantity output of firm j producing product m, MC_{jm} are the marginal costs for firm j, and f is fixed costs. The implicit production function is linear in the input good with a coefficient equal to the inverse of the marginal cost.

We focus on one submarket to illustrate the model in the discussion that follows. Let N_m^* be the number of firms producing $q_{nm} > 0$. The Cournot-Nash equilibrium outcomes yield the profits for firm j as

$$\pi^*_{jm} = \max \left\{ \begin{array}{c} -f, \\ \theta_m P_y Y \frac{\left(\sum_{n=1}^{N^*_m} MC_{mn} - (N^*_m - 1) MC_{jm}\right)^2}{\left(\sum_{n=1}^{N^*_m} MC_{mn}\right)^2} - f \end{array} \right\}. \tag{2.6}$$

Firms choose to produce if:

$$\left(\sum_{n=1}^{N^*_m} MC_{mn} - (N^*_m - 1) MC_{jm} \right) \geq 0. \tag{2.7}$$

Equation (2.7) simply states that a firm will choose not to produce if its marginal costs are too high relative to its competitors.

The Cobb-Douglas specification generates a Cournot solution for the intermediate goods firms in which profits are homogenous of degree zero in the vector of marginal costs across firms once we normalize expenditures on the final good, $P_y Y$, to unity. Thus, a proportional change in the vector of marginal costs leaves profits the same despite falling marginal costs through process innovation (described below). Moreover, it allows for *continuously* declining marginal costs as opposed to the Ericson-Pakes framework where marginal costs are restricted to take on values in a finite set. The reason is that for any given vector of marginal costs, once the policy functions specifying R&D expenditures, entry, and exit are determined, these decisions will not vary provided the vector of marginal costs changes proportionally. Hence, policy functions for a finite subset of possible vectors of marginal costs are sufficient to characterize the long-run equilibrium as marginal costs continuously decline with process innovation.

However, the functional form of the demand system does create a problem in the case of an intermediate goods industry containing a monopolist. Because the price elasticity of market demand is unity, the monopolist's solution is not well defined. We assume that there is a minimum scale level of operations for a monopoly.[4] Let $\underline{q}$ be the minimum amount that a monopoly must produce in order to engage in the market. The assumption has two effects. First, it immediately defines a solution for the monopoly problem with a positive level of output while still providing the monopoly with incentives to invest in order to lower its costs. Furthermore, provided $\underline{q}$ is sufficiently small, there remain strong incentives for firms to strive to become monopolists. The minimum scale chosen for the simulations of the next section, while small enough to generate large monopoly profits, is such that in equilibrium firms always have sufficiently strong incentives to remain in the market or enter the market when the number of firms is small. Those incentives are discussed in the next subsection. Given that true monopolies without regulatory protection are exceedingly rare, the focus on markets where the probability of a monopoly emerging is quite small seems realistic and appropriate for the questions at hand regarding the distribution of firms.

2.3.2 Evolution of Market Structure

The number of firms operating in each product market and their relative levels of marginal cost determine the market structure at any point in time. The market structure evolves through process R&D which lowers a firm's marginal cost when

[4]There are other assumptions that could be made here instead, but do not significantly affect the results. For example, it would be more natural to think of the minimum scale assumption applying to all firms whether or not there is a monopoly. This assumption, while more plausible, only complicates the Cournot-Nash solution by changing the corner solution for output from 0 to $\underline{q}$ for affected firms. Moreover, Dixit-Stiglitz technology is a viable alternative that yields the same homogeneity of degree zero property, but it does not create a poorly defined monopoly problem as in Laincz (2004b). That extension introduces a more complicated problem to solve without adding much in the way of additional insights for the present inquiry.

R&D is successful. We track the level of marginal costs by accounting for the number of innovations that each firm j in market m has available at time t and denote it as i_{jmt}. The mapping from innovations to marginal costs is:

$$MC_{jmt} = \frac{1}{Z}\exp(-\eta i_{jmt}). \tag{2.8}$$

Marginal costs fall at the rate η with each additional innovation. Z is a scale parameter on costs which we use below to calibrate the model to match the mean employment level of firms observed in the data. Z captures the unit labour costs of firms relative to the price of the final output good. Firms with a greater number of innovations enjoy a cost advantage over rivals. The cost advantage generates higher profits and the motive for engaging in process R&D.

The total number of innovations accessible by firm j is the sum of publicly available innovations for product m, labelled I_{mt}, and each firm's private innovations, ip_{jmt},

$$i_{jmt} = I_{mt} + ip_{jmt}. \tag{2.9}$$

Private innovations of incumbents diffuse to the public stock at a constant rate, δ. Thus, I_{mt} increments by one with probability δ in every period. We interpret δ as the strength of lead-time, secrecy, and patent protection within the industry. However, for incumbent firms an increment in the stock of public innovations also means a reduction in the stock of private innovations. Thus, in the absence of successful R&D in any period (discussed below), diffusion of an innovation to the public stock leaves the total stock of innovations for an incumbent unchanged. Consider an industry with, e.g., three firms simultaneously active. If, at a given moment in time, there is a diffusion process, although all firms loose a private innovation, the public stock of innovations available increases by one such that the incumbents' total innovation stock is unaltered. However, contrary to the private stock of innovations, the public stock is not private information and it is also available to the entrants, such that with the diffusion process the gap between the total innovation stock of the incumbents and the entrants is reduced. Thus, the effort that potential entrants need to pursue in order to catch up with the incumbents is lower than before the diffusion process. In section 2.5, we explore how the diffusion rate, δ, alters the observed firm size distribution.

The constantly growing public stock of innovations allows potential entrants to remain viable.[5] Completely new firms in a particular market do not have to invest to learn all of the innovations that have taken place in an industry since the beginning of time. Rather, we assume that most innovations are in the form of readily available public knowledge, while more recent innovations are held privately by incumbent firms. Existing firms have access to all of the publicly available technological innovations and have discovered some new ones through process R&D which is temporarily private information. It is through this process of knowledge diffusion that industries are prevented from becoming permanently monopolized.[6]

We assume that new firms generally enter at relatively lower efficiency levels than incumbents to capture the fact that hazard rates of exit decline with the age of the firm (See Dunne, Robertson, and Samuelson, 1988). Specifically, new

[5]If all information was permanently private, a leading firm could innovate a sufficient number of times such that the cost to a new firm of acquiring enough innovations to generate positive profits would make entry costs prohibitively high.

[6]In the specification presented here, there are no spillovers between active firms which contrasts with the empirical evidence (e.g. Griliches, 1992). The spillover from private to the public stock of knowledge is necessary for continual growth because it enables new firms to enter at levels competitive with incumbents. The model can be adjusted to account for diffusion between incumbent firms. Doing so would enable analysis of the role of secrecy and lead time and how they interact with market structure. Overall, we do not believe it would change the main results presented in the next section, but we do believe it is worthy of exploration in future work.

firms will enter, on average, with fewer private innovations than incumbents:

$$0 < ip^{EN} < \widetilde{ip} = \int \left(\sum_{m=1}^{M} \sum_{n=1}^{N_m} \frac{1}{N_{mt}} ip_{nmt} \right) dt, \tag{2.10}$$

where ip^{EN} represents the number of private innovations of a new entrant. If new firms entered at higher levels than incumbents, then incumbents would be more likely to die than entrants producing the counterfactual result that incumbents have a higher hazard rate of exit than new firms. The left side of the inequality implies that new firms are bringing some new ideas while the right-side, $\widetilde{ip}$, is the equilibrium average (over the long-run) number of private innovations held by incumbent firms. This assumption then creates the possibility that new firms can immediately establish themselves as the new leader if incumbents repeatedly fail to innovate.[7]

The stock of private innovations held by incumbent firm j increases through successful R&D. The role of R&D is given by:

$$ip_{jm,t+1} = ip_{jmt} + v_{jmt} \tag{2.11}$$

where:

$$\Pr(v_{jmt}) = \left\{ \begin{array}{l} \frac{ax_{jmt}}{1+ax_{jmt}}, \text{ for } v_{jmt} = 1 \\ \frac{1}{1+ax_{jmt}} \text{ for } v_{jmt} = 0 \end{array} \right\}. \tag{2.12}$$

x_{jmt} is the level of R&D undertaken by a firm at time t. Note that the function does not vary with firm size, i.e. large firms do not possess an inherent advantage in successfully conducting R&D. We do not need to assume advantages owing to size to generate R&D spending distributions that match the highly skewed

[7]This outcome occurs only rarely. Most of the time new firms will enter with a small market share relative to existing incumbents.

distributions in the data (see Laincz 2004a). This assumption is consistent with the arguments of Cohen and Klepper (1992) among others that there are no differences in the productivity of research investment owing to firm size.

The parameter a governs the productivity of R&D and is interpreted as measuring the technological opportunity and basic state of science. We assume this to be constant across firms and product markets. Clearly, the level of R&D productivity will be an important parameter for variation in the firm size distribution. Higher levels of a generate greater potential for any one firm to extend its technological advantage and generate greater variance in firm sizes. We explore the relationship between technological opportunity and the firm size distribution in section 5.

The combination of the two stochastic variables, R&D and diffusion, in conjunction with the solution to the dynamic equilibrium results in an upper bound on how much of a lead firm will actually gain over potentially new firms in equilibrium. Because returns to investment are decreasing when marginal costs are relatively low, firms will enter a "coasting" state and choose not to invest because the gains eventually become outweighed by the costs.[8]

This specification for the evolution of marginal costs and innovation has several notable features. First, it is the relative marginal costs that matter to firms' profits as shown in (2.6); the absolute level of the marginal costs (or total stock of innovations) is irrelevant to the decisions of a firm. Second, in contrast to Ericson-Pakes, the spillover process does *not* change the marginal costs of active firms, but it does lower the costs of potential entrants because the stock of publicly available innovations continually grows. This feature allows for potential entrants to remain within striking distance of the incumbents. Hence, the con-

[8] See Ericson and Pakes (1995) for a discussion of the coasting states.

tribution of private innovations to the public stock is an externality that benefits the pool of potential entrants.

2.3.3 Dynamic Equilibrium

Let s_{nm} be the number of firms with ip private innovations producing product m and define the vector $s_m = [s_{nm}]$ which describes the market structure at any point in time. There are two types of firms facing different problems: incumbents and potential entrants. Incumbents are either producing for the market or choosing to exit. Their problem is characterized by comparing the expected net present value of investment in R&D against a positive liquidation value given by ϕ. Potential entrants compare an outside alternative, ψ, against the net present expected value of entering minus sunk costs of establishing production facilities denoted by χ. Both ϕ and χ are assumed constant across time and equal across firms.

An incumbent's intertemporal decision can be described by the following Bellman equation where time subscripts are replaced with a prime indicating a future value and all others are current:

$$V_{jm}^{I}(ip_{jm}, s_m) = \max\left\{\phi,\ \pi(i_{jm}, s_m) - cx_{jm} + \left(\frac{1}{1+r}\right) E\left[V^{I}(ip'_{jm}, s'_m | ip_{jm}, s_m)\right]\right\} \tag{2.13}$$

where the I superscript refers to the value of an incumbent. If the firm chooses to exit it receives the liquidation value ϕ, otherwise the firm receives current period profits minus its investment level in R&D, x_{jm} at a cost of c per unit plus the discounted expected value conditional on future market structure. The future market structure depends on the firm's current number of private innovations and the current market structure. $1/(1+r)$ is the common discount factor facing all firms. The expectation sign reflects the fact that the firm is assigning probability weights via the transition matrix of the market structure moving from its current state to all possible states. These include the probability of a spillover, the probability the firm itself will be successful in R&D, the probabilities of other firms being successful, and the probabilities of entry and exit.

A potential entrepreneur may enter a submarket, incur sunk entry costs, and establish production and R&D facilities. Production and sales do not begin until the following period. The Bellman equation resembles that for incumbents with few changes:

$$V_{jm}^{EN}(ip^{EN}, s_m) = \max\left\{\psi,\ -\chi - cx_{jm} + \left(\frac{1}{1+r}\right) EV^I(ip'_{jm}, s'_m | ip^{EN}, s_m)\right\} \tag{2.14}$$

where the EN superscript refers to entrants and the future value corresponds to that of being an incumbent in the next period. ψ measures the opportunity cost of entering and χ represents the sunk entry costs. By endogenizing entry and exit, we can observe how turnover rates respond to changes in structural parameters as we observe changes in the firm size distribution in the analytical section of the paper.

The Bellman equation can be written as follows:

$$V_{jm}^{I}(ip_{jm}, s_m) = \max\left\{ \begin{array}{c} \phi,\ \pi(i_{jm}, s_m) - cx_{jm} + \\ \left(\frac{1}{1+r}\right)\left[\frac{ax_{jm}}{1+ax_{jm}} C_1(ip'_{jm}+1, s'_m) + \frac{1}{1+ax_{jm}} C_2(ip'_{jm}, s'_m)\right] \end{array} \right\} \tag{2.15}$$

The investment strategy of firms derives from the first order conditions on the above. Let $C_1(ip'_{jm}+1, s'_m)$ denote the expected value of the firm conditional on successful innovation and $C_2(ip'_{jm}, s'_m)$ the expected value if it fails to innovate[9]. From this, the first-order condition yields the following policy function:

$$x_{jm}(ip_{jm}, s_m) = \max\left\{0, \frac{-1+\sqrt{\frac{a(C_1-C_2)}{(1+r)c}}}{a}\right\}. \tag{2.16}$$

The firm chooses the value maximizing level of R&D investment subject to a non-negativity constraint on R&D expenditure. Investment in R&D rises with the expected marginal gain in value, $C_1 - C_2$, and falls with the discount rate, r, and the cost of investment, c. The productivity of R&D, captured by a, has offsetting effects. As a rises it increases the probability of successful R&D and an incentive to increase investment. However, the higher the level of a the lower the marginal product for any given level of investment which lowers R&D investment.

Overall, the industry exhibits growth in total output and thus, Y in equation (2.3), grows over time while the innovations constantly reduce the cost of inputs. This continually growing industry can exhibit a great deal of change over time in terms of the identities of firms, their relative sizes, and the degree of entry and exit. The model provides us with the ability to generate a long-run firm size distribution based on the ergodic distribution of the model *and* the ability to examine the growth-size relationship at the individual firm level. The numerical algorithm uses value function iteration to solve for the space of values given by all possible combinations of firms and private innovations. We use a code for finding the Markov Perfect Nash Equilibria of the game which is the original C programming language version of the code by Pakes, Gowrisankaran and McGuire

[9]This notation follows the convention used in Pakes, Gowrisankaran and McGuire (1993) "Implementing the Pakes-McGuire Algorithm for Computing Markov Perfect Equilibria in Gauss", pag. 17, available at the authors webpage. Details on this notation can be found in the appendix to this chapter, which presents an explanation for the algorithm and code used to find the equilibrium of the game presented.

for implementing the Pakes and McGuire algorithm. A Gauss and Matlab versions of this code are made available at the authors' webpage for this code is written in [10]. We extract the policy functions including R&D expenditure as well as the entry and exit decisions. From the solutions, we can simulate our product markets and industry for comparison with the results found in the empirical literature. We now turn to that analysis.

2.4 Firm Size Distribution

2.4.1 Simulation

Table 2.1 presents the baseline parameterization of the model. We set the discount rate to 1/1.08 as an approximation of the average cost of capital for firms following Ericson and Pakes (1995). The rate of technological spillovers, δ, is set to 0.7 such that knowledge enters the public pool roughly one-and-a-half years after discovery. This fits with the empirical estimates of Mansfield (1981) on imitation time. Cost of a unit of R&D spending is set to one unit of the final good. The liquidation value and outside opportunity cost are chosen to be small to prevent them from dominating the incentives firms face. The liquidation value is about 7.5% of the average firm value, while the opportunity cost is roughly 15% of average firm value. We set both fixed and sunk costs equal to the outside opportunity cost.

The parameters a and η interact to determine the incentives for investment in R&D and ultimately the growth rate of the industry as measured by the rate of cost reduction. These parameters are set to 3 and 10% respectively. The latter

[10]Details on the algorithm and code used to find the Markov Perfect Nash solutions to the entry/exit and R&D investment problem stated here can be found in the appendix to this chapter.

implies that successful R&D will reduce marginal costs by 10%, but the former governs the incentives to engage in R&D such that we find at the mean level of R&D investment, the expected rate of cost reduction is just over 2%, which is approximately the industry growth rate. Both of these two parameters, plus the rate of knowledge diffusion, fixed and sunk costs will be allowed to vary in the following section to analyze their relationship to the moments of the firm size distribution.

The state space constraints we use have a maximum of six firms per submarket and each firm can hold up to 30 private innovations. To ensure that the state space boundaries do not drive the results we choose our demand parameters such that when six symmetric firms are in a submarket they are making negative profits. For the maximum number of private innovations, we checked in the simulations whether any firm attempted to obtain more private innovations than exist in the state space and made adjustments accordingly which led to our choice of 30.

Because we had no priors on how to vary the submarket sizes[11], we choose to use a simple, transparent linear function as follows:

$$\theta_m = \theta_1 + (m-1)\,b \tag{2.17}$$

where θ_1 is the market share of the smallest submarket and each submarket increments by b[12]. Upon simulating the model, we use the state space constraints

[11] An example on empirical work along these lines is the article by Buzzacchi and Valletti (1999), where they develop a test to the independent submarkets model proposed in Sutton (1998) for the Italian motor insurance industry. The independence between opportunities in the sector, both due to spacial and administrative reasons, provides an ideal setting to test Sutton's independent submarkets model. The authors have found that the size of the submarkets affects the skewness of firm size distribution. Further investigation concerning the impact of the size of the submarket on skewness and the other higher moments of the firm size distribution in the model presented here would be of interest.

[12] We also considered using a random process, or possibly demand shocks, but opted for the

to determine θ_1 and determine m by matching the model to the empirical results on the firm size distribution. b then follows from $\sum_{m=1}^{M} \theta_m = 1$.

The choices on the market size parameters, θ_1 and b, were determined as follows. We started with 10 submarkets, $m = 10$, where the smallest market share was determined by the lowest level of the market size that still produced positive levels of investment in R&D. At the tenth submarket firms began to invest at the upper bound of the state space. Therefore the increments in market share per submarket were determined to be 0.0318. For the analysis below, we then choose the number of submarkets, m, to analyze by matching the general shape of the log-normal distribution which closely, but not perfectly, resembles the empirical firm size distribution across industries. We found that if there are too few markets, the distribution is skewed left instead of right. Thus, for very narrowly defined markets with only one or two submarkets, the model generates a high frequency of average sized firms and a small number of tiny firms. On the other hand, as the number of submarkets expanded the model generated a bimodal distribution, in accord with some of the findings in Bottazi et. al. (2003b) for some industries. For the general log-normal distribution, we found that specifying five submarkets, $m = 5$, was the closest match to the results reported in Hart and Oulton (1996) discussed in the next subsection. As a further check on the validity of the results, beyond matching the general pattern of the firm size distribution, we then conduct cross-sectional regressions to see if the growth-size relations match the empirical literature in section 2.4.3.[13]

simple linear function on account of its transparency.

[13]Clearly, it would be more appealing to have the submarket sizes determined endogenously. This additional feature could perhaps be accomplished by specifying a Dixit-Stiglitz demand function. However, it would still require additional assumptions on how firms interact across sectors in terms of both price-taking (or not) behavior as well as specifying how innovations in one sector affect the other. The additional complications introduced would detract from the main task of this paper, which is to understand how the overall firm size distribution changes with underlying structural parameters. Moreover, because our model captures strategic

2.4.2 Distribution Results

In our first comparison of the model with the data, we compute the ergodic distribution by simulating the model.[14] From the distribution found in the simulations, we weight the observed outcomes by their probability of occurrence to generate the ergodic distributions for various size measures.[15]

Table 2.2 shows the results of the baseline parameterization compared with the statistics found in Hart and Oulton (1996) who use subsets (50 to 80 thousand companies) of a large UK database that includes very small firms in the sample. We calibrate the cost parameter, Z, to match the mean log size of employment reported in Hart and Oulton (1996). Because their data set has a good representation of small firms we felt that it was the most appropriate for comparison with the model. They find that the distribution of the natural log of various size measures (employment, sales, and net assets) exhibit positive skewness (long right-tails) and peaked (leptokurtic) distributions relative to the log normal distribution. We report analogous measures based on our model.[16] Sales are computed by extracting the quantities and prices while we use firm values, V^I, for net assets. All values below are reported in natural logs.

interaction within each submarket, it is likely that most forms of strategic interaction across submarkets would be of second order importance. Our assumption of no strategic interaction across submarkets fits with the arguments of Sutton (1997), mentioned earlier, for blending strategic interaction with the independent opportunities assumed in the older literature on Gibrat's Law.

[14]The simulation runs the model for one million periods. In order to avoid any bias caused by the specification of the initial market structure, we simulate it first for 10,000 periods and find the modal market structure. The main simulation then uses the modal market structure as its starting point.

[15]It is important to note that the comparison here with the data is not direct. We take advantage of the fact that through simulations we can generate the probability distribution of the market structures. Empirical studies use a cross-section of firms at a point in time (we turn to this analysis later) while the ergodic distribution shows the probabilities of a market structure occurring at a point in time. That is, the ergodic distribution is generated as a time series, but it reveals what the *expected* cross-section would look like.

[16]All employment calculations add one in levels to represent the manager which we view as part of fixed costs.

Figures 2.2 to 2.7 show the distributions in levels and in logs. The distributions in levels, for all the three size proxies considered, exhibit long right tails, especially for the net assets distribution. The size range accumulating the higher probability mass lies to the left of the mean size in all the distributions. The distributions for both the log of sales and log of net assets are roughly bell shaped, but exhibit thicker tails and higher peaks than the standard normal. The distribution of the log of employment exhibits less variance, but shows some skewness and leptokurtosis.

There are two noticeable differences between the model and the data. First, the standard deviations of the size measures are considerably smaller. This discrepancy is not surprising since the model is designed for a particular industry whereas the empirical estimates cover a large range of industries which would generate greater size variation.

Of greater concern is the slight negative skewness in the natural log of employment generated by the model versus the positive skewness observed in the data. Upon careful examination of the results, it turns out that the negative skewness is being driven by a tiny fraction of extremely small firms. These are firms with less than one employee which constitute about 0.1% of all firms and only 0.0025% of total employment. Those firms are to the left of the vertical line in Figures 2.2 and 2.3. If we eliminate them from the distribution and recalibrate Z, the skewness in employment goes slightly positive and the negative skewness in sales is cut in half as shown in Table 2.3. Moreover, the high kurtosis value in sales comes down considerably and is much more in line with the data. If we drop more small firms, less than 5 employees (0.02% of total employment), the skewness for employment rises to approximately 0.41. In fact, we found that we can match the Hart and Oulton skewness figure almost exactly if we eliminate all firms of less than 10 employees (0.14% of total employment).

The negative skewness for sales remains even after eliminating the small firms, although the skewness value for sales reported by Hart and Oulton, while positive, is the smallest of the three. This result of our model is being driven by the strong implications of Cournot oligopoly pricing with homogenous goods in each submarket. For example, when multiple firms produce large quantities, and hence have substantial employment, the direct competition between them drives the price down significantly. What we find is that the model often generates 3 or 4 firms in a given submarket with marginal costs that are very close. Although quantities are reasonably high for these firms, the low level of the price accounts for the reduced skewness when comparing employment and sales.[17] Overall, the ergodic distribution of the model reasonably matches the observed data in terms of deviations from a log-normal distribution

When we turn to the growth-size relationship in the next subsection, we extract a balanced panel which eliminates exiting firms. These exiting firms include these extremely small firms that generate the negative skewness in employment. Thus, in our summary statistics on the balanced panel below, the skewness measures increase significantly. These results suggest an interesting hypothesis. First, although the skewness is generally smaller than that observed in the data, it is important to bear in mind that data sets rarely include the full population of small firms. Second, the model accounts, in some sense, for part-time workers while data typically do not. These differences may be relevant empirically for testing distributions against the log-normal distribution. For example, if data collected round workers upward it would imply an underweighting of the left-side

[17]Note that the statistics we report exclude monopolies altogether so as not to be affected by the minimum quantity assumption. In the baseline, monopolies account for less than 0.000001% of all observations.

of the distribution which would bias skewness upwards.

2.4.3 Cross-Sectional Growth-Size Properties

We examine the growth-size relationships of the simulated model and compare them with the empirical literature on Gibrat's Law. To extract cross-sectional data comparable to that used in the empirical literature, we simulate the model five times for 5,000 periods each and extract the final periods from each run.[18] That provides us with simulated panel data to test the growth-size relationship. We ran these simulations ten times to check the robustness of these results.

The average number of total firms observed in each simulation was 138.8 (range of 132 to 143) with an average of 43.4 new entries over two periods and 25.9 exits.[19] We eliminate all the new entrants, who do not produce in their initial period and all firms that exit to generate a balanced panel for analysis. Table 2.4 shows the average distributional characteristics for size measures across these simulations of the balanced panel of firms for the initial period. The measures are similar to those shown in Tables 2.2 and 2.3, but note the substantial increases in the skewness values when small exiting firms are eliminated.[20]

Table 2.5 provides the results of the regressions on each of the ten simulations of the following form:

$$y_t = \beta y_{t-1} + \epsilon_t \tag{2.18}$$

where y_t is the log of the various size measures. Of interest in terms of Gibrat's

[18]To prevent the variance of the size of the firms from being dominated by the overall growth process, we shut down the increments to the public stock of knowledge except for the periods we extracted for analysis.

[19]Note that entries and exits would match almost exactly if we included those firms that exited in the previous period.

[20]The negative skewness in sales persists, but becomes even smaller in absolute value than when we eliminated the smallest firms outright in the preceding subsection.

Law is whether the coefficient is significantly different from unity. Hall's (1987) estimates for β as applied to employment across three different samples were consistently 0.99 and significantly less than one. Evans (1987) found values for β that range between 0.93 and 0.97 for employment. The model here also generates a coefficient less than one, below Hall's estimates, but in accord with those of Evans'.[21] The last columns report the percentage of times the null hypothesis of $\beta = 1$ was rejected at the 10% level, followed by the percentage of times it was rejected at the 5%, and 1% levels, respectively.

We use the median size values to split our sample into large and small firms. Empirical evidence suggests that Gibrat's Law works better for the large firms. Our results show a similar pattern. The estimated β's are consistently closer to one for the large firm sample for each size measure and in every simulated sample. In fact, for the employment size measure we cannot reject $\beta = 1$ nine out of ten times and then only at the 10% level.

For the three size measures we also tested for the equality of the coefficients between the large and small firm subsets reported in the last rows of Table 2.5. For employment we reject equality in all cases at the 5% level or better and 80% of the time at the 1% level. For sales, the differences weaken somewhat and we reject equality 60% of the time all of those t-statistics at the 5% level or better. Equality of values is rejected in nearly all of the subsamples. It is worth emphasizing, however, that in all subsamples, the estimated beta for large firms was greater than that of small firms for all three size measures. Given the small sample size we draw, the large number of significant rejections of $\beta_{LARGE} = \beta_{SMALL}$,

[21] We report the results using robust standard errors, but even without using them the results are hardly changed. The R-squared's are exceedingly high typically between 0.95 and 0.99. However, since there are only two simple stochastic processes in the model and nothing akin to demand shocks, it is not surprising in the least that past size is a good predictor of size in the short-run.

indicates the model matches the empirical findings.

In order to test for serial correlation we reduce our sample to those firms surviving in three consecutive periods for a balanced panel. The previous tests only had 69.5 observations on average and after one more period of eliminating exiting firms to retain a balanced panel, we were left on average with 46.9 firms. The test specification is similar to Kumar (1985) where growth is the dependent variable (instead of the log of size):

$$\left(\frac{y_t}{y_{t-1}}\right) = (1+\beta)\, y_{t-1} + \gamma \left(\frac{y_{t-1}}{y_{t-2}}\right) + \epsilon_t. \tag{2.19}$$

Persistence in growth will show up as a positive value for γ. We find that the coefficients for β and γ are significant at the 5% level for the majority of the ten samples. Their average estimated values are 0.9817 and 0.577 respectively. γ is positive and significant in all but one of the samples at the 1% level.

The positive and significant value of γ indicates serial correlation which comes as no surprise given the design of the model. There are several contributing factors to serial correlation in our model. First, successful firms seek to build on and protect any technological advantage and thus invest more heavily than small firms. In addition, a growing firm pushes rival firms closer to the exit threshold. Thus, the growing firm will get a subsequent additional increase in market share with the increase in the likelihood of rivals' exit. These processes of firm dynamics effectively embed serial correlation in error terms that do not control for innovative behaviour and expected future changes in market share conditional on them. The results suggest that serial correlation should weaken in empirical studies if appropriate controls for own and rival R&D expenditure and innovations are included. We leave this hypothesis for future empirical work.[22]

Finally, we look at the variance in growth rates across firm sizes. A number of studies find that the variance of the growth rate is larger for small firms (e.g. Dunne and Hughes (1994)). The work by H. Stanley and his co-authors[23] has also found, for a dataset comprising information concerning all publicly traded U.S. companies between 1974 and 1993, that the fluctuations in the growth rates , as measured by the width of the distribution, decrease with company size and increase with time. They find a scaling relationship between the variance of growth rates and company size which is the same for all size measures they have considered. In fact, they find that the spread in the distribution of rates decreases with increasing sales as a power law over seven orders of magnitude.

Again, we separate our simulated samples by the median size. Table 2.6 shows the average standard deviations in growth rates across the ten samples and for large and small firms according to the three size measures. By all three measures the variance in the growth rate of the small firms is larger than that of the large firms and across all ten samples. The final columns report the percentage of rejections based on the F-statistic for the variance ratio test for equality of the standard deviations. We reject equality at the 1% level based on the employment and sales measures in seven out of ten samples, but in only half

[22] One note on the magnitude of serial correlation is required here. Our estimate of γ is larger than that found in either Singh and Whittington (1975) or Kumar (1985) who find values of approximately 0.3 and 0.12 respectively. The distinguishing feature is in the difference in time periods. Those authors use a much longer time frame, 10 to 12 years, compared with our simulated data which corresponds to roughly three years based on the user cost of capital we specify. Because we know that the model will predict serial correlation that declines over time due to the Markov perfect nature of the equilibrium, we do not pursue that issue any further here. Suffice it to say, that the model does generate serial correlation in the errors when using the basic regression model found in the growth-size regressions related to Gibrat's Law. See Pakes and Ericson (1998) for the empirical implications of the Markov Perfect feature embedded in the model.

[23] "Scaling Behaviour in Economics: The Problem of Quantifying Company Growth", Physica A (1997), and "Scaling Behaviour in the Growth of Companies", Nature (1996), among other articles.

the samples for net assets. The latter also has the highest level of the standard deviation. Overall, the results are encouraging in the sense that, again the model replicates empirical findings.

To summarize the section, we find that the model is able to replicate empirical studies of Gibrat's Law in two ways. First, it can generate a firm size distribution with the higher moments deviating from the log-normal distribution in the same direction as actual distributions. Secondly, in the cross-section the model generates a negative firm size-growth relationship, decreasing variance in the growth rate with firm size, and serial correlation, all found in the data. Based on the above we feel reasonably confident in using the model to understand how underlying structural parameters affect the overall firm size distribution.

2.5 Variation in the Firm Size Distribution

The previous section established that the model reasonably matches the data in terms of the firm size distribution and in its cross-sectional empirics. Now we ask how the moments of the firm size distribution change with underlying structural parameters suggested in the literature. Specifically we vary the following parameters: sunk costs, fixed costs, productivity of R&D, rate of spillovers, and the rate of decline in marginal costs.[24] The goal of this section is to generate a set of hypotheses that can be examined empirically. We do not carry out that examination in this paper, but view the contribution of this analysis as setting an

[24] We do not vary the outside alternative parameter, ψ, because it enters in much the same way as sunk costs, and we do not vary the liquidation value because the parameter must be constrained to be less than $X_e/(1+r)$ such that firms cannot enter, produce nothing, and exit with a net gain. We also do not vary the discount rate because even if the discount factor varied uniformly across firms, such variations basically mean interest rate variations and those variations are typically short-run fluctuations rather than long-run characteristics inherent to an industry that shape the firm size distribution.

empirical research agenda on the firm size distribution guided by theory. How the model fares when taken to the data should provide insights for improvement in the model itself and a deeper understanding of the empirical work. All of the analysis below shows the distribution including all levels of employment, i.e. all firms are included no matter how small.[25][26]

2.5.1 Fixed Costs

We start with fixed and sunk costs. In the baseline fixed costs were set to 0.2 and we allow that to vary from 0 to 0.25, when translated via the unit cost of labor, the range of the fixed costs then go from about 12% (for the smallest non-zero value, 0.025) to just over 50% of total costs of production excluding R&D costs for the mean sized firm in the sample.

We find that *increases in the level of fixed costs lead to a larger mean size of firms, but lower variance, skewness and kurtosis of the firm size distribution.* Figure 2.8 shows how the first four moments of the firm size distribution change relative to the baseline while Figure 2.9 shows the same measures when we look at the firm size distribution in natural logs. The x-axis shows the level of fixed costs and the y-axis shows the percentage change from the baseline. Figure 2.10 shows the baseline distribution in levels against the low and high value of fixed costs.[27] In the latter, Figure 2.10, low levels of fixed costs are associated with

[25] We also examined the behavior of the first four moments when eliminating small firms as in the previous section, but found no qualitative differences. The only notable difference was that as we eliminated small firms from those of less than 1 employee, to less than 5, to less than 10, the effects became even more pronounced. By that we mean that the percentage changes in any moment of the firm size distribution were larger when eliminating small firms, but the direction of the effects was stable.

[26] We also analyzed the changes in the distributions of both sales and net assets, but we do not report those results here. Qualitatively they are very similar to the effects on the size distributions by employment.

[27] In Figures 9 through 12 for the graphs showing the shape of the firm size distribution, we standardize the x-axis to maximums of 160 and 6, in levels and logs. However, in many

a greater mass of the distribution at lower levels of employment, but also with a longer right-tail and, hence, a higher skewness. The high levels of the fixed cost exhibit greater mass further to the right and there is a small, but noticeable, second mode emerging to the right of the peak.

Figure 2.11 shows the same distribution but in natural logs and we see similar changes. The mean size rises as the mass of small firms shrinks while the mass of larger firms grows. The variance falls as the distribution becomes more leptokurtic (in logs) as firms become more concentrated around the mean size. The skewness increases relative to the baseline at lower levels of the fixed costs, but then declines at the higher levels. The initial increase stems from an increase in the frequency of firms above the mean size creating more mass on the right-side of the distribution. The decrease follows from the flattening of the left-side of the distribution as smaller firms become more dispersed in their scale of operations.

When we set fixed costs to zero, the mean firm size is more than 11% below the mean size of firms in our baseline. As fixed costs rise the mean size of firms also rises with a more rapid increase at higher levels. Although the pattern exhibits some non-monotonicity, the variance generally falls with increases in fixed costs, while both skewness and kurtosis decline.

Intuitively, in the model higher fixed costs make it more likely that small firms will choose not to produce as in equation (2.7) and increase the likelihood of exit because future profit values are smaller for the same level of output. Thus, by reducing the fraction of small firms in the sample the mean size increases. Moreover, with small firms more likely to exit, higher fixed costs create greater incentives to innovate for incumbent firms to distance themselves from the exit

cases the maximum sized firm exceeds those values. We choose to standardize the x-axis to facilitate comparison in the regions showing the bulk of the mass and how the parameters alter the distribution.

threshold which further increases the mean. The variance, however, declines and is related to the decrease in kurtosis. With a reduced fraction of small firms, the frequency of firms near the mean size increases, but there is also an increase in the mass of firms to the right of the peak reflecting higher R&D investment among incumbents which decreases the kurtosis. In natural logs we see more or less the same pattern, however, the higher moments behave differently. Skewness displays an inverted-U shaped pattern, while the kurtosis displays a generally increasing pattern. The reason is that the fattening of the tails in levels is primarily just to the right of the primary mode such that in logs the effect blends with the original mode and the tails remain relatively flat.

Skewness, in levels, falls because the region of small firms becomes smaller while the frequency of mid-sized firms increases. This effect flattens and lengthens the tail on the left-side which reduces the skewness. Thus, fixed costs act in a way that is fairly straightforward by making it more difficult for very small firms to survive, essentially requiring a minimum scale of operations to produce non-negative profits and survive. These results are consistent with the findings of Machado and Mata (2000) who use a Box-Cox quantile regression model to characterize the effect of covariates on the firm size distribution of Portuguese firms. They find that minimum efficient scale had a consistently positive impact on the size of firms, a negative effect on the skewness, and an ambiguous effect on kurtosis.[28]

2.5.2 Sunk Costs

Figures 2.12 to 2.15 show the results from varying the sunk costs of entry. The range here starts from 0.1 such that $\left(\frac{1}{1+r}\right) X_e > \psi$ continues to hold. The upper

[28] Machado and Mata (2000) do not report measures of variance.

bound here is much higher than for fixed costs to capture industries for which sunk entry costs will take, in expectation, significant time to recover. These values can be understood as a ratio to the value of the mean sized incumbent firm. The range is from 10% to 35% and equals approximately 21% at the baseline. The vast majority of firms that enter the market do not recover their full sunk costs. However, those that survive and grow ultimately reap substantial rewards. At low levels of sunk costs, a firm does not need to survive for a long period of time before making entry optimal. However, at high levels of the sunk costs, firms require a substantial likelihood of sustained success to induce entry.

At low levels of the sunk costs we see little change in the mean size of firms, but a negative effect on all the higher moments. In fact, we find that *industries characterized by lower levels of sunk and/or fixed costs will more nearly match the log-normal distribution and the strong form of Gibrat's Law.* In Figure 2.14 its clear that these changes are fairly small when comparing the shape of the baseline distribution to the low end for sunk costs. However, once sunk costs reach 0.25 (or approximately 26% of the value of the mean sized firm), the mean size of firms rises rapidly, while the variance increases though somewhat non-monotonically. The entry barrier discourages new firms reducing the mass of small firms. Markets become more concentrated with fewer firms, but of greater average size. Thus, *industries with high levels of sunk entry costs will exhibit greater average size of firms, higher variance in the size, but a flatter distribution potentially with multiple modes.*

Higher sunk costs have offsetting effects for incumbents which increases the variance but continues to reduce skewness and kurtosis. With smaller firms less likely to enter and pose a threat to incumbents, firms have less competition reducing the benefits of engaging in R&D which flattens the far end of the right-tail. However, at the same time, among incumbents, because the sunk costs help extend the expected lifetime of any one firm, competition in terms of R&D intensifies. Thus, once a firm does enter it has strong incentives to try to develop a technological lead over its rivals. This incentive leads to an increase in the mass of firms in the mid-sized range. Once a sufficient cost advantage has been established the first effect comes to dominate and discourages firms from establishing an even larger technological lead because the threat of entry has been reduced.

The model thus suggests that industries with large sunk costs should have a larger mean size, greater variance and a flatter distribution overall. The flatter distribution and Figures 2.14 and 2.15 suggest that bi- or multi-model distributions are quite likely for industries characterized by high sunk entry costs. The sunk entry costs protect incumbents such that once a firm reaches a sufficiently large size, it seeks to maintain that size by investing in R&D to maintain its advantage but with less incentive to increase that advantage.

Audretsch et. al. (2004) argue on the basis of some studies on particular industries that the service sector approximates better the strong form of Gibrat's Law and therefore the lognormal distribution because the link between firm size and survival rates is weaker in industries with lower sunk costs and where capital intensity and scale economies play less of a role[29]. While evidence on this point is limited, the Audretsch claim is consistent with the production of the present model reported above (page 54, italicised). In our analysis, we find that to be true particularly for the fixed costs which imply a higher requirement for scale in order for net profits to be non-negative. We also find that the distribution more closely approximates the log-normal distribution at the low end of the sunk costs.

[29] Audretsch et al (2004) study the Dutch hospitality sector.

2.5.3 Rate of Cost Reduction

The rate of cost reduction is captured by the parameter η. More specifically, η, which first appears in (2.8), is the percentage decline in marginal costs of production conditional on a successful innovation. Thus increases in η will translate into a faster industry growth rate for the same level of investment as measured by output of the final good. We think of η as a key parameter in governing the rate of technological progress which in the context of the model is the rate at which costs fall.

The parameter ranges from 0.07 to 0.20 with our baseline value set to 0.1 (10%), but to make better sense out of this specification, we convert it to the expected cost reduction at the mean level of investment. At the low end, few firms are actually engaged in R&D and thus the mean expected rate of cost reduction is only 0.32%, an anemically growing industry with little innovation. However, at the upper bound, there is a fair amount of R&D and the mean rate rises to 12.13%.[30]

Looking at Figures 2.16 to 2.19, we see that *increases in the rate of technological progress lead to an increase in both the mean and variance of the firm size distribution.* The higher levels of cost reduction lead to greater incentives to engage in R&D and capture market share from rivals which leads to increases in both of the first two moments. At the same time, *variation in the rate of technological progress has non-monotonic effects on both the skewness and the kurtosis.* Both exhibit convexity as η rises. Skewness falls initially because at low levels of

[30] Jorgenson and Stiroh (2000) report industry growth rates for highly aggregated industries with the fastest growing industry, electronic and electric equipment, growing at an annual rate of 5.457% from 1958 to 1996. That would suggest an upper bound for η of approximately 0.14 or 0.15. However, since that growth rate is for an industry at roughly the 2-digit SIC level, it therefore averages across more detailed sectors. Thus, we examine the effects for even faster rates of growth.

cost reduction, there is a very small percentage of extremely large firms. These are firms that established a technological advantage and raced ahead to cement their leading position. With increases in the rate of cost reduction and, therefore, greater incentives for small and medium sized firms to use R&D, the scope for business stealing rises. As a result the industry becomes more competitive leading to more firms and more competition with fewer truly giant firms which reduces the overall skewness.

As the rate of cost reduction increases, around $\eta = 0.15$ or a mean expected cost reduction of 7.1%, the skewness begins to increase as a larger mass of firms emerges in the mid-sized range as can been seen in Figures 2.18 and 2.19. Kurtosis undergoes a similar change. In fact, the whole distribution almost completely flattens out at our extremely high range. This effect occurs because the range of relative marginal costs throughout the incumbent firms increases along with the strong incentives to engage in R&D to defend existing market share as well as capture market share from rivals. Thus, rapid growth should lead to a high variance and a flatter distribution. This leads to the hypothesis that *industries with high rates of technological progress are more likely than those with low rates to exhibit multi-modal distributions.*

Machado and Mata (2000) also measure empirically the marginal effects of industry growth rates on the firm size distribution in their paper. They find that firms in faster growing industries have a higher mean, but more rapid growth reduces the skewness. For kurtosis they also find a negative effect, but it is not statistically significant.

2.5.4 Technological Opportunity

The productivity of R&D which we think of as the technological opportunity facing an industry is captured by a in equation (2.12). It is related to the rate of cost reduction, η, in the sense that those two parameters jointly affect the equilibrium rate of cost reduction. a governs the incentives to engage in R&D and η defines the gains of success in terms of cost reduction. Higher levels of a imply a higher probability of success for any given level of R&D expenditure. However, the marginal impact of an increase in R&D expenditure falls with higher levels of a. Moreover, the solution for the optimal level of investment based on the first-order condition of the value function shown in (2.16) implies that changes in a will have countervailing effects. Thus, as a prior, we expect to find non-monotonic behaviour as a varies.

The range of a that we used went from 2 to 4, centred around the baseline value of 3. At both the lower and upper limits the computational algorithm began to generate extreme results. At the lower level, we found that virtually no firms were investing in R&D while at the upper bound firms began to exceed the limitations of the state space. To provide some economic interpretation of these values, a firm investing at the average level from the baseline, would expect success in R&D with a probability of 20.7% and thus an expected cost reduction by the following period of 2.07%. At the lower bound of a, 2, those values fall to less than 10% and under 1% while at the upper bound they are slightly less than double the baseline.

Figures 2.20 through 2.23 show the effect of varying a. *Both the mean and variance of firm sizes rise with productivity of R&D.* In natural logs the pattern is similar, but there is some concavity at the higher levels of a with respect to the variance. *Increases in the productivity of R&D have non-monotonic effects* *on both the skewness and the kurtosis.* Skewness and kurtosis both exhibit concavity, which contrasts with the results for the rate of cost reduction. The difference between these two parameters appears to come from the left-side of the distribution and their effects on smaller firms. Changes in the productivity parameter affect the right-tail of the distribution in much the same way as an increase in the rate of cost reduction. In both we observe a steady increase in the mass of firms to the right of the peak while the peak itself shrinks which eventually lowers the skewness and the kurtosis. In natural logs the pattern is similar with, again, another mode emerging on the right.

At very low levels for a, we found that the right-tail was much shorter and thinner than for the higher values. This follows from the much lower productivity in R&D which stunts the incentive to engage in R&D and the mean size of firms is considerably smaller as a result. Thus, at the lowest levels of a, as R&D expenditures yield greater returns with the higher marginal product, larger firms emerge and stretch the right-tail initially leading to increases in skewness and the kurtosis. As a rises beyond 2.5, more firms engage in R&D leading to the increase in the variance and hence a flatter distribution overall with less skewness.

Of the structural parameters we investigate, this parameter is almost certainly the most difficult to capture empirically. However, we do wish to emphasize the strong non-monotonicity in this variable and in the rate of cost reduction. The main conclusion we draw here is that empirically we should not expect proxies for either a or η to have clear monotonic effects on the higher moments of the distribution and caution should be exercised in generalizing results found in studies of the firm size distribution for a selected group of industries.

2.5.5 Rate of Spillovers

In the model, the parameter δ governs the rate at which the public stock of knowledge grows. The faster it grows the easier it is for entrepreneurs drawing on the public stock to enter the industry and challenge incumbents for market share. If $\delta = 0$, it would imply that no knowledge enters the public stock and over time no entrant would be able to challenge existing incumbents. At the other extreme, if $\delta = 1$ then all new innovations enter the public stock in the following period which would be similar to Klepper (2002) where all R&D is costlessly imitated.[31]

Mansfield (1981) reports imitation times that range from about 6 months to nearly three years. Thus, we allow δ to vary from 0.3 to 0.9 which generates an expected lifetime for a single innovation to remain private from just over one year to more than three years. Low values of δ can be interpreted as pertaining to an industry where incumbent firms possess strong advantages through secrecy, patent protection, and/or lead time to implement their innovations.

Figures 2.24 through 2.27 show the results which are quite striking. Changes in the rate of spillovers generate an enormous impact on both the mean and variance. *Industries with stronger patent protection (secrecy, or lead time) will have a higher mean and variance in the size of firms.* For example, at the low end of $\delta = 0.3$ the mean firm size is nearly six times that of the baseline! Intuitively the stronger the protection for private innovations, the greater their value to any one firm. Therefore firms will accumulate a great number of private innovations and establish a large presence in the market making it difficult for

[31] In Klepper (2002), he assumes randomly assigned R&D productivities which allows for survival of the more productive firms while generating high rates of exit during the product life cycle. Here we do not allow R&D productivity to vary by firm but allow the innovations to diffuse slowly which generates the advantage of size because large firms will hold more private innovations that smaller firms.

any new entrant to mount a successful challenge. However existing incumbents will compete fiercely in the R&D arena which contributes to both the high mean and the large variance.

Looking at Figures 2.26 and 2.27, we show the distributions when we move away from the baseline of $\delta = 0.7$ by ± 0.2. The changes, particularly as δ falls, are more substantial here than for other parameters examined. When δ is increased the peak mode becomes more pronounced with less variance in firm sizes. Because private knowledge passes quickly from any firm to the public stock, there is less ability and incentive for firms to engage heavily in R&D to separate themselves from rivals. Firms in an industry with a high rate of spillovers are facing an uphill battle on a slope that is nearly vertical.

At the low level of δ, the distribution has no obvious peak and shows great variation over the mid-sized range. For the same distribution, a small, but noticeable mode emerges to the far right (around 250 in levels and 5.6 in natural logs) which we do not see in other distributions. In fact the distribution generated by the model fails to resemble the empirical distributions. Turning to the higher moments there does not appear to be any straightforward effect and no discernible pattern. We draw no conclusions regarding the effects on skewness and kurtosis here other than to say they appear to be highly non-monotonic.

Based on the analysis here clearly the diffusion rate plays a critical role in shaping the firm size distribution. While δ represents the rate at which knowledge becomes available to new firms, it does not capture spillovers between incumbents. The extreme changes in the shape of the firm size distribution that follow from modifying δ at levels that are empirically plausible, suggest that our measure is simply too crude to capture all that secrecy and lead time entail. Extending the framework to account for spillovers, imitation costs, and absorptive capacity between active firms seems a highly fruitful avenue for further work.

2.6 Conclusions

Understanding the forces that generate differences in the firm size distribution enables us to identify the forces that generate more or less concentration across industries. This study provides a model for undertaking this task. We show that the model replicates the characteristics of the firm size distributions reported in the literature and reproduces the empirical growth-size relationships. The model generates a substantial list of empirical hypotheses for testing the effect of various structural parameters on the first four moments of firm size distributions. In addition the model suggests that serial correlation in firm growth should weaken in empirical studies if appropriate controls for R&D expenditure and innovations of existing and rival firms are included.

It is worth emphasizing that the model is quite flexible and can be adapted to serve as a baseline for analyzing particular industries by matching the parameters and the moments of an observed industry level firm size distribution. With that baseline, counterfactual experiments can be conducted and the effects of policies on the firm size distribution can be analyzed, such as subsidization of R&D or regulations that affect barriers to entry.

We note some missing elements in our framework that could be incorporated in future work. First, merger activity is one of the major concerns in the empirical literature (for example see Kumar, 1985, and Dunne and Hughes, 1994). Our model can incorporate mergers by combining it with Gowrisankaran's (1999) extension to mergers of the Ericson-Pakes framework. Second, the model here relies on stochastic R&D success and diffusion of knowledge to generate entry, growth, survival, and exit. More could be done to capture other risks that entrepreneurs face such as uncertainty of true costs as in Jovanovic (1982). That would enable an exploration of how the rise of venture capital and lowering of entry barriers, other than the sunk costs discussed here, affect the firm size distribution.

We view the work presented here as a step forward in the interplay between the theory and empirics of the firm size distribution. While the motivation for the theory comes from a host of empirical observations, the theory provides us with a list of hypotheses that can be examined empirically across industries. We would be surprised to find that all of the hypotheses generated apply to all industries and it is highly likely that the model may serve well for some industries but not for others. That probable outcome would lead to both further refinements of the model and, we hope, a better understanding of the forces that shape the firm size distribution across industries and their consequences.

Table 2.1: Parameter values

Parameter	Symbol	Value
Discount Rate facing firms	$1/(1+r)$	1/1.08
Rate of Technological Spillover	δ	0.7
Productivity of R&D investment	a	3
Sunk Entry Costs	X_e	0.2
Unit cost of R&D Spending	c	1
Fixed Costs	f	0.2
Liquidation Value	φ	0.1
Rate of Decrease in Marginal Costs	η	0.1
Liquidation Value	ϕ	0.1
Outside Alternative Value	ψ	0.2
Smallest Submarket Market Share	θ	10.136
Increments in Submarket Size	b	0.0318
Unit Cost of Labour	Z	131.12

Table 2.2: Distribution Statistics for Baseline (Natural Logs)

Measure	Normal Distribution	H&O Emp	Model Emp	H&O Sales	Model Sales	H&O Net Assets	Model Net Assets
Mean	-	3.1582	3.1582	7.2015	5.1321	5.5539	4.9010
Std. Dev.	-	1.5197	0.2803	1.6628	0.3697	1.9635	2.1468
Skewness	0	0.7487	-0.1114	0.1932	-1.0220	0.4366	0.7825
Kurtosis	3	4.5794	4.7265	6.1876	11.7373	4.835	2.9123

Note: H&O refers to the results reported in Hart and Oulton (1996).

Table 2.3: Distribution Statistics for Baseline Excluding Firms with <1 Employee (Natural Logs)

Measure	Normal Distribution	H&O Emp	Model Emp	H&O Sales	Model Sales	H&O Net Assets	Model Net Assets
Mean	-	3.1582	3.1582	5.1825	5.1321	5.5539	4.8469
Std. Dev.	-	1.5197	0.2718	1.6628	0.3516	1.9635	2.1499
Skewness	0	0.7487	0.0777	0.1932	-0.5251	0.4366	0.7777
Kurtosis	3	4.5794	3.9254	6.1876	5.6798	4.835	2.9031

Note: H&O refers to the results reported in Hart and Oulton (1996).

Table 2.4: Average Summary Statistics in Natural Logs

Size Measures	Employment	Sales	Net Assets
Observations	69.5	69.5	69.5
Average	3.521	5.027	5.189
Std. Dev.	0.521	0.493	1.410
Skewness	0.867	-0.402	0.717
Kurtosis	4.183	3.819	2.714

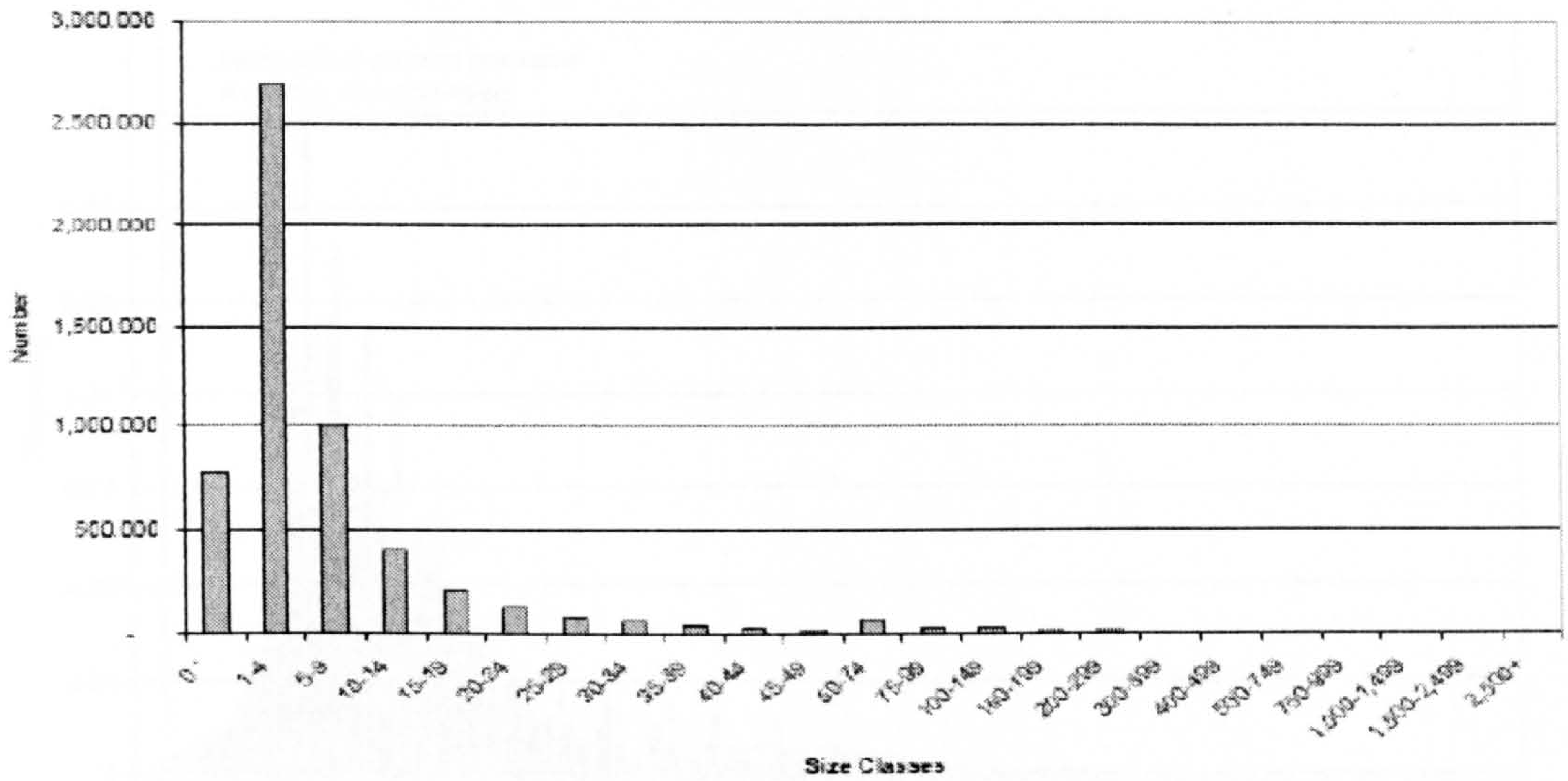
Table 2.5: Regression Results

Firms	Size Measures	Average Coefficient	$H0: \beta = 1$ Rejection Rate 10%	5%	1%
All	Employment	0.971	100%	90%	80%
	Sales	0.966	100%	100%	100%
	Net Assets	0.944	100%	100%	100%
Large	Employment	0.998	10%	0%	0%
	Sales	0.979	100%	90%	80%
	Net Assets	0.970	70%	70%	50%
Small	Employment	0.929	100%	100%	90%
	Sales	0.947	100%	100%	90%
	Net Assets	0.883	100%	100%	100%
		Average Difference	$H0: \beta LARGE = \beta SMALL$ Rejection Rate 10%	5%	1%
Equality	Employment	0.069	100%	100%	80%
	Sales	0.019	60%	60%	20%
	Net Assets	0.061	90%	90%	70%

Table 2.6: Tests of Standard Deviation of Growth Rates by Size Class

Size Measures	Total	Large	Small	$H0: \sigma_1 = \sigma_2$ Rejection Rate 10%	5%	1%
Employment	0.301	0.159	0.361	80%	80%	70%
Sales	0.321	0.195	0.401	70%	70%	70%
Net Assets	0.542	0.438	0.584	50%	50%	30%

Figure 2.1: US Firm Size Distribution by Employment, 2002



Source: Statistic of US Business (SUSB), Small business Administration

Figure 2.2: Employment Distribution

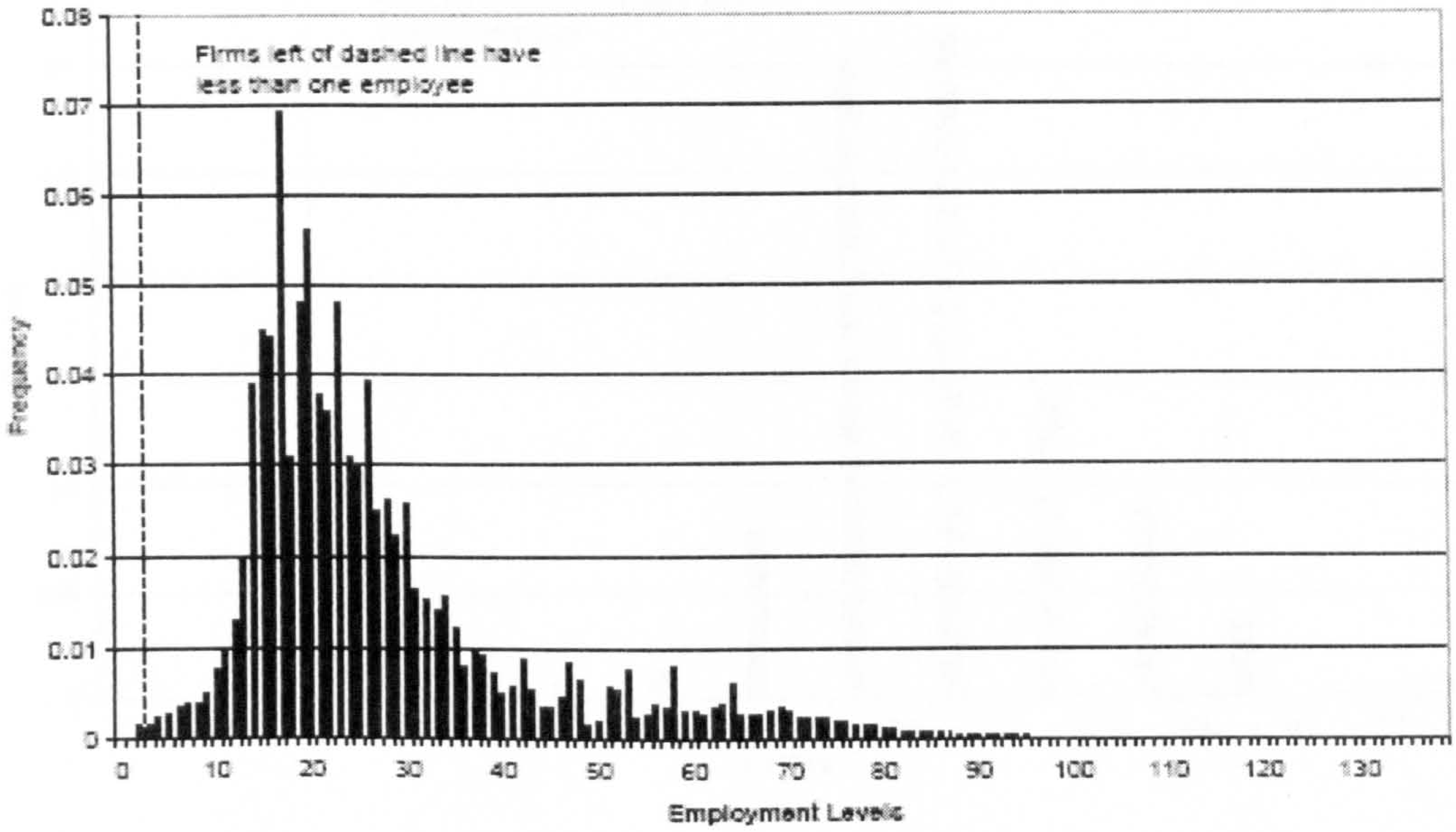


Figure 2.3: Log of Employment Distribution

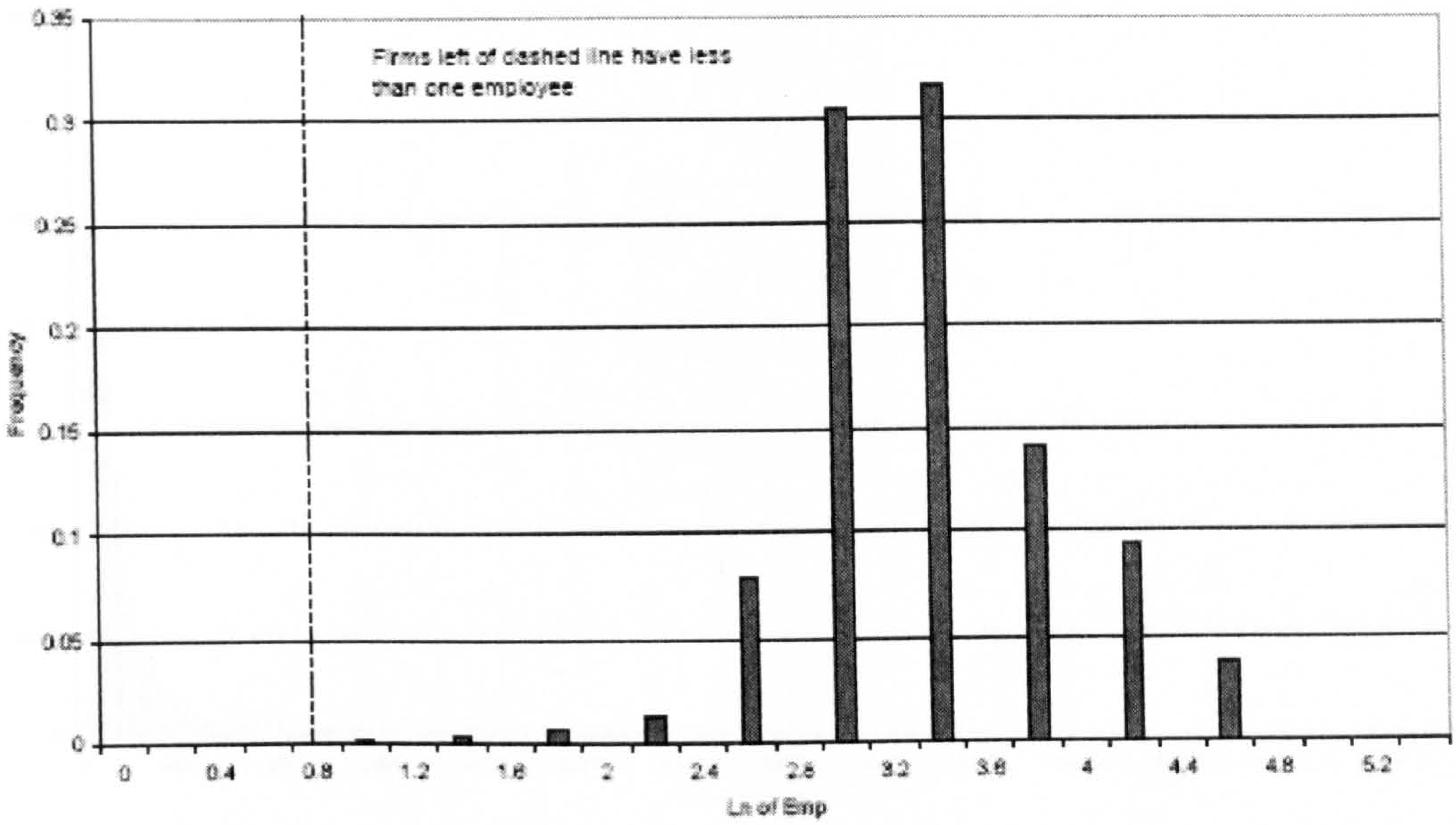


Figure 2.4: Values Distribution

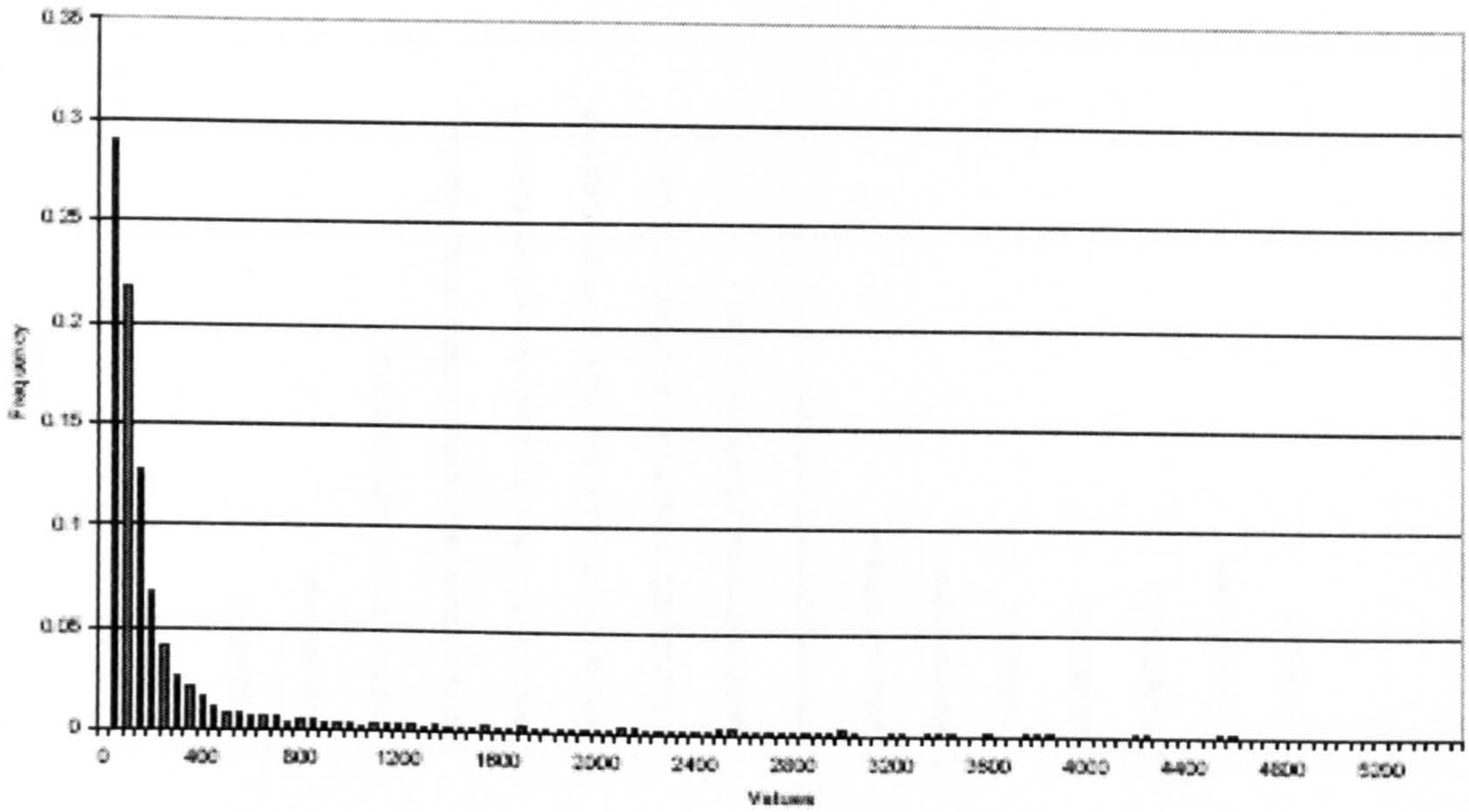


Figure 2.5: Log of Values Distribution

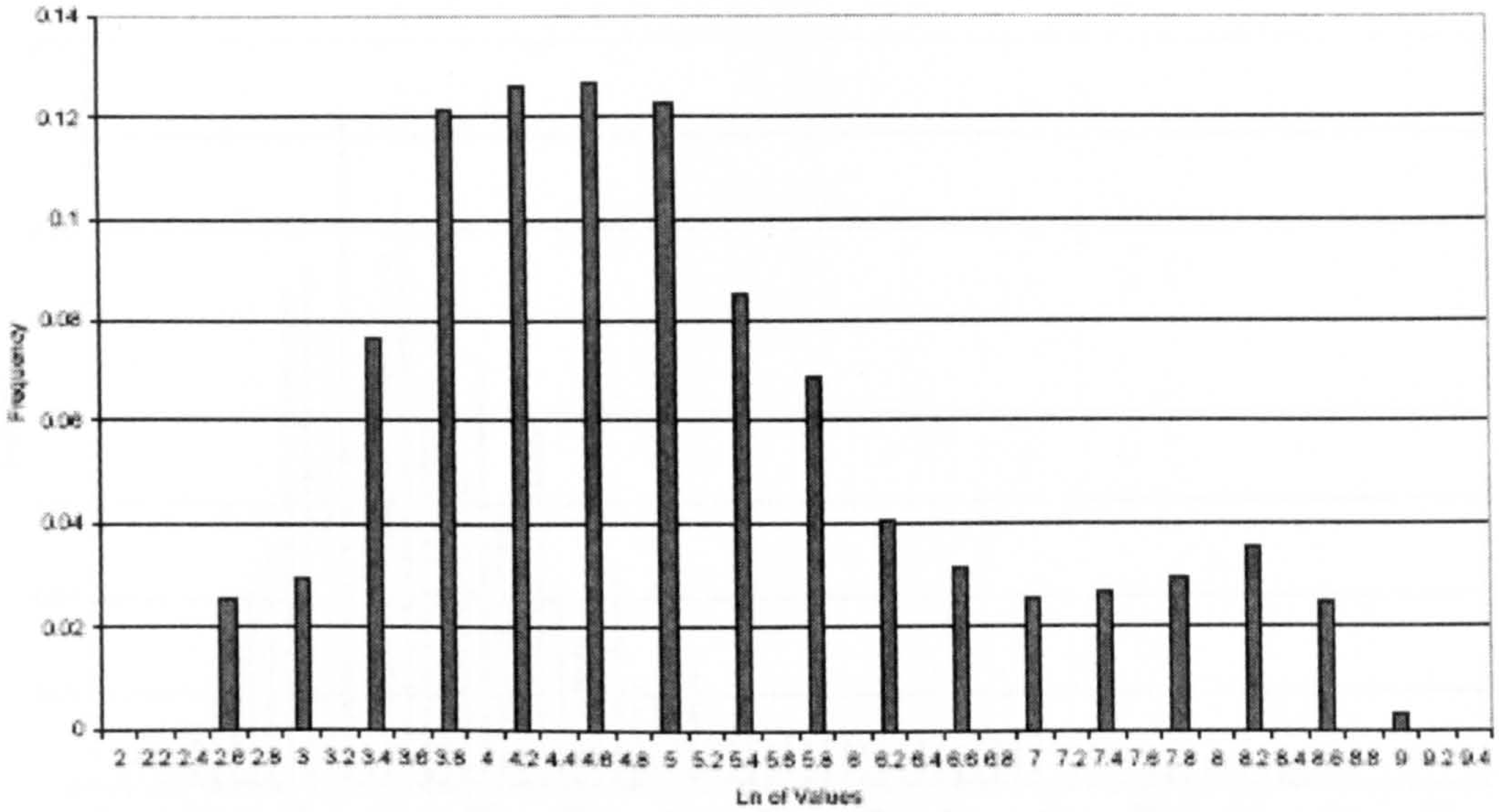


Figure 2.6: Sales Distribution

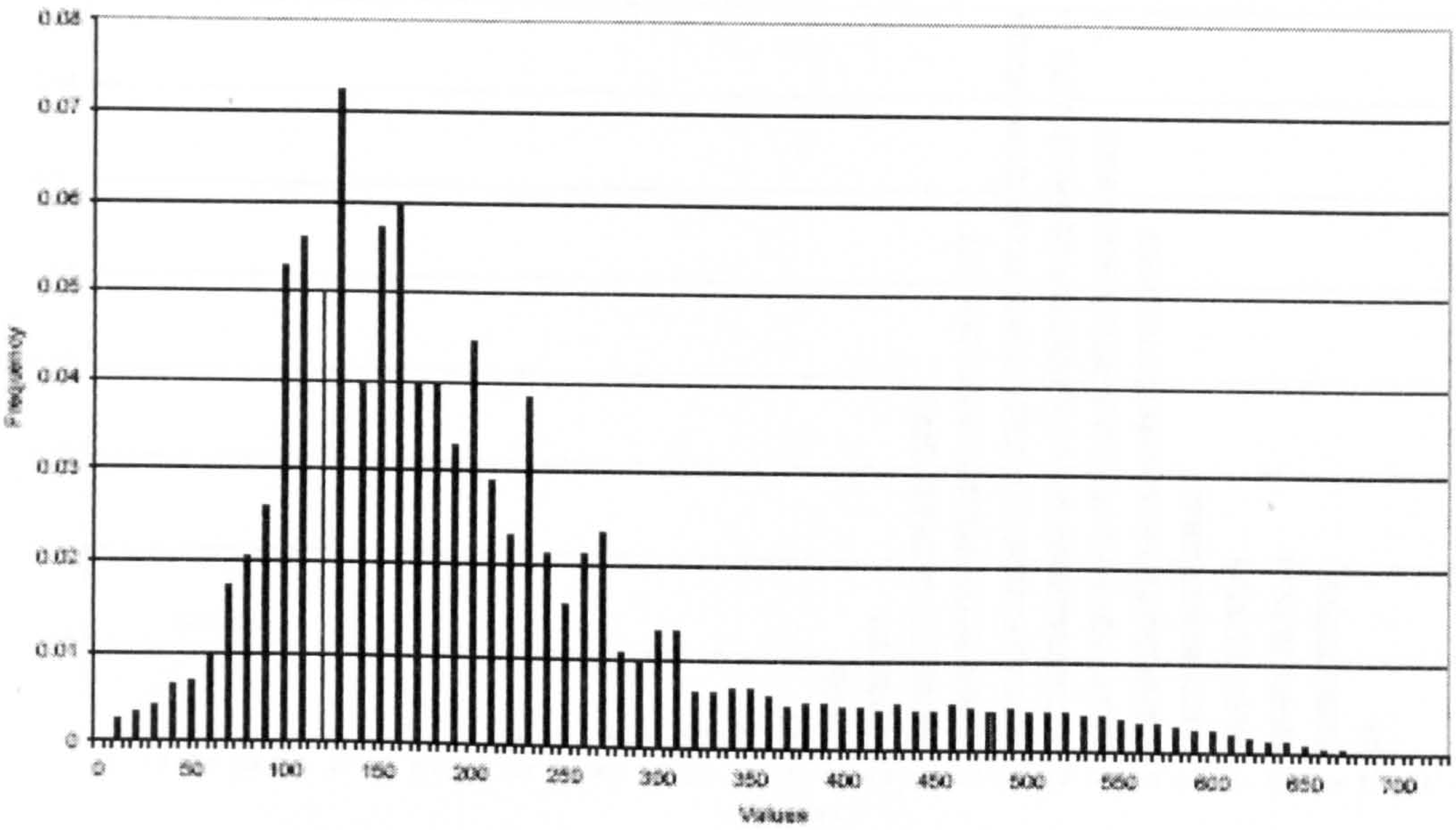


Figure 2.7: Log Sales Distribution

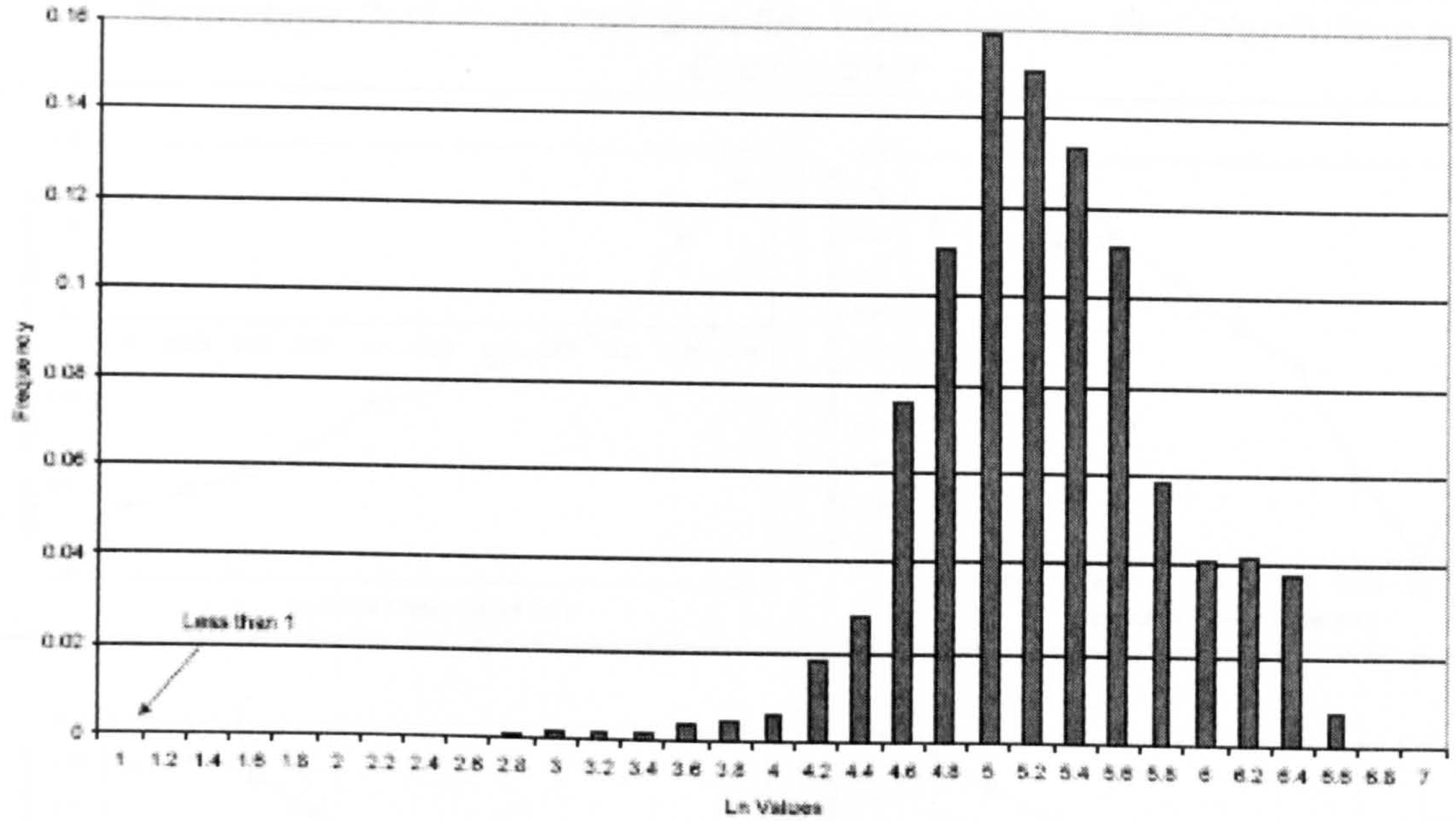


Figure 2.8: Fixed Costs Effects (Levels)

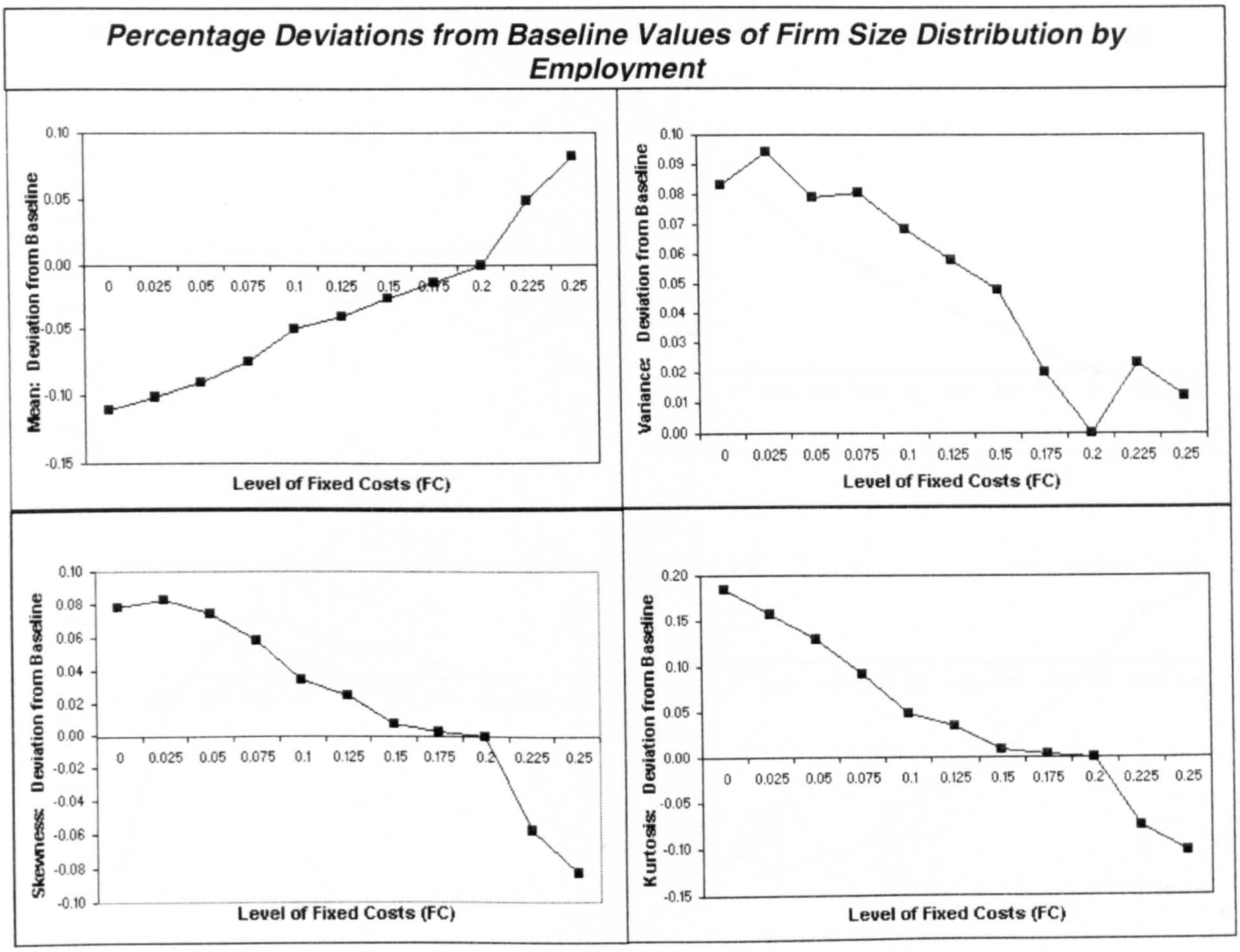


Figure 2.9: Fixed Costs Effects (Logs)

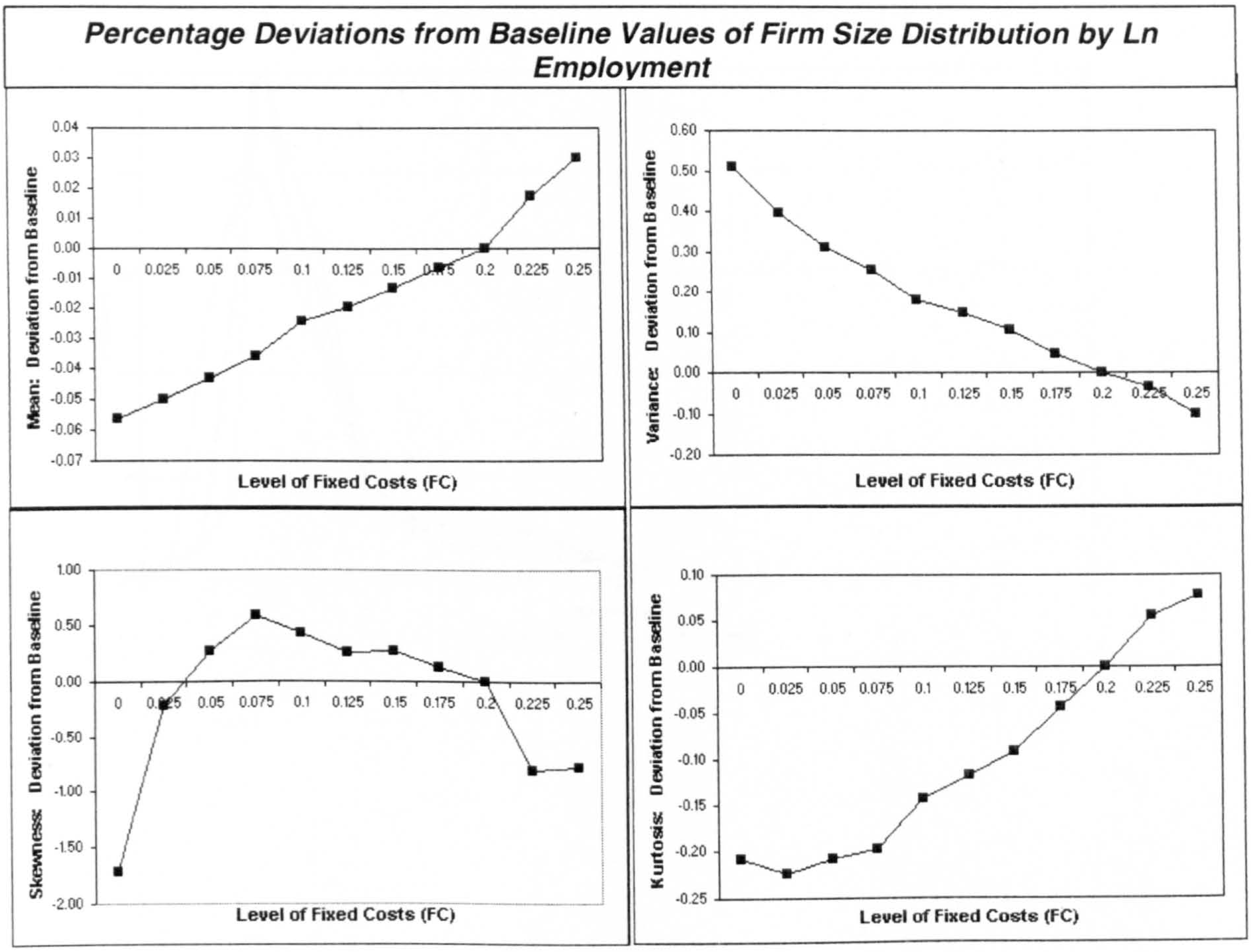


Figure 2.10: Effects of Fixed Costs on Firm Size Distribution (Levels)

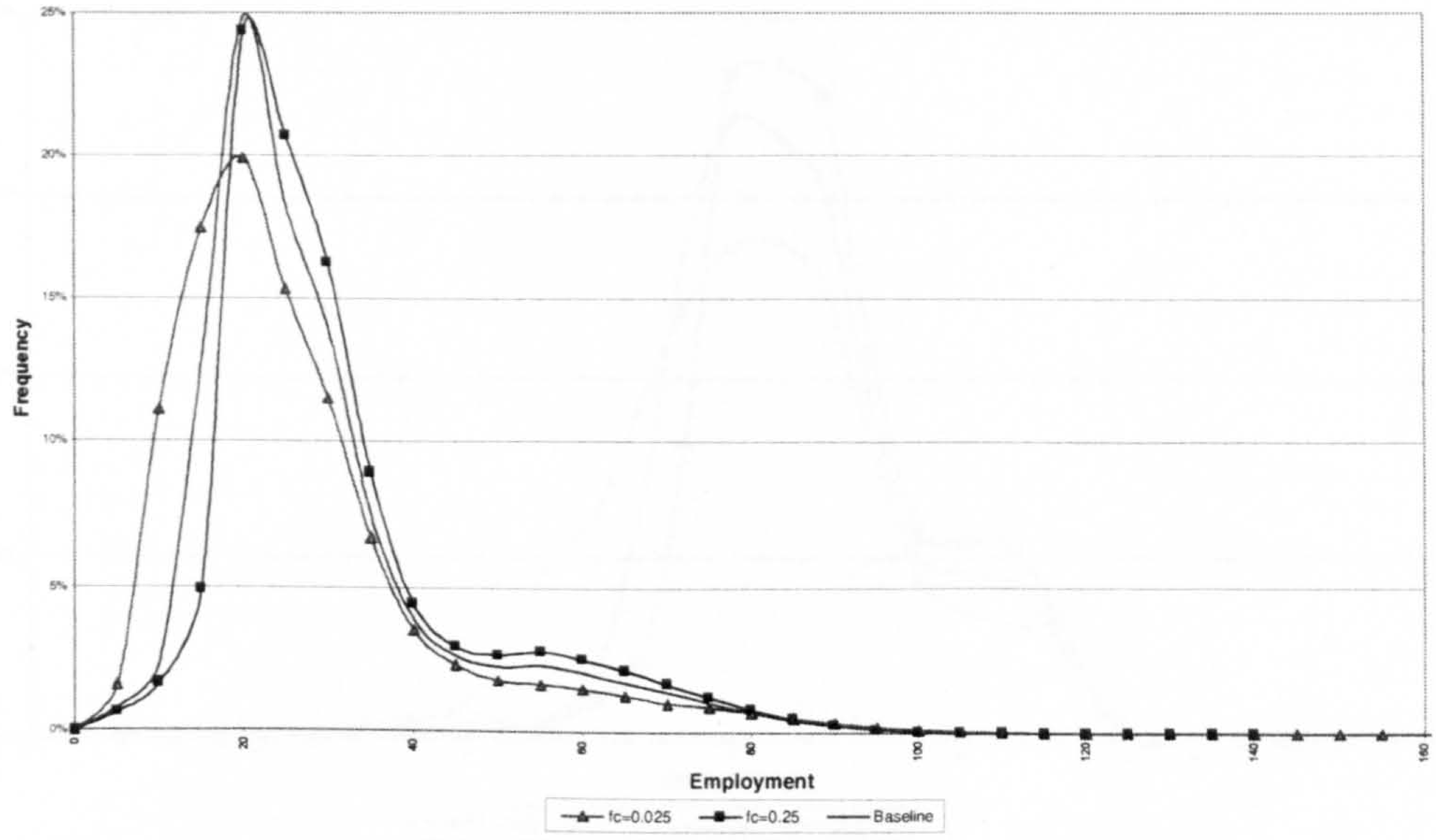


Figure 2.11: Effects of Fixed Costs on Firm Size Distribution (Logs)

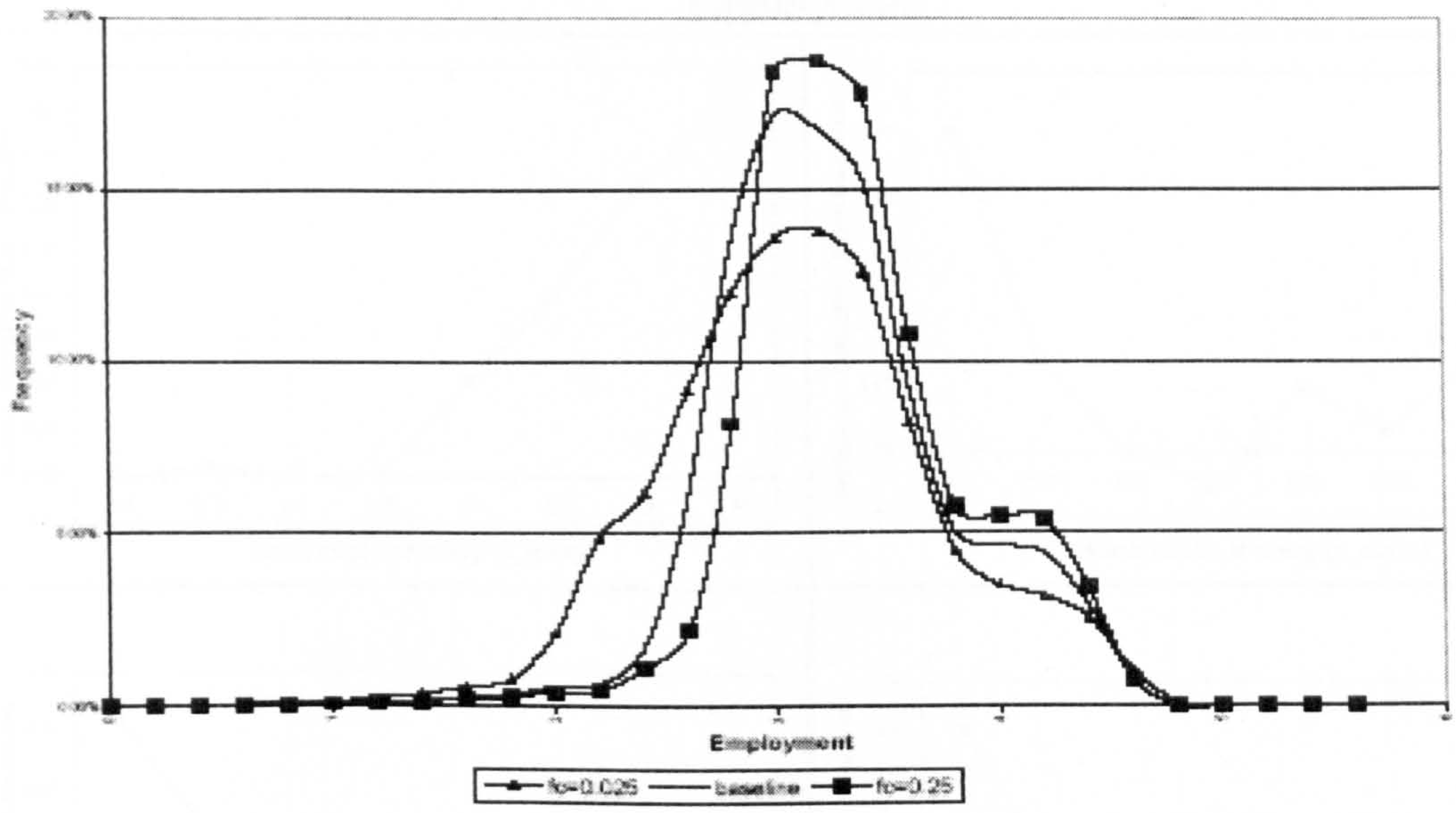


Figure 2.12: Sunk Costs Effects (Levels)

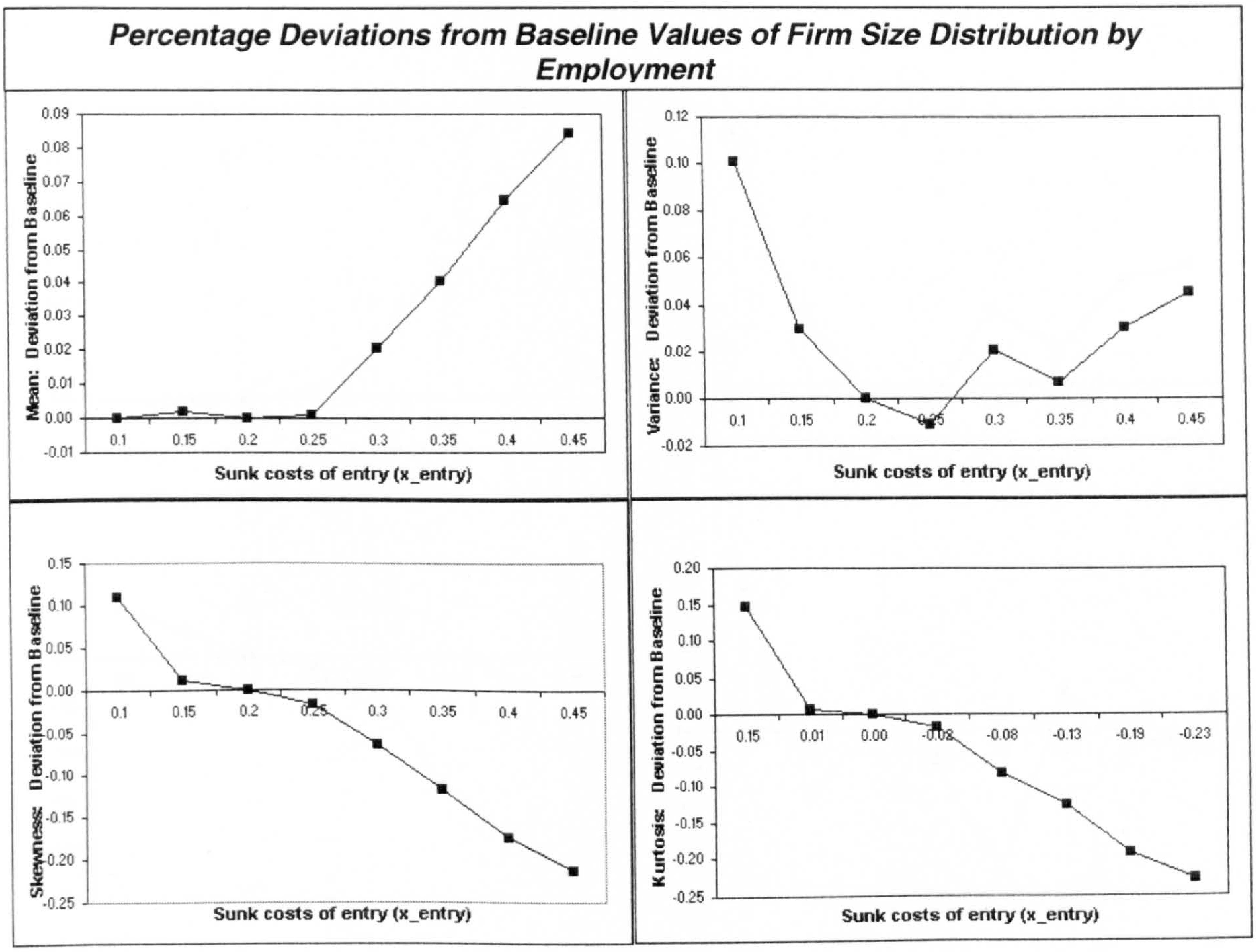


Figure 2.13: Sunk Costs Effects (Logs)

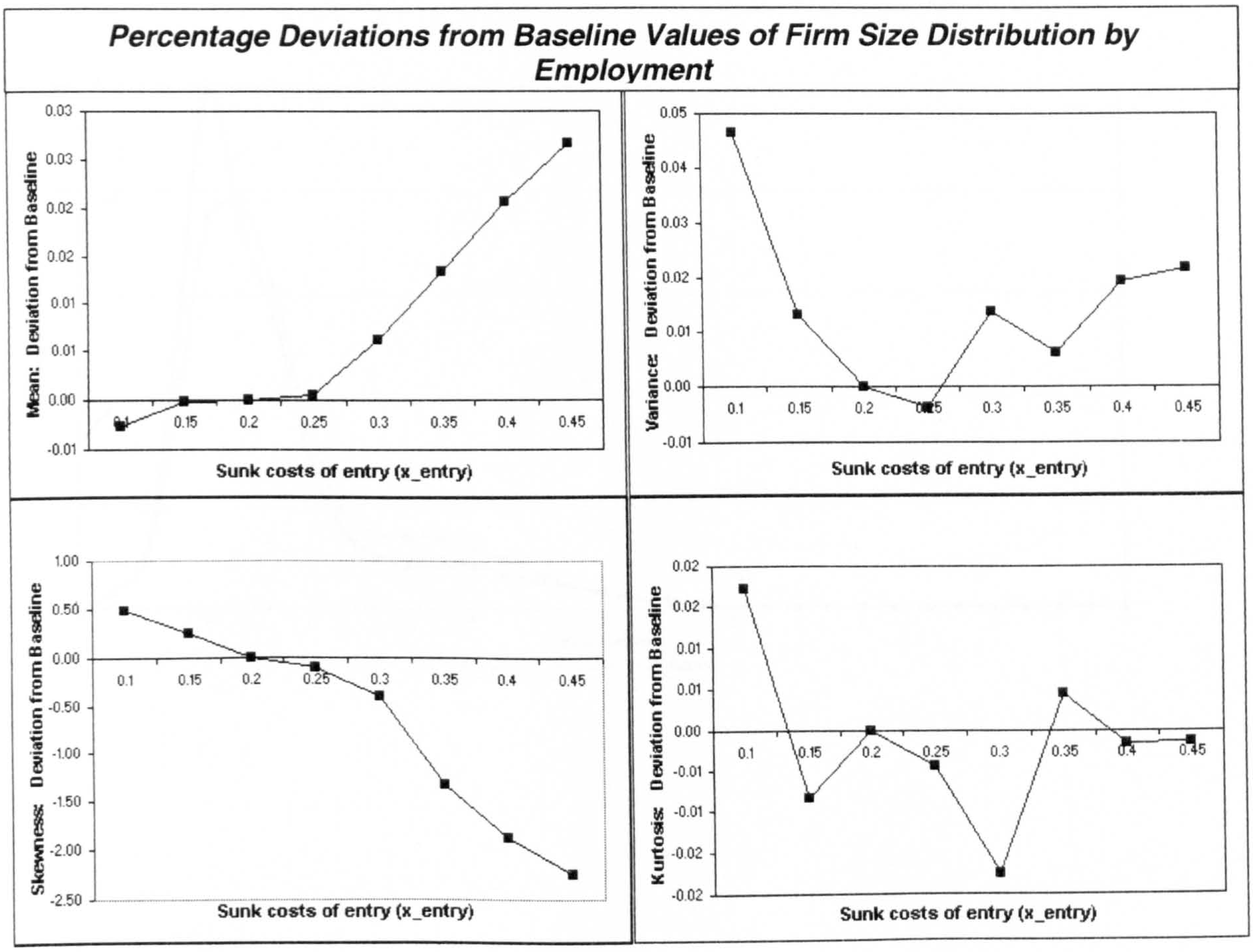


Figure 2.14: Effects of Sunk Costs on Firm Size Distribution (Levels)

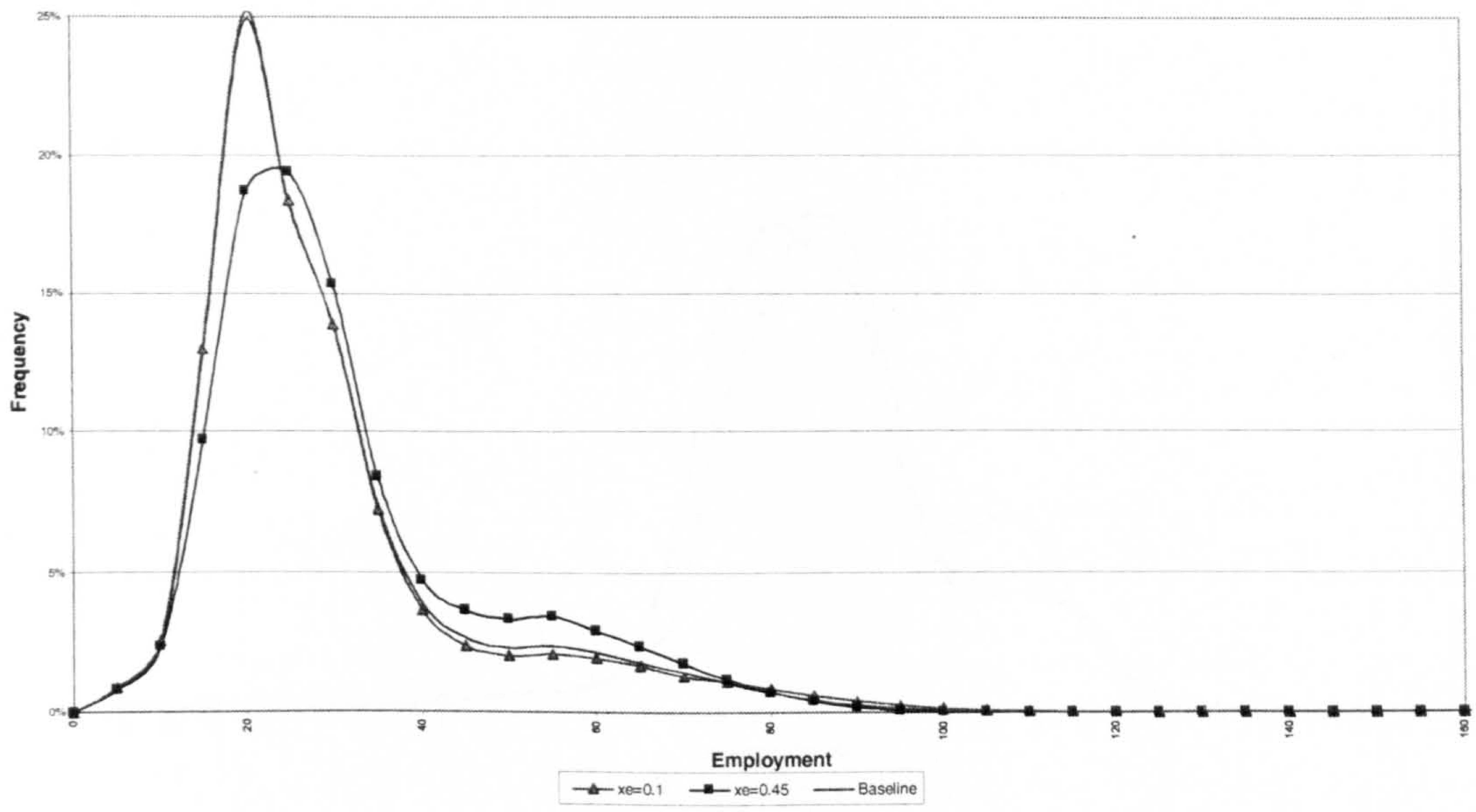


Figure 2.15: Effects of Sunk Costs on Firm Size Distribution (Logs)

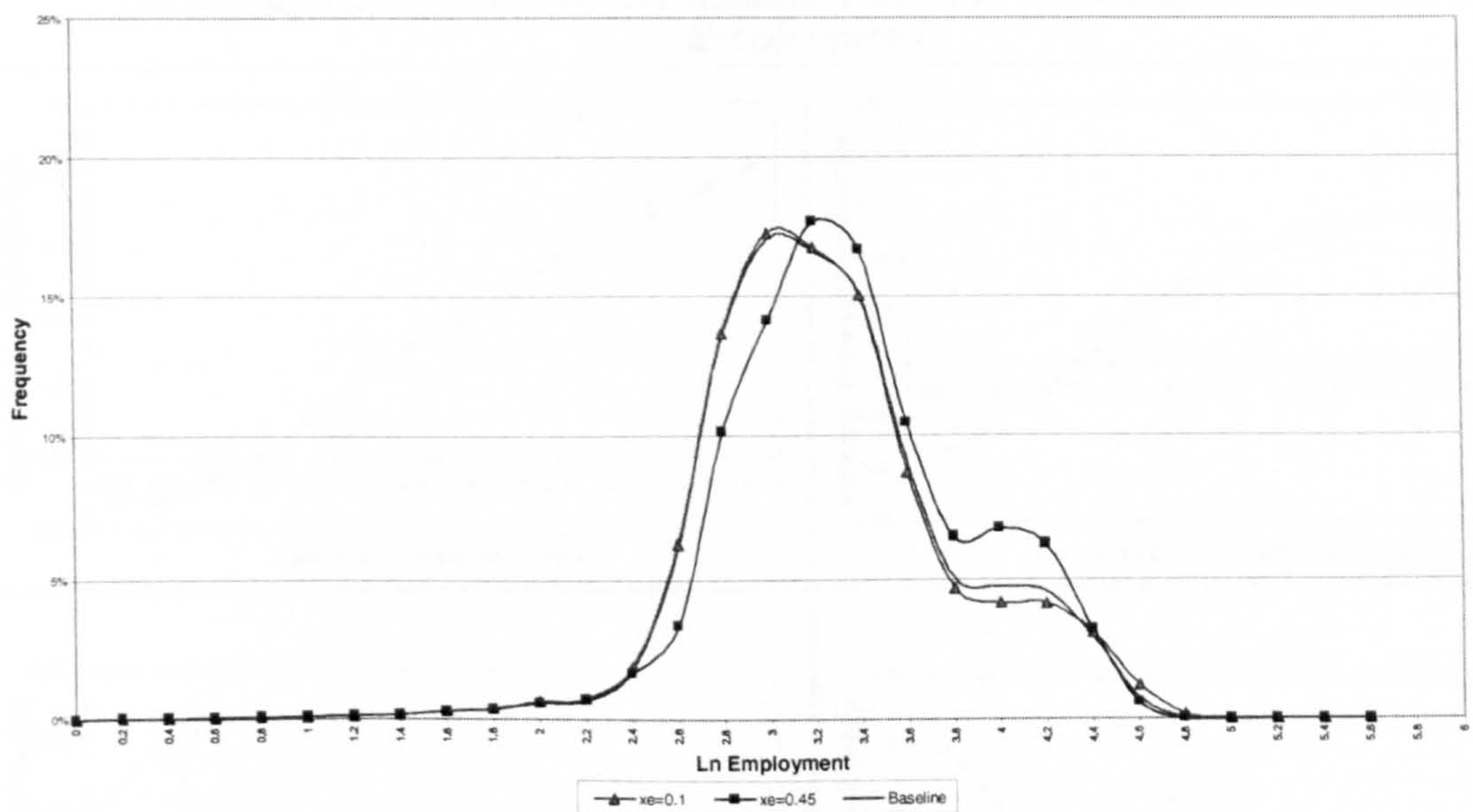


Figure 2.16: Rate of Cost Reduction Effects (Levels)

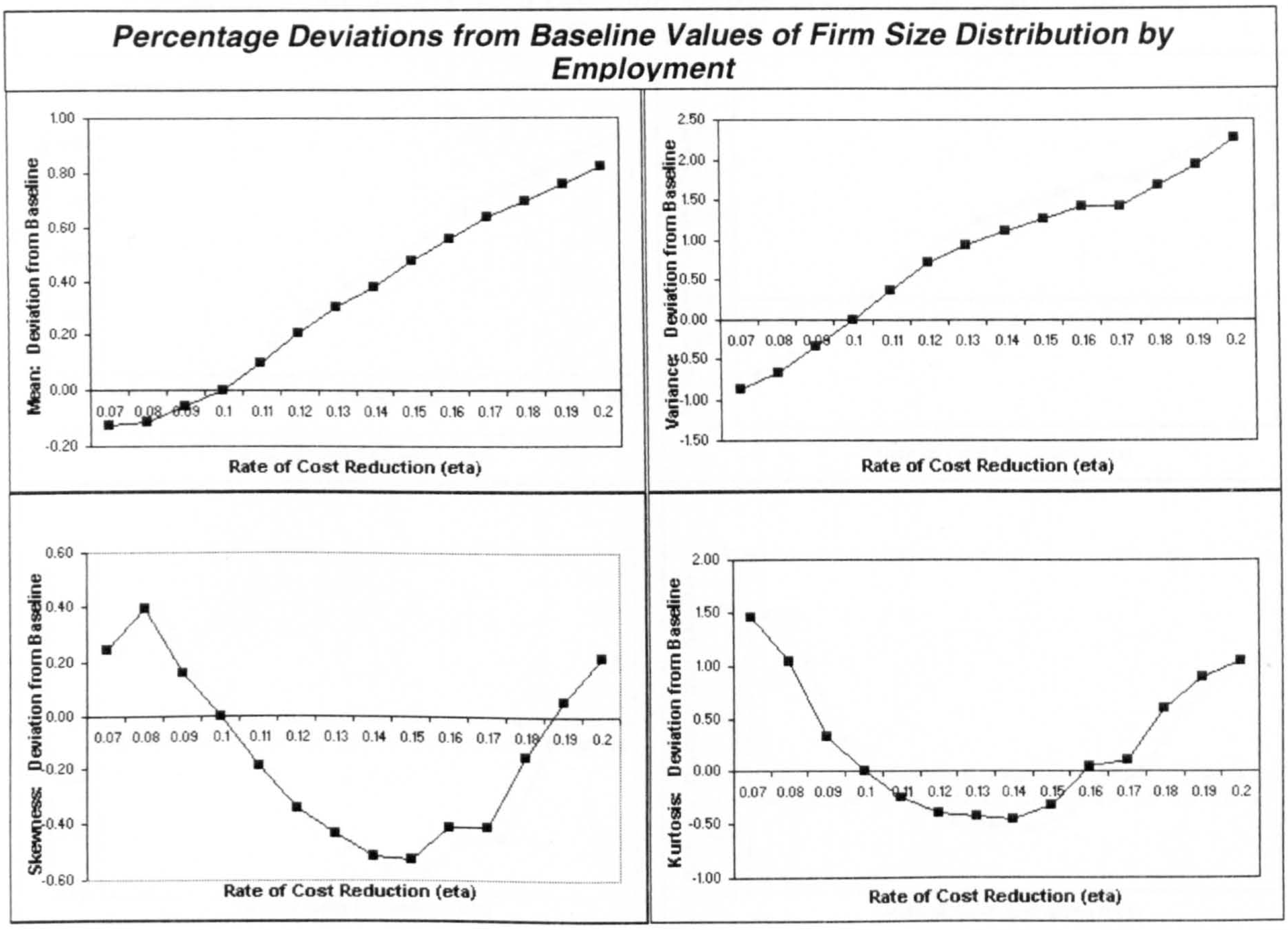


Figure 2.17: Rate of Cost Reduction Effects (Logs)

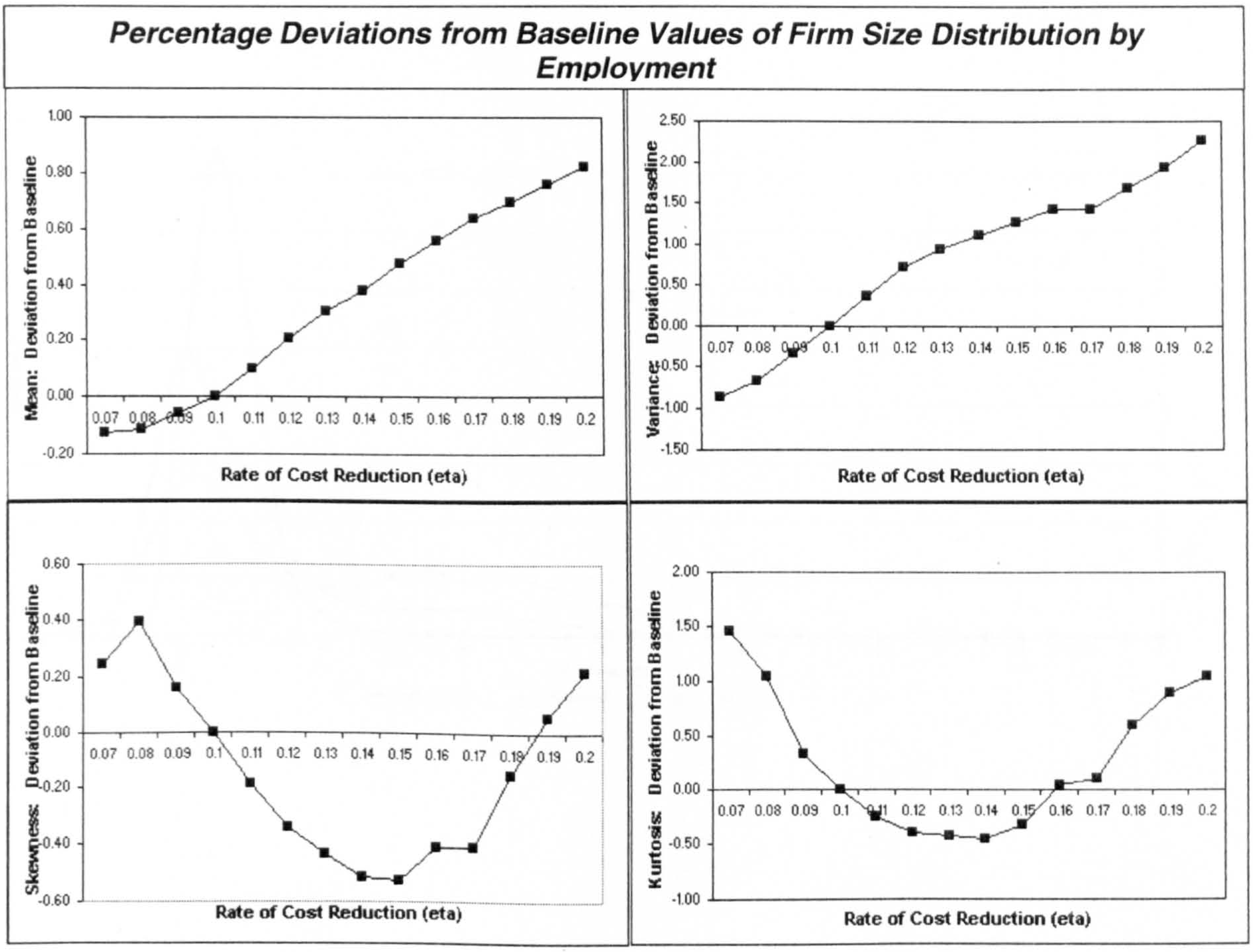


Figure 2.18: Effects of the Rate of Cost Reduction on Firm Size Distribution (Levels)

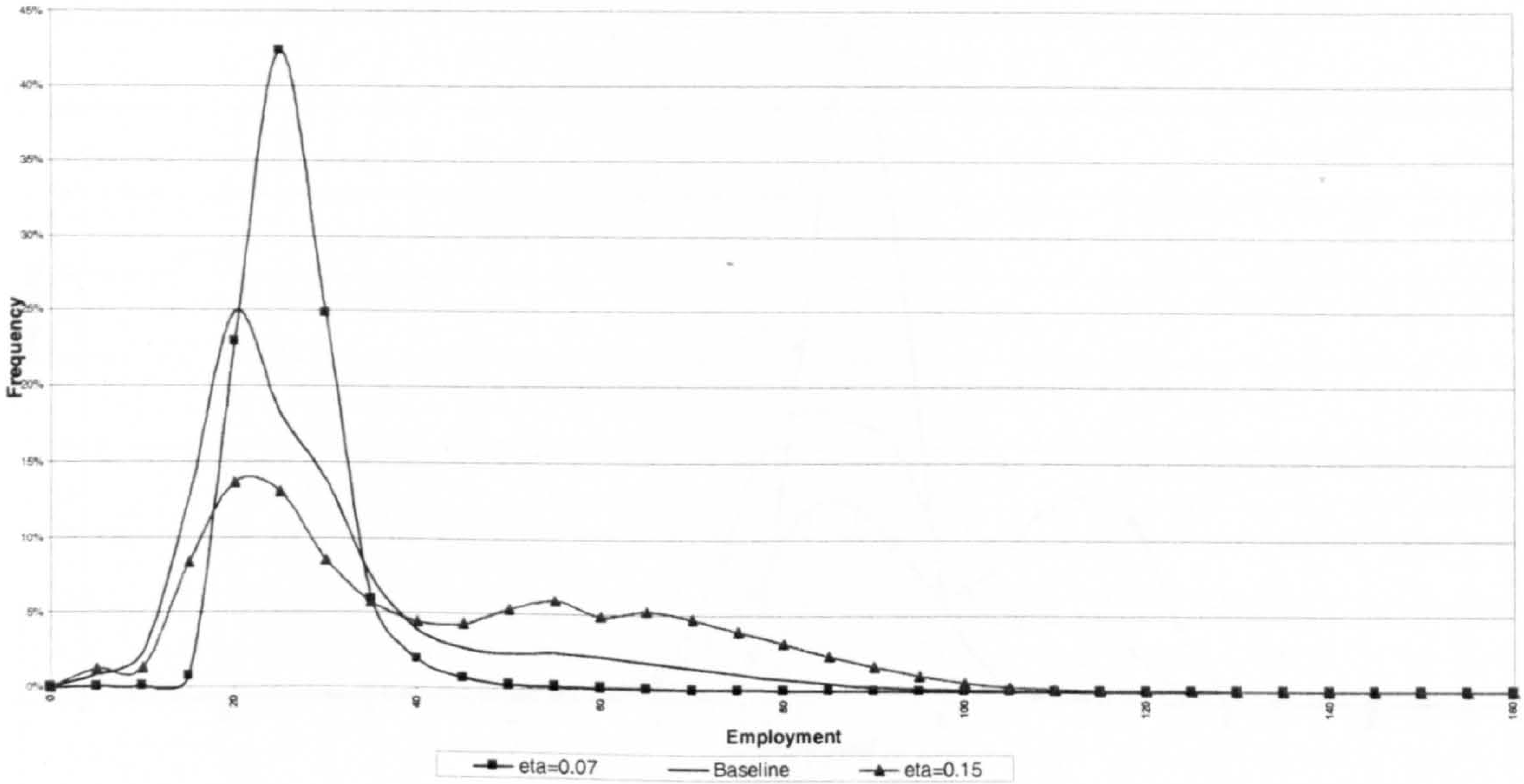


Figure 2.19: Effects of the Rate of Cost Reduction on Firm Size Distribution (Logs)

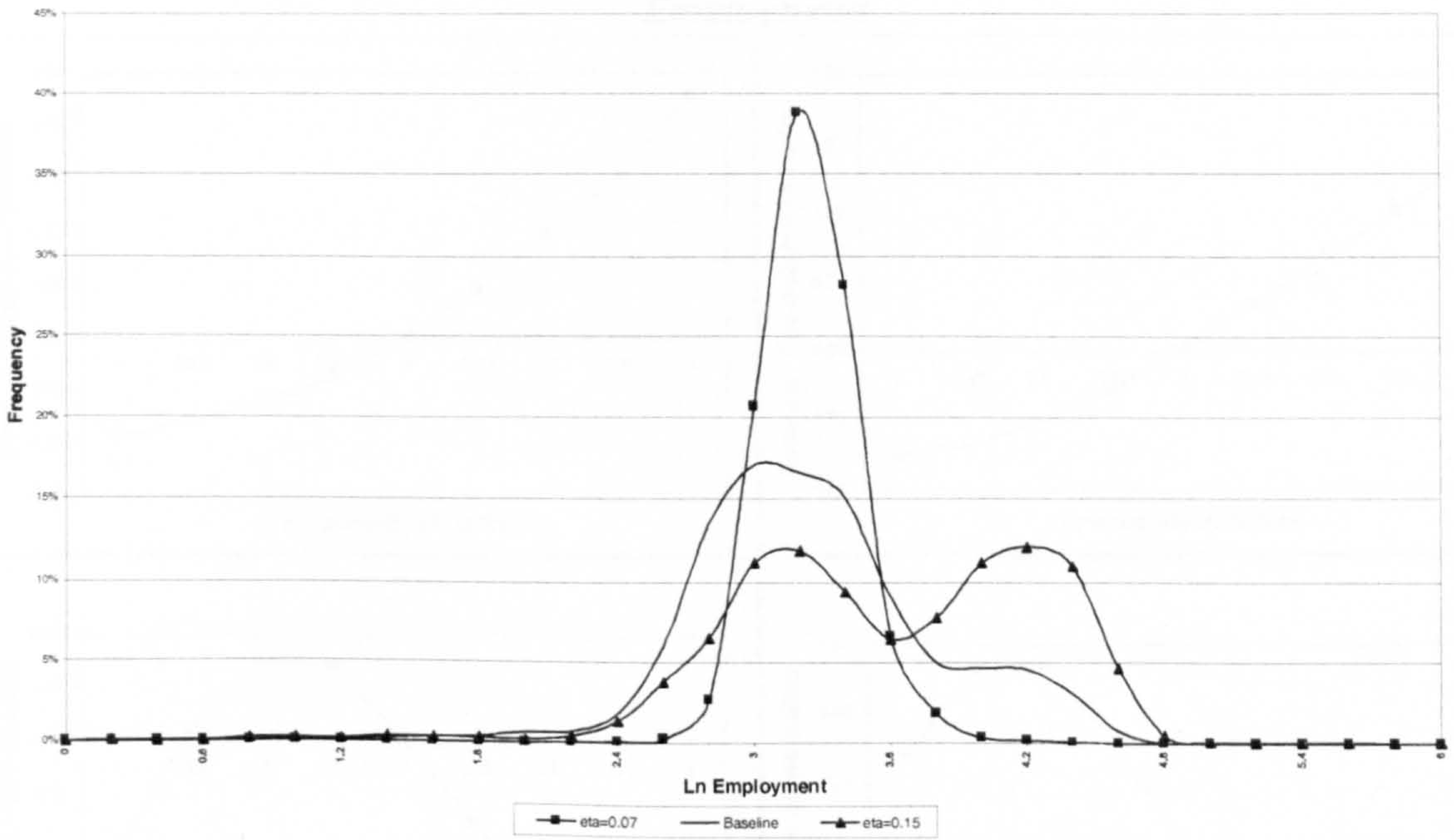


Figure 2.20: Technological Opportunity Effects (Levels)

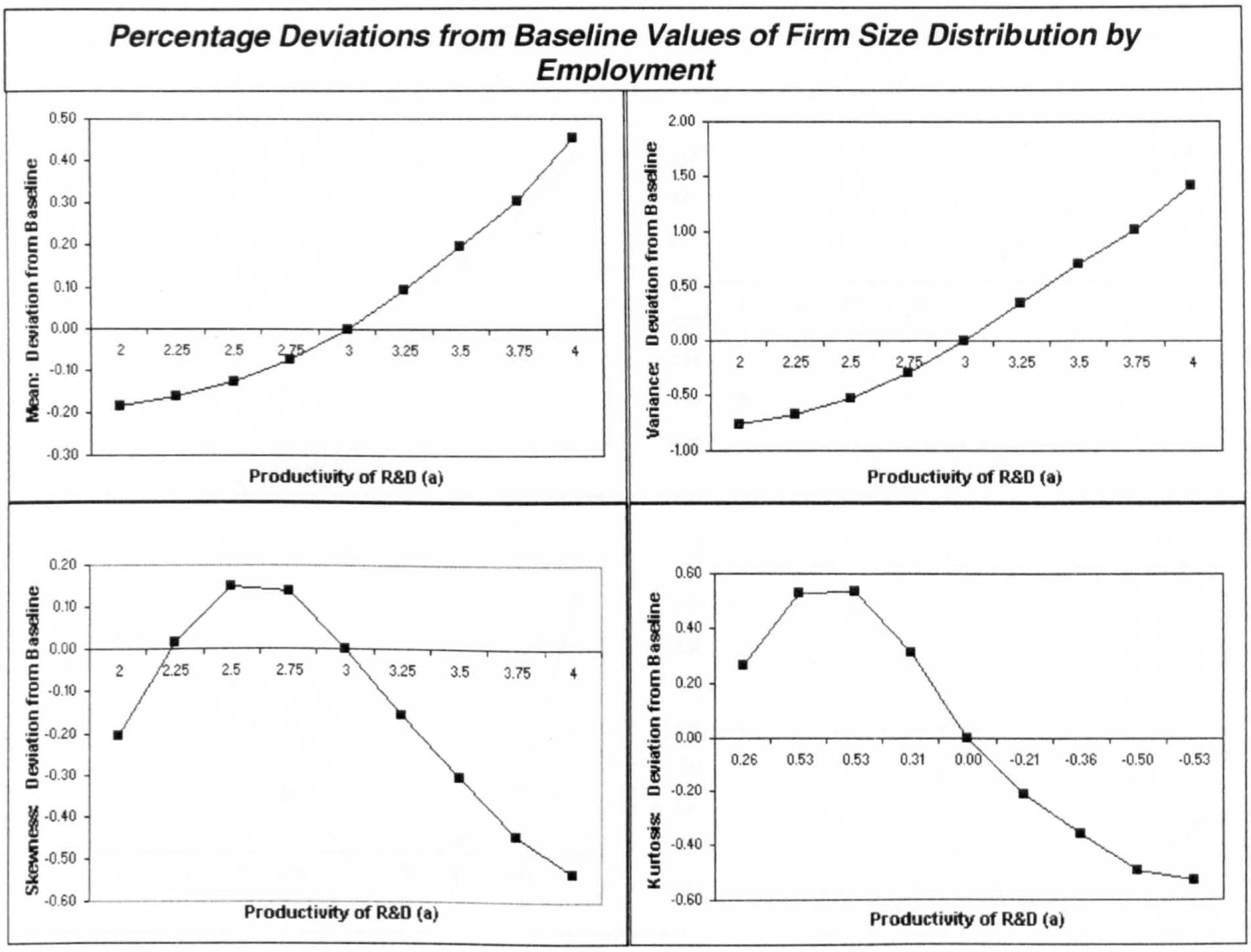


Figure 2.21: Technological Opportunity Effects (Logs)

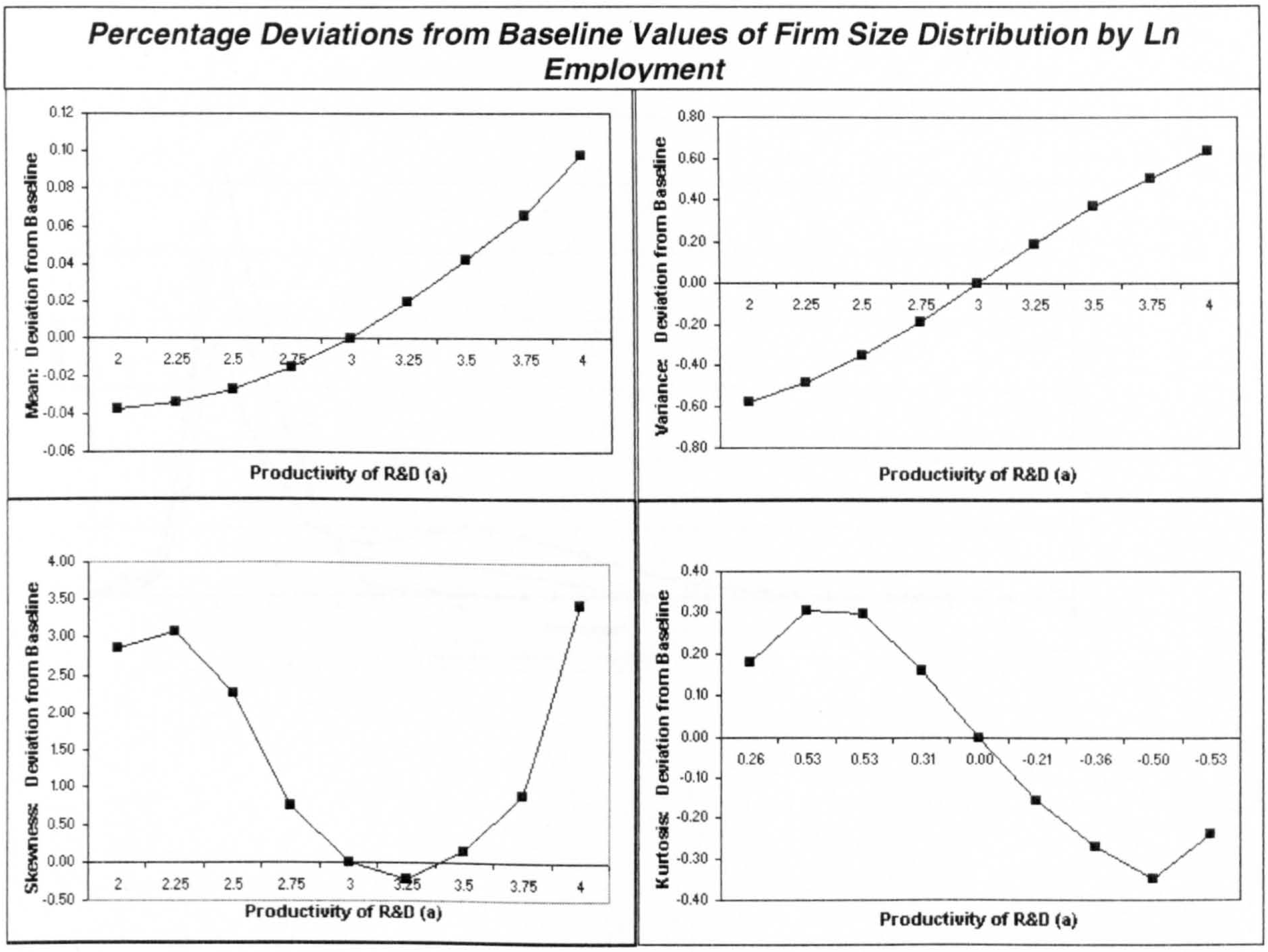


Figure 2.22: Effects of the Productivity of Investment on Firm Size Distribution (Levels)

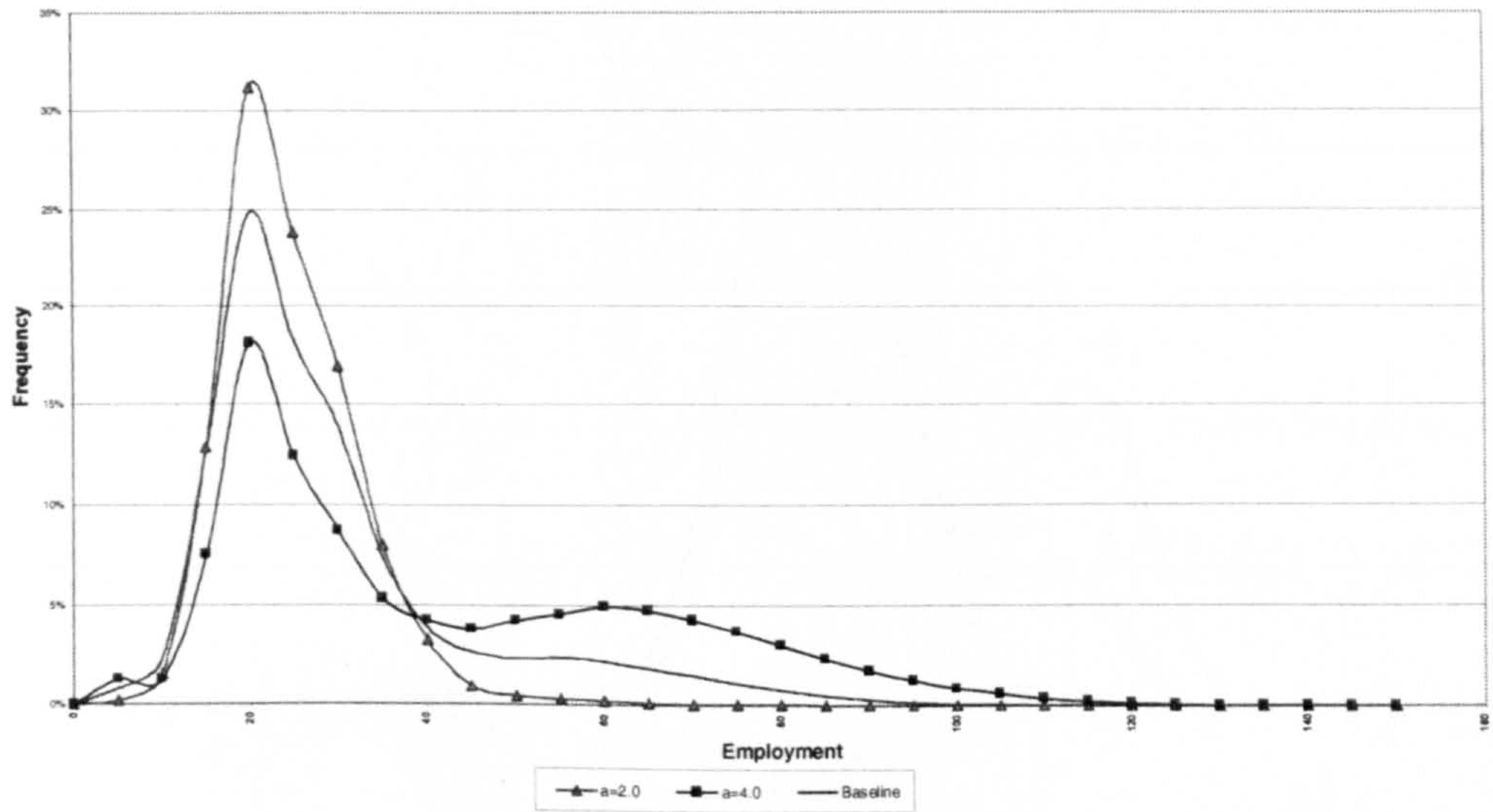


Figure 2.23: Effects of the Productivity of Investment on Firm Size Distribution (Logs)

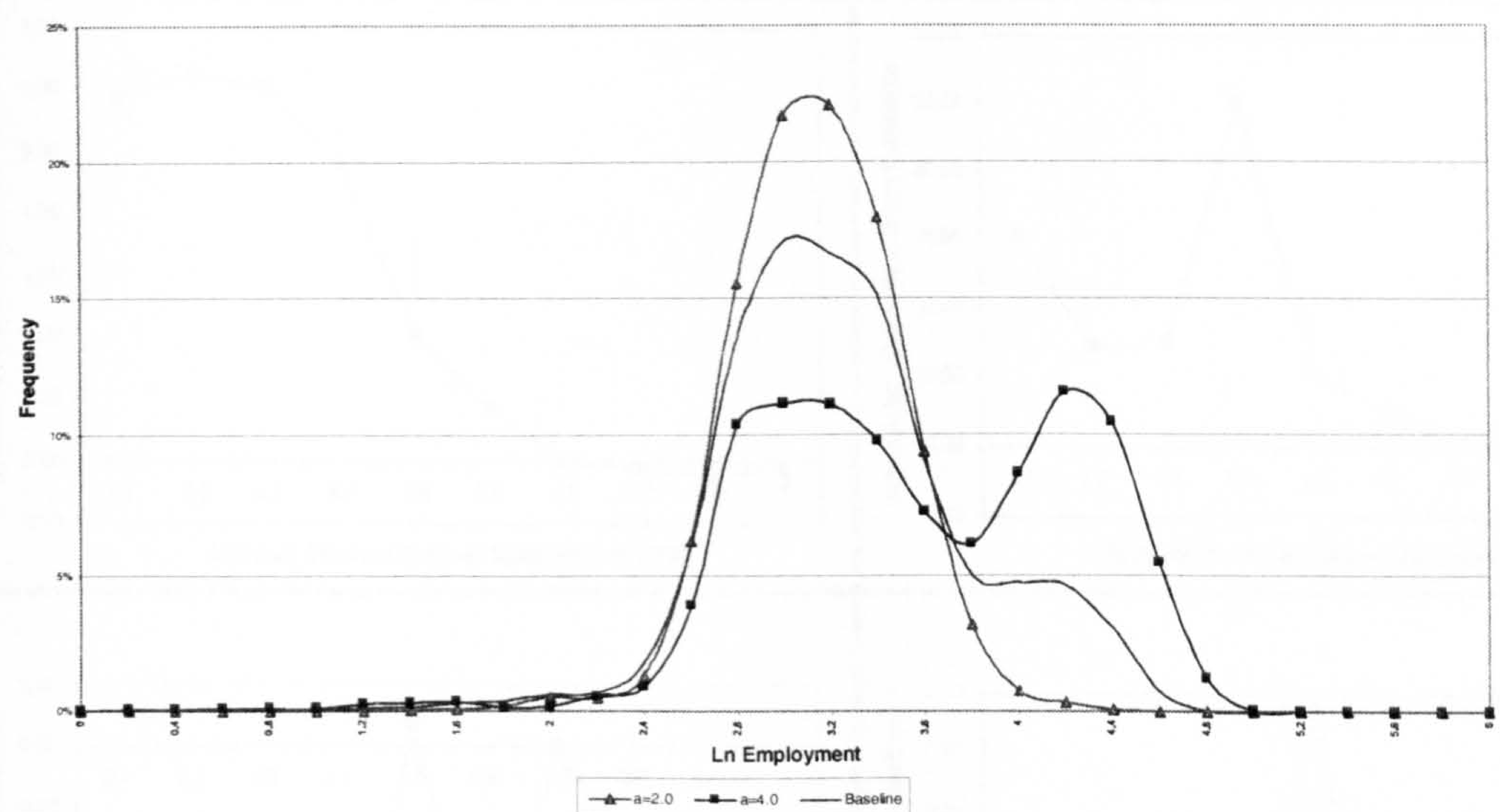


Figure 2.24: The Rate of Technological Spillovers Effects (Levels)

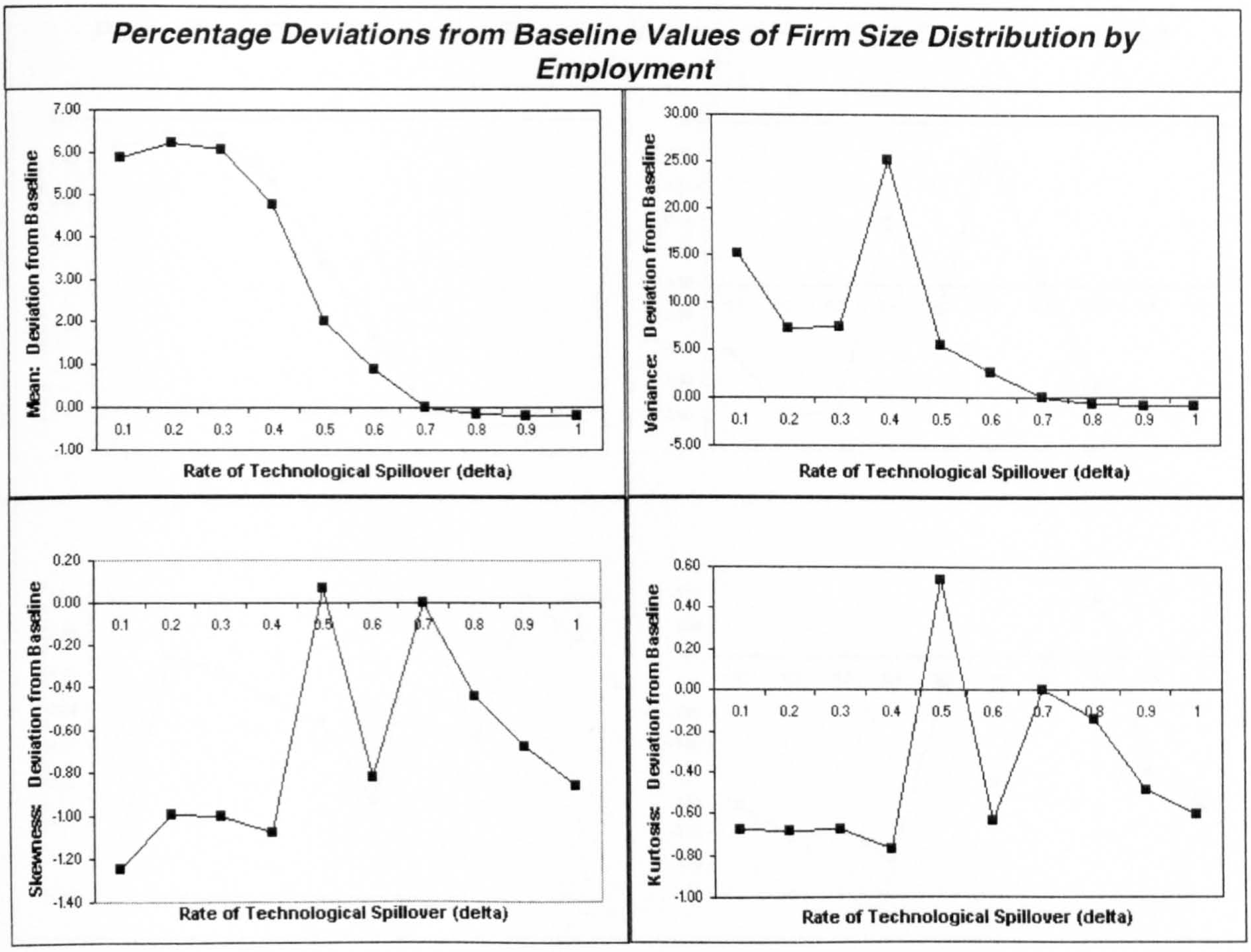


Figure 2.25: The Rate of Technological Spillovers Effects (Logs)

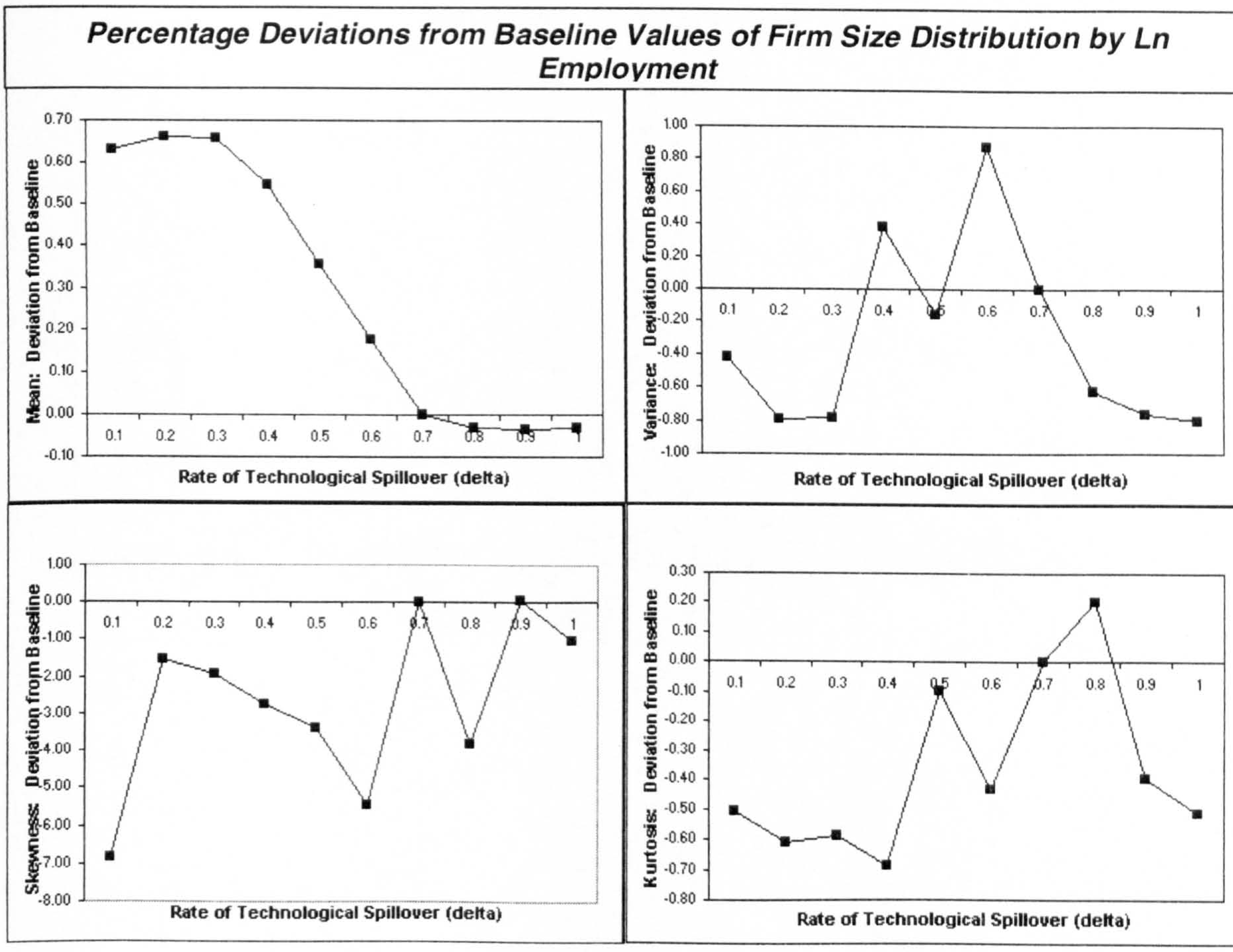


Figure 2.26: Effects of the Rate of Spillovers on Firm Size Distribution (Levels)

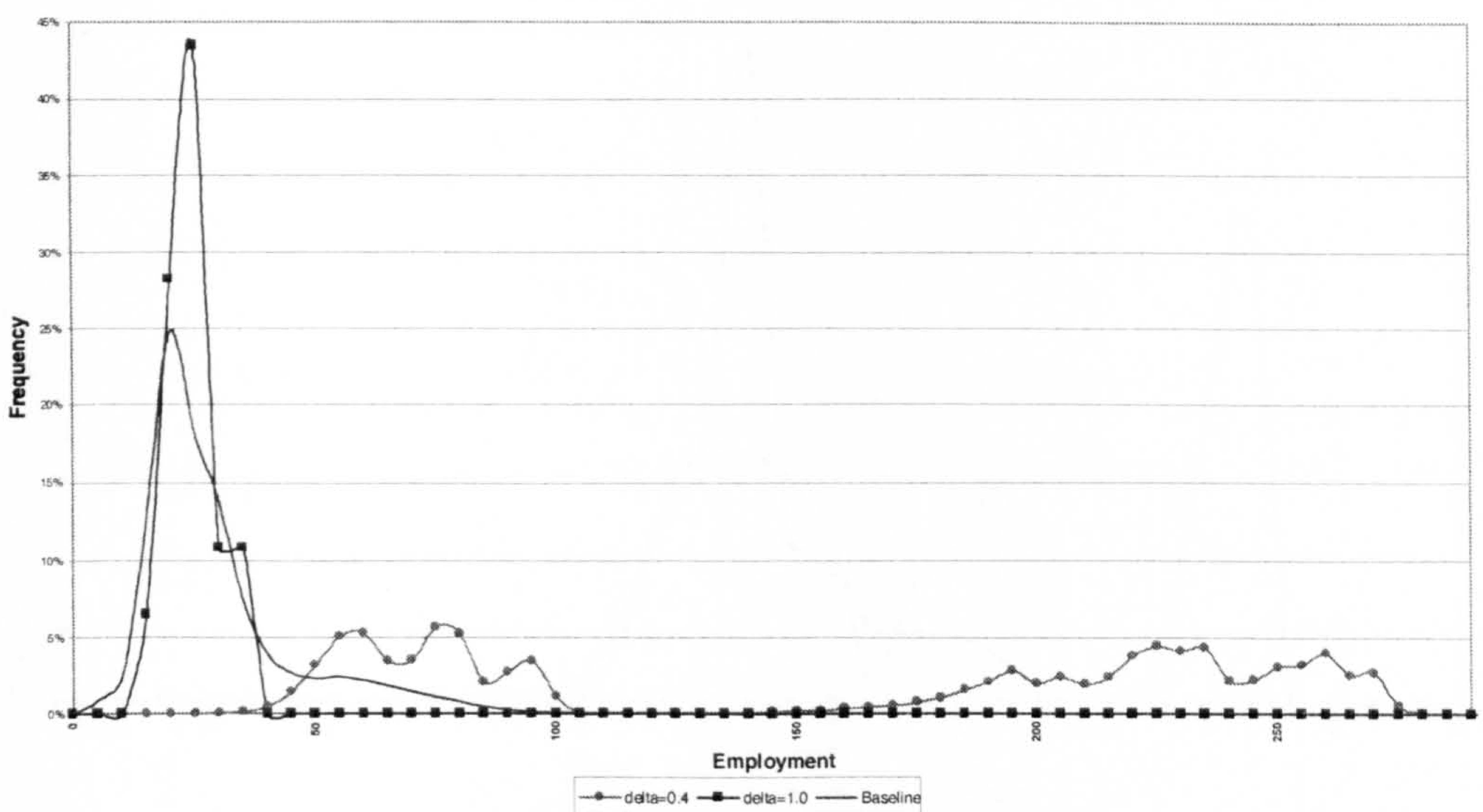
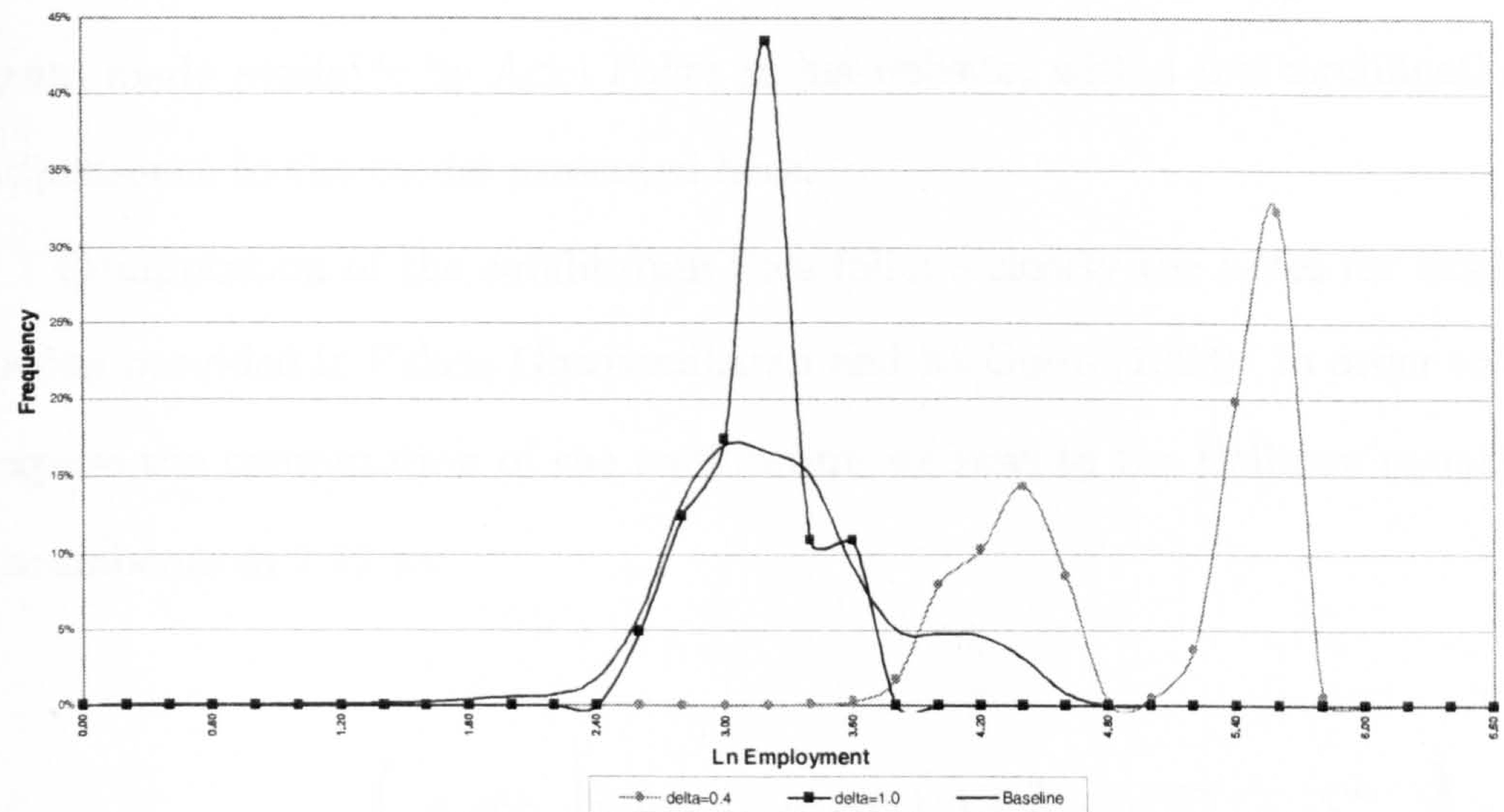


Figure 2.27: Effects of the Rate of Spillovers on Firm Size Distribution (Logs)



2.7 Appendix

In finding the Markov Perfect Nash Equilibrium of the game presented in Chapter 2 we follow the Pakes McGuire algorithm. The code utilized to implement the algorithm is a C programming language version of the original Gauss program made available by Ariel Pakes in his website, with a few modifications for adjustment to the model presented here.

Computation of the equilibrium thus follows closely the notes for implementation provided in Pakes, Gowrisankaran and McGuire (1993). In order to better expose the computation of the equilibrium we rewrite the Bellman equation for incumbents in 2.13 as:

$$V(i,s_m) = \max \left\{ \begin{array}{c} \phi, \sup_{x \geqslant 0} \Big[\pi(i,s) - cx + \left(\frac{1}{1+r}\right) P^E(s_m) \sum_{\nu_1=0,1} \cdots \sum_{\nu_N=0,1} \\ V[i', s'_m] \Pr[\nu_1 \,| x_1^{j-1}, \Delta] \ldots \Pr[\nu_N \,| x_N^{j-1}, \Delta] \Pr[\Delta] \\ + \left(\frac{1}{1+r}\right) [1 - P^E(s_m)] \sum_{\nu_1=0,1} \cdots \sum_{\nu_N=0,1} V[i', s'_m] \\ \Pr[\nu_1 \,| x_1^{j-1}, \Delta] \ldots \Pr[\nu_N \,| x_N^{j-1}, \Delta] \Pr[\Delta] \Big] \end{array} \right\} \quad (2.20)$$

where primes stand for next period values, N is the number of firms active in the industry, $P^E(s_m)$ stands for the probability of entry conditional on the current market structure on submarket m, ν is as defined in equation (2.11), i.e., captures the outcome of the R&D process, Δ corresponds to the realizations of the diffusion process which occurs with probability δ and j stands for iteration.

The value of the firm is the maximum between the value associated with shut down (exit) and the continuation value in the industry. The continuation value is given by the firm's current profits, minus its investment costs plus the discounted expected value next period, considering all possible market structures to which the current market structure might evolve, weighted by their probability of occurrence. These probabilities are those associated with the outcomes of success, for the firm and its rivals, the probability of a spillovers process (diffusion of a private innovation) and the probability of entry, given the current market structure. If we exclude, from this expression, the probabilities of success for the firm making the decision, and define:

$$\begin{gathered} C(i,s_m) = \left(\tfrac{1}{1+r}\right) P^E(s_m) \sum_{\nu_1=0,1} \cdots \sum_{\nu_N=0,1} V[i',s'_m] \Pr[\nu_1 \left| x_1^{j-1}, \Delta\right] \ldots \\ \ldots \Pr[\nu_N \left| x_N^{j-1}, \Delta\right] \Pr[\Delta] + \left(\tfrac{1}{1+r}\right) \left[1 - P^E(s_m)\right] \sum_{\nu_1=0,1} \cdots \sum_{\nu_N=0,1} V[i',s'_m] \\ \Pr[\nu_1 \left| x_1^{i-1}, \Delta\right] \ldots \Pr[\nu_N \left| x_N^{i-1}, \Delta\right] \Pr[\Delta] \end{gathered} \tag{2.21}$$

as the value of the firm given the outcome of its R&D process. This notation is similar to that used in the Ericson Pakes literature (e.g. Doraszelski and Pakes (2006), page 14). After setting this notation, one can write the Bellman equation as:

$$V^I(i,s_m) = \max \left\{ \begin{array}{c} \phi, \ \sup_{x \geqslant 0} \left[\pi(i,s_m) - cx + \right. \\ \left(\tfrac{1}{1+r}\right) \tfrac{ax}{1+ax} C(ip'+1, s'_m) + \left(\tfrac{1}{1+r}\right) \tfrac{1}{1+ax} C(ip', s'_m) \left. \right] \end{array} \right\} \tag{2.22}$$

The first order condition with respect to x yields the following solution for R&D investment, i.e., the optimal firm's action:

$$x(ip', s'_m) = \frac{-1 + \sqrt{\frac{a(C_1 - C_2)}{(1+r)c}}}{a} \tag{2.23}$$

where $C_1 = C(ip'+1, s'_m)$ denotes the value associated with the successful innovation outcome (when the firm increases its private innovation stock, ip) and $C_2 = C(ip', s'_m)$ denotes the value of the firm in case it fails to develop a further innovation. The expression derived above is an analytical formula for R&D investment given that it remains active and the nonnegative constraint on x. The algorithm uses value function iteration for the state space comprising the firm's own and rivals' positions. The algorithm starts with an initial guess for next period values and all firms' investment levels. The choice of R&D by the firm with the highest efficiency level is obtained directly from the optimal level of investment above eq.(2.23). This new information is used to calculate the level of investment for the firm next in the rank of private innovations, and similarly for the next firm and so forth. Each iteration of the algorithm starts with the values and investment outputted from the last iteration. Thus, the optimal level of investment is obtained directly from the analytical expression for investment, by deriving $C(ip', s'_m)$ and rival's investment using the last iterations' value and investment (except for a firm's own cohorts of higher efficiency). This iterative procedure is computed until convergence in values occurs, i.e., when the changes in values between iterations is bellow some tolerance level. A last note to say that the values and investment are outputted from each iteration in matrix form, each row containing the positions of firms in a given market structure, the matrix containing all possible market structures.

Therefore, the algorithm iterates on the values that result from firms' actions to find the optimal set of actions for each firm in a given market structure.

The code used to find the equilibrium policy functions and simulate the model to extract the ergodic distribution of market structures in this chapter makes modifications to the original Pakes, Gowrisankaran and McGuire code in order to adjust for the functional forms utilised in the model stated. However, the numerical methodology to find equilibrium policy function remains fairly the same, and thus, no changes to the core of the Pakes McGuire algorithm are made, and the notes for implementation provided in Pakes, Gowrisankaran and McGuire (1993) apply fully to the solution methodology used to find the equilibrium policy functions of the model presented here. The most important difference that should be highlighted is the fact that the stochastic exogenous process in Ericson-Pakes (1995) deteriorates every firm's states upon positive realisations. In the model presented here, this stochastic exogenous process is a diffusion process of a private innovation to the public innovation stock, therefore decreasing firms' private stock of innovation, but leaving the total number of innovations that determine the level of innovations unchanged because of the increase in the stock of public innovations. The impact of the spillover process, however, is that the gap in the innovation stock of incumbents and potential entrants is reduced with the diffusion process. This difference is introduced into the code to implement the algorithm to find firms' equilibrium policy functions and simulate the evolution of the industry's market structure.

Existency

Ericson and Pakes (1995) set out a number of conditions for the existence of the Markov Perfect Nash Equilibrium of their model, which hold in the model presented here. However, contrary to what they claim in the paper, the existence of equilibrium in their model of industry dynamics requires admissibility of mixed entry/exit strategies. Computing mixed strategy equilibrium increases the computational complexity of the algorithm immensely, and existing algorithms cannot cope with that. Doraszelski and Satterthwaite (2007) discuss this problem intensively and propose an incomplete information argument to overcome the problem. They propose a model analogous to that of Ericson Pakes (1995), but with a further element of heterogeneity in the form of randomly drawn, privately known scrap values and setup costs into the model. Although agents are assumed to have pure strategies, their rivals perceive them as following mixed strategies due to incomplete information. The proof of existence requires a continuous mapping from policies into themselves. So, one of the reasons for adding random scrap values/setup costs to the original E-P framework is that they allow to treat the continuous exit and entry probabilities as the policies, in contrast to the discrete entry and exit decisions we adopted in this chapter, such that there will always be an equilibrium in cutoff entry/exit strategies. Even if a formal proof for the existence of equilibrium cannot be provided in the context of the model presented here, the solutions we found are approximations to equilibriums that do exist, and our results hold. However, we can only ensure and prove the existence of this equilibrium in our model, namely under other sets of parameters, by making the changes proposed by Doraszelski and Satterthwaite (2007). Thus, so as to ensure that there exists a computationally tractable Markov Perfect Equilibrium of the model presented here in Chapter 2, for whatever parameterization chosen, similar changes should be introduced in the model presented here. The model presented in the third chapter, however, is not affected by these issues as it abstracts from entry and exit.

Multiplicity

Doralszelski and Satterthwaite (2007) develop a model analogous to that of Ericson and Pakes, but where the conditions for the existence of equilibrium are met. These conditions, however, are, as they show, sufficient but not necessary. As Pakes and Doraszelski (2006) state, just as the problem of the computational burden associated with the model, multiplicity can effectively limit the extent to which we can analyse particular applied problems. The extent to which this problem affects an application of the Ericson Pakes framework depends on the aim of the analysis. According to Pakes and Doraszelski (2006), "undoubtedly the reasonableness of any selection depends on what the subsequent analysis is to be used for (...). If we were trying to demonstrate the feasibility of a theoretical proposition or the feasibility of a particular change in outcomes as we vary some parameter, then probably any of the equilibria would be fine", which is the kind of analysis developed here.

Chapter 3

The Endogenous Determination of Market Structure and R&D Spillovers

3.1 Introduction

Initially treating technological change as an exogenous process, economic theory has moved towards a better understanding of the economic forces behind this fundamental source of growth. Technological progress is not a random event, but rather a product of the incentives for individual firms in developing new products and production processes. The negative impact of the lack of appropriability of the outcome of the R&D effort in firms' incentives to invest generates the *classical external economies problem* in the innovative activity highlighted by Nelson (1959) and Arrow (1962). Even if static efficiency is enhanced when knowledge is treated as a common pool (having almost fully appropriable R&D may lead to duplication of effort and costs and socially redundant R&D), knowledge externalities are detrimental to private incentives. We argue here that this does not necessarily lead to dynamic inefficiency.

The initial literature on the problem of free riding investigates the trade-off between private incentives to innovate and efficient use of knowledge, and tries to determine the net effect for total welfare and firms' performance. Spence (1984) develops a deterministic model where symmetric firms invest in cost-reducing R&D, and finds that the absence of appropriability leads unambiguously to sub-optimal investment in R&D, but he stresses the fact that appropriability is not the solution as market performance will remain inefficient[1]. He concludes that the highest market performance is achieved in an industry with high spillovers and subsidies to restore incentives to invest and improve allocative efficiency. In the context of a static, deterministic, duopoly model, D'Aspremont and Jacquemin (1988) show that, independently of any subsidies policy, cooperative research allows for the spillovers to be internalised by firms leading to an increase in the investment undertaken in the industry. Later work by Henriques (1990), Suzumura (1992), Simpson and Vonortas (1994) and Ziss (1994) extend the analysis of the role of joint research ventures in mitigating the classical externalities problem.

Levin and Reiss (1988) analyse R&D choices in the context of a static, deterministic model where both cost-reducing and demand creating R&D are imperfectly appropriable and find that diminished appropriability does not necessarily lead to a decrease in firms' investment effort. They argue that their model, although providing some understanding "about the relationship of R&D and market structure, it is but a small step towards understanding the more complete dynamic process where market structure evolves through time". Cohen and Levinthal (1989) show that treating external knowledge as a pure public

[1] The allocative inefficiency arises because the R&D costs required to achieve a given rate of industry cost reduction are higher than when there is a public good character to R&D.

good, acquired at zero cost by all firms in the industry, is in the genesis of the results pointing to subinvestment in the presence of R&D spillovers. They claim there are substantial costs associated to the learning and absorption of publicly accessible industry knowledge. By accounting for the absorptive role of R&D, i.e., R&D that enhances the firm's ability to make use of the industry pool of knowledge, they show that contrary to the existing literature, externalities might increase the amount of R&D investment undertaken in an industry. This is in line with Rosenberg (1974) and Nelson (1982) claims that, in order to be able to take advantage of externalities, the firm has to undertake some investment of its own, i.e., it has to have laboratories with research personnel that will use the industry pool of knowledge.

The literature on the impact of spillovers on R&D choice by firms is extremely vast, but it is mainly composed of analyses that rest on *static*, *deterministic* models, that depart from a *given market structure*, usually by assuming symmetry, and then evaluate the impact of spillovers on R&D incentives. But as argued by Dasgupta and Stiglitz (1980), market structure and R&D are both endogenous variables, strongly interacting in their determination. We propose a model to analyse the implications of imperfect appropriability of R&D on welfare and market structure in a dynamic context where market structure and R&D feedback on each other. We pursue this analysis by extending the Markov-Perfect dynamic industry model proposed by Ericson and Pakes (1995), through the introduction of a non-proprietary productivity component to R&D as part of a dynamic, stochastic process. The framework presented here captures the dynamic character of innovation and market structure. Furthermore it allows for uncertainty in the outcome of the R&D activity and for heterogeneity of firms. Kamien and Schwartz (1971) and Reignanum (1983) argue for the relevance of the assumption of uncertainty versus determinism when modelling rivalry. Reinganum shows that uncertainty is not innocuous for pre-emptive patenting. Here, uncertainty will play a role in the escape competition motive, under which firms invest to gain a comfortable distance with respect to the followers. The assumption of heterogeneity gains one further dimension in this context since a firm's state of development, as well as the state of development of its rivals, will affect the ratio of own R&D to external R&D.

We analyse the effects of externalities on the rates of innovation, concentration and welfare measures under two distinct scenarios. We first examine the *case of costless R&D spillovers,* where firms can freely assimilate and exploit the external pool of knowledge as in Spence (1984). We then compare these results with those of the *case of absorptive capacity,* where it is assumed that a firm can only have the ability to identify, assimilate and make usage of the industry's outside knowledge if it pursues some positive level of R&D expenditures. Therefore, firms will build their absorptive capacity through the R&D they pursue to develop their own technology, as in Cohen and Levinthal (1989).

We find that when firms can freely absorb and exploit external knowledge, spillovers have a negative impact on total welfare. There is, in fact, an increase in producers' welfare, which is fully driven by a decrease in the costs associated with the R&D activity. This decrease in costs is a result of the substitution of external R&D, which is free, for own R&D, which has a positive unit cost. However, these benefits will not be passed onto the consumers, because the marginal costs of production are not affected by the opportunity of cost saving associated with the existence of an alternative source of R&D. Furthermore, the lack of appropriability of R&D leads to a decrease in the amount of R&D undertaken in the industry. In the context of reduced appropriablility, the decision of investing in R&D, besides increasing the firm's chances of success, also enhances the R&D stock of the rival firm, and consequently, the likelihood that the rival's marginal costs will drop. This implies a loss of profitability for the firm choosing to make the costly investment in R&D. As a result of the disincentive to invest in R&D, the rate of innovation in the industry decreases. This will translate into an industry with less efficient firms and higher marginal costs than under full appropriability, which in turn implies higher prices and a loss of welfare for consumers.

If we rather assume that in order to absorb and exploit external R&D, firms have to develop absorptive capacity, which is built as a by-product of R&D, then spillovers increase the productivity of R&D. Furthermore, this increase is higher for the follower firm in the industry because the spillover pool available to it is higher. Consequently, the gap between the state of the leader and the follower firm will decrease as well as the concentration levels as the industry moves from less to more competitive market structure. The dynamic incentives lead to a higher rate of innovation in the industry, and consequently the marginal costs of production in the industry fall and consumer surplus will increase.

In the remainder of the chapter, we proceed as follows: Section 3.2 presents the analytical framework, distinguishing between the two cases considered, costless R&D spillovers and the absorptiva capacity case. Section 3.3.1 presents the methodology used to find the optimal policy surfaces, and section 3.2.3 discusses the optimal investment choices and the simulation results for the costless spillovers case and, similarly, section 3.3.3 presents and discusses the optimal policy surfaces as well as the simulation results for the absorptive capacity case. Section 3.4 concludes.

3.2 The Model

We model a duopolistic industry where firms produce homogeneous goods and engage in Cournot competition in the spot market. Firms can invest in cost reducing R&D in order to increase their likelihood of success in innovation and hence, potentially lower their marginal cost of production and increase their market share. Thus, R&D enhances the chances of exploring profitable opportunities in the spot market. There are externalities to R&D investment because part of the R&D undertaken leaks to the industry pool of knowledge. The industry is characterised, at time t, by the firms' efficiency levels. These levels capture the state of development of a firm, and evolve over time according to the stochastic outcome of its innovative activity and to an exogenous stochastic process determining the evolution of the factor price index in the industry. In each period, firms decide, through intertemporal optimization, how much to spend in R&D. Firms' decisions are based on the information set available to them in each period, and on their predictions of others' future states, given that they know the probability laws governing the industry motion. The equilibrium is Markov-Perfect in the sense of Maskin and Tirole (1988), and it's a rational expectations equilibrium, i.e., the equilibrium is reached when the expectations of the firms about future states, are the future states that are actually a consequence of those expectations (see E-P, 1995).

3.2.1 The Spot Market

The returns to innovation, and thus the incentives to invest in R&D, are determined in the spot market. We model an industry is composed of two firms producing homogeneous goods. Marginal costs of production are invariant with respect to output, but they vary across firms reflecting the differences in their stock of innovations.

The demand for the industry product is assumed linear and is given by the following inverse demand function:

$$P(Q) = A - B(Q) \quad with\ A, B > 0 \tag{3.1}$$

where P(Q) is the price of the good produced, and Q the industry output.

At the beginning of each period, and in order to decide how much quantity to supply to the spot market, firms compete in quantities and solve the following standard profit maximization problem:

$$\max \pi_i = (P(Q) - mc_i)q_i - f \tag{3.2}$$

where q_i is the quantity produced by firm i and f is the fixed cost of production.

The Cournot-Nash equilibrium will determine the following optimal quantity choices:

$$q_i^* = [A + mc_j - 2mc_i] / 3B \tag{3.3}$$

which will jointly determine the following equilibrium price:

$$P^* = (A + mc_j + mc_i) / 3 \tag{3.4}$$

Given equilibrium price and quantities, equilibrium firms' profits are:

$$\pi_i^* = \max \left\{ -f, \frac{[A + mc_j - 2mc_i]^2}{9B} - f \right\} \tag{3.5}$$

The firm will choose whether to produce or not to do so by comparing its marginal costs to the rival's. If a firm succeeds in innovating less often than its competitor, such that its marginal costs are higher than that of its rival, then the firm will instead choose not to produce and just pay the amount of fixed costs.

3.2.2 Dynamic Processes

The profitability of a firm, at a given moment in time, is determined by its own marginal cost relative to its rival's. Marginal costs are, in turn, fully determined by the efficiency level of firms, which capture their accumulated success in innovation. Therefore, the efficiency levels determine firms' profitability. However, its the relative position of the firm that determines its profits, not the absolute value of its efficiency level.

Let $w_i \in \mathbb{Z}^+$ stand for the efficiency level of firm i. The mapping between the efficiency level and marginal costs of production of the i^{th} firm at time t is given by:

$$mc_i = mc_0 e^{-\eta w_i} \tag{3.6}$$

where $\eta > 0$ captures the rate at which marginal costs decrease with a unit increase in the efficiency level. Higher efficiency levels deliver lower marginal costs of production.

The market structure of the industry given by the efficiency levels of the firms evolves over time as a result of the stochastic outcomes of two processes: 1) an exogenous process determining the evolution of the factor price index in the industry and 2) the investment process undertaken by firms. Given the efficiency level of firm i at time t, the realisations of the two stochastic outcomes will determine efficiency levels at time $t+1$. Letting primes refer to the values of the variables next period

$$w_i' = w_i + \nu_i - \xi \qquad \text{for } i = 1,2. \tag{3.7}$$

The outcome of the first of these processes, namely the one determining the evolution of the factor price index in the industry, is captured by the random variable ξ . The realisation of this variable is common to all firms in the industry, inducing some degree of correlation in firms' fates, as it changes all firms' costs simultaneously. ξ takes value 1 with probability $\delta > 0$, and zero with probability $1 - \delta$. Positive realisations of ξ increase marginal costs in the industry by exactly the same amount of the decrease in marginal costs driven by a successful innovation and we interpret them as increases in the factor price index. The second process is the firms' innovative activity, whose outcome is captured by the binary random variable ν_i. Firms invest in R&D in an attempt to enhance the likelihood of decreasing their marginal costs of production by means of successful innovation. The realisations of this random variable are firm specific, and are the source of heterogeneity in the industry. Since the realisations of ξ are industry wide, the differences in the efficiency levels (and therefore marginal costs) are due to the independent outcomes of firms' innovative activity. The efficiency levels of the firms are bounded from above by means of the positive realisations of ξ. This dynamic process also generates continuous pressure for firms to invest in R&D.

3.2.3 R& D and Appropriability

Let x_{it} be the level of R&D expenditures of firm i at time t. following the specification in Levin and Reiss (1988) and Cohen and Levinthal (1989), we express the total amount of innovative R&D that firm i can utilise to pursue its innovative purposes, m_i, as follows:

$$m_i = x_i + \gamma_i(x_i)bX_{-i} \tag{3.8}$$

where $0 \leq b \leq 1$ is a parameter capturing the extent of the intra-industry spillovers, i.e., the fraction of rivals' R&D made available to the firm. $X_{-i} = \sum_{j \neq i} x_j$ is the amount of R&D undertaken by all firm i's rivals at a given moment in time. In this duopoly set up, X_{-i} will simply be given by the amount of R&D pursued by the rival firm, x_j. We drop time subscripts for simplicity reasons.

The extent of spillovers depends on the ease of imitation, patent policies, worker mobility, the amount of knowledge embodied in the output of innovation process, etc. Larger b implies that a higher proportion of firms' R&D investment is spilling into the industry pool of knowledge. If b is set to zero, then R&D is perfectly appropriable and we obtain the specification in E-P (1995). If b equals unity, then all the R&D in the industry becomes publicly accessible, but not necessarily used because how much of it the firm is able to absorb depends on the absorptive capacity, which might be costly. When $b > 0$, some R&D is becoming public knowledge and bX_i will be the portion of others' R&D that the firm will potentially use.

The function γ_i determines the absorptive capacity of the firm and lies in the range $0 \leq \gamma_i \leq 1$. The functional form for the absorptive capacity of firm i is given by:

$$\gamma_i = f(x_i) = \frac{\gamma x_i}{\phi + \gamma x_i} \quad \text{where} \begin{cases} \phi = 0 & \text{Costless Spillovers Case} \\ \phi = 1 & \text{Absorptive Capacity Case.} \end{cases} \tag{3.9}$$

where the parameter γ governs the productivity of the firm's R&D investment in improving its ability to assimilate and utilize external R&D. This functional form accounts for the two cases of limited appropriability. In the Costless Spillovers case $\gamma_i = 1$, $\forall x_i$ and hence external R&D is a pure public good, costlessly acquired by the firms in the industry. The amount of usable R&D of firm i becomes:

$$m_i = x_i + bx_j.$$

This specification implies that a firm can develop a process innovation, even if it does not pursue any R&D of its own. Higher b implies that the firm can rely more on rival's R&D and substitute it for own R&D and it also entails a detrimental effect on a firm's incentives to engage in R&D expenditures that arises from the possibility that the rival free rides on the firm's own R&D. This is the case addressed by the initial literature on the appropriability of R&D such as Spence (1984), D'Aspremont and Jacquemin (1988).

In the absorptive capacity case, firms require research personnel, laboratories or some technological knowledge in order to learn from the external pool of R&D. As in Cohen and Levinthal (1989), the fact that the firm engages in R&D for innovating has the side effect of enhancing its absorptive capacity. Equation (3.9) implies that the absorptive capacity of firm i is a monotonically increasing concave function of R&D investment ($f_x > 0$ and $f_{xx} < 0$). Furthermore, if the firm does not engage in R&D of its own, it will not be able to succeed in innovation ($\gamma(0) = 0$). The parameter γ in the equation above captures the productivity of investment in increasing firm i's absorptive capacity and is related with the ease of learning associated with the specificities of the technological knowledge, e.g. highly sophisticated and complex in contrast to easily recognisable knowledge.

Hence, while parameter b (extent parameter) captures the R&D made available in the industry, γ_i (absorptive capacity parameter) determines the fraction of it that the firm will effectively assimilate and use. For positive levels of spending in R&D, the firm will be able to absorb and exploit a portion γ_i of the spillover pool (bX_i), adding $\gamma_i bX_i$ to its usable R&D. As implied by (3.8), the total amount of R&D a given firm i will have at a given moment of time will be the sum of its own R&D and the fraction of the rival's R&D that it can absorb from intra industry spillovers.

The amount of usable R&D devoted to innovation increases the probability of success and hence the chances of improving its efficiency level, w_i. The outcome of firm i's innovative activity, captured by the random variable ν_i, is given by:

$$\nu_i \begin{cases} = 1 \text{ with probability } p(\nu_i) \\ = 0 \text{ with probability } 1 - p(\nu_i) \end{cases} \tag{3.10}$$

where

$$p(\nu_i) = \frac{am_i}{1 + am_i}. \tag{3.11}$$

thus, when the binary variable ν_i takes value 1, the firm is successful in innovating. In equation (3.11), a indicates the productivity of usable knowledge in increasing the likelihood of developing an innovation. This is basically the technological opportunity in the industry, i.e., the difficulty of innovating in the industry, and it is related to the stage of development of scientific knowledge and other knowledge specific characteristics.

The probability of an increase in the efficiency level of firm i can be expressed as:

$$Probability(w_i' = w_i + 1) = \frac{(1-\delta)am_i}{1+am_i} = (1-\delta)\frac{a\,(x_i + \gamma_i(x_i)bx_j)}{1+a\,(x_i+\gamma_i(x_i)bx_j)}$$

Figure 3.1 illustrates a parameterized example of how the probability of innovation changes with own and rival's R&D investment. The graphs show the probability of innovation associated with a combination (x_i, X_i) of own R&D (in the x-axis) and rival's R&D (in the y-axis). For the purposes of this illustration, productivity parameters a and γ are set to 0.5 and 1, respectively. Fig 3.1 a) depicts a situation of fully appropriable R&D, which is the same as saying that b is set to zero, such that the only input to the innovative activity is the knowledge the firm develops by pursuing R&D spending.

This figure makes clear the decreasing marginal returns of R&D to the probability of innovation showing a monotonically increasing concave relationship between R&D and the probability of innovation. Under full appropriability, rival's R&D is irrelevant for the probability of innovation, since it does not translate into external R&D.

Fig 3.1b) and 3.1c), on the other hand, depict a situation where 20% of the R&D undertaken in the industry spills to the common pool of knowledge (b = 0.2). Figure 3.1b) captures the costless spillovers case, where firms can achieve success in innovation even if they do not pursue positive amounts of R&D spending because they allocate to the innovative activity the external R&D that they absorb. Figure 3.1c) illustrates the absorptive capacity case and shows how a firm can experience an increase in its probability of developing an innovation due to the increase in the amount of R&D undertaken by the rival firm. Under this scenario, and as can be seen in the figure, an increase in the public pool of knowledge will only translate into higher probability of innovation if the firm conducts some R&D of its own. The higher the firm's own investment effort, the higher the impact of a unit of external R&D in its probability of innovation. There are also decreasing marginal returns to rivals R&D, as shown by the concavity of the probability of innovation with respect to the common pool of knowledge. In the case of absorptive capacity, the impact of a small increase in rival's R&D is large[2], but smaller than the impact for the probability of success of an increase of the same magnitude in the expenditures in own R&D.

3.2.4 Dynamic Equilibrium

In each period of time, firms choose the level of R&D expenditures that maximize the expected present value of their future stream of profits. This optimal choice requires the perception of rivals' future states. Let $s = [w_i]$ be a vector which describes the market structure by listing the efficiency level of each firm at a given moment in time, and $pr(w_i^{'}, s^{'} \mid x, w_i, s)$ denote a firm's perceptions of the joint probability that its efficiency will evolve to $w_i^{'}$ in the next period, and that the market structure it faces will be $s^{'}$, conditional on the vector of R&D investment for the two firms, x, its current state and current market structure. Then the optimal investment choice is the one solving the following Bellman equation:

$$V_i(w_i, s) = \max_{x_i \geqslant 0} \{\pi(w_i, s) - cx_i + \beta E_t V_i(w_i^{'}, s^{'})\} \tag{3.12}$$

where $\beta = 1/(1+r)$, with r standing for the interest rate, is the discount factor, common to all firms, and c is the unit cost of R&D investment. Given

[2]Except for the absorptive capacity case when the level of own R&D investment is zero, in which case the external pool of R&D cannot be used by the firm, and consequently leaves the probability of innovation at zero.

the optimal investment choice x_i^*, the firm receives the maximum value of the sum of current period profits minus its investment expenditures in R&D, plus the discounted value of the future payoffs conditional on future market structures. In order to solve for value maximization, firms sum over the values of all possible future states weighted by their probability of occurrence. Given the transition matrix of the probabilities that the market structure moves from its current state to all other possible states, we can write (3.12) as:

$$V_i(w_i, s) = \max_{x_i \geqslant 0}\{\pi(w_i, s) - cx_i + \beta V_i(w_i', s')pr(w_i', s' \mid x, w_i, s)\} \tag{3.13}$$

To understand the impact of the extent of spillovers, we need to understand the basic shape of the value function. Figure 3.5 shows the solution to the value function for firm 1 using the parameters listed in Table 3.1 in appendix with $b = 0$. The x-axis represents firm 1's efficiency level, the y-axis is the rival firm's (firm 2) efficiency level, and the z-axis is firm 1's value. Higher efficiency levels mean lower marginal costs. Thus, firm 1's value increases in its own efficiency level due to higher current and expected profits, while its value decreases in its rival's efficiency level. Holding the rival firm's efficiency level fixed, the value function takes a distinctly convex shape at low levels of efficiency, and after some point it then becomes concave. At the extreme ends, the value function is relatively flat. In these regions, firms enter coasting states where they cease R&D as they hit the non-negativity constraint. Firms too small to compete cease R&D because the marginal gains are too small relative to the costs given that its rival has such a large advantage. However, as the gap in efficiency levels shrinks, the value increases at an increasing rate creating larger marginal increments in value. Eventually, with higher relative efficiency levels, the firm enters the concave portion of the value function. This region represents a firm with most or all of the market that achieves little benefit from reducing costs because it captures little, if any, additional market share from its rival and it faces diminishing returns in the product market.

Let $C_1(w_i'+1, s_m')$ denote the expected value of the firm conditional on success in innovating, and $C_2(w_i', s_m')$ the expect value in the case it fails to develop an innovation. Then one can rewrite (3.13) as:

$$V_i(w_i, s) = \max_{x_i \geqslant 0} \left\{ \begin{array}{c} \pi(w_i, s) - cx_i + \beta \left[pr(w_i', s' \mid x, w_i, s) C_1(w_i' + 1, s_m') + \right. \\ \left. \left[1 - pr(w_i', s' \mid x, w_i, s) \right] C_2(w_i', s') \right] \end{array} \right\} \tag{3.14}$$

Furthermore, both the expected values of the firm conditional on success (C_1), and its expected value conditional on failure, (C_2), can be expressed in terms of the two possible outcomes for the rival firm (firm j):

$$\begin{aligned} C_1(w_i' + 1, s_m') &= pr(w_j', s' \mid x, w_j, s) C_{11}(w_i' + 1, s_m' = (w_i' + 1, w_j' + 1)) \\ &+ \left[1 - pr(w_j', s' \mid x, w_j, s) \right] C_{12}(w_i' + 1, s_m' = (w_i' + 1, w_j')) \end{aligned} \tag{3.15}$$

and

$$\begin{aligned} C_2(w_i', s_m') &= pr(w_j', s' \mid x, w_j, s) C_{21}(w_i', s_m' = (w_i', w_j' + 1)) \\ &+ \left[1 - pr(w_j', s' \mid x, w_j, s) \right] C_{22}(w_i', s_m' = (w_i', w_j')), \end{aligned} \tag{3.16}$$

where C_{11} is the firm's value conditional on both its own and its rival's success; C_{12} is the firm's value when it is successful and its rival fails to develop an innovation; C_{21} is the firm's value conditional on its own failure and its rival's success in innovating; and finally, C_{22} is the firm's value associated with the failure of both itself and its rival. Hence, the value of the firm conditional on its success, (C_1), and its lack of success in innovating (C_2) will be the weighted average of the two possible outcomes for the rival firm (success/failure), where the weight of each of these outcomes is given by the rival firm's probability of success/failure. Given that a firm's value is monotonically increasing in it's efficiency level, and decreasing in that of its rival, the following relationship holds: $C_{12} > C_{11}$ and $C_{22} > C_{21}$. A higher value is attached to an outcome involving the firm's success and the rival's failure, with the lowest value being associated with its failure and its rival's success. The profit function has those relations and the value function heavily reflects the profits, so the relationship between the values associated with different outcomes follow from that.

The problem can therefore be written as:

$$V_i(w_i, s) = \max_{x_i \geq 0} \left\{ \begin{array}{c} \pi(w_i, s) - cx_i + \\ \beta \frac{a(x_i+\gamma_i(x_i)bx_j)}{1+a(x_i+\gamma_i(x_i)bx_j)} \left[\frac{a\left(x_j+\gamma_j(x_j)bx_i\right)}{1+a\left(x_j+\gamma_j(x_j)bx_i\right)} C_{11} + \frac{1}{1+a\left(x_j+\gamma_j(x_j)bx_i\right)} C_{12} \right] + \\ \beta \frac{1}{1+a(x_i+\gamma_i(x_i)bx_j)} \left[\frac{a\left(x_j+\gamma_j(x_j)bx_i\right)}{1+a\left(x_j+\gamma_j(x_j)bx_i\right)} C_{21} + \frac{1}{1+a\left(x_j+\gamma_j(x_j)bx_i\right)} C_{22} \right] \end{array} \right\} \tag{3.17}$$

The difference between this dynamic problem and the one which departs from the assumption of full appropriability resides on the fact that, due to the externalities of R&D investment, the outcome of the rival's R&D process is not independent of the firm's R&D investment. Fully rational firms take this strategic effect of their R&D spending into account when deciding the optimal level of their R&D investment.

Taking the derivative of (3.17) with respect to x_i:

$$\frac{\partial V_i}{\partial x_i} = \frac{-c}{\beta} + \frac{a\frac{\partial m_i}{\partial x_i}}{(1+am_i)^2}(C_1 - C_2) + \frac{a\frac{\partial m_j}{\partial x_i}}{(1+am_j)^2}\left[\frac{am_i}{1+am_i}C_{11} + \frac{1}{1+am_i}C_{21}\right] - \frac{a\frac{\partial m_j}{\partial x_i}}{(1+am_j)^2}\left[\frac{am_i}{1+am_i}C_{12} + \frac{1}{1+am_i}C_{22}\right] \quad (3.18)$$

where $(C_1 - C_2)$ is simply the marginal value gain from innovating. Let $(C_1^* - C_2^*)$ be the differential value between the outcome in which the rival innovates, and that in which the rival experiences failure in innovation, i.e.:

$$C_1^*(w_i', s_m') = \frac{am_i}{1+am_i}C_{11} + \frac{1}{1+am_i}C_{21} \quad (3.19)$$

$$C_2^*(w_i', s_m') = \frac{am_i}{1+am_i}C_{12} + \frac{1}{1+am_i}C_{22}. \quad (3.20)$$

Then we can rewrite the first order condition in a simplified form:

$$\frac{\partial V_i}{\partial x_i} = \frac{-c}{\beta} + \frac{a\frac{\partial m_i}{\partial x_i}}{(1+am_i)^2}(C_1 - C_2) + \frac{a\frac{\partial m_j}{\partial x_i}}{(1+am_j)^2}(C_1^* - C_2^*) = 0 \quad (3.21)$$

The firm's value function is monotonically increasing in its own efficiency level, and monotonically decreasing in the efficiency level of the rival. Thus, the firm experiences a value increase when it sees its position relative to the rival improve, i.e., $(C_1 - C_2)$ is positive, and the firm is better off if the rival fails to innovate then when it is successful, i.e., $(C_1^* - C_2^*)$ is negative. Again, these reflect the higher current and expected profits associated with successfull innovation outcomes for the firm and with rival's faillure, as both translate into an improvement in the firm's relative state.

While in the second chapter, it was not possible to ensure the existence of equilibrium in pure strategies due to entry and exit decisions, these problems do not arise here, as we abstract from entry and exit[3]. Concerning potential multiple equilibrium, since entry and exit decision, which are discrete, are an important source of these multiple equilibrium, having a fixed number of firms reduces the likelihood that multiple equilibrium arise. Furthermore, this reduction is stronger in the context of an industry with a very small number of agents, which is a feature of the model presented here (duopoly).

For each level of R&D spending, the probability of innovation when spillovers are costless differs from that of the absorptive capacity case, which entails different formulations for the Bellman equation. This justifies the separate treatment we assign to the dynamic equilibrium in each of these two scenarios.

Costless R&D Spillovers

Myopic R&D Choice In order to disentangle the economic forces at work in the model, as well as the sources of the impact of spillovers on the incentives to conduct R&D, we start by presenting the R&D choices which firms would pursue if they behaved myopically, i.e., if they did not internalize the strategic effect of their investment decisions on their rival's transition matrix. Myopic firms are assumed to ignore the (positive) impact that their decision of engaging in R&D has on promoting the rival's chances in being successful in innovation, which implies, given the structure of the model, a deterioration in the firm's profits and, consequently, its value. Analytically, this implies ignoring the third term in

[3]More details on these issues can be found in the appendix to Chapter 2, where the Pakes et al. algorithm is briefly presented and discussed.

the first order condition set out in equation (3.21):

$$0 = \frac{-c}{\beta} + \frac{a}{(1 + a(x_i + bx_j))^2}(C_1 - C_2). \tag{3.22}$$

Solving the equation with respect to x^*, we obtain the optimal choice of R&D investment when there are *costless* R&D spillovers and firms behave *myopically*:

$$x_i^* = \frac{1}{a}\left[\sqrt{\frac{a(C1 - C2)\beta}{c}} - 1\right] - bx_j. \tag{3.23}$$

Substituting the analytical formula for the optimal choice for the rival's investment, we obtain the best response function for investment by firm i:

$$\begin{aligned} x_i^* &= \frac{1}{a}\left[\sqrt{\frac{a(C1 - C2)\beta}{c}} - b\sqrt{\frac{a(C1^{\S} - C2^{\S})\beta}{c}} + b - 1\right] \quad (3.24) \\ &= \frac{1}{a}\left[\sqrt{\frac{a\beta}{c}}\left(\sqrt{C1 - C2} - b\sqrt{C1^{\S} - C2^{\S}}\right) + b - 1\right] \quad (3.25) \end{aligned}$$

where $C1^{\S} - C2^{\S}$ denotes the marginal value gain from innovating for the rival. The relationship between $C1 - C2$ and $C1^{\S} - C2^{\S}$ depends on the relative slope of the value function at the states where both the firm and its rival are.

The solution in (3.23) is similar to the optimal R&D investment under full appropriability (the original Ericson-Pakes (1995) set up), the only difference residing on the last term ($-bx_j$), which reduces the R&D choice. Since firms benefit from rivals' R&D, and provided $x_j > 0$, they can lower their level of R&D expenditures without reducing the rate of innovation. This negative effect of spillovers in R&D investment has been widely discussed in the literature and we refer to it as the *substitution effect.* Holding the level of rival's R&D fixed, the larger the extent of spillovers, the higher the reduction in own R&D due to the substitution effect. However, x_j is an endogenous variable, and the last term captures both the direct effect of spillovers in decreasing a firm's own R&D, as well as the indirect effect through the impact of b on the rival's R&D, x_j. The first term of the R&D choice in (3.23) might also differ from the full appropriability choice to the extent that the marginal value gain from innovation $(C_1 - C_2)$ is endogenous with respect to b. A change in the extent parameter may change the slope of the value function by means of the positive cost effect and its impact on the distribution of the incentives to invest, which in turn depend on the amount of R&D undertaken by the rival and the productivity of own R&D in decreasing marginal costs.

If we substitute the expression for the optimal amount of own R&D spending into the expression for the optimal level of total usable R&D, we obtain:

$$\begin{aligned} m_i^* &= x_i^* + bx_j \\ m_i^* &= \frac{1}{a}\left[\frac{a(C1 - C2)\beta}{c}\right]^{1/2} - \frac{1}{a} \end{aligned} \tag{3.26}$$

The expression for the total level of usable R&D (m_i^*) by a myopic firm in a setting with costless R&D spillovers is analogous to the one presented in E-P (1995) for optimal R&D investment (x_i^*) in the case of fully appropriable R&D. Consequently, the expression for the optimal probability of innovation will also be analogous to that of the full apropriability case. The firm will be choosing the combinations of in-house and outside R&D to input into the knowledge production function (3.11) that allow incurring in the lowest possible investment cost (cx_i) while delivering the optimal amount of total usable R&D, m^*. In this R&D program, own and external R&D are substitute investments since increases in bx_j^*, ceteris paribus, decrease x_i^* in order to maintain m^* at its optimal level. Therefore, the share of external R&D in m^* is bigger the more favourable the conditions for spillovers in R&D, and the higher the unit cost of own R&D.

This does not imply, however, that m_i^* in this specification is equal to the optimal amount of own R&D investment in E-P (1995), since the presence of spillovers will have an impact on the incremental value from innovation $(C_1 - C_2)$. Spillovers affect the incremental value from innovation to the extent that they affect the slope of the value function, and consequently the incentives to invest. This impact is made both through the cost saving emerging from the existence of a pool of R&D, as well as through the changes in the transition matrix. We can not conclude about the direction of this effect, but the simulations we performed showed it to be small relative to the other effects[4]. Thus, we do not expect costless spillovers to significantly affect the rate of innovation when firms are myopic with respect to the strategic effect of their R&D investment choices. The relative efficiency levels, which are determined by the rates of innovation of the two firms in the industry, shape the market structure in this model. Consequently, the level of concentration should remain fairly unchanged with variations in the degree of appropriability of R&D.

The price paid by consumers reflects the marginal costs of production, which depend on firms' efficiency levels. What matters in terms of consumers' welfare is not the level of own R&D undertaken by the firm (x_i), but rather the rate of innovation which depends on the total amount of R&D (m_i) the firm is able to devote to the research process, independently of whether it's the firm's own R&D

[4]These results are presented and discussed in section 3.3.

investment or the rival's R&D. Firms utilise outside knowledge as a substitute for their own efforts and reduce costs which leads to lower R&D expenditure, but the use of outside knowledge allows the rates of innovation to remain the same. Consequently, when firms are myopic, costless R&D spillovers are not expected to have a significant impact on market structure and consumer welfare.

Fully Strategic R&D Choice We now abandon the hypothesis of bounded rationality, and assume that firms fully account, in their R&D decisions, for the impact of their investment spending on the rival's chances of success. The first order condition for finding the optimal policy function for R&D investment can then be written as follows:

$$0 = \frac{-c}{\beta} + \frac{a}{(1 + a(x_i + bx_j))^2} (C_1 - C_2) + \frac{ab}{(1 + a(x_j + bx_i))^2} (C_1^* - C_2^*) , \quad (3.27)$$

which differs from the first order condition for the myopic case (eq (3.22)) in the last term, which captures the *strategic effect* of spillovers on a firm's R&D choice. All the parameters in this term are positive, there is a non-negativity constraint on the levels of x_i and x_j, and $(C_1^* - C_2^*) < 0$, henceforth the term is negative. Consequently, a lower level of R&D is required for optimality. This is so because a firm's efficiency level will be stochastically increasing in its own R&D investment as well as in the external R&D it absorbs, but its profitability, and hence its value, will be stochastically decreasing in the amount of R&D the rival inputs to its innovative activity, which is higher the larger the pool of external R&D made available. When choosing its R&D expenditure the firm will anticipate the positive impact on the rival's probability of success. The strategic effect driven by the imperfect appropriability of R&D investment generates a free rider problem and decreases a firm's incentive to invest even further with respect to the myopic R&D choice. Note that in the myopic case of costless spillovers, despite the reduction in the level of own R&D expenditures due to the substitution effect, the total amount of R&D devoted to innovation remains fairly similar to the full appropriability level. Consequently, we expect the level of R&D investment, and thus the rate of innovation in the industry, to be lower than in the myopic case.

Solving implicitly for the optimal level of total usable R&D, we obtain:

$$m_i^* = \frac{1}{a}\left[\frac{a(C1-C2)\beta}{c} + \frac{ab\beta}{c}\left(\frac{1+am_i}{1+am_j}\right)^2 (C_1^* - C_2^*)\right]^{1/2} - \frac{1}{a} \tag{3.28}$$

This solution differs from the myopic case in the last term inside brackets, which is negative. Holding the marginal increment in value $(C_1 - C_2)$ constant, the total amount of usable R&D that fully rational firms devote to the process of developing new innovations is smaller in the presence of costless R&D spillovers than in the set up where firms are able to fully appropriate their R&D investment. Since the total amount of R&D inputted to the innovative activity is expected to be smaller, the rate of innovation associated with an industry with fully strategic (rather then myopic) firms will be smaller. With less frequent innovations, and a less favourable innovation path, the marginal costs in the industry will be higher than under full appropriability, and consequently, spillovers will lead to a decrease in consumer welfare.

The first order condition in (3.27) has no analytical solution. The optimal amount of firms' R&D investment is determined numerically within the value function iteration algorithm. In section 3.3.2, we analyse the results of simulating this set up in order to evaluate whether the magnitude of the change in $(C_1 - C_2)$ is how we expect and to conclude about how costless R&D spillovers affect the optimal probability of innovation, and thus market structure and welfare measures.

The Case of Absorptive Capacity

Myopic R&D Choices When the absorption and exploitation of spillovers requires the firm pursuing some positive level of R&D investment, the amount of usable R&D differs from the one in the costless spillovers case. The absorptive capacity case introduces a new role to a firm's own R&D expenditures, the learning incentive of R&D investment. The free riding problem of knowledge externalities might eventually be compensated by the role of R&D in increasing the ability of the firm to assimilate and exploit the knowledge which is made available to it at the expense of rivals' costs. As before, we start by addressing the myopic case. The first order condition for the dynamic problem of a myopic firm, in the case of absorptive capacity, is as follows:

$$\frac{\partial V_i}{\partial x_i} = \frac{-c}{\beta} + \frac{a\left(1 + \frac{\gamma}{(1+\gamma x_i)^2} b x_j\right)}{\left(1 + a(x_i + b\frac{\gamma x_i}{1+\gamma x_i} x_j)\right)^2} (C_1 - C_2) = 0 \qquad (3.29)$$

Solving implicitly for the optimal level of total usable R&D of myopic firms we obtain:

$$m_i^* = \frac{1}{a}\left[\frac{a(C1 - C2)\beta}{c}\left(1 + \frac{\gamma}{[1+\gamma x_i]^2} b x_j\right)\right]^{1/2} - \frac{1}{a} \qquad (3.30)$$

Holding the marginal value gain constant, this solution differs from the full appropriability case (and thus the costless myopic) in the second term inside the curly brackets, which is positive. This term captures the increase in the productivity of R&D in the absorptive capacity case. An extra unit of own R&D not only adds a further unit to total usable R&D through its linear contribution to m^*, it also increases the level of rival's R&D that the firm is able to add to its R&D stock. Thus, we expect to find that R&D spillovers in the absorptive capacity case increase the rate of innovation in the industry with respect to both the full appropriability and the costless spillovers cases. Myopic firms take the learning effect of R&D into account when making their decisions and increase their R&D spending as a response to the increased productivity associated with the absorptive capacity.

Fully Strategic R&D Choices Fully strategic firms in the absorptive capacity case will fully account for the impact of their R&D decisions on the chances of the rival being successful in innovating. In the absorptive capacity case, we allow firms to take into account, in their choices of x_i, the learning incentive of R&D investment, as well as the fact that they also contribute to the rivals' innovative activity by expanding the industry pool of R&D becoming available in the industry. The free riding problem of knowledge externalities might eventually be compensated by the role of R&D in increasing the ability of the firm to assimilate and exploit the knowledge which is made available to it at the expense of rivals' costs. The absorptive capacity requirement will change the impact of a firm's R&D choice on its rival's rate of innovation.

The R&D program of fully strategic firms in the absorptive capacity case will be the one solving the following first order condition:

$$
\begin{aligned}
0 &= \frac{-c}{\beta} + \frac{a\left(1+\frac{\gamma}{(1+\gamma x_i)^2} b x_j\right)}{\left(1+a(x_i + b\frac{\gamma x_i}{1+\gamma x_i} x_j)\right)^2}(C_1 - C_2) \qquad (3.31)\\
&+ \frac{am_i}{1+am_i}\frac{ab\frac{\gamma x_j}{1+\gamma x_j}}{\left(1+a(x_j + b\frac{\gamma x_j}{1+\gamma x_j} x_i)\right)^2}[C_{11} - C_{12}]\\
&+ \frac{am_i}{1+am_i}\frac{ab\frac{\gamma x_j}{1+\gamma x_j}}{\left(1+a(x_j + b\frac{\gamma x_j}{1+\gamma x_j} x_i)\right)^2}[C_{21} - C_{22}]
\end{aligned}
$$

which yields the following implicit solution for total usable R&D:

$$
\begin{aligned}
m_i^* &= \frac{1}{a}\left[\frac{a(C1 - C2)\beta}{c}\left(1+\frac{\gamma}{[1+\gamma x_i]^2} b x_j\right)\right. \qquad (3.32)\\
&\left.+\frac{ab\beta}{c}\left(\frac{1+am_i}{1+am_j}\right)^2\left(\frac{\gamma x_j}{1+\gamma x_j}\right)(C_1^* - C_2^*)\right]^{1/2} - \frac{1}{a} \qquad (3.33)
\end{aligned}
$$

When compared with the solution in (3.28) for the fully strategic costless case, this solution differs in the two terms associated with the absorptive capacity, $\gamma(x_i)$. The first term, as we already discussed in the myopic case, is related with the increase in the productivity of R&D. In the second, $\frac{\gamma x_j}{1+\gamma x_j}$,the expression for the absorptive capacity of the rival (firm j), appears multiplicatively in the strategic effect. The requirement of building absorptive capacity to explore external knowledge increases the productivity of own R&D, on the one hand, and restricts the portion of the firm's R&D that the rival is able to absorb, on the other. However, since the strategic term is negative, the expression in (3.32) allows concluding that, holding marginal value gains constant, the amount of R&D devoted to innovation in the absorptive capacity case is lower when firms behave rationally than when they choose R&D myopically.

Since there is no explicit analytical solution for the optimal level of R&D investment in the fully strategic case of absorptive capacity, we analyse the effects of R&D spillovers in this case in section 3.3.3.

3.3 Numerical Results

3.3.1 Methodology

In what follows we address the results concerning the impact of R&D spillovers on concentration and welfare by simulating the evolution of industry market structure. We interpret the results obtained in order to conclude about the relative magnitude of the effects involved, given that the limitations of the analytical analysis only allowed us to infer about the expected direction of these effects.

The numerical algorithm solves for the Markov-perfect Nash equilibrium policy functions for investment and values associated, providing the equilibrium strategies $\{x_i(w_i, s), V_i(w_i, s)\}$ for all $w_i \in W$, and all $s \in S$. The simulation program then uses these equilibrium policies to stochastically generate the evolution of the market structure of the industry, which is an ergodic process. As noted in the appendix to Chapter 2, where a description of the core of the Pakes et al. algorithm is presented, multiple equilibrium cannot be ruled out in the original framework proposed by Ericson Pakes. However, multiple equilibria in the model are associated with entry and exit decisions which is not a feature of the model presented here. We simulate the costless and absorptive capacity cases of the model 100.000 times each for various values of parameter b, the extent of spillovers, to obtain the numerical results for the ergodic distributions of market structures and the expected discounted value of the welfare measures for different levels of R&D appropriability.

We allow for the extent of spillovers to go from fully appropriable R&D to complete spillovers, with jumps of 10% in the extent of knowledge becoming available in the industry. In a few cases, when the spillovers parameter reached $b = 0.9$ and 1.0, firms began exceeding the maximum efficiency level. Thus, we only present the results for the range between $b = 0$ and a situation where 80% of the R&D firms undertaken leaks to the public stock of knowledge, as these are enough to demonstrate the predictions of section 3.2. The parameter values used for the simulation of the model are as in table 3.1 at the end of the chapter.

3.3.2 Costless R&D Spillovers

Optimal Policy Functions

In order to evaluate the impact of imperfect appropriability of R&D on market structure, we analyse how the optimal policy functions for all possible market structures change when spillovers increase. Figure 3.2 depicts the optimal policy surface for R&D investment for fully rational firms at different levels of externalities, namely full appropriability ($b = 0$), 20%, 50% and 80% of R&D spillovers. In the figures, each point in the grid formed by the x and the y axis corresponds to one possible combination of firms' efficiency levels, i.e., one possible market structure. The vertical axis gives a firm's (say, firm 1) optimal level of investment as a function of its own efficiency level (w_1) and its competitor's efficiency level (w_2).

As depicted in figure 3.2, firm 1's investment initially increases with the firm's own level of efficiency, reaches a maximum at the inflection point (around $w_1 = 5$), and thereafter it decreases with w_1.

These properties of the optimal policy function for investment derive from the fact that the value function as defined in E-P (1995) is bounded above and below, such that it is initially a convex and then a concave function of the firm's own efficiency level. Consequently, the incremental value from innovating when the firm has a high efficiency level is very small, and the firm has little incentive to invest in R&D.

As costless spillovers increase, there is a noticeable slump in the amount of investment of firm 1 around the market structures where the rival has efficiency level 5. At these market structures, the competitor is investing heavily in R&D and the external pool of R&D potentially available to firm 1 is at its maximum. The higher b, when the rival is investing heavily, the higher the reduction in firm 1's R&D. Therefore, as b increases, the firm will substitute external for own R&D. This substitution occurs more intensively when the rival is investing heavily in R&D, which happens when $w_2 = 5$ because of the convexity of the rival's value at that state. We refer to the top left panel of figure 3.2 (page 148), where it is seen that the maximum investment by the rival occurs around an average of about 5, i.e., in the middle of the range shown. Figure 3.2 shows how the composition of the R&D program is changing with the increase in the extent of spillovers, but in order to evaluate the impact on market structure it is necessary to analyse what happens to the value function and the probability of success when spillovers increase. Figures 3.3 and 3.4 depict the change in the probability of success and firm's value when there are externalities to R&D. The vertical axis in figure 3.3 gives the probability of success when there are spillovers to R&D minus the probability of success when R&D is fully appropriable for each market structure. According to Figure 3.3, for some market structures, changes in b lead to very small changes in the optimal probability of innovation. However, the optimal probability of innovation for firm 1 experiences a significant increase for most market structures where the rival firm has an efficiency level, w_2, around 5. From figure 3.2, it is made clear that at efficiency level 5 firms are strongly engaging in R&D investment. This increment in the probability of innovation of firm 1 is especially significant for market structures where firm 1 has a very high efficiency level, and those where it has a particularly low efficiency level. At those market structures, firm 1 has little incentive to perform R&D (as made clear in figure 3.2), and so all the R&D it devotes to innovation is outside R&D steaming from the intense R&D activity of firm 2. Furthermore, the increases in the probability of innovation of firm 1 are higher when it holds a dominant position in the market (to the left of the diagonal) than when the firm is the follower firm in the industry (market structures to the right of the diagonal). The fact that the probability of reducing marginal costs is higher when the firm has a strong lead suggests that costless R&D spillovers increase the concentration level. In market structures where firm 1 has a high efficiency level and firm 2 has a very small efficiency level, the probability of innovation of firm 1 decreases as spillovers increase. At these market structures, the incentives for firm 1 to invest are minimal because this firm is in the upper end of the value function, where further innovations will generate very small increments to the firm's value, and moreover the rival can increase its value significantly by enhancing the amount of R&D devoted to innovative activity and thus highly benefit from spillovers. Figure 3.4 depicts the change in value that emerges from an increase in the extent of costless spillovers.

Figure 3.4 illustrates that for many of the market structures, spillovers lead to an increase in value mainly due to the fall in the costs of financing R&D that follows from the emergence of an alternative type of investment. This increase in value is particularly significant at market structures where the rival firm is strongly engaging in R&D. However, the market structures for which figure 3.3 show a higher increase in probability do not correspond to those where the value increase is higher. This apparent contradiction is related to the shape of the value function. Figure 3.5 depicts the value function, for $b = 0$. At market structures for which firm 1 has a very high efficiency level, the incremental value for innovation is very small due to the concavity of the value function at those efficiency levels. The higher increases in value with the extent in spillovers shown in figure 3.4 correspond, in turn, to market structures for which, in spite of having a smaller increase in the optimal probability of innovation, the marginal value gain from innovation is higher due to the convexity of the value function at those market structures, as depicted in figure 3.5. However, figure 3.4 shows an exception to this situation at market structures where firm 1 has efficiency level of approximately 5, and for which the slope of the value function is at its highest but there is no matching increase in value from spillovers. This slump in firm 1's value differential is due to the fact that at those efficiency levels, firm 1 is strongly engaging in R&D and firm 2 is free-riding on this R&D spending. This effect lessens the value gains from spillovers for firm 1 at those market structures.

Simulation Results

In our search for the impact of costless R&D spillovers on the evolution of market structure, we now turn to the probability of occurrence of each of the market structures associated with the ergodic distribution of market structures we obtained from simulating the model. We also register mean investment, mean value and mean probability for each value of b, i.e., for each degree of R&D spillovers. While previously we addressed the optimal policy functions for every possible market structure, in what follows we look at the results of simulating the evolution of firms' states, and thus market structure, departing from an initial condition and given the optimal policy functions discussed above. Figure 3.6 depicts the ergodic distribution of the market structures obtained from the simulations, i.e., it informs about the frequency with which a certain market structure was visited out of the hundred thousand periods of industry dynamics.

The grid represents all possible combinations of the efficiency levels of the two firms in the industry. The vertical axis registers the frequency of occurrence of each of these market structures. The market structures are represented as weakly descending tuples of the efficiency levels of the firms such that there are no events to the right of the diagonal of the grid because there cannot be a market structure where the leader has a lower efficiency level then the follower firm. As spillovers increase, the symmetric mode where both firms are relatively efficient disappears. The mass of the distribution moves towards the asymmetric mode, at the same time as the likelihood of occurrence of higher efficiency levels decreases for both firms. Thus, as spillovers increase firms compete less fiercely but, on the other hand, the likelihood of more asymmetric market structures increases. The results for the myopic case are not presented here, but it suffices to say that the market structure stays fairly unchanged as costless spillovers increase. This implies that the changes we described for the market structure of the fully strategic case are all attributable to the strategic effect. Figure 3.7 shows the percentage change in the concentration index C_1 when we allow for externalities in R&D with respect to the complete appropriability case. The solid line corresponds to the results of simulating the set up assuming rational firms and the dashed line illustrates the results of simulating the model departing from the assumption of myopic firms. We perform this exercise since it helps elucidating which impacts arise from the strategic effect of spillovers. The levels of concentration do not seem to change much with the increase in costless R&D spillovers. When firms are myopic, given that most variations in C_1 are below the tolerance level used for finding the Markov-Perfect Nash equilibrium of the model, it seems fair to say that the level of concentration in the industry stays fairly unchanged in the presence of spillovers. When firms behave fully rationally, the level of concentration initially decreases and then from $b = 0.5$ onwards it increases with the extent of spillovers, although the variations are of small magnitude throughout. The difference in the impact of b on the concentration index is due to the strategic effect. For b bellow 0.5, given that the follower will free ride on its R&D expenditure, the leader will experience a disincentive to invest. The leader's probability of innovation will decrease together with an increase in the follower's likelihood of success, causing a decrease in concentration. However, for high enough spillovers, the leader starts increasing its investment. In spite of the fact that the follower is free riding, the decrease in its R&D investment widens the gap between the follower and the leader, such that concentration starts rising.

Figure 3.8 shows mean R&D investment, mean probability of innovation and mean firm's value obtained from the simulations for the different levels of the extent of spillovers. The graphs on the left, with the solid lines, correspond to the results for rational firms and the graphs on the right, with the dashed lines, illustrate the results of simulating the behaviour of myopic agents. Additionally, in all graphs, the lines with the triangular markers represent the leader (the firm with the highest efficiency level and thus with the highest market share) and the ones with no markers represent the follower. This correspondence will be used in all the graphs throughout this section. The figure shows the general tendency for a decrease in R&D investment as the extent of costless R&D spillovers increases, which takes place for both myopic and fully strategic firms. When firms ignore the strategic effect of their R&D choice, according to the analytical results obtained in section 3.2.4, they choose their investment as to maintain their probability of innovation at the optimal level, which is not expected to differ substantially from the full appropriability case. As a result, the graphs for the myopic case illustrate a set up where there is a fall in investment while the rate of innovation stays fairly stable. Nevertheless, there is an increase in firm's value reflecting the decrease in the investment cost involved in the R&D activity. In fact, firms spend less than in the fully appropriable R&D set up by substituting own for external R&D (which comes at no cost), and achieve a similar level of probability of innovation. The difference between these results and those of the simulations for rational firm's lies in the strategic effect. Fully rational firms may reduce further their R&D investment, anticipating the impact of their R&D in their rival's odds of innovating. For lower values of b, the reduction in the leader's investment is stronger than that of the follower. Whence, the gap between the leader and the follower decreases leading to a decrease in the value of the leader and increase in the value of the follower. For high levels of spillovers, the rate of decay of the leader's investment get's smaller until investment actually starts to increase, while the follower continues reducing its R&D spending. This allows the leader to gain market share again. This gain in market share is not reflected in a value gain most probably because of the investment costs in which the firm is incurring.

In order to analyse the effects of the costless R&D spillovers on social welfare, we compute the statistics for the present discounted value of consumer surplus and producer benefits obtained in the simulation of the ergodic equilibrium. We use firm's value, $V_i(w_i, s)$, as a proxi for producer benefits. We register the observed outcomes for both the welfare measures and weight them by their probability of occurrence. Figure 3.9 depicts the percentage change of the expected discounted value of consumer surplus for the rational firms case (solid line) and the myopic case (dashed line). For myopic firms, the impact of the extent of spillovers on consumer surplus is negligible. This is so because consumer surplus is fully determined by the rate of innovation, which stays fairly unchanged in the costless myopic case as depicted in figure 3.8. However, if firms behave rationally, an increase in costless R&D spillovers leads to a decrease in consumer surplus. The direct comparison of these two cases allows us to identify the strategic effect as the source of the decrease in consumer welfare. The incorporation of the strategic effects in firms' decisions adds a further disincentive to invest in R&D which will translate into lower rates of innovation and consequently higher prices and lower consumer welfare.

In what concerns the impact of costless spillovers in the EDV of Producer benefits for rational (solid lines) and myopic firms (dashed lines), illustrated in figure 3.10, the results are quite surprising. The direct comparison of these two figures suggests that, in general, in the presence of costless R&D spillovers, firms would be better off interacting in a myopic manner than if they behave fully rationally. If firms ignore the strategic effect, there are welfare gains to both firms due to the extra source of R&D, which is free. However, as previously shown in figure 3.9, these gains are not passed onto the consumers. If firms behave rationally, while the follower continues to benefit due to cost minimization from free riding in the leader's R&D, the leader registers a value loss as a result of the decrease in its own investment for strategic reasons. Consequently, its losses in terms of the probability of innovation more than compensate for the gains in cost saving from costless R&D spillovers.

3.3.3 Absorptive Capacity

Optimal Policy Functions

In what follows we argue that spillovers do not necessarily deliver market structures which are detrimental for consumers. We present the results obtained for the optimal policy functions and the simulation of the model under the assumption that the absorption of R&D spillovers requires absorptive capacity. Figure 3.11 illustrates the evolution of R&D investment with the increase in the extent of spillovers.

As b increases, firm 1's R&D investment decreases for market structures where the competitor has a high investment level (around $w_2 = 5$). In these market structures, the effectiveness of own R&D increases by allowing the firm to absorb a substantial amount of external R&D. Therefore, when there is an external pool of R&D and the firm has some positive level of absorptive capacity, R&D investment decreases because firms can now achieve the same probability of innovation with lower levels of R&D spending.

In their static, deterministic model with symmetric firms, Cohen and Levinthal (1990) find that externalities in R&D might increase the amount of R&D investment undertaken in an industry. The predictions implied by the model presented here suggest their results cannot be extended to a dynamic, stochastic framework with firm heterogeneity. However, the fall in R&D investment does not imply that the efficiency level of firms will decrease.

In fact, the analysis of figure 3.12 suggests that the probability of success of the follower increases, in spite of the decrease in R&D investment. In turn, the leader's rate of innovation will decrease for market structures in which it has very high efficiency levels and its rival has a very low efficiency level. At these market structures the follower benefits much more from each unit of R&D investment.

Figure 3.13 illustrates how the changes in the rate of spillovers affect the value of the firm. Spillovers have a positive impact on firm 1's value in most market structures. This increase is higher when the rival firm is investing more heavily in R&D ($w_2 = 5$). Given the convexity of the value function for lower levels of efficiency, the increment in value is higher when firm 1's efficiency level is lower, and decreases with w_1. The only market structures where firm 1 experiences a value decline due to the increase in spillovers are those in which the rival has a small efficiency level, especially if, concomitantly, firm 1 has an efficiency level around its maximum for investment ($w_1 = 5$). This is so because the rival is engaging in few R&D, while the firm's own R&D investment is large. In these market structures, the leader firm has virtually nothing to gain from spillovers and the rival is free riding (as it emerges from figure 3.12 as well).

Simulation Results

In order to analyse how the externalities actually affect the evolution of the market structure in the industry, we now turn to the analysis of the results obtained when simulating the model assuming firms are required to have absorptive capacity to benefit from external R&D. Figure 3.14 is analogous to figure 3.6.

As the extent of spillovers increase there is a clear shift in the likelihood of occurrence from highly concentrated market structures to more competitive ones. The peak on the lower left hand corner of the graphs captures the high probability of occurrence of market structures where the firm with the highest efficiency level has a very strong lead over the follower. As b increases, this peak becomes smaller, and the likelihood of market structures closer to the diagonal, the ones where the gap between the efficiency levels of the leader and the follower is at its minimum, rises. The decrease in the mean level of concentration in the industry, as measured by the concentration index C1, with the increase in the extent of spillovers is confirmed by figure 3.15.

The fall in concentration is even higher when firms behave strategically than when they are myopic. The difference between the two lines is attributable to the strategic effect and hence, the disincentive to invest in R&D. Since this effect is particularly strong for the leader, the gap between the two firms decreases further than in the myopic case.

Figure 3.16 shows the change in mean R&D investment, mean probability of success and mean value when b increases. Analogously to figure 3.8, the graphs on the left illustrate the results for rational firms, and the ones on the right show the results for myopic firms.

For both rational and myopic firms, an increase in spillovers decreases the amount of R&D undertaken in the industry, and this reduction is larger for the leader firm in the industry. The firm with the highest efficiency level will be engaging in the amount of R&D that will deliver fairly the same probability of innovation as in the full appropriability case, as made clear by the stability of the leader's probability of innovation. The follower, however, experiences an increase in the likelihood of enhancing its efficiency level. Whence, the follower will experience an increase in value, while the impact of spillovers on the leader's value is much less significant, and slightly non-monotonic. The amount of R&D investment and the probability of success of the follower firm actually overcome those of the leader at intermediate values of b.

The comparison of the graphs for rational and myopic firms suggests that, when the absorption of R&D spillovers requires absorptive capacity, the strategic effect of R&D speeds up the reduction in the gap between the leader and the follower. In fact, the disincentive to invest is higher for the leader. Although the level of R&D of the follower also decreases with spillovers, the rate of decay is smaller than that of the leader, such that for high enough levels of spillovers the follower's level of R&D actually overcomes that of the leader.

The welfare implications of R&D spillovers for consumers in the case of absorptive capacity are illustrated in figure 3.17.

As opposed to the results obtained for costless R&D spillovers, there is a clear positive impact on consumer welfare in the case of absorptive capacity. This result follows from the rise in the follower's likelihood of developing a new innovation and consequently experiencing an increase in its efficiency level. This benefit will translate into lower marginal costs of production and hence be passed on to consumers through a decline in prices.

The evolution of producer benefits is illustrated in figure 3.18. The changes in producer surplus reflect two types of effects, namely the cost saving in R&D spending steaming from the efficient choice of in-house and external R&D, and the effects of R&D spillovers on the relative position of the two firms in the industry. A firm experiences a value increase, ceteris paribus, if its probability of success increases more than that of its rival. Figure 3.18 shows a sharp rise in the follower's value, which is higher when firms interact strategically. In turn, the leader's value remains fairly constant when there is no strategic effect of R&D spillovers, but it experiences a clear-cut value decrease if firms take into account, in their investment decisions, the effect of their R&D in the rival's position. The disincentives to invest arising from the strategic effect are higher for the leader, leading to a sharper decrease in its level of R&D investment. Consequently, the follower will see its relative position improve. Even if the probability of innovation of the leader stays fairly stable, as made clear in figure 3.16, its value decreases since the losses from the increase in the followers probability of innovation preside over the gains from R&D cost saving.

3.4 Conclusions

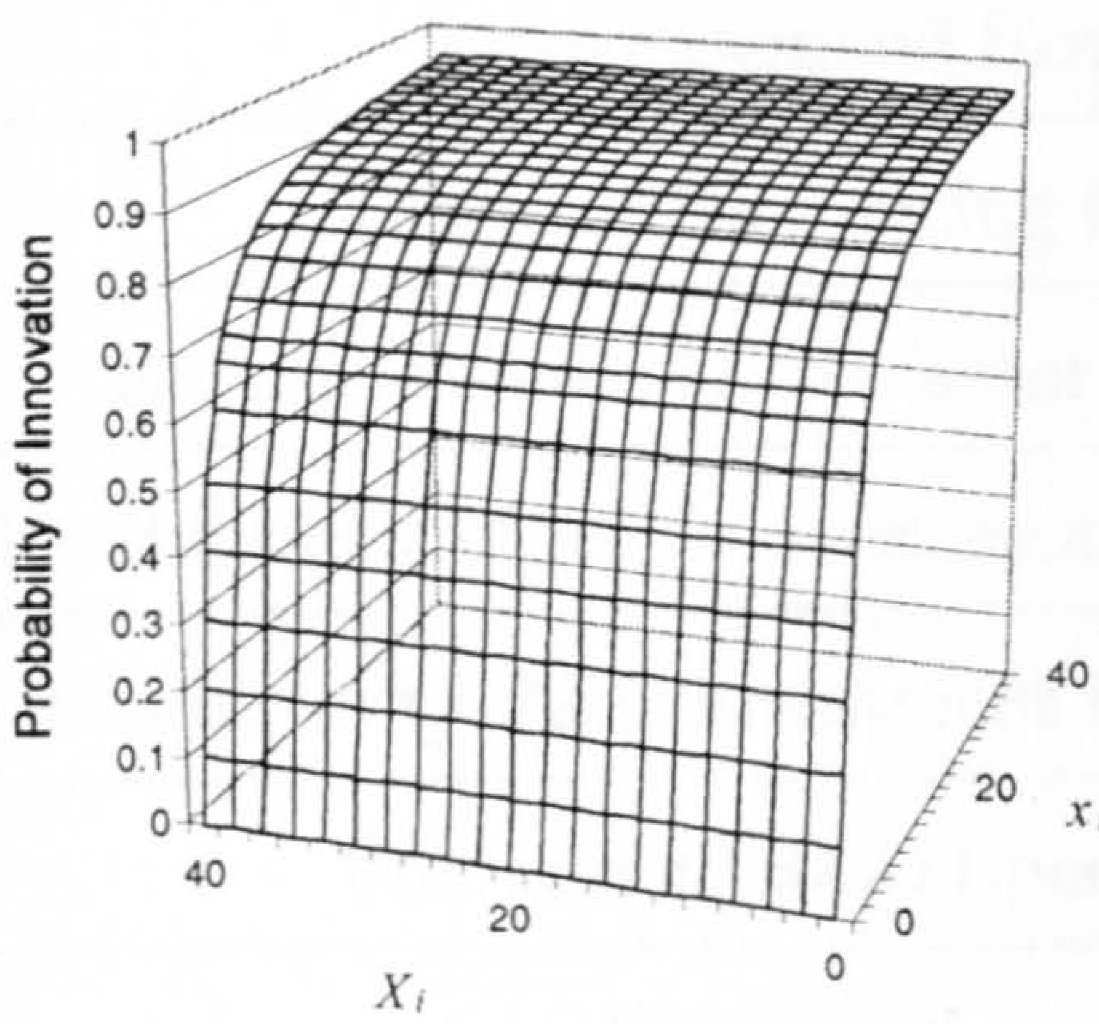
Apart from the oligopolistic interactions in investment decision-making embedded in the Ericson-Pakes (1995) framework, the model presented here allows for knowledge to spill inside the industry. Therefore, when a firm invests in R&D, it enhances its probability of developing a new idea, increasing the competitive pressure on the rival firm, but through knowledge externalities it will also improve the likelihood of the rival being successful in its innovative process.

Our simulations show that the degree of appropriability plays a fundamental role in determining the degree of asymmetry between firms. However, the impact of spillovers on the incentives to conduct R&D depends on the process through which firms acquire external knowledge. If external knowledge is a pure public good, i.e., if there are no knowledge requirements to learn from outside R&D, the strategic effect lowers the levels of R&D bellow those which would guarantee a stock of usable R&D that would deliver the rates of innovation obtained under full appropriability. Firm's R&D choices translate into a lower rate of innovation in the industry, implying higher marginal costs and higher prices for consumers. If firms would ignore the positive feedback of their R&D on the rival's rate of innovation, then they would engage in the amount of R&D that together with the external R&D available to them, delivers an optimal probability of success which is fairly similar to that of full appropriability. If this would be the case, then the market structure would remain unchanged and consumers would experience no gains or losses of welfare. Firms, on the other hand, would gain from the emergence of a alternative, free source of R&D.

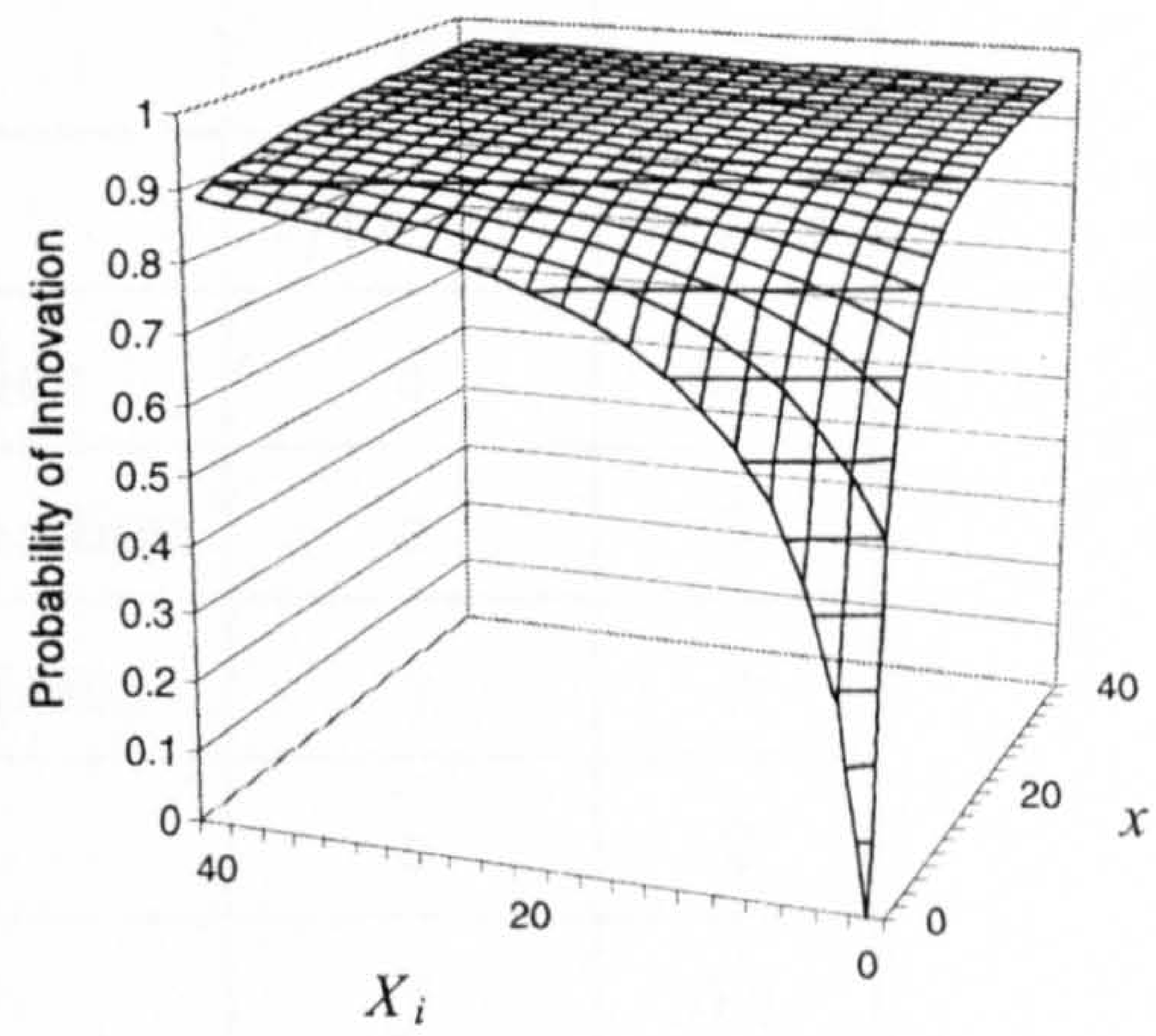
When the absorption of external R&D requires absorptive capacity, an increase in spillovers delivers more competitive market structures. There are welfare gains for consumers due to the increase in the amount of innovations developed in the industry. There are also welfare gains from externalities for the follower in the industry, because this firm able to improve its position relative by investing more heavily than the leader along with the R&D cost saving it is able to incur.

The main substantive finding relates to the way in which market structure changes as spillovers of a pure public good kind increase. It might seem intuitively that such spillovers would deter R&D and so lower the fixed costs incurred by the incumbent and thereby lower equilibrium concentration. What we show, however, is that the strategic interaction in our present model is as to lead a rise in concentration (paragraph 3, page 136).

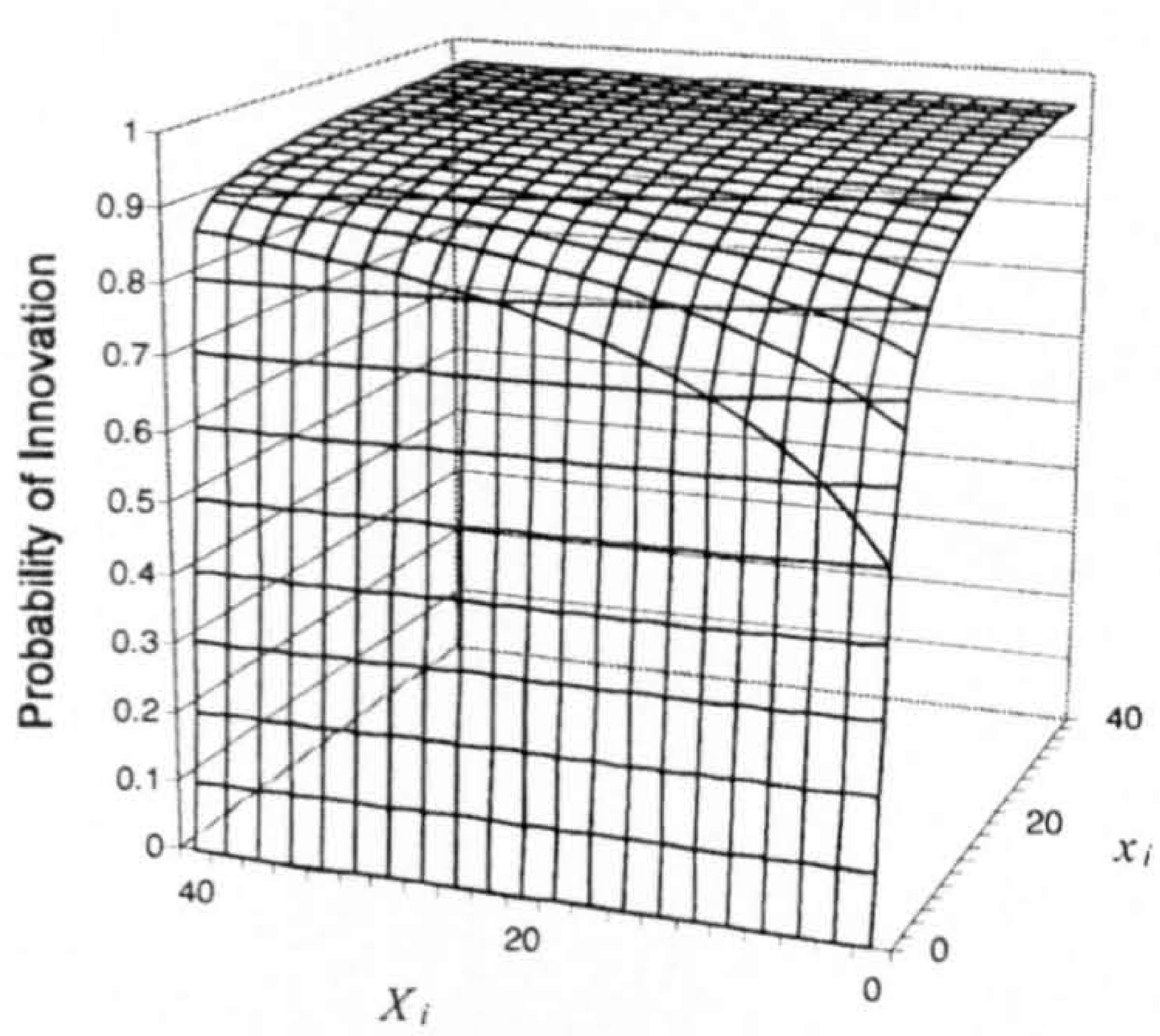
Figure 3.1: The Impact of Spillovers on the Probability of Innovation Under the Different Appropriability Scenarios



a) Full Appropriability



b) Costless R&D Spillovers

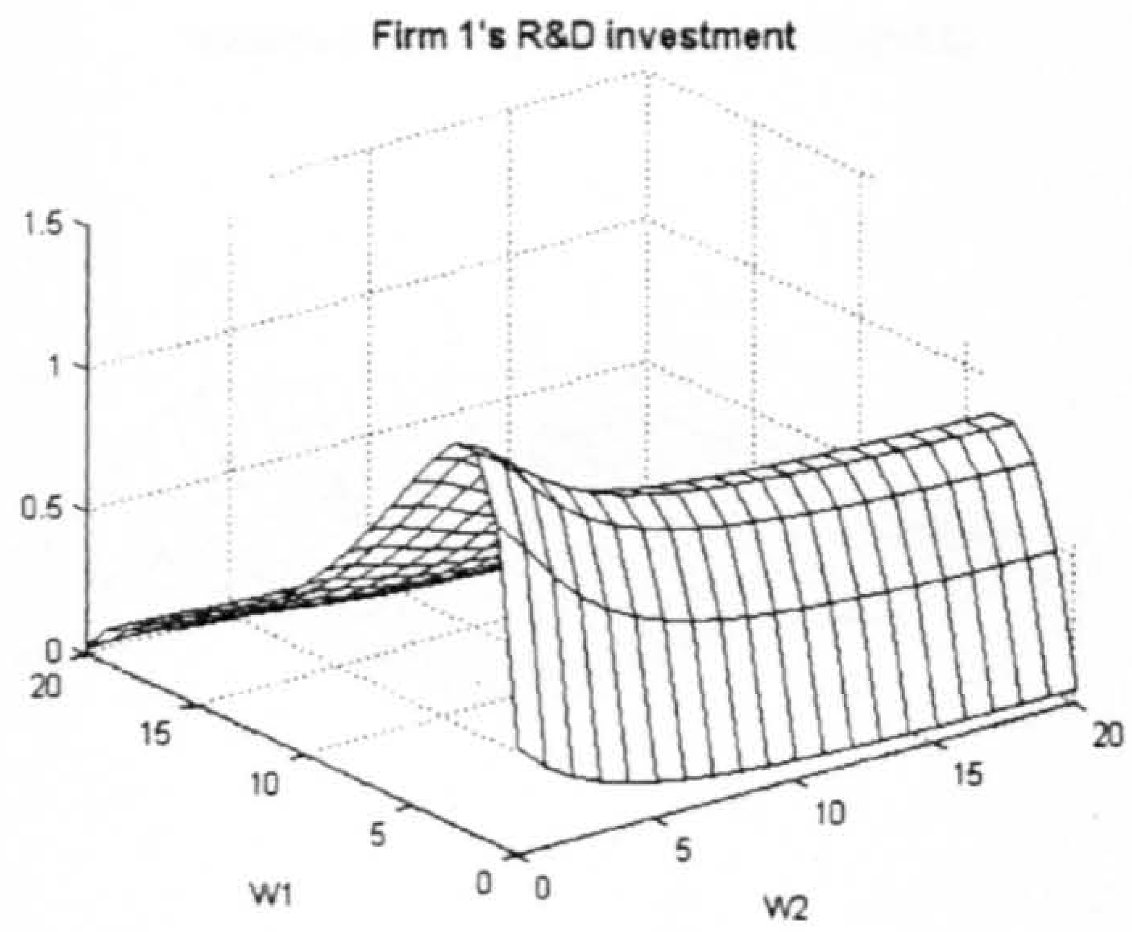


c) Absorptive Capacity

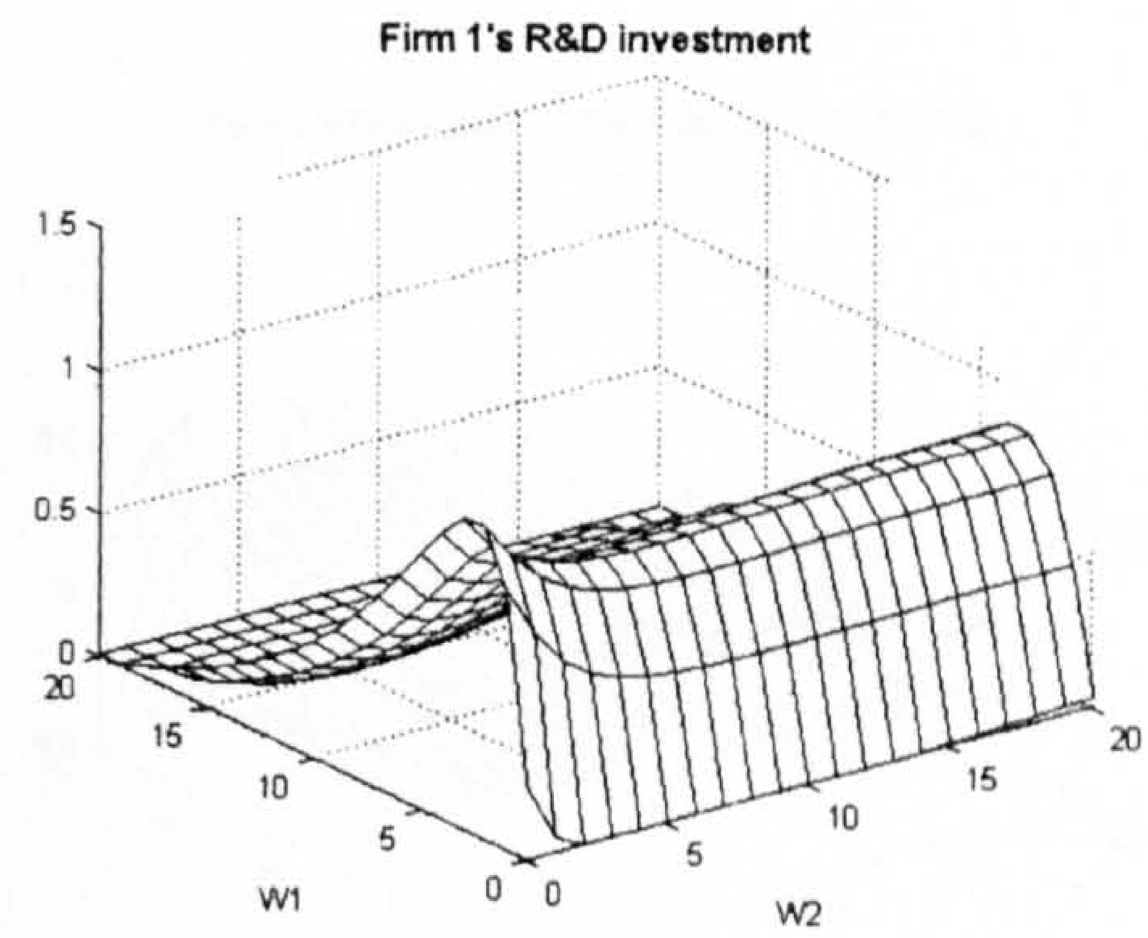
Table 3.1: Parameter Values

Linear Demand Intercept	A	8
Linear Demand Slope	B	1
Discount Rate facing firms	$1/(1+r)$	1/1.08
Rate of increase of the factor price index	δ	0.7
Productivity of R&D investment for innovation	a	3
Productivity R&D investment for absorption	γ	3
Unit cost of R&D Spending	c	2
Fixed Costs	f	0.1
Rate of Decrease in Marginal Costs	η	0.3

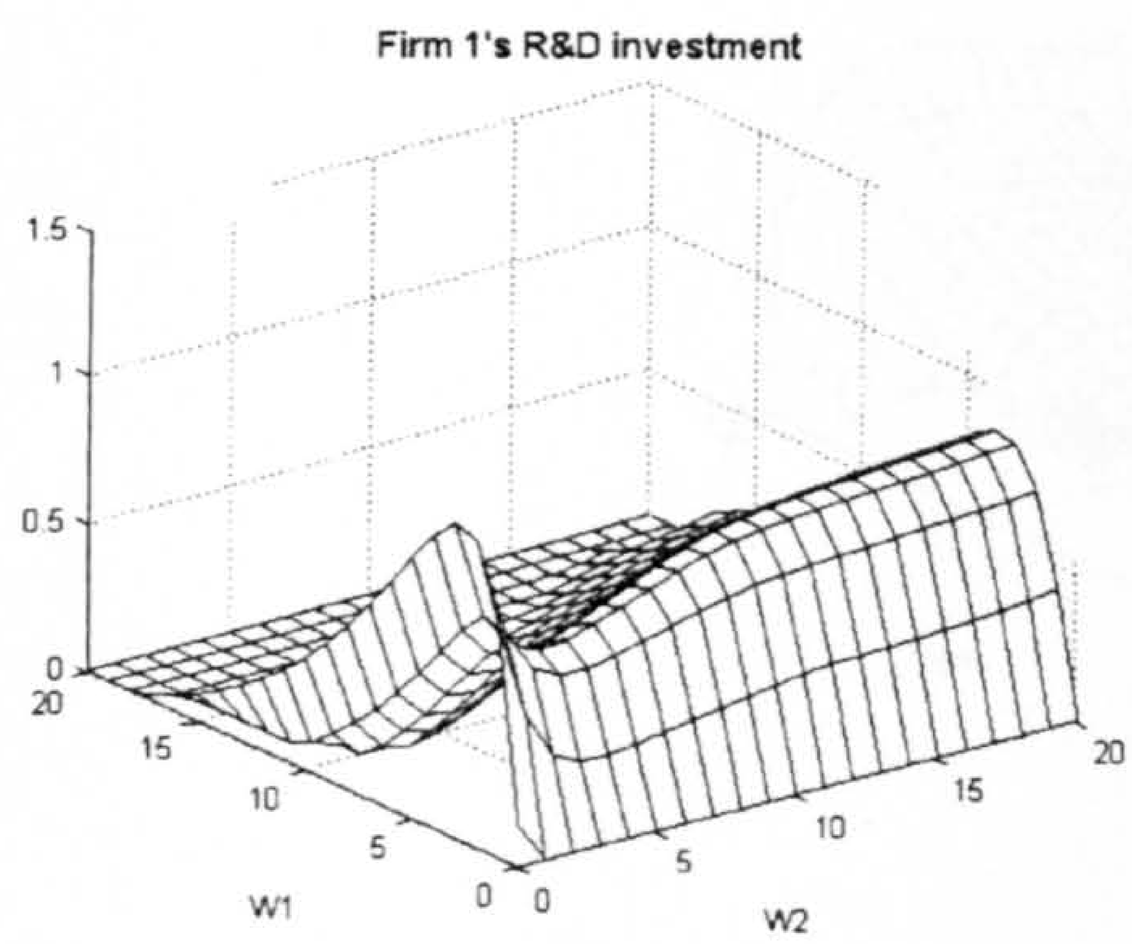
Figure 3.2: The Impact of Costless Spillovers on the Optimal Investment Policy Function



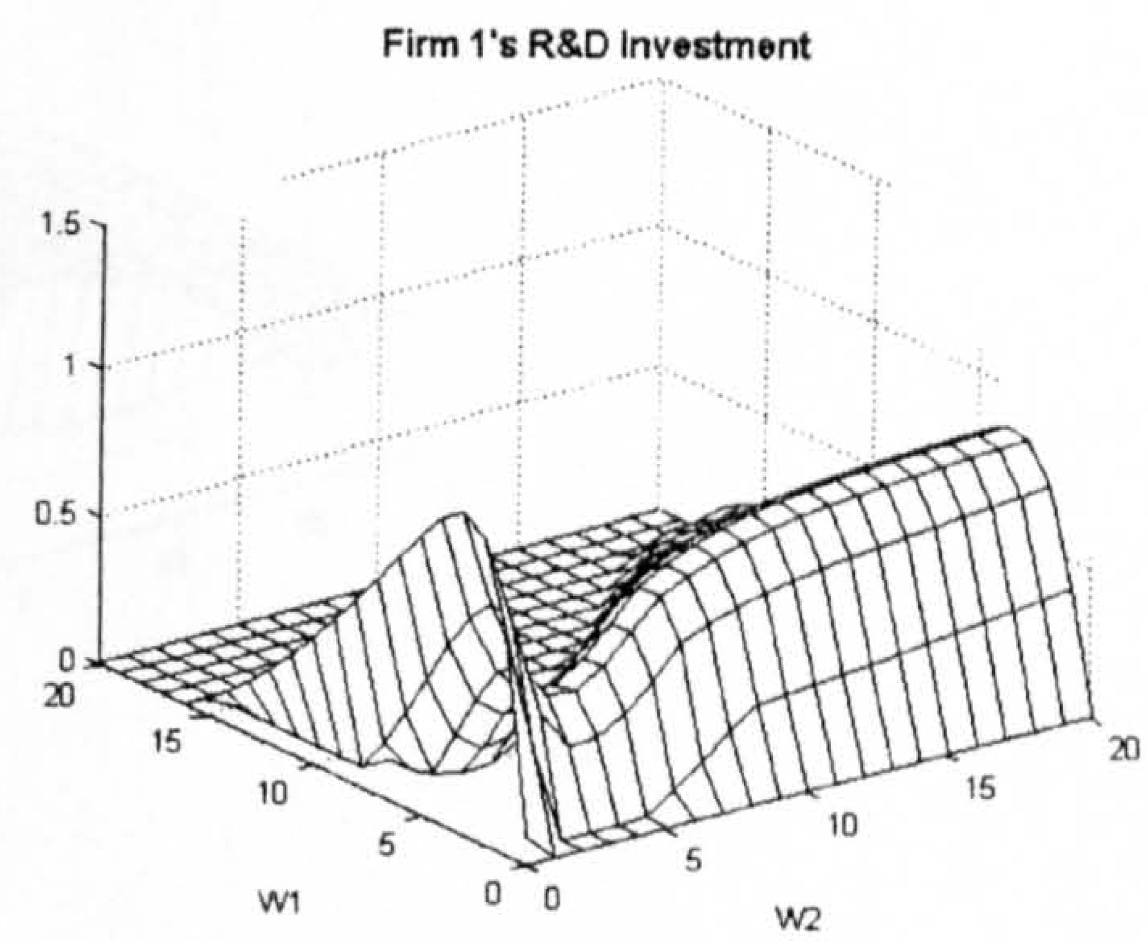
b=0.0



b=0.2



b=0.5



b=0.8

Figure 3.3: The Impact of Costless Spillovers on the Optimal Probability of Innovation

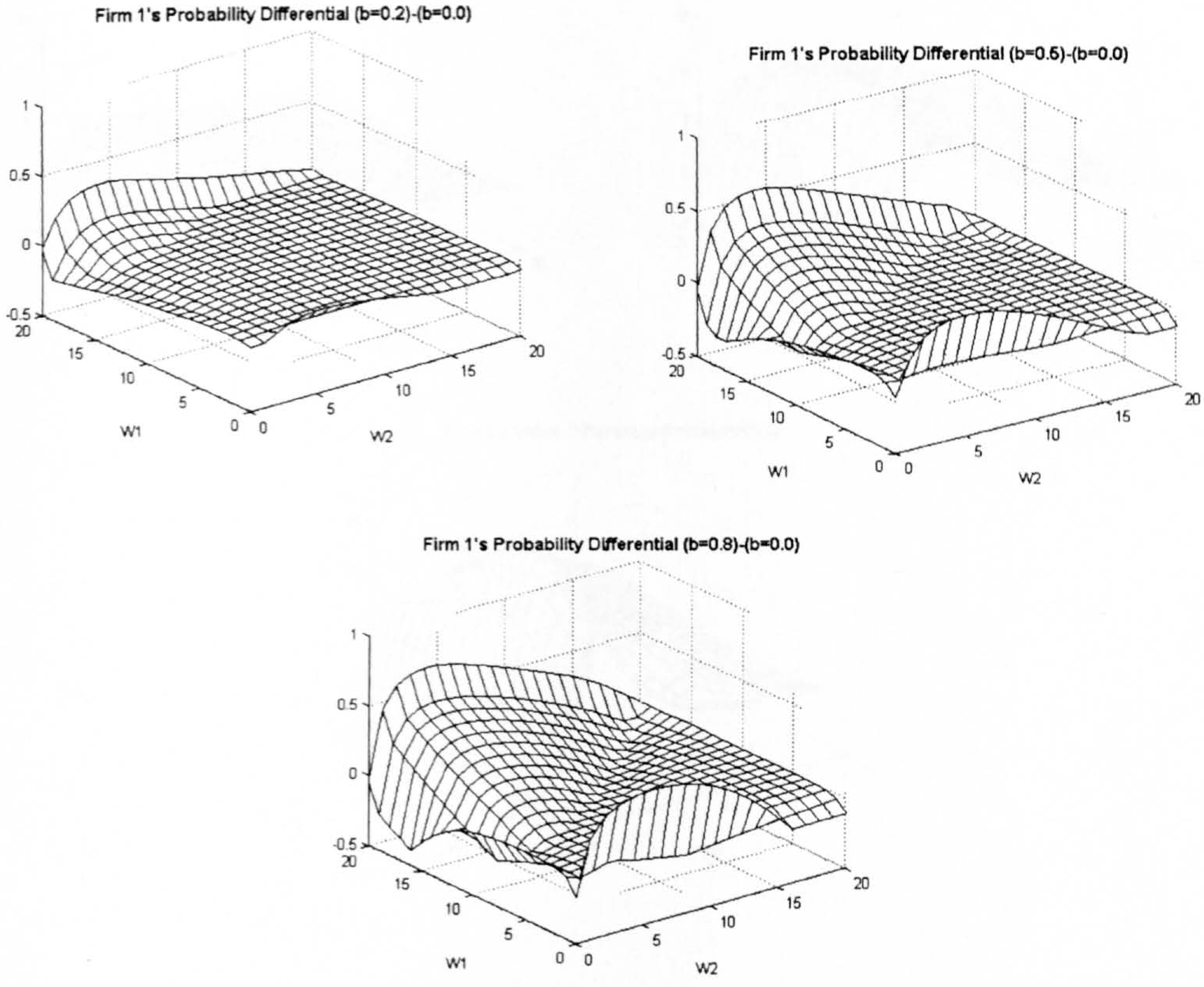


Figure 3.4: The Impact of Costless Spillovers on the Optimal Firm's Value

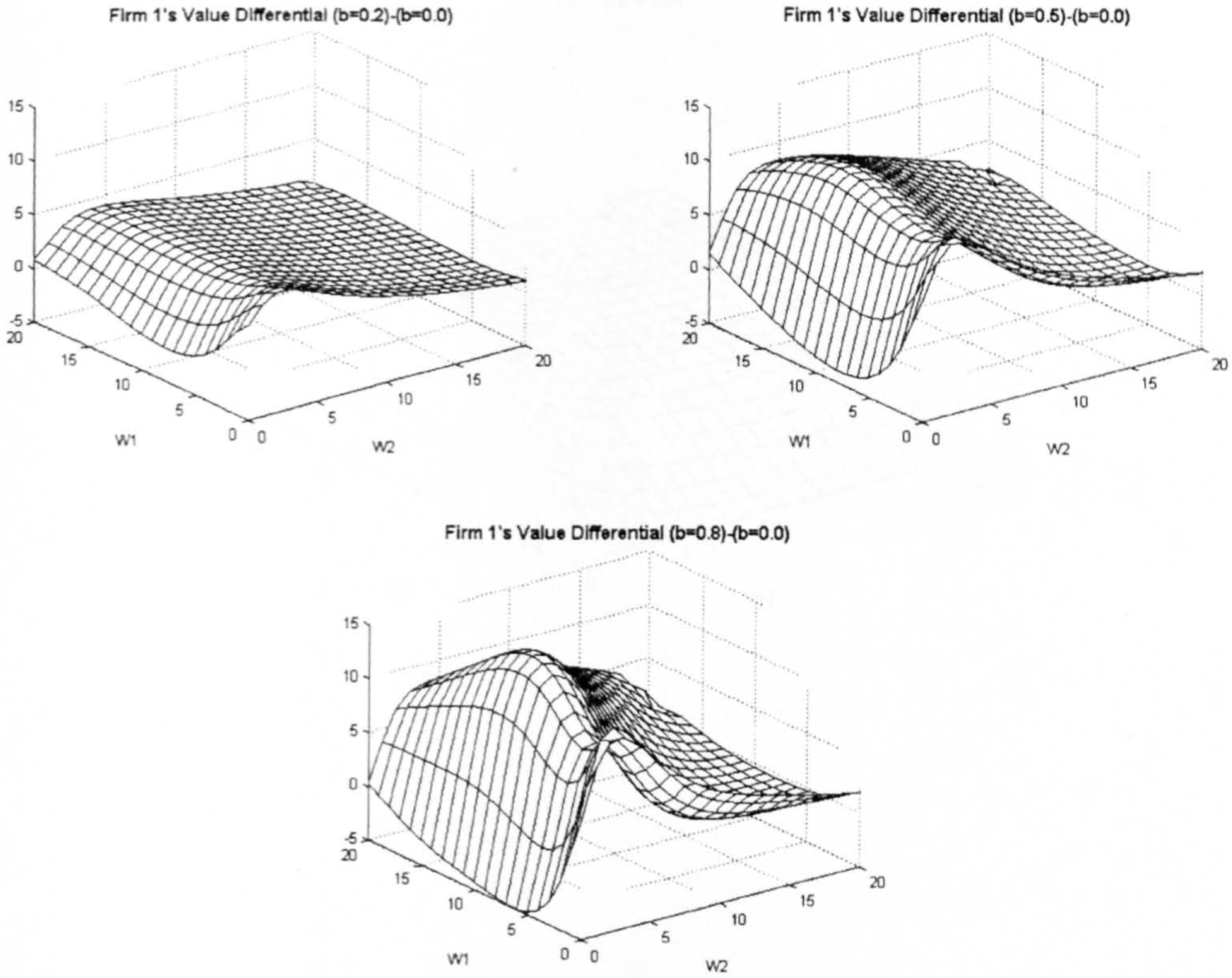


Figure 3.5: The Value Function for b=0.0

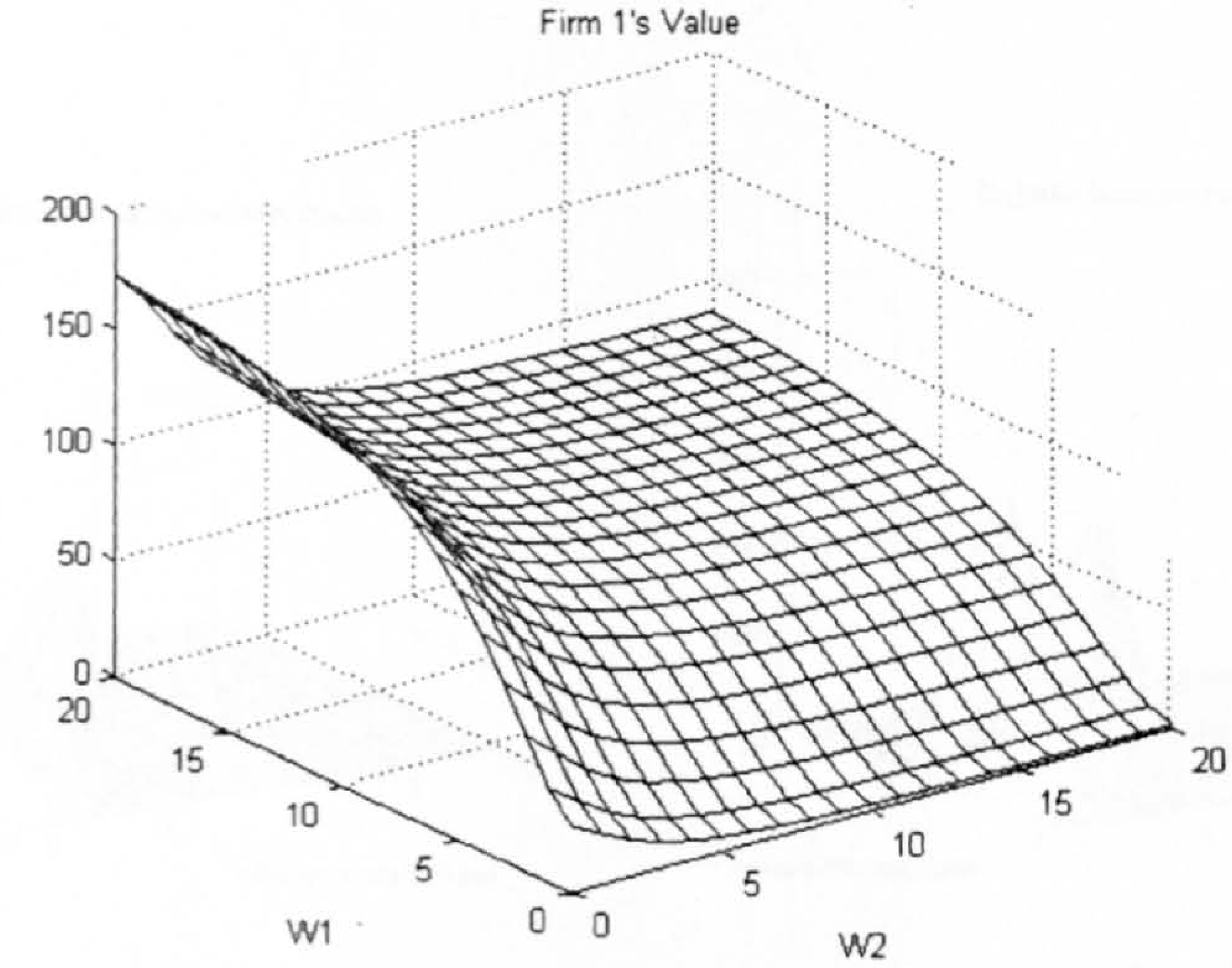
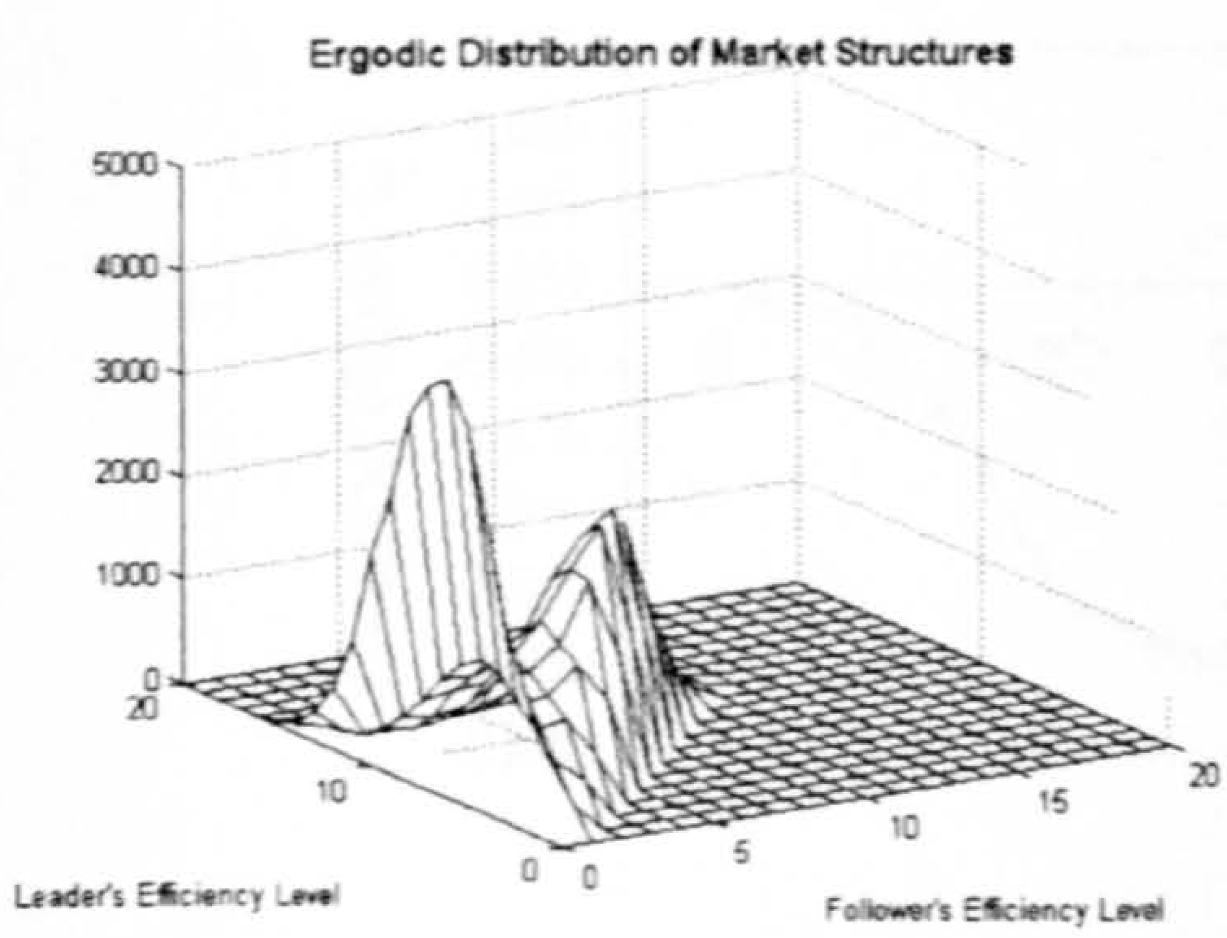
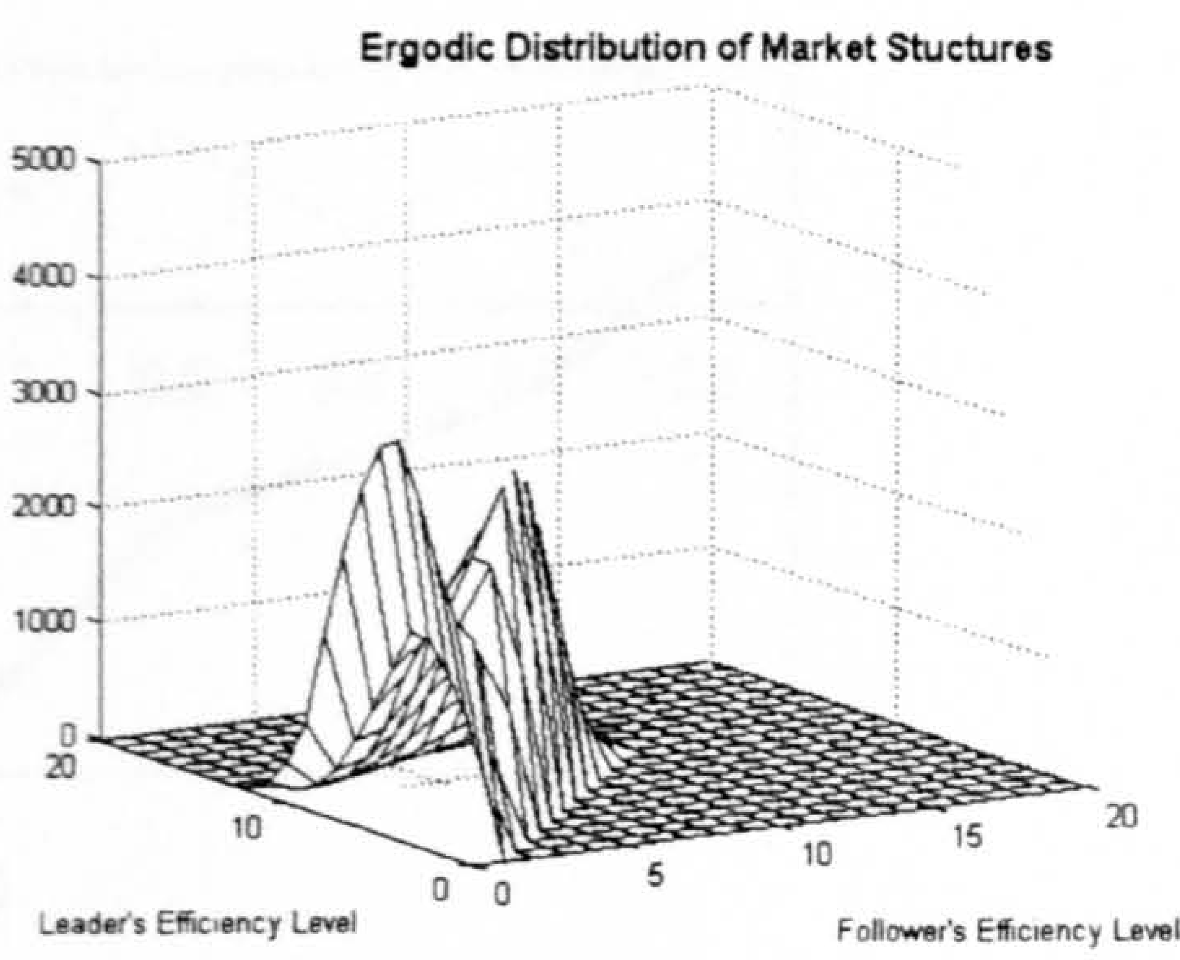


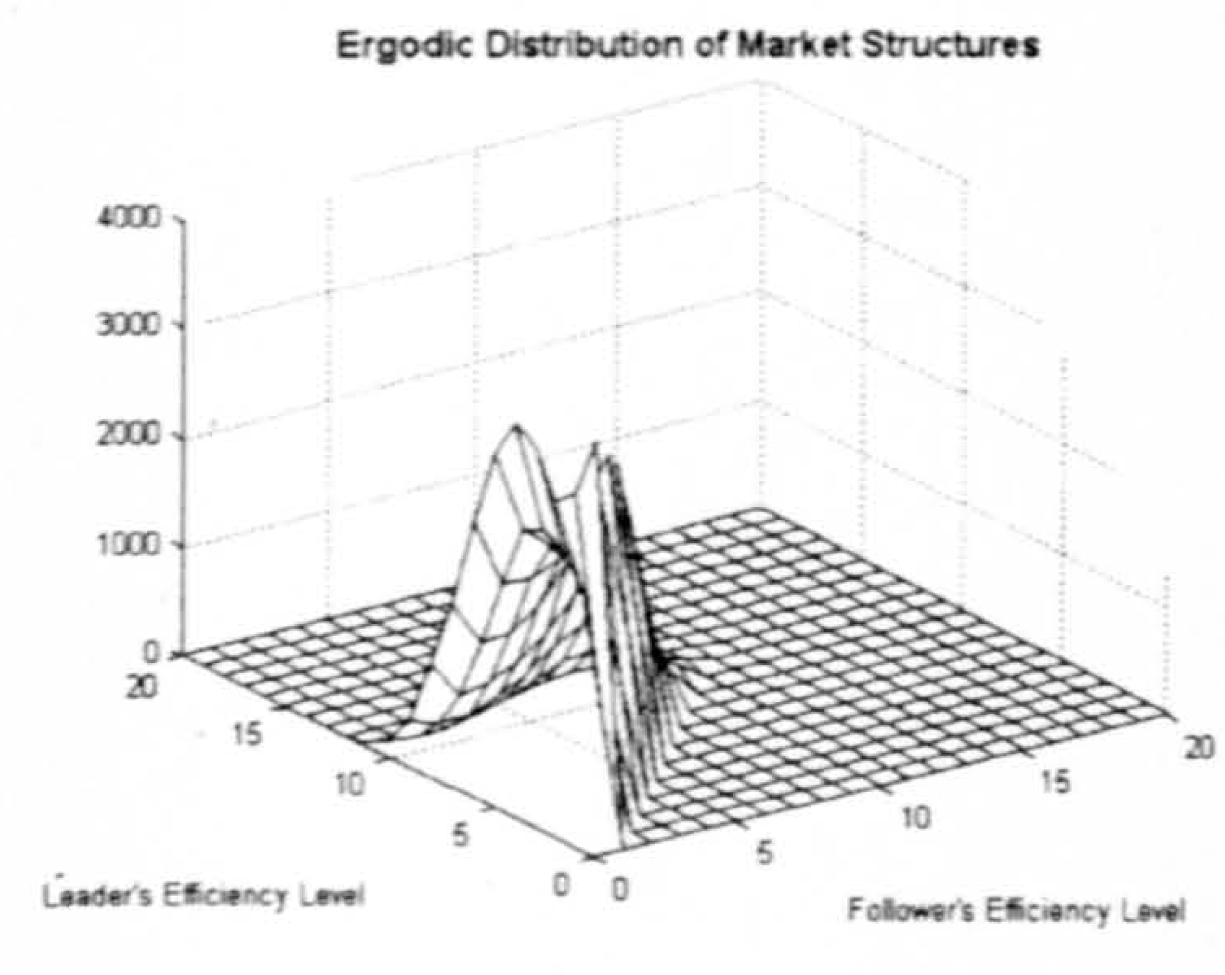
Figure 3.6: The Impact of Costless Spillovers on the Ergodic Distribution of Market Structures



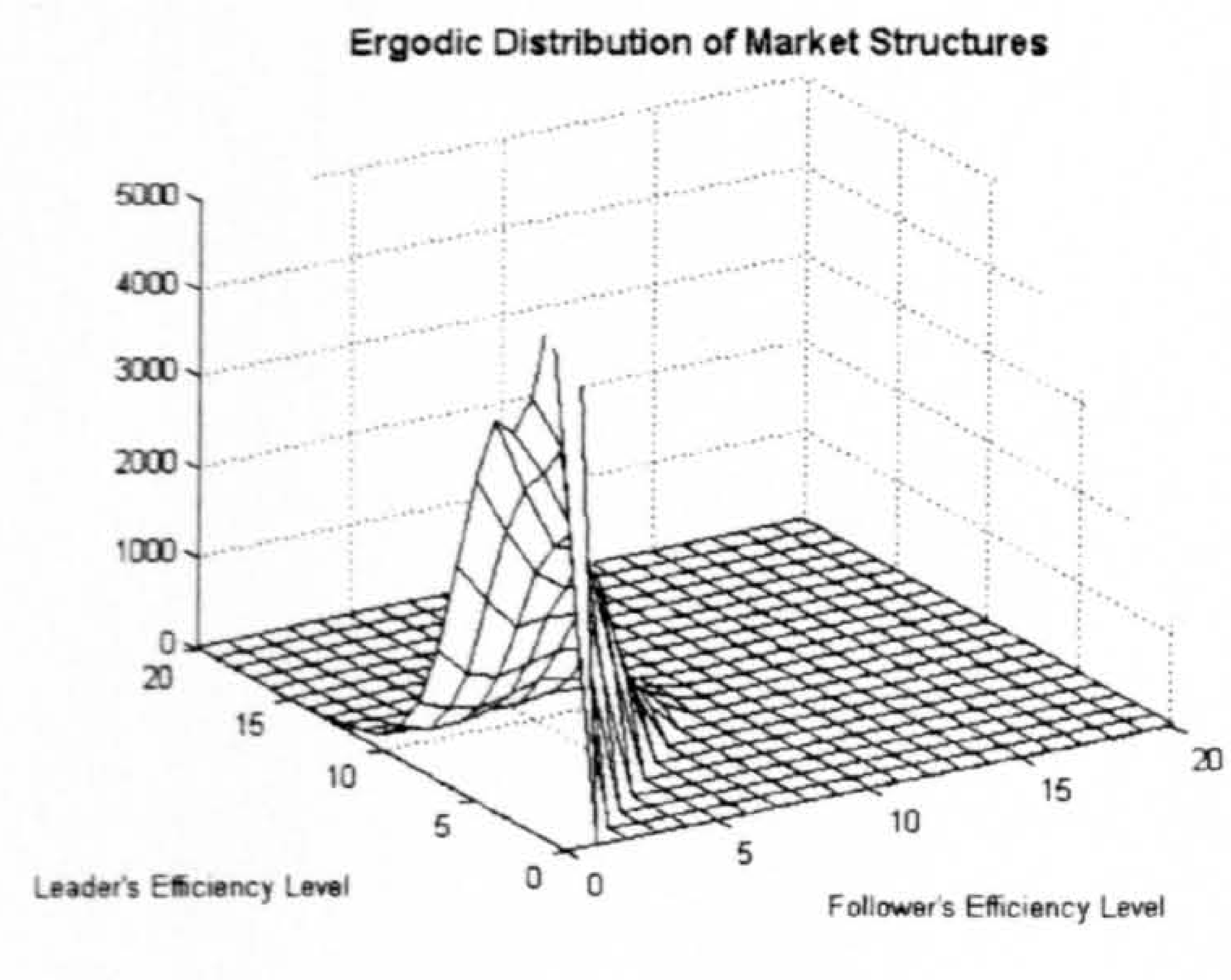
b=0.0



b=0.2



b=0.5



b=0.8

Figure 3.7: The Impact of Costless Spillovers on the Concentration Index C1

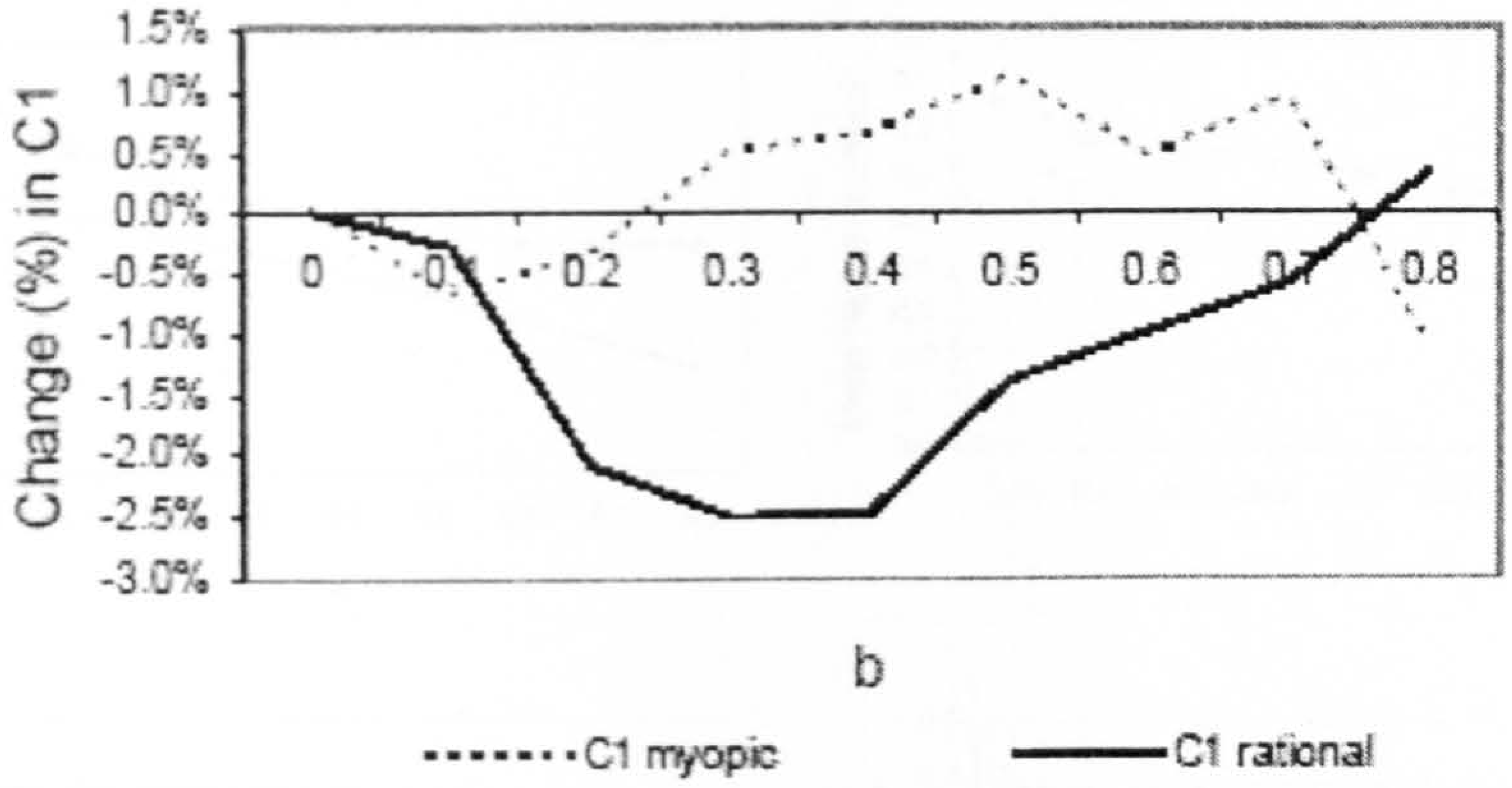


Figure 3.8: The Impact of Costless Spillovers on Investment, Probability of Success and Firm's Value

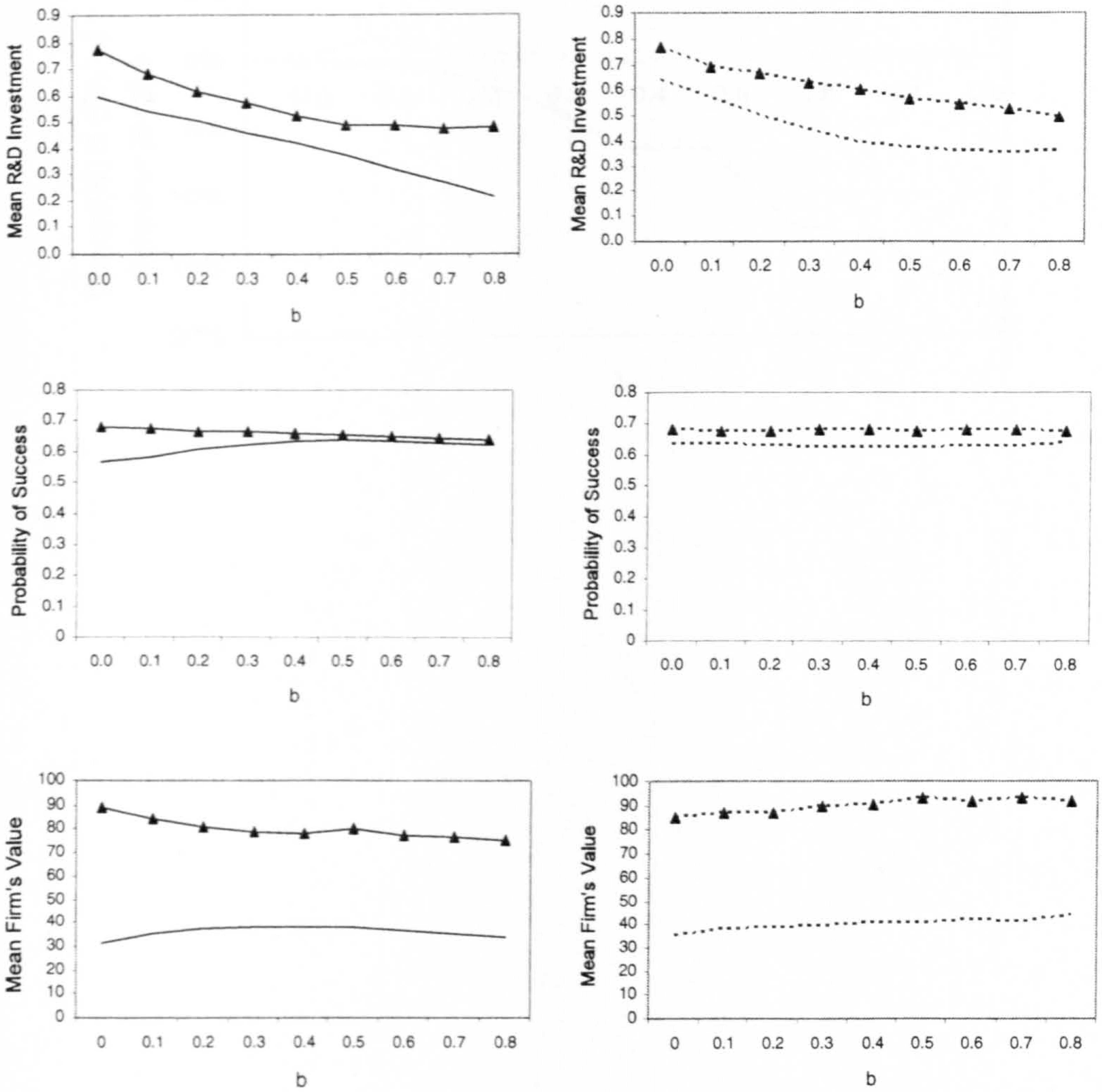


Figure 3.9: The Impact of Costless Spillovers on Consumers' Surplus

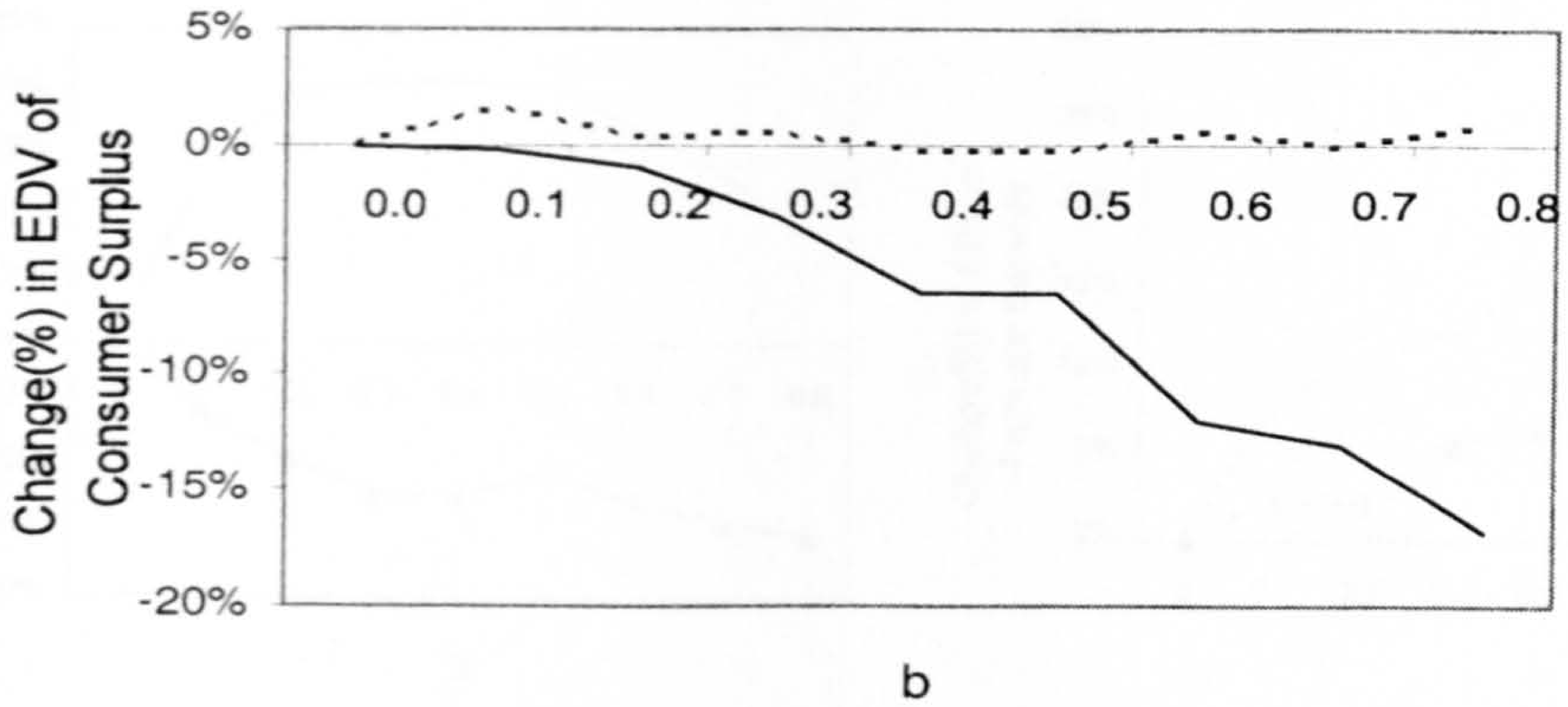


Figure 3.10: The Impact of Costless Spillovers on Producer Benefits

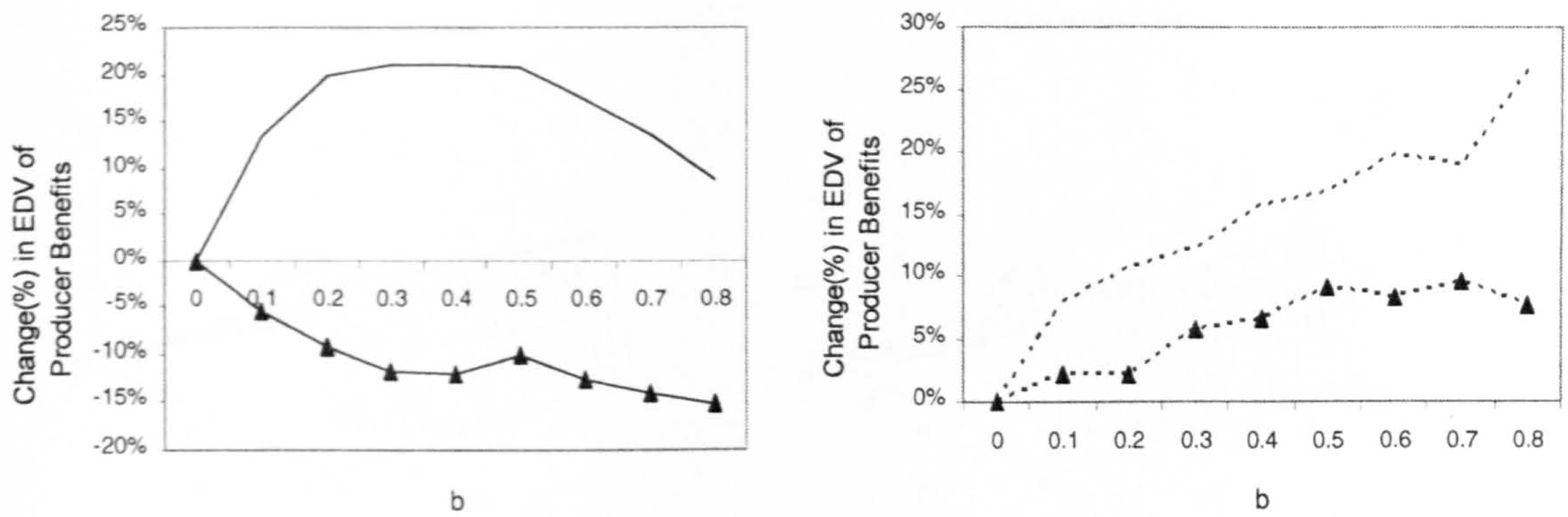
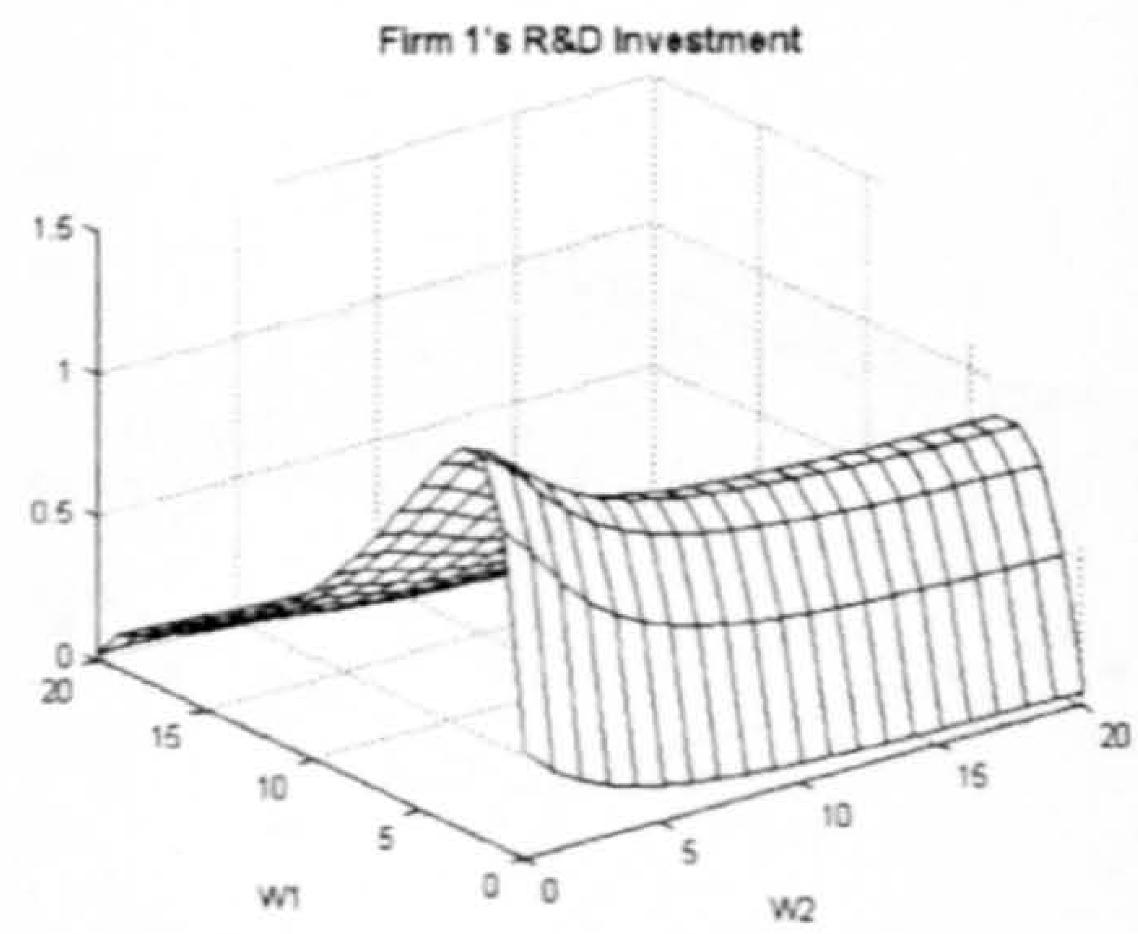
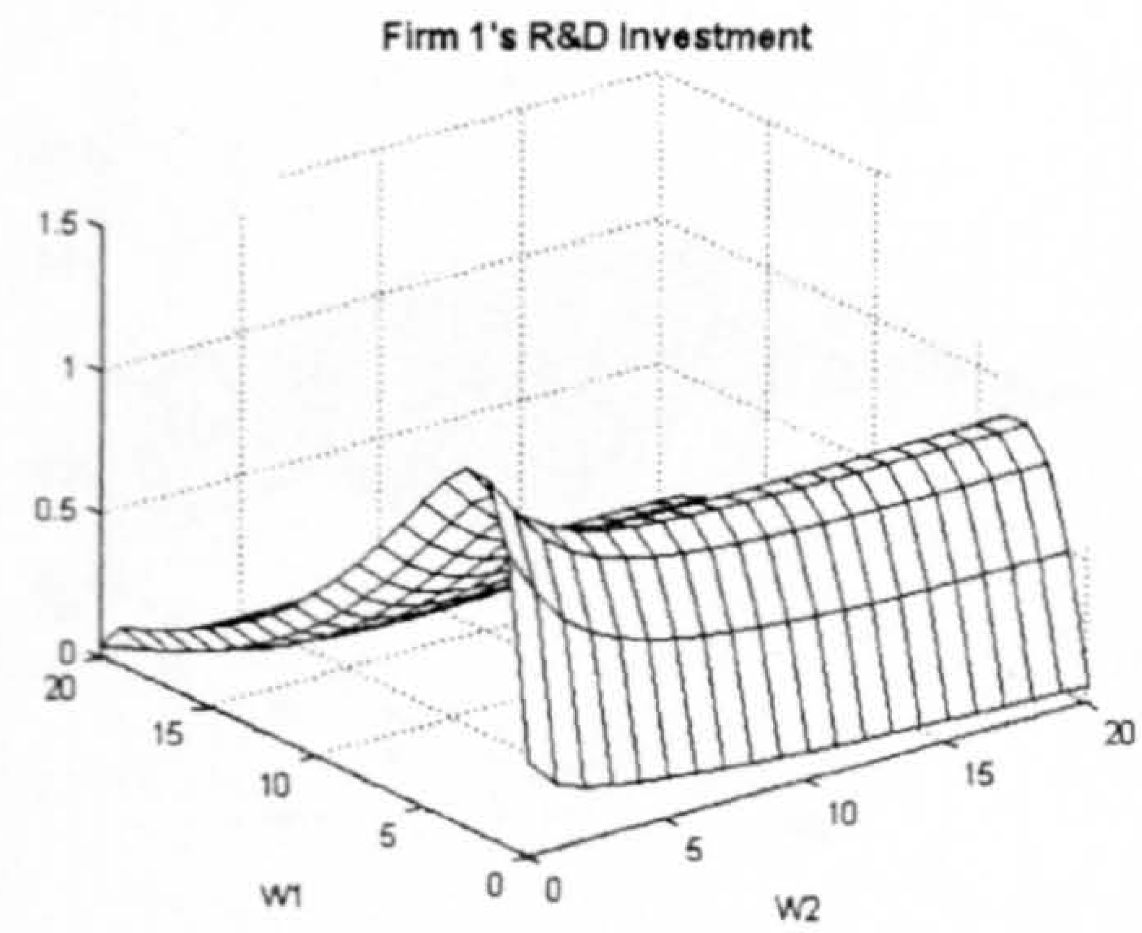


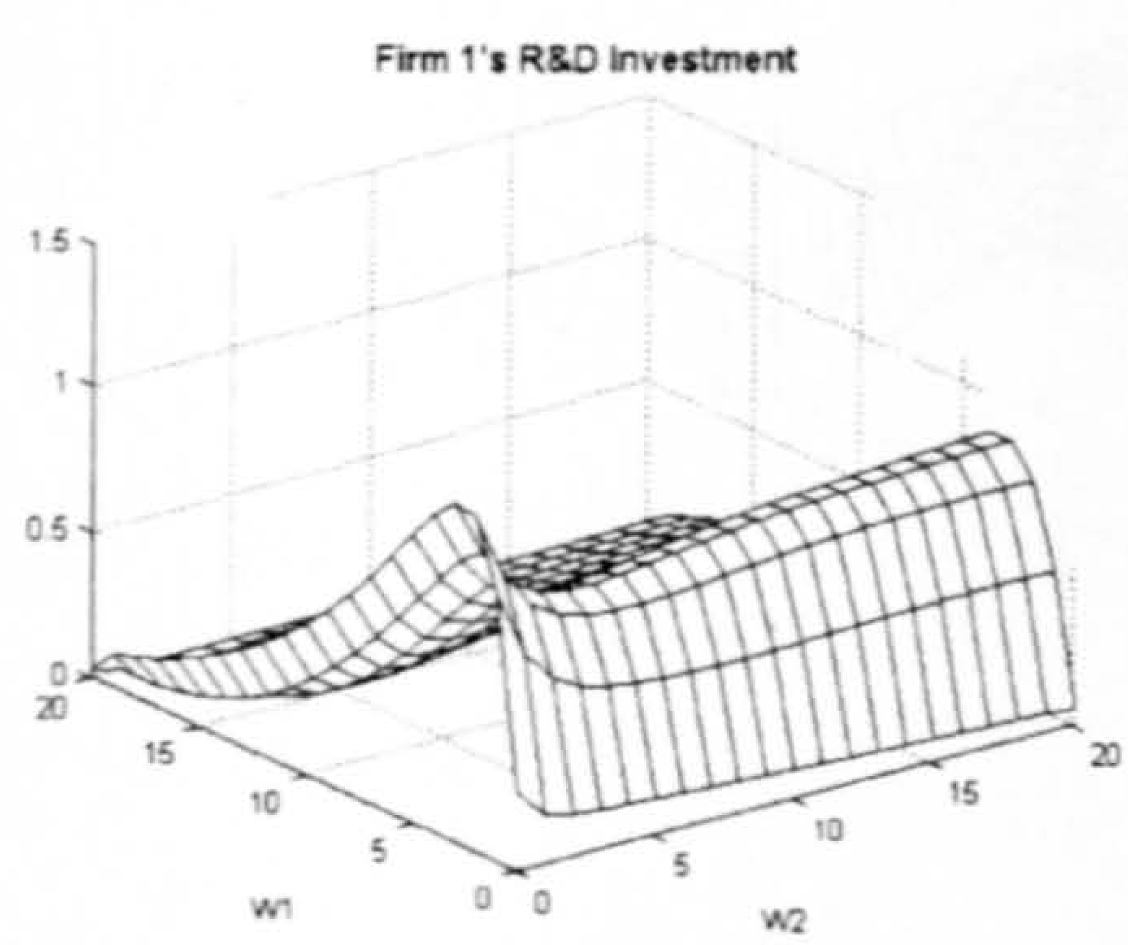
Figure 3.11: The Impact of Spillovers on the Optimal Investment Policy Function in the Case of Absorptive Capacity



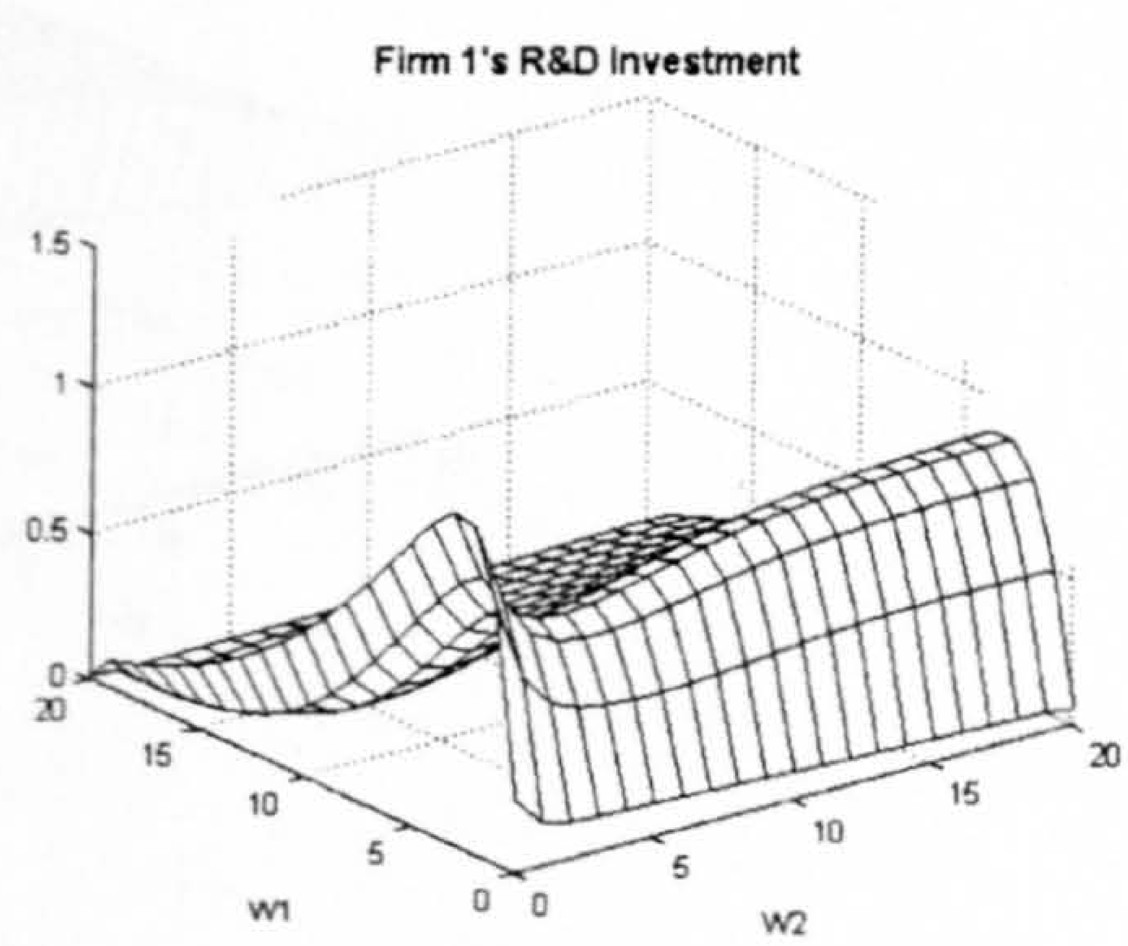
b=0.0



b=0.2

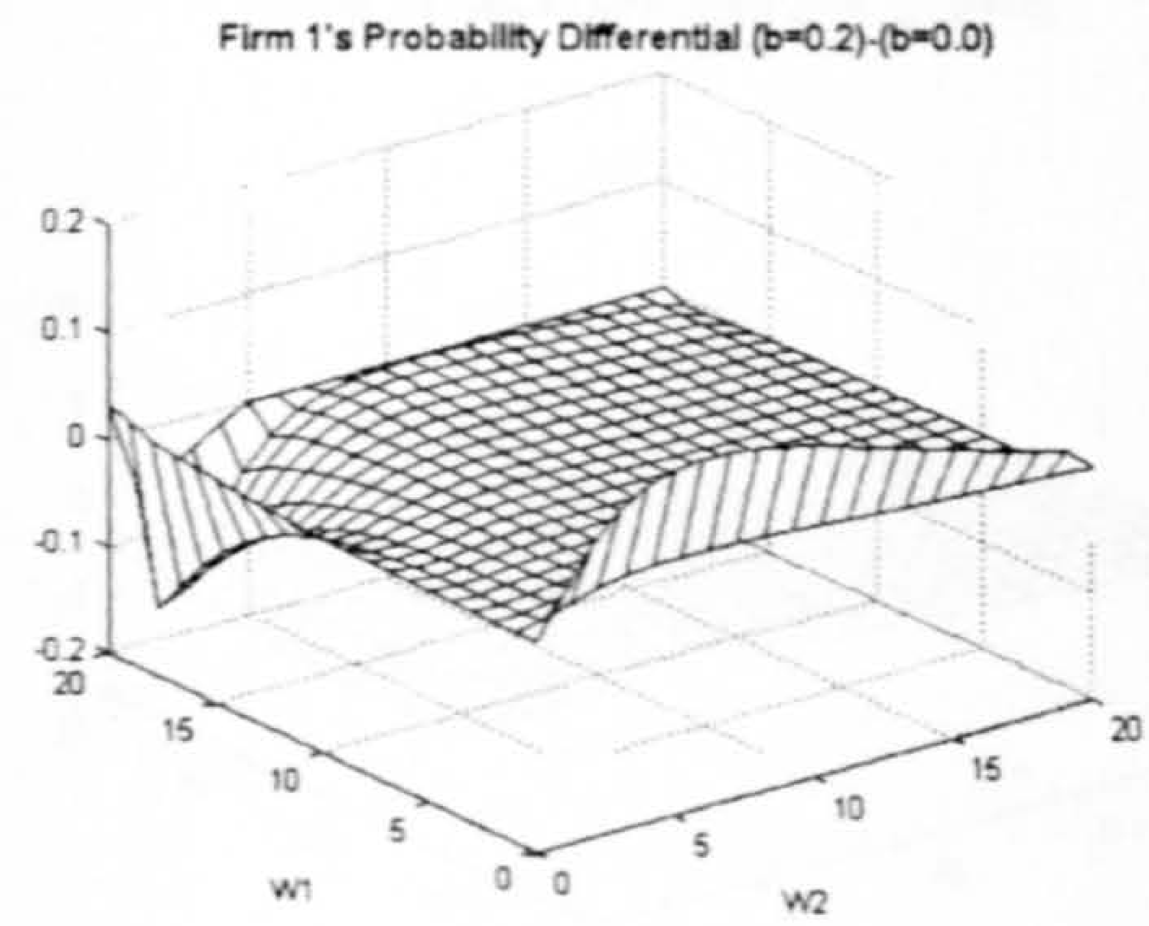


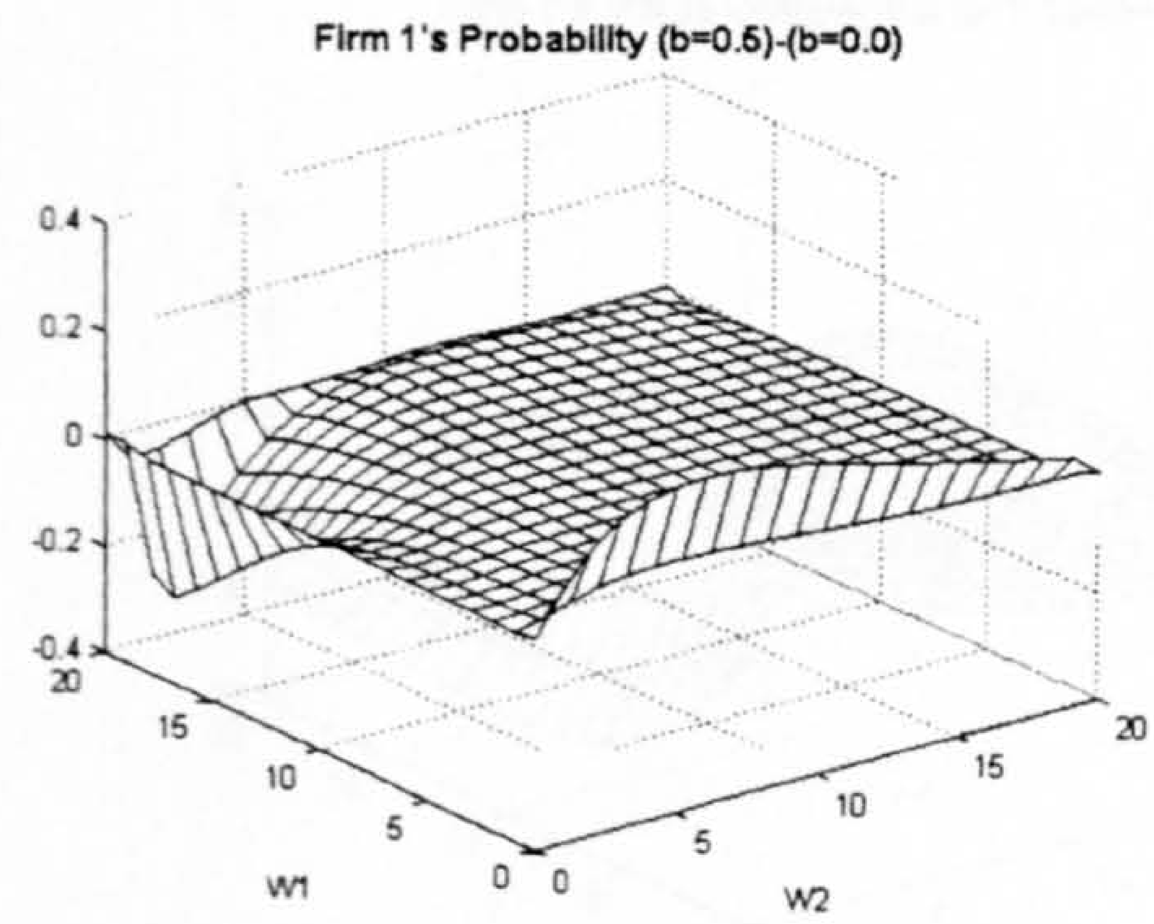
b=0.5



b=0.8

Figure 3.12: The Impact of Spillovers on the Optimal Investment Probability of Success in the Case of Absorptive Capacity





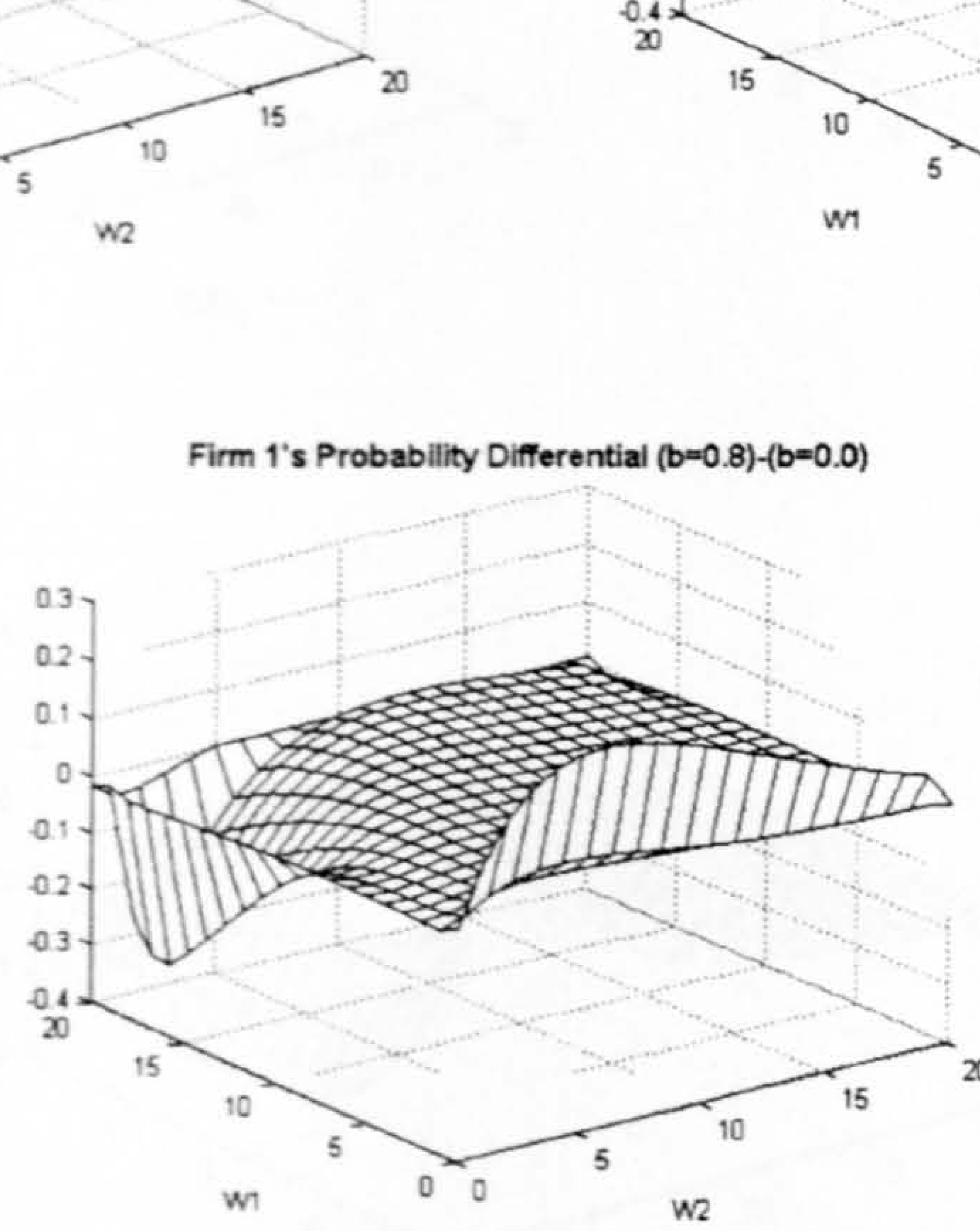


Figure 3.13: The Impact of Spillovers on Firm's Value in the Case of Absorptive Capacity

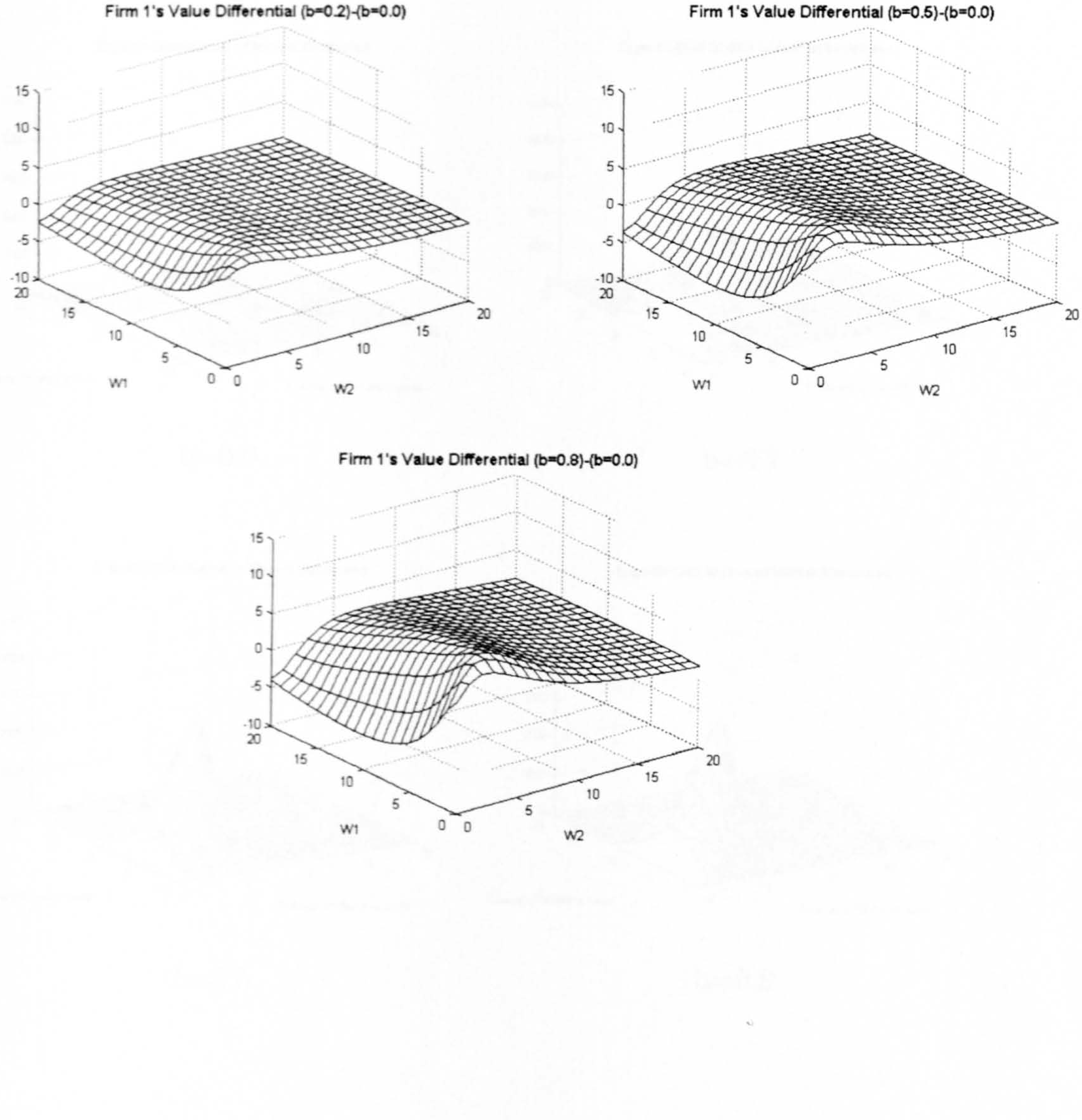
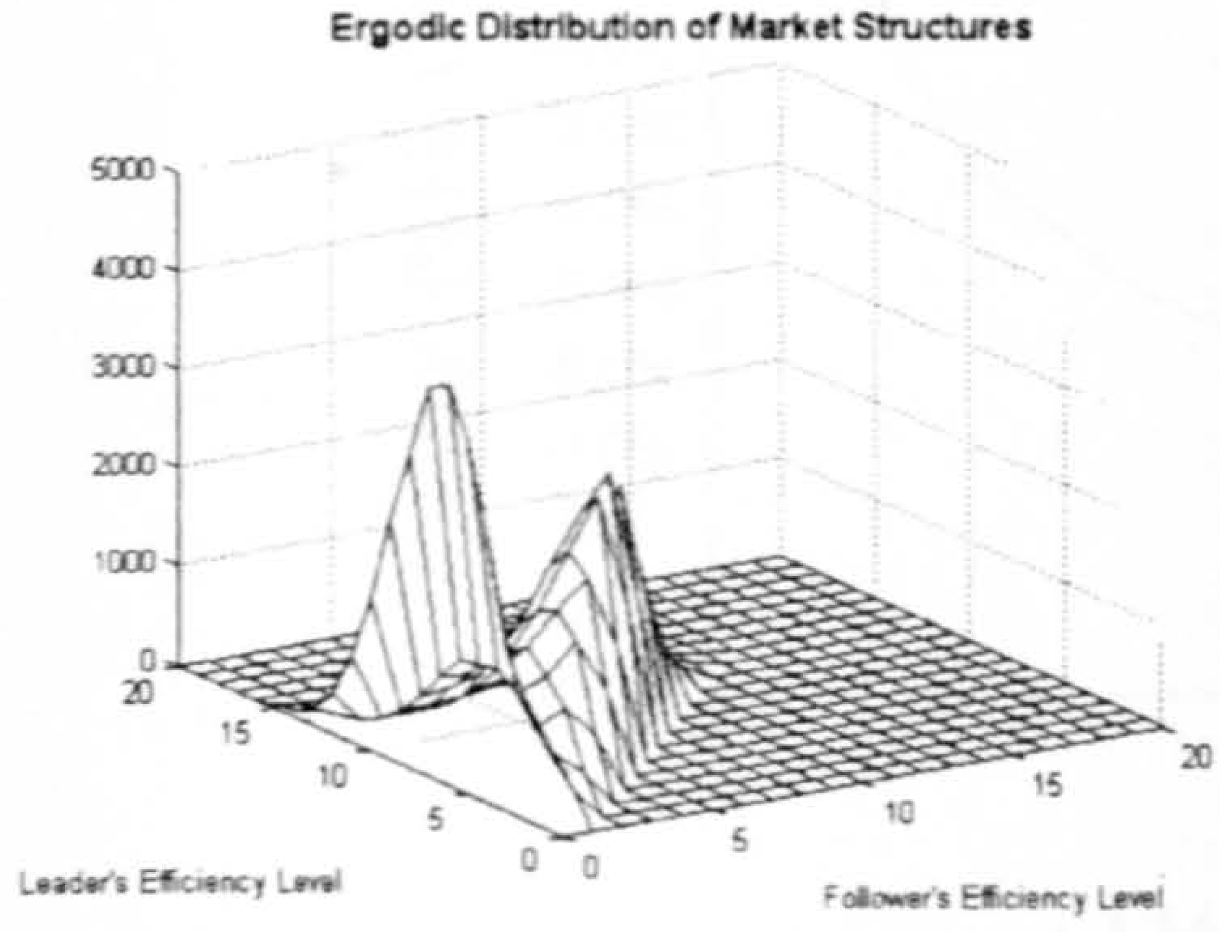
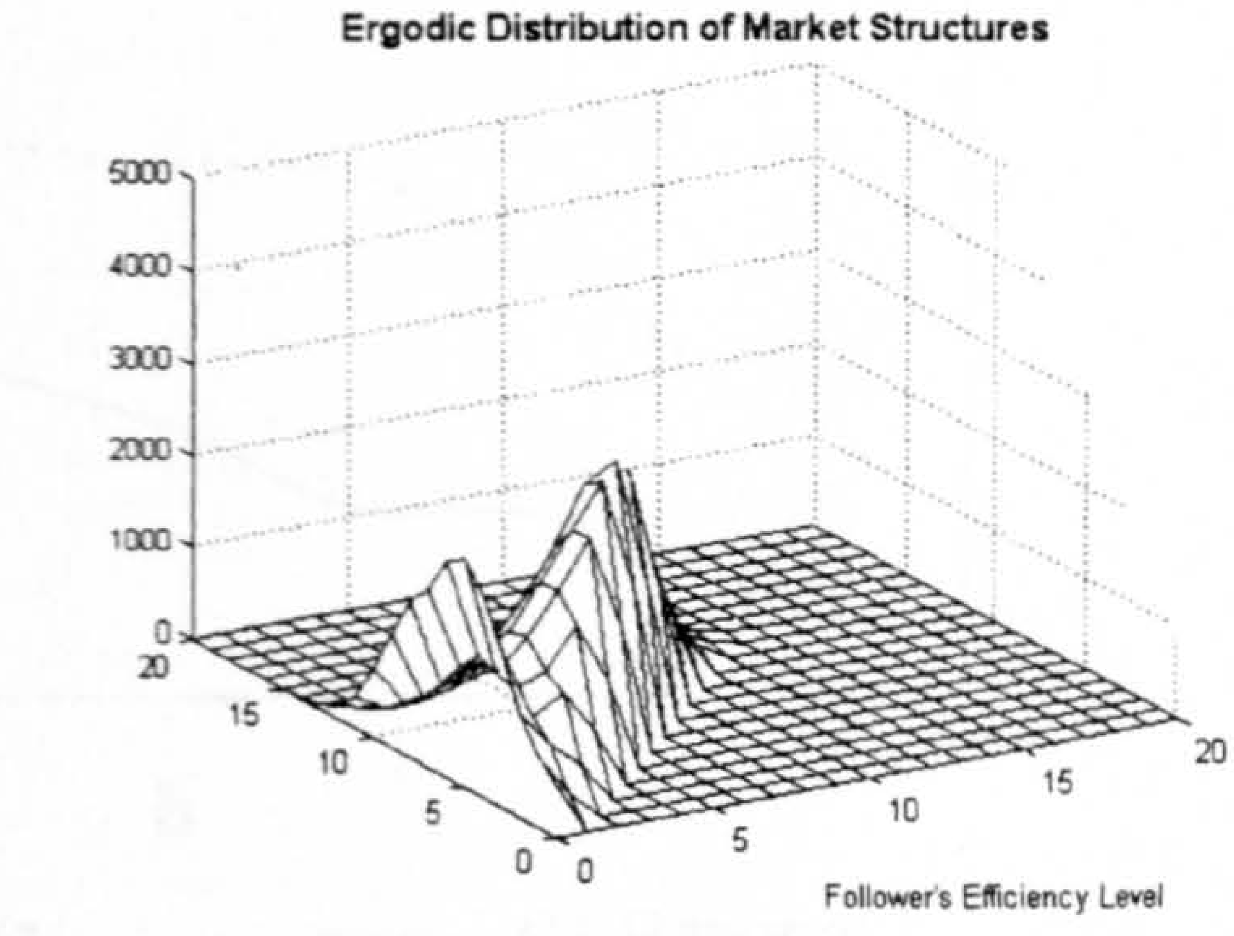


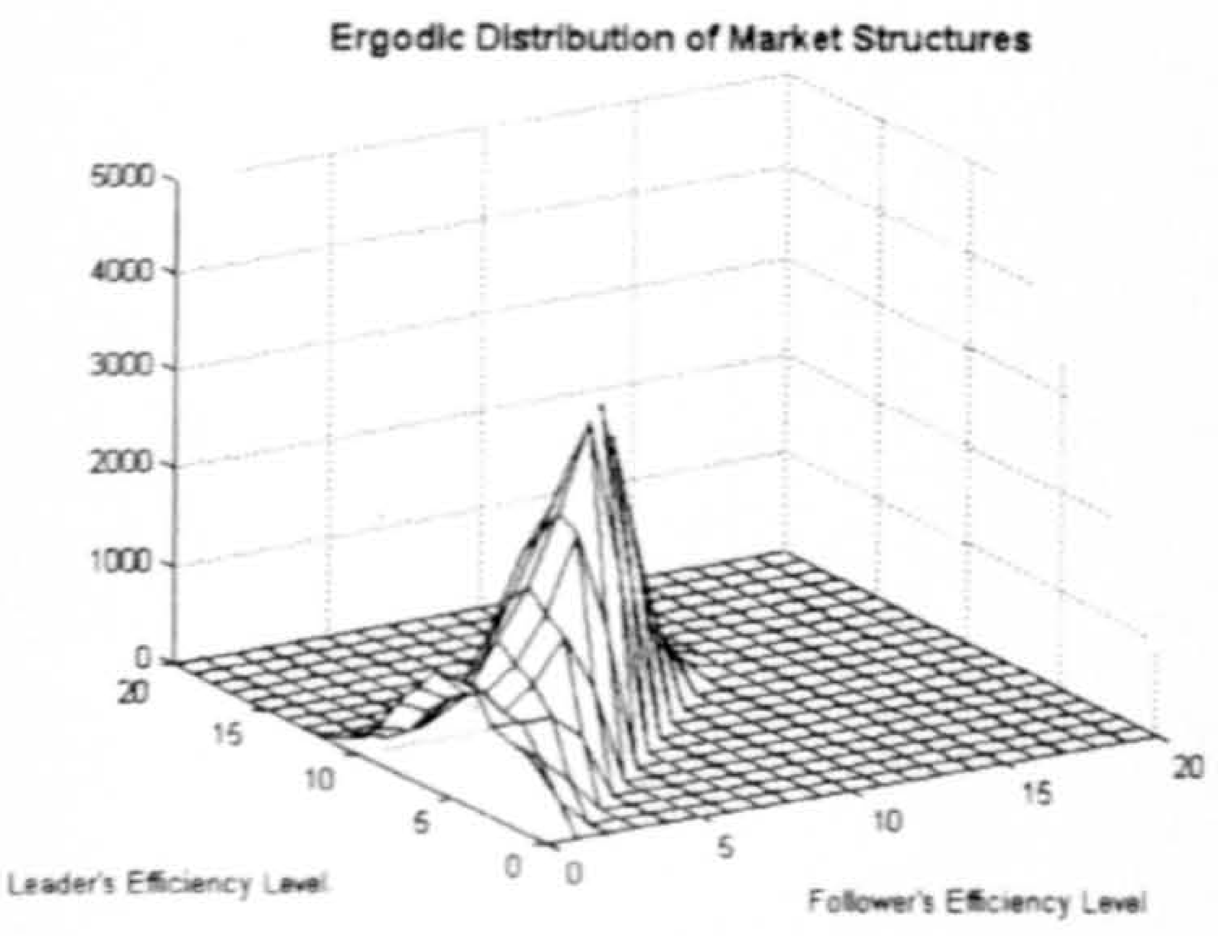
Figure 3.14: The Impact of Spillovers on the Ergodic Distribution of Market Structures in the Case of Absorptive Capacity



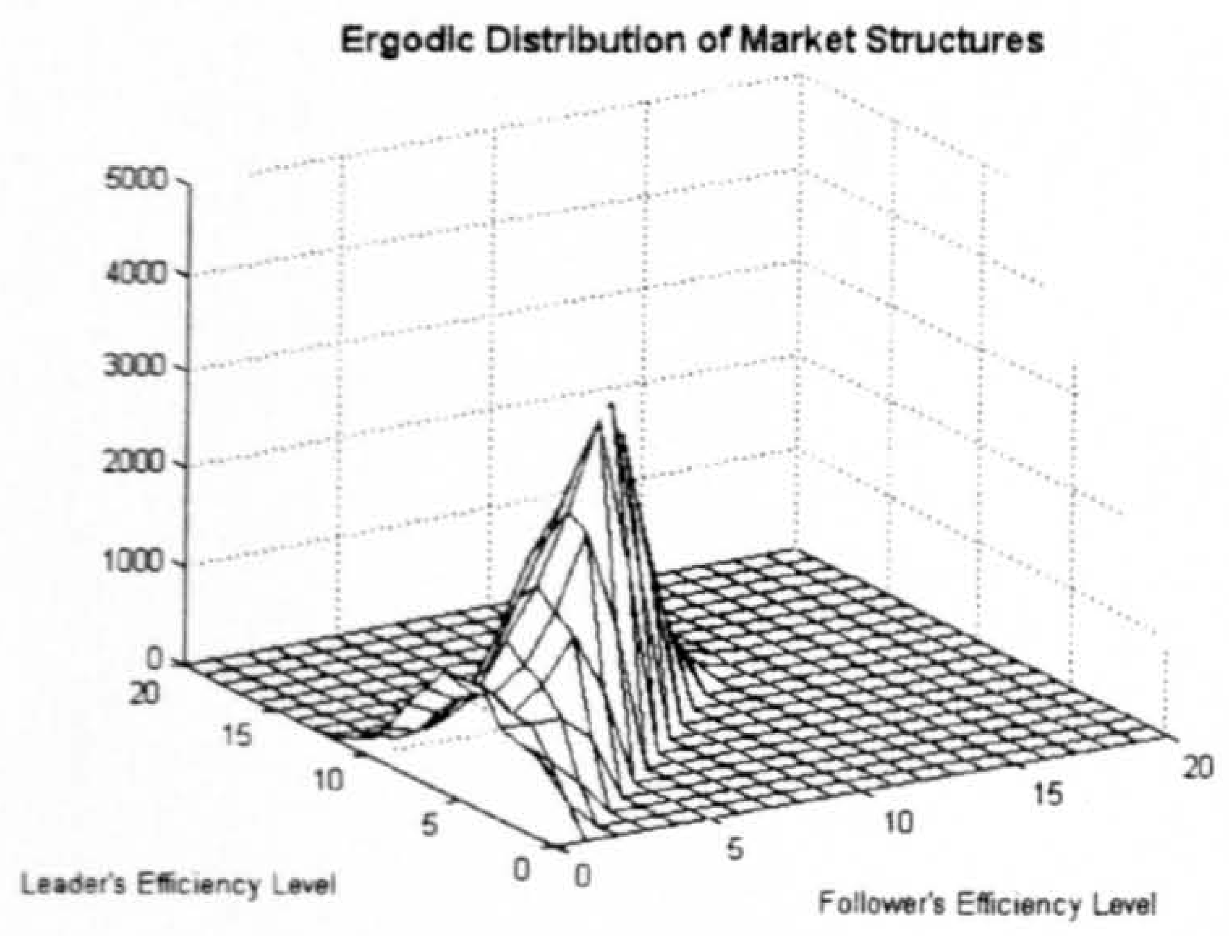
b=0.0



b=0.2



b=0.5



b=0.8

Figure 3.15: The Impact of Spillovers on the Concentration Index C1 in the Case of Absorptive Capacity

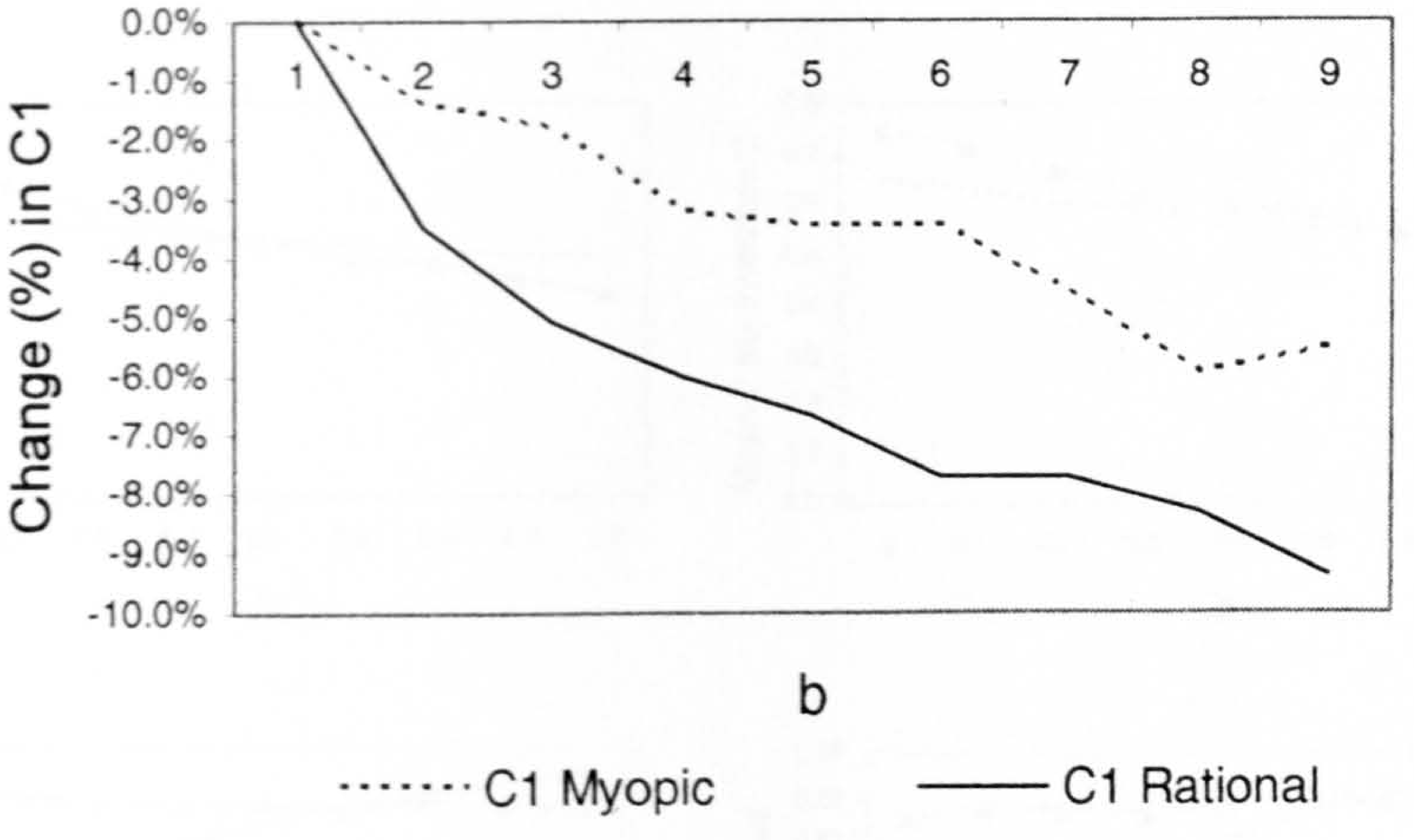


Figure 3.16: The Impact of Spillovers on Investment, Probability of Success and Firm's Value in the Case of Absorptive Capacity

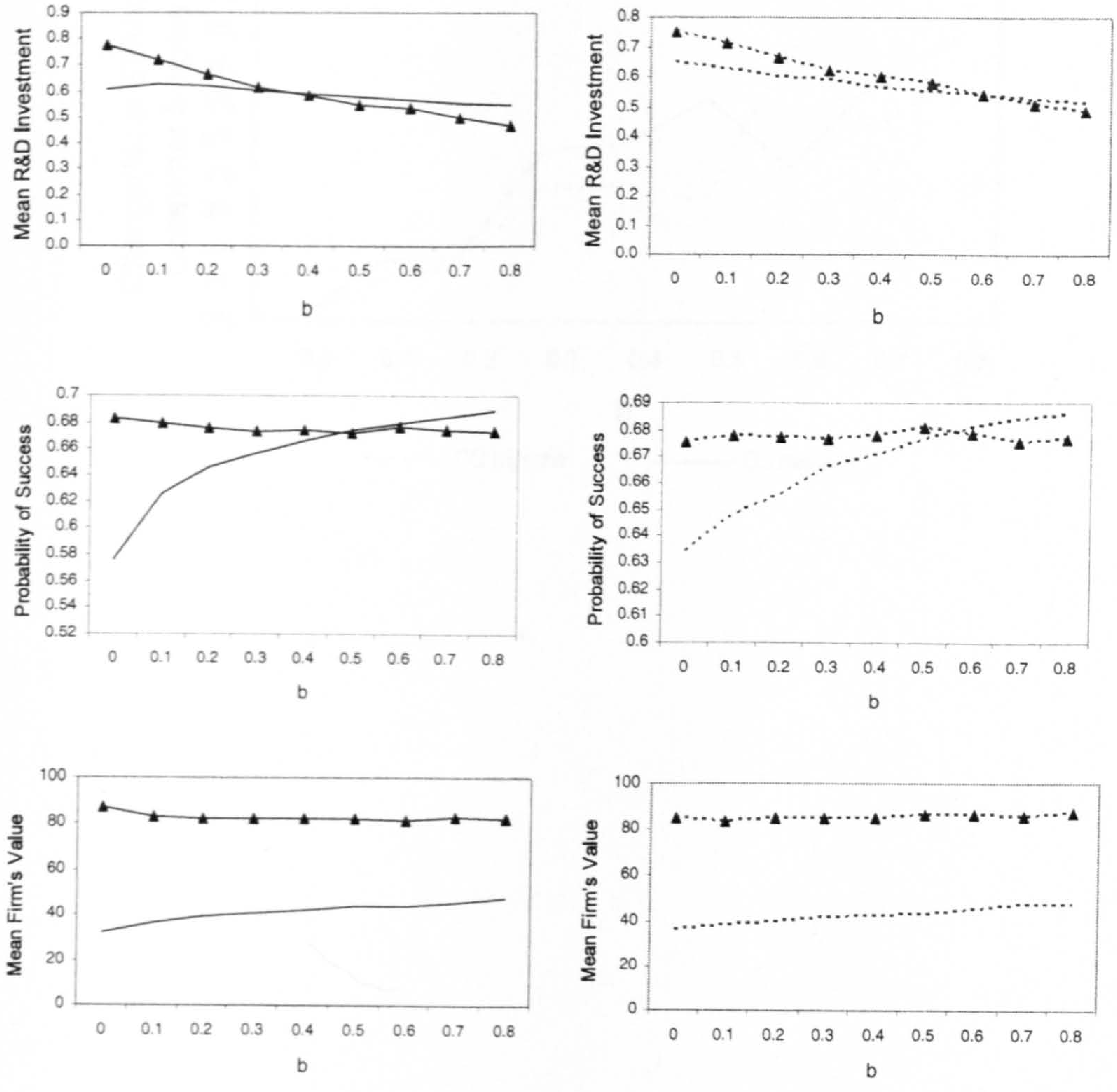


Figure 3.17: The Impact of Spillovers on Consumers' Surplus in the Case of Absorptive Capacity

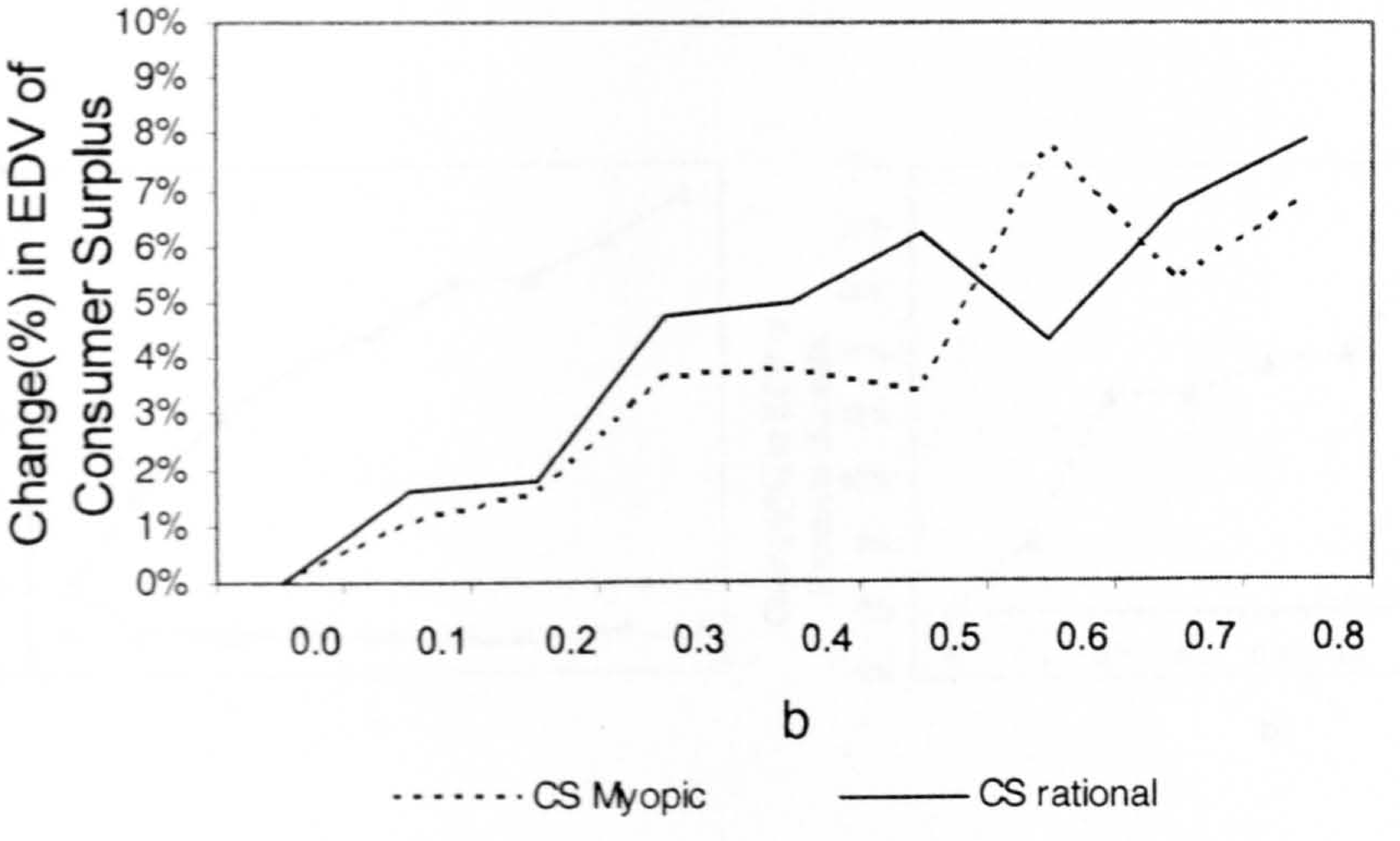
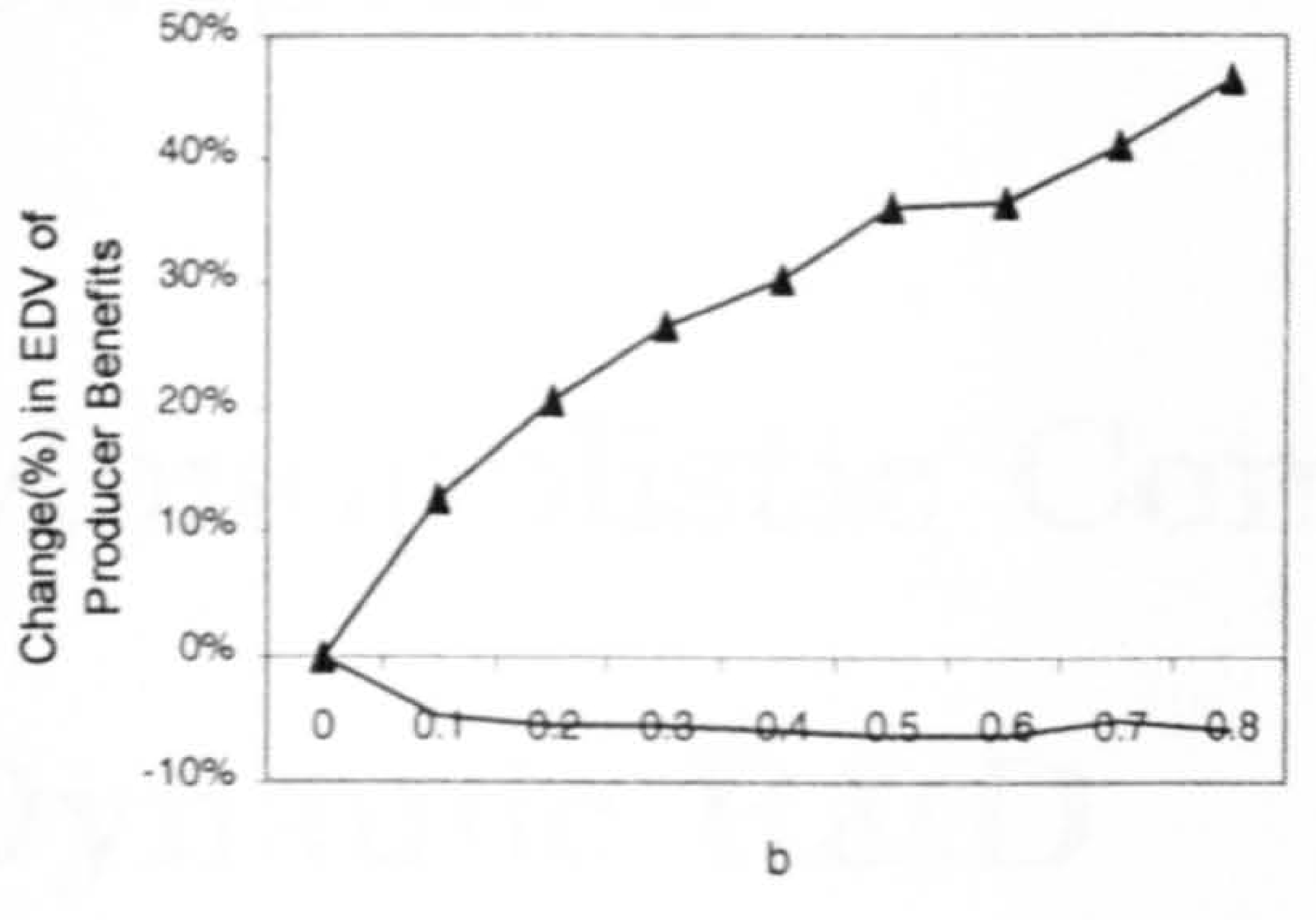
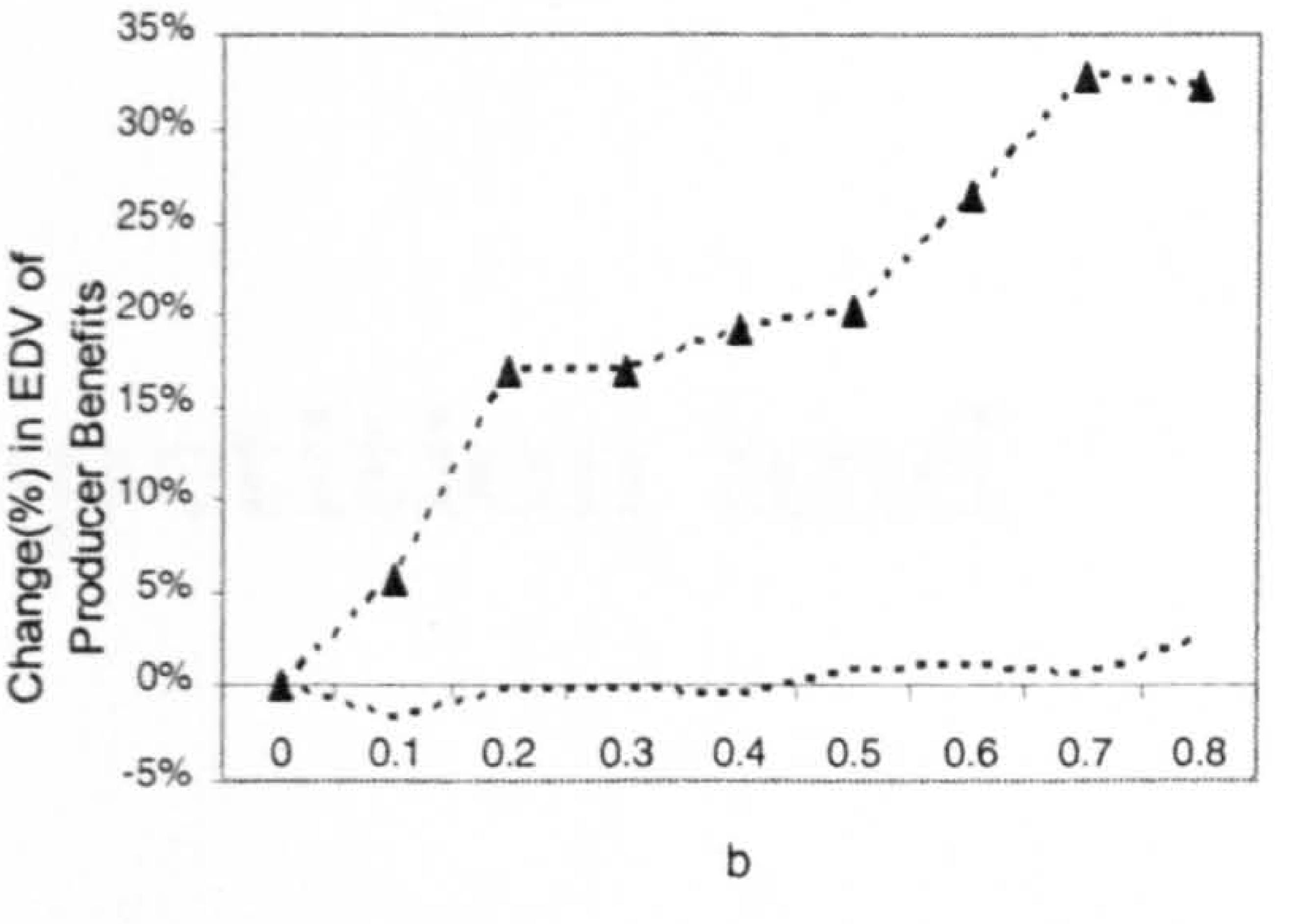


Figure 3.18: The Impact of Spillovers on Producer Benefits in the Case of Absorptive Capacity





Chapter 4

Monopolistic Competition and Dynamic R&D

4.1 Aims and Motivation

We propose an imperfect competition model *a la* Dixit and Stiglitz, with many firms, instead of the oligopolistic framework proposed in E-P. In fact, there are very few examples of rational expectations, multi-agent, stochastic models, and most methods for computing stationary equilibrium in these models have been doomed at facing dimensionality problems, both in terms of memory and time requirements, that restrict their applicability. In particular, the computation of equilibrium in discrete-time stochastic games, such as that proposed by Ericson-Pakes, can be solved only for an industry with a few firms, since increasing the number of players would increase exponentially the possibilities for the future states of the game, such that computing the expectations over future outcomes becomes infeasible.

Many attempts where made to overcome this limitation, such as Dorazelski and Judd (2003) continuous-time stochastic version of their dynamic game and Pakes and MacGuire (2001) method of calculating the expectations over the states on the ergodic path only, ignoring the rest of the off path possible market structures. As Ericson and Pakes (1995) highlight, there is a descriptive and policy need for dynamic models accounting for heterogeneity both in firms as well as in their response, and we believe that the computational burden associated with this dynamic problem needs to be overcome such that we can model variability of firms' fates for an industry with a large number of firms simultaneously active. Consequently, we propose a dynamic model analogous to that of Ericson and Pakes (1995) (E-P from henceforth) with a few important differences.

Firms are characterized by the state variable which is their efficiency level or "index of success". The degree of market competition a firm faces is determined by the market structure, and it is entirely captured on what we define as the *overall efficiency index* in the industry. The profitability of each firm depends on the ratio of the firm's own efficiency level to the overall efficiency level in the industry. We assume that firms cannot access (or rationally don't make use of[1]) all relevant information to form expectations concerning the evolution of their rivals' efficiencies. However, they are able to form rational expectations concerning the long run efficiency index in the industry, which is enough to determine their profitability in the stationary equilibrium. This is a convenient assumption that we believe to represent an improvement in terms of the realism in the representation of actual industries. Furthermore, it allows to solve the model for a large number of agents and overcome the "curse of dimensionality" highlighted by Bellman in 1961, when referring to the computational problem of the excess burden

[1]We don't specify the rationality in choosing not to use the information concerning the evolution of the relative position of each of the competitors, but it's easy to imagine situations where accessing this information might come at a cost.

associated with dynamic stochastic problems with a large number of states. By having agents taking decisions based on their rational expectations concerning the expected long run level of competition, as given by the efficiency index in the industry, we transform a multidimensional problem into a one dimension one. We propose an algorithm to find the optimal policy function for R&D investment under these assumptions. Upon simulating the model, we show that it delivers a market structure, in equilibrium, which accounts for a substantial degree of heterogeneity.

4.2 Background

Theories of endogenous growth explain economic growth through the accumulation of knowledge that derives from the R&D decisions made by profit maximising innovating firms. Yet, this decision process is operated, within these models, in an extremely stylized environment that has embodied little of the theoretical and empirical unfoldings of the Industrial Organization (I.O.) literature. In the quality ladder models of Grossman and Helpman (1991), Aghion and Howitt (1992) and Barro and Sala-i-Martin (1995), incumbent firms do not perform R&D and each innovation is introduced by a new "courageous" entrepreneur. Growth arises as a consequence of investment incentives that generate the dynamics of "creative-destruction" due to the "winner takes it all" assumption. As a result, these models deliver market structures, in equilibrium, which are implausible. For tractability reasons, symmetry is also a common assumption in endogenous growth models (Romer (1986)) that widens the gap between their micro-structure and that of the I.O. models. Modelling growth through decisions at the corporate level departing from strong assumptions on the market structure is questionable, especially given the widely accepted strong relevance of market structure to firms' decisions. As Dasgupta and Stiglitz (1980) highlight, the importance of this issue is clear cut *"because a recognition of the importance of technical progress raises serious doubts about the adequacy with which traditional micro-economic models allow one to understand the functioning of modern market economies, and develop policy prescriptions"*.

Empirical work on firm level data has uncovered the high degree of firm heterogeneity within industries as well as variability in similar firms' fates over time. This heterogeneity manifests itself in firms' market shares, their investment strategies and therefore, in the evolution of a firm's position relative to its rivals. The stylized facts of the I.O. literature concerning market structure are now well established in economic theory. One of the oldest and probably most studied empirical regularity in the I.O. literature is the fact that most industries depict a highly skewed asymmetric distribution of firm sizes, and that the Gibrat's Law, according to which a firm's growth rate is independent of its size, is a good, although imperfect approximation to firm growth (Mansfield (1962), Hall (1987) and Evans (1987)). However, smaller firms appear to experience more volatile growth rates than large firms (Dunne and Hughes (1994) and Hart and Oulton (1996)). These were the regularities uncovered by a large stream of empirical literature concerning the scaling relationship, among which we highlight only a few.

Up to the moment, there are only a handful of frameworks that explain growth through the modelling of more complete micro-market structures. Thompson (2001) proposes a R&D based growth model with product differentiation and stochastic quality growth as a way to explore the micro-structure of endogenous growth models. In his model, however, each new innovation makes the previous one obsolete just as in the creative-destruction models. Klette and Griliches (2000) propose a theoretical model that accounts for firm heterogeneity and where the evolution of the market structure and growth is driven by quality improvements that result from the stochastic outcome of the R&D activity. This model takes inspiration on the patent race literature and the firms' main competitive focus is to prevent the entry of any other firm in their product line. In the model by Klete and Kortum (2002), each firm produces a certain number of differentiated products and can experience size growth by engaging in R&D and adding further new products to its portfolio. Aghion et al (2001) analyse the effect of product market competition on growth in the context of a duopoly model where firms invest in R&D to improve the attractiveness of their product. Finally, Laincz (2005) proposes an endogenous growth model with a dynamic market structure and entry and exit analogous to that of E-P (1995) and this model is shown to deliver results which are consistent with the results of the empirical I.O. literature (see Laincz and Rodrigues (2006)). However, this model shares with the Ericson and Pakes methodology the drawback of dynamic, discrete stochastic games, the problem of the "curse of dimensionality", that arises from the fact that the burden of computing players expectations over all possible future states increases exponentially in the number of state variables.

The model we propose here is a dynamic industry model with many, heterogeneous agents with idiosyncratic uncertainty. Firms allocate resources to R&D aiming at developing new, improved production processes and reducing their marginal cost of production. This assumption is in line with the increased attention devoted to corporate R&D as a source of growth in the past decades, as invention and innovation by manufacturing firms has arisen as the main source of R&D in the economy. A firm's returns to R&D are determined not only by the firm's performance but also by the outcome of the research effort of its rivals. Growth is driven by a spillover effect of the innovations developed by the firms in the industry. When firms are successful in developing a technical improvement, it is assumed that given the limitations of intellectual property rights protection and secrecy, this innovation will contribute to the increase of the public stock of knowledge and consequently improve the marginal cost conditions for all firms in the industry. The evolution of the size of each firm is determined endogenously according to the stochastic outcome of the R&D process and an exogenous stochastic process of depreciation. We aim at obtaining a stationary equilibrium, where growth is driven by a process of knowledge diffusion, and which delivers a firm size distribution with a substantial degree of heterogeneity. This distribution of firm sizes or market structure is the endogenous outcome of an R&D investment process in which firms engage in an attempt to enhance their likelihood of developing innovations. Firms' decisions are driven by the ambition of moving forward in terms of their state variable in order to ultimately exploit profitable opportunities.

Therefore, we represent growth, at the industry level, through a rich representation of the microfoundations of firms' decisions, inspired in the set up of Ericson and Pakes (1995), but allowing for heterogeneity among a large number of firms operating in the same industry. However, in order to better to approximate better the statistical properties of the real firm size distribution the degree of heteregeneity should be more skewed. Therefore, avenue for future work would be to introduce entry and exit of firms in the model. In fact, turbulence in the market as given by entry and exit would increase the mass of small firms. At each period of time, there would be a number of firms that recently entered the industry which would still be struggling to gain market share, and firms which would be getting small and closer to their exit threshold value. Also, increasing the variability in firm's fates in the model would improve the match of the distribution of firm sizes. We believe that introducing a further element of heterogeneity in the model would allow approximating better the higher moments of the actual firm size distribution found for real data. An example of such elements are randomly drawn set up costs and scrap values, that would vary the exit and entry threshold among firms, and assuming that the degree of risk aversion varies across agents and that the private discount factors are randomly drawn from a common distribution. The latter would vary firm's decisions reflecting different degrees of "impatience". These two assumptions would increase the variability in firm's outcomes and consequently the skewness of the firm size distribution for our simulated data.

4.3 The Model

In this section we propose an endogenous growth model with a dynamic market structure. We model an industry with many firms, each producing a slightly differentiated product. Consumers have *CES* preferences over a set of monopolistic commodities. We allow firms to differ in their technologies of production, such that marginal costs are asymmetric amongst firms. Ceteris paribus, each firm's profits are increasing in the price of rival commodities and decreasing in the price of the good produced by the firm. In an attempt to increase their market share, firms invest in cost-reducing R&D to improve the likelihood of experiencing a marginal cost reduction that will allow them to decrease the price of their good relative to the price of the other goods in the industry. Investment decisions are intertemporal, and the firm will choose the level of R&D expenditure that maximizes its expected future stream of profits. As opposed to the original dynamic industry framework proposed by E-P, we assume that rather than assigning probability weights to all possible future market structures, firms form rational expectations on the long-run overall efficiency index in the industry, and will compute their expected future stream of profits assuming that this long-run level of efficiency index in the industry is constant. This allows us to decrease tremendously the computational burden of this dynamic game and, consequently, solving the model for environments where a large number of firms are operating simultaneously in the industry.

Growth, in this model, arises from the increases on the public stock of knowledge which are a stochastic consequence of firms' innovative activity, as in Laincz (2004). Increases in the public stock of knowledge represent positive aggregate productivity shocks in the marginal costs of all firms in the industry. Given the specification of the model presented here, a proportionate improvement in the marginal cost conditions of all firms in the industry will not alter the profitability of firms as every firm's ratio of price index to own price will remain unchanged, and therefore so will the R&D program. Only changes in firms' efficiency levels which cause non-proportionate changes in the vector of marginal costs of the active firms in the industry will alter firms' profitability. However, as a result of the aggregate productivity shocks, the total output produced within the industry will increase giving rise to continuous industry growth.

4.3.1 The Spot Market

We model a monopolistically competitive industry composed by n firms producing a set of commodities $1, 2, 3...n$. We abstract from entry and selection process in this model, so the number of firms, and therefore the number of commodities produced in the industry, is fixed. We assume consumers have Cobb Douglas preferences over the output of the monopolistically competitive sector, denoted by y; and a good zero which represents the numeraire and aggregates the rest of the economy.

$$u = U(x_0, y) = x_0^{1-\alpha} y^{\alpha} \tag{4.1}$$

Each of the monopolistic goods is an imperfect substitute of the other $n-1$ differentiated goods produced in the industry, such that consumer preferences over these goods can be represented by a CES sub-utility function.

$$u = U\left(x_0, \left\{\sum_i^n x_i^{\frac{1+\varepsilon}{\varepsilon}}\right\}^{\frac{\varepsilon}{1+\varepsilon}}\right) \tag{4.2}$$

U is an homogeneous, concave function, increasing in the consumption x of good i and $\varepsilon < -1$ is a parameter representing the price elasticity of demand for good i. We can write down the budget constraint of the consumer as:

$$x_0 + \sum_i^n x_i p_i = I \tag{4.3}$$

where p_i is the price paid and x_i is the quantity consumed of good i, respectively, and I is the income of the consumer in terms of the numeraire. We denote the amount of income or total expenditure devoted to the n monopolistic commodities by Y:

$$Y = \sum_i^n x_i p_i \tag{4.4}$$

The overall price index and the quantity index in the industry are q and y, respectively:

$$q = \left\{ \sum_{i}^{n} p_i^{1+\epsilon} \right\}^{\frac{1}{1+\epsilon}} \tag{4.5}$$

$$y = \left\{ \sum_{i}^{n} x_i^{\frac{1+\epsilon}{\epsilon}} \right\}^{\frac{\epsilon}{1+\epsilon}} . \tag{4.6}$$

The Cobb Douglas specification for preferences over y and the rest of the economy ensures a constant budget share allocated to the monopolistically competitive sector [2], such that:

$$y = \frac{Y}{q} = \frac{\alpha I}{q}$$

We further define the overall marginal cost index in the industry, which we denote by m. Letting mc_i stand for the marginal cost of production of firm i:

$$m = \left\{ \sum_{j}^{n} mc_j^{1+\epsilon} \right\}^{\frac{1}{1+\epsilon}} \tag{4.7}$$

The optimal level of consumption for good i is given by:

$$x_i = \frac{Y}{q} \left(\frac{q}{p_i} \right)^{-\epsilon} \tag{4.8}$$

The number of firms operating in the industry is assumed large enough to consider the effects of a firm's price choice, p_i, on the overall price index, q, as negligible and consequently, we have aggregate price taking behaviour. The spot market condition for profit maximisation is given by the *marginal revenue=marginal cost* equation and the price charged by firm i will be a constant markup σ over its marginal cost, mc_i :

[2] This assumption of a constant budget share is crucial for the solution methodology as we discuss later on in the paper.

$$p_i = \sigma mc_i = \frac{\varepsilon}{1+\varepsilon} mc_i \tag{4.9}$$

At the aggregate level, this implies that the price index can be expressed as a function of the marginal cost index: $q = \sigma m$.

If fixed costs of production are set to zero, then optimal profits can be written as:

$$\begin{aligned} \pi_i &= (p_i - mc_i)x_i \\ &= -\frac{1}{\varepsilon} Y \left(\frac{q}{p_i}\right)^{-\varepsilon-1} \\ &= -\frac{1}{\varepsilon} Y \left[\frac{m}{mc_i}\right]^{-\varepsilon-1} \end{aligned} \tag{4.10}$$

A firm's profits are increasing in the expenditure devoted to the monopolistically competitive industry and the marginal cost index in the industry, and decreasing in the firm's own marginal cost. This expression for profits is homogeneous of degree zero in the vector of marginal costs. This property, which is also present in Laincz (2005 a) and b)) endogenous growth model, implies that proportionate changes in the vector of marginal costs leave profits unaffected. What is relevant for a firm's profitability is not the absolute value of its marginal cost, but rather its relative value with respect to the marginal cost index in the industry. Hence, the lower the ratio of own marginal cost to marginal cost index, the higher the profits of the firm.

It is important to note that although profits are homogeneous of degree zero in the vector of marginal costs, the choice of quantity to produce is not:

$$x_i = \frac{Y}{q}\left(\frac{q}{p_i}\right)^{-\varepsilon} = \frac{Y}{\sigma}\frac{m^{-\varepsilon-1}}{mc_i^{-\varepsilon}} \tag{4.11}$$

The quantity produced by firm i is increasing in the overall marginal cost index, m, and decreasing in its own marginal cost. An improvement on the marginal cost conditions in the industry causing a proportionate decrease in the vector of marginal costs will increase x_i, the quantity produced, although profits remain unchanged.

4.3.2 The dynamic environment

The previous subsection showed how firms' profits are determined by their marginal cost of production relative to those of the rivals. In the framework presented here, marginal costs are a negative function of the state variable, which is the firm's efficiency level. The efficiency level of a firm is basically its index of success, and high efficiency levels imply the firm is in a favourable position relative to its rivals and thus able to obtain higher payoffs. In this model, we assume that efficiency levels affect payoffs by reducing the marginal cost of production. Firms differ in terms of their efficiency levels and therefore, they incur in different levels of marginal cost of production. Let $w_{i,t} \in \mathbb{Z}^+$ stand for the efficiency level of firm i at time t. The efficiency levels are bounded from above, such that $w_{i,t} \in \{1, 2, ...w^{\max}\}$. Dropping time subscripts for simplicity, we further define the overall efficiency index in the industry at each moment in time as w:

$$w = \left\{\sum_j^n w_j^{(1+\varepsilon)}\right\}^{\frac{1}{1+\varepsilon}} \tag{4.12}$$

The mapping between the efficiency levels and marginal costs is given by:

$$mc_{i,t} = mc_{0,t} w_{i,t}^{-\mu} \tag{4.13}$$

where μ is the rate of cost reduction and $mc_{0,t}$ is a scalar on marginal costs. The more efficient a firm is relative to its rivals, the lower the firm's marginal cost relative to the marginal cost index in the industry which implies a higher ratio of the overall price index in the industry relative to the price of the commodity produced by the firm and thus, a higher market share for that firm. This functional form for marginal costs complies with two important conditions to hold for the marginal costs, namely that there is a smooth relationship between efficiency levels and marginal costs (see figure 4.1 at the end of the chapter), and that profits are bounded from below and above. Given the setup used here, if the functional form for the mapping between the efficiency levels and the marginal costs was the one used in Pakes and McGuire (1995), i.e. $mc_{i,t} = mc_{0,t} e^{-\mu w_{i,t}}$, profits would increase exponentially in the efficiency level. The upper boundary condition on profits would not be satisfied and firms would invest more the higher their efficiency level implying that the state space for each firm would not be finite and the problem would have no solution.

The marginal cost index can be written as a function of firms' efficiency levels:

$$m = mc_0 \left\{ \sum_j^n w_j^{-\mu(1+\epsilon)} \right\}^{\frac{1}{1+\epsilon}} = mc_0 f(w_t) \quad \text{with} \quad f'(w_t) < 0. \tag{4.14}$$

Equation (4.13) implies that the marginal cost index in the industry decreases with the overall efficiency index. We further assume that mc_0 decreases exponentially with the total number of innovations developed in the industry:

$$mc_{0,t} = e^{-\gamma . PS_t}. \tag{4.15}$$

PS_t is the total stock of innovations developed in the industry until time t, and γ captures the cost reducing impact of the public knowledge in the industry. Increases in PS scale down the marginal cost curve for all firms in the industry. Figure 4.1 below depicts the marginal cost curve as a function of the efficiency level associated with two different public stocks of knowledge, namely when there were 100 past successful innovations in the industry, and when the public stock is composed by 300 innovations. In calibrating the marginal cost curves depicted in the figure we use the same parameters that we later use for the simulations, i.e., we set $\gamma = 0.002$ and $\mu = 0.3$.

The market structure of the industry, at a given moment in time, is given by a vector indicating the efficiency levels of the firms in the industry. These efficiency levels evolve, over time, as a result of the stochastic outcomes of two processes: the investment process undertaken by firms and the outcome of the idiosyncratic shocks that deteriorate the firm's state. Given the efficiency level of firm i at time t, the realizations of the two stochastic outcomes will determine efficiency levels at time $t + 1$. Letting primes refer to the values of the variables next period, then:

$$w_i' = w_i + \nu_i - \xi_i \qquad \text{for } \ i = 1, 2. \tag{4.16}$$

The random variable ξ_i captures the outcome of the depreciation process that affects negatively a firm's efficiency level, leading to loss of competitiveness for the firm and hence a deterioration of profits. The realizations of this random variable are independent across firms such that for each firm i there is a positive probability

δ that ξ_i takes value 1, and a probability 1-δ that ξ_i is zero[3]. Positive realizations of ξ_i increase the marginal cost of the firm by exactly the same amount of the decrease in marginal costs driven by a successful innovation. The second process is the firms' innovative activity, whose outcome is captured by the random variable ν_i, which is also firm specific. Firms invest in R&D in an attempt to enhance the likelihood of being successful in innovating and consequently decreasing their marginal costs of production.

The outcome of firm i's innovative activity, captured by the random variable ν_i, is given by:

$$\nu_i \begin{cases} = 1 \text{ with probability } p(\nu_i) \\ = 0 \text{ with probability } 1 - p(\nu_i) \end{cases}$$

where

$$p(\nu_i) = \frac{az_i}{1 + az_i}. \tag{4.17}$$

In the above set of equations, z_i stands for the level of investment undertaken by firm i and a is a parameter denoting the productivity of R&D investment in increasing the likelihood of developing an innovation. This is basically the technological opportunity in the industry, i.e., the difficulty of innovating in the industry, and it is related to the stage of development of scientific knowledge and other knowledge specific characteristics.

[3]In E-P the negative shocks to firms' efficiency levels are industry wide. Here, in order to allow for more variability in firm sizes, and also because we find it more adequate in terms of realism, we model these shocks as idiosyncratic. Furthermore, an aggregate shock to the efficiency levels would be incompatible with a constant long run market structure.

4.3.3 The R&D Choice

In each period of time, firms choose the level of R&D expenditures that maximises the expected present discounted value of their future stream of profits. In the E-P formulation, this optimal choice requires the perception of rivals' future states. They propose a numerical algorithm, for which a code is made available, which solves this dynamic problem. The simulation of the market structure, given the optimal policy functions, delivers an ergodic distribution of firms' states. In the model presented here, we assume that in making their intertemporal choices concerning R&D investment, firms do not assess the probability of all possible future market structures, they rather form an expectation on the long-run efficiency index in the industry, and optimise accordingly. When a large enough number of firms are present in the industry, then the industry's efficiency index will converge to a constant value, $\bar{w}$:

$$\text{Let } \tau > 0. \ \exists\, n_0, \text{ s.t. } n > n_0 : \lim_{t \longrightarrow \infty} (w_t - \bar{w}) < \tau$$

Firms have rational expectations concerning this long-run value of the efficiency index in the industry, and their profitability will be fully determined by the evolution of their own efficiency level and the expected value for the long-run industry state, which they treat as constant.

The firm's intertemporal problem will then be written as a function of this expected value:

$$V_i(w_i, w, \bar{w}) = \max_{z_i \geq 0} \{\pi(w_i, w_t) - cz_i + \beta E_t V_i(w_i', \bar{w})\}, \tag{4.18}$$

where β is the discount factor, common to all firms and c is the unit cost of

R&D investment. The Bellman equation above simply states that a firm's value at each moment in time is given by its current profits minus its expenditures in R&D investment plus the discounted expected value next period. If we denote by $pr(w_i' \mid z_i, w_i)$ the probability that firm i faces of moving from efficiency level w_i to w_i', conditional on firm i's level of investment, z_i, we can write (4.18) as:

$$V_i(w_i, w, \bar{w}) = \max_{z_i \geqslant 0}\{\pi(w_i, w_t) - cz_i + \beta V_i(w_i', \bar{w})pr(w_i' \mid z_i, w_i)\} \qquad (4.19)$$

Since firms take the long-run efficiency level as constant, and form rational expectations on this value, the likelihood that the firm's present discounted value next period is $V(w_i', \bar{w})$ depends only on the evolution of its own efficiency level. In the original Markov-Perfect Dynamic Industry model proposed by E-P, the firm would also have to work out the probability of facing each possible market structure. Having firms' decision process taking into account only the long-run industry's efficiency index as proposed here greatly simplifies the method for finding the optimal policy functions as there is only one dimension to this problem, which is simply the evolution of the state variable of the optimising firm.

Let $C_1 = V_i(w_i' + 1, \bar{w})$ denote the expected value of the firm conditional on success in innovating, and $C_2 = V_i(w_i', \bar{w})$ the expected value in the case it fails to develop an innovation. Then one can rewrite the equation above as:

$$V_i(w_i, w, \bar{w}) = \max_{z_i \geqslant 0}\left\{ \begin{array}{c} \pi(w_i, w_t) - cz_i + \\ \beta\left[\frac{az_i}{1+az_i}V_i(w_i' + 1, \bar{w}) + \frac{1}{1+az_i}V_i(w_i', \bar{w})\right] \end{array} \right\}. \qquad (4.20)$$

The first order condition for this dynamic Bellman equation is given by:

$$\frac{\partial V}{\partial z_i} = -c + \beta \left[\frac{a}{(1+az_i)^2} (C_1 - C_2) \right] = 0$$

which yields the following optimal policy function for R&D investment:

$$z_i^* = \max \left\{ \begin{array}{c} 0, \\ z_i^* = \frac{1}{a}\sqrt{\frac{a\beta(C_1-C_2)}{c}} - \frac{1}{a} \end{array} \right\} \tag{4.21}$$

The expression above shows that a firm's optimal choice of R&D investment is decreasing in the unit cost of investment ($\partial z_i^*/\partial c < 0$) and increasing in the discounted value gain from innovating ($\partial z_i^*/\partial (C_1 - C_2) > 0$). An increase in the discount factor β increases the level of investment because the firm attaches higher weights to future payoffs ($\partial z_i^*/\partial \beta > 0$). The effect of an increase in the productivity of investment, a, in the optimal level of investment is ambiguous because higher values of a imply more R&D ascribed to the innovative activity, but the higher a the more quickly decreasing returns settle down.

In order for the upper limit on the efficiency levels not to drive the results from the optimization process, we ensure that a firm at the highest efficiency level does not have an incentive to invest and therefore $x_i^*(w = w^{\max}) = 0$. Furthermore, it is shown in the appendix to this chapter that the convexity of the profit function in (4.10) holds if $-\mu(1+\varepsilon) < 1$.

We constructed an algorithm to solve for the equilibrium policy functions and a code for simulating the model. The algorithm is described in section 4.3 ahead.

4.3.4 Endogenous Growth

As in the New Growth theory models, the source of growth in this model are the externalities that arise from the innovative activity in which firms engage. In this framework, each new innovation developed enhances the knowledge base of the industry. The public stock of knowledge at a given moment in time will then be given by the sum of all the innovations generated until then by the set of firms in the industry. Equation (4.22) below describes the change in the public stock of knowledge from one period to the next:

$$PS_{t+1} = PS_t + \sum_{i=1}^{n} \nu_{i,t}, \tag{4.22}$$

where $\sum_{i=1}^{n} \nu_{i,t}$ is the sum of all positive realizations of the firm specific random variables capturing the outcome of the investment process. It is this process of knowledge spillovers that will allow for marginal costs to continuously decrease in the industry. The motion of the public stock of knowledge given by equation (4.22) will generate a continuous improvement through time of the marginal cost conditions in the industry as mc_0 decreases with the cumulative effect of innovations. However, changes in mc_0 imply proportionate changes in the vector of marginal costs in the industry and, as in implied by equation (4.10) above, proportionate changes in the vector of marginal costs do not affect firms' profitability:

$$\begin{aligned} \pi_i &= -\frac{1}{\varepsilon} Y \left[\frac{m}{mc_i} \right]^{-\varepsilon-1} \\ &= -\frac{1}{\varepsilon} Y \left[\sum_j^n \left[\frac{w_j}{w_i} \right]^{-\mu(1+\varepsilon)} \right]^{-1} \end{aligned}$$

The set of equations above implies that profits remain unchanged by the increasing motion of the public stock of knowledge, and consequently so will the optimal R&D program. If there are no changes in the optimal investment decisions, the evolution of the market structure will not be affected and the long-run efficiency index in the industry will be constant in spite of the continuous amelioration of the marginal cost conditions. This property of the homogeneity of degree zero of profits in the vector of marginal costs is crucial for solving for the R&D program under the assumption that firms form rational expectations on the long-run index in the industry and treat it as constant. If this was not the case, then the long-run efficiency index in the industry would not be constant. Rational firm behaviour implies that firms would take into account the impact of the motion of the public knowledge base in their R&D investment decisions.

When finding the equilibrium policy functions for this problem, we can simply ignore the motion of PS. However, quantity choices at each period of time will be affected by the aggregate productivity motion since the homogeneity of degree zero does not hold for quantities:

$$\begin{aligned} x_i &= \frac{Y}{\sigma} \frac{m^{-\varepsilon-1}}{mc_i^{-\varepsilon}} \\ &= \frac{Y}{\sigma} \frac{\left[mc_0 \left\{ \sum_j^n w_j^{-\mu(1+\varepsilon)} \right\}^{\frac{1}{1+\varepsilon}} \right]^{-\varepsilon-1}}{\left[mc_0 w_i^{-\mu} \right]^{-\varepsilon}} \\ &= \frac{Y}{\sigma} \frac{f(w)^{-\varepsilon-1}}{mc_0 w_i^{\varepsilon\mu}} \end{aligned} \tag{4.23}$$

It is straightforward from the equation above that the quantity produced by each firm will increase as mc_0 decreases with the increasing motion of the public stock of knowledge. The negative idiosyncratic shocks to the efficiency level of firms will maintain the competitive pressure on them to pursue R&D investment, and thus the number of innovations developed in the industry will be ever increasing. As a result, the overall quantity produced in the industry will continuously increase as the vector of marginal costs decreases through time. This process will generate continuous growth in the total quantity produced within the industry.

The assumptions of no substitution between the monopolistic goods and the rest of the economy, together with the property of homogeneity of degree zero of the profit function with respect to the vector of marginal costs in the industry are essential for the validity of the methodology used to solve this dynamic, stochastic problem. The endogenous growth process in the industry leads to ever decreasing marginal costs in the industry. If profits were sensitive to proportionate changes in the vector of marginal costs, the optimal R&D investment policy would be changing over time. Furthermore, as marginal costs decrease, so will the overall price index in the economy, q. If we were to allow for the substitution between the monopolistic commodities and the rest of the economy, the budget share allocated to the former would increase. Allowing for Y to be negatively related to q would imply a continuous increase in the budget share allocated to monopolistic competitive sector, due to the decrease in the price index for the sector relative to the rest of the economy. A quick look at the expression for profits (equation 4.10) shows that the budget allocated to the monopolistically competitive sector affects firms' profits positively. Thus, the level of profits for a given efficiency level would be increasing over time. Dropping either of these assumptions would imply that profits would change as a consequence of the continuous improvement in the marginal cost conditions, and hence so would the optimal investment strategies, preventing us from being able to find an optimal solution for the problem at question.

In the framework presented here there is no mechanism that can generate the empirically observed correlation in firm's fates. Having the budget share allocated to the monopolistically competitive sector increasing as the marginal cost index in the industry decreases over time, as described above, would generate an elegant source of correlation in firms' profits, but this assumption would render the solution methodology inapplicable. Still, there is some degree of correlation in the quantities produced which increase as a result of the advances in the public stock of knowledge.

4.4 Solution Methodology

In order to find the Markov equilibrium of the model previously described, we design an algorithm to find the optimal R&D investment choices and correspondent firm value for every state in which a firm can be. This is possible because there are a finite number of these states as there is an (endogenous) bound on the maximum efficiency level a firm can achieve in the industry. The algorithm is composed of two main loops: a major loop on the expected value for the overall efficiency level in the industry; and a smaller loop, inside the major one, iterating on the value function to determine the optimal policy function for a given expected long-run efficiency index in the industry.

Once the primitives of the model were inserted into the algorithm, namely the fixed number of firms, the maximum efficiency level and the parameter choices, the algorithm starts with an arbitrary guess for the long-run overall efficiency index in the industry, say w^{g_1}. Then, the vector for the profits $\pi(w_i, w^{g_1})$ associated with each efficiency level $w_i \in \{1, 2, ...w^{\max}\}$, is calculated . Given the profit vector and the parameters choices, the algorithm finds the optimal level of R&D investment for each efficiency level in which a firm can find itself, i.e., the vector of optimal investments that solves the bellman equation in (4.18), given the expected long run efficiency index $\bar{w}$.

The algorithm then involves the simulation of the market structure for the number of periods of time chosen, using the optimal policy function that emerged from the value function iteration. The market structure at the last period of time is registered and the overall efficiency level associated with this distribution of efficiency levels or market structure is computed. So as to impose rational expectations equilibrium, i.e., that equilibrium is reached when the outcome of firms' strategies is precisely equal to the expected outcome used to compute these optimal strategies, we compare the overall efficiency level generated from the market structure simulation and the expected value of the overall efficiency level used in the last iteration to compute the optimal R&D investment choices. If the difference between the two is lower than some tolerance level chosen by the researcher, then we found the optimal policy function for a given parameterisation. Otherwise, a new guess for the long-run efficiency index is calculated according to a smoothing function and the algorithm repeats the previous steps. In what follows, we present the structure of the algorithm in a more schematic way.

Start with an arbitrary choice for $\bar{w}$, say $w^g = w^{g_1}$.

Step 1: Computes $\pi(w_i, w^g)$ for all $i = \{1, 2, ...w^{\max}\}$.

Step 2: Set $j = 0$

2.1 : Value Function iteration; iteration$= j$

2.2 : Starts with an arbitrary guess for the 1^{st} iteration $V(w_i, w^{g_1})_{j=0}$, $z_i(w_i, w^g)$ and

$pr(w_i + 1|\, z_i, w_i)_{j=0}$ (e.g, set them all to zero),

2.3 : Performs one iteration for the Bellman equation in (4.18) for all $i = \{1, 2, ...w^{\max}\}$.

2.4 : If $V(w_i, w^g)_j - V(w_i, w^g)_{j-1}$ <tolerance level, go to step 3;

Otherwise $j = j + 1$ and step 2.2 is repeated.

Step 3: Starts with an initial guess for the market structure, and uses the optimal $z_i(w_i, w^g)$ computed according to equation (4.21) to simulate the market structure until the maximum number of periods of time chosen by the researcher is reached.

Step 4: w is computed according to equation (4.12).

If $w - w^g$ <tolerance level, stop;

Otherwise, $w^g = \frac{1}{\lambda}w + \frac{\lambda-1}{\lambda}w^g$ and the procedure is repeated starting again in step 1.

4.5 Firm Size Distribution

In this section, we analyse the characteristics of the firm size distribution that is generated by the dynamic, stochastic model proposed in section 4.2. We also analyse the equilibrium investment strategies for the parameterisation chosen.

We solve for the equilibrium policy functions, which involves the simulation of the model as described in the methodology section. In order to ensure the convergence of the overall efficiency index and have a rich distribution of firm sizes, we simulate an industry with two hundred firms simultaneously active. Our choice of parameters is specified in table 4.1 .We choose the budget share of the monopolistically competitive sector large enough to ensure that given the demand and cost conditions in the industry, all firms will be making positive profits. The discount rate is set at 0.925 as in Pakes and McGuire (1995). The depreciation rate is 0.7 such that innovations become obsolete after approximately

1.5 years after they were developed. After this time, the innovation will still have an effect on the firm's marginal costs through its contribution to the public stock of knowledge. However, this knowledge is common to all firms in the industry and it does not affect their relative positions and hence their profits. While mc_0 leaves profits unaffected, the quantity choice decreases with the scalar on marginal costs. Thus, the rate of growth of total output produced in the industry is determined by the rate of decay of marginal costs, which stems from the improvements in the public stock of knowledge.

Firms' profitability depends only on their efficiency level and the overall efficiency index, which they take as given due to the assumption of aggregate price taking behaviour. The impact of each further innovation in decreasing a firm's marginal costs is determined by μ, which is set to 0.3. The elasticity of substitution is set to -2.0 such that the condition for the existence of a steady state solution to the R&D program, as stated in proposition 1, is satisfied. The unit cost of R&D investment is set to one and a half units of the numeraire.

The solution methodology involves the simulation of the evolution of the market structure for $100,000$ periods, given the initial condition or initial market structure, the optimal policy functions previously computed by the numerical algorithm and a random number generator. Figures 4.2 and 4.3 depict the profits and firm's value associated with each efficiency level given the long-run efficiency index. Figure 4.2 shows how profits evolve in a concave manner with the efficiency level, such that the incentives to improve the index of success decrease as the firm becomes more efficient. Figure 4.3 illustrates how firm value changes with the efficiency level. The concavity of the value function implies that firms at lower levels of efficiency will invest in R&D more heavily because their incremental value from developing an innovation is higher. Firms at the maximum efficiency level will not be engaging in any R&D investment. It is precisely this concavity of the value function that ensures the existence of a solution to the dynamic problem. If this condition fails to hold, which happens when the condition in proposition 1 is not met, the state of the industry explodes. Given the probability δ a firm faces of experiencing a depreciation in its efficiency level, and the small amount of R&D pursued at very high efficiency levels, there's a small probability that these firms will increase their index of success.

The optimal solution for R&D investment, depicted in figure 4.4, reflects the firms' incentives emerging from the relationship between both profits and value and the efficiency level. The highest level of investment is undertaken by the smaller firms, who have more to gain from developing innovations, than larger firms. Firms at very high efficiency levels do not invest because they have little to gain from innovation. These results show that the endogenous solution for R&D investment ensures that the bound on the maximum efficiency level is not exceeded.

The stochastic outcome associated with the R&D decision and the exogenous shocks to the efficiency levels govern the evolution of the market structure in the industry. By simulating the model presented here, we generate a firm size distribution with a substantial degree of heterogeneity. We compare the results we obtain with the properties of real firm size distributions. Figure 4.5 shows the size distribution for British firms in terms of their capitalization, illustrating the conventional, empirically documented properties of firm size distribution, found for the majority of industries. The distribution of firm sizes is highly skewed, with most of the firms being very small, and only a few firms ever achieving very large sizes. Figures 4.6 and 4.7 that follow illustrate the firm size distribution in levels and in logs, respectively, generated by the model.

The model generates a firm size distribution with some degree of heterogeneity, although the degree of heterogeneity and skewness amongst firms is far less than the one registered for actual firm size distributions. Similar results were found by Krusell and Smith (1998) for the distribution of income and wealth. They propose an extended version of the stochastic growth model with uncertain aggregate productivity shocks and idiosyncratic income shocks to include substantial heterogeneity in income and wealth and found it "to display far less cross-sectional dispersion and skewness in wealth than the U.S. data. (...) When the representative agent model is altered only by adding idiosyncratic, uninsurable risk, the resulting stationary wealth distribution is quite unrealistic: there are too few very poor agents and much too little concentration of wealth among the very richest". Our model also generates too few small firms compared with the size distributions found in most industries, and a smaller degree of asymmetry than expected.

Nevertheless, we are able to generate a population of heterogeneous agents in the context of this micro-based model of firm choices. Reproducing precisely the moments of the firm size distribution is definitely a way to progress in this line of research. We believe that introducing further elements of heterogeneity, such as heterogeneity in firms' discount factors, would allow to better approximating actual firm size distributions. Introducing turbulence into the model by adding entry and exit, and having randomly drawn sunk costs and scrap values, might be another way of bringing this model closer to reality.

4.6 Conclusions

There are very few macroeconomic models at present which include a rich representation of the micro structure. However, these issues are of fundamental importance given the need to develop policy prescriptions that do not rely so heavily on representative agent abstraction. As Ericson and Pakes (1995) state, "there is a policy and descriptive need for such models", given the importance of heterogeneity in firms' responses to assess the implications of the aggregate market response to a given policy.

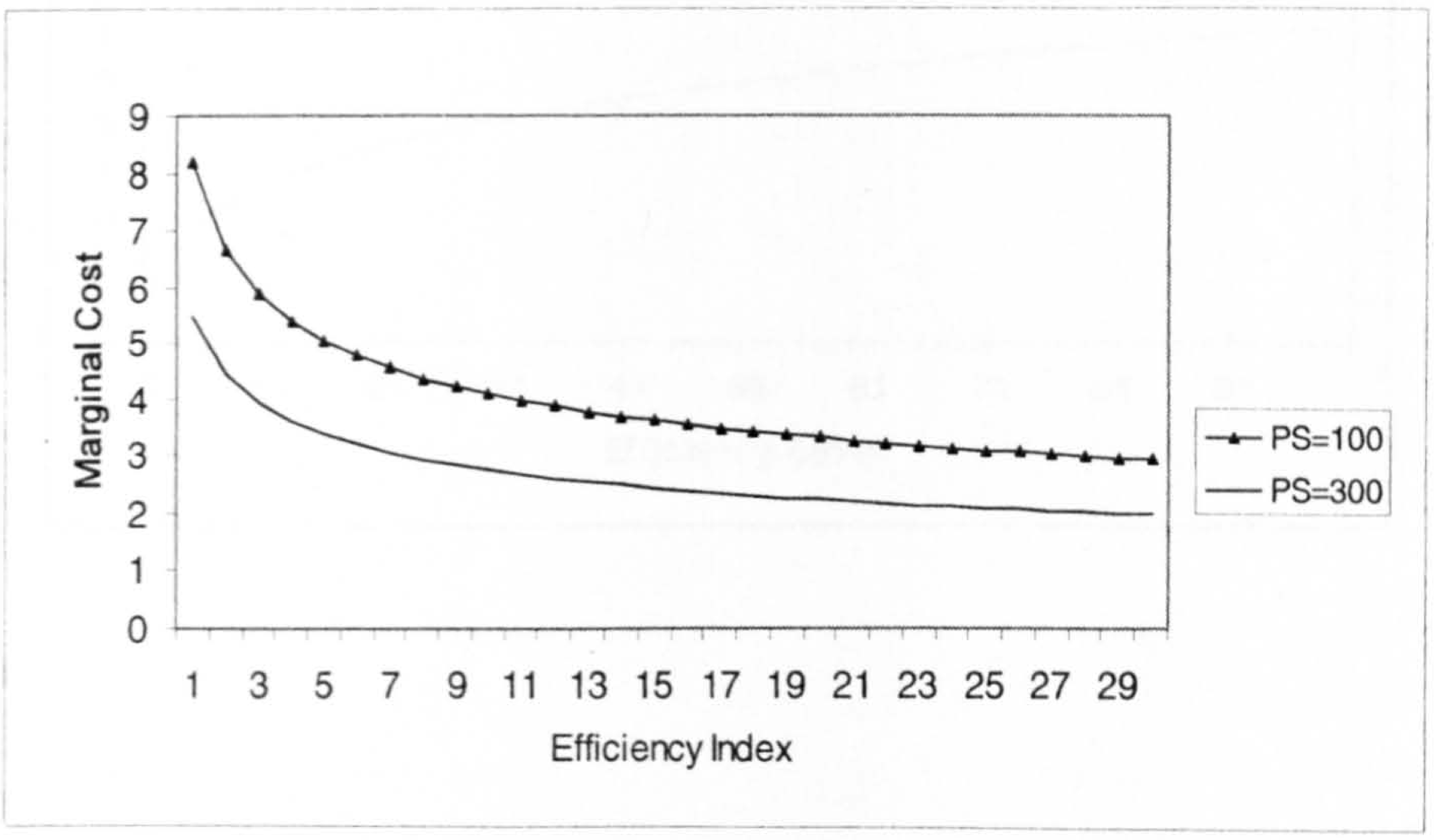
The model presented here is a model of endogenous industry growth, which arises from a process of knowledge diffusion that takes place automatically every time a firm succeeds in developing a new innovation. The micro setup of this model is a rich one, where firms invest in R&D in order to improve their relative position in the market structure and the dynamics are driven by the stochastic outcomes of firms' innovative activity.

Furthermore, we propose a method of firm decision making where firms form rational expectations concerning the overall long run average state of the industry rather than weighting the probability of the industry being in all possible future market structures. Growth arises from aggregate productivity shocks on the marginal costs conditions in the industry but, by leaving firm's profitability unchanged, these improvements have no effect on the optimal investment policy. Thus, we can treat the long run state index of the industry as constant even if the output produced in the industry is increasing over time. If the aggregate productivity shock would affect the firms' profitability, and consequently firms' optimal R&D program, we would not be able to solve for the equilibrium policy function for this model. This framework, contrary to most of the methodology proposed for solving dynamic models with multi-agent heterogeneity, does not suffer from the curse of dimensionality and can be solved for very high efficiency levels and for a large number of firms. Having price taking firms form expectations on the long run average industry state is an elegant and simplifying manner to overcome dimensionality problems, and the model allows for that while still being compatible with a growth process arising from changes in the aggregate conditions in the industry. After simulating the model for an industry with a large number of firms, we show that we are able to generate a firm size distribution with a substantial degree of heterogeneity.

Table 4.1: Parameter Values

Productivity of R&D Investment	a	1.5
Discount Factor	β	0.925
Probability of Exogenous Depreciation	δ	0.7
Unit Cost of R&D Investment	c	1.5
Rate of Marginal Cost Reduction from efficiency level	μ	0.3
Rate of Marginal Cost Reduction from Public Knowledge	γ	5.10^{-6}
Fixed Costs of Production	fc	3
Budget Share for Monopolistically Competitive Sector	Y	2000
Elasticity of Demand	ε	−2.0

Figure 4.1: The Change in the Marginal Cost Curve with PS



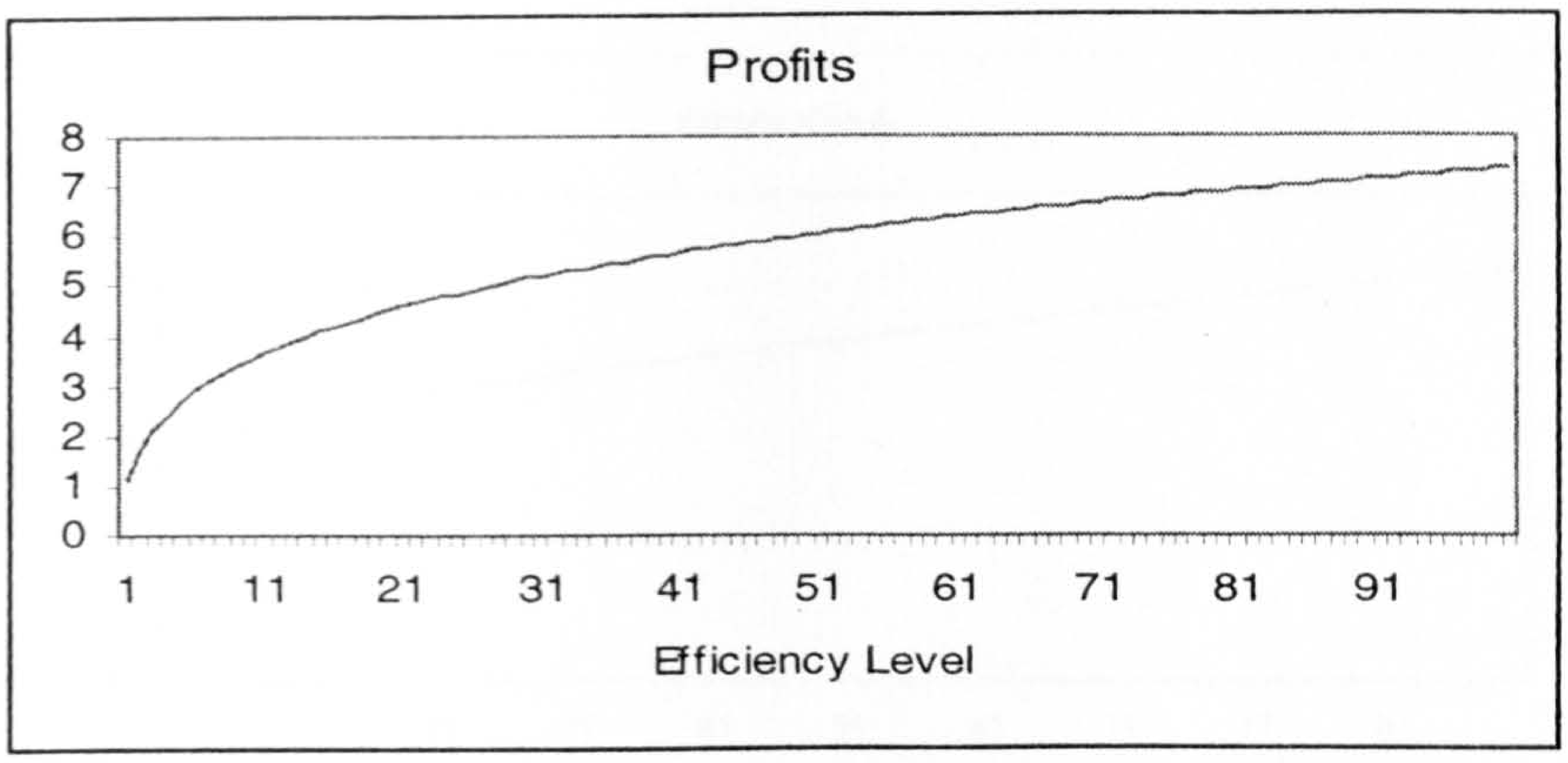

Profits
8
7
6
5
4
3
2
1
0
1
11
21
31
41
51
61
71
81
91
Efficiency Level

Figure 4.2: Firm's Value as a function of the Efficiency Level

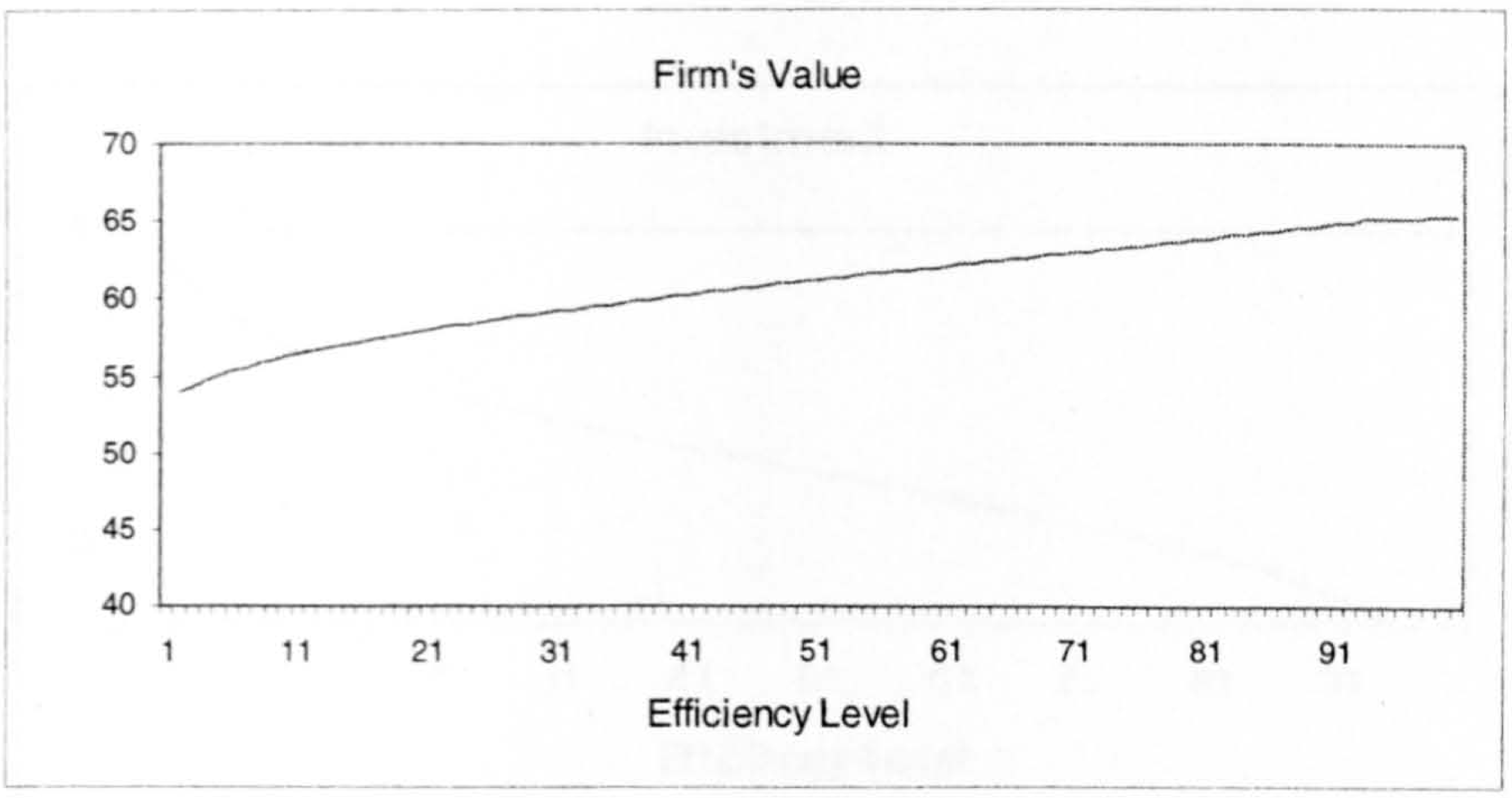


Figure 4.3: Investment as a function of the Efficiency Level

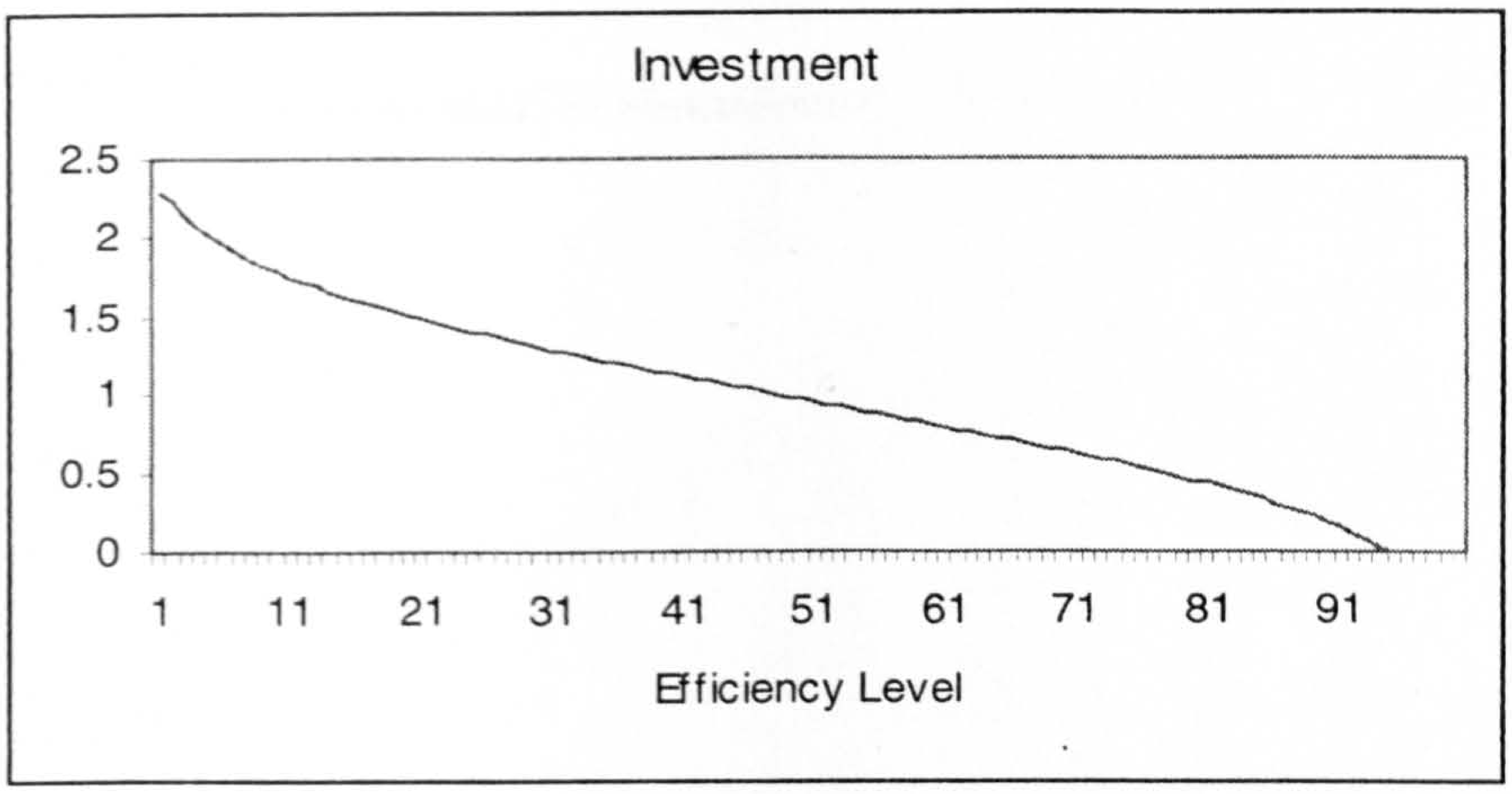


Figure 4.4: The Distribution of Firm Sizes for a Sample of British Firms

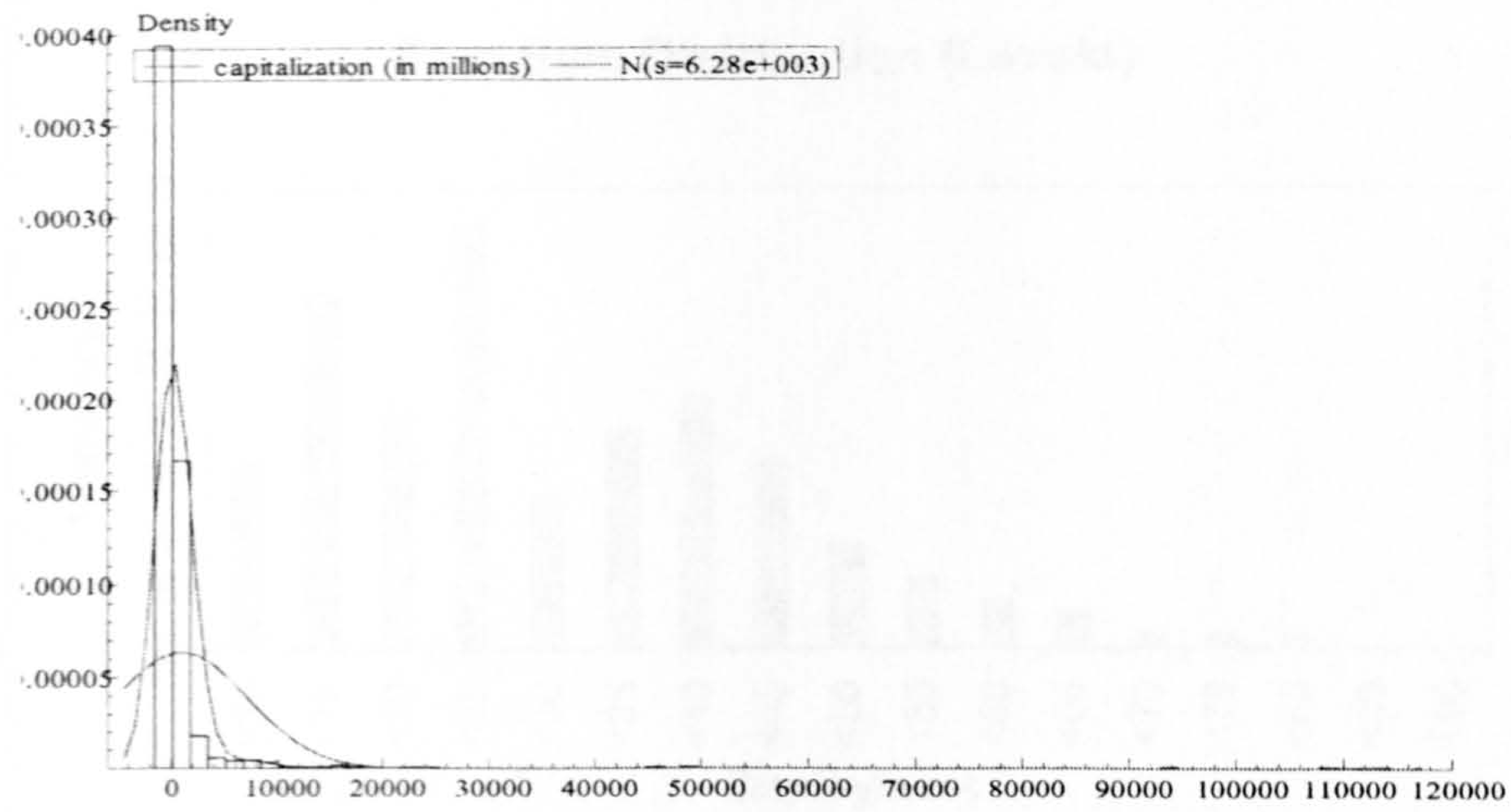


Figure 4.5: The Distribution of Employment

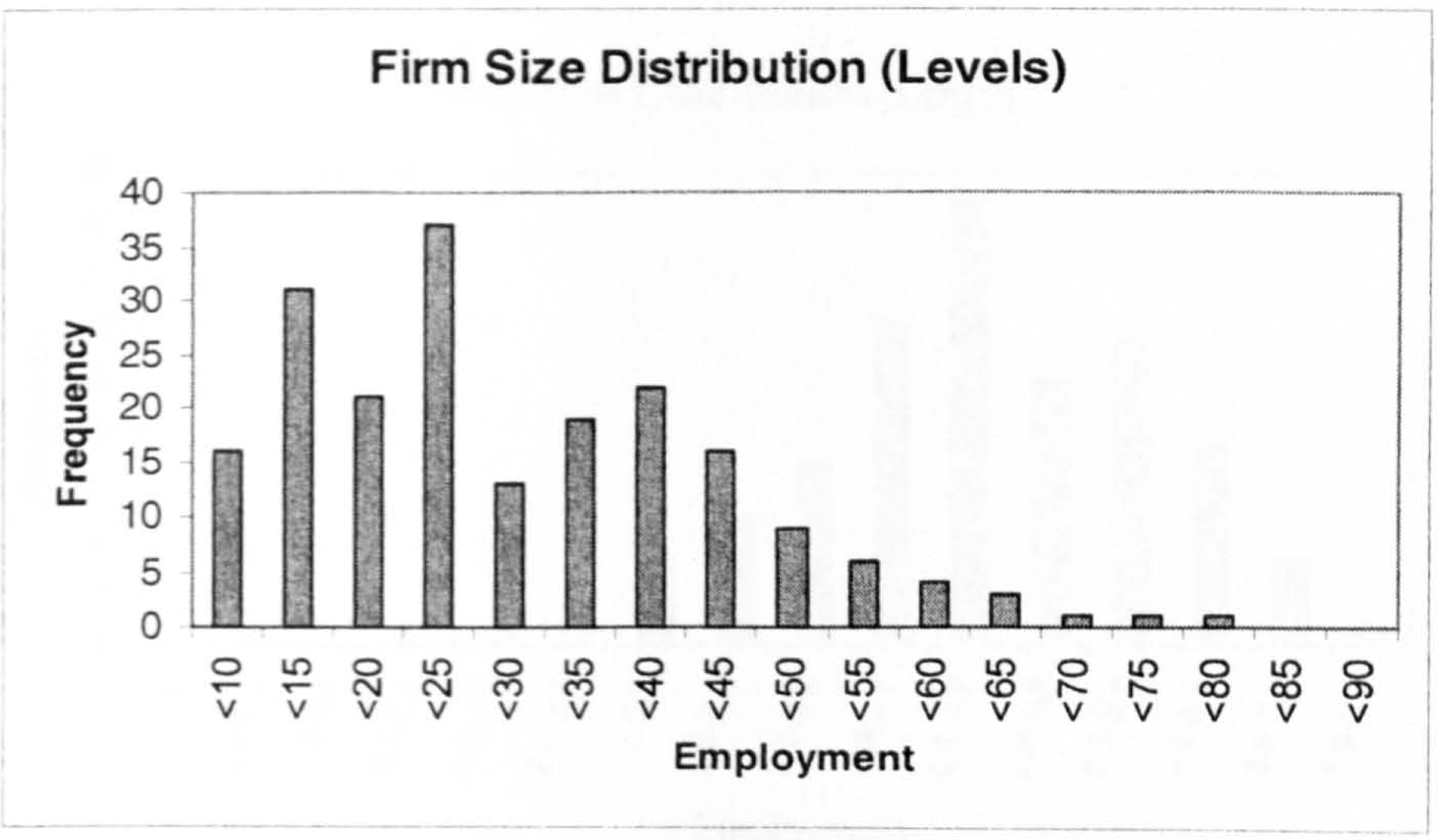
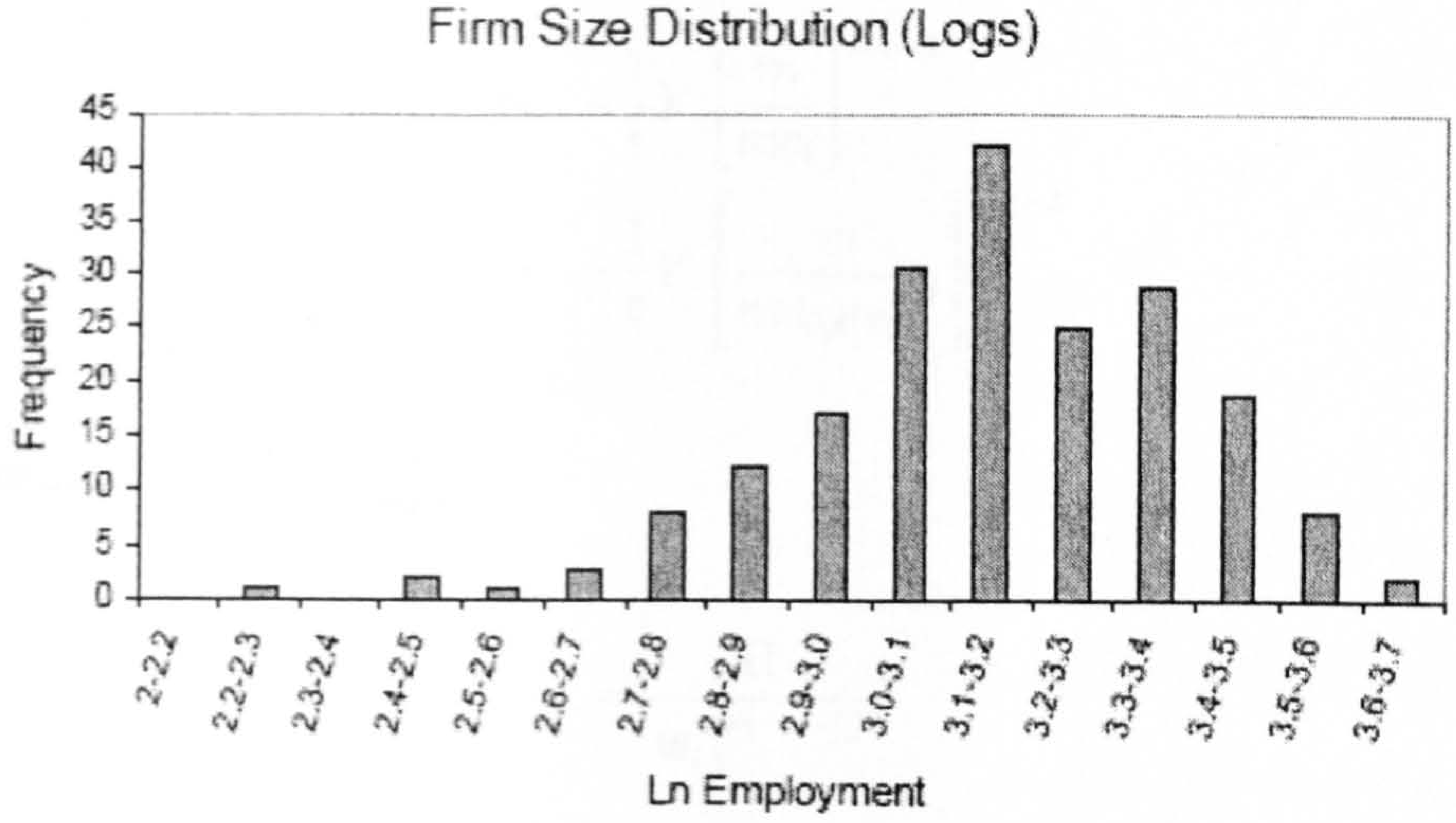


Figure 4.6: The Distribution of Ln Employment



4.7 Appendix

(Proff of proposition 1)

$$\begin{aligned}\pi_i &= -\frac{1}{\varepsilon}Y\left[\frac{m}{mc_i}\right]^{-\varepsilon-1} \\ &= -\frac{1}{\varepsilon}Y\left[\frac{m}{mc_{0,t}w_{i,t}^{-\mu}}\right]^{-\varepsilon-1}\end{aligned}$$

denoting $\Pi = -\frac{1}{\varepsilon}Y\frac{m^{-\varepsilon-1}}{mc_{0,t}}$

$$= \frac{\Pi}{w_{i,t}^{-\mu(-\varepsilon-1)}}$$

The fist order conditon is positive:

$$\partial\pi_i/\partial w = \Pi\mu(-\varepsilon-1)w_{i,t}^{-\mu(-\varepsilon-1)-1}$$

The second order condition:

$$\partial\pi_i^2/\partial^2 w = \Pi\mu(-\varepsilon-1)\left[\mu(-\varepsilon-1)-1\right]w_{i,t}^{-\mu(-\varepsilon-1)-2}$$

The condition to ensure the concativity of the profit function:

$$\begin{aligned}\mu(-\varepsilon-1)-1 &< 0 \\ -\mu(\varepsilon+1) &< 1.\end{aligned}$$

Chapter 5

Final Remarks

Throughout this thesis, we built on the Markov Perfect Dynamic Industry framework proposed by Ericson and Pakes (1995) to explore various dimensions of the dynamic strategic interaction between the active firms in an industry. We model an industry where firms differ in their efficiency level, which reflects the cumulative outcome of their innovation path, and determines their marginal costs of production. Firms invest in R&D in order to enhance the likelihood of developing a cost reducing innovation, and lowering their unit costs of production. The marginal cost each firm faces conditions its competitiveness in the product market, where firms engage in repeated competition. Profits are maximised in each period of time conditional on the firm's current marginal cost relative to that of its rivals. It's the relative positions in the industry that determine the spot market profits, not the absolute level of marginal costs. Firms are heterogeneous with respect to their efficiency levels, and thus are subject to different incentives to engage in R&D, which will generate heterogeneity in R&D efforts across firms. Within this framework, we studied how industry characteristics are correlated with the higher moments of the firm size distribution. These issues are manifestly important given the empirical evidence revealing cross industry differences in the statistical properties of the firm size distribution. It is our belief that the investigation carried out in this thesis represents a contribution to the understanding of the relationship between several industry characteristics (such as fixed costs, sunk costs of entry, the rate of technological progress, the productivity of R&D and the rate of technological spillovers) and the degree of concentration in the industry. To our knowledge, this is the first attempt to propose a theoretical model able to explain the disparity on firm asymmetry across industries. Furthermore, the framework proposed is micro based, i.e., it explains cross industry differences in the firm size distribution based on the outcome of rational firms' optimizing behaviour with respect to R&D investment, in an environment that is a good approximation for the market turbulence in which firms carry out their activity. We found that the model reasonably matches the real industry data both in terms of the statistical properties of the firm size distribution and in terms of cross sectional empirics. After validating the model for the task at hand, we analysed how the moments of firm size distribution change with industry characteristics identified as relevant in the literature. Our model predicts that industries with high fixed costs and higher technological growth are likely to have a higher mean size of firms, because they reduce the mass of small firms in the industry. By increasing the mass of medium sized firms and reducing the frequency of smaller firms, fixed costs flatten and lengthen the tail on the left side, reducing the skewness. With respect to the degree of technological progress in the industry we found that the higher the cost reduction in the industry the higher the incentives to invest in order to enhance the efficiency in the product market relative to the other players as well as defending their existing positions. Thus, rapid growth should lead to a high variance and a flatter distribution.

Analogously to fixed costs, sunk costs also reduce the mass of small firms, but their impact is rather at the level of potential entry, due to its discouraging effect of reducing the expected discounted value of entry. Industries with higher sunk costs will exhibit greater average firm size, higher variance in size but a flatter distribution with multiple modes. Similarly to the empirical results of Audrestch et. al. (2004), we also found that industries where entry is less costly will more nearly match the log normal distribution and the strong form of Gibrat's Law, and this is particularly true in industries less capital intensive and where scale economies play less of a role. Increases in the productivity of R&D increase the mass of firms to the right of the peak of the distribution, therefore increasing the mean size of firms, but, similarly to the effects of technological progress, we should not expect the productivity of R&D to have clear monotonic effects on the higher moments of the firm size distribution. Concerning the rate of spillovers, we found tat industries where incumbents can protect their advantages through secrecy, patent protection and/or lead time will have a higher mean and variance in the size of firms, because the higher the appropriability of an innovation, the higher its value to the firm and the higher the incentives to invest. The impact of the rate of spillovers on the skewness and kurtosis are highly non-monotonic and depict no discernible pattern, which prevents us from drawing any conclusions regarding the effects of this structural parameter on the higher moments of the firm size distribution.

Avenue for future work would be to examine empirically the hypotheses generated by the theoretical model. This exercise would most certainly deliver orientation for further refinements to the model. Applications of the model to particular industries by calibration of the parameter set would provide an excellent framework for analysing the effects of policies, such as R&D subsidization or changes in entry regulation, on the firm size distribution.

This thesis also investigated the impact of knowledge externalities on firms' incentives to perform R&D. We extended the E-P (1995) dynamic industry framework to allow for the fact that the R&D investment each firm pursues adds into a public stock of knowledge that improves the knowledge conditions in the industry. We showed that the lack of appropriability has an unambiguous negative effect on the amount of R&D undertaken in the industry. That does not imply, however, that the rate of innovation in the industry falls given that, with knowledge spillovers, the same level of R&D spending delivers a higher innovation rate than under full appropriability since firms also benefit from the external pool of R&D. Upon simulating the model, we found that when R&D spillovers are costlessly absorbed, the rate of innovation actually decreases with respect to the case of full appropriability due to the strategic effect of R&D spillovers. Firms combine own R&D investment and the external knowledge available in the industry in order to achieve the optimal rate of innovation, and take advantage of the scope for cost saving in R&D spending that follows from R&D spillovers. The fact that each firm anticipates that the rival will free ride on its own investment leads to a further decrease in the level of its R&D expenditure. Due to this strategic effect, the rate of innovation in the industry will decay, and this will translate into lower consumer welfare. The level of concentration in the industry initially decreases as the gap between the leader and the follower falls but for higher degrees of spillovers this gap starts widening up and concentration starts increasing. Overall, we found the effect of costless spillovers on concentration to be non-monotonic and not very strong, even if there is a clear change in the ergodic distribution of market structures.

By comparing these results, obtained under the assumption of public knowledge being a pure public good, with those obtained under the assumption that firms that have their own R&D are better at making use of outside knowledge, we find that the Schumpeterian relationship between appropriability and innovation does not hold in the context of this stochastic dynamic industry environment. These results are in line with those obtained by Cohen and Levinthal (1989) in the context of a static model, and which they support in another paper with a review of the empirical evidence, Cohen and Levinthal (1990). In fact, when imposing the requirement of research effort for assimilating external R&D, the productivity of R&D increases because besides increasing the rate of innovation by adding a further unit to the level of own R&D, it also improves the amount of external R&D devoted to innovation. As a result, the increased incentive for R&D investment that follows from the increase in the productivity of R&D compensates for the disincentive to invest that arises from the free riding effect. We also found, in our duopoly model, that the follower firm experiences an improvement in its position relative to the leader, which implies that its value increases and that the level of concentration in the industry decreases.

This framework can be brought closer to reality by introducing entry and exit in the model and by expanding the number of firms interacting in the industry. Allowing for the R&D externalities to affect the entry decision of potential competitors reinforces the strategic effect of R&D spillovers and may offset the welfare gains that arise in the absorptive capacity case. Another way of pursuing further this line of research would be to analyse the robustness of the results to different assumptions concerning the creation of absorptive capacity. An example would be to analyse the impact of R&D spillovers when the firm's absorptive capacity is path dependent, e.g. a function of R&D stock rather than R&D flow, or as a function of the firm's past successes, i.e., the firm's stock of innovations. The contestability of the leader's position in this scenario would be lower, and the incentives to invest by the follower firm would decrease.

In chapter 4, we propose a framework to approximate an industry with several firms. We move away from the oligopolistic setting by assuming that there are many heterogeneous firms, each firm being too small to have a strategic effect on its rivals' choices. In the original E P framework, the oligopolistic market structure follows an ergodic stochastic process and settles down into some recurrent pattern, but with a limited number of firms. At each moment in time, decisions concerning R&D investment are made taking into account the firm's perceptions concerning the probability associated with each rival's outcome. In the model proposed here, by having many firms simultaneously active where each firm is too small to have a strategic effect on its competitors' decision, firms do not form perceptions concerning the probability outcomes for each of its rivals, they rather only care about the long run average industry state. We are therefore able to overcome the *"curse of dimensionality"*, and solve the stochastic dynamic industry game with many heterogeneous agents. We developed an algorithm and code to find the optimal policy functions for the model presented, and simulated the market structure for a given set of parameters for 100.000 periods of time. We plotted the simulated data in size bins and showed that the model is able to deliver a firm size distribution with a substantial degree of heterogeneity. This model also accounts for productivity growth, maintaining the tractability of the solution methodology. We concluded, however, that in order to approximate the firm size distribution to that obtained with real industry data, there should be more heterogeneity across firms.

This framework is a first step and can be improved in many ways. Introducing selection would enrich the model, and bring it closer to reality by adding turbulence to the environment. Furthermore, entry and exit will most likely increase the mass of small firms in the industry, thus increasing the skewness of the distribution. Further improvements would involve introducing more elements of heterogeneity amongst firms, such as firm specific randomly drawn scrap values and/or time discount factors.

Bibliography

[1] Arrow, K (1962), "Economic Welfare and the Allocation of Resources for Invention" in Richard Nelson, ed., *The Rate and Direction of Inventive*

[2] Aghion, P., C. Harris, P.Howitt and J. Vickers (2001), "Competition, Imitation and Growth with Step-by-Step Innovation". *Review of Economic Studies*, 68, pp. 467-492.

[3] Aghion, P. and P. Howitt (1992) *A Model of Growth throught Creative Destruction Econometric,* Vol 60(2), pp323-351.

[4] Audretsch, D., L. Kompf, E. Santarelli, and A. Thurik, 2004, "Gibrat's Law: are the Services Different?" forthcoming in *Review of Industrial Organization.*

[5] Arrow, K (1962), "Economic Welfare and the Allocation of Resources for Invention" in Richard Nelson, ed., *The Rate and Direction of Inventive Activity,* Princeton University Press.

[6] Barro, R. and X. Sala-i-Martin (1995) *Economic Growth* McGraw-Hill, New York.

[7] Bottazi, G. and A. Secchi, 2003a, "Sectoral Specificities in the Dy-

namics of US Manufacturing Firms." *LEM Working Paper* 2003/18, http://www.sssup.it/~LEM/.

[8] Bottazi, G., E. Cefis, G. Dosi, and A. Secchi, 2003b, "Invariances and Diversities in the Evolution of Manufacturing Industries." LEM Working Paper 2003/21, http://www.sssup.it/~LEM/

[9] Buzzacchi, L. and Valletti (2006), "Testing the "Independent Submarkets Model" in the Italian Motor Insurance Industry". International Journal of Industrial Organization, vol. 24 (4) pp. 809-834.

[10] Caves, R., 1998, "Industrial Organization and New Findings on the Turnover and Mobility of Firms." *Journal of Economic Literature*, 36 (December, 1998), pp. 1947-198

[11] Cohen, W. and S. Klepper (1992) "The Anatomy of Industry R&D Intensity Distributions," *American Economic Review,* Vol. 82(4), pages 773-99, September

[12] Cohen, W and D. Levinthal (1989), "Innovating and Learning: The two faces of R&D". *Economic Journal*, vol. 99.

[13] Cohen, W. and D. Levinthal (1990), "Absorptive capacity: a new perspective on learning and innovation", *Administrative Science Quarterly,* 35, pp.128-52.

[14] D'Aspremont, C. and A. Jacquemin (1988) "Cooperative and Noncooperative R&D in Duopoly with Spillovers" *The American Economic Review,* vol. 78, No 5, pp 1133-1137.

[15] Dasgupta, P. and J, Stiglitz (1980) "Industrial Structure and the Nature of Innovative Activity" *The Economic Journal*, Vol. 90, No. 358 pp. 266-293.

[16] Dorazelski,U. and K. Judd, (2005) "Avoiding the Curse of Dimensionality in Dynamic Stochastic Games" NBER Technical Working Paper No. 304

[17] Dunne, P. and A. Hughes, 1994, "Age, Size, Growth and Survival: UK Companies in the 1980s." *Journal of Industrial Economics*, 42 (2), pp. 115-140.

[18] Ericson, R. and A. Pakes, 1995, "Markov-Perfect Industry Dynamics: A Framework for Empirical Work." *Review of Economic Studies*, 62, pp. 53-82.

[19] Evans, D., 1987, "The Relationship Between Firm Growth, Size, and Age: Estimates for 100 Manufacturing Industries." *Journal of Industrial Economics*, 35 (4): 567-581.

[20] Gilbert, R. and Newberry, D. (1982) "Preemptive Patenting and the Persistence of Monopoly", *The American Economic Review*, vol. 72, No 3, pp 514-526.

[21] Gowrisankaran, G., 1999. A Dynamic Model of Endogenous Horizontal Mergers. *RAND Journal of Economics*, 30, pp. 56-83.

[22] Griliches, Z. (1992) "The Search for R&D Spillovers," *NBER Working Papers* 3768, National Bureau of Economic Research, Inc.

[23] Grossman, G. and E. Helpman (1991) *Innovation and Growth in the Global Economy*. MIT Press, Cambridge MA.

[24] Hall, B., 1987, "The Relationship Between Firm Size and Firm Growth in the US Manufacturing Sector." *Journal of Industrial Economics*, 35 (4): 583-606.

[25] Hart, P. and N. Oultan, 1996, "Growth and Size of Firms." *The Economic Journal*, 106 (438): 1242-1252.

[26] Hart, P. and S. Prais, 1956, "The Analysis of Business Concentration: A Statistical Approach." *Journal of the Royal Statistical Society*. Series A, 119 (2): 150-191.

[27] Henriques, I. (1990), "Cooperative and Noncooperative R&D in Duopoly with Spillovers: Comment." *American Economic Review*, Vol. 80, No. 3, pp. 638-640.

[28] Ijiri, Y. and H. Simon, 1964, "Business Firm Growth and Size." *American Economic Review*, 54 (2), pp. 77-89.

[29] Jorgenson, D.W. and K. Stiroh, 2000, "Industry-Level Productivity and Competitiveness Between Canada and the United States." *American Economic Review*, 90 (2), pp. 161-167.

[30] Jovanovic, Boyan, 1982. "Selection and the Evolution of Industry." *Econometrica*, 50 (3), pp. 649-670.

[31] Jaffe (1986),"Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value". *The American Economic Review*, Vol. 76, No 5.

[32] Kamien, M. and Schwartz, N. (1971) "Limit Pricing under Uncertain Entry" *Econometrica*, Vol. 39, No 3, pp 441-454.

[33] Kamien, M. and Schwartz, N. (1972) "Timming of Innovations Under Rivalvry" *Econometrica*, Vol. 40, No 1, pp 43-60.

[34] Klepper, S., 2002. "Firm Survival and the Evolution of Oligopoly." *RAND Journal of Economics* 33 (1), pp. 37-61.

[35] Klette, T. and Z. Griliches, (1998). "Empirical Patterns of Firm Growth and R&D Investment: a Quality Ladder Model Interpretation," Economic Journal, Royal Economic Society, Vol. 110(463), pages 363-87

[36] Klette,T. and S. Kortum, (2003). "Innovating Firms and Aggregate innovation." *Journal of Political Economy,* Vol 11285) pp. 986-1018.

[37] Kumar, M. (1985), "Growth, Acquisition Activity and Firm Size: Evidence from the United Kingdom." *Journal of Industrial Economics,* 33, pp. 327-338.

[38] Krusell, P. and A. Smith (1998) *"Income and Wealth Heterogeneity in the Macroeconomy"* Journal of political Economy Vol 106(5) pp. 867-896

[39] Laincz, C., 2004a, "Market Structure and Endogenous Productivity Growth: How do R&D Subsidies Affect Market Structure?" *Journal of Economic Dynamics and Control,* 29 (1-2), pp. 187-224.

[40] Laincz, C., 2004b, "A Model of Growth with Dynamic Market Structure" Working Paper.

[41] Laincz, C. and A. Rodrigues, 2005, "The Impact of Cost-Reducing R&D Spillovers on the Ergodic Distribution of Market Structures." Working Paper.

[42] Lotti, F. and E. Santarelli, 2004, "Industry Dynamics and the Distribution of Firm Sizes: A Nonparametric Approach." *Southern Economic Journal* 70 (3), pp. 443-466.

[43] Levin and Reiss (1988), "Cost Reducing and Demand Creating R&D with Spilloovers ". *Rand Journal of Economics*, vol. 99.

[44] Loury, G. (1979) "Market Structure and Innovation". *The Quarterly Journal of Economics.* Vol 93, No 3, pp 3495-410.

[45] Machado, J. and J. Mata, 2000, "Box-Cox Quantile Regression and the Distribution of Firm Sizes." *Journal of Applied Econometrics*, 15, pp. 253-274.

[46] Mansfield, E., 1962, "Entry, Gibrat's Law, Innovation, and the Growth of Firms." *American Economic Review*, 52 (5), pp. 1023-1051.

[47] Mansfield, E., M. Schwartz, and S. Wagner, 1981, "Imitation Costs and Patents: An Empirical Study." *The Economic Journal*, 91, pp. 907-918.

[48] Mansfield, E., 1985. "How Rapidly Does New Industrial Technology Leak Out?" *Journal of Industrial Economics*, 34, 2, pp. 217-223.

[49] Maskin, E. and Tirole, .J (1988), "A Theory of Dynamic Oligopoly: I & II ". Econometrica Vol 56, pp 549-600

[50] McCloughan, P., (1995), "Simulation of Concentration Development from Modified Gibrat Growth-Entry-Exit Processes." *Journal of Industrial Economics*, 43 (4), pp. 405-433.

[51] Nelson, R. (1962) "The Rate and Direction of Inventive Activity", *Princeton University Press.*

[52] Pakes, A. and Doraszelski,U. (2006)" Framework for Applied Dynamic Analysis in IO". Chapter Draft for the Handbook of I.O., available at the authors' homepages.

[53] Pakes, A. and R. Ericson, (1998). "Empirical Implications of Alternative Models of Firm Dynamics." *Journal of Economic Theory*, 79, pp. 1-45.

[54] Pakes,A. and McGuire (1994) "Computing Markov Perfect Nash equilibria: Numerical Implications of a Dynamic Differentiated Product Model". *RAND Journal of Economics, RAND*, vol. 25(4).

[55] Pakes,A. and P. McGuire (2001), "Stochastic Algorithms, Symmetric Markov Perfect Equilibria, and the 'Curse' of Dimensionality",Econometrica, vol. 69, no. 5, pp. 1261-81.

[56] Prais, S., 1976, *The Evolution of Giant Firms in Britain*, Cambridge, Cambridge University Press.

[57] Reinganum, J. (1983) "Uncertain Innovation and the Persistency of Monopoly". *The American Economic Review*, vol. 73,No 4, pp 741-748.

[58] Rosenberg,N. (1974) "Science, Invention and Economic Growth". *Economic Journal, vol. 84.*

[59] Schumpeter, J. (1942) *"Capitalism, Socialism and Democracy".*

[60] Simon, H. and C. Bonini (1958), "The Size Distribution of Business Firms." *American Economic Review*, 48, pp. 607-617.

[61] Simpson, R.D. and N. Vonortas (1994), "Cournot Equilibrium with Imperfectly Appropriable R&D." *Journal of Industrial Economics*, Vol. 42, No. 1,. pp. 79-92.

[62] Singh, A. and G. Whittington (1975), "The Size and Growth of Firms." *Review of Economic Studies*, 42 (1): 15-26.

[63] Spence, M. (1984) "Cost Reduction, Competition, and Industry Performance." *Econometrica*,Vol. 52, No. 1, pp. 101-122.

[64] Stanley et al (1997), "Scaling Behaviour in Economics: The Problem of Quantifying Company Growth", Physica A pp. 1-24.

[65] Stanley et al (1996), "Scaling Behaviour in the Growth of Companies", Nature vol. 379, 804-806.

[66] Sutton, J., 1997, "Gibrat's Legacy." *Journal of Economic Literature*, 35 (1), pp. 40-59.

[67] Sutton, J., 1998, Technology and Market Structure. MIT Press: Cambridge, MA.

[68] Suzumura, K. (1992), "Cooperative and Noncooperative R&D in an Oligopoly with Spillovers." *American Economic Review*, Vol. 82, No. 5, pp. 1307-1320.

[69] Thompson, P. (2001) "The Microeconomics of an R&D based Model of Endogenous Growth". *Journal of Economic Growth,* Vol 6(4) pp. 263-283

[70] Ziss, S. (1994), "Strategic R&D with Spillovers, Collusion, and Welfare." *Journal of Industrial Economics*, Vol. 42, No. 4, pp. 375-393.