Behaviour and Ownership

in the Theory of Competition and Regulation

by

Michael Hermann Hardt

London School of Economics.

A dissertation submitted to the University of London for the degree of Doctor of Philosophy.

But he that is an hireling, and not a shepherd, *whose own the sheep are not*, seeth the wolf coming, and leaveth the sheep. The hireling fleeth, because he is an hireling and *careth not*...

The Bible, Gospel of John, chapter 10, verses 12 and 13. KJV.

Abstract

Ownership matters. It affects residual rights under incomplete contracts and, therefore, incentives. The first chapter of this thesis analyzes in how far ownership can be substituted by other economic factors. Contrary to an assumption found in the literature market foreclosure can be achieved without vertical integration in the following scenarios: repeated games, reputation games, and also in a finitely repeated game when there are switching costs.

The main chapter is concerned with implications of ownership in regulated industries where a monopolistic supplier of an essential input is required by a regulator to charge cost based prices. Our analysis focuses on the impact of ownership on the monopolist's incentives to exploit informational asymmetries about production costs. We conduct a comparative study of vertical integration, vertical separation, and joint ownership. Effects on welfare, investments incentives, and entry are analyzed for each ownership structure. Joint ownership performs best. Accounting separation is shown to be generally ineffective as regulatory instrument.

We use an alternative model which allows to take into account network duplication. Starting from a free market analysis of equilibrium pricing and entry decisions we explore the relation between ownership and the degree of regulation required in order to ensure efficient outcomes. Two part tariffs, network duplication, price discrimination and a long-term commitment to fixed input prices induce reductions of final prices.

The final part of this thesis investigates results in the theory of competitive market equilibrium. Many of these results rely on restrictive assumptions on consumer behaviour. We analyze in how far traditional equilibrium theory is robust against a relaxation of underlying assumptions. We do not assume agents to be rational in the sense that their choices arise from maximisation. Randomly fluctuating demand is allowed for and consequences for predictions made by traditional competitive equilibrium theory are re-examined.

Acknowledgements

I am indebted to no less than four supervisors. During my studies in Bonn, Werner Hildenbrand's enthusiasm for General Equilibrium Theory proved to be a factor that influenced strongly my research interests. At the London School of Economics, Tony Horsley repeatedly encouraged me to finish and generalise the work I had started in Bonn. John Sutton's lectures as well as some seminar series at the LSE introduced me, in a very stimulating way, to the theories of firms, industries, contracts and related issues. During this period, Patrick Bolton gave me a most valuable supervision. After each weekly appointment I was impressed. His comments did not only prove extremely helpful but, to a large degree, shaped my way of thinking about economic problems. Although, due to Patrick's illness and subsequent academic leave, this period was very short, his constructive criticism and encouragement, very soon, led to my first published paper. I would like to thank him very much. Further, I am particularly indebted to John Moore for taking over the supervision of a thesis large parts of which, by then, had already been written. He not only gave valuable academic guidance in the final stage but very kindly introduced me to a number of seminars and workshops to which I would not have had access otherwise. Apart from that I was often impressed by his knowledge, commitment and, last not least, his speed of reasoning. I profited from all.

There is a large number of individuals to whom I am grateful for helpful comments and support. I met them at the LSE, at the University of Bonn as well as in Berlin, Cambridge, Jerusalem, Louvain-la-Neuve, Paris and other places. I would like to thank Urs Schweizer for his commitment in organising the European Doctoral Programme which opened many doors that proved to be of considerable value.

Financial help from the Graduiertenkolleg, from the German Academic Exchange Service, the London School of Economics and from the Royal Economic Society is gratefully acknowledged.

Apart from academic guidance and financial help there are other, possibly less tangible, factors that contributed to the completion of this thesis. Friends, family and, last not least, Nelly, have not only borne much of the social costs involved but also encouraged and helped me. I thank them all for their, often prayerful, support.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

CHAPTER I

1 Introduction

Concerning the organisation of this thesis a few preliminary remarks may be in order. Chapters are numbered in Roman numbers. Sections, subsections etc. within these chapters are numbered in Arabic numbers, separately for each chapter. The same applies to footnotes. Each chapter contains a short abstract, introduction, conclusion, and references. In the introduction to the thesis as a whole we aim at presenting ideas, intuition, and links between the different chapters. Therefore, most references will be given in the separate introduction of each chapter. The few references we do give in the introduction are included in the reference list of the chapter referred to.

The size of a firm is well known to be among the issues that the neoclassical theory of the firm, despite of all the useful explanations it has provided in relation to many production process oriented questions, has not been able to explain. Neither from an individual firm's point of view nor from the point of view of social welfare has traditional theory provided answers to the question of what determines or should determine the boundaries of the firm. It has not even provided a framework within which these questions can be addressed because it leaves the extent of the firm undefined. Some modern market economy phenomena like horizontal or vertical mergers and takeovers can therefore not be explained on the basis of neoclassical theory and even less can policy recommendations on these issues be founded on neoclassical theory.

An early attempt to explain why certain transactions are carried out within the firm and others are 'contracted out' is the approach suggested by Coase (1937). The central underlying idea is the importance of transaction costs. This drives the conclusion that the market will be used to carry out a given transaction unless the same can be done at lower costs within the firm. Although this approach provides an important rationale for the existence of an *asymmetry* between those transactions carried out in the market and those carried out within the firm it does not define a difference *in nature* between these two kinds of transactions. Rather, both can be viewed as transactions in an 'internal' market known as firm and an 'external' market depending on which possibility appears to be advantageous. A similar criticism applies to the suggestion that the firm should be viewed as a certain type of contract (see Jensen and Meckling (1976)). While it is possible to define and describe such possible types of contracts that can be viewed as 'firm' this would not help to explain the amount and kind of transactions that should be included in the standard type of contract called 'firm' and those that should be agreed in separate contracts.

An argument that does address the question has been put forward by Williamson (1985), namely the idea that the degree to which opportunistic behaviour is profitable for a particular party depends on whether or not the transaction concerned is carried out within one and the same firm. Williamson argues along the following lines. In the case of two separate firms where relationship specific investments need to be undertaken in the first stage and surplus is realised and divided in the second stage, the following problem may occur. Informational asymmetries or the large number of situations that may, in all possible states of the world, arise during a contractual relationship make a complete specification of all actions to be carried out under the contract either impossible or too costly. Therefore, a hold-up problem results in the final stage and will, in general, induce inefficient investment levels to be chosen in the first stage. But this would no longer be the case if firms merged before investment decisions are taken.

This is essentially the idea that has been taken up and formalised in the ownership rights approach (see [1] Grossman and Hart (1986)) in the following way. All decisions that remain to be taken after an

[1]An outline of this idea as well as various other views on the theory of the firm are found in Hart (1989). For further and more recent references (and implications for corporate governance) see Hart (1995a) and Hart (1995b).

(incomplete) contract has been agreed will be taken by the owner of the assets needed to carry out the purpose of the contract. Therefore, ownership crucially affects the incentives of parties involved and there are clear cut reasons for choosing particular boundaries of a firm. Horizontal or vertical mergers or break-ups become worthwhile even though substantial transaction costs may be involved in the respective process.

During the past decade a fast growing literature based on the importance of ownership rights has developed and started to explore the costs and benefits of particular firm boundaries. Research has been conducted with a view to shedding light on many important phenomena in modern market economies. Vertical and horizontal firm and industry structure, principal-agent relationships, antitrust policy and regulated industries are just some of the areas in which ownership has been identified as a crucial factor and interesting insights have been won in all these areas.

In particular, an extensive literature on determinants and effects of vertical integration has been built on ownership rights. We do not attempt to survey[2] this literature here but just point to a most concise summary of two classes of effects resulting from vertical integration which has been given by Armstrong, Cowan and Vickers (1994). They point out that most issues discussed in this context can be subsumed under one of the following two main purposes for vertical integration, namely (1) overcoming externalities between the parties involved, and (2) creating strategic advantages with respect to third parties (p.147). Much of the literature on vertical integration can be seen as an investigation of the question which of the above effects dominates in a given setting. The conclusions as to whether or not vertical integration is desirable rely on the fundamental assumption that residual rights of control are allocated to the owner of the assets in question.

There is, however, a need to balance the import of ownership with the import of other economic considerations that may be relevant to the situation in question. An attempt to highlight effects of owner-

[2]Surveys are for example Perry (1987) or Waterson (1993).

ship may bear a danger of neglecting other factors affecting the issue. Assumptions made may then reflect the convenience of modelling a given situation rather than the situation itself. Demonstrating the degree to which ownership matters has proved to be a fruitful line of research and will partly be followed in this thesis. But conclusions derived from the ownership rights approach should be considered in the light of the following question. To what extent can the effects that have been attributed to ownership (and rightly so) be generated by alternative devices? Paraphrasing, one may ask whether or not results which are driven by the fact that, under contractual incompleteness, ownership entitles to residual rights, should be attributed to ownership exclusively.

This question is taken up in the chapter on vertical integration (chapter II) and may be regarded as an attempt of viewing ownership implications in a broader framework encompassing other economic factors such as dynamic effects, reputation, relationship specific investments, to name just a few.

It would be a misunderstanding to interpret this chapter as a criticism of the literature based on ownership rights. It rather attempts to shed light on the interpretation of results derived within the ownership rights approach. For pointing out alternative strategies that lead to an effect attributed to vertical integration, for instance, can as well be interpreted as pointing out strategies that can be substituted by vertical integration.

A particularly clear case demonstrating how the importance attributed to ownership may rely on an assumption deserving re - examination is found in a paper on market foreclosure and vertical integration by Hart and Tirole (1990). The authors show that, in a vertically related markets setting, an upstream monopolist has incentives to integrate with one downstream firm in order to foreclose another downstream firm from the input supply market. The argument is that a non-integrated monopolist, having supplied one downstream firm, always has ex post incentives to supply a further downstream firm as well.

Due to the upstream firm's inability to commit to exclusive-dealing no firm in the downstream market will anticipate to make monopoly profits and therefore no downstream firm will sign an exclusive-dealing contract that would be based on these profits. Any such contract would have to be non-enforceable because, firstly, it does not comply with antitrust regulations and, secondly, there may be informational asymmetries about quantities supplied to other firms (this could happen, for instance, via third parties, i.e. firms operating in different markets). This contractual failure of preventing opportunistic behaviour, Hart and Tirole argue, can be overcome with the help of vertical integration.

While this conclusion is certainly true the question arises whether or not this failure can be overcome by vertical integration only. From a policy oriented point of view one may ask whether or not an antitrust policy that forbids mergers which are suggestive of foreclosure motives is sufficient to prevent foreclosure. In other words, the question is which alternative strategies are available to the upstream firm that would enable it to successfully commit to exclusive-dealing while remaining a separate legal entity.

Apart from shedding light on the question how unable to precommit a non-integrated supplying monopolist is, i.e. which alternative commitment devices exist, this analysis may serve as a solution for a puzzle created by some antitrust cases. Following Hart and Tirole's assumption it is hard to explain why certain (non-integrated) suppliers not only manage to sustain successfully an exclusive-dealing oriented policy but are even willing to spend large amounts of resources on defending their exclusive-dealing practices before court. Such casual evidence can, on the other hand, be well explained once various commitment devices that are at the disposal of vertically separated firms are taken into account.

One very obvious scenario in which commitment arises as equilibrium outcome is the following simple modification of Hart and Tirole's basic model. Considering a multiperiod game in which, in each period, there is a positive probability that the stage game will be repeated one finds that, unless discount rates are too high, expected future profits are sufficient to prevent the monopolist from breaking an exclusive-dealing contract. This simple model may well capture the fact that firms take into consideration the chance of dealing with a given trading partner again in the future. Corporate strategy, especially in oligopolistic or even semi-monopolistic industries, can hardly be assumed to be fully captured by a one shot game.

Confining attention to finitely repeated games in order to escape from the usual criticism folk theorem type of results are exposed to, one finds that it is sufficient to introduce (small) costs incurred by the upstream firm in case of switching to another downstream firm. Such costs may result from the necessity of modifying the production process due to any firm specific input requirements, from contractual negotiations or many other factors that are likely to affect the situation.

A third and probably most obvious alternative is commitment sustained by reputation effects. An upstream monopolist may invest in building up a reputation for exclusive-dealing as long as there is a sufficiently long future. We also point out briefly how various enforceable contracts can be used to implement mechanisms that enable the upstream firm to precommit to restrict output ex post.

This chapter concludes with a case study that demonstrates how TI Raleigh Ltd. defended their exclusive-dealing practices before court. Some of the arguments brought forward by the company can be linked to commitment devices presented in this chapter.

Given the crucial role of ownership right allocations one would expect these considerations to be vital in all industries. Moreover, industries in which the production process consists of various vertically related stages would appear to be key industries for the application of vertical integration theory. However, this apparently trivial statement does not seem to apply to vertically related industries which are regulated (or, rather, in much of the literature ownership has not been taken into account to the degree it deserves to be). There is, in fact, a large literature on regulated industries that, to a large extent, takes ownership structures to be given exogenously. Analytical difficulties may present one reason for this development, but other reasons may be present as well. There may be assumptions that have been used as substitutes for a thorough economic analysis. In particular, it has often been assumed that vertical separation may serve as alternative to regulation of intermediate goods prices and that, therefore, vertical separation is desirable. On the other hand, it has been assumed that economies of scope outweigh inefficiencies such as discriminatory practices of an integrated firm and that, therefore, vertical integration is to be favoured. Moreover, the fact that in the US as well as in several European countries decisions about ownership structures have already been taken may have further discouraged a detailed thorough analysis of ownership in regulated industries. Chapter III of this thesis attempts to emphasise the role of ownership in regulated industries and to show how the implementation of particular ownership structures can reduce the regulatory burden considerably. The interplay between regulation and liberalisation has long been recognised and corresponding steps have been taken by regulatory authorities throughout the world. The interplay between regulation and ownership structures may deserve attention to a similar degree, but most certainly it deserves more attention than it has been granted in much of the literature of this field of economics.

But, surprisingly, the above criticised feature of economic analysis of regulated industries may be found to be symptomatic for a much more general tendency in this field, namely a lack of examining a question one would have thought to arise naturally. This question is simply in how far features that are characteristic for regulated industries prevent general results to apply. Instead of identifying an industry's particular segments needing regulation and proposing adequate remedies for these very segments there appears to be a tendency towards assuming that market failure, or at least the need for regulation, automatically extends to every segment of the industry. While it may be true that there are 'spillover' effects between different industry segments the emphasis of regulatory effort should be on the problematic industry segment. Once this segment has been regulated satisfactorily the other segments should be left to competition which suggests that liberalisation is the suitable policy for all potentially competitive segments. The following quotation may serve as an example for a different view often held: 'A discussion of an access rule without reference to the rest of the regulatory environment has limited interest. The quality of an access pricing rule depends on the determination of prices for the final products.' (Laffont and Tirole (1994b) p.27). We are inclined to argue, on the other hand, that regulation in segments other than the bottleneck may be the legitimate result of complications arising in these segments[3] like, for instance, high fixed costs, but one should not start the analysis from the outright assumption that all segments need to be regulated.

Our focus is on regulated industries in which production relies on the use of an essential facility, often a network. In telecommunications, electricity, gas and water supply, networks are needed for the production or delivery of the final product to the customers. The essential feature of such industries therefore is (not only that the production process involves several stages which are vertically related but) the fact that an essential input, namely network access, is provided by a monopolistic supplier. As long as the network operation involves considerable economies of scale network duplication is socially undesirable and monopolistic provision should be maintained. *This* is the distinguishing feature of network related industries. Other potential differences such as final goods price regulation, rate-of-return regulation, entry restrictions or assistance etc. are superimposed by regulatory regimes rather than inherent industry characteristics.

Historically, these industries tended to be organised as state-owned monopolies. Efficiency considerations led to privatisation in many countries. In order to prevent the now privatised monopoly from ex-

[3]Any further complications, on the other hand, may require further regulatory tools.

ploiting its position as sole supplier it has to be constrained by regulation, possibly in conjunction with liberalisation, i.e. the introduction of competition to the non-monopolistic segments.

But in designing a suitable regulatory regime the emphasis should be placed on the simple goal of regulating effectively the very domain that is affected, in our case, access provision. Once this is achieved the essential intermediate good is provided at prices that are socially desirable all other aspects of the industries will in no way need to be treated differently from other industries (in which production involves stages that are vertically related). Instead of designing an ever increasing number of regulatory tools most of which interfere with each other in highly non-trivial ways one would have to attempt a satisfactory regulation of the very aspect that distinguishes these industries from such where one would fully rely on market forces. This key aspect is the pricing of network access in a non-discriminating and socially desirable way.

These considerations, however, must not be overemphasised. We will describe access pricing rules, for example, that make sense only if final goods prices are regulated. Our point is that instead of assuming informational asymmetries to be too vast to allow for any successful regulation of the bottleneck per se one should identify ownership structures that minimise, or even eliminate, these problems. Once such ownership structures are found one can attempt to develop a regime of minimal regulation in the sense that regulation is, as much as possible, confined to segments that cannot be run efficiently in the absence of it.

Having identified the existence of (and the necessity of access to) a bottleneck segment as the particular industry characteristic deserving regulation, the overall problem is, of course, far from being resolved. Even in a setting of complete information about costs the determination of an optimal access price is not trivial. Allocative efficiency is achieved when marginal cost pricing is implemented which, however, would not allow the network operating firm to break even due to considerable fixed costs involved in the operation of the network. Covering the fixed costs by a transfer payment from the state would imply that the particular industry is subsidised by consumers of other goods or by income earners generally. Distortions of one kind or another will result and be more or less undesirable depending on the way in which the subsidy is raised (shadow costs of public funds). Two-part-tariffs, firstly, have potential to foreclose low-volume users from the market and, secondly, may encourage inefficient resale. A menu of two-part-tariff contracts can only ease these problems but not abolish them in principle. Second best (Ramsey) pricing, on the other hand, allows the access provider to break even but imposes excessive informational requirements on the regulator as all elasticities (including cross elasticities) enter the calculation of the optimal prices. Therefore, theorists as well as regulators have turned to cost based pricing rules. But pricing at average costs or fully distributed costs not only involves thorny issues related to cost allocation (especially for multiproduct firms) but also destroys allocative efficiency and gives little incentives for cost minimisation. Nor can rate-of-return regulation be an ideal solution because this would imply productive inefficiency due to minimal incentives for the firm to undertake (effort or any other) non-contractible investments with a view to achieving productive efficiency.

This problem has contributed to the relative popularity of the efficient component pricing rule (ECPR) or Baumol-Willig rule (see Baumol and Sidack (1994) and Willig (1979)) which states that the access price should be set equal to the sum of the direct costs of providing access and the opportunity cost to the incumbent. [4] Based on a contestable markets framework this concept assumes that, due to potential competition, final goods prices are at the competitive level. Although, under this assumption, the ECPR is optimal this rule has been criticised for various reasons. If the assumption about optimal regulation of final goods prices does not hold the ECPR simply protects the incumbent's monopoly rent. Further, it abstracts from incentive issues

[4]This concept, obviously, makes sense only in a setting with an integrated incumbent because a network operator not active in the final market would not incur any opportunity costs apart from production costs.

such as the incumbent's incentive to inflate the marginal cost of access provision. Further, adequate incentives for cost reduction exist only if there is a sufficiently long time lag between cost reduction and adjustment of the pricing rule. Laffont and Tirole (1994b) argue that the ECPR is not optimal in case there are asymmetries between incumbent and competitor such as brand loyalty (captive customers), cost asymmetries, etc. But, as Armstrong and Vickers (1995) show, this criticism applies to the simplest form of the ECPR ('margin rule') only, but not to the general interpretation using the opportunity costs in a general sense. They demonstrate how the opportunity cost term has to be adjusted in order to take into account various complications (bypass, no contestability, relaxed assumptions on production).

The paper by Laffont and Tirole quoted above suggests a further alternative. It is pointed out there that, under a global price cap, i.e. including access charges, a firm has incentives to charge Ramsey prices. Under this regime, therefore, the task of determining elasticities etc. is delegated to the firm that may not have complete information but it certainly has more than the regulator and may therefore engage in some form of price discrimination. More importantly, they argue, the problem of regulatory capture that, due to unknown elasticities, results from the regulator's discretion in fixing the access price, disappears. This usage based approach is certainly attractive in that it implements approximately efficient prices, a feature that all approaches that are based merely on costs must lack. But caveats such as remaining informational requirements about demand and the price cap concept based on weights given by exogenous demand forecasts [5] remain.

However, there is a serious drawback which all access pricing concepts discussed so far have in common. All these concepts are based, in one way or another, either on network operating costs or, as in Laffont and Tirole (1994b), on a profit constraint. But these costs or profits are generally not entirely known to the regulator because, due

[5] A similar criticism applies, of course, to any kind of price cap regulation per se because price caps are taken to be exogenous but will, in general, depend on past performance.

to accounting and cost allocation procedures, considerable degrees of freedom may be involved. The network operating firm not only has a considerable informational advantage but also an incentive to exploit this asymmetry and to extract as high an informational rent as it possibly can. Not only does such a strategy increase revenues but it also creates an asymmetry between the incumbent and other firms in the downstream market by raising rivals' costs.

Summarising our remarks so far we can state that, firstly, access pricing is a key problem in network related industries and successful regulation of access provision would pave the way for a policy of liberalisation in all other, i.e. potentially competitive, areas of the industry in question. Secondly, the determination of the access price that should be implemented is a rather complicated issue, theoretically as well as practically, especially because of the informational asymmetries that may highly compound the difficulties outlined above. These two facts can be regarded as the starting point of the chapter on network access (chapter III).

In a setting in which access prices are based on production costs we focus on the network operator's incentives to exploit his informational advantage. These incentives, we argue, depend on the ownership structure that has been implemented in the industry. In particular, we compare vertical integration (with and without accounting separation) of the access provider with one downstream firm, vertical separation, and various forms of joint ownership. While the welfare ranking between some of these ownership structures depends on industry characteristics we establish the existence of one particular ownership structure (namely joint ownership in conjunction with an optimal capital adjustment rule) which always yields efficient outcomes in the sense that the access price fixed by the regulator will not be distorted at all by informational asymmetries. The particular joint ownership concept we suggest is designed to annul entirely any incentives to distort information about costs. Following the spirit of the above introduction, namely that liberalisation is adequate once the access problem has been resolved we do not assume regulation of final goods prices, entry assistance or restriction or other policies along these lines.

There are issues which, although not directly related to access pricing, are vital in this context. For instance, there may be (non-verifiable) investment decisions to be taken by the network operator that would generate a surplus part of which accrues to downstream firms. Similarly, positive externalities result from investments that are taken by individual downstream firms because such investments may affect the profit of the upstream firm and of downstream rivals. We include these issues in our analysis and demonstrate that the ownership structure that is optimal with respect to access pricing also performs best as far as investments are concerned. In regulated industries the case for joint ownership is, therefore, twofold: fair access provision and efficient incentives for investments.

Entry, or rather entry deterrence, is another issue we analyze in our model. Again, joint ownership with optimal adjustment of capital shares performs best, the reason being that incentives for entry deterrence vanish entirely.

Note again the difference in spirit between our entry analysis and much of the regulatory literature. While, following the paradigm of liberalisation after successful access price regulation, we measure welfare as (possibly weighted) sum of consumer surplus and producer surplus which, under the assumptions of our model, amounts to aiming at low prices of the final goods. In other words, the downstream industry should be as competitive as possible and entry is therefore regarded as desirable which, in turn, implies that we favour ownership structures that give least incentives for entry deterrence. (De-)regulation literature, on the other hand, often includes entrants' fixed costs in the welfare function which paves the way for excess entry results. This procedure is, of course, perfectly legitimate as long as the results are applied to industries in which entry sunk cost play an important role in the downstream (or non-monopolistic) sector. Our analysis, on the other hand, focuses on industries with a potentially competitive segment in which entry sunk costs are small and entry is desirable. Assuming that entry sunk costs are relatively small implies that entry restrictions in competitive segments of regulated industries are as unnecessary as entry restrictions for any other industries that are considered to be competitive. This difference in approach only illustrates what has been stated earlier. Having realised that industries with a naturally monopolistic segment need to be regulated economists seem to infer that regulation must necessarily be extended to all segments of the industry. However, if some segments of the industry in question are potentially competitive one would expect the introduction of competition into these segments to be the suitable objective.

In other words, we argue that pursuing the aim of *optimal* regulation may well lead to *over*-regulation. The end in view should rather be *minimal* regulation that simply achieves the task of preventing the firm acting as supplier in the monopolistic segment from exploiting its market power, not more and not less.

There are, of course, qualifications to be made. Once we leave behind the idealised world of the economic model it may be less clear whether or not access price regulation has been successfully implemented to such a degree that no distortions arise. Real life may confront regulators as well as other parties involved with effects not taken into consideration in the model. Further informational or other imperfections may necessitate further regulatory tools. Without denying the relevance of such effects we rather want to make the point that regulation should concentrate on the efficient provision (and pricing) of bottleneck facilities and should then introduce competition into 'potentially competitive' segments.

Another issue tackled in this chapter may be mentioned here by way of introduction. Discussing the question of a monopolists's vertical structure it has been argued that the costs and benefits of a break-up are hard to be quantified and the former may or may not outweigh the latter. Economists have generally hesitated to take an unambivalent stand on the issue and regulatory bodies have taken different decisions on this matter. However, it has also been argued in the literature that structural separation may be closely (or even fully) substituted by a separation of accounts. Accounting separation has been assumed to ban discrimination and therefore to procure the benefits of structural separation without imposing the costs thereof on the industry (or on anybody else). But this conclusion can only develop due to a lack of consideration for implications of ownership. We show in our chapter on network access that the argument outlined above does not hold. On the contrary, an imposition of accounting separation may be completely neutralised by the incumbent firm's strategy. The crucial step in the argument is simply that ownership matters. Residual rights exist and depend on ownership of assets and not on published accounts. Again, one of our central points can be subsumed under the principle outlined earlier. The occupation with specific industries or classes of industries may bear a danger of neglecting to take into account principles that apply more generally.

Another central point that is related to the issue is that a comprehensive study of access pricing must aim at taking into account features such as asymmetric information and ownership. There are various studies that take into account either information or ownership which make significant contributions to the theory but fail to analyze the interplay between both factors. Relatively few studies deal with both issues at the same time. Although this may involve analytical complications it is worthwhile to analyze both issues in one model because ownership affects considerably the incentives to exploit informational asymmetries. Again, this is simply a special case in which a more general insight must be taken into account, namely that, in a world of incomplete contracts, ownership affects incentives.

Chapter IV 's focus is similar to the one of chapter III, namely vertically related industries in which a bottleneck may exist. Many of the introductory remarks apply to both chapters. However, it is not exactly a generalisation or continuation of the preceding one. It rather sheds light on related issues in a slightly different framework.

We use an alternative model which allows to take into account other settings including more than one upstream firm. Apart from vertical integration and vertical separation as defined in chapter III we consider the case of an upstream competitor that may or may not be integrated with a downstream firm.

A price setting game is analyzed which allows to draw conclusions as to the import of different classes of contracts or tariffs allowed for. Further, we analyze other issues relevant to this setting such as the impact of a monopolistic input provider exerting price discrimination between different downstream firms. We also point out how a change in the timing structure affects the outcome. If access prices are chosen before entry occurs then more entry and lower prices will result. A long-term commitment of access providers may reduce hold-up problems and therefore constitute a suitable aim for regulatory policy. Entry of a competitor that is more efficient than the incumbent does not imply that the final outcome is more efficient. Implications of ownership structures for pricing and entry decisions are analyzed and compared.

This chapter does not model explicitly informational asymmetries (although, as we argued in detail, this is a relevant point). The idea rather was first to examine a setting without regulation but with features of regulated industries as far as the number of firms and the vertical relation of markets are concerned. Starting from this free market analysis of equilibrium pricing and entry decisions, we then explore the relation between ownership and the degree of regulation required in order to ensure efficient outcomes. At this stage then, we can draw conclusions as to the informational requirements for regulation. In a setting in which, without any kind of regulation, the efficient access price tends to be charged in equilibrium, the informational requirements are obviously very small and, even without making asymmetric information explicit in the model, one can argue that such a setting lends itself for regulated industries.

Our final chapter (chapter V) addresses a rather fundamental issue that is relevant to the theory of competitive markets. The focus of this chapter is on the demand side and addresses the question in how far human (choice) behaviour lends itself to a formal mathematical description and, more importantly, in how far results derived on the basis of strict behavioural requirements are robust against a relaxation of these assumptions.

Traditionally, economic theory has assumed that individuals are able to order all available alternatives according to their tastes, i.e. preferences, and always select the very best alternative. In consumer theory alternatives are usually bundles of goods, preferences are assumed to satisfy regularity conditions such as continuity, completeness and, in most cases, transitivity and convexity, and all individuals are rational in the sense that they choose the bundle which is optimal according to their preferences.

An extreme example for strands of economic theory relying on the utility (or preference) maximisation paradigm is the theory of competitive markets or, from now on, general equilibrium theory as first formalised by Debreu (1959). Not surprisingly, this field of economics also stands out for a remarkable precision and theoretical appeal of its results. Agents maximise utility independently, firms maximise their own profits, all agents act in a purely self-interested way without even attempting to coordinate their actions but, despite of all this, there exists an equilibrium price vector which, if implemented by the 'Walrasian auctioneer', leads to simultaneous clearing of all markets. Moreover, the resulting allocation of the competitive equilibrium turns out to be Pareto optimal. These results can be obtained for arbitrarily many agents, goods and, in the case of a multi-period setting with uncertainty, even for arbitrarily many periods and states of nature. Conversely, any Pareto optimal allocation can be obtained as equilibrium outcome of free trade simply by ensuring an adequate distribution of initial endowments prior to trade.

Undoubtedly, general equilibrium theory is a field of economics in which most powerful tools have been used and most precise results have been obtained in return. But there can be as little doubt that these very results rely on equally precise assumptions, particularly about individual choice behaviour. It may, for instance, seem demanding to require of all agents the ability to order all bundles of goods, especially in a modern market economy in which many consumers would not even be aware of the existence of numerous (perhaps lately developed, or imported) goods. The assumption that agents know and take into account all states of nature that may occur in any future period of the relevant time horizon and optimally transfer income between these states and periods seems strong if not heroic.[6] In recent years it has been argued that individual choice behaviour is rather marked by inconsistencies even in extremely simple situations. Various avenues have been pursued to explain this kind of human choice behaviour all of which abandon the assumption that agents can or do maximise preferences or at least they assume that agents may not choose exactly the optimal alternative.

This scenario begs the following question. In how far are precise results derived in equilibrium theory sensitive to deviations of human behaviour from the ideal world as defined by the assumptions of many theoretical models and, in particular, by the utility maximisation paradigm. General equilibrium theory has been criticised for both, its assumptions as well as its predictions. It therefore appears to be a prime candidate for a sensitivity analysis that would attempt to shed light on the questions how sensitive results derived in general equilibrium theory are and which of the results are most sensitive. The latter question is important because a break-down of some of the less plausible predictions in case the utility maximisation paradigm is substituted by alternative assumptions would be an argument in favour of a substitution along these lines whereas it would be rather worrying if predictions that are in line with empirical observations can no longer be obtained.

In order to provide some basic insights along these lines we proceed as follows in chapter V. We abandon the assumption that all individ-

[6] As pointed out earlier, this fact is one of the motivations for incomplete contracts.

uals' choices necessarily result from maximisation of preferences. Demand is modelled as stochastic and may result from moods, random influences, or some coincidence that may induce an agent to buy a certain good. On the other hand, we allow for choices to be the result of state dependent preferences in order to capture the other extreme of the whole spectrum of rationality. We start from a general equilibrium model in which utility functions or, equivalently, continuous preferences, are substituted by demand densities defined on the budget hyperplanes. As one would expect, in this model markets generally do not clear. We then define an expected equilibrium in which expected total excess demand equals zero. Such equilibria exist but are not unique unless some rather restrictive assumptions are made about the agents' consumption distributions. As markets clear on the average only and not for actual realisations of demand (i.e. with probability one) we introduce rationing into the model and argue that rationing is probably a more realistic result than precise market clearing. The amount of rationing, however, is shown to become small for economies with many agents. For any finite size of economy we can give a bound for the probability that the amount of rationing for a certain good exceeds a given limit.

All results outlined above, as well as some other results obtained in the final chapter, appear to point into the same direction in the following sense. The predictions of general equilibrium theory which are not robust against a relaxation of utility maximisation are the least plausible ones. Those that obtain in some modified way become more, rather than less, plausible due to these modifications. And those predictions that obtain in an unmodified form are the most obvious where, for the moment, obvious is defined by 'in harmony with empirical observation'. There are even new results that can be obtained due to the assumption of random demand and that may serve as motivation for assumptions made in some areas of economic analysis. An example for this latter type of (new) results would be rationing that induces agents to (randomly!) buy and sell assets for liquidity reasons which is exactly the assumed behaviour of 'noise traders' in the financial markets literature.

More realistic in the sense of more descriptive assumptions may be less convenient for formal analysis and may lead to less precise and less ideal results. The payoff of using them nonetheless lies in a higher degree of realism of the predictions obtained under these assumptions.

The statements made above, in our view, imply neither that general equilibrium theory is wrong or useless nor that it cannot be successfully applied to various economic questions. Our interpretation of the sensitivity analysis outlined above is rather that general equilibrium theory provides results which are not only theoretically attractive due to their content and precision but also a useful borderline case from which economic analysis may start. To find more final answers to specific problems, however, it may prove useful to relax some of the assumptions that are commonly made, especially those that concern the least calculable, precise, and predictive element contained in it, namely human behaviour of choice.

CHAPTER II

Market Foreclosure Without Vertical Integration [1]

Abstract

This chapter challenges the assumption that vertical integration is the sole way for upstream and downstream firms to monopolise the downstream market. Future trade, switching costs, and reputation may constitute alternatives for vertical integration and lead to dominant outcomes.

1 Introduction

Following our remarks on the role of ownership in the general introduction this chapter analyses the plausibility of a particular contractual assumption used in a paper by Hart and Tirole (1990), (H-T) from now on, on the behaviour of integrated and non-integrated upstream and downstream firms.

In the above paper it is demonstrated that vertical mergers can have anticompetitive effects and that there are foreclosure incentives leading to vertical integration. Many of the conclusions reached are certainly interesting and shed light on questions that are of practical interest. The aim of our analysis is not to question the possibility of anticompetitive effects arising from vertical integration, but rather it is to show how foreclosure and monopolisation can be achieved with the help of contracts which are not based on profit sharing and which are

[1] A summary of the main ideas of this chapter can be found in Hardt (1995).

not necessarily enforceable. In fact, the results of the ex-post monopolisation version of the model by Hart and Tirole are, to a large extent, driven by the assumption that nonenforceability of exclusive-dealing contracts implies the inability of firms to commit to monopolisation.

While, on the one hand, their analysis provides good arguments for foreclosure motives underlying vertical integration it also presupposes, on the other hand, that non-integrated upstream and downstream firms cannot commit to monopolise the downstream market.

This supposition, however, is vulnerable theoretically as well as practically. Theoretically, this assumption implies that a monopolist cannot behave as a monopolist given the sequential structure of moves in the model. Having sold output to consumers whose reservation price exceeds the monopoly price he will, due to ex post incentives, sell additional amounts to other agents. Practically, the supposition incurs problems when confronted with casual evidence. Frequently, cases have been made against upstream firms because of their refusal to supply downstream firms. In many of these cases, however, the upstream firms are not integrated with downstream firms but, still, they stick to agreements made with certain downstream firms not to supply others. Special attention will be drawn to an antitrust case against Raleigh in 1981.

We describe the setting in section 2. Section 3 contains models of some scenarios in which non-integrated firms are able to commit to monopolising a downstream market. Section 4 discusses briefly some implications. A case study is analyzed in section 5. Section 6 concludes this chapter.

2 The Commitment Problem

The basic features of the model by Hart and Tirole are the following. The concept of ownership is the one based on residual rights (under incomplete contracts) along the lines of Grossman and Hart (1986). There are two potential upstream firms, U_1 and U_2, and two downstream firms, D_1 and D_2. Firms do not know ex ante which type of intermediate good will be traded. Vertical integration between U_i and D_i allows for profit sharing (which would not be feasible otherwise). Costs of integration are summarised by a fixed amount E. There is competition in the downstream market for final goods. The intermediate good is produced by U_1 and U_2 at constant marginal costs c_1 and c_2.

Here, we focus on the special case of the ex-post monopolisation variant of the model when the less efficient upstream firm has infinite marginal costs. [2] In this framework, (H-T) show that, in spite of other claims that had been made, U_1 (U from now on) might have incentives to merge with a downstream firm in order to monopolise the downstream market. Our point is to show how monopolisation can be achieved even without integration.

The reason why this is not possible in H-T's analysis is the following: U, it is argued, after selling the monopoly quantity q^m to a non-integrated downstream firm, D_1 say, for a lump-sum payment equal to monopoly profit, π^m, would always have ex-post incentives to supply D_2 as well. D_1, anticipating U's cheating would not sign an exclusive-dealing contract [3] in equilibrium. Analogously, after supplying each downstream firm with a share of the monopoly quantity for the respective share of the monopoly profit, the upstream firm would have incentives to supply further amounts. Again, in equilibrium, no firm would accept such an offer.

We wonder in how far this argument applies to practical situations. Firms considering vertical integration usually have a long history of mutual trade relationships and, more importantly, they intend to en-

[2] Hart and Tirole analyze mergers between a relatively efficient upstream firm, U_1, and a downstream firm, D_1. In their constant returns to scale framework this means $c_1 < c_2$. So the special case corresponds to $c_2 = \infty$. But our point could also be made for the general case c_1, c_2, with $c_2 < \infty$. The difference is simply that the efficient upstream firm, U_1, then must commit to supply a positive but low quantity to other downstream firms, i.e. the quantity that, otherwise, would be supplied by the inefficient upstream firm, U_2.

[3] The underlying assumption here is that cheating is observable but, for legal or for informational reasons, it is not verifiable.

gage in future trading. Thus, a one-shot game seems to be of limited applicability in this context. In the following section we will explore several ways to take these effects into account.

3 Commitment in Dynamic Settings

As indicated in the previous section we hold effects arising from long-term relationships to be important in these situations. Cheating today most certainly affects trading tomorrow. In what follows we demonstrate this view with the help of some simple game-theoretic models, and contracts.

3.1 Possibility of Future Trade

One of the most straight-forward ways to think about the situation would be the following setting: in any given period firms expect that, with probability α, there will be trade between the respective firms in the following period. Firms maximise their future expected payoff discounted at rate $\tilde{\delta}$. This setting corresponds [4] to an infinite horizon game with discount factor $\delta = \alpha \cdot \tilde{\delta}$.

Consider the following situation: suppose that, in the industry in question, it has been common for the efficient upstream firm to trade the Cournot quantity with each of the downstream firms. However, U suggests to monopolise the downstream market.

In period t, U proposes to D_1 to monopolise the market, i.e. sign the contract (q^m, π^m). Consequent to this offer D_1 can either accept ('A') or buy the Cournot quantity, q^c, for a lump sum payment equal to Cournot profit, π^c ('C'). If D_1 has played 'A', U decides to cheat 'Ch' and receive π^{ch}, or to stick to the monopolisation contract, i.e. to be honest ('H'). If D_1 has not accepted the monopolisation contract, the Cournot quantities are traded.

We further assume that this game is common knowledge and that players have perfect information about past moves.

[4] Here, we assume risk neutrality of firms.

In this context the following pair of strategies form a subgame-perfect equilibrium provided δ is large enough:

D_1: play 'A' in every period if no cheating has occurred so far, but always play 'C' if cheating has occurred ('grim' strategy).

U: play 'H' in each period.

U's continuation payoff in period t is given by

$$\pi(\text{'H' in all periods}) = (1-\delta)\sum_{\tau=t}^{\infty}\delta^{\tau}\pi^{m} = \delta^{t}\cdot\pi^{m}.$$

But for deviation in period t we obtain

$$\pi(\textit{'Ch' in period } t) = (1-\delta)[\delta^{t}(\pi^{m}+\pi^{ch}) + 2\sum_{\tau=t+1}^{\infty}\delta^{\tau}\pi^{c}].$$

As one can easily verify U is better off playing 'H' if the following condition holds:

$$\delta > \frac{\pi^{ch}}{\pi^{m}+\pi^{ch}-2\pi^{c}}.$$

It is important to notice that, as $\pi^{m} > 2\pi^{c}$ (monopoly profits exceeds sum of cournot profits), there always exists a $\delta < 1$ that satisfies the above condition.

The interpretation of this condition is straight-forward: the restriction on δ becomes stricter when cheating becomes more worthwhile or when, ceteris paribus, the Cournot profits increase or monopoly profits decrease.

For D_1 it is an equilibrium to accept the monopolisation contract because, given U does not cheat, it has no incentives to deviate by playing Cournot.

The subgame-perfect equilibrium we have thus established faces, of course, the counter argument that it is not unique, in particular, it would also be a subgame-perfect equilibrium for U always to cheat and for D always to play Cournot. But note that the monopolisation equilibrium is on the Pareto frontier and, given some degree of rationality of firms, it is more likely to be selected than an inefficient equilibrium.

This is an argument often used in cartel theory. See, for instance, the paper by Porter (1983). [5]

3.2 Switching Costs

Maintenance of mutual trade relationships might be important because of other effects. Some trade relationships are only worthwhile if maintained in a minimum number of periods. We formalise this idea in the following game.

Fixed switching costs

Suppose U incurs a fixed cost c of supplying a further firm with the input good. [6] Suppose further that this cost is such that U would not make this investment if there was just a single period in view in which U could supply the extra quantity to D_2 (i.e. $c > \pi^{ch}$). A much weaker condition will be provided in the end of this subsection.

This assumption which appears to be a rather plausible for a number of cases is sufficient to destroy the standard backward induction argument that would yield cheating as the unique subgame-perfect equilibrium in a finite repetition of the stage game.

Consider the following game that is to be played in period t, $t = 1, \dots T$, $T < \infty$. The strategies of U and D_1 are the same as in the previous section with the only modification that U has to pay c if it wants to supply D_2. Then the continuation payoff for U in any period $t = 1, \dots, T$ is given by

$$\pi(\text{'H' in all periods from t on}) = (1-\delta)\sum_{\tau=t}^{T} \delta^{\tau}\pi^{m}$$

[5] In a more extensive analysis one would have to consider the possibilities of the upstream firm to use the threat of making an exclusive-dealing contract with D_2 if D_1 does not accept. As it would be an equilibrium for D_2 to accept this offer this threat is credible and gives another argument for D_1 to accept the offer.

[6] One could, for instance, think of c as the cost of designing a machine for the production of the particular input good needed by a downstream firm.

But for deviation in period t we obtain

$$\pi('Ch' \text{ in period } t) = (1-\delta)[\delta^t(\pi^m + \pi^{ch} - c) + \sum_{\tau=t+1}^{T} \delta^\tau 2\pi^c].$$

In period T, it is certainly optimal for U not to cheat. Now consider period t, $t < T$. Given it is optimal in periods $t+1, \dots, T$ not to cheat the payoff for U in period t is as given above. Thus we obtain the following condition for U not to cheat in period t:

$$\frac{1-\delta^{T-t}}{1-\delta}(\pi^m - 2\pi^c) > \frac{\pi^{ch} - c}{\delta}$$

which is, in fact, true for every $0 < \delta < 1$ given our assumption on c and the fact that $\pi^m - 2\pi^c > 0$ always holds.

From a game-theoretic point of view this result is, of course, rather obvious. But the point we want to make is an economic one: the specific investment is undertaken only once and it only exceeds the one-period profit from cheating, but not the discounted sum of profits from future cheating. The reason why this investment is not undertaken is not that it would not be profitable in itself but that the other downstream firm will punish the upstream firm in subsequent periods.

Another way to interpret this model would be to think of rental of technology: U can produce a technology for D_2 at cost c and then rent it for a fixed payment each period. Given U and D_1 have agreed to monopolise the market, the maximum payment D_2 will be willing to make for rental of the technology in period t is the profit from cheating in that period, π^{ch}. But, because of $c > \pi^{ch}$, U would not produce the technology for D_2 if rental could be agreed upon for a single period only. Thus the usual backward induction argument does not apply here. This case of supply of a more efficient technology to one of the firms might be one in which monopolisation is more plausible than in other cases: supply of a particular technology might be directly observable, not only via price reaction on the final product market.

In the situation modelled above 'cheating' by U can be excluded even though the game is finite and the discount factor may be arbitrarily small.

Initially small switching costs

As indicated earlier the condition $c > \pi^{ch}$ can be replaced by a much weaker condition if switching costs do not have to be constant over time. Denoting switching costs in period t by c_t, U will cooperate in period $T-t$ if

$$c_{T-t} > \pi^{ch} - \left(\sum_{\tau=1}^{t} \delta^{\tau}\right)(\pi^m - 2\pi^c) \quad , \; 0 \leq t \leq T.$$

Using the convention $\sum_{\tau=1}^{0} = 0$, this condition implies that $c > \pi^{ch}$ is necessary in the very last period only. In earlier periods switching costs may be smaller. Arbitrarily small values of switching costs are sufficient for any given value of π^{ch}, for any $\delta > 0$, and for any $T \geq 1$ it is optimal for U to cooperate in each period, given $(\pi^m - 2\pi^c)$ is sufficiently large ('monopolisation worthwhile').

The condition on c_t is weaker for larger values of T or δ. Note that c_t might be negative initially. A nice interpretation for negative switching costs would be side payments made by the other downstream firm or legal costs resulting from supply refusal. Even in the presence of such additional incentives to deviate from cooperative behaviour U will cooperate for the sake of future profits.

3.3 Reputation Effects

The idea formalised in this section is that the upstream firm U might have incentives to build up a reputation for being honest and that the downstream firm D_1 might have incentives to build up a reputation for not rejecting monopolisation contracts. Instead of proving the existence of a sequential equilibrium as in Kreps et al. (1982) we will use a simplified version similar to the one in Tirole (1988 p. 256 f). The reason for this is that this seems to be the simplest framework which allows to make the following point: we will show that, even in a finitely repeated game, it is sub-optimal for both firms to behave non-cooperatively from the first period on, given that they face a sufficiently long future and that the discount factor is sufficiently high.

In what follows we are going to assume that there is a small amount of uncertainty about the agents' types. The intuition for this assumption is that an upstream firm is not sure about whether a downstream firm trusts the offer to monopolise the market. A downstream firm, on the other hand, faces uncertainty in so far as the upstream firm might either stick to an exclusive-dealing contract or cheat by supplying additional amounts to other downstream firms. If one assumes the latter to be very unlikely this can be reflected by arbitrarily small probabilities.

The stage game of section 3.1 is now to be repeated in a finite number of periods t=1,...,T. We introduce the following possible types of agents. The payoffs given in section 3.1 are valid if the players are 'sane' which happens with probabilities $1-\alpha_U$ and $1-\alpha_D$ for U and D_1, respectively. With probability α_i, player i is 'crazy' in the sense of deviating from the norm as defined by the payoffs of the stage game. To simplify terminology we will say that U cooperates if it does not cheat and that D_1 cooperates if it accepts the contract. For player i the crazy type plays the grim strategy:

- cooperate in period 0

- cooperate in period $t \geq 1$ unless player j has not cooperated in any of the previous periods.

The discount factor will be denoted by δ. In what follows we need to assume that a downstream firm has nonzero bargaining power [7] and thus receives a positive share of the profits. We will thus denote U's bargaining power by β, where one might want to think of β close to one because of U's monopoly position as sole supplier in the market. In what follows we analyze the behaviour of the sane type in the presence of uncertainty about the other player's type.

If U cheats in period 0 this implies that he is sane. In this case there will be no cooperation in any period because D_1 will retaliate if crazy and will not accept the contract anyway if sane. The only equilibrium in this case is that the Cournot quantity is traded in all subsequent

[7] This assumption can easily be justified by taking into account delays caused by disagreement in the bargaining process.

periods until T. The payoff for U is then given by

$$\begin{aligned}\pi_U(\text{'Ch' in period }0) &= \alpha_D\beta[\pi^m + \pi^{ch} + 2\pi^c(\delta + \cdots + \delta^T)] \\ &+ (1-\alpha_D)2\beta\pi^c(1 + \ldots + \delta^T).\end{aligned}$$

We now show that U can do better by not cheating in the first period. Suppose, for a moment, that U uses the grim strategy: 'cooperate until T or until cheating occurs'. The worst case for him then would be that D rejects the contract even in the very first period. This case provides a rough lower bound on U's payoff which is sufficient to derive a condition for cooperation, provided the indicated restrictions on the parameters are satisfied. Using

$$\pi_U(grim) = \alpha_D[\beta\pi^m(1+\delta+\cdots+\delta^T)] + (1-\alpha_D)2\beta\pi^c(1+\delta+\cdots+\delta^T)$$

and simplifying the algebra yields that $\pi^U(grim) > \pi^U(\text{'Ch' in period }0)$ if

$$\frac{\delta - \delta^{T+1}}{1-\delta} > \frac{\pi^{ch}}{\pi^m - 2\pi^c} =: K. \tag{1}$$

For $\delta \to 1$, the left hand side of (1) converges to T so that the condition is satisfied for T sufficiently large. For $T \to \infty$ the we obtain the same condition as in the infinite horizon case (3.1)

$$\delta > \frac{\pi^{ch}}{\pi^m + \pi^{ch} - 2\pi^c}.$$

More interestingly, for any given $\delta \in (\frac{K}{K+1}; 1)$, we can give a minimum number of periods $T_1(\delta)$ such that, given $T > T_1(\delta)$, it is not optimal for U to cheat before period $T - T_1(\delta)$:

$$T_1(\delta) \geq \frac{\ln\left(1 - \frac{1-\delta}{\delta}K\right)}{\ln\delta}, \ \delta < 1.$$

The right hand side of this expression is positive for all $\delta < 1$, and the logarithm has a positive argument for $\delta > \frac{K}{K+1}$. Comparative statics yields the expected results.

A similar analysis holds for the downstream firm. Accepting the contract in period 1 implies the risk of making a loss, $-l$ say, in this period. In subsequent periods then the Cournot quantity is traded, i.e. D_1 obtains $(1-\beta)\delta^t\pi^c$ in period t. As a condition for D_1 to accept the contract one then obtains

$$\begin{aligned} \alpha_U & \ [(1-\beta)\pi^m(1+\delta+\cdots+\delta^T)] \\ + & \ (1-\alpha_U)[-l+(1-\beta)\pi^c(\delta+\cdots+\delta^T)] \\ > & \ (1-\beta)\pi^c(1+\cdots+\delta^T), \end{aligned}$$

which simplifies to

$$T_2(\delta) \geq \frac{\ln\left(1-\frac{1-\delta}{\delta}\tilde{K}\right)}{\ln\delta},$$

for

$$\tilde{K} = \frac{\pi^c + (1-\alpha_U)\tilde{l} - \alpha_U\pi^m}{\alpha_U(\pi^m - \pi^c)},$$

where $\tilde{l} = l/(1-\beta)$. Again, comparative statics results can be easily obtained.

So, for $T > \max\{T_1(\delta,\alpha_D), T_2(\delta,\alpha_U)\} =: M(\alpha_U,\alpha_D,\delta)$, the firms will cooperate in all periods $t \leq T - M(\alpha_U,\alpha_D,\delta)$.

We have thus established in this section that arbitrarily small probabilities for the existence of an honest upstream firm and a cooperating downstream firm are sufficient to induce cooperation in a finitely repeated game. Only when the number of future periods is too small for at least one agent i (given δ and α_j) cooperation breaks down. This result corresponds to the intuition provided in the beginning of this section, namely that in a world where cooperation is possible, even though unlikely, firms can achieve monopolisation of markets even though they might be using contracts which are not enforceable. Thus, vertical integration is not the only way round the firms' 'inability' to commit themselves to monopolise the market[8].

[8]Alternatively, one might think of a sequence of contracts where, in each period, U sets

Additionally, we would like to emphasise that commitment induced by reputation is not only an additional way to achieve monopolisation, but it is rather an outcome that even dominates vertical integration. The costs of integration, which are represented by a fixed amount of E in H-T's model are not incurred in the setting of this section . Thus, privately as well as socially, non-integrated cooperation dominates vertical integration.

3.4 Enforceable Contracts Replacing Nonenforceable Ones

Hart and Tirole, in their paper, assume that vertical integration and profit sharing are equivalent in the following sense. The former implies the latter and, due to possible profit misrepresentations etc., the latter is not attainable without the former. Starting from this notion of equivalence they conclude that the upstream firm is not able to restrict the output in the downstream market other than by integrating with a downstream firm. In contractual terms the main source of incompleteness of possible contracts to restrict outputs is nonenforceability. In this section we would like to point out that alternative contracts can be written which are enforceable and yield the required output restriction.

One possibility is to trade the input good for a price p_I that is conditioned on the price of the final good p_F in the downstream market: $p_I = p_I(p_F)$, provided the downstream market price is verifiable which will be the case in a large number of markets. If p_I depends crucially on p_F then the upstream firm has sufficient incentives not to supply other downstream firms. Demand uncertainty would not change this result qualitatively, but just lead to more complicated strategies, e.g. trigger-price strategies.

Another possible scenario is to take into account the possibility of bankruptcy. Here, payment occurs if and only if D_1 has not been

the price in a way that D_1 accepts the offer. By building up reputation, U can charge higher and higher prices each period. Ultimately, prices will converge to the monopoly price.

forced to exit the market because of big losses in the previous trading period. This might induce U to supply very little, if anything, to other downstream firms.

Alternatively, one could think of specific forms of payment as, for instance, payment in shares of the downstream firm.

These hints are just presented very informally here but this might suffice to indicate some commitment devices which arise when enforceable contracts are written to replace the nonenforceable ones.

4 Implications

The various counter arguments presented in the previous sections show that, in many settings, it cannot be taken for granted that a non-integrated pair of an upstream and a downstream firm cannot commit to monopolise the downstream market. To what extent does this affect the analysis presented by Hart and Tirole?

As output contraction occurs in Variant I ('ex-post monopolisation') of their model only the results in that section are concerned: it is no longer true that firms have to integrate in order to achieve output contraction in the downstream markets, and that non-integration implies that Cournot quantities are traded. Total profits are highest when firms are not integrated but behave cooperatively. The reason why the second ('scarce needs') and third variant ('scarce supplies') are not affected is that there, output contraction is not at stake: firms have incentives to integrate because this enables them to channel scarce input goods (or scarce orders) to the integrated firm that would be rationed otherwise. An important difference is that in variants 2 and 3 no additional profits are generated by vertical integration because output contraction does not occur. Thus, non-integrated cooperation is not equivalent with integration (even when we neglect the efficiency loss 'E''). Cooperation is likely to arise with a view to output contraction and generate additional profits, but not with a view to solving problems of scarcity of needs or supply where profits could be redistributed but not increased.

5 Empirical Evidence: a Case Study

> 'Several types of evidence are available to enlighten economists... One type of evidence is casual observation which, although not terribly scientific, is better than no observation at all.'
>
> Dennis W. Carlton (1989)

At this stage it might be interesting to investigate cases of supply refusal. An especially interesting case seems to be the one against TI Raleigh Industries Limited and TI Raleigh Limited in 1980. As a basis for our studies we largely rely on a report of the Monopolies and Mergers Commission (1981).

This case seems to illustrate the question we are addressing in this note. Raleigh had refused to supply several discount stores. But, interestingly enough, its distribution network consisted of bicycle dealers the vast majority of which (87 %) were not integrated with Raleigh.

Much of this report is concerned with questions of public interest and questions of applicability of certain competition laws, the relevance of market shares etc. But what we will focus on is the argumentation presented by Raleigh for their refusal to supply discount stores. The reason why, among the arguments presented by Raleigh, we will focus on the purely economic ones as effects on prices, sales, profits etc. is that we are concerned with the firm's incentives to refuse supply, not the possible ways of justifying it in various legal frameworks.

In what follows we quote from a report by the Monopolies and Mergers Commission (1981) and give in brackets the numbers of the passages quoted from.

Some of the arguments presented by Raleigh were:

1. A valuable brand image was developed and sustained by the selective distribution system (3.10).

2. A policy of competing on price (and producing to a low specification) cannot be successfully combined with a policy of selling primarily through specialist dealers (and producing to a high specification) (3.15).
3. Selling to discount stores would undermine the Raleigh brand image both with consumers and with its other retailers (3.19).
4. In Raleigh's view, it was not uncommon for discount stores to sell goods as 'loss leaders' [...]. Other retailers could not match these prices, therefore they would switch to selling other goods [...] (3.20).
5. Raleigh did not think it anti-competitive to seek to preserve existing specialist and technical distribution channels by trying to maintain general price levels consistent with the provision of good service [...] (although they emphasised the difference between this and retail price maintenance (3.22).
6. Raleigh said that it hoped, but could not guarantee, that the expanded network [of '5 Star' specialist dealers] would increase its total sales (3.28).
7. The safety image, Raleigh maintained, was reinforced if the bicycle was sold through a specialist outlet [...] - and this had implications for the price (3.29).

The arguments quoted so far can essentially be summarised as follows: the selective distribution policy sustains the brand image. This enables dealers to charge relatively high prices. If, on the contrary, discount stores sell Raleigh bicycles as well these prices can no longer be maintained.

As argument no. 6 suggests the hope of an increase of sales due to a specialised distribution network is a rather vague one. The main point is that delivery to discount stores would result in price cuts. Why should Raleigh be concerned about these price cuts in the downstream market? Paragraphs (3.36) to (3.42) describe the reaction of specialised dealers to a change of distribution policy. This seems to be the main drawback.

Again, we quote the main points.

1. Raleigh represented that, if it were forced to supply discount stores [...] it would lose much of its dealers' loyalty and goodwill. Many would desert Raleigh out of a feeling that Raleigh had deserted them [...]. Raleigh had built up its good will on the basis that it was pursuing the policy of selling through specialist dealers (3.37).

2. Raleigh argued that the Office of Fair Trading (OFT) had underestimated the probable reaction of Raleigh dealers and that more than 10 per cent of them would drop its bicycles if there were a change in distribution policy (3.38).

3. Such a reaction by dealers would be understandable because (apart from their feeling of being let down) they would face what they would regard as predatory selling (3.39).

4. When, in 1965, Raleigh sold bicycles to catalogue mail order houses Raleigh's decision affected its trading relationship with, and its sales to, its dealers. Dealers' allegiance had had to be carefully recultivated over the years (3.40).

Although our models do not address the quality issue raised by Raleigh [9] some of its features seem to be reflected by the main arguments used by Raleigh that essentially amount to the following line of reasoning: if prices cannot be maintained in the downstream markets bicycle dealers will no longer buy Raleigh bicycles, at least they will not buy them at the prices Raleigh would like to ask for. Raleigh,

[9]The quality or brand image issue will, in fact, be raised by the majority of firms alleged of supply refusal. Hardly any of them would admit that their refusal is based on an output contraction strategy aiming at higher prices because such a statement would be detected immediately as anti-competitive and resulting in a reduction of consumer surplus and welfare. Defensive arguments used by firms before courts should be interpreted as such.

expecting this reaction, does not supply discount stores although this would be, in itself, a profitable transaction.

We will now try to relate this reasoning to the question of an upstream firm's commitment possibilities. Raleigh's motivation for not 'cheating' can be related in particular to two approaches we modelled in this note. First, the possibility of future trading periods seems to be taken into account by Raleigh quite seriously. The situation is not viewed as a one shot game. Instead, future periods are vital for Raleigh's strategic considerations. This makes 'not cheating' a subgame-perfect equilibrium (section 3.1). Secondly, as the argumentation put forward by Raleigh shows time has been needed and used to build up a reputation for not cheating (section 3.3). This reputation would suffer from a change in distribution policy and, more importantly, the desire to acquire and maintain this reputation clearly is an important motivation for Raleigh to refrain from supplying other downstream firms.

In the light of the assumption made in (H-T) one would have expected Raleigh to supply the discount stores after having supplied the specialist dealers. The models presented in section 3 of this note suggest that, in various settings, this should not happen. The case presented goes farther than this. Raleigh is not only able to commit to a policy of monopolisation in the sense of supply refusal but it holds this strategy to be so much more profitable than supplying discount stores, that it undertakes much effort to defend this policy before the courts.

6 Conclusion

The assumption that vertical integration is the only way for an upstream firm to monopolise the downstream market is confronted with several specific settings that yield cooperative outcomes in the sense that firms can and do commit to monopolise the market without integrating.

Cooperation arises as a subgame-perfect equilibrium if players always expect a future trading period with positive probability. A setting in which cooperation occurs as (unique) equilibrium in a finitely repeated game is the situation where the upstream firm incurs sufficiently high cost c if it wants to supply other downstream firms (specific investments). In a setting of a finitely repeated game with incomplete information (commitment types) we show that firms may have incentives to build up a reputation of being cooperative. The outcome obtained is cooperation in all periods except for a limited number of periods towards the end of the game. Finally, it is pointed out that nonenforceable exclusive-dealing contracts can be substituted by various other contracts which are enforceable.

Implications for the model by Hart and Tirole (1990) primarily concern the results in the 'ex-post monopolisation variant' where integration is presented as the only way for firms to achieve output contraction in a downstream market. However, our models do not imply that integration is equivalent to collusion of non-integrated firms. Such a claim would not hold for situations of scarce needs (variant 2 of H-T's model) or scarce supply (variant 3).

For empirical evidence we analyzed a case against Raleigh. Some of Raleigh's most important arguments can be related to models discussed in our paper. Possibilities of future trade and investments in reputation appear to be key elements determining Raleigh's distribution policy.

7 References

Armstrong, Mark, Simon Cowan, and John Vickers (1994) : *Regulatory Reform: Economic Analysis and British Experience;* MIT Press.

Carlton, D.W. (1989) : *'The Theory and the Facts of how Markets Clear: Is Industrial Organization Valuable for Understanding Macroeconomics?' in Handbook of Industrial Organization;* North-Holland.

Coase, R. (1937) : The nature of the firm; *Economica 4: 386-405.*

Grossman, Sanford J., and Oliver D. Hart (1986) : The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration; *Journal of Political Economy 94: 691-719.*

Hardt, Michael (1995) : Market Foreclosure without Vertical Integration; *Economics Letters 47: 423-429.*

Hart, O. (1989) : An Economist's Perspective on the Theory of the Firm; *Columbia Law Review 89: 1757-1774.*

Hart, O. (1995a) : Corporate Governance: Some Theory and Implications; *The Economic Journal 105: 678-689.*

Hart, O. (1995b) : *Firms, Contracts, and Financial Structure;* Oxford University Press.

Hart, O. and J. Tirole (1990) : Vertical Integration and Market Foreclosure; *Brookings Papers on Economic Activity, Microeconomics, 205-276.*

Jensen and Meckling (1976) : *Theory of the Firm: Managerial Behaviour, Agency Costs and Ownership Structure;* 3 Journal of Financial Economics: 305-310.

Kreps, D., P. Milgrom, J.Roberts, and R. Wilson (1982) : Rational Cooperation in the Finitely Repeated Prisoner's Dilemma; *Journal of Economic Theory 27: 245-252.*

Monopolies and Mergers Commission (1981) : Bicycles, A Report on the application by TI Raleigh Industries Limited and TI Raleigh Limited of certain criteria for determining whether to supply bicycles to retail outlets; *London, HMSO.*

Perry, Martin K. (1989) : Vertical Integration: Determinants and Effects; *in Schmalensee, R. and R.D.Willig (Eds.), Handbook of Industrial Organization, Amsterdam: 183-255.*

Porter, R. (1983) : Optimal Cartel Trigger Price Strategies; *Journal of Economic Theory 29: 313-338.*

Tirole, Jean (1988) : *The Theory of Industrial Organization;* MIT Press.

Waterson, Michael (1993) : Vertical Integration and Vertical Restraints; *Oxford review of economic policy, Vol 9: 41-57.*

Williamson, Oliver (1985) : *The Economic Institutions of Capitalism;* New York: Free Press.

CHAPTER III

Network Access and Ownership Structure [1]

Abstract

Competition can be introduced to network related industries by giving competitors access to the network. When regulation requires that access prices are based on network operating costs there are generally incentives to misrepresent these costs. We analyze how these incentives depend on the ownership structure of the industry, namely vertical integration, vertical separation, and joint ownership. With a fixed number of downstream network users vertical separation is dominated by vertical integration which, in turn, is dominated by joint ownership. With entry this welfare ordering is reversed when incentives to deter entry become sufficiently strong. But the following result holds irrespective of whether or not entry is allowed for. If capital shares of joint owners are adjusted endogenously and optimally incentives to distort costs disappear entirely and highest welfare is obtained. Our analysis of the incentives for upstream and downstream investments amplifies these results. Finally, in the vertical integration case accounting separation turns out not to be an effective regulatory tool unless some rather restrictive assumptions are satisfied.

1 Introduction

We now turn from considerations of ownership and vertical integration per se to the impact of ownership in industries that are regulated due to the existence of a bottleneck facility.

In various industries like, for instance, telecommunications, gas or electricity supply, and railway transportation a single network is superior to competing network capacity because of the natural monopoly

[1]This chapter is based on Hardt and Stürmer (1995a).

character [2] of the network related production or transmission process. Therefore, an attempt to introduce competition into these industries requires that network access is granted to third parties.

So far, the relevant theory and politics have primarily been concerned with the determination of the optimal access price that provides fair access conditions for new entrants and cost coverage for the incumbent (see Baumol (1994), Armstrong and Doyle (1994), Cave and Doyle (1994) etc.). Research along these lines has mainly treated access pricing in settings of complete information and vertical integration. [3] However, moral hazard (effort decision) and adverse selection (asymmetric information about costs of the network operator) induce inefficiencies that regulation is designed to minimise. And such inefficiencies may depend to a large degree on the vertical structure of the industry. Laffont and Tirole (1993, 1994a), among others, take into account informational asymmetries but concentrate on vertically integrated settings.

Relatively little research, however, has been done on implications of ownership structures on access pricing and on the possibilities and incentives to discriminate among network users. This is somewhat surprising as network related industries which have been opened for competitive access actually show various patterns of vertical structures, i.e. they differ in the allocation of ownership rights between upstream (network operation) and downstream activities (e.g. provision of telecommunication services, power generation, train operation). These patterns differ not only across industries in the same country (e.g. telecommunications and railway industries in the UK: the former is vertically integrated, the latter is structurally separated) but they

[2] In some of the above industries the natural monopoly character may be questionable. In telecommunications, for example, new technological developments (optical fibre networks or radio links) now allow for competing networks even in the local area. For a more detailed discussion see, for instance, Baumol and Sidack (1994). In the discussion of this chapter our attention will be focused on the access problem in single network industries. Network duplication will be considered in chapter 4.

[3] See Armstrong, Cowan and Vickers (1994) and Cave and Doyle (1994). For a survey on access pricing with special reference to ownership structures and informational asymmetries see Hardt and Stürmer (1995b).

also differ across countries in one and the same industry (e.g. telecommunications in the UK and in the US). These differences do not appear to be the result of well understood differences in industry specific or country specific characteristics. Rather, there seems to be a considerable amount of uncertainty as to the question of which ownership structure facilitates regulation of network related industries and serves best the aim of enhancing efficiency.

Moreover, the existing work taking into account different ownership structures has rarely considered informational asymmetries.

Among the few contributions treating implications of ownership there is a model in Vickers and Yarrow (1988). They distinguish vertical separation - corresponding to non-integration in our terminology - and interconnection, i.e. partial integration, and analyze regulatory measures to prevent the incumbent from charging excessively high access prices. It is argued that vertical separation does not enhance welfare by itself because separation may lead to double marginalization. This problem does not appear under partial integration which may thus be more favourable if effective access price regulation leads to equal access conditions for non-integrated downstream firms. Asymmetric information about network operating costs is not taken into account.

Byg and Hardt (1995) analyze a simple network with links that may be owned by different firms which, in turn, may be separate or integrated. In this setting they investigate the effects of ownership and of the class of contracts permitted on equilibrium prices and on entry. Starting from this free market analysis they draw conclusions as to the informational requirements for the regulator in various settings. But private information about costs is not taken into account explicitly in the model.

Armstrong, Cowan and Vickers (1994, p. 135 ff.) show in a simple model that, in a complete information setting, the optimal access price under vertical integration exceeds the one under vertical separation. It is argued that under asymmetric information generally the same principles apply as under complete information; however, informational asymmetries may have negative effects in the case of vertical integration because incentives to raise rivals' costs arise. But a higher access price may, on the other hand, lead to less entry and therefore less duplication of fixed costs. The overall welfare effect of vertical integration is therefore ambiguous. It is also pointed out that some other aspects are of importance: the quality of access may depend on the vertical ownership structure; and so may the quality of information available to the regulator. Therefore, the careful conclusion drawn is that 'if vertical conduct regulation is difficult and if the benefits of competition are thought to be substantial, then structural remedies even at some cost in terms of scope economies may deserve examination' (p.162). Part of this analysis is based on (an earlier version of) a paper by Vickers (1995) that analyzes the effect of vertical integration in a setting of imperfect information about a (regulated) monopolist's costs and imperfect competition in the downstream industry. The welfare effect of vertical integration is ambiguous because of the reason outlined above: the incentive for raising rivals' costs may or may not be offset by the resulting lower degree of fixed cost duplication.

Alger and Braman (1993) [4] and Alger (1993) suggest a different ownership structure in the context of gas industries. They argue that competitive joint ventures induce competition sufficient to regulate some natural monopolies. The arguments showing efficiency of competitive joint ventures are based on the assumptions of open ownership, independent marketing, and the 'use or lose' rule. These assumptions prevent owners from restricting capacity in either stage of the production process (network capacity or downstream services).

Their approach differs from the joint ownership version of this chapter's model in several ways. When we compare a setting of joint ownership to other ownership structures, we do not use the above com-

[4] Gale (1994) analyses a model which is similar to Alger and Braman, but a complete analysis of a price setting game capturing the main features of joint ventures is given. Informational asymmetries about costs of production, however, are mentioned among the caveats that remain. Lewis and Reynolds (1979) discuss informally a set of rules for competitive joint ventures which are similar to those analyzed by Alger and Braman.

petitiveness assumptions. The network, in our joint ownership case, is owned by a fixed number of firms that are active in the downstream market. Also, we formally incorporate informational asymmetries about the cost of running the network. Alger and Braman argue that a new owner is better informed than the regulator and that, therefore, the problem of estimating the average costs (which are the basis for the access price) is 'not worse' than the regulator's problem of determining the access price. We show how, in a setting where network operating costs are not enforceable, incentives to misrepresent costs can vanish completely.

In this chapter we analyze the impact of such informational asymmetries under different ownership structures. In doing this we partly follow a suggestion by Laffont and Tirole to analyze the costs and benefits of breakups in the light of the foreclosure literature such as the models by Bolton and Whinston (1993) and Hart and Tirole (1990). [5] Our approach differs from the models by Laffont and Tirole in that we take into account different ownership structures of the network related industry. Further, we do not design optimal contracts for the regulated firm. The question we analyze is rather in how far problems of moral hazard and adverse selection can be minimised by implementing an appropriate ownership structure: suppose the regulated monopolist is required to provide network access on the basis of per unit costs.[6] Then, given the assumed informational asymmetries, the network operating firm has incentives to exploit its informational advantage by misrepresenting costs with the help of accounting manipulations ('cost padding', etc.) in order to defend higher access charges. If there is an ownership structure under which these incentives are small or even zero (and we will show that such cases do exist) the role of the regulatory authority in fixing and controlling the access price would be marginalised, once the appropriate ownership structure has been established.

In contrast to the vertical integration models by Hart and Tirole

[5] See Laffont and Tirole (1993), p.267, footnote 9.

[6] We do not deal with the question of the optimal allocation of total costs to network users. For an overview on cost allocation see Cave and Mills (1992).

and by Bolton and Whinston cited above we assume the upstream monopolist to be regulated and also we allow for more general ownership structures than non-integration and partial integration. Joint ownership in our analysis means that the downstream firms own capital shares (adding up to one) in the upstream firm.

The rest of this chapter is organised as follows. Section 2 gives an overview of the ownership structures we are going to consider and examples for their occurrence in various industries. The basic model is presented in section 3. Section 4 solves for equilibrium quantities, profits, and access prices under the assumption that the regulatory tool of accounting separation is sufficient to prevent the incumbent firm from discriminating against a competitor by means of unequal access (controllability). This assumption is dropped in section 5. Welfare implications of partial integration, non- integration and joint ownership are given in both of these sections. Sections 6 and 7 contain extensions of the model: in section 6 the roles of downstream and upstream investments are analyzed; section 7 investigates the impact of entry on the equilibrium allocation and on welfare. Section 8 contains some concluding remarks on implications of our findings for regulatory policy and suggestions for further research. In the appendix we discuss some arguments concerning the effectiveness of accounting separation as a regulatory tool.

2 Ownership Structures in Partly Deregulated Industries: Some Examples

Although open access provision is a widely used means for the introduction of competition to network related industries there are significant differences in the ownership structures that have been implemented in these industries by the regulatory authorities. We observe the following possible ownership structures.

2.1 Non-integration

In order to guarantee equal network access for the downstream firms the former vertically integrated monopolist has to divest from its downstream activities. The most famous example is the deregulation of the American telecommunications market where the Modification of Final Judgement (MFJ) led to the divestiture of AT&T from the local telephone companies in 1984.

2.2 Partial integration

A less drastic step is to keep the vertically integrated structure intact, i.e. the upstream monopolist competes on the downstream market while giving at least one competitor access to the network.

Existing deregulation patterns applied in such settings of partial integration can be distinguished by the extent to which the upstream firm is able to discriminate against competitors by charging unequal access prices. The European Commission's approach, for instance, is to prescribe access prices to be cost based and non-discriminatory which requires separate accounting for the upstream and downstream activities of the integrated firm. This applies not only to telecommunications but also to other industries like railway and the energy sector.

The British telecommunications regulator Oftel, in 1984, intended to follow a more liberal approach: the determination of access prices and conditions of using British Telecom's local telephone network were subject to negotiations between BT and Mercury. Only now the regulatory authority is departing from this approach after continuous complaints by Mercury about discriminatory pricing and non-pricing activities by BT. Now BT is forced to account for its network business separately from its retail activities. Further, BT must publish cost-based charges for access to specific services needed by competitors that require access to the BT network (see Oftel (1994)).

We will refer to these two ways of organising a partially integrated firm by the term partial integration with accounting separation and without accounting separation. Both structures will turn out to differ in their welfare implications only if a rather strong assumption is satisfied, namely controllability. We denote the case of partial integration with effective accounting separation (i.e. accounting separation under controllability) by 'PI' and distinguish it from the case without accounting separation 'PI^N' which includes the case of accounting separation without controllability. We define the notion of controllability in the following way.

Definition 1 *Controllability of access prices is given if and only if the regulator can induce the integrated firm to base the choice of downstream output on the officially published access price.*

Assuming controllability implies that firms always act as if access costs were identical for both downstream firms. In this setting, firms are symmetric not only as far as accounting is concerned but also as far as output choices are concerned. We show later that the controllability assumption plays a crucial role in the assessment of accounting separation as a policy tool to provide fair access conditions and also in the assessment of the welfare implications of partial integration.

2.3 Joint ownership

Another possible concept for the introduction of downstream competition is the common ownership of the network by the downstream firms. A few examples can be found in private sector industries. The oil pipeline industry in the United States as well as in Europe are examples where there are joint ventures in which several firms, usually oil companies, own a percentage of the pipeline stock often proportional to anticipated usage or joint ventures in which each participant owns directly, and therefore has the right to make use of, a percentage of pipeline capacity (see Hillman (1991)). A similar arrangement was introduced to deepwater ports in the US (see Lewis and Reynolds (1979)).

But this kind of ownership structure does not yet seem to have been adopted by regulatory authorities. [7]

One of the aims of the model we are going to present is to contrast welfare implications of joint ownership with those of other ownership structures. We identify conditions under which joint ownership socially dominates partial integration and/or non-integration, and vice versa.

Joint ownership, of course, allows for different ways of allocating control over the assets to the joint owners. In our specific context, control will be equivalent to the decision about the access charge (cost signal). Here, we will use the majority rule (see Hart and Moore (1990)) where the downstream firm with the highest capital share in the network operating firm chooses the access charge. Alternatively, the charge could, for example, be determined jointly by the weighted average of the downstream firms' choices. For m downstream firms the signal would then be given by $s = \sum_{i=1}^{m} \mu_i s_i$ where μ_i is the share of downstream firm i in the upstream firm, $\sum_{i=1}^{m} \mu_i = 1$, and s_i is the cost signal given by firm i. The interpretation in the latter case would be that a firm's relative bargaining power in the determination process of the access charge largely corresponds to the share in the network operator's capital held by this firm.

A question we want to analyze is whether or not joint ownership structures similar to those in the private sector present a desirable allocation of ownership rights in regulated industries.

3 The Model

The presentation of the model will be followed by a number of justifications of some of the assumptions.

[7] Although it is true that the National Grid Company (NGC) in the UK is jointly owned by the Regional Electricity Companies (REC) this does not constitute a counter example to the above claim. This is because generators other than the REC are the main group of firms that need access to the NGC's transmission network (high tension) but do not have ownership shares in the NGC. Further, the REC need access to the NGC's network only for a part of their business, namely for the supply of large customers in other areas. For a detailed presentation of the electricity industry see Armstrong et al. (1994), Littlechild (1994), and Newbery (1994).

The network is operated by the upstream firm U. Downstream Cournot duopolists D_1, D_2 have access to the network at a price that depends on the extent of usage of the network (s per unit). We compare three different ownership structures:

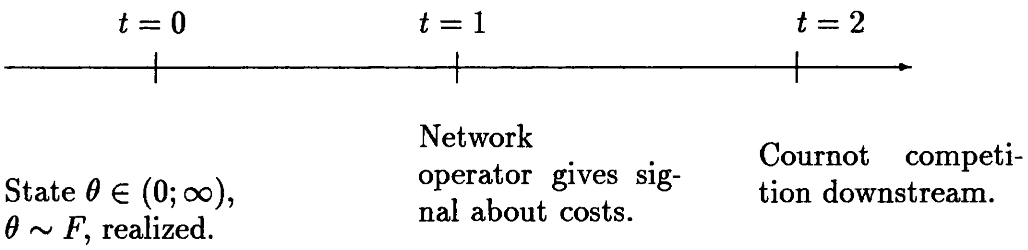
1. Non-integration (NI): the three firms are owned separately (complete divestiture of the downstream activities).

2. Partial Integration (PI): the upstream firm is also active in the downstream market, i.e. integrated with one downstream firm, D_1 say. Effective accounting separation (i.e. the controllability assumption holds) is denoted by PI while the cases of non - accounting separation and accounting separation without controllability are denoted by PI^N.

3. Joint Ownership (JO): the upstream firm is owned by the two downstream firms (proportions $\mu, 1-\mu$). The majority rule applies for the choice of the cost signal.

We do not need to analyze explicitly the monopolistic case where an integrated firm supplies the whole market. The timing of the game is as follows:

Figure 1



Upstream per unit operating costs are given by a random parameter θ which is distributed on $(0;\infty)$ according to an arbitrary distribution function, $F(\cdot)$. Cost parameter θ is realised in period $t=0$ and is observable by upstream and downstream [8] firms but not verifiable.

The network operating firm is required to give access to the network at per unit costs.[9] In $t = 1$, a signal about cost parameter θ is given by the network operating firm, i.e. U under NI, $\{U, D_1\}$ under PI, and D_i with $\mu_i \geq \frac{1}{2}$ under JO. We assume this signal can be either the true one, or a fixed amount $\Delta > 0$ can be added to the incurred costs: $s(\theta) \in \{\theta, \theta + \Delta\}$. If a signal $s \neq \theta$ is given firm U incurs costs $K \geq 0$ (accounting manipulations). This assumption merely aids intuition but is in no way essential; the whole analysis remains true for $K = 0$.

Firms D_1, D_2 choose their output levels in $t = 2$. Downstream per unit production costs are denoted by d_1 and d_2 respectively. The final goods market is imperfectly competitive which is reflected by Cournot competition in our model. Demand for the downstream product is given by a linear downward sloping function $p(Q) = a - bQ$.

A few remarks may be in order as further justification for some of the above assumptions.

Firstly, cost misrepresentations may be achieved by allocating common costs or even downstream costs to the network. A building may be used for the administration of upstream and downstream operation. The same technical or managerial staff may be involved in the provision of services in the competitive as well as in the monopolistic segment. Therefore, there are always some degrees of freedom in the determination of the network operating costs, a phenomenon that is well known to all those who are involved in, or acquainted with, the discussion about access deficit contributions etc.

Further, The parameter Δ can be interpreted as kind of a proxy variable for the seriousness of the moral hazard problem. Allowing for

[8]Observability of θ by downstream firms is necessary for the Nash equilibrium under PI^N (section 5). The intuition is that all firms active in the industry know the true costs of operating the network but they cannot prove this to a third party (regulator).

[9]Different suggestions as to the optimal access price have been made in the literature. These range from average or marginal costs, or even lower values as in Vickers and Yarrow (1988), to concepts including opportunity costs (efficient component pricing by Baumol and Sidack (1994)); see also Armstrong et al. (1994) on this point. What we need to assume here is only that some fixed compensation per unit can be charged and that this compensation is required to be based on the network operator's costs.

optimal cost misrepresentation in the sense that the network operator is able to choose Δ from a continuum would complicate the analysis considerably without any obvious advantage or additional insight. We aim at choosing the simplest model that allows us to analyze the incentives for cost misrepresentations. Giving the network operator a binary choice accomplishes this task.

We chose an additive distortion although there is a variety of ways of modelling cost misrepresentations. One could, for instance, think about multiplicative distortions. Then, firms would turn out to cheat for high values of θ (given K is fixed) instead of for low values as in our model. But results remain qualitatively the same in that the same ownership structures as in our model turn out to be more vulnerable to non-truth-telling, i.e. the welfare ordering we establish is not affected. If K depends on the size of the distortion the range of parameters for which cheating occurs depends crucially on the specification of the cost function $K(\Delta)$. This seems to be rather arbitrary. As no additional insight seemed to be obtained from these modelling alternatives we decided to use the fixed additive distortion for our model.

Regulation is modelled in a similarly simple way. Our concern here is *not* to inquire how the regulator may best use the available tools (menus of contracts etc.) in order to minimise the moral hazard problem but rather to analyze the impact or ownership given the regulatory policy of cost-based pricing. Underlying this argument there is kind of a continuity assumption in the sense that the welfare ordering we will establish for a given policy remains the same when a larger set of tools is available. A richer set of tools may enable the regulator to extract some of the informational rent but there appears to be no plausible reason to assume that the welfare *ordering* would be affected. However, we will provide some arguments later on, i.e. in conjunction with the relevant propositions, why we think that some ownership structures are inherently more biased towards non-truth-telling strategies than others.

We now analyze the model conditional on ownership structures.

4 Equilibrium without Discrimination

In this section we assume that under partial integration the integrated firm is required to publish access prices and to publish separate accounts for its upstream and for its downstream activities. We further assume that the regulator is able to make sure that the integrated firm produces the output which is profit maximising under the published access price, i.e. we assume controllability. This assumption is made here in order to demonstrate what one needs to assume if some of the arguments frequently put forward in the discussion and informal literature on access pricing and, especially, on accounting separation are to be true. We will argue in section 5 that the assumption is unrealistic and analyze the model without assuming controllability.

4.1 The game

Under any of the above ownership structures, in equilibrium, the Cournot quantities are traded in the downstream market. Total downstream operating costs for D_i are given by $c_i = d_i + s(\theta)$. [10] The unique Nash equilibrium in $t = 2$ is characterised by output levels

$$q_i^* \quad = \quad \frac{a - s(\theta) - 2d_i + d_j}{3b} \tag{1}$$

which implies that profits are given by

$$\pi_i^* = \frac{1}{b}\left(\frac{a - s(\theta) - 2d_i + d_j}{3}\right)^2$$

in $t = 2$. For a given ownership structure this equilibrium is unique. But equilibrium quantities and profits vary across ownership structures because the cost signal s does, as we will see.

Signals about costs are given in $t = 1$. We can now determine the conditions under which it is optimal for the upstream firm to misrepresent costs. This is the case whenever the total profit from the non-truth-telling strategy exceeds the total cost incurred by adopting

[10] Note that under the assumptions of this section both firms, D_1 and D_2, are charged the same access price: $s_1(\theta) = s_2(\theta) = s(\theta)$.

this strategy. The latter is given simply by K for the case of NI but with PI or JO downstream losses have to be taken into account. Therefore, for the ownership structures in question, one obtains the following conditions: [11]

$$\begin{aligned}
\mathrm{NI}: &\quad \Delta(q_1+q_2) &>& \quad K & (2)\\
\mathrm{PI}: &\quad \Delta(q_1+q_2) &>& \quad K+\pi_1(\theta)-\pi_1(\theta+\Delta) & (3)\\
\mathrm{JO}: &\quad \mu_1\Delta(q_1+q_2) &>& \quad \mu_1 K+\pi_1(\theta)-\pi_1(\theta+\Delta) & (4)
\end{aligned}$$

where ($\mu_1 \geq \frac{1}{2}$) denotes the larger capital share of the two downstream firms, D_1 say (w.l.o.g.). $\pi_1(s)$ denotes firm D_1's downstream profit when cost signal s is given. In $t=1$, therefore, $s=\theta+\Delta$ is chosen whenever the relevant condition (2)-(4) is satisfied, and $s=\theta$ obtains otherwise. The natural question arising here is if it is possible to infer from conditions (2) - (4) whether or not non-truth-telling strategies are more worthwhile under some ownership structures than under others. We pursue this question in the following subsection.

4.2 Welfare implications

In what follows we adhere to the welfare notion based on the sum of producer and consumer surplus (all results remain true for a weighted sum) which implies that welfare is higher for higher output levels or, equivalently, for lower output prices. As output and prices in our model are contingent on the cost parameter θ we use expected prices and output as welfare indicators [12].

In this setting the following proposition can be stated.

[11] We assume that upstream and downstream profits are valued equally although upstream profits are not payable to shareholders in the form of dividends and do not appear on the balance sheet as they are 'hidden' by cost misrepresentations. This equivalent treatment, however, might be justified if one takes into account the 'dividend puzzle': the value of a firm and therefore of its shares does not depend on the amount paid as a dividend (see Black (1984)).

[12] This welfare definition is equivalent to that of expected consumer surplus only if the variances of the prices $p(\sigma)$ (given some ownership structure σ) satisfy some regularity conditions. For the sake of tractability, however, we identify highest welfare with lowest expected output prices.

Proposition 1 *Given that the integrated downstream firm is not too inefficient ($d_1 < d_2 + \frac{3bK+\Delta^2}{3\Delta}$) the following welfare ranking holds. Under joint ownership welfare is higher than under partial integration for any distribution of ownership shares $\mu \in [\frac{1}{2}, 1)$. Under partial integration welfare is higher than under non-integration . For symmetric firms ($d_1 = d_2$) highest welfare is attained for JO with $\mu_1 = \frac{1}{2}$.*[13]

Proof. The difference in downstream profits in (3) and (4) can be written as

$$\pi_1(\theta) - \pi_1(\theta + \Delta) = \frac{2\Delta}{3b} \cdot \frac{a - \theta - 2d_1 + d_2}{3} - \frac{\Delta^2}{9b}.$$

Substituting equilibrium outputs q_i^* from (1) into (2) - (4) yields that non-truth-telling occurs if and only if θ does not exceed a critical limit $\bar{S}(\sigma)$ for ownership structure $\sigma \in \{NI, PI, (JO; \mu_1)\}$:

$$\begin{aligned}
\text{NI}: \theta &< \bar{S}(NI) := a - \Delta - \frac{1}{2}(d_1 + d_2) - \frac{3bK}{2\Delta} && (5)\\
\text{PI}: \theta &< \bar{S}(PI) := a - \Delta - \frac{1}{4}(5d_2 - d_1) - \frac{\Delta}{4} - \frac{9bK}{4\Delta} && (6)\\
\text{JO}: \theta &< \bar{S}(JO, \mu_1) && (7)
\end{aligned}$$

where

$$\bar{S}(JO, \mu_1) := a - \Delta - \frac{1}{2(3\mu_1 - 1)}[\Delta + d_1(3\mu_1 - 4) + d_2(3\mu_1 + 2) + \frac{9\mu_1 bK}{\Delta}].$$

It is easy to see that $\bar{S}(NI) > \bar{S}(PI)$ obtains whenever

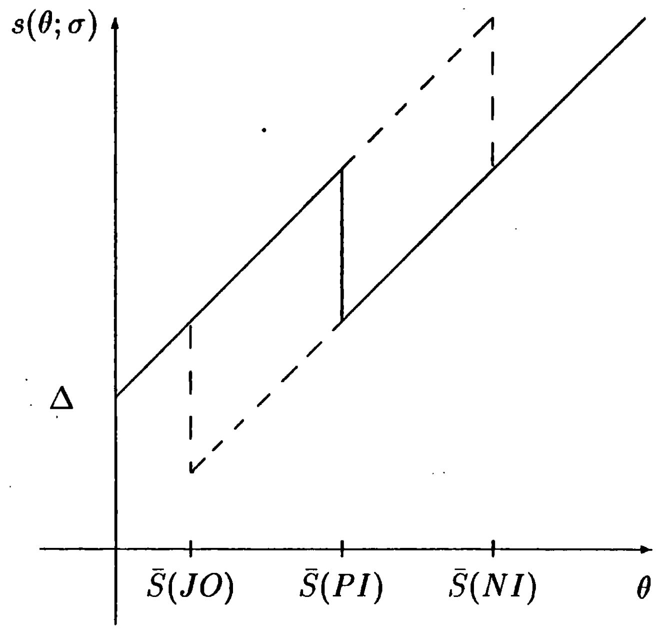
$$d_1 < d_2 + \frac{3bK + \Delta^2}{3\Delta}.$$

Similarly, $\bar{S}(PI) > \bar{S}(JO, \mu_1)$ holds for all μ_1 in the relevant range ($\mu_1 \in [\frac{1}{2}, 1)$) and, not surprisingly, equality holds for $\mu_1 = 1$.

Let $s(\theta, \sigma)$ denote the signal that is given in state θ when the ownership structure is $\sigma \in \{NI, PI, (JO, \mu_1)\}$. These signals and the critical values $\bar{S}(\sigma)$ are illustrated in figure 2.

[13] For obvious reasons of symmetry any of firms D_1, D_2 can choose the cost signal.

Figure 2



Because of $\bar{S}(NI) > \bar{S}(PI) > \bar{S}(JO;\mu_1)$ the probability of misrepresenting costs are highest under non-integration, lower under partial integration, and lowest under joint ownership.

Denoting by $F(\cdot)$ the distribution function of θ, the probability of misrepresenting costs under ownership structure σ is given by $F(\bar{S}(\sigma))$. From $\bar{S}(NI) > \bar{S}(PI) > \bar{S}(JO;\mu_1)$ it follows immediately that

$$F(\bar{S}(NI)) > F(\bar{S}(PI)) > F(\bar{S}(JO;\mu_1)).$$

This implies that expected access charges satisfy:

$$E[s(\cdot)|\ NI] > E[s(\cdot)|\ PI] > E[s(\cdot)|\ JO].$$

Using (1) we obtain

$$E[q_i^*|\ NI] < E[q_i^*|\ PI] < E[q_i^*|\ JO], \quad i = 1, 2.$$

This is because output levels are linear in $s(\cdot)$. But, as higher expected total output implies lower expected prices, the result follows.

It remains to be shown that, in case of joint ownership and symmetric firms, highest welfare is attained for $\mu = \frac{1}{2}$. To see this is true it is sufficient to note that for $d_1 = d_2$ (7) simplifies to

$$\bar{S}(JO, \mu_1) := a - d_1 - \Delta - \frac{9\mu_1 bK + \Delta^2}{\Delta(6\mu_1 - 2)}$$

and therefore

$$\frac{\mathrm{d}\bar{S}(JO, \mu_1)}{\mathrm{d}\mu_1} = \frac{18bK\Delta + 6\Delta^3}{[\Delta(6\mu_1 - 2)]^2} > 0$$

always holds. This implies that $\bar{S}$ is minimised for the lowest value of μ_1 in the relevant parameter range, i.e. at $\mu_1 = \frac{1}{2}$.

Q.E.D.

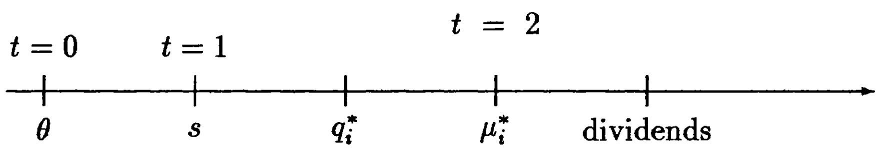
At first sight it might be surprising that such a clear cut welfare ranking can be established and, in fact, different conclusions will be obtained once we allow for entry. But the intuition for the result in this setting here is plain. Basically, there is a conflict between upstream interests, i.e. high revenues from access charges, and downstream interests, i.e. low access charges. A non-integrated network operator has upstream interests only and has thus highest incentives to charge high access prices. An integrated firm has to balance the upstream incentives to charge high prices with the downstream interest in having low access charges and therefore low marginal costs. Under joint ownership, finally, there are full downstream interests to be taken into account but only a share μ_1 in the upstream profits is being considered. Therefore the setting of joint ownership performs better than partial integration for any $\mu_1 < 1$ and best for smallest shares of the firm choosing the cost signal, i.e. $\mu_1 = \frac{1}{2}$. The full intuition for this last point will become clear after the analysis in the section on efficient outcomes.

Efficient outcomes

A question arising here is in how far ownership structures can induce efficient behaviour in the sense that distortions disappear entirely and the welfare maximising access price θ is charged. We will, in fact, describe such a setting in what follows.

For the sake of clearer intuition we now allow for m downstream firms. Each of them initially holds a capital share μ_i, $\sum_{i=1}^{m} \mu_i = 1$, in the upstream firm. Further, we modify our basic model in the following way. Firstly, in order to avoid zero share prices, we assume that the regulator allows the upstream firm to charge an access price that equals the costs of giving access plus a limited amount of profit per unit, say $\theta + \epsilon$. Such arrangements are actually quite common in regulated industries[14]. Secondly, capital shares will be adjusted after competition. Timing, therefore, is as follows. The cost parameter θ is realised in period $t = 0$. In $t = 1$, the cost signal is given. The majority rule is used in the sense that the downstream firm with the highest capital share gives the cost signal.[15] The access price now consists of the sum of the cost signal and the allowed profit: $s + \epsilon$, $s \in \{\theta, \theta + \Delta\}$.

Figure 3



Period $t = 2$ is subdivided into three stages as shown in figure 3. In stage 1 output decisions are taken. In stage 2 the capital shares of the downstream firms may be adjusted according to a rule that has been specified by the regulator in advance. [16] This rule specifies the amounts of capital to be transferred between firms and the price at which these shares are traded, both, of course, conditional on realised

[14]Average revenue regulation. See Armstrong et al. (1994), pp. 69-72 and p. 178.

[15]The choice of decision rule here will turn out not to be crucial because, as we will show, none of the downstream firms has any incentive for non-truth-telling. Therefore, any firm can be asked to give the signal.

[16]An arrangement along these lines has been advocated by Lewis and Reynolds (1979) (for a setting of complete information). They describe a situation where capital shares are adjusted on a yearly basis according to 'throughput shares'. This is cited as one out of a set of competitive rules which the Department of Justice designed for deepwater ports.

output. Finally, the upstream firm pays dividends to the downstream firms. Total dividends equal the amount of profit the upstream firm has been able to make under the pricing constraint imposed by the regulator.

Note that for $m = 2$ and $\epsilon = 0$ our basic model evolves as a special case with the only difference that we now allow for an ex post adjustment of capital shares.

In this setting the following efficiency result is obtained.

Proposition 2 *If the regulator adjusts capital shares according to the following rule*

$$\mu_i^* := \frac{q_i}{q_1 + \cdots + q_m}, \ i = 1, \cdots, m,$$

and if the upstream firm's capital is traded at a price that equals the profit it is allowed to make, i.e. $p_U = \epsilon \cdot (q_1 + \cdots + q_m)$

then

1. *no firm has incentives for non truth-telling strategies, i.e. to choose an access price different from* $\theta + \epsilon$, *and*
2. *if an access price higher than* $\theta + \epsilon$ *were charged this would not affect the firms' output decisions.*

Proof. To see why part 2 of the proposition holds one just has to write down the maximisation problem of firm i. Suppose a signal $s \in \{\theta, \theta + \Delta\}$ is charged. Denote by $\pi_i^m(s)$ the profit of firm i when there are m firms active in the market. Then, using the notation $Q = \sum_{i=1}^m q_i$, firm i maximises

$$\begin{aligned} \max_{q_i} \quad \Big[& \quad \pi_i^m(s) + \mu_i \cdot (s - \theta + \epsilon)Q \\ - & \quad (\frac{q_i}{Q} - \mu_i)\epsilon Q + (\frac{q_i}{Q} - \mu_i)(s - \theta + \epsilon)Q\Big], \end{aligned}$$

where the first two terms represent downstream profit and the share in the upstream profits given the initial capital share μ_i and the last two expressions show the price paid for the adjustment of capital and the dividend received in return. We now simplify the algebra taking into account that $\pi(s) = (a - bQ - s - d)q_i$ implies

$$\pi_i^m(\theta + \Delta) = \pi_i^m(\theta) - \Delta q_i.$$

Then, for the case $s(\theta) = \theta + \Delta$, the above maximisation problem turns out to be equivalent to the firm's maximisation problem in the absence of cost misrepresentations, i.e. $s(\theta) = \theta$:

$$\max_{q_i}\left[\pi_i^m(\theta) + \mu_i \epsilon Q\right].$$

The cost signal has cancelled out. Therefore, output decisions are no longer affected by cost misrepresentations.

We now show that no firm has an incentive, if entitled to do so, to choose an access price different from $\theta + \epsilon$ (part 1). Suppose firm i is choosing the cost signal $s(\theta)$. It knows that the choice of this signal will not affect output decisions. Taking into account the capital adjustment in period 2, choosing $s = \theta + \Delta$ would be profitable if and only if

$$\frac{q_i}{Q} \cdot \Delta \cdot \left(\sum_{j=1}^{m} q_j\right) > \mu\ K + \pi_i^m(\theta) - \pi_i^m(\theta + \Delta).$$

The left - hand side of this inequality equals $\Delta \cdot q_i$; the right - hand side equals $\mu K + \Delta \cdot q_i$. This is because quantity choices are not affected by the imposition of higher access charges $s(\theta) = \theta + \Delta$ (follows from part 2 of this proposition). Therefore, it is never profitable for any downstream firm to charge higher access prices. If accounting manipulations are costless ($K = 0$) the firm would be indifferent between $s(\theta) = \theta$ and $s(\theta) = \theta + \Delta$ (and, in fact, any other access charge).

Q.E.D.

Remark 1

1. *Note that the price at which shares must be traded is verifiable. It depends on ϵ only, the component of U's profit that is known to and fixed by the regulator.*

2. *The cost signal cancels out but ϵ does not. Equilibrium output is given by*

$$q_i = \frac{a - \theta - d}{m+1} - (\frac{2}{m+1} - \mu_i)\epsilon$$

which implies that a higher ϵ would restrict output. But ϵ is chosen by the regulator who is aware of this effect.

3. *Output is restricted by an amount $\frac{m-1}{b(m+1)}\epsilon$ and is therefore independent of the initial distribution of capital shares μ_i.*

The intuition behind the above proposition is the following: if the capital shares equal the respective output shares a downstream loss due to a higher access charge would exactly be recovered via the share in the upstream profit.

This is the borderline case in which the role of a regulator in choosing and controlling access prices would be marginalised.

An argument put forward frequently against a joint ownership concept is that it might facilitate collusive behaviour (see[17] Reitman (1994) and references given there). One might suppose that firms agree to choose a high cost signal in order to reduce total output. The following corollary states that this is not the case in our setting.

Corollary 1 *Under joint ownership with optimally adjusted ownership shares firms cannot commit to restrict total output by choosing a high cost signal. Joint ownership, therefore, does not induce collusion.*

The proof of this corollary is simply this: suppose firms D_i agree to choose a high cost signal. (It is irrelevant here if one takes $s = \theta + \Delta$ or, for instance, s such that a share of the monopoly quantity is produced by each downstream firm D_i). Then it follows immediately from part (2) of the previous proposition that downstream output is not affected (as the additional costs are recouped via access charges).

Therefore, to substantiate the claim that joint ownership induces collusion one would have to find different mechanisms that may be

[17]The antitrust literature on essential facilities pursues this point, see for instance Reiffen and Kleit (1990) and references given there.

activated by JO. But JO does not represent in itself a commitment device helping firms to monopolise the market. But it is essential that capital shares are adjusted optimally. Otherwise, collusion arises in equilibrium and incentives to misrepresent costs exist. In this case of exogenous shares one would have to find different mechanisms to solve the problem.

On the other hand, although the adjustment rule must be implemented, i.e. capital shares must be endogenous this does not imply that adjustments take place in equilibrium. In fact, as the following corollary states this is not the case in a symmetric setting.

Corollary 2 *If there are m symmetric downstream firms (identical production costs) with initial capital shares $\mu_i = \frac{1}{m}, i = 1, \cdots, m$, and the regulator implements the rule described in the previous proposition then no capital adjustments occur in equilibrium.*

Interestingly, the findings of this chapter correspond closely to arrangements adopted in the following two settings. Firstly, as reported in Hillman (1991), the guideline often agreed upon in the oil industry is that each firm using the pipeline network should hold a capital share in the network operator that would equal the anticipated output share. Secondly, similar arrangements have been described for deepwater ports by Lewis and Reynolds quoted above. So the above proposition provides a welfare based rationale for such practices and establishes the exact timing (of competition, adjustment, and dividends) under which these arrangements are optimal.

5 Equilibrium with discrimination

There are two obvious cases from which discriminatory access pricing could arise, namely the cases where there is either accounting separation but no controllability or no accounting separation at all. The former case seems to be the more subtle but also the more interesting one and will therefore be fully analyzed in what follows. We will then be able to argue that both cases, perhaps surprisingly, lead to equivalent outcomes.

5.1 Accounting separation without controllability

Dropping the assumption of controllability implies that, under PI, the integrated firm is now free to choose its profit maximising output. Taking into account that, due to accounting separation, the official access prices for both firms are identical and given by the cost signal $s \in \{\theta, \theta + \Delta\}$ the maximisation problem for firm $\{D_1, U\}$ in $t = 2$ is given by

$$\max_{q_1} \Big[(a - b(q_1 + q_2) - s - d_1)q_1 + (s - \theta)(q_1 + q_2) \Big] \tag{8}$$

where the second term represents the additional (upstream) profit that arises whenever cost signal $s = \theta + \Delta$ is given in period 1. Firm D_2 maximises profits given cost signal s:

$$\max_{q_2}(a - b(q_1 - q_2) - s - d_2)q_2.$$

Equilibrium output levels are then given by

$$q_1 = \frac{a + s - 2d_1 + d_2 - 2\theta}{3b}$$

and

$$q_2 = \frac{a - 2s - 2d_2 + d_1 + \theta}{3b}.$$

This implies that, for $s = \theta$, we have $q_i = \frac{a-\theta-2d_i+d_j}{3b}$ and, for $s = \theta + \Delta$, we obtain

$$q_1 = \frac{a - \theta - 2d_1 + d_2}{3b} + \frac{\Delta}{3b}$$

and

$$q_2 = \frac{a - \theta - 2d_2 + d_1}{3b} - \frac{2\Delta}{3b}.$$

Remark 2 *Under PI without controllability a higher access charge induces asymmetric output levels in favour of the network owner and reduces total output.*

In $t = 1$ the integrated firm chooses the signal about the cost observed in $t = 0$. It is profitable to misrepresent costs whenever

$$\Delta q_2 \quad > \quad K + \pi_1(\theta, \theta) - \pi_1(\theta, \theta + \Delta) \tag{9}$$

where $\pi_1(s_1, s_2)$ denotes firm $D_1's$ profit when the signal s_i is charged to firm D_i.

Welfare implications

In the absence of effective accounting separation new incentives for charging higher access prices arise. The possibility of discrimination provides a second motive for raising access charges above true costs. Not only does this strategy generate profits for the upstream division but it also raises the rival's operating costs. Unless the competitor is too inefficient discrimination can be shown to lead to lowest welfare. This is made precise in the following proposition.

Proposition 3 *In a setting where downstream firms' marginal production costs d_1, d_2 satisfy $d_2 < d_1 - \frac{bK}{\Delta}$, partial integration without accounting separation yields lowest welfare:*

$$PI^N \prec NI.$$

The intuition for this result and for the parameter restriction imposed on costs are as follows. As the controllability assumption is not satisfied the integrated firm $\{U, D_1\}$ is able to charge $s = \theta + \Delta$ and thereby to collect additional revenues Δq_2 and, at the same time, to obtain an advantage in the downstream market by raising the rival's costs. Only if the rival's market share is too small the revenue from high access charges become too small compared to the revenue in the non-integration case. This case is excluded by the restriction on costs which should not be viewed as too strong. In fact, the same result obtains in various different versions of the model. It is sufficient, for instance, to model the cost of misrepresentation as a per unit cost in which case (9) becomes $(\Delta - K)q_2 > \pi_1(\theta,\theta) - \pi_1(\theta,\theta+\Delta)$ and the same conclusion obtains. Alternatively, one may consider a symmetric setting with more than one rival firm.

Proof. Under PI^N downstream profits contingent on the signals charged are given by

$$\pi_1(\theta,\theta) = \frac{1}{b}\Big(\frac{a-\theta-2d_1+d_2}{3}\Big)^2$$

and

$$\pi_1(\theta,\theta+\Delta) = \frac{1}{b}\Big(\frac{a-2d_1-\theta+d_2+\Delta}{3}\Big)^2$$

which implies

$$\pi_1(\theta,\theta) - \pi_1(\theta,\theta+\Delta) = -\frac{2\Delta}{3}\cdot\frac{a-\theta-2d_1+d_2}{3b} - \frac{\Delta^2}{9b}.$$

Again, condition (9) can be shown to be satisfied if and only if θ does not exceed a critical value $\bar{S}(PI^N)$ where

$$\bar{S}(PI^N) = a - \Delta - \frac{4d_2+d_1}{5} - \frac{9bK}{5\Delta}.$$

A comparison with (5) shows that $\bar{S}(NI) < \bar{S}(PI^N)$ under the assumptions on costs made above. The remainder of the proof is identical to the last part of the proof of proposition 1: a higher critical value implies higher probabilities of cost misrepresentation which, in turn, implies higher expected cost signals and therefore lower expected output and higher price level, i.e. lower welfare.

Q.E.D.

This confirms that, under regularity conditions, discrimination has the expected negative effects on welfare.

5.2 Equilibrium without accounting separation

If there is no accounting separation at all there would be no reason for the integrated firm to calculate its profits on the basis of the cost signal (possibly $s = \theta + \Delta$) but it would always assess its own costs by using the observed parameter θ:

$$\max_{q_1}\ (a - b(q_1+q_2) - \theta - d_1)q_1 + (s-\theta)q_2.$$

But this problem is equivalent to (8). Therefore, accounting separation does not affect the integrated firm's decision problem as long as there is no controllability.

The intuition behind this result is simple: the integrated firm realises that accounting separation will result in a mere transfer payment from D_1 to U without affecting the downstream behaviour at all. Therefore, the welfare obtained under accounting separation without controllability is the same as in the case without accounting separation.

The previous proposition demonstrates how crucial the controllability assumption is for political recommendation: only if the regulator can enforce that $\{U, D_1\}$ base their output decision on the published access price, i.e. controllability holds, partial integration dominates NI. Otherwise, partial integration is equivalent to PI^N and thus worse than NI which would, of course, be a strong argument in favour of divestiture.

The fact that the profit of the integrated firm is not affected by accounting separation may well explain the way in which BT reacted to the introduction of accounting separation. Instead of the usual opposition to Oftel's regulatory demands BT approved claims to separate accounts. "BT condemned Oftel's 'costly increase in regulation'. But Mr Michael Hepher, BT managing director, has assured Oftel of BT's co-operation in the introduction of accounting separation which will apply from this financial year."[18]

Given the above findings this consent is not surprising.

6 Investment Decisions

The following two subsections deal with investment decisions taken by either upstream or downstream firms. We analyze how incentives to invest vary under different ownership structures.

[18] Financial Times 9/3/1994.

6.1 Downstream Investment

We model downstream investments as follows: firm D_i chooses an investment level e_i which reduces its constant marginal costs: $c_i = s_i + d - e_i$ [19] where s_i is the signal given by the network owner to firm i. In order to concentrate on asymmetries resulting from ownership induced differences in investment incentives we assume that downstream firms are symmetric as far as initial production costs are concerned: $d_1 = d_2 = d$. D_i incurs investment cost (or 'disutility in monetary units') given by an increasing and convex function $\psi(e_i)$. In what follows we take $\psi(e) = \frac{1}{2}e^2$.

We assume that the cost signal is known before investment decisions are taken. This timing reflects that access charges are determined for long periods and are regarded as given when firms invest. The timing is therefore as follows:

Figure 4

$t=0$	$t=1$	$t=2$	$t=3$
θ realized.	Network operator gives signal about costs.	Levels of (downstream) investment chosen.	Cournot competition downstream.

We first solve the last two stages of this game for NI. In $t = 3$ we obtain

$$q_i^{NI} = \frac{a - s - d + 2e_i - e_j}{3b} \tag{10}$$

and, in $t = 2$,

$$e_i^{NI} = e_j^{NI} = \frac{4}{9b-4}(a - s - d) \tag{11}$$

resulting in output levels

$$q_i^{NI}(e_i^{NI}) = \frac{3}{9b-4}(a - s - d). \tag{12}$$

[19]This assumption of constant marginal costs being reduced by the amount of effort undertaken is quite common in the literature, for instance Laffont and Tirole (1993).

We now analyze the case of JO which will yield the results for PI as a special case, i.e. by setting $\mu_1 = 1$.[20] Given JO firm D_i's decision problem in $t = 2$ is given by

$$\max_{e_i} \pi_i(e_i, e_j) + \mu_i(s - \theta)[q_i(e_i, e_j) + q_j(e_j, e_i)] - \frac{1}{2}e_i^2, \quad i = 1, 2, \ i \neq j.$$

This results in

$$e_i^{JO} = \frac{4}{9b-4}(a - s - d) + \frac{3}{9b-4}(\mu_i - \mu_j)(s - \theta).$$

For the case of PI, i.e. $\mu_1 = 1$ and $\mu_2 = 0$, this implies

$$e_1^{PI} = \frac{4}{9b-4}(a - s - d) + \frac{3}{9b-4}(s - \theta)$$

and

$$e_2^{PI} = \frac{4}{9b-4}(a - s - d) - \frac{3}{9b-4}(s - \theta)$$

Given these results we can conclude that

1. If truth-telling occurs ($s = \theta$) incentives to invest are identical under all ownership structures.

2. Non-truth-telling ($s = \theta + \Delta$) results in lower levels of total investment levels under all ownership structures. This makes ownership structures under which cheating tends to occur even more inefficient.

3. In case of PI: even though there may be controllability and therefore output levels based on the given cost signal (instead of true costs) there are asymmetric investment incentives. The reason is simple: although, in $t = 3$, $\{U, D_1\}$ cannot take into account the upstream profit (arising from 'cheating') it can do so when choosing investment levels. By investing more in $t = 2$ a higher output (due to higher investment and therefore lower marginal costs) will be justified in $t = 3$. Thus, the barrier to cheating that

[20] Formally, one also obtains the above results for NI by setting $\mu_1 = \mu_2 = 0$ in the solutions for JO.

was implemented via accounting separation and controllability can be circumvented via investment decision. Even a regulator who believes that controllability of output levels is a reasonable assumption must realize that these arduously erected barriers can be at least partly neutralized by other strategic decisions of the integrated firm where controllability does not apply.

The only way out would be to introduce the assumption of *controllability of investment levels* which is simply not possible for non-verifiable investment decisions.

6.2 Upstream Investment

We consider a setting in which an effort or investment level, e, is to be chosen by the network operator. At least part of the benefits of these investments accrues to downstream firms in one of the following possible ways:

1. $p = (a + e) - bQ$ (marketing, higher network capacity, etc.)
2. $c_i = d - e + s$ (providing better network quality and thus reducing downstream marginal costs).

These two specifications, although quite different in interpretation, will turn out to be formally identical. In what follows we assume that investment decisions are taken in $t = 2$, i.e. the timing is the same as in the section on downstream investments (see figure 4). The equilibrium of this game is as follows.

In $t = 3$, the equilibrium output levels under any given ownership structures are given by

$$q_i = \frac{a + e - s - d}{3b}.$$

This result is obtained for demand increasing investments ($p = a + e - bQ$) as well as for cost reducing investments ($c_i = d - e + s$). In $t = 2$, the levels of upstream investments are chosen by the firm operating the network. Here, we have to analyze the different ownership structures separately.

Non-integration:

The problem of choosing investment levels upstream in $t = 2$ is given by

$$\max_e \Big[(s(\theta) - \theta) \cdot (q_1(e) + q_2(e)) - \frac{1}{2} e^2 \Big].$$

The resulting first order condition implies

$$e^*_{NI} = \frac{2(s - \theta)}{3b}.$$

Partial Integration:

$$\max_e \Big[(s - \theta) \cdot (q_1(e) + q_2(e)) + \pi_1(e) - \frac{1}{2} e^2 \Big].$$

The first order condition

$$\frac{2(s-\theta)}{3b} + \frac{2}{9b}(a + e - s - d) = e$$

implies

$$e^*_{PI} = \frac{6(s - \theta) + 2(a - \theta - d)}{9b - 2}.$$

Joint Ownership, $\mu \geq \frac{1}{2}$:

$$\max_e \Big[\mu(s - \theta) \cdot (q_1(e) + q_2(e)) + \pi_1(e) - \frac{1}{2} \mu e^2 \Big].$$

The first order condition here implies

$$e^*_{JO,\mu} = \frac{6\mu(s - \theta) + 2(a - \theta - d)}{9\mu b - 2}.$$

which, in turn, implies that for $\mu^* = \frac{1}{2}$ we obtain

$$e^*_{JO,\mu^*} = \frac{6(s - \theta) + 4(a - \theta - d)}{9b - 4}.$$

Efficient investment levels

Efficient levels of investment that would take into account the effect of an investment decision on *all* firms would maximise the total surplus given by

$$S = (s - \theta)(q_1(e) + q_2(e)) + \pi_1(e) + \pi_2(e) - \frac{1}{2} e^2.$$

The first order condition implies

$$e^* = \frac{6(s-\theta)+4(a-\theta-d)}{9b-4}.$$

Given these results it is easy to see that there is underinvestment in all cases except for the joint ownership case with $\mu = \frac{1}{2}$. Under non-integration investment levels are lowest, and partial integration is an intermediate case. All this follows from simple algebra.

Writing down the conditions under which cost misrepresentation in $t = 1$ is profitable given the different ownership structures would be a repetition of the analysis in section 4.1. Instead, we formulate our main points here in the following proposition and its corollary.

Proposition 4 *Suppose the network operating firm chooses investment levels that either augment market demand or decrease marginal costs in the ways specified above, then there is underinvestment under all ownership structures $\sigma \in \{NI, PI, (JO,\ \mu \in (\frac{1}{2}, 1])\}$. Investment levels are lowest for NI and highest for (JO, μ):*

$$e^*_{NI} < e^*_{PI} < e^*_{JO,\mu} < e^*.$$

Corollary 3 *In the setting of the preceding proposition efficient investment levels are obtained for the case of joint ownership with capital shares $\mu^*_i = \frac{1}{2}(= \frac{q_i}{q_1+q_2})$.*

$$e^*_{JO,\mu^*} = e^*.$$

This corollary provides an additional rationale for the implementation of joint ownership with capital shares corresponding to production shares: not only do incentives to misrepresent costs or to choose distorted output levels disappear (for $\mu^*_i = q_i/Q$ endogenous) but also efficient levels of upstream investment are induced (for $\mu^*_i = q_i/Q = 1/m$[21] exogenous). This applies to investments that increase market demand as well as to investments that reduce downstream marginal costs of both firms.

[21] Due to symmetry.

Remark 3 *These results partly rely on the assumption that, in the case of joint ownership, the cost incurred is only a share* μ *of the investment costs:* $\mu\frac{1}{2}e^2$*. This would not be plausible if* e *is thought of as an effort decision but only if the investment costs reduce the hidden profits of the upstream firm. Otherwise, in the above calculations, we would have* $\mu = 1$ *and therefore* $e^*_{NI} < e^*_{PI} = e^*_{JO,\mu} = e^*_{JO,\mu^*} < e^*$.

But even in this case JO would not be dominated by any other ownership structure. The intuition behind these results is simply that, for each party taking an investment decision, the share in marginal costs of investment should correspond to the share in marginal profit generated by these investments. This is the case when capital shares correspond to market shares ($\mu_i = \frac{q_i}{Q}$) or, for symmetric firms, when capital shares are equal: $\mu_i = \frac{1}{m}$, $i = 1, \ldots, m$.

7 ENTRY

In the preceding sections we analyzed a setting in which the network is used by a downstream duopoly (or, as in proposition 2, by m downstream firms). The welfare implications of ownership structures derived there rely on the assumption that the number of downstream firms is fixed. Liberalisation of formerly monopolistic or duopolistic (like telecommunications in Britain) industries, however, requires that new competitors can enter the market. Therefore the question arises for us which of the ownership structures discussed so far performs best in settings where entry is possible. We thus allow for entry of additional downstream firms into the industry. This is of particular interest because the ownership structure of the industry will turn out to have an impact on the motivation to deter entry from the industry. Such effects should be incorporated into a welfare analysis of these ownership structures.

Among the traditional ways of deterring entry there are low prices, high capacities (or even low capacities, following Benoit (1991)) or high levels of R&D or advertising outlays aiming at comparative advantages in cost structure or in demand for the advertised good. In our framework the upstream firm has additional opportunities. If it manages to claim higher than the true costs for the network operation it can influence a potential entrant's marginal costs. This, in turn, affects the entrant's profit and, possibly, the entry decision.

We assume that two firms, D_1 and D_2, are active in the downstream market and there are potential entrants, $D_3, \cdots, D_I$, that incur entry sunk costs $E_3 < \cdots < E_I$. Allowing sunk costs to vary across firms would reflect that there may be firms that are already active in related industries or foreign companies active in the same industry and thus have different prior knowledge and incur different levels of R&D costs. Alternatively, one could allow for different levels of marginal costs of new entrants but this would only complicate notation and analysis without changing results qualitatively. Further, we assume controllability (definition 1).

The firm controlling the network (U in case of NI, $\{U, D_1\}$ in case of PI, and D_i with $\mu_i \geq \frac{1}{2}$ in case of JO) gives a signal about observed upstream per unit costs. Given this signal entry decisions are taken by firms D_i, $i \in \{3, \cdots, I\}$. The equilibrium number of firms is assumed to be the largest integer number of firms for which each firm's profit covers the entry costs incurred. There is Cournot competition in the last stage. The timing is shown in figure 5:

Figure 5

$t = 0$	$t = 1$	$t = 2$	$t = 3$
θ realized.	Network operator gives signal about costs.	Entry decisions taken by firms D_i, $i = 3, \cdots, I$.	Cournot competition downstream.

Equilibrium without discrimination

For $t = 3$ the Cournot equilibrium is given by

$$q_i^* = \frac{a - 2c_i + c_j}{b(m+1)} = \frac{a - s - d}{b(m+1)}$$

and

$$\pi_i^* = \frac{1}{b}\left(\frac{a - s - d}{m+1}\right)^2,$$

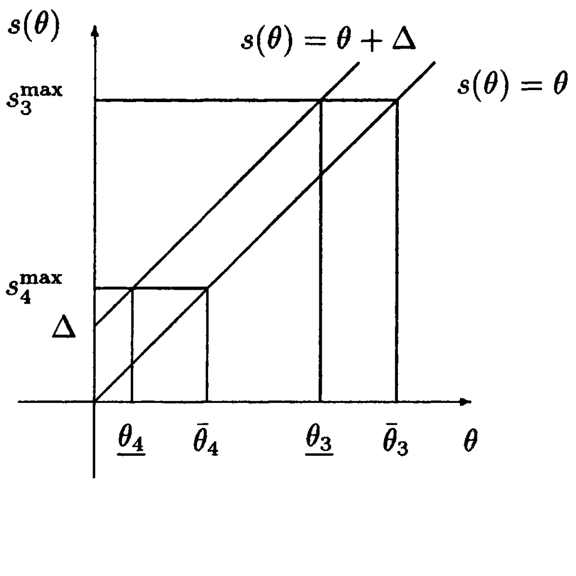
where m is the number of firms in the market.

In $t = 2$, firm D_i enters if and only if $\pi_i \geq E_i$. The equilibrium number of firms is thus given by

$$m^* = \max\left\{2; \max\{m | m \leq \frac{a - s - d}{(b \cdot E_m)^{1/2}} - 1, m = 3, \cdots, I\}\right\}.$$

In $t = 1$, the firm operating the network chooses a signal $s \in \{\theta, \theta + \Delta\}$ about its per unit costs observed in $t = 0$. This choice may affect the equilibrium number of downstream firms. Figure 6 shows the intervals $\Theta_i = [\underline{\theta}_i, \bar{\theta}_i]$ in which entry of firm i can be deterred by choosing cost signal $s(\theta) = \theta + \Delta$.

Figure 6



The maximum signal that would make entry profitable for D_i is given by

$$s_i^{\max} = a - d - (m+1)(bS_m)^{1/2}.$$

The intervals Θ_i are defined by

$$\bar{\theta}_i = s_i^{\max} = \underline{\theta_i} + \Delta$$

and, to simplify the analysis, we assume they are disjoint. We denote by Θ the union of the intervals Θ_i.

For $\theta \notin \Theta$ entry cannot be deterred. The conditions for profitability of cost misrepresentations are qualitatively the same as those in the case without entry.

For $\theta \in \Theta$ it must be true that $\theta \in \Theta_m$ for some m. Thus entry of firm D_i can be deterred by choosing the high signal, i.e. $s(\theta) = \theta + \Delta$. But this is not necessarily profitable. Given an ownership structure $\sigma \in \{NI, PI, JO\}$ it is profitable to misrepresent costs under the following conditions:

$$\begin{aligned}
NI : \Delta(q_1+, \cdots, +q_{m-1}) &\geq K. \\
PI : \Delta(q_1+, \cdots, +q_{m-1}) &\geq K + \pi_i^m(\theta) - \pi_i^{m-1}(\theta + \Delta). \\
JO : \mu\Delta(q_1+, \cdots, +q_{m-1}) &\geq \mu K + \pi_i^m(\theta) - \pi_i^{m-1}(\theta + \Delta).
\end{aligned}$$

There are quantities of $m - 1$ firms on the left-hand side because, for $\theta \in \Theta_m$, entry of firm m is deterred.

Denote by $\epsilon(\theta)$ the difference in downstream profits in state θ, i.e. $\epsilon(\theta) := \pi_i^m(\theta) - \pi_i^{m-1}(\theta + \Delta)$. We distinguish the following three cases:

1. $\epsilon(\theta) > 0$: in this case entry deterrence is not profitable in the sense that downstream profits under the true signal θ with m firms would exceed downstream profits under the high signal $\theta + \Delta$ with $m - 1$ firms. This leads us to exactly the same conclusions as in the analysis where the number of firms in the market was fixed: incentives for misrepresentation of costs are highest for NI, second highest for PI, and lowest for JO.

2. $\epsilon(\theta) < 0$: In this case entry will be deterred. The order of the critical values $\bar{S}(\sigma)$ is reversed: $\bar{S}(JO) > \bar{S}(PI) > \bar{S}(NI)$. So is the welfare ranking.

3. $\epsilon(\theta) = 0$. In this (borderline) case incentives to misrepresent costs are the same under all ownership structures $\sigma \in \{NI, PI, JO\}$.

There are two opposite effects on a firm already active in the market which are connected with entry deterrence. Firstly, the firm suffers from the higher signal (same effect as in the 2-firm case). Secondly, there is the gain arising from a reduction in the number of firms (here by 1, i.e. to $m-1$, as the intervals Θ_i are disjoint). The inequality

$$\epsilon(\theta) = \pi^m(\theta) - \pi_i^{m-1}(\theta + \Delta) < 0,$$

i.e. case 2, holds if and only if the second effect dominates the first. This condition is satisfied if $\Delta < \frac{a-\theta-d}{m+1}$. We summarise the welfare implications in the following proposition.

Proposition 5 *If entry deterrence is possible and profitable then the welfare ranking of proposition 1 is reversed: non-integration dominates partial integration and partial integration dominates joint ownership. Otherwise, i.e. if entry deterrence is not possible or not profitable, the results of the previous sections (no entry case) remain unchanged.*

The intuition for this result is clear. The network operator is interested in high output levels downstream and therefore in entry but downstream firms are interested in high downstream profits and therefore in entry deterrence. This implies that deterrence of potential entrants is more profitable for ownership structures that involve a higher emphasis on downstream activities. This entry deterrence effect may or may not (depending on θ) dominate the ownership structure effect worked out previously.

Efficient outcomes with entry

Having established the above welfare rankings in a setting with entry two important questions arise. Firstly, JO is usually best but, with entry, it may be worst. Does this risk imply that JO should be rejected as concept for regulated industries? Secondly, we may ask just as in the framework with a fixed number of downstream firms: is there an ownership structure under which truth - telling is always the equilibrium outcome? Both of these questions will actually be answered in one by the following argument.

It is possible to show that the optimal JO concept of proposition 2 can be generalised such as to allow for entry. The timing would be as follows. In $t = 0$, θ is realised. Initial capital shares are $\mu_1, \cdots, \mu_m$ with $\mu_1 + \mu_2 = 1$ and the shares of potential entrants equal zero. In $t = 1$, the cost signal is given by the firm with the largest capital share. In period $t = 2$, entry decisions are taken. Period $t = 3$ corresponds exactly to the last period in figure 3: output decisions are taken, then capital shares of all firms in the market are adjusted, and dividends are paid. In this setting the following corollary to proposition 2 and to corollary 1 holds.

Corollary 4 *In an industry with free entry as described above joint ownership with optimal adjustment of capital shares yields efficient outcomes in the sense that no downstream firm has incentives to charge an access price different from the true costs of access. Output levels do not depend on the access price. Firms cannot use a higher signal about costs as a commitment device for collusion.*

Proof. Exactly as in the proof of proposition 2 the profit maximisation problem of firms $D_1, \cdots, D_m$ in $t = 3$ can be shown not to depend on the cost signal. Therefore, the entry decisions do not depend on the signal either ($t = 2$). Therefore, the choice of the signal in $t = 1$ is irrelevant, i.e. there are no incentives for non-truth-telling strategies.

Remark 4 *The above corollary implies that also the second motive for charging high access prices, namely the deterrence of entry, disappears if capital is adjusted optimally at the end of the trading period.*

8 Conclusions and Extensions

Access pricing theory has often treated the access problem in isolation from either informational asymmetries or differences in ownership structures or both. We believe, however, that both factors are highly relevant to the problem. In this chapter we have shown how ownership affects incentives to exploit informational asymmetries.

In the basic model with two downstream firms the following welfare ranking is established. Joint ownership dominates partial integration which dominates non-integration. But the latter result is true only if the regulator can ensure that the integrated firm bases its output decision on the published access price (which, as we have argued, is highly unlikely to be the case). Otherwise, partial integration performs worst because of discriminatory practices. These results are underlined by the fact that the same order turns out to describe the degree to which underinvestment (in the case of upstream investment) occurs. The case of downstream investment shows how controllability of output levels can be circumvented with the help of investment decisions. This provides a further argument against the hypothesis that accounting separation can rule out discrimination.

In a setting with entry, however, this ranking is preserved only if incentives to deter entry do not become too strong. Otherwise this ranking is reversed: non-integration dominates partial integration which, in turn, dominates joint ownership. We therefore argue that in assessing welfare implications of ownership structures in regulated industries it is generally important to take into account industry specific features such as desirability of entry and possibilities and profitability of entry deterrence.

Joint ownership with optimal capital adjustment (such that capital shares equal output shares) turns out always to yield efficient outcomes in the sense that incentives to misrepresent costs disappear entirely. Collusion cannot be sustained by choosing higher cost signals. This result holds irrespective of whether or not there is entry into the industry. Further, it provides a welfare based rationale for forms of joint ownership that occur in some private industries.

Regulators should consider possibilities to implement such joint ownership structures.

Among the extensions one could consider there are, for instance, product differentiation and network effects (see Willig (1979)) both of which might provide arguments for lower access prices and more entry. An especially interesting and relevant line of research would be an extended model that takes into account competing networks. Due to the development of new technologies in some industries it becomes increasingly relevant to deal with upstream oligopolies. Our conjecture here would be that competing networks may mitigate some of the problems that are associated with the pricing of access to a monopolistic network (under non-integration or partial integration). But a single network owned jointly would outperform a situation of competing network capacity under any of the other vertical structures.

Another interesting question arising from the analysis contained in this chapter is related to the use of mechanisms. One of the items on our research agenda is to investigate in how far the various inefficiencies arising in some of the settings discussed may be overcome if appropriate mechanisms are implemented. A mechanism avoiding collusive behaviour in the case of joint ownership with exogenous or non-optimally determined capital shares would be of special interest.

Appendix to Chapter III

The Nonequivalence of Accounting Separation and Structural Separation as Regulatory Devices [22]

Introduction

Suggestions supported by conventional wisdom not always coincide with theoretical analysis.

The issue of accounting separation seems to provide another example for this observation. There are properties which may be wrongly attributed to accounting separation by regulators as well as by economists. We feel that, firstly, economists should be as clear as possible about the true impact of ownership structures and supplementary arrangements. Secondly, it is vital for regulators and for all other parties involved in the discussion to have a clear understanding of the effects of regulatory tools. We therefore include, in the form of an appendix, the discussion of a paper presenting a view on accounting separation we tend to disagree with and we analyse the arguments presented there in the light of the analysis conducted in chapter III of this thesis.

In a recent article Cave and Martin (1994) discuss current regulatory policies with regard to accounting separation in Australia and the UK. They assess the costs and benefits of accounting separation and give a detailed account of the practical measures adopted in Australia and in the UK. Their article compares accounting separation with the alternative of structural separation.

Cave and Martin point out that there are significant institutional and political barriers to structural separation. The supposed effect

[22] In this appendix we summarise the arguments presented in Hardt (1995). Some aspects have already been indicated in the main part of this chapter. However, in policy oriented literature different results have been claimed to hold. We analyze some, rather informal, arguments presented there and propose counter-arguments on a similarly informal level, though corroborated by predictions from the model presented in this chapter.

of accounting separation is as follows. 'Proper accounting separation may achieve the same regulatory aims as structural separation $\cdots$ ' and 'there are potentionally significant welfare benefits arising out of economies of scope $\cdots$; however, unless accounting separation is clearly shown not to be a viable regulatory alternative to structural separation why risk the potential for economies of scope?' (p.14).

The simple economic analysis of chapter III provides results that are surprising in the light of what accounting separation is supposed to achieve. Theory [23] predicts that, unless an assumption of controllability is satisfied, accounting separation has no effect on the dominating firm's behaviour, accounting separation does not effectively prevent discrimination of a competing network user, and accounting separation cannot effectively be used to promote entry either. In many ways accounting separation is not equivalent to structural separation. Although both may look equivalent at first sight, their way of functioning economically and their implications (in terms of access prices, output levels and prices, and entry possibilities for potential competitors) differ considerably.

Finally, Cave and Martin outline the dangers of accounting separation caused by informational asymmetries. The extent of these dangers, they conclude, is likely to become clearer over the next few years.

We would like to link this suggestion with the model we have discussed in this chapter.

The impact of accounting separation

Without accounting separation the network operating firm is able to discriminate. It will charge the competitor a high access price and charge its downstream division a lower access price. The high access price charged to competitors leads to

1. increased revenues from access charges

[23] We will point out later that our conclusions are based on a very general argument (namely joint profit maximisation) and not on peculiarities of our model.

2. downstream cost advantage compared to the competitor which, in turn, implies
3. a relatively high share of downstream output and profits accruing to the incumbent.

Now, accounting separation is supposed to deal with these discriminatory reactions, but we would like to argue that this is not the case: suppose the network operator must publish access prices and separate accounts for network operating activities distinguished from downstream activities. This will create a transparency that prevents discrimination, defenders of accounting separation would argue.

In our model, however, the following happens. The dominant firm will continue to charge high access prices to the rival. Further, it will charge its own downstream division the same high access price. But it will not condition the output decision on the published price. The incumbent's downstream division will behave as if the low access price were charged and will be able to offer services cheaper than the competitor (i.e. produce high output in terms of the Cournot model). This strategy maximises the sum of profits of the dominant firm's up- and downstream activities. And this is how an integrated firm behaves: maximising joint profits.

So, what is the effect of the high access price? The high (non-discriminatory) access price plays the role of a mere transfer payment between the incumbent's downstream division and the incumbent's upstream division. It does not affect the incumbent's total profits in any way. The effect of accounting separation is completely neutralised by the rationally behaving incumbent firm.

It is immediately clear that the arguments we applied to accounting separation do not apply to structural separation: the difference is that an integrated firm (with or without accounting separation) maximises joint profits of upstream and downstream activities. Structurally separated firms each maximise their own profits.

These arguments show that accounting separation does not achieve the same ends as structural separation.

We do not conclude, on the other hand, that separating accounts and enforcing the publication of access prices has no effect at all. It does complicate discriminatory practices slightly. In the case of several network users the network operator cannot discriminate between two of the nonintegrated users, but this is not the main problem of discrimination. The problem is discrimination between an independent, i.e. nonintegrated, network user and the network operator's own downstream division. Accounting separation cannot abolish these discriminatory practices at all. Enforcing separate accounts simply means to erect barriers that, with some effort, can be circumvented.

Remedies

From a theoretical point of view, are there any remedies that would make accounting separation an effective tool in regulation? Yes, there are, but the practical applicability is rather doubtful and must be considered carefully.

Theoretically, it is sufficient to assume 'controllability', i.e. to assume that the regulator is able to control the dominant firm's decision in the following way: the regulator must make sure that the output (or price) decision of the dominant firm's downstream division is based on the published access price, not on the true (nonverifiable) costs.

But distinguishing these two strategies will be extremely difficult for a third person (regulator). And what will be even more difficult is the verification of a conclusion reached by the regulator to a firm that has private information about its true costs. How can the regulator prove that the firm would have acted in a different way (lower output, higher prices) if the published access prices were the true costs? These difficulties are considerable. There are, however, some industry features we would expect to evolve if the integrated downstream firm bases its output decision on a lower value than the published access price.

1. If the incumbent bases output decisions on the low access price

but pays the high transfer price the result must be that the upstream profits are high and downstream profits are low, possibly negative.

2. Output of the integrated firm will always remain considerably higher than the rival's.

3. Market prices will remain relatively high, competition will have little effect.

The first of these points would be the clearest indication if we observed it. But the incumbent will have ways to avoid such a clear picture. The upstream firm could undertake investments that might otherwise have been carried out by the downstream division and thus prevent losses downstream. Also, costs accruing in the downstream division might be attributed to upstream activities, again reducing downstream losses (or, equivalently, raising downstream profits).

Items (2) and (3) are features that can actually be observed in several regulated industries. But these facts, although they may be a clear indication to an unbiased party that some discriminating practices are going an, can hardly be seen as a conclusive proof. The incumbent would rather argue, they must be attributed to different sources: the output shares are asymmetric just because of the incumbent's technological and managerial superiority, marketing, etc. Prices are relatively high because it is costly to provide telecommunications services, not because of inherent inefficiencies in the production process or in the determination of access prices.

Therefore, none of the above three items could enable the regulator to ensure that the output level produced by the incumbent is based on the published price and not on a different value. So, there might be some empirical evidence, especially along the lines of items (2) and (3), that would support the thesis that a high transfer price does not affect the incumbent firm's strategy. But such evidence can, at best, provide an indication for us that predictions obtained from our model are realistic. But it will hardly be sufficiently clear-cut to ensure controllability. And without controllability accounting separation has no effect on the incumbent firm as shown in the previous paragraph. Accounting separation can therefore hardly be a viable and effective regulatory policy in this context.

Further issues: investment, entry and evidence

Even if one still assumes controllability there are further complications that can easily destroy the derived effects of accounting separation. As mentioned earlier, investment decisions provide further means for the integrated firm to circumvent accounting separation (even in the presence controllable output prices) and to create asymmetries. This is because incentives for investments depend on the true marginal costs and not on published access prices. Trying to introduce an assumption of 'controllable investment levels' is not only highly unrealistic in general but positively impossible as far as non-verifiable investments are concerned.

In a setting with entry, discrimination has further ramifications. High access prices discourage entry. If entrants perceive that, under accounting separation, there is still scope for an integrated firm to practice discrimination they will (correctly) anticipate low profits. Entry is discouraged by discrimination.

A policy that is designed to encourage entry and competition can hardly be backed up by enforcing accounting separation. There are, in fact, doubts as to the effectiveness of current policy in achieving a reduction of prices on the end user market: 'It remains to be seen whether the new entrants to the liberalised telecommunications market will force down prices for residential users ...' [24]. The above arguments provide a theoretical foundation for such doubts.

If regulatory measures are used which cannot effectively prevent discrimination, the impact of new entrants will remain very limited.

Some kind of casual evidence may be contained in positive reaction

[24] Blackman (1994).

of BT we alluded to earlier. In the light of our economic analysis this reaction is not surprising. It fits exactly into the picture of a rationally behaving firm: the rational firm will not suffer from accounting separation, as accounting separation has no effect (except for a transfer payment within the firm). So, why protest? A rational firm will, instead, give in as far as this matter is concerned and concentrate its bargaining power and efforts on negotiating regulatory measures that do hurt.

Concern

Often, concern has been expressed as to the effectiveness of attempts to achieve lower prices for telecommunications services and to make these services available for a larger share of the world population. Experts have often pointed out [25] that there is the danger that

1. the technologically driven progress in the information sector may not become available to a majority of society but only to a small proportion of it, an informational elite.

2. privatisation and competition do not bring about the effects we hoped for.

In other words, what we are winning on the technological side of things might be lost on the economic side. High prices prevent a huge majority either from using these services at all (in some countries) or, in other countries, from using these to a larger extent and at more affordable conditions.

Our claim is that a wrong assessment of accounting separation as a regulatory tool is part of the problem. Actually, a critical review on the effectiveness of accounting separation was to be carried out in

[25] One might, for instance, quote a recent note by Colin Blackman (1994): 'There is also continuing talk of convergence - between telecommunication, information and broadcasting and between business and domestic use - and the creation of conglomerate telecommunications and cable companies to provide multimedia services... is cited as evidence that we are at last entering the much-vaunted Information Society. I am not so sure: the impact on the vast majority of people has been rather limited...'

Australia at the end of last year (1994) because the opinion emerged that the introduction of accounting separation in Australia has not had the desired effects on Telecom [26].

These remarks shed light on the importance for regulators to be aware of the full effects of regulatory measures. If accounting separation is wrongly regarded to achieve the same goals as structural divestiture this might slow down competition and a more universal use of information technology considerably.

Conclusion

In the presence of informational asymmetries a vertically integrated network operator can engage in discriminatory access pricing.

Accounting separation does not, although this has often been assumed in the literature, solve this problem. With and without accounting separation the same output quantities are produced by the integrated firm. We have shown that the effect of separate accounts merely amounts to a transfer payment within the integrated firm. Under structural separation, on the other hand, the access price is no longer a mere transfer payment but represents the true costs for the separated downstream firm.

While Cave and Martin suppose that accounting separation is either equivalent to structural separation (in the absence of economies of scope) or dominating structural separation (in the presence of economies of scope) we have shown why accounting separation cannot be used effectively to prevent or diminish discrimination while structural separation can.

The arguments summarised in this appendix do not imply that structural separation is the best of all worlds. In fact, as shown in the main part of chapter III, it is not. But it does show that there are benefits of structural separation that need to be taken into account.

It is important for regulators to be fully aware of the economic implications of the measures adopted in a policy aiming at nondiscrimina-

[26] I owe this hint to Ian Martin.

tory access pricing. An incorrect assessment of the effect of accounting separation will lead to higher consumer prices and lower welfare.

9 References

Alger, Dan (1993) : Competitive Joint Ventures For Natural Gas Pipelines; *Mimeo.*

Alger, Dan and Susan Braman (1993) : Competitive Joint Ventures and Reducing the Regulation of Natural Monopolies; *Mimeo.*

Armstrong, M., S. Cowan and J. Vickers (1994) : Regulatory Reform - Economic Analysis and British Experience; *Cambridge: MIT Press.*

Armstrong, Mark and Chris Doyle (1994) : *Access Pricing , Entry and the Baumol-Willig Rule;* Discussion paper no. 9422, University of Southampton.

Armstrong, Mark and John Vickers (1995) : The Access Pricing Problem; *Mimeo. University of Southampton and University of Oxford.*

Baumol, William J. and J. Gegory Sidack (1994) : *Toward Competition in Local Telephony;* MIT Press.

Benoit, Jean-Pierre (1991) : Entry deterrence and dynamic competition; *International Journal of Industrial Organization 9, 477-495.*

Black, F. (1984) : *The Dividend Puzzle;* in Jenson, M.C. and Smith, C.W. (eds.) 'The Modern Theory of Corporate Finance'. McGraw-Hill New York.

Blackman, Colin (1994) : To have and have not; *Telecommunications Policy 18 1, p. 3.*

Bolton, Patrick and Michael Whinston (1993) : Incomplete Contracts, Vertical Integration, and Supply Assurance; *Review of Economic Studies 60: 121-148.*

Byg, Torkild and Michael Hardt (1995) : Access Pricing, Ownership and Contracts; *mimeo, University of Bonn and London School of Economics.*

Cave, Martin and Chris Doyle (1994) : Access Pricing in Network Utilities in Theory and Practice; *Utilities Policy 4: 181-189.*

Cave, Martin and Ian Martin (1994) : The costs and benefits of accounting separation; *Telecommunications Policy 18 1: 12-20.*

Cave, Martin and Roger Mills (1992) : Cost Allocation in Regulated Industries.; *Centre for the Study of Regulated Industries.*

Gale, Ian (1994) : Price competition in noncooperative joint ventures; *International Journal of Industrial Organization: 53-69.*

Hardt, Michael (1995) : The Nonequivalence of Accounting Separation and Structural Separation as Regulatory Devices; *Telecommunications Policy 19 No.1: 69-72.*

Hardt, Michael and Susanne Stürmer (1995a) : Network Access and Ownership Structure; *Mimeo. London School of Economics and Freie Universität Berlin.*

Hardt, Michael and Susanne Stürmer (1995b) : Access Price Regulation, Ownership, and Information. A Survey; *Mimeo. London School of Economics and Freie Universität Berlin.*

Hart, O. and John Moore (1990) : Property Rights and the Nature of the Firm; *Journal of Political Economy 98: 1119-1158.*

Hart, O. and Jean Tirole (1990) : Vertical Integration and Market Foreclosure; *Brookings Papers on Economic Activity, Microeconomics: 205-276.*

Hillman, J.J. (1991) : *Oil Pipeline Rates: A Case for Yardstick Regulation;* Competition and the Regulation of Utilities, Michael A. Crew (Ed.) pp. 71-95. Kluwer Academic Publishers.

Laffont, Jean-Jacques and Jean Tirole (1993) : *A Theory of Incentives in Procurement and Regulation;* MIT Press.

Laffont, Jean-Jacques and Jean Tirole (1994a) : Access pricing and competition; *European Economic Review: 1673-1710.*

Laffont, Jean-Jacques and Jean Tirole (1994b) : Creating Competition Through Interconnection: Theory and Practice; *Mimeo. IDEI and MIT.*

Lewis, Lucinda M. and Robert J. Reynolds (1979) : *Appraising Alternatives to Regulation for Natural Monopolies;* in Edward J. Mitchell, ed., 'Oil Pipeline and Public Policy'. Washington, D.C.: American Enterprise Institute: 135-140.

Littlechild, S. C. (1994) : Competition in Electricity: Retrospect and Prospect; *Lectures on Regulation 1994, The Institute of Economic Affairs.*

Newbery, David (1994) : Regulatory Policies and Reform in the Electricity Supply Industry; *DAE Working paper No. 9421, University of Cambridge.*

Oftel (1994) : Interconnection and Accounting Separation: The Next Steps, March 1994; *London: Oftel.*

Reiffen and Kleit (1990) : Terminal Railroad Revisited: Foreclosure of an Essential Facility or Simple Horizontal Monopoly?; *Journal of Law and Economics, vol.XXXIII: 419-438.*

Reitman, David (1994) : Partial Ownership Arrangements and the Potential for Collusion; *Journal of Industrial Economics XLII: 313-322.*

Vickers, J. and Yarrow, G. (1988) : *Privatization: An Economic Analysis: 73-75;* MIT Press.

Vickers, J. (1995) : Competition and Regulation in Vertically Related Markets; *Review of Economic Studies 62: 1-17.*

Willig, Robert D. (1979) : The Theory of Network Access Pricing; *in: H.Trebing, Ed., Issues in Public Utility Regulation, Michigan State University Press: 105-152.*

CHAPTER IV

Access Pricing, Ownership and Contracts [1]

Abstract

In vertically related industries an integrated incumbent firm owning a bottleneck facility does not necessarily have incentives to grant access to a competing downstream firm, even if the competitor is more efficient. If the competing firm's productivity advantage is small the incumbent firm can obtain its monopoly profit by leaving the entire market to the rival and by charging the optimal access price. A competitor with a large productivity advantage will be left to serve the market alone but the resulting prices are even higher.

The burden of regulation depends on the class of contracts permitted. Two-part-tariffs increase welfare but do not solve the problem of market power. It is also analyzed how ownership affects equilibrium pricing and entry. Duplication of the bottleneck facility, under Bertrand competition, leads to marginal cost pricing. Under vertical separation, price discrimination leads to a reduction in final goods prices. If the access price must be chosen before entry decisions occur this leads to more entry and lower access prices than the reverse of this timing structure.

1 Introduction

This chapter analyzes pricing and entry decisions in markets in which a firm supplies an essential input to other firms and may compete with these firms in the final goods (downstream) market. Typical examples are transport and utility industries.

Final products are regarded as perfect substitutes. This case is es-

[1] This chapter is based on Byg and Hardt (1995).

pecially relevant for basic services in telecommunications (long distance calls, faxes, etc.) where consumers largely regard service providers as equivalent. Electricity and gas are other obvious examples where different firms can hardly engage in product differentiation except for modifications in related services such as itemised bills etc. The assumption of perfect substitutes is therefore a good proxy for a large class of goods provided by utility industries.

An interesting question arising in this context is whether or not a full service provider has incentives to supply a competing downstream firm with the essential input (produced upstream). This question is discussed for various ownership structures and for different relative productivities. Further, the upstream firm's decisions concerning pricing and access provision will be affected drastically by the existence of a second upstream firm.

We analyze simple network structures where different segments may be owned by different firms. Among the arrangements we consider there are vertical integration (with a downstream competitor), vertical separation, and a setting in which there is a second supplier of the essential input, either as separate firm or integrated with the second downstream firm.

In these settings we analyze a game with three stages. First, downstream firms decide whether or not to enter the market. Then,[2] access prices are set by the upstream firm(s). Downstream competition, in the form of simultaneous price setting, takes place in the final stage. Equilibrium pricing and entry decision are derived for various cases that differ in the relative productivity of the firms.

We first investigate how equilibrium pricing and entry are affected by network duplication, ownership structures, and pricing regimes allowed (structure of tariffs and possibility of price discrimination). Starting from this free market analysis we point out implications for regulation of such industries by comparing informational requirements in

[2]An alternative timing structure, namely a reversal of the first two stages is also considered.

these various cases.

Two strands of the literature are especially relevant to the theme of this paper, namely vertical integration and regulation. For the former, surveys are, for instance, Perry (1989) and Waterson (1993). For the latter, Armstrong et al. (1994) provide a comprehensive overview of regulation in theory and practice in various industries. Baumol and Sidack (1994) focus on the telecommunications industry and propose the ECPR (efficient component pricing rule). Laffont and Tirole (1994) provide a comparison of the ECPR with alternative pricing rules.

As pointed out in the introduction and in chapter 2 of this thesis, relatively few contributions focus on the impact of ownership on regulatory requirements. A paper that does discuss various ownership structures in network related industries is Economides and Woroch (1992). Their analysis, however, differs from ours in several aspects. Firstly, Economides and Woroch assume that downstream goods are imperfect substitutes. [3] Further, they consider the case of an upstream monopoly only. More importantly, they do not analyze the impact of ownership, tariff structures, timing etc. on regulation but restrict their analysis to private incentives. Unlike Economides and Woroch, we show that in some cases there are incentives to foreclose a competitor from the market and that divestiture may be socially advantageous because it leads to productive efficiency.

The rest of this chapter is organised as follows. Section 2 presents the basic model and analyzes the case of vertical integration with a downstream competitor when only simple tariffs are allowed. In section 3 general tariffs are shown to reduce the regulatory burden. Section 4 considers the impact of alternative ownership structures on equilibrium pricing and entry and on the regulatory burden. Section 5 concludes.

[3] Their model cannot handle perfect substitutes as a special case because the assumption of perfect substitutes implies a discontinuity of their demand function.

2 Vertical Integration and Simple Tariffs: a Borderline Case

In this section we analyze a situation in which a vertically integrated firm acts as sole supplier of a downstream competitor. This case is relevant due to its frequent occurrence in regulated industries but we use it as borderline case only. The impact of other contracts and allocations of ownership will be investigated in sections 3 and 4.

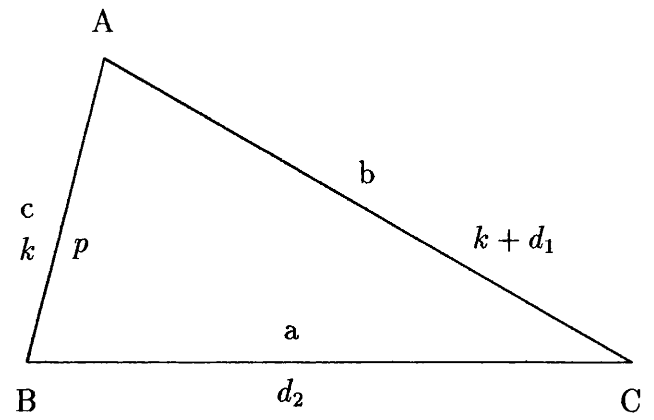
Interconnection charges are given by linear tariffs in our basic model. There are several arguments why it is of interest to look at simple tariffs. This case of linear tariffs plays an important role in some industries as well as in much of the literature in this area. One theoretical reason for the use of linear tariffs appears to be the fact that they rule out arbitrage opportunities that would exist under two-part-tariffs, for instance. Even if regulators allow for two-part-tariffs to be used they will generally not allow the network operator to choose an arbitrary fixed component. Rather, many regulators follow a paradigm of cost based access charges and would therefore limit the fixed component to an amount representing the amount of fixed costs allocated to this particular service.

2.1 A Model

We analyze a network that consists of points A, B, C and the links a,b,c connecting those points with each other.

Figure 1

This simple network structure can be used to analyze various problems in different industries. Most of our interpretations will be phrased as referring to the telecommunications or transport industries. But applications to other network related industries such as gas and electricity supply can be made quite easily although the vertical structures of



those industries are not exactly identical. [4]

One possible interpretation will be that links b and c represent the assets owned by an integrated incumbent (local network and long distance operations) while a is owned by a long distance operator that needs access via c to the local network. Alternatively, c can be viewed as a network that can bypass the incumbent's local network (i.e. second cable network or alternative technologies, e.g. microwave based). The long distance operator may or may not be integrated with this firm. A slightly modified version of the network will be used in section 4.1. This will allow for an analysis of vertical separation of the incumbent.

There are two firms offering 'transport' on this network. In the standard version of the model D_1 is the incumbent running the network lines b and c and can be thought of as being vertically integrated with an upstream firm U owning these links. Firm D_2 owns and runs link a and needs access to the bottleneck facility c.

D_1 operates link b at constant marginal costs $k + d_1$ and link c at marginal costs k. D_2 operates link a at constant marginal costs d_2. D_1 chooses the access price p at which D_2 gets access to link c. Firms D_1, D_2 incur fixed costs F_1, F_2 respectively. In this framework F_1 is given by the sum of fixed costs attributable to the operation of links b

[4] In our analysis it is the upstream activity that is considered to be monopolistic. In the electricity or gas industry the 'upstream' sector may be (potentially) competitive but, still, access to a network is needed. Therefore the conclusions apply although minor rephrasing may be necessary.

and c, $F_1 = F_b + F_c$, and $F_2 = F_a$.

There is demand only for 'transport' from A to C, where the demand schedule is given by $D(p) = \alpha - p$. Transport via B ($\overline{ABC}$) is a perfect substitute for direct transport along $\overline{AC}$. Firms compete by setting prices simultaneously for transport from A to C, i.e. we assume Bertrand competition. D_1 faces demand

$$d_1(p_1, p_2) = \begin{cases} \alpha - p_1 & : \quad p_1 < p_2 \\ \frac{1}{2}(\alpha - p_1) & : \quad p_1 = p_2 \\ 0 & : \quad otherwise \end{cases}$$

The timing of the game is as follows. In stage $t = 0$ a competitor (e.g. long distance operator) decides whether or not to enter the market. In case of entry, sunk costs $S \geq 0$ are incurred. In $t = 1$ the incumbent fixes the access charge. Bertrand competition takes place in the final stage ($t = 2$).

Figure 2

$t = 0$	$t = 1$	$t = 2$
Entry of competitor D_2 in section a of the network. Entrant incurs sunk costs $S \geq 0$.	Access charge fixed.	Bertrand competition of D_1 and D_2 in the market for transport from A to C.

In what follows we analyze equilibrium outcomes in the stages of this game starting in $t = 2$.

2.2 Competition in $t = 2$

In stage $t = 2$ the access price p is given. If D_2 did not enter in $t = 0$ then D_1 will simply charge monopoly prices on link b. We therefore focus on the case in which both, D_1 and D_2, have entered the market (in $t = 0$).

Firm 1

Firm D_1's profit is given by

$$\pi_1(p_1,p_2) = \begin{cases} (\alpha - p_1)(p_1 - k - d_1) - F_1 & : \; p_1 < p_2 \\ \frac{1}{2}(\alpha - p_1)(p_1 - k - d_1) + \frac{1}{2}(p-k)(\alpha - p_1) - F_1 & : \; p_1 = p_2 \\ (\alpha - p_2)(p-k) - F_1 & : \; p_1 > p_2 \end{cases}$$

For any given price p_2 chosen by firm D_2 for the operation of link a firm D_1 decides whether to stay in the market and sell at $p_1 = p_2$ or not to operate link b at all and to make profit simply from giving access to link c. In the former case D_1's profit is given by

$$\pi_1 := \frac{1}{2}(p_2 - k - d_1)(\alpha - p_2) + \frac{1}{2}(p-k)(\alpha - p_2) - F_1.$$

But, as D_1 will just slightly undercut D_2, the profit would be

$$\pi_1 := (p_2 - k - d_1)(\alpha - p_2) - F_1.$$

In case D_1 leaves all transport to D_2 the profit is generated by access charges:

$$\hat{\pi}_1 := (p-k)(\alpha - p_2) - F_1$$

It is easy to see that, no matter which of the two alternative definitions of π_1 is chosen, one always obtains

$$\pi_1 > \hat{\pi}_1 \quad \Longleftrightarrow \quad p_2 > p + d_1. \tag{1}$$

Firm 2

D_2's profit is given by

$$\pi_2(p_1,p_2) = \begin{cases} -F_2 & : \; p_1 < p_2 \\ \frac{1}{2}(p_2 - p - d_2)(\alpha - p_2) - F_2 & : \; p_1 = p_2 \\ (p_2 - p - d_2)(\alpha - p_2) - F_2 & : \; p_1 > p_2 \end{cases}$$

if D_1 is active in the market. D_2 exits if $p_1 < p + d_2$. Otherwise, D_2 will choose

$$p_2 = \min\{p_2^m(p), p_1\}.$$

If D_1 is not active in the market D_2 maximises

$$\max_{p_2} \pi_2 = (p_2 - k - p)(\alpha - p_2) - F_2.$$

These considerations lead to

Proposition 1 *In $t = 2$, there exists a unique Nash Equilibrium given by*

case 1: $d_1 < d_2$

$$p_1 = \min\left\{\frac{\alpha + k + d_1}{2}, p + d_2\right\} \tag{2}$$

$$p_2 = p + d_2. \tag{3}$$

case 2: $d_1 \geq d_2$

$$p_1 = p + d_1. \tag{4}$$

$$p_2 = \min\left\{\frac{\alpha + p + d_2}{2}, p + d_1\right\} \tag{5}$$

Proof.

case 1: $d_1 < d_2$

Given $p_2 = p + d_2$ is charged by firm D_2 firm D_1 will maximise its profits by undercutting or charging the monopoly price in case this is lower, i.e. $p_1 = p + d_2$ if $p + d_2 < p_1^m$ and $p_1 = p_1^m = \frac{\alpha+k+d_1}{2}$ otherwise. Firm D_2 cannot raise its price above $p + d_2$ because in that case D_1 would undercut and D_2 loses its market. Nor can D_2 lower the price because this would imply losses. For uniqueness simply note that the standard Bertrand argument implies that there cannot exist an equilibrium in which $p_2 > p + d_2$. Furthermore, D_2 will never choose p_2 such that $p_2 < p + d_2$.

case 2: $d_1 \geq d_2$

D_1 will not price below $p + d_1$ because of a simple opportunity cost argument: it would always earn margin $p - k$ and save $d_1 + k$ for a unit sold by D_2 to final customers. Nor can D_1 raise its prices above $p + d_1$ because this would be undercut by D_2. In analogy to D_1's behaviour in case (1) D_2 will undercut if necessary and charge the monopoly price otherwise. Again, uniqueness follows.

Q.E.D.

The analysis of $t = 2$ can be interpreted as follows. If D_1 is efficient ($d_1 < d_2$) D_2 will exit because D_1 will undercut. If D_1 is inefficient ($d_1 \geq d_2$) D_1 will want to induce D_2 to produce an output quantity which is as large as possible. D_1's credible threat is to sell at $p + d_1$. So $p_2 = p + d_1$ will obtain since D_1 cannot credibly commit to choose a lower price.

2.3 Access pricing in $t = 1$

We here have to analyze the following four cases. In case (i) firm D_1 is efficient ($d_1 < d_2$). Otherwise ($d_1 \geq d_2$) we distinguish the following subcases: in case (ii) the condition $3d_1 < \alpha - k + 2d_2$ holds. This is the case in which D_2's productivity advantage $d_2 - d_1$ is small which induces D_1 to stay in the market. More specifically, this condition arises from setting D_2's monopoly price greater than D_1's opportunity cost $p + d_1$ and then substituting the optimal access charge p^* derived below. If, on the other hand, $3d_1 \geq \alpha - k + 2d_2$ holds then firm D_1 can no longer induce firm D_2 to serve the market at marginal costs. Here two subcases have to be distinguished: D_1 maximises its profits either by acting as a supplier of firm D_2 (case (iii)) or by serving the market alone (case (iv)). The former happens whenever $d_1 \geq (\alpha - k)(1 - \frac{1}{\sqrt{2}}) + \frac{d_2}{\sqrt{2}}$ as will be demonstrated below.

case (i): $d_1 < d_2$**:** ('Efficient incumbent')

D_1 sets $p = \infty$ and charges the monopoly price in $t = 2$.

case (ii): $d_1 \geq d_2$ **and** $3d_1 < \alpha - k + 2d_2$ **:** ('Competitor with small productivity advantage')

D_1's optimal choice is to pick the maximum of its monopoly profit π^m

(which results by setting $p = \infty$) and the maximum profit obtained from access charges:

$$\max\left\{\max_p\{(p-k)(\alpha-p-d_1)\}, \pi_1^m\right\}$$

Interestingly, the first order condition implies

$$p^* = \frac{\alpha - d_1 + k}{2}$$

which leads to

$$\pi_1(p^*) = \left(\frac{\alpha - d_1 - k}{2}\right)^2 = \pi_1^m.$$

Substitution of p^* into (5) yields

$$p_2^* = p^* + d_1 = \frac{\alpha + d_1 + k}{2} = p_1^m.$$

This result is due to Bertrand competition. It does not depend, as we will state in corollary 2, on the assumptions made concerning the demand schedule. We now analyze

case (iii): $d_1 \geq d_2$ and $3d_1 \geq \alpha - k + 2d_2$ and $d_1 \geq (\alpha - k)(1 - \frac{1}{\sqrt{2}}) + \frac{d_2}{\sqrt{2}}$: ('Efficient competitor serves market alone').

D_1 : solves

$$\max\left\{\max_p\{(p-k)(\alpha-p_2^m(p))\}, \pi_1^m\right\}$$

The first order condition implies

$$p^* = \frac{\alpha + k - d_2}{2}.$$

Substitution yields

$$p_2^m(p^*) = \frac{3\alpha + k + d_2}{4}$$

will be charged in $t = 2$. Firm D_1's profit therefore will be

$$\pi_1 = (p^* - k)(\alpha - p_2^m) = \frac{1}{8}(\alpha - d_2 - k)^2.$$

This profit π_1 exceeds the monopoly profit $\pi_1^m(d_1 + k)$ when

$$\frac{1}{8}(\alpha - d_2 - k)^2 > \frac{(\alpha - d_1 - k)^2}{4}$$

or, equivalently,

$$d_1 \quad > \quad (\alpha - k)(1 - \frac{1}{\sqrt{2}}) + \frac{d_2}{\sqrt{2}} \qquad (6)$$

Further, we had

$$d_1 \quad > \quad \frac{\alpha + 2d_2 - k}{3} \qquad (7)$$

Either of inequalities (6) or (7) may be binding.

case (iv): $d_1 \geq d_2$ and $3d_1 \geq \alpha - k + 2d_2$ and $d_1 < (\alpha - k)(1 - \frac{1}{\sqrt{2}}) + \frac{d_2}{\sqrt{2}}$: ('Efficient competitor foreclosed')

Like in case (iii) firm D_1 is so inefficient that it can only maximise its profit by choosing an appropriate access price if firm D_2 is in the market. Due to the last condition characterising case (iii) now being reversed the maximal profit it can obtain with firm D_2 active is lower than its monopoly profit. Firm D_1 can make the monopoly profit by precluding firm D_2 from entering the market.

The above discussion is condensed in

Proposition 2 *In $t = 1$ firm D_1's optimal choice of access price is given by*

$$p^* = \begin{cases} \infty & : \quad \textit{case (i) and case (iv)} \\ \frac{\alpha - d_1 + k}{2} & : \quad \textit{case (ii)} \\ \frac{\alpha - d_2 + k}{2} & : \quad \textit{case (iii)} \end{cases}$$

The analysis of stage $t = 1$ suggests

Corollary 1

1. *If the incumbent firm is efficient ($d_1 \leq d_2$) it has no incentive to open link c.*
2. *Deterrence of socially desirable entry may occur.*
3. *Depending on the size of a competitor's productivity advantage the incumbent may choose to serve solely as a supplier to D_2. The access price is higher in case (iii) than in case (ii). For α large, market prices are highest in case (iii).*

Proof. Part (1) simply corresponds to case (i) in which link c is not opened. Part (2) describes the outcome of case (iv). The access price in case (iii) must exceed the one in case (ii) because $d_1 \geq d_2$ in both cases. To see that the market price is highest in case (iii) note that $p_2^m(p^*) = \frac{3\alpha+k+d_2}{4}$ exceeds the monopoly price $p_1^m = \frac{\alpha+k+d_1}{2}$ if and only if $\alpha - k - d_1 + (d_2 - d_1) \geq 0$ which, for α large, is always true (in other words: the productivity differential must not exceed the monopolist's 'maximum margin' $\alpha - k - d_1$).

Q.E.D.

In case (iii) p^* is higher than in case (ii). The intuition for this outcome is that firm D_1 cannot credibly threaten to sell cheaper than $p + d_1$. Therefore firm D_1 has to maximise its profit by the choice of a high access price.

Corollary 2 *Given the parameter restrictions of case (ii) firm D_1 can extract monopoly profits either by operating as a monopolist on link b or by giving the whole market to D_2 and charging an optimal access price p^*. This result holds for arbitrary demand schedules.*

Proof. Consider a general demand schedule $D(p)$. If firm D_2 charges price $p_2 < p_1$ and firm D_1 demands access price p then D_1's profit is given by

$$\pi_1(p_2) = (p_2 - k - d_1) \cdot D(p_2)$$

if firm D_1 undercuts firm D_2 or by

$$\hat{\pi}_1(p_2) = (p - k) \cdot D(p_2)$$

if she confines herself to the role of a supplier to firm D_2. D_2's best reply is then given by

$$p_2^*(p) = \min\{p + d_1, p_2^m(p)\}.$$

Let $p + d_1 \leq p_2^m(p)$. Then $\pi_1(p) = (p - k)D(p + d_1)$. Therefore, the first order condition is

$$D(p + d_1) + (p - k)D'(p + d_1) = 0.$$

Substituting $p = p_2^* - d_1$ yields

$$D(p_2^*) + (p_2^* - d_1 - k)D'(p_2^*) = 0$$

which is exactly the monopolist's unique first order condition.

Q.E.D.

Corollary 2 shows that D_1, by charging access price p^*, can obtain exactly her monopoly profit without being directly active in the market.

2.4 Entry in $t = 0$

We now have to analyze stage $t = 0$. Due to our analysis of stages 1 and 2 the following obtains. In cases (i) and (iv) entry never occurs. In case (iv) this leads to inefficiency provided that the entry cost S is not too large. In case (ii) entry occurs if and only if $S \leq \pi_2$. But as

$$\begin{aligned} \pi_2 &= (p_2^* - p^* - d_2)d(p_2^*) - F_a \\ &= \left(\frac{\alpha + d_1 + k}{2} - \frac{\alpha - d_1 + k}{2} - d_2\right)\left(\frac{\alpha - d_1 - k}{2}\right) - F_a \end{aligned}$$

entry occurs if and only if the following condition holds.

$$S \leq (d_1 - d_2)(\frac{\alpha - d_1 - k}{2}) - F_a, \tag{8}$$

which means that entry is more likely for larger cost differentials $d_1 - d_2$ as long as the conditions defining case (ii) are not violated. Analogously, in case (iii) entry occurs if and only if

$$\begin{aligned} S \leq \pi_2 &= (p_2^* - p^* - d_2)d(p_2^*) - F_a \\ &= \left(\frac{3\alpha + d_2 + k}{4} - \frac{\alpha - d_2 + k}{2} - d_2\right)\left(\frac{\alpha - d_2 - k}{2}\right) - F_a \end{aligned}$$

or, equivalently,

$$S \leq \frac{(\alpha - d_2 - k)^2}{8} - F_a. \tag{9}$$

This condition is independent of d_1 because, in case (iii), D_1 cannot credibly threaten D_2 to charge a lower than its monopoly price for the optimal access price. The following proposition summarises the main findings of this section.

Proposition 3 *Inefficient entry does not occur. Efficient entry occurs if the competitor's productivity advantage is sufficiently large. Otherwise, socially desirable entry will not be individually rational. In the latter case inefficient production obtains.*

2.5 Regulatory measures

In the setting described above there might be a need for the following regulatory measures:

1. Force firm D_1 to charge such an access price that firm D_2 is willing to enter the market whenever D_2 is efficient because, in some cases, D_1 might, inefficiently, prefer not to do so. This would improve welfare in that case. For access pricing it would be sufficient for the regulator to know k. Then she could set p equal to k. This implies equilibrium prices $p_1 = p_2 = k + \max\{d_1, d_2\}$.

2. Price regulation of the final goods: in order to obtain efficient final goods prices, however, the regulator would have to fix the final goods prices at $\min\{p_1, p_2\} = k + \min\{d_1, d_2\}$ in addition to the regulation of the access price. This implies, of course, that in case $d_1 \neq d_2$ the inefficient firm does not produce which is a socially desirable result.

Effects of regulation on entry

With the two kinds of regulation described above entry occurs whenever this is efficient i.e., in cases (ii), (iii), and (iv).

Effects of regulation on investment decisions

We now assume that, preceding competition, firm D_1 can choose levels of investment that lead to reductions in per unit costs of either k or d_1.

Without regulation there are full incentives to invest in cases (i), (ii), and (iv). In case (iii) incentives are discontinuous in the following sense. There are zero incentives to invest as long as $d_1 > \frac{\alpha-k+2d_2}{3}$ and $d_1 > (\alpha - k)(1 - \frac{1}{\sqrt{2}}) + \frac{d_2}{\sqrt{2}}$ remain true. There are full investment incentives whenever either of these conditions does not hold.

With full scale final goods price regulation in the sense that

$$p = k + \min(d_1, d_2),$$

on the other hand, there are zero incentives for any of the firms to invest since their profit, due to regulation, is independent of their costs. This changes, of course, when we allow for a time lag between investment decision and regulation. If there is a period in between in which profits can be realised this situation would, to a certain degree, mimic competitive markets as has been pointed out, for instance, in Baumol (1994).

Concluding, one can state that with simple i.e., linear, contracts there is a need for both regulatory measures mentioned above: access price regulation and final goods price regulation. Informational requirements for an implementation of these measures are considerable.

3 General Tariffs

In this section we consider the consequences of allowing for a wider class of tariff contracts in the access pricing stage. We will focus on resulting differences in the equilibrium outcomes and in the need for regulatory intervention.

3.1 A Model

We use the model presented in section 2.1 with the modification that the access price for link c does not need to be a fixed per unit payment but can be any contract agreed in $t = 0$, i.e. before the entry decisions are taken. [5]

[5] If the contract is agreed in $t = 1$ the outcome depends crucially on the modelling of the bargaining game: if firm D_1 makes a take-it-or-leave-it offer to D_2 it tries to extract $\pi_2^m + F_a$ and therefore D_2 will not enter in equilibrium (which may be a socially undesirable outcome). If, on the other hand, D_2 makes the take-it-or-leave-it offer it offers π_1^m. In equilibrium this offer is accepted by D_1.

Interestingly, the following proposition allows us to limit our discussion to the class of two-part-tariffs i.e., contracts (p, z) specifying a payment function of the form $\hat{p}(p, z) = z + pq_2$, where z is a lump sum payment made by D_1 to D_2, p is the per unit access price, and q_2 is D_2's output level.

Proposition 4 *If $d_2 \leq d_1$ then there exists an optimal contract of the form $\hat{p}(p^*, z^*)$ namely, $p^* = k$ and $\pi_1^m \leq z^* \leq \pi_2^m$.*

Proof. The maximisation of the sum of the firms' profits implies $p = k$. Further, $d_2 \leq d_1$ implies that $\pi_2^m \geq \pi_1^m$ which implies that there exists a $z^* \in [\pi_1^m, \pi_2^m]$. The lump sum payment $z \in [\pi_1^m, \pi_2^m]$ is chosen such that signing the contract is individually rational.

Q.E.D.

Proposition 5 *Two-part-tariffs increase welfare, but do not implement the first best solution.*

Proof. The sum of profits is higher than with simple, i.e. linear, contracts because there is no double marginalisation. Consumers' surplus is higher because the market price is lower. But there is still market power which leads to the implementation of the price that would be charged by an efficient monopolist.

Q.E.D.

3.2 Regulation with general tariffs

In this setting link c is opened whenever this is efficient $(d_2 < d_1)$ and therefore productive efficiency obtains (as long as efficient entry occurs). Also, the efficient access charge is chosen in equilibrium: $p^* = k$. Therefore the thorny issue of access price regulation does not arise here. Nor does any necessity of behavioural regulation. The problem of market power, however, remains. On the final goods market monopoly prices are charged unless regulation successfully solves this problem. For this task downstream production costs will have to be known and therefore the informational needs of regulation are still far from negligible. But the essential point of this section was to show that two part tariffs may lead to efficient access provision and to efficient production in the vertical integration setting.

Arbitrage will not be a large problem as long as the number of firms for whom firm D_1 acts as a supplier is small which, for a number of industries, seems to be a reasonable assumption.

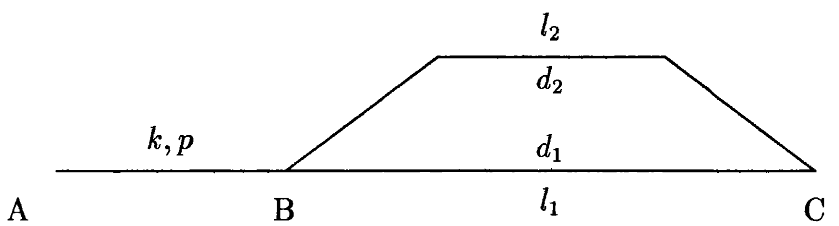
4 Ownership Structures

We will now focus on different allocations of ownership rights with a view to determining their effects on equilibrium prices for access and for final goods.

4.1 Access to complementary links

In order to be able to analyze the case of a vertically separated access provider we first define a slightly modified network structure. We modify our model by making link c a complementary unit of the network in the sense that access to this link is needed as a complementary input to the use of links a and b. We thus obtain, in the simplest case, a network structure like in the following diagram which is similar to an example in Baumol (1994).

Figure 3



If firm D_2 owns link l_2 and firm D_1 is integrated as before, i.e. owns the two other links, the analysis is very similar. We consider the same game structure and timing as in the preceding sections with the only modification that firm D_1 also takes a decision whether or not to open link c in stage $t = 0$. It is simple to consider such an entry decision in the original model as well: in the initial model firm D_1 would only enter, i.e. open link c, if this is worthwhile. There are cases when it is socially desirable but not individually rational to open this link. A comparable problem does not occur in the network structure modelled in this section since the link is necessary for $D_1's$ own production process. For $t = 2$ and for $t = 1$ we obtain exactly the same equilibria as in the model with substitutive links. But in $t = 0$ link $\overline{AB}$ is always opened in equilibrium which is not the case with link c $(= \overline{AB})$ in the other model.

The setting in which link $\overline{AB}$ is owned by a separate firm is one of the cases we take into account in what follows. We distinguish the following four possible allocations of ownership.

1. D_1 owns links b and c while D_2 owns link a in figure 1. D_1 gives access to link c. ('Vertically integrated incumbent with a downstream competitor').

2. An independent firm U owns link c and may grant D_2 access to it. D_1 owns link b and D_2 owns link a. ('Vertically integrated incumbent facing competing network operator and downstream competitor').

3. D_1 owns link b and D_2 owns links a and c ('Integrated duopoly').

4. U owns $\overline{AB}$ in the model of section 4.1 (figure 3) and gives access to firms D_1 and D_2 that own links l_1 and l_2 respectively ('Vertically separated upstream monopoly').

The first of these four cases has already been discussed in detail in section 2. We briefly analyze the other cases in order to demonstrate the influence of the allocation of ownership on pricing and entry decisions.

4.2 Vertical integration with upstream and downstream competition (case (2))

If $d_1 < d_2$, firm D_1 serves the market alone at $p_1 = \min(p + d_2, p_1^m)$. This holds for any access price $p \geq k$ charged in $t = 1$. Due to Bertrand competition in $t = 2$, U will, in equilibrium, offer access at $p = k$[6] which implies zero profits for U and for D_2. Unless D_2 is extremely inefficient the final goods price $p_2 = \min(k + d_2, p_1^m)$ is lower than $D_1's$ monopoly price because the foreclosure possibilities of case (i) in section 2.3 do not exist here.

If $d_1 \geq d_2$ prices $p_1^* = p_2^* = \min\Big\{\max\{p+d_2, k+d_1\}, p_1^m, p_2^m(p)\Big\}$ are charged in $t = 2$. Firm U, in $t = 2$, chooses, in equilibrium, p such as to annul productivity differentials, i.e. $p + d_2 = k + d_1$ or, equivalently, $p = k + (d_1 - d_2)$ unless D_1 is so inefficient that the resulting market price $p_2^*(p^*)$ would exceed p_1^m. Substitution yields final goods prices

$$p_1^* = p_2^* = \min(k + d_1, p_1^m, p_2^m).$$

U thus extracts D_2' s profit. Entry occurs in $t = 0$ if and only if $[d_2 \leq d_1$ and $S = 0]$.

Note that this outcome depends on the timing of the game. If the entry decision follows the determination of the access price the following obtains. The price equilibrium in $t = 2$ is not affected. Entry occurs in $t = 1$ if and only if entry sunk costs can be recovered, i.e.

$$k + d_1 \geq p + d_2,$$

i.e. D_2 can undercut by setting $p_2 = k + d_1$, and the equilibrium profit of firm D_2 is larger than her entry costs,

$$(\alpha - k - d_1)(k + d_1 - d_2 - p) - F_a \geq S. \tag{10}$$

In $t = 0$ firm U will choose p such that (10) is satisfied with equality unless this implies a loss:

$$p^* = k + \max\Big(0, d_1 - d_2 - \frac{S + F_a}{\alpha - k - d_1}\Big).$$

[6]One could allow for firm D_2 offering access to D_1 as well. But, as $p = k$ holds anyway this would not affect the equilibrium prices in $t = 2$ and we can therefore assume w.l.o.g. that D_2 buys access from U only.

The intuition for this result is that more entry occurs and lower access prices tend to be charged if network operators must commit to an access price before entry occurs or, equivalently, if network operators are not able to extract all of the downstream firm's profit subsequent to the entry decision. A policy recommendation would therefore be that the regulator requires access providers to specify access prices in advance, i.e. to freeze[7] access prices over a sufficiently long period of time.

4.3 Integrated Duopoly (case (3))

In this setting the prices charged in $t = 2$ are $p_1 = p_2 = k + \max\{d_2, d_1\}$. Here the issue of access pricing does not arise because D_2 owns c. We can therefore set $p = k$. Entry by D_2 occurs in $t = 0$ if and only if

$$d_2 \leq d_1 \text{ and } S \leq (d_1 - d_2) \cdot (\alpha - k - d_1) - F_a, \qquad (11)$$

where we neglect the case that firm D_1 is so inefficient that firm D_2, in equilibrium, can charge her monopoly price. Case (3) does not yield comparatively more entry than case (2) [with 'reversed timing'] because there firm U chose such an access price that entry occurred whenever efficient.

4.4 Vertical separation (case (4))

Firm U owns $\overline{AB}$ and firm D_i owns link l_i, $i = 1, 2$, in the model of section 4.1.

In this setting we distinguish two cases, namely whether or not the upstream firm is allowed to discriminate between the two downstream firms by charging different access prices. We first assume discrimination is allowed, i.e. U may charge access prices $r_1 \neq r_2$ to firms D_1 and D_2 respectively.

In $t = 2$ there exists a unique Nash equilibrium. It is given by

[7] It is, of course, only a price ceiling one has to fix over time. This would give firms some downward flexibility in pricing and enable them to respond to positive exogenous shocks such as new technologies etc. What is needed is a commitment that prices will not be *raised* subsequent to entry.

$$p_1 = \min\left\{\max\{r_1+d_1, r_2+d_2\}, p_1^m\right\} \quad (12)$$

$$p_2 = \min\left\{\max\{r_1+d_1, r_2+d_2\}, p_2^m\right\} \quad (13)$$

In $t=1$ consider the case $d_1 \leq d_2$ (w.l.o.g.). Firm U will set r_1, r_2 with a view to inducing large quantities of final goods sold which is achieved by equating

$$r_1 + d_1 = r_2 + d_2.$$

(Again we omit the trivial case that firm D_1 is very inefficient compared to firm D_2.) Let $d_1(p_1,p_2) = \alpha - p_1$ if $p_1 = p_2$ for existence. Then, as Bertrand competition implies $p_1 = p_2 = d_1 + r$, the upstream firm's profit is given by

$$\pi_U = (\alpha - d_1 - r_1)(r_1 - k)$$

which is maximal for

$$r_1 = \frac{\alpha + k - d_1}{2},$$

i.e. the efficient firm serves the market. Substituting this into (12) and (13) yields prices

$$p_1 = p_2 = \frac{\alpha + k + d_1}{2}.$$

If entered in t=0, D_2 would serve the market only if $d_2 < d_1$. But, due to Bertrand competition in t=2, the margin per unit is zero: under $d_2 \leq d_1$ we obtain $r_2 = \frac{\alpha+k-d_2}{2}$ and prices $p_1 = p_2 = \frac{\alpha+k+d_2}{2}$. D_2's profit is therefore given by

$$\pi_2 = \left(\alpha - \frac{\alpha+k+d_2}{2}\right) \cdot \left(\frac{\alpha+k+d_2}{2} - \frac{\alpha+k-d_2}{2} - d_2\right) - F_2 = -F_2.$$

Therefore entry occurs in t=0 if and only if $S=0$ and $F_2=0$.

We now analyze the non-discrimination case $r_1 = r_2 = r$. In $t=2$ prices charged in equilibrium are given by

$$p_1 = p_2 = \min\left\{r + \max\{d_1, d_2\}, p_1^m, p_2^m\right\}.$$

Let $d_1 \leq d_2$. Then, in $t = 1$, the access price is chosen such as to maximise $\pi_U = (r - k)\ (\alpha - r - d_2)$ which implies

$$r^* = \frac{\alpha + k - d_2}{2}.$$

Substitution yields final goods prices

$$p_1 = r^* + d_2 = \frac{\alpha + k + d_2}{2} > \frac{\alpha + k + d_1}{2}.$$

D_2 enters the market if and only if sunk costs as well as fixed costs are zero. These findings lead to

Corollary 3 *Price discrimination undertaken by an upstream monopolist leads to lower access prices, lower downstream profits and higher upstream profits. The prices charged for the final good correspond to the price of an efficient monopolist if price discrimination is allowed and to the price of an inefficient monopolist otherwise.*

The intuition for this result is that price discrimination enables the monopolist to induce tough competition. By using the access charges in order to offset downstream productivity differentials the upstream firm induces lower prices in order to raise output levels and thereby the total amount of access charges that will be collected.

4.5 Comparison between different ownership structures: implications for equilibrium prices and for entry

Summarising results from sections 2 and 4 the following can be stated (using cases 1 to 4 as defined in section 4.1).

Vertical integration with a downstream competitor generally leads to monopoly pricing. Only if the competing firm has a sufficiently large productivity advantage it will serve the market alone, but the final price charged in this case, at least for α sufficiently large, exceeds the monopoly price p_1^m.

If a second network operating firm is present, either separate (case 2) or integrated with D_2 (case 3), final prices are given by the maximum of the marginal costs, i.e. $p_1 = p_1 = k + \max(d_1, d_2)$.

Vertical separation (case 4), on the other hand, leads to prices an efficient monopolist would charge. This is true as long as price discrimination is allowed. Otherwise the monopoly price of the inefficient firm obtains.

Entry decisions are affected by ownership in the following way. With vertical integration (case 1) a downstream competitor enters only if sufficiently more efficient than the incumbent and not foreclosed (i.e. not in cases (i) and (iv) and in cases (ii) and (iii) only if conditions (8) and (9) hold respectively.

In case a second (but non-integrated) supplier of the essential input exists (case 2) entry occurs if $d_2 \leq d_1$ and $S = 0$ which means that upstream entry encourages downstream entry. If the access price is fixed before the entry decision is taken then entry obtains whenever (10) holds, i.e. entry is no longer ruled out by positive sunk or fixed costs. If the second supplier is integrated with a downstream firm (case 3) entry obtains under (11). This implies that a vertically integrated competitor enters when his productivity advantage is sufficiently large to cover the sum of sunk and fixed costs. Under the original timing, i.e. if entry precedes access pricing, case (3) therefore leads to more (socially desirable) entry than case (2) which may be taken as an argument for allowing vertical integration of competitors in the presence of a vertically integrated incumbent.

Under vertical separation (case 4), again, entry occurs if $d_2 \leq d_1$ and if sunk and fixed costs are zero.

4.6 Regulatory measures

From the preceding subsection it is clear that access price regulation is crucial whenever there is a monopoly in the bottleneck, i.e. in cases 1 and 4. Setting $p = k$ leads to Bertrand competition and therefore prices $p_1 = p_1 = k + \max(d_1, d_2)$ whenever both downstream firms are active in the market. Entry occurs and leads to productive efficiency as long as the productivity advantage is sufficiently large compared to the entry sunk costs which can therefore be recovered. In cases

(2) and (3), on the other hand, the issue of access price regulation does not arise. Whenever industry features allow for duplication of the essential facility (or bottleneck) this avenue should be pursued unless the associated sunk costs are prohibitively high.

5 Conclusions

In much of the literature on regulated industries ownership and regulation are treated in isolation from each other. In this paper we have started by considering a setting without regulation. We analyzed the behaviour of firms in an industry with features that are typical for industries that tend to be regulated: an essential input is required by all downstream firms and provided by one or more integrated or separate upstream firms.

We have shown how equilibrium pricing depends on productivity differentials and on ownership. Productive efficiency by no means implies efficient outcomes. Production by a more efficient firm may even lead to higher prices as long as a monopolistic supplier controls the bottleneck. Duplication of the bottleneck changes the situation dramatically even if the second supplier merges with a downstream firm. Regulation of industries with a monopolistic upstream segment should therefore focus on the provision of efficient access to the bottleneck. Further, it was shown how a long term commitment of a monopolistic supplier to fix access charges over a longer period leads to lower access prices and more entry.

Vertical separation may be advantageous because it leads to productive efficiency while an efficient competitor may well be foreclosed by a vertically integrated incumbent. If the incumbent is vertically integrated then vertical integration of competing firms should not be banned. This policy may encourage efficient entry in the downstream industry which would not occur otherwise.

Starting from these results we have pointed out how the burden of regulation is affected by the industry characteristics described above (ownership, tariff structure, timing, and relative efficiency of firms). We identify the kind of regulation (access price, final goods price etc.) required in different scenarios and indicate implications for situations with asymmetric information.

A major problem with regulation along these lines, of course, will be the informational advantage of the network operating firm. It has been argued by Hardt and Stürmer (1995) that the incentives to exploit these informational asymmetries vary between ownership structures as well.

Further analysis along these lines is needed. In our view a first step is to find the best ownership structures for each specific industry. Network duplication, for instance, may cause prohibitively high sunk costs in some industries (rail industry etc.) but may be desirable in other industries (telecommunication services) due to new technologies. One can then analyze which are the best tools to solve problems of informational asymmetries starting from a structure that minimises the overall regulatory burden. In other words, we regard regulation with complete information as one borderline case and the free market analysis as the other borderline case. Regulation with asymmetric information may then be interpreted as intermediate case. An ideal ownership structure would then be one under which both borderline cases coincide. Such a case was found in chapter III.

6 References

Armstrong, Mark, Simon Cowan, and John Vickers (1994) : *Regulatory Reform: Economic Analysis and British Experience;* MIT Press.

Baumol, William J. and J. Gregory Sidack (1994) : *Toward competition in local telephony;* MIT Press.

Byg, Torkild and Michael Hardt (1995) : Access Pricing, Ownership and Contracts; *Mimeo. University of Bonn and London School of Economics.*

Economides, Nicholas S. and Glenn A. Woroch (1992) : Benefits and Pitfalls of Network Interconnection; *Discussion paper EC-92-31, Stern School of Business, New York University.*

Hardt, Michael and Susanne Stürmer (1995) : Network Access and Ownership Structure; *Mimeo. London School of Economics and Freie Universität Berlin.*

Perry, Martin K. (1989) : Vertical Integration: Determinants and Effects; *in Schmalensee, R. and R.D.Willig (Eds.), Handbook of Industrial Organization, Amsterdam: 183-255.*

Waterson, Michael (1993) : Vertical Integration and Vertical Restraints; *Oxford review of economic policy, Vol 9: 41-57.*

CHAPTER V

Consumer Behaviour and Competitive Equilibrium [1]

Abstract

We consider a general equilibrium model which is modified in so far as consumption behaviour is characterized by probability distributions instead of preferences. In this setting markets generally do not clear. We define the concept of an expected equilibrium and examine the properties of expected market demand. These results are used to prove the existence of expected equilibria. Uniqueness of expected equilibria obtains under rather restrictive assumptions on consumption distributions. Quantity rationing is introduced as a way in which excess demand or supply can be dealt with. For the case of proportional rationing schemes it is shown that rationing becomes insignificant (arbitrarily small) when the economy becomes large. We finally provide a result on the size of disequilibrium in a finite economy. A sharp upper bound is given for the probability that excess market demand for some good exceeds a given bound ϵ.

1 Introduction

> 'It is [...] necessary to show how assumptions of bounded rationality could replace assumptions of optimisation in actual economic reasoning.'
>
> Herbert A. Simon (1992)

[1]This chapter is based on Hardt (1995).

1.1 General equilibrium and bounded rationality

In the literature on general equilibrium theory as well as in many other areas of economics individual choice behaviour has commonly been described with the help of preferences or utility functions. Although valuable insights can be derived with the help of utility maximisation models the scope of their application and their inherent realism are strongly limited. In recent years more and more attention has been paid to bounded rationality (see Simon (1982) and Simon (1992)). It has been argued that individuals do not at all follow a calculus of maximisation when making their choices. Abundant empirical evidence for this claim has been obtained from various laboratory experiments. For an overview in experimental economics see Tietz, Albers and Selten (1988).

But, on the other hand, economists carry on building models based on assumptions of utility maximisation and perfectly rational behaviour. The reason for this phenomenon seems to be a conjecture that without utility maximisation all the commonly used models break down and hardly anything can be said about the functioning of a market economy. This chapter attempts to abandon this commonly chosen avenue of complete rationality.

It is hard to define in general terms what is meant by complete rationality but it seems even more difficult to give a formal definition of bounded rationality. One could even argue that it is characteristic for the notion of bounded rationality that it cannot be fully captured by any formal definition. So we will not pretend to have found a general definition that does. Various approaches for modelling weaker rationality have been suggested in the literature. Magill and Quinzii (1992), for instance, argue that in a general equilibrium setting incompleteness of the asset market structure may arise from a lack of rationality. Geanakoplos (1992) develops the impacts of different rationality levels on speculation, betting, and 'agreeing to disagree'. Among other approaches there is the concept of qualitative reasoning which would suggest that agents take qualitative decisions only, revising variables up or down, instead of attempting to find the parameter values that would lead to a maximisation of the outcome.

Here, we will choose randomness of demand as one possibility to replace utility maximisation as a description of human behaviour. Special cases of random demand have been suggested by Becker (1962) and by Mirrlees (1986) to formalise irrationality or bounded rationality respectively. A motivation for our approach is given in the following two subsections. We do not inquire which of the above approaches is fitting best for which purposes because the question we want to ask in this chapter is rather the following one: how essential is the assumption of utility maximisation really or, more precisely, how sensitive is general equilibrium analysis to this assumption? Thus the main part of this work is a re-examination of classical issues and questions in an alternative framework. It will in fact turn out that some of the predictions from standard general equilibrium theory can be obtained without assuming that individuals maximise utilities. Other conclusions differ from the standard results which, however, does not imply that they are less plausible.

1.2 A motivation for randomness of demand

It appears appropriate to provide some arguments why stochastic demand appears to be a suitable alternative way of modelling consumption behaviour.

1. It allows for a simple model capturing a wide range of rationality levels, starting from irrationality on the one hand to full rationality on the other hand: even though we do not make the assumption that consumption is in any way a result of utility maximisation such an interpretation is possible. One could start, for instance, from stochastic endowments or from stochastic preferences: after a certain state of nature has been realised preferences (or endowments, respectively) are fixed and agents maximise their utility subject to these realisations (see Block and Marshak (1960)

and Hildenbrand (1971)). In such a model agents are, of course, somewhat 'hyperrational'. They are not only able to maximise utility, but they do so any time a new state of nature occurs. On the other hand, in our model agents are not necessarily rational at all. They could, for instance, be characterised by a uniform distribution associating the same probability (i.e. density) with each consumption bundle satisfying the budget constraint. This case is usually meant to describe 'irrational agents' (see Becker (1962)). [2]

2. The random demand model we will use contains the standard general equilibrium model as a special case: under standard assumptions on probability distributions consumption behaviour can be derived from maximisation of stochastic preferences. Thus a one point distribution on individual preferences corresponds to the usual model.

3. Some of the results obtained from this model seem less counterintuitive than classic predictions like, for instance, boundary behaviour.

4. Stochastic consumption allows for an analysis of stochastic interaction. Choice behaviour of an economic agent is obviously influenced by other agents' choices. Random demand provides a possibility of analyzing such interaction. Research along these lines has been conducted for example by Föllmer (1974). Although we do not tackle this problem explicitly in this chapter we would like to emphasise that the model we present establishes a rather general framework in which stochastic interaction can be dealt with. In fact, we do not make any assumption of independence except for two subsections where the issues of uniqueness and large economies, respectively, are addressed (propositions 7,

[2] For the special case of deterministic utility but stochastic decision rules (see section 1.3) it has already been argued by McFadden (1981) and by Mirrlees (1986) that random demand well captures the idea of bounded rationality.

8 and 9).

5. It is sometimes argued that, due to the Law of Large Numbers, randomness of demand simply 'cancels out'. This is true in the limit, i.e. when the number of agents tends to infinity, but we do not focus primarily [3] on limit economies. Instead, we are interested in analyzing markets we observe: they are characterised by a finite number of agents and by fluctuations in demand rather than by a continuum of agents or by 'zero variance'. [4]

6. Random demand is used in a wide variety of models in the literature. Thus a treatment of some of these approaches in a unified framework might be desirable. A short overview of this area is given in the following subsection.

1.3 Random demand in economic theory

There [5] is in fact a huge literature in different fields varying from Mathematical Psychology and Theory of Choice to applications in Industrial Organisation, Learning, and many other areas of economics where consumption is assumed to be stochastic. Although the resulting randomness is traced back to rather different causes all these models have in common that consumption is stochastic. Thus they can be treated as special cases of, or even as motivation for, the assumptions of our model. We will now quote some of these approaches in order to make precise what has been said above and in order to provide some references.

The thought that human behaviour in a given situation may differ from one time to another and thus exhibit inconsistencies like intran-

[3] We do analyze a limit economy in proposition 8. But we do so with the intent to examine the effect of economies becoming larger. We then analyze the finite case and show that the probability that the amount of rationing exceeds a given limit becomes small for large economies. This is made precise in proposition 9.

[4] We still assume price-taking behaviour of the agents: this assumption is more appropriate to describe boundedly rational agents than an assumption of strategic behaviour would be.

[5] The following outline owes a lot to the first chapters of Anderson, de Palma and Thisse (1992) where a detailed survey of the literature on discrete choice is given.

sitivities, for instance, is rather old. The history of random utility models can be traced back to Thurstone (1945). Psychological experiments led to the statement that the utility of each alternative is in fact measured by a random variable. This approach is equivalent to assigning random preferences to the set of alternatives as was shown by Block and Marshak (1960). Thus the central idea of this approach is the maximisation of a 'momentary utility'. The kind of utility relevant at a certain point of time is determined by unknown parameters, moods, or 'psychological states' (Thurstone). In more economic terminology this would correspond to agents maximising state dependent preferences. Machina (1985) shows that the random utility model is a special case of his model in which agents are supposed to maximise deterministic utilities on lotteries. The choice probabilities are then given by the optimal lottery (i.e. they are determined endogenously in contrast to the theory of choice under uncertainty).

Another approach has been followed by Luce (1959) and Tverski (1972a,b), namely the approach of deterministic utility and stochastic decision rules (see Tverski (1972a) p.281): an individual's choice does not have to coincide with the alternative that would yield a maximum utility (see our remarks on bounded rationality in section 1.1).

This concept of random preferences or random choice behaviour has been applied in various fields of economics: there are, for instance, learning models using random demand. One example from the learning literature is Bray and Savin (1986). Many models in Industrial Organisation assume stochastic market demand in order to analyze firms' behaviour under uncertain demand conditions. Among a large number of models [6] dealing with oligopolistic or monopolistic markets with uncertain demand one could quote Novshek-Sonnenschein (1982). The book by Anderson et al. quoted above contains many applications of random demand in the theory of oligopoly in differentiated product markets. In most of the models of these areas the motivation for de-

[6] Demand uncertainty is not necessarily a complicating factor introduced to deterministic models. An evident example is Klemperer and Meyer (1989) where demand uncertainty is shown to reduce the multiplicity of Nash equilibria.

mand uncertainty is that suppliers lack information about the demand they face. This corresponds to the interpretation of stochastic human behaviour by Manski (1977 p.229) and McFadden (1974). The alternative interpretation, namely intrinsically probabilistic behaviour, was originally adopted by Quandt (1956) and later followed by psychologists.

For our equilibrium analysis based on the notion of random demand, however, we do not need to take a clear position towards these different interpretations of randomness, nor do we confine ourselves to any specific class of models mentioned above. We rather regard them as approaches which differ from each other in many aspects but all have the common feature of random demand thus providing motivations for replacing the utility maximisation paradigm of general equilibrium analysis by stochastic demand.

The rest of this chapter is organised as follows: in the following section we present our random consumption model and show that, with probability one, markets do not clear. In section 3.1 the concept of an expected equilibrium is defined and properties of expected market demand are examined (propositions 1 and 2). These results are used to prove the existence of expected equilibria (propositions 3 and 4). The problem of uniqueness is examined in Section 3.2. It is pointed out there that monotonicity of expected partial market demand is implied by a property of stochastic dominance of individual demand (proposition 7) and that monotonicity of expected demand holds for price independent distributions. Section 4 introduces quantity rationing as a way in which excess demand or supply can be dealt with. In subsection 4.1 we present the model of quantity rationing and define the notion of a rationing - allocation. It is shown in subsection 4.2 that rationing becomes the more insignificant the larger the economies get. Finally, we provide a result in subsection 4.3 about the size of disequilibrium in a finite economy (proposition 9). Conclusions follow in section 5.

2 Model: An Exchange Economy with Stochastic Consumption

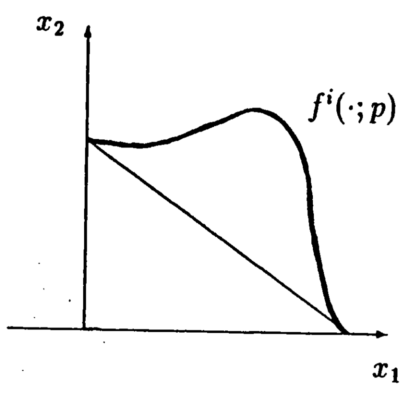
An exchange economy is usually defined by a commodity space and by the agents' characteristics, i.e. their endowment vectors and their preferences or utility functions:

$$\mathcal{E} = [I\!R^L_+, (U^i, e^i)_{i=1,\dots,I}].$$

We adopt this model as far as the commodity space and the endowments are concerned but we do not consider explicitly preferences that might induce the agents' consumption choices. Instead, we start with the choices themselves: an agent i chooses stochastically a consumption bundle X^i, $i = 1, \dots, I$, thus $X^i : \Omega \to I\!R^L$ is a random vector. We assume that agents spend their total income. Thus, as illustrated in the figure below, an agent's consumption behaviour is characterised by a probability distribution on his budget hyperplane and a stochastic exchange economy is given by

$$\mathcal{E} = [I\!R^L_+, (X^i, e^i)_{i=1,\dots,I}].$$

Figure 1



The budget sets are $I\!B(p, e^i) = \{x \in I\!R^L_+ : p \cdot (x - e^i) \leq 0\}$ for $p \in S^{L-1}_{++} := \{p \in I\!R^L_{++} : \sum_{l=1}^L p_l = 1\}$. We denote by S^{L-1}_+ the analogous set where prices are not required to be strictly positive: $p \in I\!R^L_+$.

Let $\tilde{\mathbb{B}}(p, e^i) = \{x \in \mathbb{R}^L_+ : p \cdot (x - e^i) = 0\}$ denote the respective budget hyperplane. As we assume the budget identity to hold the probability distributions have dimension $L - 1$. The shape of an agent's consumption distribution depends, of course, on the relevant price vector $p \in S^{L-1}_{++}$. A density function describing agent i's consumption is denoted by $f^i(x_1, \cdots, x_L; p)$ or, in short, $f^i(x, p)$. It is non-negative on the budget hyperplane and satisfies

$$\int_{\tilde{B}(p,e^i)} f^i(x, p)dx = 1.$$

For $p \notin S^{L-1}_{++}$ we define $f^i(x; p) := f^i(x, p/\sum_{l=1}^L p_l)$. Marginal densities for good l are then given by

$$f^i_l(x_l; p) = \int_{\{y \in \tilde{B}(p,e^i) | \ y_l = x_l\}} f^i(y; p)dy$$

and marginal distribution function are defined as

$$F^i_l(x_l; p) = \int_0^{x_l} f^i_l(x; p)dx.$$

really?

We omit the index for the good and simply write $F^i(x_l; p)$ and $f^i(x_l; p)$ respectively when no ambiguities are involved.

In this framework markets generally do not clear as the following lemma states.

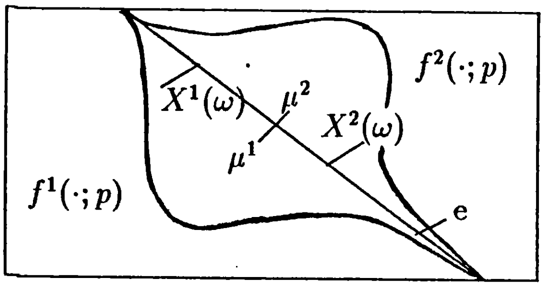
Lemma 1 *Let $\mathcal{E}$ be a stochastic exchange economy. Let agent i's random consumption bundle X^i have a continuous density $f^i(x; p)$ on the budget hyperplane. Then for any price vector $p \in S^{L-1}_{++}$ markets do not clear with probability one, i.e.:*

$$\sum_{i=1}^I X^i(\omega; p) \neq \sum_{i=1}^I e^i \qquad \text{w.p.1.}$$

This result obtains because $\sum_{i=1}^I X^i = \sum_{i=1}^I e^i$ describes a hyperplane of Lebesgue measure zero. In the Edgeworth box below indifference curves have been substituted by demand densities. Realisations of demand are unlikely to coincide with supply (even if the respective expected excess demands add up to zero).

Figure 2

Edgeworth box with demand densities



This result, although not quite in line with the spirit of general equilibrium theory, may well be more descriptive than the prediction of market clearing. A less demanding notion of equilibrium is introduced in the following section.

Remark 1 *Lemma 1 states that one cannot implement ex ante a price that will lead to market clearing. It is simple to construct random demand models that yield market clearing in the following way. A realisation of a state of the world fixes preferences and therefore a demand function for each individual. Then the usual tâtonnement process leads to a market clearing price. The difference with our model is that the notion of a demand function simply does not exist. Thus there is no way to predict how individual consumption changes with prices and thus there is no way to define a tâtonnement process.*

3 Expected Equilibria

3.1 Existence

The equilibrium concept we define in this section formalises the idea that markets do not clear for every state $\omega \in \Omega$, but they may clear on the average, i.e. in expected terms. One can imagine the Walrasian auctioneer [7] modified in the following way: he does not know which state of the world will be realised, but he does know the probability distribution of agent i's consumption $X^i(\cdot;p)$ (instead of the preferences $\succeq^i$, $i = 1,\ldots,I$). Thus he is able to calculate expected excess demand and aims at choosing a price p such that, for each good, expected demand equals supply.

Definition 1 *An expected equilibrium $(p^*, (E[X^1(p^*)],\ldots,E[X^I(p^*)])) \in \mathbb{R}^L_+ \times \mathbb{R}^{IL}_+$ of a stochastic exchange economy is defined by a price vector $p^* \in S^{L-1}_+$ and an expected allocation $(E[X^1(p^*)],\ldots,E[X^I(p^*)])$ satisfying*

$$\sum_{i=1}^{I} E[X^i(p)] = \sum_{i=1}^{I} e^i.$$

Remark 2 *An optimality property of p^* will become clear in section 4 (see propositions 8 and 9).*

In order to address the issue of existence of an expected equilibrium we will now examine the structure of expected market demand. The following proposition shows that, under regularity assumptions, the properties of homogeneity, Walras Law, and continuity can be established for expected market demand.

Proposition 1 *Let $\mathcal{E} = [\mathbb{R}^L_+, (X^i, e^i)_{i=1,\ldots,I}]$ be a stochastic exchange economy. Let individual demand $X^i(\cdot;p)$, $i = 1,\ldots I$, depend continuously on prices p, i.e. $f^i(x;p)$ is continuous in p $\forall x$. Then individual demand satisfies the following conditions:* Meaning?

1. *Homogeneity:* $E[X^i(\lambda \cdot p)] = E[X^i(p)]\ \forall\ \lambda \in \mathbb{R}_+$
2. *Walras Law:* $p \cdot (\sum_{i=1}^{I} E[X^i(p)] - \sum_{i=1}^{I} e^i) = 0\ \forall\ p \in S^{L-1}_{++}$
3. *Continuity:* $E[X^i(p)]$ *is continuous in p.*

[7] We are well aware that the old story about the auctioneer is not very satisfactory. This problem has not been resolved in general equilibrium analysis and we shall not attempt to solve it here. But the modification seems obvious: in the case of uncertainty the auctioneer aims at market clearing in the sense that expected demand equals expected supply.

Proof.

1. For prices $p \notin S_+^{L-1}$ the density functions are given by $f^i(x,p) = f^i(x, \frac{p}{\sum_{l=1}^L p_l})$. Thus for $\lambda \in I\!R_+ : f^i(x, \lambda p) = f^i(x, \frac{\lambda p}{\sum_{l=1}^L \lambda p_l}) = f^i(x,p)$.

2. As the distributions are defined on the budget hyperplane the budget restrictions hold (with equality) for every possible realisation $\omega \in \Omega$. Thus they hold for expected demand so that Walras Law is satisfied in expected terms.

$$\begin{aligned} p \cdot X^i(p)(\omega) &= p \cdot e^i \; \forall \, \omega \in \Omega \\ \Longrightarrow p \cdot (\sum_{i=1}^I E[X^i(p)] - \sum_{i=1}^I e^i) &= 0. \end{aligned}$$

3. Due to integration by parts we obtain:

$$E[X_l^i(p)] = \int_0^\infty (1 - F^i(x_l;p))dx_l - \int_{-\infty}^0 F^i(x_l;p)dx_l.$$

Distribution functions depend continuously on p because densities are continuous. Therefore, the integrands of both integrals are continuous. Thus, due to Lebesgue's theorem, expected demand is continuous in p.

Q.E.D.

Thus some of the main properties of demand under utility maximisation can be established in our framework for expected demand. But this is not the case for boundary behaviour. The problem with boundary behaviour is that the usual analysis relies essentially on maximisation of unsatiated preferences. But as we do not assume any form of maximisation here it is not possible (without making further assumptions) to derive an 'explosion' of demand in case one price tends to zero. This can be illustrated by the following

Example 1

Let an agent's consumption of good k be characterised in the following way. Consider first a density $f(\cdot;p)$ on $[0;\infty]$ which has a finite mean μ. Now let the agent's consumption of good k be described by the truncated normalised density defined as follows:

$$\hat{f}(x_k;p_k) = \frac{f(x;p_k)}{1-c(p_k)} \text{ for } 0 \leq x_k \leq \frac{b}{p_k} \text{ and } 0 \text{ otherwise,}$$

where income is denoted by b [8] and

$$c(p_k) := \int_{\frac{b}{p_k}}^{\infty} f(x;p_k)dx_k$$

is the appropriate constant to achieve normalisation. Then the expected value of the truncated distribution converges to the expected value of the original (non-truncated) distribution and is therefore finite:

$$\begin{aligned} E_{\hat{f}}[X] &= \frac{E_f[X]}{1-c(p_k)} - \frac{1}{1-c(p_k)}\int_{\frac{b}{p_k}}^{\infty} xf(x;p_k)dx_k \\ &\to E_f[X] \qquad \text{for } p_k \to 0 \\ &= \mu < \infty. \end{aligned}$$

Thus we have constructed a simple example where the assumptions made as to the distribution of the agents' consumption do not seem to be too counter-intuitive but a lack of boundary behaviour can be derived. This result is in line with the observation that only Walras Law and continuity can be empirically verified. In what follows we examine classes of distributions for which various forms of boundary behaviour can be derived. meaning?

Definition 2 *A distribution is called symmetric if the marginal densities $f^i(x_l;p)$ are symmetric around $\frac{b}{2p_l}$, i.e.* ?

$$f^i(\frac{b}{2p_l} + x_l, p) = f^i(\frac{b}{2p_l} - x_l;p), \qquad 0 \leq x_l \leq \frac{b}{2p_l}.$$

like Cobb Douglas

[8] Notation is a bit sloppy here because we treated b as a constant. But the argument does not change when one writes income explicitly as $b = p \cdot e$ because then $\frac{b(p)}{p_k} = e_k + \sum_{l \neq k}^{L} \frac{p_l}{p_k} \cdot e_l$ where p_l is non-decreasing for $p_k \to 0$ such that $\frac{b(p)}{p_k} \to \infty$ for $p_k \to 0$.

Definition 3 *A probability distribution has a price independent shape if the following condition holds for the marginal distributions $F^i(\cdot;p_l)$:*

$$F^i(\frac{y}{p_l};p_l) = F^i(\frac{y}{p_l'};p_l') \;\forall\; y \in I\!R \;\forall\; p_l, p_l' \text{ such that } p, p' \in S^L_{++}.$$

The geometric intuition for this definition is that the shape of the marginal distribution is invariant under price changes. This does not mean that the distributions $f(\cdot;p_l)$ are all of the same type (e.g. exponential) but rather that the transformation of the distribution induced by a price change is either an expansion or a contraction. To see this let $y_l = p_l x_l$, denote the expenditure level for good l, then the above condition becomes

$$\tilde{F}^i(x_l;p_l) = \tilde{F}^i(x_l \cdot \frac{p_l}{p_l'};p_l'),$$

where the ratio $\frac{p_l}{p_l'}$ is the expansion or contraction factor.[9]

The following proposition shows that various kinds of boundary behaviour are implied under various distributional assumptions and shows how they are related to each other in the sense of logical implications.

Proposition 2 *Let $\mathcal{E} = [I\!R^L_+, (X^i, e^i)_{i=1,\ldots,I}]$ be a stochastic exchange economy where the individual consumption distribution have (continuous) densities. Then for conditions 1 to 6*

1. $\exists\; \epsilon > 0 :\; E[X_l^i(p)] - e_l > 0$ for $p_l < \epsilon, l = 1, \ldots, L$.
2. $E[X_l^i(p)] \to \infty$ for $p_l \to 0$.
3. The marginal density $f^i(x_l;p)$, $1 \leq l \leq L$, is symmetric.
4. The distribution of $X^i(p)$ has a price-independent shape and $f^i(x,p)$ > 0 on the budget-hyperplane.
5. $P(X_l^i(p) < e_l) \to 0$ for $p_l \to 0$.

[9]We consider price changes in component l : $p_l' > p_l$. As prices are normalised the other components of p are determined by the change in component l. Therefore we can write $F^i(\cdot;p_l)$ instead of $F^i(\cdot;p)$.

6. Agent i chooses $X^i := \arg\max\ u^i(X^i) + \tilde{\epsilon}$ where $u(x)$ is a monotonic utility function and $\tilde{\epsilon}$ is an error variable with expected value zero.

the following relations hold:

$$\begin{array}{ccccc} & & (4) & & (5) \\ & & \Downarrow & & \Downarrow \\ (3) & \Longrightarrow & (2) & \Longrightarrow & (1) \\ & & \Uparrow & & \\ & & (6) & & \end{array}$$

Proof. The implications '(2) $\Longrightarrow$ (1)' and '(6) $\Longrightarrow$ (2)' are immediately evident. For the rest just note:

'(3) $\Longrightarrow$ (2)': symmetry of the marginal density $f_l^i(x;p)$ on $[0; \frac{b}{p_l}]$ implies that $E[X_l^i] = \frac{b}{2p_l}$. Thus we obtain $E[X_l^i(p)] \to \infty$ for $p_l \to 0$.

'(4) $\Longrightarrow$ (2)': using a simple substitution $y := p_l x_l$ one obtains

$$\begin{aligned} E[X_l^i] &= \int_0^{\frac{b}{p_l}} (1 - F(x_l; p_l))dx_l \\ &= \int_0^b (1 - F(\frac{y}{p_l}; p_l))dy \cdot \frac{1}{p_l} \end{aligned}$$

But, due to the assumption of price independence, $F(\frac{y}{p_l}; p_l)$ is constant in p_l. Further, $F(\cdot) = 1$ cannot hold on the entire interval $[0, \frac{b}{P_l}]$. So the integral is bounded away from zero: $\exists\ \epsilon > 0$ such that $\int_0^b (1 - F(\frac{y}{p_l}; p_l))dy > \epsilon\ \forall\ p_l$ which implies that $E[X_l^i(p)] \to \infty$ for $p_l \to 0$.

'(5) $\Longrightarrow$ (1)': $P(X_l^i = e_l) = 0$ because densities are continuous. So, if $P(X_l^i < e_l) \to 0$ for $p_l \to 0$, then $P(X_l^i > e_l) \to 1$ for $p_l \to 0$. But this implies that $E[X_l^i(p)] > e_l$ for $p_l \to 0$.

Q.E.D.

Thus individual consumption satisfies some boundary behaviour for special classes of distributions, for example distributions satisfying either (3), (4), or (6). But for general distributions one does not know whether any boundary behaviour is satisfied. It is now straight forward to state the following two propositions.

Proposition 3 *Given that the assumptions of Proposition 1 are satisfied for all agents $i = 1, \ldots, I$, market demand satisfies the properties of homogeneity, Walras' Law, and continuity.*

Boundary behaviour, i.e. any condition of (1) - (6) of proposition 2, is satisfied for market demand, whenever the respective condition is satisfied for at least one consumer $i \in \{1, \ldots I\}$.

Proof. Given propositions 1 and 2 the proof of proposition 3 follows immediately from proposition 3.2 in Hildenbrand and Kirman (1988).

Q.E.D.

Proposition 4 *Given that the assumptions of proposition 1 are satisfied, an expected equilibrium $(p^*, E[X(p^*)])$ exists. If in addition to the above properties boundary behaviour is satisfied for at least one agent i then there exists an expected equilibrium $(p^*, E[X(p^*)])$ such that $p^* >> 0$.*

Proof. Given propositions 1 and 2 the proof of the first part of proposition 4 corresponds to the proof of proposition 3.4 in Hildenbrand and Kirman (1988), the second part of our proposition is proved by Theorem 3.1 of the same book.

Q.E.D.

3.2 Uniqueness

3.2.1 A rationality requirement (individual level)

In the analysis of this section we impose an assumption on the agents' consumption behaviour, namely that of stochastic dominance. In vague terms this means that with a rising price of a good l the probability for low consumption levels of this good rises.

Definition 4 *Let F, G be distribution functions. F is said to dominate G stochastically, $F \succ G$, at first order if $F(x) < G(x)\ \forall\ x \in \{x \in$*

$\mathbb{R}|F(x), G(x) \in (0,1)\}$ *(first order stochastic dominance). We speak of weak first order stochastic dominance if the inequality holds weakly.*

Assumption 1 *Let $p, p' \in S^{N-1}_{++}$ such that $p'_k > p_k$ and the relative prices of the goods $l \neq k$ are unchanged: $\frac{p_l}{p_m} = \frac{p'_l}{p'_m}$ $\forall\, l, m \neq k$. Then the marginal distribution function $F^i(\cdot, p)$ satisfies the following condition:*

$$F^i(x_k; p) \leq F^i(x_k; p').$$

Remark 3

1. *Assumption 1 still is a weak rationality requirement: we do not exclude that, despite of risen prices, some individuals consume a higher quantity of this good, only the probability for higher consumption levels is expected to decrease.* [10]

2. *For the assumption of stochastic dominance empirical tests are available for which it is not necessary to make assumptions as to the type of distributions considered (nonparametric methods; for instance the test of the sum of ranks by Wilcoxon and Mann-Whitney).*

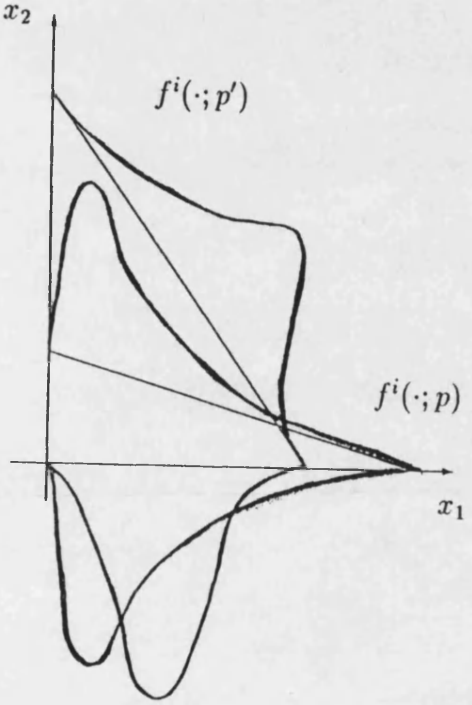
Definition 5 *Let $X^i(p)$ be the random vector of agent i's consumption when price p is relevant. Let $F^i_l(\cdot; p)$ be the corresponding marginal distribution function for good l. $X^i(p) \sim F^i(\cdot, p)$. Then individual demand of agent i is called stochastically monotonic if assumption 1 holds for $F^i_l(\cdot; p)$ $\forall\, l = 1, \ldots, L$.*

One might wonder whether the property of stochastic monotonicity is implied by the assumption that individuals spend all income, i.e.

[10] An alternative interpretation of this assumption is obtained for (deterministic) distribution economies where I subeconomies $E_1, ..., E_I$ are considered. The densities f^i are in this case interpreted as frequencies of consumption levels in the subeconomy E_i. The requirement of stochastic dominance then means that the frequency for lower demand levels of good l rises with price p_l, i.e. there are more individuals who consume less of good l. So for any subeconomy we allow for individuals who behave arbitrarily irrational, it is only the demand *distribution* of the subeconomy which we assume to shift to the left with rising price.

choose from the budget line. But the counter-example in the following figure shows that this is not the case.

Figure 3



On the other hand, stochastic monotonicity as defined above turns out to be satisfied for a large class of distributions. In fact, when the distributions are price independent then stochastic monotonicity can be derived.

Proposition 5 *Let $X^i(p)$ denote agent i's consumption given price p. Let the corresponding distribution function be price independent. Then individual demand $X^i_k(p)$ of agent i for good k is stochastically monotonic in its price p_k.*

Proof. Consider prices $p,\ p' \in S^L_{++}$ such that $p'_l > p_l$ and $\frac{p_m}{p_k} = \frac{p'_m}{p'_k}$ $\forall\ m, k \neq l$. Then, because of monotonicity of $F^i(x_l; p_l)$ in x_l, we get

$$F^i(x_l;p_l) = F^i(\frac{b_l}{p_l};p_l) = F^i(\frac{b_l}{p'_l};p'_l) \leq F^i(\frac{b_l}{p_l};p'_l) = F^i(x_l;p'_l)\ \forall x_l \in I\!R_+$$

$$\Longrightarrow F^i(\cdot;p_l) \succeq F^i(\cdot;p'_l),$$

where $b_l = p_l x_l$ denotes the amount of money spent on good l.

If the density of consumption is strictly positive on the interior of the budget hyperplane the marginal distribution function is strictly monotonic. Thus a strict inequality holds and we obtain $F^i(\cdot;p_l) \succ F^i(\cdot;p'_l)$.

Q.E.D.

The notions of stochastic dominance and price independent shape of distributions can be related to Becker's example and its discussion in Hildenbrand (1994) as follows. Becker assumes that agents' consumption is characterised by a uniform distribution on the budget line. If the consumer has income b this implies that the marginal distribution is price independent. To see this for $L = 2$ one just has to realise that

$$F(x_1, p_1) = \frac{p_1}{b}x_1, \quad \forall\ x_1 \leq \frac{b}{p_1}.$$

This immediately implies $F(\frac{x_1}{p_1};p_1) = \frac{p_1}{b} \cdot \frac{x_1}{p_1} = \frac{p'_1}{b} \cdot \frac{x_1}{p'_1} = F(\frac{x_1}{p'_1};p'_1)$.

Thus the Becker example is a special case of price independent distributions and price independent distributions are a special case of distributions satisfying the assumption of stochastic dominance.

As Hildenbrand (1994) remarked the crucial point in Becker's example is not uniformness of distribution but price independence of budget shares of mean (or expected) demand, i.e. the fact that $E[\frac{X^i_l(p_l)p_l}{b}] = E[\frac{X^i_l(p'_l)p'_l}{b}]\ \forall\ p_l,\ p'_l,\ \forall\ l$. As we will see this condition is implied by price independence of the distribution's shape. The converse, however, does not hold in general.

uniform distribution	$\Longrightarrow$	price independent shape of distribution	$\Longrightarrow$	stochastic dominance
		$\Downarrow$		
		price-independent budget shares of (mean) demand		

The last implication is shown in the following

Proposition 6 *Let $X^i(p)$ denote agent i's consumption given price p and income b. Then the following holds: if the distribution of consumption $X^i(p)$ satisfies price independence then agent i's expected budget shares $b^i_l(p_l) = E[\frac{X^i_l p_l}{b}]$ of good l are price independent. The converse does not hold.*

Proof.

1. We show that for prices $p_l \neq p'_l$ of good l the budget share b^i_l of agent i is still the same:

$$\begin{aligned}
E[b^i_l(p_l)] &= E[X^i_l(p_l) \cdot \tfrac{p_l}{b}] = \frac{p_l}{b} \cdot \int_0^\infty (1 - F^i_l(u; p_l))du \\
&= \frac{p_l}{b} \cdot \int_0^\infty (1 - F^i_l(\frac{x_l}{p_l}; p_l))\frac{1}{p_l}dx_l \\
&= \frac{1}{b}\int_0^\infty (1 - F^i_l(\frac{x_l}{p'_l}; p'_l))dx_l \\
&= \frac{p'_l}{b} \cdot \int_0^\infty (1 - F^i_l(u; p'_l))du \\
&= E[b^i_l(p'_l)]
\end{aligned}$$

   The third of the above equalities follows from price independence.

2. It is obvious that the converse does not hold (equality of integrals does not imply equality of integrands).

The question we now want to consider is what can be said about aggregate demand $X(p) := \sum_{i=1}^I X^i(p)$ if assumption 1 holds for all agents. In fact, stochastic dominance will turn out to be a rationality assumption which is robust under aggregation.

3.2.2 Implications of stochastic dominance for market demand

Stochastic monotonicity of partial individual demand implies that the same property holds for partial market demand:

Proposition 7 *Let assumption 1 hold for $X_l^i(p), i = 1, ..., I,\ l = 1, ..., L$, $p \in S_+^{N-1}$. Let $X_l^k(p)$ be independent of $\sum_{i=1}^{k-1} X_l^i(p),\ 2 \le k \le I$. Then partial market demand $X_l(p) = \sum_{i=1}^{I} X_l^i(p)$ for good l is stochastically monotonic.*

Proof. Proposition 7 is a straight forward application of the following

Lemma 2 *Let $X_1, \dots, X_n, Y_1, \dots, Y_n$ be random variables, X_i stochastically dominated by Y_i ($Y_i \succeq X_i$), $i = 1, \dots n$. Let X_k be independent of $\sum_{i=1}^{k-1} X_i,\ 2 \le k \le n$. Then $Y := \sum_{i=1}^{n} Y_i$ stochastically dominates $X := \sum_{i=1}^{n} X_i$, i.e.: $Y \succeq X$.*

Proof. We give an inductive proof of this lemma.

$n = 2$: let $Y_1 \succeq X_1$ and $Y_2 \succeq X_2$, i.e. $F_{Y_i}(x) \le F_{X_i}(x)\ \forall\ x \in I\!R,\ i = 1, 2$. Let X_1 be independent of X_2 and let Y_1 be independent of Y_2. Then the distribution of the sum of these random variables is given by the convolution of the respective individual distributions:

$$\begin{aligned}
F_{Y_1+Y_2}(t) &= F_{Y_1} * F_{Y_2}(t) = \int_0^t F_{Y_1}(t-z)dF_{Y_2}(z) \\
&\le \int_0^t F_{X_1}(t-z)dF_{Y_2}(z) = \int_0^t F_{Y_2}(t-z)dF_{X_1}(z) \\
&\le \int_0^t F_{X_2}(t-z)dF_{X_1}(z) = F_{X_1} * F_{X_2}(t) \\
&= F_{X_1+X_2}(t)
\end{aligned}$$

$$\Longrightarrow \quad Y_1 + Y_2 \succeq X_1 + X_2.$$

We now assume the claim is true for some $n \ge 2$ and infer from this that it also holds for $n + 1$. Let $X_1, \dots, X_{n+1}, Y_1, \dots, Y_{n+1}$ be random variables, X_i stochastically dominated by Y_i ($Y_i \succeq X_i$), $i = 1, \dots n+1$. Let X_k be independent of $\sum_{i=1}^{k-1} X_i$, $2 \le k \le n+1$. Then $\tilde{Y} := Y_1 + \dots + Y_{n+1}$ stochastically dominates $\tilde{X} := X_1 + \dots + X_{n+1}$, $\tilde{Y} \succeq \tilde{X}$. To see this we write

$$\begin{aligned}
F_{Y_1+\dots+Y_{n+1}}(t) &= F_{\tilde{Y}} * F_{Y_{n+1}}(t) = \int_0^t F_{\tilde{Y}}(t-z)dF_{Y_{n+1}}(z) \\
&\le \int_0^t F_{\tilde{X}}(t-z)dF_{Y_{n+1}}(z) = \int_0^t F_{Y_{n+1}}(t-z)dF_{\tilde{X}}(z) \\
&\le \int_0^t F_{X_{n+1}}(t-z)dF_{\tilde{X}}(z) = F_{\tilde{X}_{n+1}+\tilde{X}}(t) \\
&= F_{X_1 \dots X_{n+1}}(t).
\end{aligned}$$

$$\Longrightarrow \quad Y_1 + \dots + Y_{n+1} \succeq X_1 \dots X_{n+1}.$$

The first of the above inequalities holds because of the assumption that the claim holds for some $n \ge 2$ (and we had shown that this is true for n=2).

Q.E.D.

Note that this result holds rather generally. In proposition 7 we do not make any assumption as to the type of distributions. Individual consumption distributions do not even need to be identical. Additivity of stochastic dominance also holds for multi-dimensional distributions. The proof of the lemma does not change but here, however, we need the one-dimensional case only because we use marginal distributions.

Remark 4

1. *(strict monotonicity) Let $X_l^i(p)$ be agent i's demand for good l with price p. Let partial individual demand for good l be stochastically monotonic, i.e.*

$$\tilde{F}^i(\cdot, p_l) < \tilde{F}^i(\cdot, p_l'),\ p_l' > p_l,\ i = 1, \dots, I$$

then expected partial market demand for good l is strictly decreasing in p_l. This can be shown with the help of integration by parts.

2. *1.) implies that all members of the main diagonal of expected market demand's Jacobian are strictly negative.*

3. *To establish uniqueness of an expected equilibrium one has to find properties of the individual consumption distributions which imply that the off-diagonal elements of expected demand's Jacobian are such that a criterion of diagonal dominance holds.*

4. *One example where the uniqueness conditions of the Jacobian are satisfied is the case of distributions with price independent shape.*

Besides the role of stochastic dominance for uniqueness of expected equilibria it has been applied in Industrial Organisation. Examples are Rob (1991) where the impact of stochastic dominance of market demand on the entry of firms is investigated, and Rob (1992).

In this section we have dealt with expected demand only. Actually realised demand will be treated in the following section.

4 Formulated and Actually Realised Demand

4.1 Quantity rationing

We have shown in section 3.1 that under the imposed conditions an expected equilibrium always exists, but once a state of nature is realised excess demand is distinct from zero with probability one (lemma 1). This is quite in line with the empirical observation that suppliers may set prices and that they may expect to sell exactly the amount of products they offer, but actual demand often exceeds supply or vice versa.

In this section we formalise one way of dealing with nonzero excess demand. The model we use is basically the one by Benassy (1982). In contrast to an alternative model by Drèze (1975) this model allows for agents formulating their net trade offers without taking into account their true constraints whereas in Drèze's model agents choose from their constrained budget sets. In the context of our model, where as little rationality requirements as possible should be made, we allow for agents making trade offers from their whole budget set (which, in the language of Benassy, actually means that they do not even need to take into account any 'perceived constraints').

We use the following notation[11]: $\tilde{z}_l^i \in I\!R$ is a net trade offer for commodity l sent to the market by agent i, i.e. $\tilde{z}^i = X^i(p,\omega) - e^i \in I\!R^L$. Thus the vector of trade offers of all agents is $\tilde{z} = (\tilde{z}^1, \ldots, \tilde{z}^I) \in I\!R^{IL}$. We denote the sum of trade offers for good l made by agents $1, \ldots, I$, by $\tilde{Z}_l = \sum_i \tilde{z}_l^i$.

Since in general markets do not clear we denote the actually realized transaction of agent i in commodity l by z_l^i. These actual transactions are determined by a rationing scheme $G_l^i(\tilde{z}^1, \ldots, \tilde{z}^I) = z_l^i$ satisfying the following conditions, the interpretation of which is straight forward.

1. $\sum_i G_l^i(\tilde{z}^1, \ldots, \tilde{z}^I) = 0$ for all l
2. $|z_l^i| \leq |\tilde{z}_l^i|$ and $z_l^i \cdot \tilde{z}_l^i \geq 0$
3. $\tilde{Z}_l \cdot \tilde{z}_l^i \leq 0 \implies z_l^i = \tilde{z}_l^i$
4. All functions G_l^i are continuous in their arguments.

As Benassy (1982) has pointed out these conditions are satisfied for a large number of rationing schemes, such as queuing, rationing tickets, priority systems, and proportional rationing. The crucial assumption made by Benassy, namely that an agent i takes into account his constraints on all markets except for the one he is trading on is not needed here, but only for a fixed point argument in Benassy's proof of the existence of an equilibrium. We will now define the concept of a rationing-allocation (which is much weaker than, though defined in analogy to, a Benassy equilibrium).

Definition 6 *A rationing - allocation is a set* $\{(\tilde{z}_l^i, z_l^i)_{l=1,\ldots L}^{i=1,\ldots,I}, p^*\}$ *of trade offers, actual transactions, and prices such that*

1. $\tilde{z}_l^i$ *are the net trade offers that result from a random realisation of demand, i.e.* $\tilde{z}_l^i = X_l^i(\omega; p^*) - e_l^i$
2. $z_l^i = G_l^i(\tilde{z}^1, \ldots \tilde{z}^I)$

[11] To keep notation as simple as possible we will just write $\tilde{z}^i$ for the realisation $\tilde{z}^i(p,\omega)$ as well as for the random variable $\tilde{z}^i(p,\cdot)$ if no danger of ambiguity is involved.

3. $G_l^i(\cdot)$ *satisfies the conditions of a rationing scheme.*

Property 1 of a rationing scheme is market clearing. Therefore the concept of a rationing-allocation is well defined.

The concept of a Benassy equilibrium has thus been weakened mainly in two aspects: firstly, agents do not take into account perceived constraints and, secondly, a rationing-allocation is in no way motivated as fixed point of a tâtonnement process. Although the main importance of our concept will become clear at the end of this section we would like to point out some features observed in reality that are captured by this concept:

1. Goods markets do not necessarily clear.

2. Rationing is rarely anticipated by consumers.

3. Feasibility problems can arise in the sense that an agent i cannot provide enough goods in exchange for the goods $z_l^i > 0$ she actually receives because she is rationed on some other market: a consumer might, for instance, not be able to sell her labour endowment so that she cannot pay for her actual transaction. Or, similarly, in a production economy a firm might have bought inputs but it is rationed in demand for its production good so that it might not be able to meet its liabilities at once. A way to formally close the model would be to allow agents to carry their liabilities over to a following period in which they would have to take them into account ex ante (by subtracting their liabilities from their endowment vector before they formulate their trade offers). But a multi-period model along these lines is beyond the scope of this chapter.[12]

[12]This would just complicate notation without shedding further light on the point we want to make in this section: we will show in the following proposition that the rationed amounts become more and more insignificant when the economy becomes large. But we will come back to this issue when we point out possible extensions of the model in the final section of this chapter.

4.2 Large economies

To prepare the way for our proposition about rationing in large economies we define the notion of a replica economy and prove the subsequent lemmas. Lemmas 3 and 4 are used to prove lemma 5. Lemmas 5 and 6 are used in the proof of proposition 8.

Definition 7 *With a stochastic exchange economy $\mathcal{E}(I)$ we associate the N-fold replica*

$$\mathcal{E}_N(I) = [\mathbb{R}^L, (X^i, e^i)_{i=1,\ldots,NI}]$$

where in the N-th replica the preferences and endowments of agent i, $i \in \{1, \ldots, I\}$, satisfy $X^i \sim X^{i+nI}$ and $e^i = e^{i+nI}$, $n = 1, \ldots, N-1$, and $X^i, 1 \leq i \leq IN$, are independent.

If there exists an expected equilibrium of $\mathcal{E}(I)$ there also exists an expected equilibrium of the $N-$fold replica $\mathcal{E}_N(I)$ as the following lemma demonstrates.

Lemma 3 *Let $(p^*, (E[X^1(p^*), \ldots, E[X^I(p^*)]))$ be an expected equilibrium of $\mathcal{E}(I)$. Then the vector $(p^*, (E[X^1(p^*), \ldots, E[X^{N \cdot I}(p^*)]))$ is an expected equilibrium of the N-th replica economy $\mathcal{E}_N(I)$.*

Proof. We assumed that in the N-th replica economy $\mathcal{E}_N(I), N \in \mathbb{N}$, all random variables X^i, $i = 1, \ldots, IN$, are independent of each other. The product probability space is well defined. Further, as distributions of X^i and X^{i+nI} coincide for all $i = 1, \cdots, I$ and for all $n = 1, \ldots, N-1$ so do the respective means: $E[X^i] = E[X^{i+nI}]$. For the price vector p^* of an expected equilibrium the following implication holds:

$$\sum_{i=1}^{I} E[X^i(p^*)] = \sum_{i=1}^{I} e^i \Longrightarrow \sum_{i=1}^{IN} E_N[X^i(p^*)] = \sum_{i=1}^{IN} e^i$$

where E_N denotes the expectations operator defined by the probability measure P_N on the N-th product probability space.

Q.E.D.

In what follows we will make use of Komolgorov's theorem which implies that the probability space and measure are also well defined in the limit economy ($N \to \infty$).

Lemma 4 *Let $(p^*, E[X(p^*)])$ be an expected equilibrium of $\mathcal{E}(I)$. Let $\beta_l^i(N,\omega) := \{j | \tilde{z}_l^j(\omega) \geq 0,\ j = i + nI \text{ for some } n = 1, \cdots, N-1\}$, $1 \leq i \leq I$, $\omega \in \Omega$, be the set of buyers of type i and let $\sigma_l^i(N,\omega) := \{j | \tilde{z}_l^j(\omega) < 0,\ j = i+nI \text{ for some } n = 1, \cdots, N-1\}$, $1 \leq i \leq I$, $\omega \in \Omega$, be the respective set of sellers of type i in a replicated economy $\mathcal{E}_N(I)$. Then, in a sequence of replicated economies $\{\mathcal{E}_N(I)\}_{N \in \mathbb{N}}$, there exist numbers i_b and $i_s \in \{1, \ldots, I\}$ such that the proportion $\frac{\#\beta_l^{i_b}(N,\omega)}{N}$ of buyers of type i_b and the proportion $\frac{\#\sigma_l^{i_s}(N,\omega)}{N}$ of sellers of type i_s of good l are bounded away from zero ($N \to \infty$) with probability one.*

Proof. Suppose that for some good $l \in \{1, \ldots, L\}$ the proportion of buyers $\frac{\#\beta_l^i(N)}{N}$ tends to zero for $N \to \infty$ with positive probability for all types of agents $i \in \{1, \ldots, I\}$:

$$P\Big(\liminf_{N \to \infty} \frac{\#\beta_l^i(N)}{N} = 0\Big) > 0 \qquad , i = 1, \cdots, I.$$

This would imply that

$$P\Big(\liminf_{N \to \infty} \frac{1}{N \cdot I} \cdot \sum_{i=1}^{NI} (X^i(p^*) - e^i) \ < \ 0\Big) > 0$$

which cannot be true because

$$\begin{aligned}
\frac{1}{N \cdot I} \cdot \sum_{i=1}^{NI} X^i(p^*) &= \frac{1}{NI} \cdot \sum_{i=1}^{I} \sum_{n=1}^{N} X^{i+(n-1)I}(p^*) \\
&= \frac{1}{I} \cdot \sum_{i=1}^{I} \frac{1}{N} \sum_{n=1}^{N} X^{i+(n-1)I}(p^*).
\end{aligned}$$

As in a replicated economy X^{i+nI}, $0 \leq n \leq N-1$, are all identically distributed the Law of Large Numbers can be applied. So, for $N \to \infty$, the above expression tends to

$$\frac{1}{I} \cdot \sum_{i=1}^{I} E[X^i(p^*)] = \frac{1}{I} \cdot \sum_{i=1}^{I} e^i$$

because p^* is the price vector of an expected equilibrium.

$$\Longrightarrow P\Big(\lim_{N\to\infty} \frac{1}{N\cdot I}\cdot\sum_{i=1}^{NI}(X^i(p^*)-e^i)=0\Big)=1.$$

This contradiction terminates the proof. For the proportion of sellers the proof is entirely symmetric to the one for the buyers.

Q.E.D.

Lemma 5 *Let $\mathcal{E}_N(I)$ be the N-th replica economy of $\mathcal{E}(I)$. Denote by*

$$D_l(p,N)=\sum_{i=1}^{I}\sum_{j\in\beta_l^i(N)}\tilde{z}_l^j(p)$$

the demand and by

$$S_l(p,N)=\sum_{i=1}^{I}\sum_{j\in\sigma_l^i(N)}\tilde{z}_l^j(p)$$

the supply of good l in $\mathcal{E}_N(I)$. Let$(p^*,E[X(p^*)])$ be an expected equilibrium of $\mathcal{E}(I)$. Then, for $l\in\{1,\ldots,L\}$,

$$\left|\frac{D_l(p^*,N)}{S_l(p^*,N)}\right|\to 1 \qquad a.s. \text{ for } N\to\infty.$$

Proof. In the following we will use the shorthand notation $D_l := D_l(p^*,N)$ and $S_l := S_l(p^*,N)$ and the corresponding shorthand β_l^i and σ_l^i for the sets of buyers and sellers unless specific reference to the arguments is necessary.

Note first that, for continuous demand densities,

$$\begin{aligned}\frac{D_l}{S_l} &= \frac{\sum_{i=1}^{I}\sum_{j\in\beta_l^i}\tilde{z}_l^j}{\sum_{i=1}^{I}\sum_{j\in\sigma_l^i}\tilde{z}_l^j}=\frac{\frac{1}{IN}\sum_{i=1}^{IN}\tilde{z}_l^i-\frac{1}{IN}\sum_{i=1}^{I}\sum_{j\in\sigma_l^i}\tilde{z}_l^j}{\frac{1}{IN}\sum_{i=1}^{I}\sum_{j\in\sigma_l^i}\tilde{z}_l^j}\\ &= \frac{\frac{1}{IN}\sum_{i=1}^{IN}\tilde{z}_l^i}{\frac{1}{IN}\sum_{i=1}^{I}\sum_{j\in\sigma_l^i}\tilde{z}_l^j}-1\end{aligned}$$

Now it is obvious from the proof of Lemma 3 that the numerator of the above fraction tends to zero for $N\to\infty$. It remains to show that the denominator is bounded away from zero. To see this is true note that

$$\frac{1}{IN}\sum_{i=1}^{I}\sum_{j\in\sigma_l^i}\tilde{z}_l^j = \frac{1}{I}\sum_{i=1}^{I}\frac{\#\sigma_l^i}{N}\cdot\frac{1}{\#\sigma_l^i}\sum_{j\in\sigma_l^i}\tilde{z}_l^j.$$

As the $\tilde{z}_l^j$ are all identically distributed ($j \in \sigma_l^i$) we apply the Law of Large Numbers to obtain

$$\frac{1}{\#\sigma_l^i}\sum_{j\in\sigma_l^i}\tilde{z}_l^j \to E[\tilde{z}_l^i|\tilde{z}_l^i < 0] \ < \ 0 \qquad N \to \infty.$$

But, according to Lemma 4, there exists an $i_s \in \{1, \ldots, I\}$ such that $\#\sigma_l^{i_s}/N$ is bounded away from zero with probability one. This implies that at least one expression of the form $\frac{\#\sigma_l^{i_s}}{N}\cdot\frac{1}{\#\sigma_l^{i_s}}\sum_{j\in\sigma_l^{i_s}}\tilde{z}_l^j$ in the above sum is strictly negative in the limit while all other terms are either negative or zero in the limit.

Q.E.D.

In what follows we use the norms $||\cdot||_1$ and $||\cdot||_2$ defined as usual: $||x||_1 := \sum_{i=1}^n |x_i|$ and $||x||_2 := \left(\sum_{i=1}^n x_i^2\right)^{\frac{1}{2}}$ for $x \in \mathbb{R}^n$.

Lemma 6 *Let $\mathcal{E}_N(I)$ be the N-th replica economy of $\mathcal{E}(I)$. Let X^i be independent, $1 \le i \le N \cdot I$, and let $E[(X^i)^2] < \infty$. Then*

$$\frac{1}{NI}\cdot\sum_{i=1}^{NI}||\tilde{z}^i||_2^2 \to \frac{1}{I}\sum_{i=1}^{I}E[||\tilde{z}^i||_2^2] < \infty \qquad a.s. \quad \text{for } N \to \infty.$$

Proof.

$$\begin{aligned}
\frac{1}{NI}\sum_{i=1}^{NI}||\tilde{z}^i||_2^2 &= \frac{1}{NI}\sum_{i=1}^{NI}\sum_{l=1}^{L}(\tilde{z}_l^i)^2 \\
&= \sum_{l=1}^{L}\frac{1}{I}\sum_{i=1}^{I}\frac{1}{N}\sum_{n=1}^{N}(z_l^{i+(n-1)I})^2 \\
&\to \frac{1}{I}\sum_{i=1}^{I}\sum_{l=1}^{L}E[\tilde{z}_l^{i\,2}] \quad \text{for } N \to \infty \\
&= \frac{1}{I}\sum_{i=1}^{I}E[||\tilde{z}^i||_2^2]
\end{aligned}$$

which is the arithmetic mean of the second moments all of which are finite by assumption.

Q.E.D.

We are now in a position to formulate the main proposition of this section stating that rationing becomes insignificant for large economies.

Proposition 8 *Let $\mathcal{E}(I) = [\mathbb{R}^L, (X^i, e^i)_{i=1,\dots,I}]$ be a stochastic exchange economy with I agents and $\mathcal{E}_N(I)$ be the N-th replica economy of $\mathcal{E}(I)$. Let X^i be independent, $1 \leq i \leq NI$, let $(p^*, E[X(p^*)])$ be an expected equilibrium of $\mathcal{E}(I)$. Let $E[(X^i)^2] < \infty$. Let $G(\cdot)$ be a proportional rationing scheme. Then*

$$\frac{1}{IN}\sum_{i=1}^{IN} \|\tilde{z}^i - G^i(\tilde{z})\|_1 \to 0 \text{ in prob. for } N \to \infty.$$

Proof. A proportional rationing scheme $G(\cdot)$ is defined in the following way: $G^i(\tilde{z}_l) = \tilde{z}^i_l \cdot \min(1, \frac{S_l}{D_l})$ for buyers ($i:\ \tilde{z}^i_l = X^i_l - e^i > 0$) and, analogously, $G^i(\tilde{z}_l) = \tilde{z}^i_l \cdot \min(1, \frac{D_l}{S_l})$ for sellers ($i:\ \tilde{z}^i_l = X^i_l(p) - e^i < 0$). Therefore, for $i \in \{1, \dots, I\}$,

$$\begin{aligned}
|\tilde{z}^i_l - G^i_l(\tilde{z}_l)| &\leq |\tilde{z}^i_l| \cdot \left(\left|1 - min(1, |\frac{S_l(p)}{D_l(p)}|)\right| + \left|1 - min(1, |\frac{D_l(p)}{S_l(p)}|)\right| \right) \\
&\leq |\tilde{z}^i_l| \cdot \left(\left|1 - |\frac{S_l(p)}{D_l(p)}|\right| + \left|1 - |\frac{D_l(p)}{S_l(p)}|\right| \right)
\end{aligned}$$

$$\begin{aligned}
\Longrightarrow \frac{1}{IN}\sum_{i=1}^{IN} \quad \| \quad & \tilde{z}^i_l - G^i_l(\tilde{z}_l)\|_1 = \\
&= \frac{1}{IN}\sum_{i=1}^{IN}\sum_{l=1}^{L} |\tilde{z}^i - G^i(\tilde{z}_l)| \\
&\leq \frac{1}{IN}\sum_{i=1}^{IN}\sum_{l=1}^{L} |\tilde{z}^i_l| \cdot \left(\left|1 - |\frac{S_l(p)}{D_l(p)}|\right| + \left|1 - |\frac{D_l(p)}{S_l(p)}|\right| \right) \\
&\leq \left(\frac{1}{IN}\sum_{i=1}^{IN} ||\tilde{z}^i||_2 \right) \cdot \left(||\alpha||_2 + ||\beta||_2 \right),
\end{aligned}$$

where α is the vector the l-th component of which is $1 - |\frac{S_l(p)}{D_l(p)}|$ and β is vector the l-th component of which is $1 - |\frac{D_l(p)}{S_l(p)}|$.

The last inequality follows from the Cauchy Schwarz inequality. In order to see that the expression we obtained tends to zero note that the first factor has a finite limit. This follows from Lemma 6. The second factor tends to zero because, due to Lemma 5, each component of α and β tends to zero.

Q.E.D.

As mentioned earlier in this chapter the reason why we examine this limit case ($N \to \infty$) is not that we attach much importance to this case itself but rather that we want to examine the impact of the size of the economy on the problem that market clearing does not hold in our model. Proposition 8 is an answer to this question in so far as it states that under the assumptions made the problem of rationing becomes arbitrarily small for sufficiently large economies.

This result, together with the non-existence result of lemma 1, explains both of the following features observed empirically: although markets do not clear entirely and rationing does occur it is still true that, on the average, agents are rationed by amounts which tend to be insignificant in large markets.

One implication of proposition 8 is that it states an optimality property of the price p^* defined by the expected equilibrium concept, namely that, given price p^*, excess demand converges to zero in large economies. The proof of our proposition relies heavily on the 'expected market clearing' property of p^*.

Another interesting question would be in how far proposition 8 can be generalised for arbitrary, or at least for some other, rationing schemes.

4.3 The finite case

An obvious question arising at this stage is what can be said about the finite case ($IN < \infty$). It seems desirable to quantify the degree of disequilibrium for a given size of the economy. This line is pursued by the following proposition which gives an upper bound for the probability that, for any good, average excess demand exceeds a given value $\epsilon > 0$.

Proposition 9 *Let $\mathcal{E}(I)$ be a stochastic exchange equilibrium. Let $(p^*, E[X(p^*)])$ be an expected equilibrium of $\mathcal{E}(I)$. Further, let the covariance matrix Σ of the random vector $\frac{1}{I}\sum_{i=1}^{I} X^i(\cdot;p^*)$ be non-singular. Then*

$$P\Big(\Big|\frac{1}{I}\sum_{i=1}^{I}(X_l^i(\cdot;p^*) - e_l^i)\Big| \geq \epsilon, \text{ for some } \mathrm{l} \in \{1,\ldots,\mathrm{L}\}\Big) \leq tr\ \bar{B}^{-1}\Pi,$$

where tr denotes trace, $\bar{B}$ is the unique positive definite matrix having ones on the main diagonal such that $\bar{B}\Pi^{-1}\bar{B}$ is diagonal, and Π is the matrix with typical elements $\Pi_{l,m} = (\frac{1}{\epsilon I})^2 \sum_{i=1}^{I} \sigma_{l,m}^i$ with $\sigma_{l,m}^i$ denoting the covariance $\mathrm{cov}(X_l^i, X_m^i)$.

Proof. The proof of our proposition relies on Theorem 3.6 in Olkin and Pratt (1958) which is a multivariate generalisation of Tchebysheff's inequality. It is shown there that for a random vector $y = (y_1, \ldots, y_p)$ with mean zero and non-singular covariance matrix Σ the following inequality holds:

$$P(|y_i| \geq k_i\sigma_i, \text{for some } i) \leq tr\ \bar{B}^{-1}\Pi,$$

where $\bar{B}$ is defined as in the above proposition and $\Pi = K^{-1}RK^{-1}$ with $R = (\rho_{ij})$ denoting the correlation matrix of y and K is a diagonal matrix with $k_i, i = 1, \ldots, p$.

We apply this result to our variable $y = \frac{1}{I}\sum_{i=1}^{I}(X^i(p^*) - e^i)$ which has mean zero because p^* is the price vector of an expected equilibrium. The covariance matrix of y is the covariance matrix of $\frac{1}{I}\sum_{i=1}^{I} X^i(p^*)$ and therefore non-singular by assumption.

It remains to calculate the typical elements of Π: we choose $k_l = \frac{\epsilon}{\sigma_l}$ where σ_l denotes the standard deviation of y_l. This standard deviation is given by $\sigma_l = \frac{1}{I}\Big(\sum_{i=1}^{I}(\sigma_l^i)^2\Big)^{\frac{1}{2}}$. It follows that

$$\Pi_{l,m} = (K^{-1}RK^{-1})_{l,m} = \frac{1}{\epsilon^2}\sigma_m\sigma_l\rho_{l,m} = \frac{cov(Y_l, Y_m)}{\epsilon^2} = (\frac{1}{\epsilon I})^2 \sum_{i=1}^{I} \sigma_{l,m}^i.$$

These equations hold because X^i and X^j, $i \neq j$, are independent.

Q.E.D.

Remark 5

1. *The inequality of the above proposition is sharp in the following sense: if the bound given is less than 1, there is a distribution for* $\frac{1}{I}\sum_{i=1}^{I} X^i$ *(with mean* $\frac{1}{I}\sum_{i=1}^{I} e^i$ *and covariance matrix* Π*) under which the bound is attained. Otherwise, there is a distribution such that* $P(|\frac{1}{I}\sum_{i=1}^{I}(X_l^i(\cdot;p^*) - e_l^i)| \geq \epsilon, \text{ for some l}) = 1$.
2. *An obvious implication of the proposition is* $P(||\frac{1}{I}\sum_{i=1}^{I}(X^i(\cdot;p^*) - e^i)||_1 \geq \epsilon, \text{ for some l}) \leq L \cdot tr\ \bar{B}^{-1}\Pi$. *This inequality is not necessarily sharp.*
3. *Again, as in proposition 8, it is essential that* p^* *is the price vector of an expected equilibrium.*
4. *The bound given in the above proposition depends negatively on* ϵ *and negatively on* I *which is a comparative statics result one would intuitively expect to hold.*

Proof. Part 1 is a direct implication of theorem 3.7 in Olkin and Pratt (1958). The rest is evident.

Q.E.D.

We formulated proposition 9 with reference to average demand. Alternatively, one can give a bound for the probability that *total* excess demand exceeds a given bound ϵ in which case the typical elements of Π no longer contain the factor $\frac{1}{I^2}$, i.e. $\Pi_{l,m} = (\frac{1}{\epsilon})^2 \sum_{i=1}^{I} \sigma_{l,m}^i$.

5 Conclusion

Economic theory has been criticised because of its lack of realism and its more or less counter-intuitive assumptions. In this chapter we have taken up the challenge and go a (certainly very preliminary) step towards a higher degree of descriptiveness of economic models by relaxing the utility maximisation hypothesis.

Another objective was to get some further understanding as to the following questions:

1. How heavily do results in general equilibrium theory rely on utility maximisation?
2. In how far are unrealistic conclusions due to unrealistic assumptions?
3. Which predictions as to the functioning of a market economy can be made without any assumption of utility maximisation?

In answer to these questions the following results have been obtained: although the realisation of excess demand will, with probability one, not equal zero, there exists a price vector for which markets do clear in expected terms. The properties of expected demand which are necessary to establish the existence of an expected equilibrium can be obtained under regularity conditions. But in order to derive boundary behaviour of excess demand one needs more restrictive assumptions. Uniqueness of an expected equilibrium is established for special classes of distributions.

Further, we have argued in section 4 that the discrepancy between non-market-clearing and expected market clearing becomes smaller and smaller when the economy becomes larger, more precisely: the average amount of rationing in the N-th replica economy tends to zero when N tends to infinity. For an economy of finite size an upper bound is given for the probability that rationing exceeds a given amount. This bound is sharp.

Summarising our sensitivity analysis one can state that some of the predictions made by utility maximisation models which break down when the assumption of preference maximisation is given up. These predictions are the least intuitive ones (e.g. boundary behaviour, instantaneous market clearing) while those predictions which correspond to salient features observed empirically (existence of equilibrium prices that lead to market clearing on the average, small amounts of rationing, properties of expected excess demand) can well be derived without any assumption of utility maximising behaviour.

A line of research that would seem to be a particularly interesting continuation of this work would be the project to

1. embed the model into a multi-period model

2. introduce asset markets to this model and to examine whether assets can reduce variance or improve allocations in the sense that they reduce the total amount of rationing

3. extend this model to a production economy and model firms' behaviour when they face demand uncertainty.

Extension no. 2 may deserve special attention for the following reason. Once assets can be traded our model can provide a foundation for an ad hoc assumption often made in financial markets models, namely the existence of noise traders. Noise traders would arise naturally here as rationing on the demand or supply side would cause agents to buy or sell assets respectively, i.e. noise trading arises from liquidity considerations.

6 References

Anderson, S.P., A. de Palma and J. Thisse (1992) : *Discrete Choice Theory of Product Differentiation;* MIT Press.

Becker, G. (1962) : Irrational Behavior and Economic Theory; *Journal of Political EconomyVol. 70:1-13.*

Benassy, J.P. (1982) : *The economics of market disequilibrium;* Academic Press.

Block, H.D., and J. Marschak (1960) : *Random Orderings and Stochastic Theories of Response;* In I. Olkin (ed), Contributions to Probability and Statistics. Stanford University Press, 97-132.

Bray, M. and N.E. Savin (1986) : Rational Expectations Equilibria, Learning and Model Specification; *Econometrica 54: 1129-1160.*

Debreu, G. (1950) : *Theory of Value. An axiomatic analysis of economic equilibrium;* New York Wiley.

Drèze, J. (1975) : Existence of an equilibrium under price rigidity and quantity rationing; *International Economic Review 16: 301-320.*

Geanakoplos, J. (1992) : Bounded Rationality; *Mimeo. Yale University.*

Grandmont M. (Ed.) (1988) : Temporary Equilibrium: selected readings; *Academic Press.*

Hardt, Michael (1995) : Random Demand replacing Utility Maximisation in a General Equilibrium Setting; *Mimeo. University of Bonn and London School of Economics.*

Hildenbrand, W. (1971) : Random Preferences and Equilibrium Analysis; *Journal of Economic Theory 3: 414-429.*

Hildenbrand, W. and A. Kirman (1988) : *Equilibrium Analysis;* North Holland.

Hildenbrand, W. (1994) : *Market Demand: Theory and Empirical Evidence;* Princeton University Press.

Klemperer, P.D. and Meyer, M.A. (1989) : Supply Function Equilibria

1 References

Anderson, S.P., A. de Palma and J. Thisse (1992) : *Discrete Choice Theory of Product Differentiation;* MIT Press.

Becker, G. (1962) : Irrational Behavior and Economic Theory; *Journal of Political EconomyVol. 70:1-13.*

Benassy, J.P. (1982) : *The economics of market disequilibrium;* Academic Press.

Block, H.D., and J. Marschak (1960) : *Random Orderings and Stochastic Theories of Response;* In I. Olkin (ed), Contributions to Probability and Statistics. Stanford University Press, 97-132.

Bray, M. and N.E. Savin (1986) : Rational Expectations Equilibria, Learning and Model Specification; *Econometrica 54: 1129-1160.*

Debreu, G. (1950) : *Theory of Value. An axiomatic analysis of economic equilibrium;* New York Wiley.

Drèze, J. (1975) : Existence of an equilibrium under price rigidity and quantity rationing; *International Economic Review 16: 301-320.*

Föllmer, H. (1973) : Random Economies with Many Interacting Agents; *Journal of Mathematical Economics 1: 51-62.*

Geanakoplos, J. (1992) : Bounded Rationality; *Mimeo. Yale University.*

Grandmont M. (Ed.) (1988) : Temporary Equilibrium: selected readings; *Academic Press.*

Hardt, Michael (1995) : Random Demand replacing Utility Maximisation in a General Equilibrium Setting; *Mimeo. University of Bonn and London School of Economics.*

Hildenbrand, W. (1971) : Random Preferences and Equilibrium Analysis; *Journal of Economic Theory 3: 414-429.*

Hildenbrand, W. and A. Kirman (1988) : *Equilibrium Analysis;* North Holland.

Hildenbrand, W. (1994) : *Market Demand: Theory and Empirical*

Evidence; Princeton University Press.

Klemperer, P.D. and Meyer, M.A. (1989) : Supply Function Equilibria in Oligopoly under Uncertainty; *Econometrica Vol 57 No 6: 1243-1277.*

Luce, R.D. (1959) : *Individual Choice Behavior: A Theoretical Analysis;* New York: Wiley.

Machina, M.J. (1985) : Stochastic Choice Functions Generated from Deterministic Preferences over Lotteries; *Economic Journal 95: 575-594.*

Magill, M. and Quinzii, M. (1992) : *Theory of Incomplete Markets;* Mimeo.

Manski, C.F. (1977) : The Structure of Random Utility Models; *Theory and Decision 8: 229-254.*

McFadden (1974) : *Conditional Logit Analysis of Qualitative Choice Behaviour;* In P. Zaremka (ed.), Frontiers in Econometrics. New York: Academic Press, pp. 105-142.

McFadden, D. (1981) : *Econometric Models of Probabilistic Choice;* In C.F. Manski and D. McFadden (eds.). Structural Analysis of Discrete Data with Econometric Applications. Cambridge: MIT Press, pp. 198-272.

Mirrlees, J.A. (1986) : Economic Policy and Nonrational Behaviour; *Nuffield College, Oxford. Mimeo.*

Novshek, W. and Sonnenschein, H. (1982) : Fulfilled expectations Cournot duopoly with information acquisition and release; *Bell Journal 13: 214-218.*

Olkin, I, and Pratt, J.W. (1958) : A multivariate Tchebysheff inequality; *Annals of Mathematical Statistics 29: 226-334.*

Quandt, R.E. (1956) : A Probabilistic Theory of Consumer Behavior; *Quarterly Journal of Economics 70: 507-536.*

Rob, R. (1991) : Bayesian Learning under Demand Uncertainty; *Review of Economic Studies 58: 655-675.*

Rob, R. (1992) : Sales, Uncertainty and the Determinants of Investment; *CARESS Working Paper 92-03.*

Simon, Herbert A. (1982) : *Models of Bounded Rationality;* MIT Press.

Simon, Herbert A. (1992) : *Economics, Bounded Rationality and the Cognitive Revolution;* Edward Elgar Publishing Company.

Thurstone, L.L. (1945) : *The Prediction of Choice;* Psychometrica 10: 237-253.

Tietz, R., Albers, W. and Selten, R. (Eds.) (1988) : *Bounded Rational Behaviour in Experimental Games and Markets;* Springer Verlag.

Tversk , A. (1972a) : Elimination by Aspects: A Theory of Choice; *Psychological Review 79: 281-299.*

Tversk , A. (1972b) : Choice by Elimination; *Journal of Mathematical Psychology 9: 341-367.*