UNIVERSITY OF SOUTHAMPTON

ESSAYS ON ENVIRONMENTAL POLICY, PLANT LOCATION AND FINANCIAL LIABILITY

Laura Valentini

Doctor of Philosophy

Department of Economics

Faculty of Social Sciences

April, 2000

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL SCIENCES

ECONOMICS

Doctor of Philosophy

ESSAYS ON ENVIRONMENTAL POLICY, PLANT LOCATION AND FINANCIAL LIABILITY

By Laura Valentini

Two major topics of discussion in the environmental economics literature are the international dimension of environmental policies and whether governments should impose strict joint liability on firms and their lenders for the environmental damage firms have caused. This thesis tries to make a contribution to these two topics by modelling multi-stage games of firms' location decisions and governments' strategic environmental policies, and by assessing empirically the impacts that a joint liability legislation enacted in the USA has had on firms' capital structure. Three chapters of the thesis are devoted to analysing the links between trade and environment and two chapters are devoted to analysing firms' capital structure and the impacts of environmental legislation on it. In Chapter 1 I briefly introduce the two topics. Chapter 2 surveys the literature on the impacts of trade liberalisation and free capital movements on environmental policies. In chapter 3 firms' location decisions and governments' environmental policies are analysed when there are incentives for agglomeration. The main implications are that there may exist a threshold in the environmental policy beyond which a country may lose its industrial base. There may be hysteresis effects in environmental policies, and last governments' competition to attract/deter plant location may cause environmental dumping, but it may cause also 'NIMBY'. Chapter 4 addresses the issue of move structure in these types of games. In the paper two main types of game structures are compared: in the Market Share Game firms locate before governments set environmental policies and in the Location Game firms locate after governments set environmental policies. It turned out that this issue must be distinguished from another issue which is the number of instruments in the Location Game. It is shown that the extent of environmental dumping in the Market Share Game may be greater or less than in the Location Game, and that there is more environmental dumping in the Location Game when governments use a single instrument than when they condition their instruments on the number of firms that locate in the country. In chapter 5 a survey of empirical studies on firms' capital structure is conducted. This survey is instrumental to Chapter 6 which analyses the impacts of different environmental regimes on the capital structure of firms, and in particular how much bank debt they will use. A theoretical model is used to show that introducing environmental liability on firms who have limited liability will cause them to increase bank debt use; extending liability also to banks has an ambiguous effect on bank borrowing. Under some assumptions bank debt will be lower than with liability only on firms but it may be lower or higher than with no liability at all. Second the model is tested empirically using US industry data. The empirical estimates confirm that bank debt increased in the period when liability was imposed only on firms, but it then decreased when liability was imposed also on bank, however consequences are not drastic. The last chapter is left for conclusions.

LIST OF CONTENTS

Preface.. 5

Acknowledgements.. 6

1. Introduction.. 7

2. Environmental Regulation, Multinational Companies and International Competitiveness

Abstract... 21

2.1 Introduction.. 22

2.2 What Does Economic Analysis Tell Us About Claims 1-5?............... 24

2.2.1 The Simplest Model.. 25

2.2.2 Allowing for Multi-National Production............................ 27

2.2.3 Allowing for Many Firms.. 29

2.2.4 Many Firms and Plants... 33

2.2.5 Agglomeration Effects... 34

2.2.6 Time Structure Revisited... 38

2.2.7 Summary... 39

2.3 Empirical Studies.. 41

2.4 Policy Implications.. 48

2.5. Conclusions.. 56

3. Plant Location and Strategic Environmental Policy with Inter-sectoral Linkages

Abstract... 63

3.1 Introduction.. 65

3.2 The Model.. 71

3.3 Simulations Results... 79

3.3.1 Base Case... 79

3.3.2 Variations of the Base Case....................................... 80

3.3.2.1 Variations in Transport Costs for Upstream Firms........ 81

3.3.2.2 Variations in Transport Costs for Downstream Firms.... 82

3.3.2.3 Variations in Importance of Intermediate Goods.......... 82

3.3.2.4 Implications... 83

3.3.3 Exogenously Specified Emission Taxes.......................... 84
3.3.3.1 Experiment A.. 84
3.3.3.2 Experiment B.. 85
3.3.3.3 Implications.. 86
3.3.4 Nash Equilibria in Taxes.. 87
3.3.4.1 Base Case... 87
3.3.4.2 Variations on Base Case..................................... 91
3.3.5 Summary... 93
3.4 Conclusions.. 94

4. Is Environmental Dumping Greater When Firms Are Footlose?

Abstract... 97
4.1 Introduction... 98
4.2 The Model.. 104
4.3 Firms Locate Before Governments Set Policies........................... 108
4.3.1 Second-Stage Game... 108
4.3.2 First-Stage Game.. 110
4.3.3 Conclusion of Market Share Game................................... 111
4.4 Firms Locate After Governments Set Policies............................ 112
4.4.1 Second-Stage Game - Single Instrument Case......................... 112
4.4.2 Second_Stage Game - Two Instrument Case............................ 113
4.4.3 First-Stage Game – Some Properties of Welfare Functions............ 114
4.4.4 First-Stage Game – Single Instrument............................... 116
4.4.5 First-Stage Game – Two Instruments................................. 118
4.4.6 Conclusions of Location Game....................................... 122
4.5 Simulation Results... 125
4.6 Conclusions.. 133
Appendix A – Analysis of Second-Stage Location Game With Two Instruments.. 135

5. Credit Rationing and Capital Structure: a Survey of Empirical Studies

Abstract... 146
5.1 Introduction... 147
5.2 Credit Supply and Firms' Characteristics............................... 149

5.3 Some Empirical Studies.. 152
5.4 Credit Rationing in The Consumers' Loan Market.......................... 159
5.5 Capital Structure and Firms' Own Characteristics.......................... 167
5.6 Some Empirical Studies.. 172
5.7 Conclusions.. 188

6. Environmental Liability and The Capital Structure of Firms

Abstract.. 191
6.1 Introduction.. 192
6.2 The Theoretical Model.. 196
6.2.1 Regime 0: No Liability... 197
6.2.2 Regime I: Liability on Firm with Compulsory Insurance with Costless Monitoring.. 198
6.2.3 Regime II: Liability on Firm with Compulsory Insurance with Costly Monitoring 201
6.2.4 Regime III: Liability of the Firm with No Mandatory Insurance... 205
6.2.5.Regime IV: Joint Liability on Firm and Bank, with no Mandatory Insurance... 208
6.2.6 Summary... 210
6.3 Empirical Analysis of the Effects of Environmental Liability on Bank Borrowing.. 211
6.3.1 Survey of Previous Studies of Firms' Capital Structure and of the Impact of Joint Liability on Cost of Capital........................ 211
6.3.2 Description of Data and Variables 217
6.3.3 Regression Results... 220
6.3.4 The Impact of CERCLA on US Bank Borrowing..................... 221
6.4 Conclusions.. 222

7. Conclusions and Directions for Further Research........................ 228
References.. 233

LIST OF TABLES AND FIGURES

Chapter 2

Table 2.1: Impacts of a European Energy Tax on Plants and Costs.. 46

Fig. 2.1: Welfare Functions: One Firm.................................. 57

Fig. 2.2: Welfare Functions: More than One Firm..................... 58

Fig. 2.3: Abatements Falls on Marginal Costs.......................... 59

Fig. 2.4: Hysteresis Effect in Environmental Policies................ 60

Fig. 2.5: Environmental Policy: Strategic Substitutes/Complements 61

Chapter 3

Table 3.1: Equilibrium Tax Rate for Tax Cutting Equilibria......... 89

Table 3.2: Payoff Matrix for Asymmetric Equilibrium................ 90

Table 3.3: Payoff Matrix for NIMBY Equilibrium...................... 91

Fig. 3.1: Hysteresis Effect in Environmental Policies................ 95

Chapter 4

Table 4.1: Equilibria in market share game............................. 138

Table 4.2: Equilibria for location game-max. welfare................ 138

Table 4.3: Equilibria of location game- average........................ 139

Table 4.4: Equilibria of location game – race-to-the-bottom.......... 139

Table 4.5: Average Increase in Emission Limits and Proportionate Welfare Losses... 140

Fig. 4.1 Welfare functions: One Firm.. 141

Fig. 4.2 Welfare functions: More than One Firm....................... 142

Fig. 4.3 Equilibria of location game with a single instrument......... 143

Fig. 4.3 Equilibria of location game for two instruments.............. 144

Chapter 5

Table 5.1: Summary of Empirical Studies.............................. 187

Chapter 6

Table 6.1: Relationship between bank debt and financial data....... 224

Table 6.2: Estimates of Models I-III Using Interest Rate.............. 225

Table 6.3: Estimates of Models I-III Using Monetary Base............ 225

Table 6.4: Impact of CERCLA on Bank Borrowing.................... 226

PREFACE

Chapters 2, 3, 4, and 6 have all been written together with my supervisor Alistair Ulph. Chapter 2 has been presented at the conference on *Internationalisation of the Economy, Environmental Problems and New Policy Options* , Potsdam, October 15-17, 1998 and is forthcoming in the Conference Volume. Chapter 3 has been published in *Resource and Energy Economics*, vol. 19, 1997. An earlier version of Chapter 4 has been presented at the first World Congress of Environmental and Resource Economists, Venice, June 1998, it has been submitted to the *Scandinavian Journal of Economics* and is currently under revision.

ACKNOWLEDGEMENTS

This thesis has been a long process, with changes and interruptions throughout. I am clearly indebted to my supervisor, Alistair Ulph. His supervision and support went well beyond that of a normal supervision; he pushed me to continue in all moments, but he also left me alone when it was the case. He has given me a lot of support and listening in bad moments and a lot of his time always. The question is where he finds it, and we (all his students) are building some models to explain this. He has been a Master in teaching rigorous thinking, and I hope that I have learnt at least a small fraction of his knowledge.

I am grateful to all my friends, collegues, professors and secretaries in the department, they have also been a support and created a warm atmosphere around me. I need to mention in particular the 'Wizard of the department' Ray O'Brien, he has always solved uncountable problems I had with programming and with econometrics. Liz Leeks, Alistair's secretary and by transitivity mine as well, has also been helpful in all possible ways and a good friend.

Many friends have been involved during this time, in particular Marzia Raybaudi-Massilia and Martin Sola have been a constant over the whole period, and in large part the completion of this work is also due to them. Roberta Piermartini has been a 'pal' in these last two years and a strong support in her own special way. I mention briefly Pina, Gianna, Alberto, Giovanni, Francesco, Penny, Surjinder Colin and Grazia for many discussions about economics and life in general. Lucy O'Shea has been almost an assistant in her friendship, and she has always asked a lot of useful questions.

A thank also to all participants to the 'The House of the People', they made many evenings and week-ends extremely enjoyable and they have been nice enough to always appreciate my ideas and cakes with enthusiasm and good humour.

I am very much grateful to my current flatmates, they have made the atmosphere in the flat that of a family.

Back in Italy I am grateful to my family, they accepted my absence with only 'few' complaints and encouraged me not to stop. I am particularly grateful to Antonella and Simone Morandini for their help. I am also grateful to Carlo Carraro and Ignazio Musu for having taught me economics first and having helped me to continue my studies.

Last but no least, grants from 'Ente Einaudi', ESRC and financial support from the department are also gratefully acknowledged.

Chapter 1

Introduction

Two topics under discussion in environmental economics literature are the international dimension of environmental policies and whether governments should impose strict financial liability on firms for the environmental damage they have caused. First I briefly analyse the links between environmental policy and trade policy and then I briefly illustrate the main issues of imposing legal environmental liability.

The international dimension of environmental policies.

A first obvious reason why environmental policies may have an international dimension is because of the presence of transboundary pollution or global pollution problems like climate change. In this case if countries do not act co-operatively then they will ignore the effect that their own pollution levels have on other countries, and some forms of international policy coordination must be enforced. However, if countries are different then to set equal emission levels across all countries may not be Pareto efficient. In fact it might be preferable to have some form of international agreements that specify a different emission level for each country. Moreover, even ignoring transboundary pollution or global pollution there is a second reason why environmental and trade policies are related, which brings about some concerns.

On the one hand policy makers and industrialists fear that the imposition of environmental restrictions would make their firms lose competitiveness in the international markets and that at the extreme these firms could decide to relocate to other countries where environmental regulation is less stringent, towards the so-called 'pollution havens'. This concern is even stronger when firms are multinationals, as it is easier for them to relocate production to another plant in a different country. On the other hand environmentalists fear that liberalisation of trade and capital flows will first of all increase world consumption and production and thus will worsen the environment to a greater extent. Second, in a world of free trade and hence absence of trade instruments, environmentalists are concerned that exactly because of the impacts that environmental regulation may have on firms' international competitiveness, governments may be induced to set too weak environmental policies compared to what they would have done in absence of this international competitiveness concern. Parallel to relocation of production by multinational firms, another concern is that the extent of environmental dumping is larger when firms are multinational. The fear is thus that countries will engage in a so-called 'environmental dumping' to prevent relocation of production to other countries or to attract plant location from other countries, and that at the extreme countries could engage in a 'race to the bottom' in environmental policies.

Chapters 2-4 of this thesis concentrate on this last aspect of the international dimension of environmental policies, i.e. the extent to which governments may use environmental policies to achieve strategic trade objectives. In what follows I concentrate more on the links between strategic environmental policies and plant location.

In chapter 2 these concerns are addressed more precisely in a survey of some recent theoretical and empirical studies literature. In this survey it is shown that following the literature these concerns cannot in general be sustained both on theoretical grounds and from the empirical evidence. However while the fear that firms may relocate to countries with weaker environmental policies and this in turn may induce governments to set too lax policies when they act strategically cannot be sustained, also the opposite result that governments will set too tough or at least not too weak environmental policies cannot be sustained either. In this literature we start by presenting a simple model by Hoel (1994) of one firm deciding where to locate between two countries, and the two countries acting strategically in setting emission limits. As a result of strategic behaviour the two countries will engage in a 'race to the bottom' setting higher limits than they would set if they acted non co-operatively or if they acted non strategically (in the sense that the emission limit is also higher than the level that equates marginal damage costs to marginal abatement costs). However, this result is not robust to changes in the assumptions, because if the damage cost is very high so that the welfare is negative then governments will set too tough emission limits in order to deter location in their countries. Some particular extensions to this simple model are subsequently considered in the survey. One possible extension is the introduction of transport costs (Markusen-Morey–Olewiler (1993)) to provide a rationale for the existence of multinational firms. The firm now may decide not only in which country to locate but also to set a plant in each country, trading off fixed location costs with transport costs. The result in this case is that indeed when there are multinational firms governments may set too weak environmental policies, but the causality runs in the opposite direction than suggested by common wisdom: it is because governments set lax environmental policies that firms have the incentive to build more plants and not that lax environmental policies are a response to the presence of multinational firms. A particular feature of these types of models is that the welfare functions are discontinuous for some levels of the environmental policy variable. This has two implications, the first is that marginal analysis cannot be used, so that results and conclusions cannot be derived analytically but must be drawn through numerical simulations. The second important implication is that a toughening of the environmental policy may not have relocation consequences until a certain threshold is reached, but once that threshold level is reached an increase in the environmental tax, say, may have disastrous consequences, with a country losing all its industrial base.

In a further extension Markusen (1997) considers what happens when there are many firms and many plants, that is firms can decide whether to locate in the home or foreign country only or to locate a plant in each country. However because of the complexity of the model Markusen does not analyse strategic government behaviour but considers only what are the effects if the foreign government introduces a unilateral environmental policy. The fact that there are transport costs, which are necessary to rationalise multinational companies, gives the foreign governments some protection in setting the environmental policies: the government is able to set some environmental regulation without production switching from its country to the other.

The empirical evidence also does not show any indication of a significant negative relationship between plant location and environmental regulation. For example Levinson (1996) analyses new plant location decisions of US firms across US states and finds that environmental stringency even if negative in explaining new plant formation is significant only for two particular measures of environmental stringency and also in these cases the impact was rather small. Also Mani Pargal and Huq (1997), who conduct a similar study for indian firms, find no significant impact of environmental stringency on plant location decisions.

A further extension to the literature is to consider the effects of environmental policies on plant location decisions when there are incentives to agglomerate. This is the content of the third chapter. In that model we assume that there are two identical countries and two sectors, an upstream industry which provides an input of production for the downstream industry; in each industry there are two firms. Given the input-output linkages between the two industries a particular feature of this model is that demand for the intermediate goods is endogenous. This is particularly interesting because it allows us to consider if environmental policies imposed on one industry have consequences also on the other even if the latter is not directly subject to environmental regulation. Transport costs are positive so that again firms may decide to set up more than one plant, trading off fixed costs of building the plant with transport costs. In the model we assume first, consistent with the literature, that pollution is only local, and second that only the production of the intermediate good is polluting and we analyse what are the impacts of introducing an environmental tax on intermediate firms on location of both upstream and downstream firms. In order to be able to determine the impacts that environmental policies have on plant location we need to assume that governments set the environmental policy prior to firms deciding where to locate. The structure of the model is thus the following: in a first stage governments act strategically and set their emission taxes, in the second stage all producers choose their location simultaneously, and in the last stage producers choose their output. Output of final goods is determined by Cournot competition given an exogenous demand; this output then determines the demand for the intermediate good and the equilibrium output for the upstream producers is determined consequently. The model even though it is very simple, cannot be solved analytically, and for this we had to rely on numerical simulations to draw our conclusions. We run three main sets of experiments: first we determined possible location configurations for different values of the parameters. Second we saw what are the consequences of using exogenously specified emission taxes and last we determined the Nash equilibrium in tax rates.

The first set of simulations was run to see what determines decisions on plant locations. We abstracted from any environmental externality and so assumed that production is not polluting as we did not want to analyse any policy intervention. In the simulations we simply varied the basic parameters (essentially the costs parameters). In all cases we had multiple equilibria and we adopted a bargaining selection criteria. In most cases agglomeration was the selected equilibrium. The alternative equilibria are also interesting to analyse. By changing the parameters we found that agglomeration in general is associated with a middle range of transport costs, while if transport costs become high then for the firms it becomes more important to be near their customers and this makes them build more plants. On the contrary if transport costs are relatively low and so are the intersectoral linkages then firms will locate plants so as to take advantage of some degree of market power.

In the second set of experiments we introduced the environmental externality, and in particular we assumed that only production in the upstream industry is polluting. We then analysed the impacts of country 1 only imposing an emission tax. In this case we found a possibility of hysteresis effects in environmental policy. In fact starting with a situation of all firms agglomerated in a country we started to increase the emission tax in that country, until a second equilibrium emerged in which firms locate in the other country, but firms still did not change their location; however beyond a certain threshold of the environmental tax all firms moved their production to the other country. When we started to decrease the environmental tax back in the first country below the critical level, this was not enough to induce firms to relocate back in the original country, but the emission tax had to be reduced further.

Last we allowed governments to act strategically in setting their emission taxes. There are a number of results depending on the values given to the parameters of the model. For low levels of the environmental damage it is true that countries engage in tax competition and both set the tax rate below the marginal damage cost. All firms locate in one of the two countries, but the levels of welfare are the same. For medium levels of the damage costs we end up with an asymmetric equilibrium in which one country sets the tax rate at the marginal damage cost while the other country sets it at a slightly lower level and gets all the production. This country also has a lower level of welfare than the first country. For higher levels of damage cost countries again set equal tax rates but now there is no tax cutting, the tax rates are equal to the marginal damage costs. Last, for very high levels of damage both countries will set so high tax rates (still not above the marginal damage cost) that firms close down production completely.

It has been objected that this structure, for which first government set their environmental policies and then firms decide where to locate is not very realistic because it implicitly assumes that environmental policies, namely taxes, are less flexible than plant location decisions. Moreover firms may decide where to locate not just on the base of past policies, but more on the base of future expected policies. In this case it would be better to model a situation where firms decide first where to locate and then governments set the environmental policies in a subsequent stage. This is the content of the fourth chapter. In that paper we compare the extent of environmental dumping under two different game structures. The first move structure is one which represents a situation where location is fixed, that is firms decide where to locate first and governments set their environmental policies in a second stage (emission limits in this case). This game has been called Market Share Game. The second move structure represents a situation where location is endogenous, that is governments set emission limits in the first stage and firms locate in the second stage. This game has been denoted Location Game. As said before the usual presumption is that the degree of environmental dumping should be stronger when firms' location is endogenous. To keep things simple we assume that there is one industry only with two firms deciding where to locate between two countries. In this model, differently from the previous one, we assume that there are no transport costs so that firms will always build one plant only. Another important assumption is that the environmental damage is strictly convex in output. In order to properly compare the results under two different game structures we assume that, since governments can condition their policies on the number of firms that have located in the country in the Market Share Game, they are allowed to do so also when firms have not yet located before they set their policies (that is in the Location Game). To see what is the impact of changing the number of instruments available, we also compare location and environmental policies when governments can use only one instrument with location and environmental policies when governments can use two instruments, the emission limits they set when only one firm will locate in the country or the emission limits they set when two firms will locate in the country.

The assumption that governments can use two different instruments depending on the number of firms that locate in their country, even if location is endogenous, allows us to compare the alternative outcomes and rank the stringency of policies when firms are footlose and when location is fixed. One result derived from the analysis is that when the outcome of the Market Share Game is the Nash equilibrium with firms locating in separate countries, this may also be the outcome of the Location Game, when governments use two instruments, and so in this case the order of the moves is not relevant. A second result is that it is possible that competition for market share leads to weaker policies than competition for location. Last, there can be no general presumption that using one policy instrument leads to more environmental dumping than using two instruments.

With the help of numerical simulations we also assess the probability of occurrence of the different possible equilibria for the Market Share Game and for the Location Game with one and two instruments. Clearly in order to do so we need to specify particular functional forms for profit, damage and welfare functions. For the Market Share Game the Nash equilibrium with firms located in separate countries is by far the dominant equilibrium for many values of the parameters. In the Location Game when we constrain governments to use a single instrument then the 'race to the bottom' outcome emerges in 90% of the cases. When we allow governments to use two instruments then we have again the result that the Nash equilibrium is the dominant outcome, and this emerges in 50-70% of the cases. This confirms our idea that it is important to distinguish between the issue of move structure and the issue of the number of instruments. An appropriate comparison of the outcomes generated by the two move structures, separating the issues of the number of instruments, reveals that in fact the order of the moves is irrelevant in determining the final equilibrium in the majority of the cases.

Strict financial liability of environmental damage.

The types of policies considered in the first part of the thesis (taxes and emission limits) are a form of *ex-ante regulation*. The purpose of their use is that of inducing polluters to reduce emissions. It has been seen that in some extreme cases these policies may even be used to prevent the damage by deterring location of polluting enterprises in a country. Recently some countries have moved to use *ex-post* regulation, for which polluters are held liable for the environmental damage they have caused. The underlying presumption is that polluters will be induced to undertake appropriate levels of effort to reduce the risk of environmental damage, knowing that they will be held liable in the future if the accident occurs. However, polluters can escape paying for the damage if they have limited financial liability by going bankrupt. Firms' owners then will substitute own equities with debt as a way of financing, beyond the efficient level, and as a result the level of effort will also be inefficient, being too little. A possible solution to this problem is then to extend environmental liability also to lenders. The rationale now is that lenders will condition their loans on the level of effort taken to prevent the occurrence of the environmental damage. This should ensure an efficient level of effort. This solution is efficient if lenders can perfectly and costlessly monitor the level of effort firms take. If on the contrary, as is more realistically the case, information between firms and lenders is asymmetric and monitoring is costly, then lenders' liability fails to be the efficient solution. First of all the level of effort will still be too low, and second some distortion in the firms' capital structure will be introduced. The argument in favour of this solution is that in any case lenders have better access to information than any other governmental institution, moreover banks already monitor firms when they establish a credit relationship, and so they only would need to add the environmental riskiness in the monitoring process.

This type of regulation was first introduced in the USA in 1980 with the enactment of the *Comprehensive Environmental Response Compensation and Liability Act* (CERCLA). According to this act all 'potentially responsible parties' are held strictly jointly and severally liable for cleaning up contaminated properties. The act applied to hazardous waste sites, and the potentially responsible parties have included current and past owners of the site, operators of the site or companies that arranged the disposal, treatment and transport of the hazardous substances. According to the law also creditors of the site operators may be held liable if they are involved in the financial management of the operator and capable of influencing the activity of the site. This part of the law was invoked in various occasions, with different results. In particular in 1990 the bank *Fleet Factors Corporation* was held liable to pay for the clean up of the damage caused by one of its borrowers after that had gone bankrupt. This case is in general considered as the main example after which lenders' became clearly liable. In the last paper we shall pick this date as the starting date of a joint liability regime of firms and lenders.

Environmental liability regulation is relevant also for European countries. At the moment countries have different liability regimes and the European Union is considering the possibility of harmonising such regimes. At the moment lenders have not been considered liable yet, but this possibility has not been tested in courts.

The purpose of the second part of the thesis is then to analyse how different liability regimes affect firms' capital structure. First, in the fifth paper I review empirical studies of credit rationing. The purpose of the survey was to see how the allocation of credit according to borrowers' characteristics had been estimated in the empirical literature. In the survey both consumers' loans and commercial loans studies have been reviewed. The analysis on the first type of loan is interesting because, since detailed data based on consumers' characteristics are available, it is possible to estimate both demand and supply functions and then see if consumers' are credit-rationed according to their socio-economic characteristics. Unfortunately this is hardly possible for the commercial loan markets, and empirical estimates are reduced form equations of demand and supply. The demand for debt is the outcome of firms' optimisation of their capital structure and it clearly depends on the other characteristics (financial characteristics first of all) of the firm. The supply of credit depends instead on suppliers' characteristics. However private financial intermediaries will also observe some characteristics of their borrowers in order to decide whether to discriminate among firms, when they cannot completely discriminate with the interest rate. Most of the studies regress the proportion of long term private (mainly bank) debt on firms' characteristics, and a list of the independent variables included and of the results is reported in the survey. One major difficulty, together with the scarcity of data, is that many of the firms' characteristics are not directly observable and so some proxies must be used instead. Moreover some proxies can be used as indicators of more than one characteristic at the same time, so in the end it may be difficult to tell with precision if a firm's attribute is relevant or not. Last, because estimates reported are reduced form equations, there is an additional difficulty: in some cases the relationship between the amount of private debt and a characteristic (for example volatility of earnings growth) is positive or negative depending on whether we are observing the demand or the supply of credit. One advantage of this is that if one sign dominates we can clearly say if demand or supply dominates, one disadvantage is that it is likely that the two effects compensate each other and we have a non significant estimated coefficient for that characteristic not because the variable is not significant but because of this 'double' effect. If on the other hand the relationship between the amount of credit and a characteristic has the same sign for both demand and supply, it is very likely that that variable is significant as the two reinforce each other, but then on the other side it is not possible to tell which function dominates. So in general empirical estimates do not yet give a very clear answer on the relationships between the proportion of private loans and firms' characteristics, as they do not yet give a clear answer on the question whether firms are credit rationed according to their attributes. This survey however was important with regard to the following chapter, where we also estimate an empirical model of bank loans in order to determine if the environmental liability has played a role in determining how much bank loans firms take on.

In the last paper we analyse the impacts of different environmental liability regimes on firms' capital structure, and in particular on firms' choice between equities, public debt and private debt. In the first part of the paper we derive a theoretical model of firms' capital structure to determine how three main different liability regimes affect bank lending. We distinguish between the following liability regimes: initially no liability is imposed on either firms' or lenders, in a second regime environmental liability is imposed only on firms (this regime is in fact divided in more than one depending on whether there are monitoring costs or compulsory insurance), and in the third regime liability is imposed on both firms and lenders. It is necessary to distinguish between these different liability regimes because they have different impacts on the demand and supply of credit and thus on the level of borrowing actually observed. In this paper we show that if firms have limited financial liability, then they will increase their demand for private (bank) debt if they are held liable for environmental damage. The reason is that by substituting equities with private debt firms can protect their shareholders from the environmental liability. However if liability is extended also to banks, so that firms and banks are jointly liable, then the effect on bank borrowing is ambiguous, and it is not possible to predict whether bank borrowing should increase or decrease. However, under particular assumptions it is possible that bank borrowing is less under joint liability than under firms' only liability. It is still not possible to tell whether bank borrowing under the third regime, liability of both firms and banks, is higher or lower than bank borrowing under the no liability regime.

In the second part of the paper we use US industry-level data to estimate a reduced form equation of long term bank debt. We regress the proportion of bank debt over total assets on industries' characteristics and on a dummy variable that represents the environmental liability legislation, as measured by clean up costs. The dummy variable is designed to capture the three liability regimes consistently with the theoretical framework. The change in environmental liability is designed in the following way: a first dummy variable takes value zero from the first observation (1973) until 1980, it then takes the amount of the clean up costs from 1981 to 1990 and then again zero. A second dummy variable takes value zero from the beginning of the sample until 1990 and then it takes the value of clean up costs from 1991 onward. These two dummies allow us to distinguish among the three time phases: a first phase that corresponds to pre-CERCLA introduction (no liability regime); a second phase that corresponds to the firms' liability regime, after the introduction of CERCLA, and a third phase that corresponds to firms' and banks' joint liability which became manifest after the *Fleet Factors* case in 1990. Our empirical analysis strongly supports the theoretical predictions and in fact the coefficient of the first environmental variable is significantly positive, suggesting that imposing environmental liability on firms has increased the extent to which they rely on bank borrowing. The coefficient of the second environmental variable is still positive (its significance depends on the model specification) but lower than the coefficient of the first variable, meaning that imposing environmental liability on both borrower and lender may have a net positive impact on the amount of credit privately borrowed. This in turn means that bank borrowing increases slightly compared to the non liability situation but it decreases compared to when liability is imposed on firms only. For some industries, such as Chemicals and Primary Metals, the impact of imposing environmental liability on firms has been that of increasing bank borrowing of 15-20%. When liability is extended also to borrowers these same industries show that bank borrowing increases of only 1-2% compared to the non liability case.

In the very last section conclusions of main results and directions for further work are outlined.

Chapter 2

Environmental regulation, multinational companies and international competitiveness

Abstract

Concerns have been expressed that in a global market place with mobile capital, national governments will have incentives to set weak environmental policies ("environmental dumping") to protect the international competitiveness of their domestic firms, that these incentives are particularly strong in industries where plants may be relatively footloose, so that governments are concerned to prevent "capital flight", and that footloose plants are particularly associated with multinational firms. It is then often suggested that appropriate policy responses would be to seek to harmonise environmental regulations or impose minimum standards for environmental regulations.

In this paper we set out these concerns in terms of a number of more precisely made claims and then review recent developments in economic analysis (including some of our own work) and empirical evidence to show that the claims cannot be generally sustained and that the suggested policies may be harmful. However, devising more appropriate policies is by no means straightforward.

2.1. Introduction.

There is considerable public concern that globalisation - the liberalisation of trade and capital flows - leads to policy competition between governments, which in the context of environmental policy may induce governments to set too lax environmental policies (environmental dumping) in order to protect their international competitiveness. It is further argued that this incentive is particularly strong when firms are footloose so that governments may worry about 'capital flight', and multinational firms are particularly prone to such footloose behaviour. The fear of environmentalists is that competition between governments to retain or attract such footloose firms is thought to trigger a 'race-to-the-bottom' in environmental policies. In the NAFTA debate the possibility of firms/plants relocating from the US to Mexico was an issue and a similar concern is expressed about the enlargement of the EU to include Eastern European countries. To counteract such incentives for environmental dumping it is often suggested that agencies such as the European Union should seek to harmonise environmental policies of member states or at least set 'minimum standards' for environmental policies.

Behind this set of concerns we can identify a number of separate claims:

Claim 1. Firms will locate production in countries with weaker environmental policies.

Claim 2. Countries with footloose firms will do less well than countries with less footloose firms.

Claim 3. Trade liberalisation gives governments incentives to weaken their environmental policies for strategic trade reasons ('environmental dumping').

Claim 4. There will be more environmental dumping when firms are footloose than when their locations are fixed.

Claim 5. Multinational firms are particularly prone to switch production between countries in response to environmental policies, in the extreme case closing plants in some countries.

Claim 6. The appropriate policy responses to environmental dumping are the harmonisation of environmental policies or the setting of 'minimum standards' for environmental policies.

In this paper we shall review recent literature, both in economic analysis and empirical work, and show that none of the above claims can be sustained as general propositions. In the next section we review what can be said in terms of economic analysis about claims 1 - 5 while in section 3 we review recent literature on empirical aspects of these claims. In section 4 we discuss the policy implications of the previous two sections, and in particular assess the status of claim 6.

2.2. What does economic analysis tell us about claims 1 - 5?

To address all the issues addressed in claims 1 - 5 would require a model with three key features:

(i) To take seriously issues of location of plants and firms, and in particular of multinational companies requires models in which there are significant increasing returns to scale, captured most simply by fixed costs of setting up plants and firms. This immediately implies that we will be dealing with imperfectly competitive markets[1]. It also means that we shall need to consider strategic interactions between firms both in output markets and in their location decisions[2].

(ii) To take seriously the notion of 'environmental dumping' we also need to assume that there is strategic competition between governments in terms of their environmental policies and one definition of environmental dumping is to contrast the policies governments would set if they act non-cooperatively with those they would set if they acted cooperatively[3]. So we shall need to model a policy game between governments.

(iii) To take seriously the notion that in setting their environmental policies governments are concerned with their impact on location decisions of firms we need to assume not only that firms will change their location decisions depending on the environmental policies set by governments but that governments can *commit* themselves to environmental policies prior to firms choosing where to locate. Taken together with points (i) and (ii) this implies that we need to consider multi-stage games and that there will be important timing issues in such games. We shall follow the literature in contrasting two extreme cases. In *Location Games* governments are able to commit themselves to environmental policies prior to

[1] For a review of what can be said about incentives for environmental dumping in competitive markets see Ulph (1997a).
[2] Modelling location choices can become complex because of the discrete nature of these choices. This causes payoff functions for firms to be discontinuous, and in turn this causes the welfare functions for governments to be discontinuous (see Markusen, Morey and Olewiler (1993), Ulph and Valentini (1997)) for further discussion.
[3] An alternative definition is to contrast the policies governments set when acting non-cooperatively with a simple 'first-best' rule for environmental policy such as equating marginal abatement and damage costs.

firms choosing their locations. We shall contrast this with what we call *Market Share Games* where firms first choose their locations and governments then set their environmental policies, and this contrast will throw light on issue of whether there is more or less environmental dumping when governments worry about capital flight[4]. This is a very crude contrast between the relative degrees of commitment of governments to their environmental policies and firms to their location choices[5].

Not surprisingly, there is no single model which captures all the features set out above so we shall review a small number of models which cast light on different aspects of the claims we wish to test[6]. We begin with an extremely simple model of location choice based which we used in a recent paper (Ulph and Valentini (1998)) and which draws heavily on the model by Hoel (1997). This has the merit of providing a setting in which many of the concerns and claims set out in the introduction can be justified. We shall then show how moving away from the very simple model invalidates the claims.

2.2.1 The Simplest Possible Model.

Assume that there is only one firm and only two countries in which it can locate. There are no transport costs or capacity constraints, so that the firm will only use one plant, so that the notion of a multinational firm cannot arise. The firm sells its product in countries other than the two in which it can locate. Production costs are the same in both countries. Production causes pollution, some of which may be abated at a cost. Unabated pollution causes damage in the country in which the firm is located, but countries have the same damage costs. Clearly the firm will

---

[4] Note that even in the Market Share Game firms may still base their location decisions on the environmental policies which they *expect* (rationally) governments to introduce after they have located.

[5] Of course a richer model would be a multi-stage game where one models explicitly what leads to different degrees of commitment by firms and governments; Feenstra (1998) conducts such analysis for the case where firms have to choose investment in capital, and shows that a multi-stage game can lead to different conclusions from a two-stage game but does not consider location decisions. D. Ulph (1995) also notes that rather than starting from the assumption that firms choose locations *de novo* in a world of liberalised trade and capital flows it may make more sense to assume that globalisation is a process so that one starts from a position of autarky in which there are already established firms located in different countries.

[6] For other surveys see Markusen (1996), Rauscher (1995), Wilson (1996).

locate in the country that sets the weaker environmental policy. We assume that the country in which the firm locates imposes a 100% pure profits tax so that profits accrue to the country of production. Governments use a simple emission limit (e) to regulate pollution. Welfare of the country where the firm locates is thus given by profits of the firm less environmental damage, shown as $V(e)$ in Figure 2.1, while welfare of the other country is zero.

If the firm chooses its location before the government sets its policy, then the government in which the firm locates will simply set the emission limit, e^*, which maximises welfare. This will be the usual first-best emission limit such that marginal abatement cost equals marginal damage cost, so there is no strategic element to environmental policy[7]. Since the policy is the same for both countries, the firm will randomise where it locates, so each country will get *expected* welfare $U(e^*) = 0.5V(e^*)$. If governments choose their policies first, then if governments set the same emission limit, e, we assume that the firm randomises where it locates so that each government gets expected welfare $U(e) = 0.5V(e)$, shown in Figure 2.1. If governments set different emission limits the firm locates in the country with the higher limit, and that country gets welfare $V(e)$. The equilibrium of the game in emission limits will be a simple 'race-to-the-bottom' in which both governments set emission limit $\bar{e}$ where welfare gets driven to zero. It cannot be an equilibrium for the governments to set the same emission limit below $\bar{e}$ because by setting a marginally higher emission limit a government could get welfare $V(e) > U(e)$, and it cannot be an equilibrium for governments to set different emission limits if the government with the higher emission limit gets strictly positive welfare, since the other government could set a marginally higher emission limit and get positive rather than zero welfare.

[7] This depends crucially on the assumptions that there are no consumers in each country and the country in which the firm locates earns all the profits. So although the firm is a monopolist the government is quite content for it to maximise profits. The only distortion as far as the government in which the firm locates is concerned is the environmental distortion. If there were consumers in the country the government would set policy weaker than first-best to offset the monopoly distortion; if the country did not capture all profits the government would set policy tougher than first-best as a way of capturing foreign profits.

This very simple model would seem to confirm most of the fears outlined in the introduction: firms choose to locate where environmental policies are weakest (claim 1); competition between governments to attract capital induces them to weaken their environmental policies (claim 3) and indeed this takes the form of a 'race-to-the-bottom'; environmental policies are weaker when governments have to worry about the location of firms than when they do not, i.e. $\bar{e} > e^*$(claim 4); and the countries are worse off with footloose firms than when firms are not footloose, in the sense that $U(e^*) > U(\bar{e}) = 0$ (claim 2). The model does not allow us to address claim 5.

However, even within the context of this very simple model, not all these claims are robust. For the argument depends on the assumption that $V(e^*) \geq 0$, so that there are some emission limits for which damage costs are less than the profits which the firm earns. But suppose environmental damage costs are so high that $V(e^*) < 0$, then while the outcome with fixed location will be as before (the government in the country with the firm sets emission limit e^*) with endogenous location both governments will set such tough environmental policies that no firm would want to locate in their countries - the NIMBY outcome. This would immediately overturn claims 2, 3 and 4.

It might be argued that the NIMBY case is only relevant for a few extreme cases (e.g. disposal of nuclear waste) so that the general conclusion from this simple model is to confirm many of the concerns and claims outlined in the Introduction. We now turn to various ways in which the simple model might be extended.

2.2.2 Allowing for Multi-National Production.

A key feature of the simple model is that there are no transport (trade) costs so that, even if we take the more general Hoel (1997) version of the model which allows for consumers in both countries, the firm would only operate one plant. Markusen, Morey and Olewiler (1995) (henceforth MMO) use a model which has a very similar structure to the simple Hoel model except that they have transport costs[8], so that the firm has to decide whether to have a single plant, and if so where to locate it, or to have a plant in each country (i.e. to go multi-national). Note that an immediate implication is that if, in the absence of any environmental policies in either country, the equilibrium choice is for the firm to locate plants in each country, because profits are strictly greater with the multi-national equilibrium than with an equilibrium with a single plant exporting to the other country, then if one country introduces an environmental policy there will be a range of values for the environmental policy before it would pay the firm to switch to the equilibrium with a plant located in the country with no environmental policy. Thus the introduction of transport costs, which is needed to rationalise a multinational pattern of production, provides a degree of protection to a country to set a tougher environmental policy than other countries without fear of losing plants to rival countries. So this contradicts claims 1 and 5.

MMO again compare the policies governments would set after the firm has made its choice of plant location, with the policies governments would set prior to the firm deciding where to locate its plant(s). MMO reach the same conclusion as Hoel about the possibility of a NIMBY outcome - with high enough damage costs both governments will set prohibitive environmental policies to deter any plant being located in their countries, despite the fact in terms of global welfare it would be desirable that the product be produced. But there is an interesting twist to the 'race-to-the-bottom' case. When governments set their policies after the firm has chosen its plant locations (so, as in the simple model, there is no strategic competition between governments) there will be two possibilities - the firm chooses a single plant and exports to the other country or it locates a plant in each country, with the first outcome being chosen when the fixed cost of setting up a plant is relatively high. In both cases when we switch to having the governments set environmental policies before the firm chooses plant locations then there will be competition to weaken environmental policies, so claims 3 and 4 remain true. But if in the non-

---

[8] There are some other differences which do not affect the general point we make; thus there is no third set of countries to which the firm can export; where the firm sets up a single plant in one country the governments can use an export tax. Because, as noted in footnote 2, of the discontinuities this introduces in the firm's payoffs and hence welfare payoffs they have to rely on special functional forms and some numerical examples to illustrate their argument.

strategic case, the outcome involved the firm setting up a single plant, then in the process of competition the firm may decide to switch to having two plants, while if the non-strategic outcome involved the firm having two plants this will remain the outcome when the governments compete strategically. Thus strategic competition may lead to the firm proliferating plants, but, except in the NIMBY outcome, not to it reducing the number of plants. The rationale is this. The firm is trading off the fixed costs of setting up plants against the transport costs of having to export. Suppose in the non-strategic case governments set tough environmental standards so that production costs are high; then output will be relatively low, transport costs will be low relative to total production costs and it will not be economic for the firm to carry two sets of fixed plant costs; however as the governments compete and weaken environmental policies this will reduce production costs relative to transport costs, expand sales in each country and make it more attractive for the firm to set up a second plant.

Thus allowing for the possibility of multi-national production means that, if again we exclude the NIMBY outcomes, in addition to governments setting weaker environmental policies when they take account of plant location decisions, this may lead to excessive numbers of plants being set up - too many multi-national plants. In other words the link between weak environmental policies and multinational firms may be the opposite of what environmentalists suppose - it is because governments set weak environmental policies that this may allow multinational patterns of production to come into being; multinational firms are a response to weak environmental policies rather than weak environmental policies being a response to multinational firms.

2.2.3 Allowing for Many Firms.

In 2.2.2 we extended the simple model by allowing a single firm to set up more than one plant. In this section we consider what happens if we revert to the assumption of no transport costs, so each firm has a single plant, but now assume there may be more than one firm. As the simplest possible extension to the simple model of section 2.2.1 we take the model of Ulph and Valentini (1998) in which there are two firms which engage in Cournot competition. The key difference this makes to the conclusions of the simple model is that even if we ignore the NIMBY case, environmental policy when governments set policy before firms choose to locate may be *tougher* than when they set policy after firms choose where to locate; i.e. claim 4 may no longer hold. There are four reasons why introducing more firms changes the conclusions of the simple model.

(i) With fixed locations there is now the possibility that the two firms locate in separate countries. In this case, given the assumptions we have made and contrary to the previous two sub-sections, there will be strategic incentives for governments to engage in environmental dumping and set weaker environmental policies than they would set if they cooperated[9] for the usual rent-shifting reasons familiar from strategic trade theory. To see why this might change the conclusion it is possible that, for some parameter values, the non-cooperative equilibrium between governments when firms have already decided to locate in separate countries, involves such intense competition that governments set such lax environmental policies that countries get negative welfare. When governments set their policies before firms locate they would never choose to end up with negative welfare; so, for these parameter values, even if there was a race-to-the-bottom in which countries got zero welfare with endogenous locations this may involve governments setting tougher policies than in the game where firms' locations are fixed.

(ii) Precisely because there is now competition between the firms, it may no longer be the case that relaxing environmental policies always leads to higher profits for firms; in the absence of environmental policies total output would be higher than that which maximises profits, so there will be cases (with low environmental damage costs) where toughening environmental policies may raise profits and this may limit the race-to-the-bottom.

[9] In particular the assumptions of Cournot competition and the fact that all profits accrue to the country in which the firm is located lead to environmental dumping (see Ulph (1997a) for more discussion.

(iii) If damage costs are strictly convex, then, for any given level of output and pollution by each firm, while having both firms locate in one country will double the profits that country can earn it will more than double the environmental damage costs that country has to bear. Now if we think of the two-firms analogue of the race-to-the bottom argument set out in Figure 2.1, we want to compare the welfare a country gets if it sets a higher emission limit than its rival and hence attracts both firms, $V(e)$, with the expected welfare it gets if it sets the same emission limit as its rival and both firms randomise where to locate, $U(e)$; in calculating this expected welfare we now include the possibility that the two firms locate in different countries. Because of the convex damage cost argument it may no longer be the case that $V(e) > U(e)$; indeed Ulph and Valentini show that there must always be some values of e below $\bar{e}$ for which $V(e) < U(e)$ and indeed this may be true for all e. This introduces the possibility of multiple equilibria for the game where firms are footloose, and while these will include the race-to-the-bottom equilibrium they may also include equilibria which have tougher emission limits than when firm locations are fixed. Hoel (1997) makes a similar argument when extending his model from one to many firms.

(iv) A further implication of the convex damage cost argument is that when we consider the environmental policies governments will set when firms' locations are fixed, it is quite natural to assume that they will set different environmental policies depending on whether one or two firms locate in their countries, in other words we can think of environmental policies being conditioned on the number of firms that locate in those countries. However the argument sketched out in (iii), and which is used by other authors when considering models of endogenous firm location with more than one firm[10], assumes that governments set a single environmental policy independent of the number of firms, so that if governments set different emission limits then all firms locate in the same country. But if we want to compare the difference in policies when firms have fixed or endogenous locations, and we do this by varying the move structure of the game, then if we allow a government to condition its policy instrument on the number of firms

[10] See Hoel (1997), Markusen, Morey and Olewiler (1993) although in the latter case the authors do not study policy competition between governments.

located in its country under one move structure we should do so under both. This means that when governments set policies before firms locate, governments can separate the emission limits they set to attract one firm to locate from that which they set to attract both firms to locate. Speaking loosely this means that governments are not driven into an all-or-nothing race-to-the-bottom but can settle for sharing the firms between them. Perhaps not surprisingly, if governments know they can secure an equilibrium where one firm will locate in each country they will set the same level of emission limits as if these locations were fixed, so we get exactly the same outcome irrespective of whether firms locations are fixed or endogenous. This occurs for a wide class of parameter values. This is illustrated in Figure 2.2 where *W(e)* is the welfare a country gets when the two firms locate in separate countries and *V(e)* is the welfare a country gets when both firms locate in a single country. e^N is the Nash equilibrium emission limits the governments would set in the Market Share Game where they set policies after the firms have chosen to locate in separate countries. If the configuration of *W* and *V* is as shown in Figure 2.2, which is true for a wide range of parameters, then e^N would also be an equilibrium of the Location Game where governments set policies before firms choose their location, because it would not pay a government to try to attract both firms to locate in its country.

In summary, introducing many firms means that even excluding the NIMBY outcome, it need no longer be the case that governments engage in more environmental dumping when firms choose their locations in response to government policies than when government policies are set after firms have fixed their locations, and indeed for a wide class of cases the move structure makes no difference. Ulph and Valentini (1998) showed that the greater the degree of substitution between firms' products (and hence the greater the degree of market competition) the more likely it was that environmental policy would be *tougher* in Location Games (endogenous locations) than in Market Share Games (fixed locations).

2.2.4 Allowing for many firms and plants.

Markusen (1996) extends the analysis of the two previous sections by considering a general equilibrium model with two identical countries (denoted *h* and *f* for home and foreign) and two sectors - a competitive sector and a non-competitive sector. Within the non-competitive sector firms may be national (i.e. a single plant in *h* or a single plant in *f*) or multi-national (type *m*). The number of each type of firm and the amount that each firm produces is determined endogenously, using a zero-profit condition. Given the complexity of the model, Markusen does not analyse the non-cooperative setting of environmental policies by the two governments but rather just studies the impact of the foreign government unilaterally introducing an environmental policy which can affect either the marginal or fixed cost of production in country *f*. However there are results from his model which allow us to say something about claim 5 - that multinational firms encourage the switching of production between countries in response to environmental policies.

Figure 2.3 shows the types of configurations of firms that exist in equilibrium for different combinations of trade costs and environmental cost penalties in country *f*. The top row shows what would happen in the absence of any environmental policy in *f*; for low transport costs only national firms exist, while for high transport costs only multinational firms exist. For low transport costs, as production costs in country *f* rise, production will be switched away from *f* firms to *h* firms and eventually only *h* firms survive. With high transport costs, as costs in country *f* rise multinational firms will also switch production between from their *f* to *h* plants, and as costs in *f* rise it is possible for national firms in *h* to enter the market and export to *f*. Because of general equilibrium effects, type *f* firms can also emerge. With high enough production costs, multinational firms shut down (this is just the converse of the argument in section 2.2 that saw multinational activity expand as environmental polices weakened) and eventually only *h* firms survive.

Note first that the cost differential at which production in *f* shuts down completely is (slightly) higher when we start with multinational production than when we start with solely national firms. Markusen also reports that a detailed analysis of production shows that production switches more slowly from *f* to *h* when we start with multinational firms than with only national firms. Both these findings would seem to contradict claim 5 - that multinational firms exacerbate the process of switching production out of countries with high environmental costs. However Markusen noted that these findings are not due to multinational firms *per se* but rather to the fact that the existence of multinational firms can only be rationalised when there are high transport costs and it is this that protects production in country f from increases in its production costs due to environmental policy. However, in welfare terms high transport costs raise the costs of tougher environmental policies in f for the obvious reason that it increases the cost to consumers in f having to buy more of their consumption from country h and this effect outweighs any production gains, at least for Markusen's model. Dampening down the loss of production in response to environmental policies may not be beneficial.

A similar point was made in Ulph (1994) where it was shown that using tax rebates to reduce the incentive for firms to relocate abroad in response to environmental policy (i.e. reducing the extent to which domestic firms are footloose) may be counterproductive if that leaves domestic firms faced with higher costs and hence domestic consumers with higher prices than would otherwise be the case; Motta and Thisse (1994) make a similar point with respect to the use of protectionist policies to reduce the incentives for firms to relocate abroad.

2.2.5 Agglomeration Effects.

In all the models studied in the previous sections, environmental policy has affected location decisions through its impact on costs of production and hence on profits. To focus attention on the impact of environmental policies on location we have assumed that countries are identical (so there are no comparative advantage factors affecting location decisions). In the models in sections 2.2.1 and 2.2.3 with no transport costs small differences in environmental policies would be sufficient to induce firms to locate in the country with weaker environmental policies. In the models in sections 2.2.2 and 2.2.4 transport costs, which are necessary to rationalise multinational production, provided a degree of protection for a country to set a tougher environmental policy than rivals without losing all its production. Indeed Figure 2.3 shows that as environmental costs rise in country *f* there will be a steady decline in the number of plants and firms located in country *f* rather than any sudden exodus of production. The reason for this is that if a small toughening of environmental policy in *f* makes it marginally profitable for a plant to switch from country *f* to country *h* then that switch will marginally raise the profitability of firms that remain in *f* and reduce the profitability of plants located in *h*, thus reducing the incentive for any other plant to switch location.

Consider what happens to this argument if what matters in making location decisions is more than just costs of production, but also proximity to markets or sources of supply. Thus suppose that because of the input-output structure of production for producers in a particular sector, a significant fraction of their market will be producers in other sectors who use the output of this particular sector as inputs to their production processes; similarly, a significant fraction of the inputs used by producers in the particular sector will be the outputs of producers in yet other sectors. Thus because sectors are linked in the structure of production, the location decisions of producers in different sectors become interdependent. This provides incentives for agglomeration of producers. Consider then what happens if a producer in a particular sector decides to close a plant in a particular location. As noted above, within the sector itself that has the usual effect that by reducing supply in that sector it will raise the profits of the producers who remain in that sector in that location. But it will also have two knock-on effects. It will reduce the demand for the products it used as inputs, and so reduce the profits of plants which supplied that producer, which will typically be plants located close to the original producer. Second it will raise the input costs of plants which used the output of the original plant as inputs, since they will now have to get their inputs from more distantly related producers. Again these customers will have been located close to the original plant. If these reductions in profits in related sectors were sufficiently strong to cause the closure of some of the plants in those sectors that would in turn have negative impacts on the profits of the plants remaining in the original sector, which, if inter-sectoral linkages were strong enough could offset the original boost to profits of those plants caused by the closure of the original plant. Venables (1994) and Ulph and Valentini (1997) analyse models of strategic environmental policy when there are agglomeration effects due to inter-sectoral linkages of production.

There are two implications of this analysis of agglomeration. First, there is the possibility that with strong inter-sectoral linkages there is the scope for quite catastrophic effects of policy on location decisions of producers when critical thresholds are reached. This can be characterised by the concept that a country can lose its manufacturing base in a particular set of related industries. Thus even if there are transport costs which would be expected to give countries a degree of isolation of their production from increases in domestic costs, agglomeration effects may reintroduce the possibility that there could be critical thresholds at which a small toughening of environmental policy in one country triggers a substantial exodus of production. Second, when agglomeration effects are strong what matters to producers is being located close to producers in related sectors; where that happens to be is less important. This can mean that for a range of parameter values, including policy parameters, there can be multiple possible equilibria, e.g. it is perfectly consistent with a particular set of parameters that a particular set of industries be located either in h or f, while outside that range of parameters there may be a single equilibrium. This has an important implication for policy that it introduces a kind of "hysteresis effect". Suppose that at a very low level of environmental taxes, say, the only equilibrium is for a set of industries to locate in f. As environmental policy in f gets stricter there may emerge another equilibrium in which the industries could locate in h. But given that the industries are already located in f, no individual producer would wish to switch to h. When environmental policy in f gets strict enough, it is no longer possible to sustain the industries in f, and production switches to h. But now if f subsequently relaxes its environmental policy, by the same argument, the producers will not switch back to f unless the environmental policies reverted to the very low level at which location in f was the only possible equilibrium. Figure 2.4, from Ulph and Valentini (1997) illustrates this possibility for a two-country, two-sector (upstream and downstream) model in which there are two firms in each sector who have to decide how many plants to locate in each country. For a wide class of parameter values, in the absence of any environmental policy agglomeration effects lead to all firms locating a single plant in country 1. Country 1 introduces an emission tax but as long as it lies below 1.5 the strength of agglomeration effects means that it remains a unique equilibrium for all firms to locate in country 1. For taxes between 1.5 and 2.0 there is a second equilibrium in which 3 firms relocate in country 2; for taxes above 2.0 this second equilibrium is unique. So an emission tax above 2.0 would trigger a rapid exodus of firms from country 1 but country 1 would have to cut its tax below 1.5 to attract these firms back again.

2.2.6 Time Structure Revisited.

As D. Ulph (1995) notes, it is not really appropriate to use the term "hysteresis" for the above discussion, for there is no real dynamics in which environmental policies change and firms relocate; it is simply a question of multiple equilibria in what we have called a Location Game with agglomeration effects where we start with essentially a blank sheet in terms of firms and their locations and consider how firms would choose the locations of their plants once and for all in response to environmental policies previously set by governments. As noted in footnote 5, a proper analysis of hysteresis would require a multi-period interaction between governments and firms in which governments can change their policies from one period to the next and firms can change their locations in response to these changes in policy. As he notes, it is then important to distinguish between sunk costs, which are incurred when a plant is initially established and would have to be incurred again if a plant was relocated, from recurrent fixed costs of producing a positive output level in any period. It is the need to incur this sunk cost which gives a degree of commitment to firms location decisions and which may mean that locations selected in previous periods may not be changed in response to policies which would have induced an alternative location choice by firms making that location choice from scratch[11]. For such a multi-period model of hysteresis to work one would also need to explain why governments would wish to change environmental policies from one period to the next, and whether there are any factors (analogous to sunk costs) which might lead governments to commit to environmental policies (reputation effects might be a candidate). In his paper, the dynamics used by D. Ulph to motivate a change in environmental policy is the opening up of countries to trade. This has the usual effects - an increased market size effect which acts to increase profits, and increased competition effect which acts to reduce profits and a relocation effect of production switching from countries with high autarkic costs to those with low autarkic costs. In addition there will be any induced change in government environmental policies. He shows

[11] Motta and Thisse (1994) also note that because firms have prior locations in which they have incurred sunk costs the extent of delocation in response to environmental policies may be less than suggested by models which ignore such sunk costs

that while the opening up of trade will lead governments to set weaker environmental policies than under autarky, for a wide range of parameters the number of firms will be exactly the same as under autarky. This is because the gain in profits due to the net effect of increased market size and weaker environmental policies less the impact of increased competition is insufficient to compensate for the sunk costs needed to establish new firms.

However it should be noted that the trade liberalisation which drives the change in environmental policy is unanticipated by firms, for otherwise this would have influenced their initial location decisions. In the context of environmental policy we know of no fully specified multi-period analysis in which governments calculate their environmental policies each period and firms calculate their location decisions each period, with firms and governments acting strategically and with rational expectations, with a proper account of what might determine the relative degrees of commitment by firms and governments which would allow a proper analysis of hysteresis effects.

2.2.7 Summary.

If, in addition to the complexities outlined in the previous paragraph, we add the desirability of including transport costs and the simultaneous determination of multinational and national firms, and the possibility of inter-sectoral linkages with agglomeration effects it is clear why we said at the start of this section that we are a long way from having a single model which can address all of the issues outlined in the Introduction. Nevertheless from our review of a number of models which throw some light on certain aspects of the issues we can reach the following conclusions about claims 1 -5 set out in the Introduction.

1. Allowing for transport costs, agglomeration effects, sunk costs all mean that even with no comparative advantage effects there could be significant differences in environmental policies between countries without inducing firms to switch plants to countries with weaker environmental policies.

2. Even if firms are tempted to relocate abroad in response to difference in environmental policies it does not follow that a country is better off having less footloose firms, because if that leaves domestic firms with higher costs and domestic producers with higher prices, welfare may be lower than if firms had relocated abroad and exported back to the home country.

3. While there are certainly circumstances under which governments have incentives to weaken environmental policies when acting non-cooperatively, as we have shown there are other cases where governments will want to set too tough policies (e.g. NIMBYISM, where governments seek implicitly to tax foreign profits).

4. While there are models where governments will set much weaker environmental policies when they are trying to influence locations than when they set policies with fixed locations, this is by no means always the case; the discussion in section 2.2.3 showed that there may be many cases where environmental policies are unaffected by the timing of government policy setting relative to firms location decisions; when policies differ, the more intense is market competition (as measured by the degree of substitution between firms products) the more likely it is that environmental policy will be *tougher* when governments take account of the location decisions of firms in setting their policies.

5. Finally the discussion in sections 2.2.2 and 2.2.4 emphasised that because multinational patterns of production are most likely to be competitive against national production when transport costs are high relative to production costs, these high transport costs will also reduce the sensitivity of domestic production to differences between domestic environmental costs and foreign environmental costs. Moreover it may be that multinational firms are a response to weak environmental policies rather than weak environmental policies being a response to multinational firms.

2.3. Empirical studies.

In this section we review some of the empirical literature which attempts to assess whether environmental legislation has a significant impact on the plant location decisions of firms. Because there are several excellent recent surveys of this literature, (Cropper and Oates (1992), Markusen (1996), Rauscher (1995) and especially Levinson (1996a)) we shall be relatively brief. We shall draw heavily on the survey by Levinson (1996a). He begins by quoting a number of sources from international organisations such as OECD, international and national industrial associations, US politicians of all parties who all believe that industrial delocation in response to stringent environmental legislation is a major issue and have proposed steps to try to limit its effect. However, he then surveys a wide range of different kinds of evidence which suggest that these concerns are not borne out in practice.

The first kind of evidence reviewed by Levinson (1996a) is surveys, from different countries, of factors which businesses say influence their international location decisions, and in the vast majority of cases environmental regulations are unimportant; there is some evidence that in particular industries, such as chemicals, environmental regulations feature more importantly (Knogden (1979) in a survey of West German firms that invested in developing countries), although these industries were also more sensitive to all cost factors. One possible explanation is an UNCTAD survey of multinational companies which suggested that such companies were concerned with environmental regulations in their home countries rather than their host countries - i.e. they applied the same (tougher) environmental regulations wherever they located.

The second form of evidence is data on trade patterns. Most of these studies use various aggregate indicators of trade to see whether they are influenced by environmental variables. For example, Grossman and Krueger (1993) studied US-Mexico trade patterns for a wide range of industries and showed, as Heckscher-Ohlin would suggest, that the US tends to import from Mexico goods that have relatively low skilled labour and capital content; they also included a variable capturing US pollution abatement costs by industry and showed this had a positive effect on imports from Mexico, but the variable was quantitatively small (and statistically insignificant) so that whatever effect US environmental legislation had in encouraging imports from Mexico was trivial. As Levinson (1996a) notes, a problem with almost all these studies is that they fail to properly control for all the other factors that might influence trade patterns. An exception is a study by Tobey (1990) of trade in 5 products which are pollution intensive; he regresses net exports of these 5 products for a range of countries against 11 factor-endowment variables (the other factors which might explain trade patterns) and a variable which measures the strictness of the countries' environmental policies on an index of 1 to 7. The environmental strictness variable is never significant, but Levinson (1996a) comments that the other variables do not have sensible patterns either so it may be that the data is just not adequate to address the question.

The final set of studies surveyed by Levinson (1996a) are studies of location decisions by US firms across US states, which again consist of survey data and econometric studies of establishment-level decisions. Of the latter, the most comprehensive to date is one by Levinson himself (Levinson (1996b))[12]. This is a study of the locations of new plants (those that appeared in the 1987 quinquennial Census of Manufactures but not the 1982). The use of new plants gets round the problem of sunk costs noted in the last section; another reason for focusing on new plants is that many environmental regulations apply specifically to new plants. A major difficulty with studies of this type is the construction of appropriate measures of environmental stringency and Levinson confronts this issue by using six different measures of stringency - three being various indicators of stringency of legislation, one being number of employees involved in state environmental agencies (to capture stringency of enforcement rather than just what is on the statute book) and two measuring abatement costs. The model follows a standard conditional logit model of plant location in which the probability of a plant in locating in a particular state is related to a whole set of state characteristics - business taxes, employment costs, energy costs, unionisation, infrastructure etc

[12] For example the study by McConnell and Schwab (1990) considers only the motor vehicle industry.

and the environmental stringency variable. Levinson first analysed the full sample of new firms and showed that the location decisions of new plants which were branch plants of large companies were more sensitive to 'manufacturing climate' than new plants in general, which would be consistent with the view that multinational firms are more sensitive to environmental legislation than other firms, and in the rest of his analysis he concentrates on the new branch plants opened by largest 500 multi-plant manufacturing firms. The results of the conditional logit analysis shows that while the environmental stringency variable is always negative, it is significant in only two cases an index of legislative stringency constructed by the Fund for Renewable Energy and the Environment and an index of abatement costs. However even where these variables were significant, an analysis of the effects of an increase in environmental stringency (by one standard deviation) shows that the quantitative impact would be small (the probability of a plant locating in a state would drop by 1.73%). Levinson then studied the location decisions by individual industries (17 different SIC codes) and compared the impact of the environmental stringency variable with the overall pollution abatement costs to see if 'dirty' industries were more likely to be adversely affected by environmental legislation than clean industries and found no significant pattern.

Similar results can be found in analyses of location decisions in other countries; an interesting recent example is a World Bank study (Mani, Pargal and Huq (1997)) of location decisions in India since studies for developing countries have been rare. They apply the same methodology as Levinson to all new large (over Rs 500 million) industrial projects in India in 1994, choosing large projects on the presumption that they would be more footloose. Stringency of environmental regulation is captured by the number of prosecutions in a state under the Air and Water Acts normalised by the number of medium and large size plants in the state. They find that environmental stringency actually has a *positive* effect on location choices, although this is not significant. Restricting attention only to the five most polluting industries confirms the result.

The accumulation of evidence from many different kinds of studies and many different data sets all points to the same conclusion that environmental regulations either have no significant deterrent effect on plant locations, and where there is such an effect it is quantitatively small. In either case there seems to be a discrepancy between the public perception that capital flight caused by environmental legislation is sufficiently serious to warrant policy action and the available evidence. Levinson (1996a, b) summarises the four explanations why the empirical evidence shows such small effects, and whether this suggests the studies have not addressed the real issue.

(i) He first dismisses the standard argument that environmental costs are too small to have an effect - in some industries in the US they can account for up to 15% of costs and this should be large enough to influence location decisions.

(ii) A second possibility is that the various studies do not properly control for the different degrees of 'footlooseness' across types of firms or industries. We noted that Levinson focused on branch plants of large companies and Mani Pargal and Huq concentrated on large projects in both cases because it was believed that these would be more footloose (Levinson had evidence to support this). But these may not really capture footlooseness. A recent study of trade data by van Beers and van den Bergh (1997) splits industries into resource-based and non-resource-based, with the later being presumed to be more footloose, and finds that there is a more significant effect of environmental regulation in footloose industries.

(iii) A third possibility, which relates to discussion of policy in the next section, is that there may be omitted political economy dimensions. Levinson (1996a) refers to his 'cynical interpretation' that it would pay industry and politicians to exaggerate the threat that environmental legislation poses to local employment to justify other forms of assistance given to industry. Frederikson (1997b) has a theoretical model which supports this view (i.e. in a political economy equilibrium tough environmental legislation is offset by other subsidies offered to industry); in a personal communication he says this is being confirmed by empirical modelling for US states. This suggests that studies may not be properly picking up all the state taxes and subsidies offered by state governments.

(iv) Finally there may be offsetting comparative advantage aspects which are not being properly picked up by the variables used to control for other factors affecting location decisions. Almost all of these other factors are cost related. A particular aspect of this argument would relate to agglomeration effects. Returning to the theoretical discussion in the last section, this could be a possible explanation of the discrepancy between rhetoric and evidence; for the theory suggested that with strong agglomeration effects there could be a wide range of differences of environmental policies which would have no effect on location decisions, but then a critical threshold could be reached where a small further difference in policies has a 'catastrophic' effect; the empirical evidence is picking up the former effect while the policy concern is picking up the latter. It is obviously impossible to test this conjecture using conventional econometric modelling. To assess whether in practice inter-sectoral linkages are strong enough to provide this kind of catastrophic effect Venables (1994) used a calibrated model of the world chemical industry, which is one which is always identified as highly polluting and thought to be particularly vulnerable to environmental legislation. The industry is split into two sectors: basic chemicals and other chemicals; and there are four country groups: North America, Far East (Japan, Australia, New Zealand), Europe (EU+EFTA), and Rest of the World (other versions of the model have used different industrial sectors and intra-EU countries). The linkage between the sectors is that basic chemicals contribute 25% and 17% of the gross costs of producing basic chemicals and other chemicals respectively, while the corresponding figures for other chemicals are 2% and 9%. The policy instrument he uses is an energy tax imposed unilaterally by Europe (energy accounts directly for about 14% of the gross costs of basic chemicals and 3.5% of other chemicals). Table 2.1 shows the impacts of different levels of taxes on the number of plants (N) and the unit operating costs of production (C) in the two industries in the three main blocks.

Table 2.1 Impacts of a European Energy Tax on Plants and Costs
(Index Form, No Tax Case 100)

Tax	BASIC CHEMICALS						OTHER CHEMICALS					
%	North America		Far East		Europe		North America		Far East		Europe	
	N	C	N	C	N	C	N	C	N	C	N	C
10	106	100	101	100	83	102	101	100	100	100	94	101
20	113	99	101	100	64	105	102	100	101	100	87	102
30	122	99	102	100	44	107	103	99	101	100	81	104
40	135	98	102	100	18	110	104	99	101	100	73	105
50	145	98	103	100	0	112	105	98	102	100	67	106

What this shows is that when the energy tax rate reaches 50%, this will close down the European basic chemicals industry, with most of the production shifting to North America. Notice that given the moderate strength of the inter-sectoral linkages, there is no catastrophic decline in the industry at a particular threshold (nor was there any hysteresis effect), but the decline in plant numbers does accelerate as the tax rises. To understand the importance of the plant location decisions note that at the 50% tax rate unit costs have risen by 12% in Europe and declined by 2% in North America, so European competitiveness has declined by 14% relative to the U.S. Just under half of this can be accounted for by the effects of the energy tax on costs (both directly and indirectly through the higher costs of intermediate inputs). The rest of the cost increase is accounted for by the fact that suppliers are relocating to North America. Thus agglomeration effects have doubled the impact of the energy tax, which supports the claim that conventional models may understate the impact on competitiveness of environmental policies.

In summary, the evidence to date provides little support for the concerns expressed by environmentalists and policy makers about the threat of delocation posed by stringent environmental legislation. We suggest that more attention to properly capturing the footlooseness of industries, to recognising the political dimensions and hence ensuring that all forms of industry assistance are captured, and allowing for agglomeration effects may help to reconcile the difference between empirical evidence and political rhetoric, although we do not exclude the possibility that popular debate is just blind to proper evidence.

2.4 Policy implications.

From the review of economic analysis in section 2.2 we concluded that if governments act non-cooperatively they may seek to manipulate their environmental policies for strategic trade reasons in markets which are imperfectly competitive, but this need not always take the form of setting too lax environmental policies, nor is it the case that concerns about footloose firms necessarily give stronger incentives for weakening environmental policies than for sectors where firms locations can be taken as fixed, nor is it the case that multinational firms exacerbate switches of production between countries. From the review of the empirical literature we concluded that there was no very strong empirical evidence that firms' location decisions were influenced by environmental policies. From this it might be concluded that policy concerns about delocation and environmental dumping are significantly exaggerated and so there is little need to do anything about it. This is reinforced by the argument that what lies behind the environmental dumping argument is essentially a 'missing instruments' problem. When markets are imperfectly competitive governments may have incentives to engage in strategic trade policies. If such policies are outlawed by trade liberalisation agreements then governments may turn to other policies, such as environmental policies, as proxies. But the same argument would apply to other policies, such as employment protection policies, and following the discussion of the empirical evidence, labour costs are in general more significant than environmental abatement costs and so environmental policy may not be the main focus of strategic behaviour.

For the purpose of this section however, let us suppose that strategic manipulation of environmental policies is an issue to be taken seriously and ask what should be done about it. We shall address this question in its most general context, without explicit reference to delocation or multinationals, since, as we have argued, if there are incentives for environmental dumping they are not confined to delocation and multinationals. Since the problem of strategic competition arises from governments setting their environmental policies non-cooperatively, the first issue is what institutions might induce international co-ordination of domestic environmental policies. Three possibilities might be considered. The first is to allow individual governments to take action against countries who they believe are engaging in environmental dumping, for example by reforming GATT articles to allow the use of countervailing tariffs against environmental dumping. There are many reasons to oppose this approach and we do not pursue it[13]. A second possibility is the use of International Environmental Agreements as proposed for transboundary pollution. However this raises a whole set of other issues we do not wish to explore here. So we will simply assume that there exists some supra-national agency which can be given the power to set national environmental policies, which for concreteness we shall refer to as the "federal government" with national governments being "state governments", but recognising that this covers arrangements which are not formally federal, such as the EU.

The next question is what form of intervention the federal government might make in state governments' environmental policies. A commonly proposed approach is harmonisation of environmental policies supported by environmentalists to prevent a "race-to-the-bottom" and by industrialists to provide a "level playing field"[14]. For the purpose of this section we shall take a strict definition of harmonisation to mean imposing either the same environmental standards or the same equiproportionate tightening of environmental standards.

Harmonisation is neither necessary nor sufficient to ensure the absence of distortions to policy. It is not sufficient because if all countries were identical they would all impose the same environmental policy but that would still differ from either the first-best or the cooperative level of policy. It is not necessary because if countries differ in marginal damage costs or marginal abatement costs then first-best or cooperative environmental policies should differ between countries. Indeed,

[13] These reasons include: it is government not firms that are "dumping"; the difficulty of one government establishing what should be the "right" environmental policies of another government against which to measure "dumping"; if governments are engaging in environmental dumping for strategic trade reasons or at the behest of industrial lobby groups, what reason is there to believe that the countervailing tariff will not be used for similar purposes? For further discussion see Rauscher (1997), Bhagwati (1996).

[14] Bhagwati (1996) critiques various arguments that have been proposed for harmonisation of environmental and labour policies, and Leebron (1996) discusses different senses in which policies might be harmonised.

if countries differ significantly, then harmonisation cannot achieve even a Pareto improvement over the non-cooperative outcome (see Kanbur, Keen and Wijnbergen (1995), Ulph (1997b)). This is illustrated in Fig. 2.5, based on Ulph (1997b) where, assuming just two states, we show the state governments' reaction functions $e_1 = R^1(e_2), e_2 = R^2(e_1)$ as well as their iso-welfare contours. The first key point to note is that we have two versions of this diagram depending on whether environmental policies are strategic substitutes (Fig. 2.5a) or strategic complements (Fig 2.5b); depending on the nature of the model being employed or the nature of the environmental policy instrument being used by governments, either of these is possible. The second point to note is that in both cases we have shown the outcome where there is "environmental dumping" - there is a set of policies in the shaded area which would make both countries better off than in the non-cooperative equilibrium (point N), and these would involve both states having tougher (lower) emission standards than in the non-cooperative equilibrium. The third point to note is that we have assumed a significant asymmetry between the two states (e.g. different damage costs) so that in the non-cooperative equilibrium between state governments emission standards are laxer (higher) in state 1 than in state 2.

If harmonisation involves setting equal emission standards for both states, a point on the 45° line, then as can be seen from the diagram any such point would make state 1 worse off than at point N. The reason is obvious. Harmonisation involves two aspects. It attempts to reduce total emissions, and since the non-cooperative equilibrium involves total emissions which are too high in terms of the total welfare of the two countries such a move will in general make countries better off. But it also involves changing market shares, with the high emissions country losing market share to the low emission country. This harms the high emission country, and if countries are sufficiently different, this second effect outweighs the first. As shown in Ulph (1997b), countries would only have to differ by about 50% in damage costs for harmonisation not to yield a Pareto improvement over the non-cooperative equilibrium. If harmonisation involves either equal absolute or proportionate reductions in emission standards, then it is clear from Figure 2.5 that this form of harmonisation may improve on the non-cooperative equilibrium.

While it is obvious that with sufficient asymmetries between countries the strict form of harmonisation of environmental policies will not work, it is sometimes thought that a policy of minimum environmental standards would be desirable, on the grounds that it would raise environmental standards in countries which fell below the minimum standard, and if other countries choose to respond by also raising their standards it would only be because they were better off by doing so. The argument is that one "ratchets up" environmental standards across nation states. But this argument will also fail to deliver a Pareto improvement over the non-cooperative equilibrium if environmental policies are strategic substitutes as shown in Figure 2.5a. If state 1 is compelled to reduce its emission standard below the level e_1^N but above $\tilde{e}_1$, state 2 would respond by *raising* its emission standard to a point on its reaction function between N and M. This clearly makes the home country worse off. Any minimum standard tougher than (i.e. lower than) $\tilde{e}_1$ would be equivalent to strict harmonisation.

However Figure 2.5b shows that if environmental policies are strategic complements then the 'ratchet effect' works, and any combination of policies on state 2's reaction function between N and M would be a Pareto improvement on the non-cooperative outcome.

Thus neither of the policies frequently discussed - strict harmonisation or minimum standards - may yield improvements over the outcome where state governments are just left to set their own policies, thus contradicting claim 6 in the Introduction. Even if they do yield Pareto improvements over the non-cooperative equilibrium there is no reason in general to believe that they will be Pareto optimal for the two states. The obvious approach is for the federal government just to impose a cooperative solution. That raises the standard question when consideration is given to moving powers from state to federal level of whether the federal government would have enough information to calculate such an equilibrium. Ulph (1997b) assumes that damage costs are private information to state governments, so that any set of environmental policies imposed by the federal government would need to satisfy incentive compatibility constraints. The obvious constraint here is to prevent countries with high damage costs pretending to be low damage cost countries in order to be allowed to set lax environmental policies and hence obtain larger market shares. As Ulph (1997b) showed this can lead to environmental policies in countries which have different damage costs being more similar to each other than would be the case if the federal government had full information. The reason is simply that the need to satisfy the incentive compatibility constraint means that countries with high damage costs have to be rewarded for revealing that information by being allowed to produce more output, and pollution, than would be the case with full information. However this falls short of full harmonisation.

So far in this paper we have assumed that all governments are welfare maximising. But there is another reason why state governments may not implement first-best environmental policies and that is because they are responding to political influence exercised by powerful lobby groups. There is now a small literature applying political economy models of electoral competition or political influence to trade and environment[15]. As with strategic trade arguments, these models can explain why, even in a small country, a government may not implement first-best environmental policies, or pursue free trade, but deviations from first-best could involve either too lax or too tough environmental policies depending on relative strengths of lobby groups. This literature also explains why environmentalists may support protectionist groups. However, this literature does not provide any support for a policy of harmonisation, for two reasons. First, even if it is true that environmental policies in some states are not first-best, that does not provide a reason to co-ordinate reforms of environmental policies; there have to be other reasons, such as those provided by strategic trade literature, for co-ordinating environmental policies. Second, the literature is entirely positive, and does not address the issue of whether or how to limit political influence on environmental policies.

[15] See Hillman and Ursprung (1992, 1994), Frederikson (1997a, 1997b), Rauscher (1997); Ulph (1998) provides an overview.

Johal and Ulph (1998) address this latter question in a model which builds on work of Boyer and Laffont (1996) and extends the model of Ulph (1997b) to include political economy elements. Thus suppose that federal and state governments can be elected to be either Green or Industrial in the sense that a Green government uses a utility function which gives environmental damages a greater weight than in a true welfare function while an Industrial government uses a utility function which gives environmental damage too little weight. Pollution is entirely local, i.e. it affects only the state in which it occurs. There is a key asymmetry of information about damage costs. We shall continue to assume that it is only the state government which knows how damaging pollution is in its state. In particular this information is not available to either the federal government or the voters in the state. Thus, it is only when the state government comes into power that it learns a key parameter of its damage cost function (which determines the level of total and marginal damage cost for any given level of emissions). If environmental policy is to be set at the federal level then the federal government will have to provide incentives for state governments to reveal this information to the federal government. There are two prior constitutional choices that the people in the two states face: whether to delegate the setting of environmental policy to the states (in which case they will act non-cooperatively, so we get environmental dumping) or to the federal government; and whether or not to 'tie governments hands', i.e. to mandate the appropriate government (state or federal) to implement a specific environmental policy, that which maximises expected social welfare. We say expected social welfare because, as already noted, at the time the constitutional choice is made voters will not know the true value of environmental damages, and so if any policies are mandated at the constitutional stage they must be based on the expected value of environmental damages. Note that another key aspect is that at this prior constitutional stage the expected value of damage costs is the *same* for both states, so that any difference between states relates to the *ex post* damage costs, not the *ex ante* damage costs. Thus society essentially has a choice between allowing governments to come into power and learn the true value of environmental damages before setting environmental policies, but recognising that elected governments will pursue objectives which reflect the interests of the party in power, not social welfare; or else mandating governments to pursue policies which maximise expected welfare, but based only on the expected value of damage costs not the actual value of damage costs. Note that in the latter case, because *expected* damage costs are the same in both countries, tying governments hands will mean that environmental policies will be harmonised. Johal and Ulph(1998) show that whether or not it is decided to tie governments hands, it is better to set environmental policy at the federal rather than state level, and that when policy is set at the federal level it is more likely that society will want to tie governments hands; this is essentially because when policy is set at the federal level the gain to having policy set by governments who know the true value of environmental damage costs is reduced by the asymmetry of information between state and federal governments. This might provide some explanation for why as trade has become liberalised and hence the need to co-ordinate domestic environmental policies has increased there have been increasing calls for harmonisation of environmental policies; the explanation provided here is that harmonisation is designed to limit the extent of political influence on federal policy making.

To summarise, in this section we have taken seriously the issue of non-cooperative nation state governments engaging in environmental dumping and asked what could be done to prevent this. We have presumed the existence of some supra-national body ('federal government') to whom powers can be given to implement policies to maximise the joint welfare of nation states. We have shown that widely canvassed policies such as harmonisation or minimum standards may not give Pareto improvements over the non-cooperative outcome depending on the degree of asymmetry between nation states and on whether environmental policies are strategic complements or strategic substitutes. In any case such policies are unlikely to be Pareto efficient, and we then investigated two possible limitations on the federal government implementing the cooperative solution - asymmetric information between state and federal governments, which does not justify policies such as harmonisation, and political influence, which, within a very special model, may provide some justification for harmonisation. However, if one is considering the possibility of restricting what governments do, then it may be more sensible to mandate state governments not to engage in strategic environmental policy in the first place (see Grossman and Maggi (1998) for related discussion in the context of strategic trade policy).

2.5. Conclusions

In this paper we have addressed a number of concerns raised in recent debates on globalisation that faced with the threat of delocation of plants to countries with weaker environmental polcies, a threat that is particulaly associated in the public debate with multinational companies, nation states will be forced to engage in a race-to-the-bottom in terms of environmental policies. To counter this possibility it is often recommended that there should be harmonisation of environmental policies, or at least the imposition of a set of 'minimum standards' for environmental polcies.

We restated the above concerns in terms of six more specific claims. We reviewed some recent economic models and have shown that there is little substance to the fears about a race-to-the-bottom and that even if there were concerns about environmental dumping the usual policy prescriptions cannot be sustained in general. Our review of the empirical literature has also cast doubt on the significance attached in the debate to fears of delocation.

Welfare Functions: One Firm

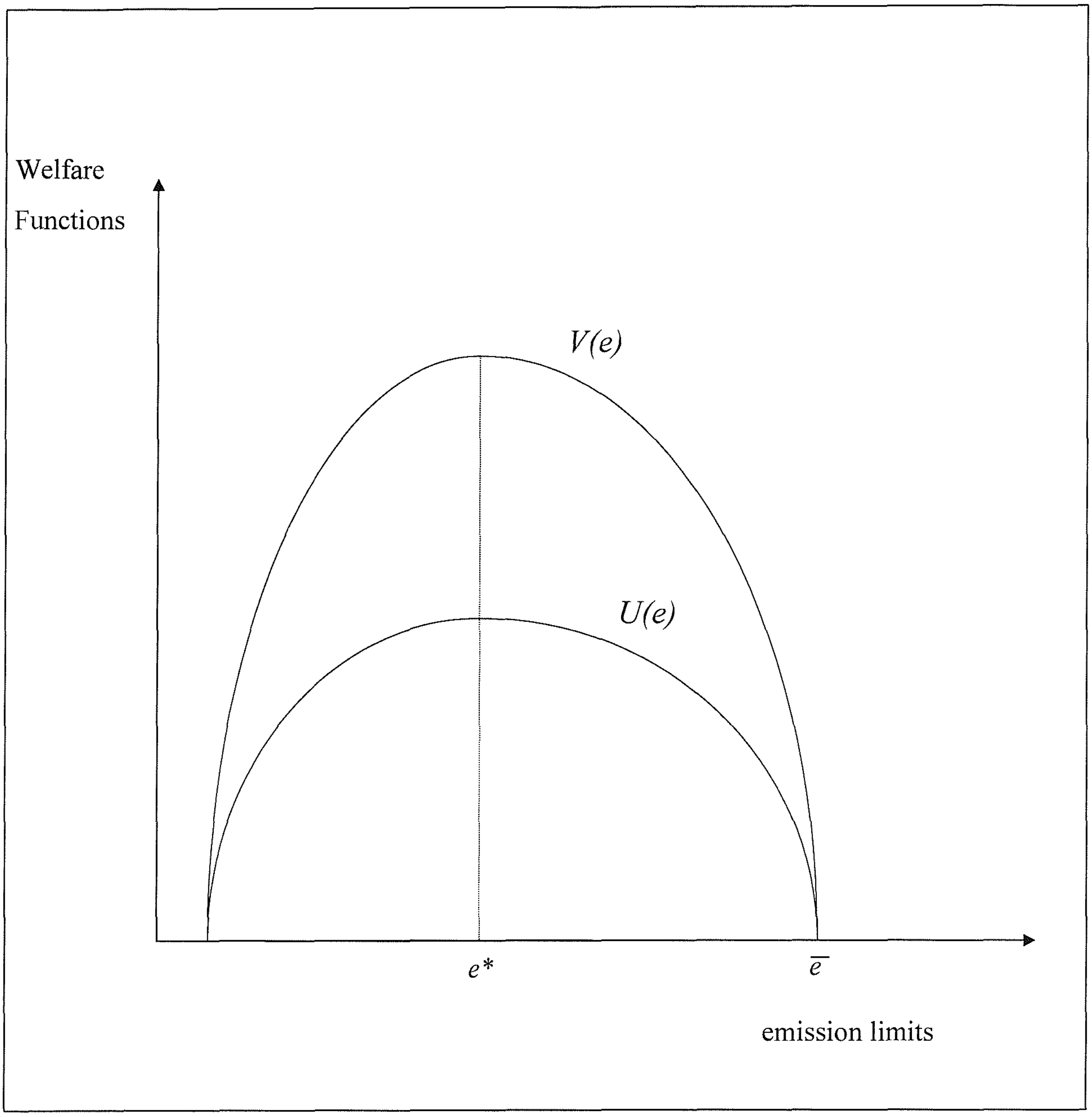

Fig. 2.1

Welfare Functions: More than One Firm

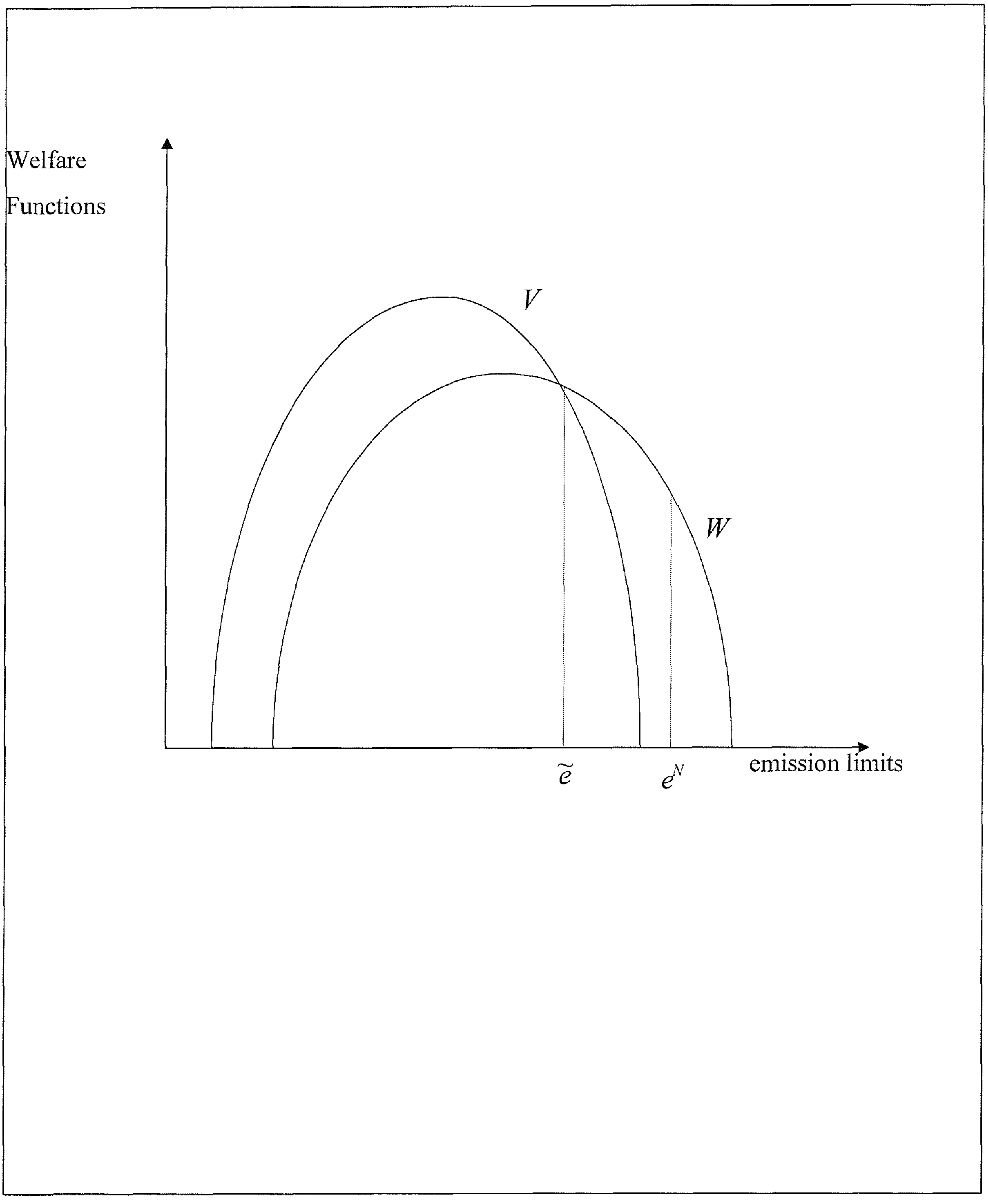

Fig. 2.2

Abatement Falls on Marginal Costs

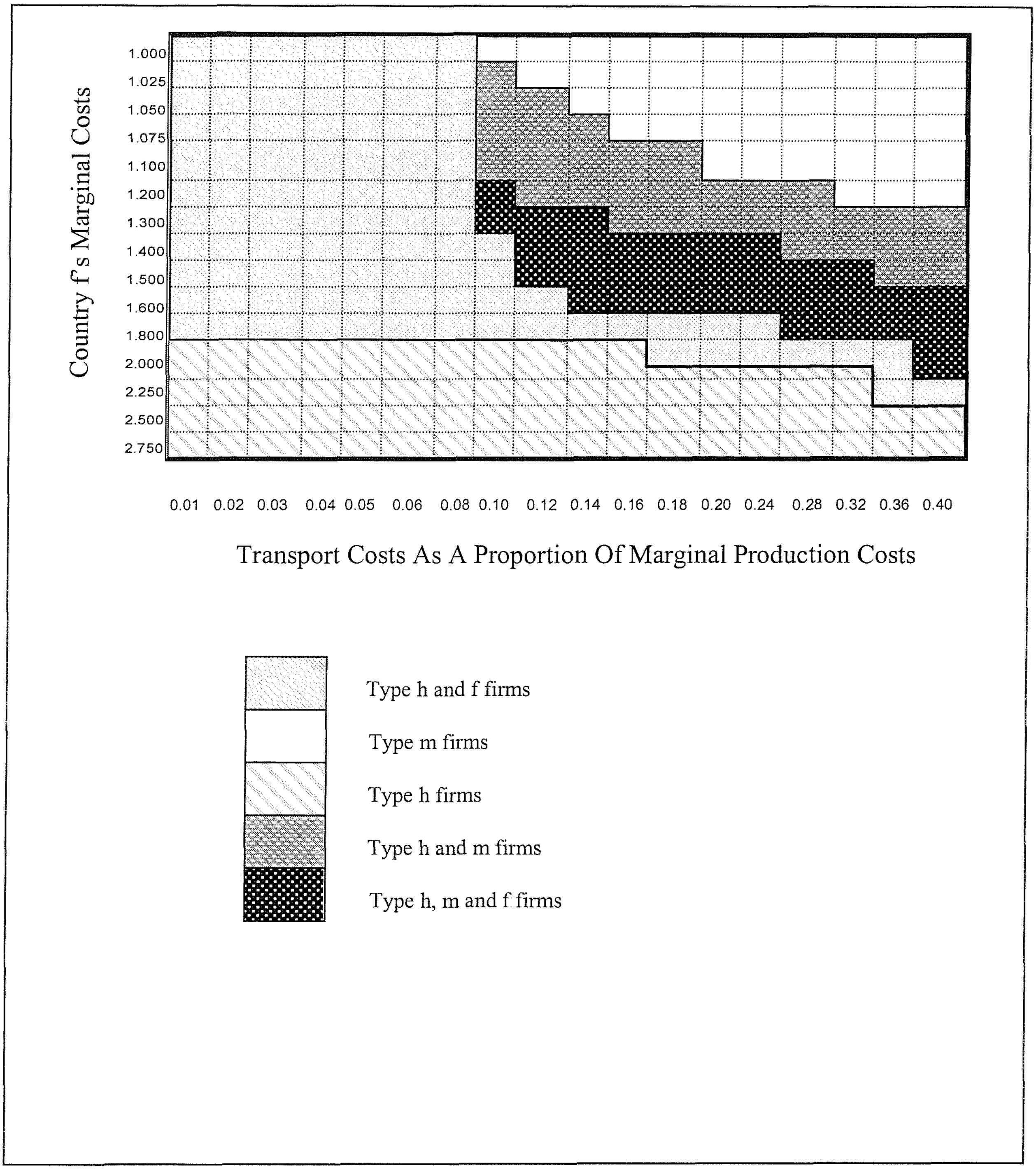

Fig. 2.3

Hysteresis Effects in Environmental Policy

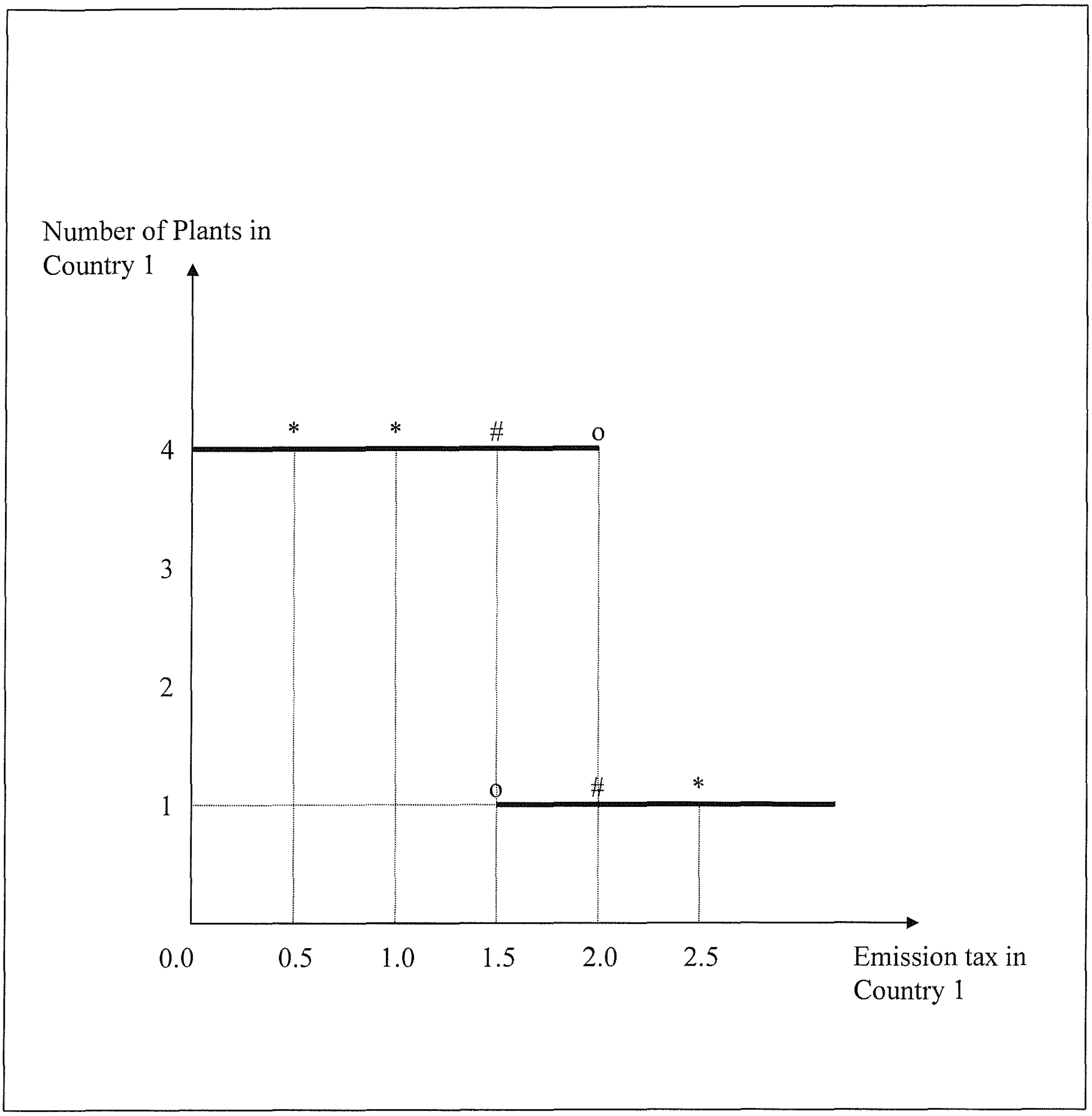

Fig. 2.4

Environmental Policies: Strategic Substitutes/Complements

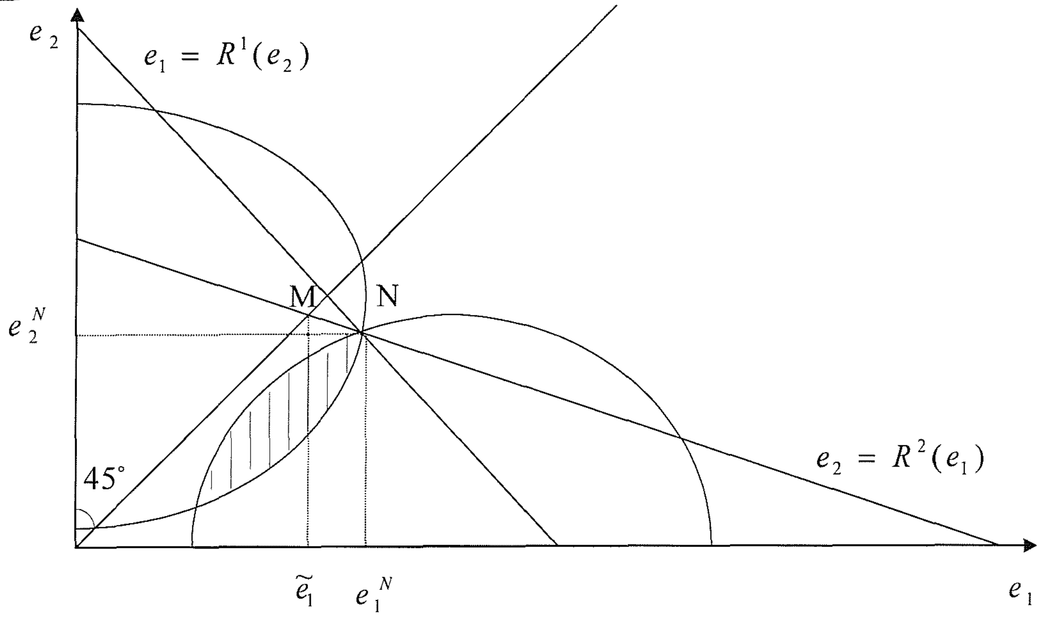

Fig 2.5.a

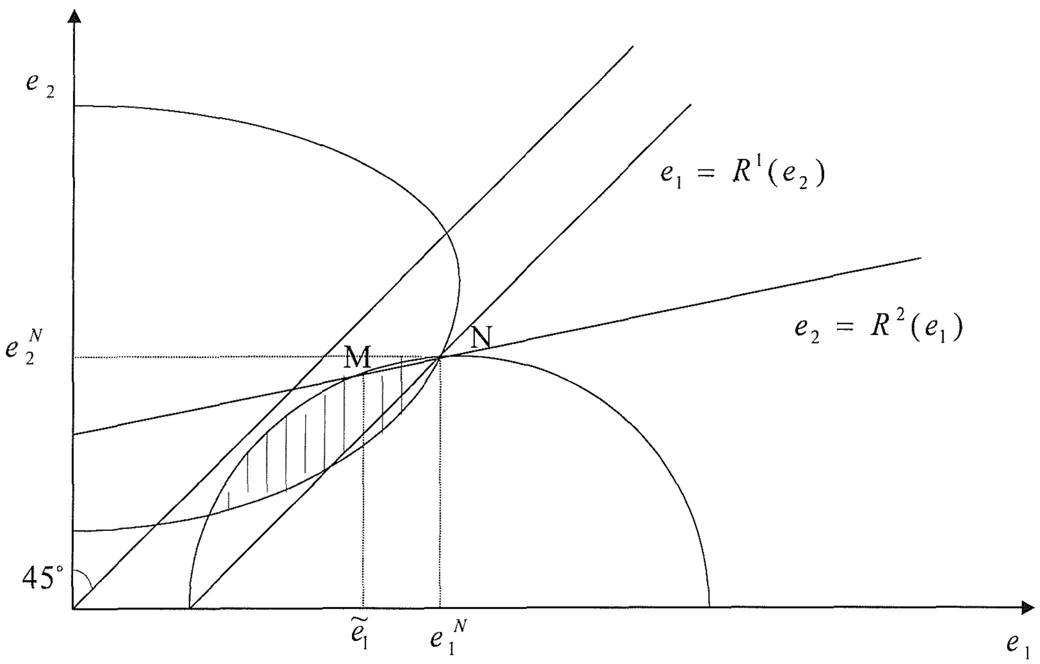

Fig. 2.5.b

Chapter 3

Plant location and strategic environmental policy with inter-sectoral linkages

Abstract.

In recent debates on trade liberalisation the concern has often been expressed that with more competitive international markets, governments will be worried that by setting tougher environmental policies than their trading rivals they will put domestic producers at a competitive disadvantage, and in the extreme case this could lead to firms relocating production in other countries. The response by governments to such concerns will be to weaken environmental policies ('eco-dumping'). In competitive markets such concerns are ill founded, but there is a small literature which has analysed whether governments will indeed have incentives for eco-dumping in the more relevant case of markets where there are significant scale economies; even here there is no presumption that the outcome will involve eco-dumping.

In this paper we extend the analysis of strategic environmental policy and plant location decisions by analysing the location decision of firms in different sectors which are linked through an input-output structure of intermediate production. The reason why we introduce inter-sectoral linkages between firms is that they introduce an additional factor, relative to those already analysed in the literature, in the plant location decision, which is the incentive for firms in different sectors to agglomerate in a single location. This has a number of important effects. First, there is now the possibility of multiple equilibria in location decisions of firms. Following from this there is the possibility of catastrophic effects where a small increase in an environmental tax can trigger the collapse of an industrial base in a country; however there is also the possibility that a country which raises its environmental tax could attract more firms to locate in that country, because of the way the tax affects the incentives for agglomeration. Finally, and again related to the previous effects, there is the possibility of a hysteresis effect where raising an environmental tax in one country can cause firms to relocate to another country, but subsequently lowering that tax will not induce firms to relocate back into the original country.

We consider a simple model with two countries, two industries, an upstream and a downstream sector, and two firms per industry. The analysis proceeds through a

three-stage game: in the first stage the governments of the two countries set their environmental policies; in the second stage the firms in both industries choose how many plants to locate and where; in the third stage firms choose their output levels, with the demand for the upstream firms being determined endogenously by the production decisions of the downstream firms. We assume that there are no limits to production capacity, so that firms do not build more than one plant in any country. However, firms may build plants in different countries because of positive transport costs. Although the model appears very simple, it cannot be solved analytically, so all the conclusions must be drawn from numerical simulations.

3.1. Introduction

In popular discussions of environmental policy and trade one frequently encounters the view that, even if pollution is a purely national concern, environmental policy should not be left to individual nations to determine. The fear is that national governments will be concerned that if they impose tougher environmental policies than their rivals, this will put their domestic producers at a competitive disadvantage; an extreme version of this is the concern that tough environmental policies will encourage a 'flight of capital' to 'pollution havens', which not only imposes economic costs on those economies but may mean that the resulting increase in pollution in the pollution havens significantly offsets the reduction in pollution in the countries which have imposed tougher environmental policies (referred to as 'carbon leakage' in the context of global warming). Such a concern by national governments, it is argued, will lead to competition between governments to relax their environmental policies (sometimes referred to as 'ecological dumping'[1]). The implication is the need for government to co-ordinate in setting their environmental policies, which is frequently interpreted to mean that governments should harmonise their environmental policies.

Economists have been largely sceptical of such views, on both theoretical and empirical grounds (see, for example, Oates and Schwab (1998); Anderson and Blackhurst (1992), Xing and Kolstad (1994) and Low (1992) provide useful overviews). Much of the theoretical analysis is based on competitive models of trade, and a more recent literature based on imperfect competition shows that it is possible to construct models which provide some support for fears that governments might engage in 'ecological dumping', though such results are by no means robust (see Ulph (1994a) for a survey which covers this recent literature). Most of the studies in this recent literature assume immobile capital but there are a number of studies which allow firms to relocate capital, (Markusen (1995),

[1] 'Ecological dumping' can have many meanings (see Rauscher (1994) for a discussion). One interpretation is that governments acting non co-operatively set environmental policies which are laxer than the policies they would set if they acted co-operatively). Another interpretation, which is the one we shall use, is that governments set environmental taxes which are below marginal damage costs.

Markusen, Morey and Olewiler (1993, 1995), Ulph (1994b), Motta and Thisse (1994), Hoel (1997), and for a survey of models with mobile capital, Rauscher (1995)). As in the literature with immobile capital, while 'ecological dumping' is a possible outcome, it is by no means the only outcome.

To understand the factors which affect the environmental taxes which governments acting non co-operatively[2] will set it is useful to consider the simple model by Hoel (1997), in which there is a single firm serving a market consisting of two identical countries. Because there are economies of scale, but no transport costs, the firm will only build one plant, and will locate that plant in the country with the lower environmental tax. The welfare of the country in which the firm does not locate is just the sum of consumer surplus and the share of the firm's profits accruing to domestic residents. For the country in which the firm does locate there is an additional element to welfare: the environmental tax revenue less environmental damages.

Begin with the case where the location of the firm is fixed, so there is no strategic competition between governments. In setting its unilateral optimal environmental tax the government of the country in which the firm is located considers the three elements of social welfare: consumer surplus, profits and environmental tax revenues minus damage costs. If there are no domestic consumers, and all profits flow to domestic residents, then the government will simply set the optimal emission tax equal to marginal damage cost (it has no reason to distort the firm's production decision other than to correct the environmental externality because the firm is choosing output to maximise profits and there are no consumers to worry about). If there are domestic consumers then the government will set its optimal

[2] There are also a number of papers which consider the impact of environmental policy on the location decisions of producers but without any strategic interaction between governments, either because only one government is assumed to vary its environmental policy (Markusen, Morey and Olewiler (1993), Motta and Thisse (1994)) or because governments have predetermined targets for emissions reduction (perhaps because they have signed an international environmental agreement) - see Ulph (1994). These papers highlighted some important points - for example that welfare in a country could be a discontinuous function of environmental taxes set by that country because of discrete changes in location and numbers of plants induced by the tax (Markusen, Morey and Olewiler (1993)). It was also noted that having firms which were not very footlose, so not vulnerable to relocation as a result of domestic environmental policy, need not be to a country's advantage (Motta and Thisse (1994), Ulph (1994)).

emission tax below marginal damage costs because the government is trading off the social cost of environmental damage against the social cost of monopoly. Finally if there are no domestic consumers, but some of the firm's profits flow out of the country, the government will set its environmental tax above marginal damage costs as a proxy for taxing the profits of the foreign shareholders. So even without any strategic considerations, the environmental tax may be above or below marginal damage costs, essentially because the environmental tax may be addressing more than one market failure.

If we now allow the firm's location decision to be endogenous, then Hoel shows that there are three possible cases:
(i) suppose environmental damage costs are relatively low, in the sense that if the country in which the firm is located could set its unilaterally optimal tax it would have higher welfare than the country without the firm. Then the Nash equilibrium involves the two governments competing to have the firm locate in their country; the tax will be set below the unilaterally optimal tax, at a level which makes each country indifferent between having the firm or not having the firm (so the environmental tax revenues just compensate for the pollution damage).
(ii) suppose environmental damages are at a level that if the country in which the firm locates sets its unilaterally optimal tax, it would be worse off than not having the firm with same tax rate. Then a Nash equilibrium has the country with the firm setting its unilaterally optimal tax, while the other country commits to a higher tax, sufficiently high that if the country with the firm set an even higher tax in order to get rid of the firm then it would end up no better off, because the high tax set by the other country would reduce consumer surplus and profits by more than the gain from losing the firm.
(iii) finally suppose that environmental damage costs are sufficiently high that the country which has the firm and sets its unilaterally optimal tax would be worse off than not having the firm at *any* tax rate. Then the Nash equilibrium has each country setting high enough tax rates to close down production by the firm (the 'Nimby' outcome).

Thus introducing strategic competition between governments for the location of a particular plant may cause emission taxes to be above, below or equal to the emission tax a government would set in the absence of any threat of relocation, and, as indicated above the unilaterally optimal emission tax may be above, below or equal to marginal damage costs depending on the distribution of consumers and shareholders across countries.

There are a number of extensions that can be made to the simple Hoel model. Introducing positive transport costs will now mean that the single firm has to decide how many plants to build as well as the question of where to locate them, trading off scale economies against transport costs . Such a model is considered by Markusen Morey and Olewiler (1995), who also derived the possibility of the first and third type of equilibria noted by Hoel. A second extension, which Hoel himself allows for, is to consider more than one firm so that there is now some form of imperfect competition between producers. Even without any possibility of firms relocating this will generate incentives for governments to distort environmental policies in order to manipulate the market outcome; there is now an additional feature influencing whether governments will want to set environmental taxes above or below marginal damage costs, namely the form of market competition, e.g. Cournot or Bertrand (see Barret (1994), Kennedy (1994), Ulph (1994, 1997c)). Markusen (1995) combines both these extensions in a model which allows the endogenous determination of the number of national and multinational firms.

There is a third extension which will be the topic of this paper. As already noted, the previous literature deals with a single industry. In this paper we address the question of environmental policy and plant location in a model where there are linkages between different industries of the economy, reflecting the input-output structure of the economy. The rationale is to allow the analysis of incentives for agglomeration of industry. In the original 'economic geography' models of Krugman (1991a, b) the driving force for agglomeration was labour mobility - industries in large markets could afford to pay higher wages because they did not need to incur transport costs, but these high wages would attract workers who form a large market, so agglomeration can become reinforcing. In this paper we follow Venables (1993) in having capital rather than labour being mobile, but the same reinforcing incentives for agglomeration can arise: if downstream industries are concentrated in a particular location, that provides an incentive for upstream industries to locate close to them; but that lowers the costs of intermediate inputs providing an incentive for the downstream firms to locate close to their suppliers. In other words, the market access reasons for upstream firms to locate close to downstream firms in turn provides a cost advantage reason for downstream firms to locate close to suppliers.

The potential implications for environmental policy are as follows. In the papers cited above location decisions were driven mainly by cost considerations; market size played a role, but that was exogenous. In the kind of model discussed here demand, at least for intermediate goods, depends on the location choices of firms, and so is endogenous. This has a number of consequences. First, environmental taxes imposed on one industry can have consequences on other industries even if they are not subject directly to environmental policy; for if emission taxes cause firms in one industry to relocate they can take with them firms in closely related industries. Thus there are going to be both cost and demand considerations to be taken into account in assessing the impact of environmental policy. Second, the literature cited above has already noted the fact there can be discontinuous responses to environmental policy; increases in emission taxes over a reasonable range will have little impact on output or welfare, but then a critical threshold can be reached at which firms relocate and a small change in emission tax can trigger a large change in production and welfare. When we add agglomeration this is going to reinforce this feature; quite large differences in environmental taxes between countries could have little effect on location because of the benefits of agglomeration; but if the differences trigger a relocation, the effect could be much larger than in a single industry model, with an entire industrial base being wiped out. However , we also show that we can generate cases where one country raising its environmental tax can cause the number of plants located in that country to increase, because the tax changes the incentives for agglomeration.

A further implication of agglomeration stressed in the economic geography literature is that it raises the possibility of multiple equilibria. If downstream firms happen to locate in country A, then the forces for agglomeration will bring upstream firms in as well, reinforcing the choice of country A by downstream firms. But if the downstream firms had happened to locate in country B the same arguments would apply. This is more than just a theoretical curiosum and can have important policy implications. As we shall show, there can be a hysteresis characteristic to environmental policy. At low levels of environmental taxes there may be a unique equilibrium in which firms agglomerate in country A, say. As country A raises its tax a second equilibrium may emerge in which all plants agglomerate in another country, but no individual firm has any incentive to defect from the original equilibrium. Beyond a certain critical level this may not induce firms to locate back in country A because no firm has any incentive to deviate from the locational equilibrium in another country.

The ideas in this paper are closely related to those of Venables (1993, 1994)[3] where the latter paper has an environmental application. Our paper differs from Venables (1994) in a number of respects:

(a) in our paper we derive explicitly the intermediate demand function for the output of one sector based on the production decisions of other sectors; this is handled in a more ad hoc fashion in Venables (1994);

(b) because this generates a considerable complexity in our model, we have had to simplify the model by assuming just two countries and two sectors (one upstream and one downstream), while Venables (1994) allows for arbitrary numbers of countries and sectors;

(c) while we fix the number of firms in each sector to two, Venables (1994) allows the number of firms to be determined endogenously by free entry;

(d) we allow for the determination of a non-co-operative equilibrium in governments environmental policies, while Venables (1994) considers only unilateral variations in policy making by a single government. [4]

---

[3] The close relationship between this paper and Venables (1994) is not surprising as they were both written as part of the same research project.

[4] There would be no difficulty in principle in introducing a strategic game between governments into the Venables (1994) model.

In the next section we set out our basic model. Despite the extremely restricted nature of the model, it is not possible to derive results analytically, for reasons we discuss in the next section. So in section 3.3 we set out the results of some numerical simulations of our model. Section 3.4 offers conclusions and directions for further work.

3.2. The Model.

There are two countries, or markets, denoted m=1,2. There are two industries, an upstream industry and a downstream industry. The downstream industry faces an inverse demand curve in market m=1,2 $p_m = A_m - b_m S_m$ where S_m is total sales in market m. The upstream industry produces a good which is an intermediate input to production of the downstream industry and we choose units so that one unit of the downstream good requires one unit of the upstream good in its production. In each industry there are 2 firms, denoted f=U1, U2, D1, D2 in an obvious notation. Firms U_m and D_m are wholly owned by shareholders located in country m, m=1,2.

Production in each industry is modelled very simply by assuming that firms face constant unit costs of production and a fixed cost of locating a plant in a particular market. For upstream firm f= U1, U2 we denote the unit cost of production of a plant located in country m by C_{fm}, while for a downstream firm, f=D1, D2, given our assumption on the input-output structure, the unit cost of production in market m is denoted $C_{fm}+q_m$ where q_m is the price of the upstream good in market m, which is to be determined endogenously. Each firm f=U1, U2, D1, D2 has to decide how many plants to build and where; firm f faces a fixed cost K_{fm} of locating a plant in market m, and given our assumptions on production costs will obviously never locate more than one plant in any market; so each firm has four possible location strategies denoted by $L_f=[l_{f1}, l_{f2}]$ where $l_{fm}=0,1$ is an index denoting whether or not firm f has built a plant in market m. As indicated in the previous section, one factor that will influence a firm's plant location will be market access. We shall refer to the costs of being located at a distance from customers as transport costs, but these stand more widely for costs of lack of proximity – warehousing and handling, loss of cargo in transit (e.g. perishable goods), lack of information about precise customer needs etc. For the upstream firms we assume that there is a constant per unit transport cost of τ between the two countries while for the downstream firms transport costs depend on the volume shipped, $0.5\theta x^2$, where x is the quantity shipped and θ is a parameter controlling the level of transport costs. The reason for the different treatment of transport costs in the two industries will be explained shortly.

Turning to pollution, we assume that there are fixed emission/output coefficients E_D, E_U in the downstream and upstream industries respectively (so there is no abatement technology). Total emission in a country will thus depend on the total amount of production in the two industries carried out in that country. We assume that pollution damage is purely domestic, i.e. there is no transboundary pollution, and country m has per unit damage costs d_m. To control pollution the government of country m uses an emission tax t_m. Finally, we have assumed that each firm is wholly owned by shareholders located in a particular country. If all the profits of that firm flowed back to the shareholders of the firm, then the government of the country in which the shareholders were located would have no direct reason for being concerned about where that firm located its plant (we do not allow for any impact of location choice on costs of unemployment, for example, although this is probably a major reason why politicians worry about such matters). To introduce a reason for governments to care about plant location we assume that the government in country m imposes a profit tax at a rate T_m on all profits earned in that country. Profits liable for tax are net of any emission tax. Welfare in each country is consumers surplus from consumption of the downstream good, plus net of tax profits of the firms owned by shareholders in that country plus emission tax revenues and profit tax revenues minus pollution damage costs.

The move structure is that in stage 1 governments set their emission taxes and profit taxes; in stage 2 producers choose their locations; and in stage 3 producers choose their outputs. In general we look for a sub-game perfect Nash equilibrium of this three stage game, but with two caveats. First profits taxes are not chosen strategically by governments, but are just set exogenously; in principle we could extend the model to allow for strategic choice of profits tax but we do not do that here. Second, for some simulations we shall also just consider the effect of exogenously specified emission taxes rather than looking for a Nash equilibrium in emission taxes, though in other simulations we do calculate such Nash equilibrium. The choice of the particular move structure here needs some comment. In particular, since we are allowing firms to consider their location choice of plants in stage 2 we are clearly modelling a long-run equilibrium, and it can be questioned whether it is reasonable to assume that governments can commit themselves to a fixed level of emission taxes for such a long period. But if the concern is that governments will set emission taxes strategically to take account of how firms will respond in terms of location choice then this move structure is the natural way to capture such a concern. For an analysis of how a different move structure could affect the results in models of plant location see Ulph (1995).

Although the model is, in many respects, extremely simple, it turns out to be difficult to obtain results analytically. To get some feel for what is involved we shall sketch out a bit more of the details of solving the three-stage game.

STAGE 3 GAME

At this stage we know the pollution taxes set by each government and the location choices of each firm. The way we solve this stage is to assume that each firm has built a plant in each country; if a firm does not actually have a plant in a particular country we just assign that plant a very high "dummy" unit operating cost, high enough to guarantee that in equilibrium that plant would never be used. In general then, denote by γ_{fm} the unit operating cost of firm f = U1, U2, D1, D2 in market m, where, for downstream firms this excludes the costs of intermediate inputs from the upstream firms; if firm f has a plant in country m, γ_{fm} will just equal C_{fm} plus the appropriate industry and pollution tax (i.e. allowing for the emission/output ratio for the appropriate industry and the emission tax for the appropriate government); if firm f has no plant in country m then γ_{fm} is set to a large number.

To determine the equilibrium in this stage we need to determine not only the usual Nash (Cournot) equilibrium between different firms in the same industry, but also how the equilibria in the two industries are related. We begin with the downstream industry.

Downstream Industry

The firms in the downstream industry take as given the prices q_m for which the intermediate good is sold in market m=1,2. Let x_{fnm} denote the sales by downstream firm f=D1, D2 from its plant located in country n to the market m. Then each downstream firm takes as given the sales by its rival and chooses its own vector of sales to maximise profits. If we consider, for example, the sales by downstream firm 1 into market 1 from its plants in countries 1 and 2, i.e. x_{111} and x_{121} respectively, we would obtain the following pair of first-order conditions:

$$\begin{aligned} &A_1 - b_1(x_{211} + x_{221}) - 2b_1(x_{111} + x_{121}) - (\gamma_{11} + q_1) \leq 0 && x_{111} \geq 0 \\ &A_1 - b_1(x_{211} + x_{221}) - 2b_1(x_{111} + x_{121}) - (\gamma_{12} + q_2) - \theta x_{121} \leq 0 && x_{121} \geq 0 \end{aligned} \qquad (1)$$

This is just the usual condition that marginal revenue in the appropriate market equals marginal cost of producing for that market, where in the second case marginal cost includes marginal transport costs. There are three other pairs of such first-order conditions, giving eight first-order conditions to determine out eight variables x_{fm}. Note that it is crucial that we allow for the possibility that some of these variables will be zero. Indeed it is straightforward to check that we could never have more than six of these variables strictly positive since our simple structure of production rules out the possibility of cross-hauling by each downstream firm. But the fact that some firms will not have plants in some markets, or if they do they may be priced out of the market will also mean that in many cases many of the x variables will be zero.

Using these first order-conditions we can add up the total amount of downstream production taking place in each country and this will give us the demand for the intermediate good in that country. Note two points. First, in general it is clear from (1) that the demand for intermediate goods in each country will depend on the prices for the intermediate good in both countries, but given the structure of the first order conditions these are going to be linear demand; i.e. we are going to have a pair of demand functions of the form:

$$Y_m = F_m + G_m \cdot q_1 + H_m \cdot q_2 \qquad \text{m} = 1,\ 2 \tag{2}$$

Second, from what we have just said, the nature of this pair of demand functions will depend on the precise configuration of positive and zero values for x_{fnm} emerging from the first-order conditions of the form in (1). It turns out that there are 144 possible configurations of pairs of demand functions for the intermediate good, and we have explicitly calculated all of them.

Upstream Industry

For the upstream industry, we assume that the firms in this industry take as given the demand functions for the intermediate good in the two countries given by (2). We invert these to obtain a pair of inverse demand functions, giving the price of the intermediate good in market m as a linear function of the total sales in the two markets:

$$q_m = f_m + g_m \cdot Y_1 + h_m \cdot Y_2 \qquad \text{m} = 1,\ 2 \tag{3}$$

Given the simpler structure of transport costs in the intermediate good industry, firm f=U1, U2 will decide which plant it will use to serve a particular market by just comparing the unit production cost of the plant it has located in that market with the unit production cost plus transport cost of serving that market from its plant located in the other country (recall that we are assuming each firm has a plant in each country, albeit the plant may be a "dummy" plant). Thus we define firm f's costs of serving market m by:

$$\Gamma_{fm} = \min\{\gamma_{fm}, \gamma_{fn} + \tau\} \qquad f = U1, U2,\ \ m = 1,2\ \ n = 1,2\ \ n \neq m$$

Defining y_{fm} as the sales by firm f into market m, in a Nash equilibrium each firm will take as given the sales of its rival and choose its own sales to maximise profits, using (3) to determine its sales revenue. There will be four first-order conditions; as an example the one for firm 1 in market 1 is:

$$f_1 + g_1 y_{21} + g_2 y_{12} + h_1(y_{12} + y_{22}) + 2g_1 y_{11} - \Gamma_{11} \leq 0 \qquad y_{11} \geq 0 \qquad (4)$$

Again we need to take careful account of the possibility that some or all of the sales quantities could be zero; this requires checking 16 possible configurations of sales by intermediate firms. Once we have determined the equilibrium pattern of sales we insert these in (3) to yield the equilibrium prices for intermediate goods. We then use the rule for deciding which plant each firm will use to serve each market to determine the production by each intermediate firm from each of its plants.

Overall Equilibrium for Stage 3.

We have defined the equilibria for the downstream firms and upstream firms; we now say how these are linked. What we do is to consider in turn each of the possible 144 configurations of intermediate demand specified in (2). For each configuration of intermediate demand we solve for the equilibrium in the upstream industry using these demand functions, and this determines a pair of equilibrium prices for the intermediate good in each country. We then use these intermediate goods prices to determine the equilibrium sales in the downstream market, using first-order conditions of the type shown in (1). If that yields the same configuration of intermediate demand that we started with, then we have an overall equilibrium. But if the configuration of intermediate demand is different from the one we started with, then that configuration is not consistent with an overall equilibrium.

There are two final points we make about this equilibrium. First there may be multiple equilibria. This can occur both at the level of the equilibrium for the upstream industry where there are 16 possible configurations and at the level of the overall equilibrium where there are 144 possible demand configurations. So we need to use an equilibrium selection criterion. What we do is to calculate the sum of the profits of the relevant firms to get an overall industry level of profits; we also calculate the product of firms' profits, which is a Nash bargaining solution. These are normalised to lie between 0 and 1, and we then take a weighted average of the two criteria. In the simulations we report we use equal weights, but it turns out that the equilibrium selected is not very sensitive to the weights used. Second, we return to the question of the different treatments of transport costs in the two industries. We began by making the same assumption in both industries, namely constant unit transport costs. But this meant that the demand functions for intermediate goods could be discontinuous, since at critical thresholds of costs downstream firms would switch production completely from one plant to the other. In some cases this caused problems in finding an equilibrium. By introducing quadratic transport costs in the downstream industry we made the demand functions continuous (although not necessarily differentiable) and this ensures that there always exists an equilibrium for the stage 3 game.

STAGE 2 GAME.

Each of the four firms has four possible plant location strategies, giving 256 possible plant configurations. For any particular plant configuration we can solve for an equilibrium of the stage 3 game. This allows us to compute the profits the firms would earn in each plant, after paying emission taxes. We then subtract the fixed costs of plant location to arrive at profits liable for profits tax, and then calculate the profits net of profit tax for each plant owned by each firm. If a firm has more than one plant we calculate total profits for each firm net of all taxes. We then search across the 256 configurations of plants to find a Nash equilibrium of plant locations.

Again there is the possibility of multiple equilibria. Indeed in the simulations we have run this is always the case because there is always a trivial Nash equilibrium in which no firm builds any plants (the reason is that if three firms do not build a plant then if the remaining firm is an upstream firm it will have no market for its product and so will not want to sink any costs of constructing a plant while if it is a downstream firm it will have no source of supply of its intermediate input, and so again will not wish to sink costs of constructing plant). Obviously one could choose parameter values or tax rates for which the trivial Nash equilibrium is the only equilibrium, but for parameter values of interest there will be at least one other non-trivial equilibrium. So again we need to employ the equilibrium selection criterion sketched above. But for purposes of later analysis we keep track of these multiple equilibria.

STAGE 1 GAME.

As indicated earlier we operate Stage 1 in two different modes. One is where we just specify exogenously the emission taxes set by the two governments. The second is where we calculate a Nash equilibrium in tax rates. There is an obvious difficulty with trying to calculate a Nash equilibrium in tax rates. As we noted in the introduction, given the discrete choices being made at subsequent stages, government welfare is a discontinuous function of emission taxes, so we cannot just write down a pair of first-order conditions and try to find a solution for these. Thus we have to search numerically for a Nash equilibrium. This leads to the further problem that tax rates are a continuous variable, and we are only able to evaluate a finite number of discrete values for tax rates. This difficulty is compounded by the large number of evaluations we have to carry out in stages 2 and 3. So we are relatively restricted in the number of tax strategies we can evaluate. If governments were not acting strategically they would set taxes equal to marginal damage costs, so we chose tax rates typically in a range from 40% to 120% of marginal damage costs, in steps of 5%. The precise tax rates used vary across the simulations since we have tried to ensure that the Nash equilibrium in tax rates is not affected by the limited range of values we have selected; for example, if the equilibrium was at the lower range of an initial set of tax rates we would use lower tax rates to check if government wished to cut taxes further.

This completes our description of the model, in the next section we report the results of the simulations we have run.

3.3. Simulation results.

The simulation results we report in this paper are based on purely fictitious parameter values; given the special assumptions of the model it did not seem sensible to attempt to calibrate the model to real world data[5]. We shall report the simulation results in three stages. First, we shall ignore environmental taxation and report some results designed to show how this model works, and in particular the factors which affect incentives for agglomeration. Second, we shall report results from simulations using exogenously specified emission taxes for the two governments. Finally we shall report results of simulations where we calculate the Nash equilibrium of tax rates by governments.

3.3.1 Base Case

The base case is completely symmetric in its treatment of countries and firms in each industry. Demand for the downstream good in each country is characterised by parameters A_m=40, b_m=1. Firms in both upstream and downstream face unit costs of production C_{fm}=1 no matter which country they locate in (these unit costs in downstream firms do not include costs of intermediate inputs; recall that we assume that a unit of the downstream good requires a unit of the upstream good as input). In each industry fixed costs of setting up a plant are K_{fm}=10 regardless of the country in which the plant is established. Unit transport costs for intermediate goods are τ=1, while for final goods the parameter determining the level of transport costs is θ=0.1. Profits tax is set at 50%. When we select between different equilibria in stages 2 and 3 the selection criterion is a weighted average of the sum of firm profits and the product of firm profits; these are given equal weight.

[5] The model by Venables (1994) has been calibrated to data from the chemical industry.

In the absence of any pollution or emission taxes there are two distinct equilibria. One involves locating all production in both industries in one country; the other involves the upstream firms locating a plant in each country and the downstream firms locating a single plant in their domestic countries, and, because of symmetry, there is no trade in either industry. This latter equilibrium shows that when there is no agglomeration of upstream firms, so downstream firms have no reason to prefer one country to another, then they prefer to locate in different countries to obtain some degree of local monopoly power. We say that these are distinct equilibria; there are actually four equilibria; for there are two equilibria corresponding to all production being located in a single country; and another equilibrium of the second type where the downstream firms locate a single plant not in their domestic countries but in the rival country. The selected equilibrium is the one with all production located in a single country, and this dominates the other type of equilibrium essentially because it economises on the costs of the intermediate firms having to build an extra plant.

In the selected equilibrium each plant produces 16.62 units of output. 16.88 units of final good output are sold in the country in which the plants are located at a price of 23.11 while 16.34 units are sold in the other country at a price of 23.66 (the price difference is less then the marginal transport costs). The intermediate good is sold at a price of 13.67. Net of tax profits in each upstream firm are 100.24 and in each downstream firm 65.72. So profits tax revenue of 331.19 accrues to the country in which production is located. Consumer surplus is 142.62 in the country in which production is located and 133.56 in the other country (lower because of the slightly lower final consumption). Welfare (excluding any allowance for environmental damages) is 640.48 in the country which has all the production and 299.53 in the other country.

3.3.2 Variations of the Base Case.

We now present some simulations which vary some of the parameters of the model to see how this affects the location of plants. All of these results still exclude any environmental effects or environmental policy and are designed simply to illustrate what drives decisions on plant locations.

3.3.2.1 Variations in Transport Costs for Upstream Firms.

In the base case we set unit transport costs in upstream firms equal to 1.0 and the resulting equilibria were either complete agglomeration (all firms locating plant in a single country) or the upstream firms locating a plant in each country with the downstream firms each setting up a single plant but in a different country from each other; the equilibrium with agglomeration was the selected equilibrium. We varied transport costs in the upstream industry to take the value 0.05, 0.1, 0.5 and 2.5. In all cases the equilibrium with agglomeration remained an equilibrium and indeed remained the selected equilibrium. This is perhaps not very surprising since in this equilibrium there is no trade by upstream firms and so varying the transport costs for upstream firms will not affect the profitability of this equilibrium. However, the changes do affect what the other equilibria might be. When transport costs are below 1.0, then the alternative equilibrium is also one in which firms build only one plant, but now they do not locate in the same country; rather each upstream firm locates its plant in a different country from that of its rival, and similarly for the downstream industry (this means that there are four such equilibria, all identical except for the name of the firm in each country). Thus with low transport costs it never pays to proliferate plants, and it also reduces the advantage of agglomeration. When upstream transport costs equal 1, then, these four alternative equilibria disappear, and, as just noted the alternative equilibrium involves the upstream firms setting up plants in both countries with downstream firms setting up a single plant, but in different countries. When transport costs for upstream firms rise to 2.5 then another equilibrium emerges, one in which the downstream firms set up plants in both countries, but the intermediate firms set up only a single plant, but again in different countries. This is clearly an alternative way in which trade in intermediate goods can be avoided.

3.3.2.2.Variations in Transport Costs For Downstream Firms.

We now vary the transport cost parameter for downstream firms. For the base case we set $\theta = 0.1$. We now set $\theta = 0.05, 0.2, 0.5, 1.0$. For $\theta = 0.05$ we get the same result as the base case: agglomeration is the chosen equilibrium but there is another equilibrium in which the upstream firms locate plants in both countries and the downstream firms each have a single plant but located in a different country. When we raise the transport costs above the base case to $\theta = 0.2$, agglomeration remains the selected equilibrium but now the alternative equilibrium is one in which the downstream firms have plants in both countries and the upstream firms each locate a single plant in a different country, and export a small amount to the other country. When we raise the transport costs in the final goods industry to 0.5 then this latter equilibrium becomes the unique equilibrium, while when we raise the transport costs even further to 1.0 then the unique equilibrium becomes one in which both downstream and upstream firms locate plants in both countries.

3.3.2.3 Variations in Importance of Intermediate Goods

The final experiment was to assess how sensitive the results of the base case would be to the importance of the linkage between the two sectors. One way of capturing this would have been to vary the size of the input/output coefficient. It turned out to be easier to do this by varying the level of costs (production and transport costs) of intermediate goods. One experiment we did was to reduce production costs of intermediate goods of 0.01, 0.1 and 0.5. In the latter two cases, the agglomerated equilibrium remained the selected equilibrium, but the second equilibrium changed. Now it was an equilibrium in which both in the upstream and downstream markets each firm built a single plant and located it in a different country from its rival, selling about half their output in both markets. When production costs fell to 0.01, this became the only equilibrium. Clearly this is an equilibrium in which there are no gains from agglomeration. A second experiment was to raise the production costs of intermediate goods from 1.0 to 2.0, combined with transport costs for intermediate goods of 0.2, 2.0 and 4.0. In all cases the agglomeration equilibrium remained the selected equilibrium. In the first case there was another equilibrium, which was the same as the other equilibrium we have just described for the case of low production costs. In the latter two cases the other equilibria were like those in section 3.3.2.1 where we had high transport costs for the intermediate good; i.e. either the upstream firms set up plants in both countries and the downstream firms set up a single plant in different countries (as in the base case) or it was the other way round- the downstream firms set up plant in both countries and the upstream firms locate a single plant in different countries.

3.3.2.4 Implications

What have we learned from these experiments? The first point is that agglomeration remains an equilibrium and indeed the chosen equilibrium over a reasonably wide range of parameters, but the variations affect what the alternative equilibria are; when we push these variations further, agglomeration ceases to be an equilibrium and one of the alternatives becomes the selected, perhaps unique, equilibrium. Second, as one might expect, when transport costs rise, the alternative equilibrium is one in which firms in one of the industries build plants in both countries to get closer to their market. On the other hand when transport costs are very low, and particularly when the links between sectors is of less importance, then the alternative equilibrium involves only a single plant, but firms locate in different countries. In short agglomeration tends to be associated with a middle range of transport costs. If transport costs become very high then the importance of being close to customers means firms proliferate plants. When transport costs are trivial and intersectoral linkages low, then firms do not worry about being close to their customers and build plants in a way which gives them some degree of local monopoly power. In the intermediate range agglomeration is attractive- transport costs are low enough to only warrant building a single plant, but still high enough for intermediate firms to be close to their customers and for downstream firms to get cost advantage from being close to their suppliers.

3.3.3 Exogenously Specified Emission Taxes.

In this section we report the results of simulations where we introduce emissions related to production of upstream or downstream products and also introduce emission taxes by governments to control such emissions. In these experiments emissions are related solely to production in the upstream firms (one unit of production in the upstream industry generates one unit of pollution). We impose an asymmetry on production costs in both upstream and downstream firms by setting unit production costs in both industries in country 2 to 2.0 while keeping them at 1.0 in country 1. This ensures that in the absence of any environmental tax all upstream production gets concentrated in country 1. We then trace the effects of imposing an emission tax only in country 1 at levels 0.5, 1.0, 1.5, 2.0 and 2.5. At this stage we do not explicitly introduce environmental damages so any welfare measures reported do not include the impact of environmental damage. We report two experiments which illustrate some interesting results.

3.3.3.1 Experiment A.

In this case we keep the remaining parameters at the values we set in the base case. The results we get are as follows. In the absence of any emission tax all production in both upstream and downstream firms is located in country 1, the low cost country, and this is a unique equilibrium. When the emission tax rises to 1.5 this remains an equilibrium, and indeed is the selected equilibrium, but another equilibrium emerges in which the intermediate firms both switch to country 2, and one of the downstream switches with them, but the other stays in country 1. At an emission tax of 2.0 the same two equilibria remain, though the selection criterion now selects the other equilibrium; when emission tax reaches 2.5 this alternative equilibrium is the sole equilibrium (Fig. 3.1 shows the number of plants in country 1 for different tax rates).

This is an illustration of the possibility of a hysteresis effect in environmental policy. To illustrate this, we need to give a dynamic interpretation of the model and the choice of equilibrium, rather than assuming that in a situation of multiple equilibria the equilibrium selection criterion is applied. Suppose then that country 1 raises its emission tax from zero. For tax rates less than or equal to 1 the agglomerated equilibrium with all plants in country 1 is unique. As the tax rises from 1 to 2 this remains an equilibrium, though not unique, so no individual firm would wish to switch its location from country 1 to country 2. When the tax exceeds 2, the original equilibrium is no longer sustainable and three plants switch from country 1 to country 2. However, if country 1 now cuts its tax again below 2, the second equilibrium remains an equilibrium, so no individual firm located in country 2 would wish to switch back to country 1, and it would not be until the tax rate falls below 1 that the second equilibrium becomes unsustainable and firms located in country 2 switch back to country 1. So for tax rates in the range 1 to 2 the equilibrium that prevails will depend on the path of tax rates prior to the emission tax moving into this range.

3.3.3.2 Experiment B.

In this experiment the structure was very similar to the above, except that we assumed that transport costs for the downstream firms are characterised by the parameter $\theta = 0.5$ not 0.1 as in the previous experiment. In the case where there are no taxes the unique equilibrium consists of having all the upstream firms locating a single plant in country 1, while the higher transport costs for downstream firms mean that they prefer to build plants in both countries. The introduction of a small environmental tax in country 1 (i.e. a tax of 0.5 or 1.0) shifts the equilibrium to one in which one of the downstream firms closes one of its plants in country 2 and keeps a single plant in country 1 because the tax is smaller than the production cost differential. The environmental tax raises the price that downstream firms have to pay for the intermediate good in both markets, and reduces the importance of transport costs of downstream goods relative to the cost of production, and increases the importance of proximity to sources of supply. Because more production is located in country 1, welfare in country 1 (neglecting any allowance for damage from emissions) is higher than it was with no tax. When the environmental tax becomes larger, i.e. greater than 1, it has a more conventional effect, driving upstream firms to relocate in country 2, while downstream firms revert to having plants in both countries.

Thus we get the rather interesting result that because of agglomeration effects, small environmental taxes imposed by one country in one sector may have the effect of attracting in production in another sector from another country, in this case because it makes a firm in another sector care more about proximity to supply than to final customers.

3.3.3.3 Implications.

To summarise this sub-section, these experiments point to two implications for the design of environmental policy. First, not surprisingly if there are intersectoral linkages then the imposition of environmental policy in one sector can spill over into other sectors. The expected spillover is a negative one –if the tax causes firms in one sector initially located in one country to move to another then that may cause firms in related sectors to move as well. But the last example shows that the effect may also be positive –if the affected firms do not move out then the tax may attract firms in related sectors to move into the country imposing the tax. Second, if the sectoral linkages mean that over some range of environmental taxes there will be multiple equilibria then there may be hysteresis effects in which there is a threshold at which a large tax differential between countries puts in train location decisions which are not reversed by a small move back below that threshold and need a substantial narrowing of the tax wedge before the location decisions are reversed.

3.3.4 Nash Equilibria in Taxes.

We now report the results of a number of simulations where we introduce environmental damages caused by emissions and also allow both governments to set their emission taxes strategically, which we model as a Nash equilibrium choice of taxes. In the Introduction we noted that, using a much simpler model, Hoel (1997) showed that there are three possible types of equilibria: (i) a "tax-cutting equilibrium" where both governments set the same tax rate below marginal damage costs in order to compete to have production located in their country; production locates in one country but welfare is the same in both countries; (ii) an "asymmetric equilibrium" where the country which does not have production sets a higher tax than the country in which production is located; (iii) a "nimby equilibrium" where both countries set taxes at such high rates that production is shut down. Our model is considerably richer than Hoel's but we shall show that our model can generate the same three types of equilibria, although the arguments are somewhat different. However there is one respect in which our model could be expected to produce similar results to those of Hoel. Hoel assumes there is a single industry with a single firm; while we have two industries with two firms in each industry, the incentives for agglomeration mean that, as we have seen, for a wide range of parameter values, the equilibrium involves all firms locating in the same country. So the strategic incentives for governments are rather similar to those where there is a single firm.

3.3.4.1 Base Case

We begin with the base case with parameters set out in section 3.3.1 above, except that, for reasons we give shortly, we assume for these experiments that there is no profits tax. Emissions are assumed to be related only to production in the upstream industry, and we considered a range of values for the marginal damage cost (MDC) parameter from 0.3 to 24; note that a value of MDC = 3.0 implies that in the absence of any policy environmental damage would amount to about 15% of welfare.

Before presenting the results, it will be useful to recall that there are three factors that will influence governments in their decisions about how to set their emissions tax rates, corresponding to the three components of welfare.

(i) First, there are profits which accrue to shareholders. In the cases we shall be considering, all firms locate in the country with the lower emission tax. Given our assumption that there is no profits tax, governments are locally indifferent about where producers locate in terms of profits; that is if we start from a position where both governments set the same tax rate and all firms locate in country 1, say, the government in country 2 has no incentive to slightly reduce its emission tax in order to attract firms to locate there since that will have a negligible effect on the profits which accrue to the shareholders located in that country.

(ii) The second consideration is consumer surplus. Unlike Hoel's model we assume positive transport costs, so that when the equilibrium involves all firms locating in one country, consumers in that country will be better off than consumers in the other country who have to pay transport costs to obtain the final product. This provides an incentive for a government to attract industry to locate within its borders.

(iii) Finally there are environmental damages net of any emission tax. Given our assumption of linear damages, if a government sets emission tax below marginal (average) damage cost, then a country makes a net loss when firms locate there.

Putting these factors together, it is clear that equilibrium cannot involve governments setting emission taxes above marginal damage costs; for if that was the case, the country which attracted all the production would get a double benefit: its consumers would gain for the reasons given above, and its emission tax revenues would more than cover its environmental damage costs. It would always pay the other country to offer a lower emission tax and attract production. Thus the incentives are for governments to set emission taxes below marginal damage costs to attract firms to locate in their countries, and they have to trade off the loss in welfare from setting taxes below damage costs with the gain in consumer surplus from having lower prices.

The numerical results confirm this intuition for our base case. There are four possible type of equilibria depending on the value of the marginal damage cost parameter:

Tax Cutting Equilibrium.

For values of MDC set at 0.3, 0.75, 1.0 and 3.0, the equilibrium involves both governments setting the same tax rate, with production located in one country, and both countries getting the same level of welfare. Table 3.1 shows the equilibrium emission tax rate as a percentage of MDC for different levels of MDC:

Table 3.1: Equilibrium Tax Rates for Tax Cutting Equilibria

MDC	0.3	0.75	1.0	3.0
Tax Rate (% of MDC)	40	65	70	95

The intuition is straightforward. The lower is the MDC the smaller is the loss from setting taxes below MDC relative to the loss of consumer surplus from not having production, so the incentives for tax cutting are greater the lower are marginal damage costs.

Asymmetric Equilibrium.

For MDC set at 6.0, 9.0 and 12.0 the equilibrium involves one country setting a tax just below marginal damage cost and the other setting it equal to marginal damage cost. To see why consider the case where MDC = 9.0. Table 3.2 gives the part of the payoff matrix for the values of emission tax rate equal to 95% and 100% of MDC.

Table 3.2: Payoff Matrix for Asymmetric Equilibrium

		Country 2	
Country 1		95%	100%
	95%	(265, 272)	(265, 272)
	100%	(272, 265)	(268, 262)

If both countries set the same emission tax, then production locates in country 1. If a country sets the tax rate at 95% of MDC and production locates in that country, then it is worse off than the country without production, because the excess of damage costs over tax revenues is greater than the gain in consumer surplus. On the other hand when both countries set taxes equal to damage costs, then the country which gets the production is better off than the country without production. So country 1's dominant strategy (in this reduced payoff matrix) is to set emission tax equal to MDC, while country 2's dominant strategy is to set the emission tax equal to 95% of MDC. The argument is a discrete equivalent of the argument given by Hoel.

Non Tax Cutting Equilibrium.

For MDC at 15.0, 18.0 and 21.0 the equilibrium involves both countries setting emission taxes equal to MDC. The structure of the payoffs is very similar to those given in Table 3.2, so that in particular setting emission tax equal to MDC remains a dominant strategy for country 1. However, given that country 1 adopts this strategy, it is better for country 2 also to set its emission tax equal to MDC, since it is better off not having any production, though that gives it a lower consumer surplus than country 1, than to cut emission tax and take the loss of emission tax revenue relative to environmental damage costs.

NIMBY Equilibrium.

For MDC set at 24.0 the equilibrium is of a NIMBY type. The payoff matrix has the following structure:

Table 3.3: Payoff Matrix for NIMBY Equilibrium

		Country 2		
		80%	90%	100%
Country 1	80%	(-30, 36)	(-30, 36)	(-30, 36)
	90%	(36, -30)	(0, 0)	(0, 0)
	100%	(36, -30)	(0, 0)	(0, 0)

Thus setting taxes at 90% of MDC or above would close down all production, and countries get zero welfare. Setting tax at 80% of MDC would mean that the country with production would suffer negative welfare. So the equilibrium involves both countries setting emission taxes at least equal to 90% of MDC with the consequences that production shuts down. However, total welfare across both countries would be higher if both set emission taxes at 80% of MDC.

3.3.4.2 Variations of Base Case.

We now show how variations in the base case parameters affect the above conclusions.

(i) Positive Profits Tax.

We took the base case with MDC = 3.0 and introduced a profits tax of 50%. This should clearly increase the incentive for tax cutting, since the country with production located is able to tax the profits of the foreign firms as well as their emissions. With no profits tax we saw that the equilibrium involved both countries setting taxes at 95% of MDC. Now the equilibrium involved both countries cutting taxes to 70% of MDC.

(ii) Emissions Arising in Different Sectors.

We assumed above that emissions were related solely to production of upstream firms. The qualitative results are largely unaffected if emissions are related to production in downstream firms or in all firms. Obviously the numerical results are affected. For example, when emissions are related to production in both countries, the NIMBY equilibrium emerges with MDC set at 16.0, for the obvious reason that production is more damaging to the environment.

(iii) Higher Transport Costs.

We thought that the significant extent to which emission taxes were being set below MDC might be due to having rather low transport costs, so that firms were fairly "footlose" and so governments would have large incentives for rent-shifting. We considered a variation of the case reported in (i) above in which we raised transport costs in the downstream industry from $\theta = 0.1$ in the base case to $\theta = 0.5$. Now as we saw in section 3.3.2.2 in the absence of any emission taxes the equilibrium would involve the downstream firms locating plants in each country, while the upstream firms set up only a single plant, but each in a different country. It turns out that the Nash equilibrium in tax rates is still for both governments to set taxes at 70% of MDC, with the equilibrium as in the case of no emission taxes. When one country imposes a higher emission tax than the other that causes both the upstream firms to locate in the country with the lower tax.

As a second experiment we kept the transport costs in the downstream industry at $\theta=0.5$ and raised the transport costs in the upstream industry from a unit cost of 1.0 to a unit cost of 2.5. It turned out that the Nash equilibrium in taxes was again for governments to set their tax rates at 70% of MDC, with the location equilibrium again being agglomeration.

3.3.5 Summary.

In this section we have shown that for a wide range of parameter values, the equilibrium choice of plant locations involves agglomeration of all firms in the same location. However, there are also multiple equilibria in plant locations for many sets of parameter values, including policy parameters such as emission taxes. This can induce hysteresis effects in environmental policy where the equilibrium location selected for a given set of tax parameters depends on the history of policy that proceeded that choice of tax rates. Finally, for plausible values of damage cost parameters Nash equilibrium in environmental policy involves "ecological dumping", i.e. both governments setting emission taxes below marginal damage costs, and this is robust to a range of variations in parameters.

3.4. Conclusions.

In this paper we have extended the analysis of the impact of environmental policy on plant locations in imperfectly competitive models of international trade by allowing for inter-sectoral linkages between industries. The reason for doing this was to introduce some aspects of economic geography into the models by providing incentives for firms in different industries to locate close to each other. The implications that this generates for environmental policy are as follows. First, environmental policy imposed in one industry can have important spillovers effects into other industries. That is scarcely surprising, but what is somewhat surprising is that the effect can sometimes be positive: i.e. a country imposing a small environmental tax can attract firms from other industries to locate in that country. Second, there may be important hysteresis effects in setting environmental policies. These arise from the multiple equilibria that are a fairly natural feature of these models; because of the incentives for agglomeration it requires a co-ordination move by firms in different industries to switch locations. Finally, for plausible parameters the model generates quite strong incentives for "ecological dumping".

There are two obvious directions for future research. The first is to introduce proper dynamic analysis into the model to better capture the hysteresis effect, and closely linked to this is the question of the appropriate move structure in a fully specified dynamic model. We are investigating this question along the lines suggested by Ulph (1995), although he uses a simpler model than ours. The second is to test models of this form using actual industry data. Preliminary work by Venables (1994) based on the chemical undustry suggests that the inter-sectoral linkages in that industry are not sufficiently strong to generate the hysteresis effects discussed in this paper.

Hysteresis Effects in Environmental Policy

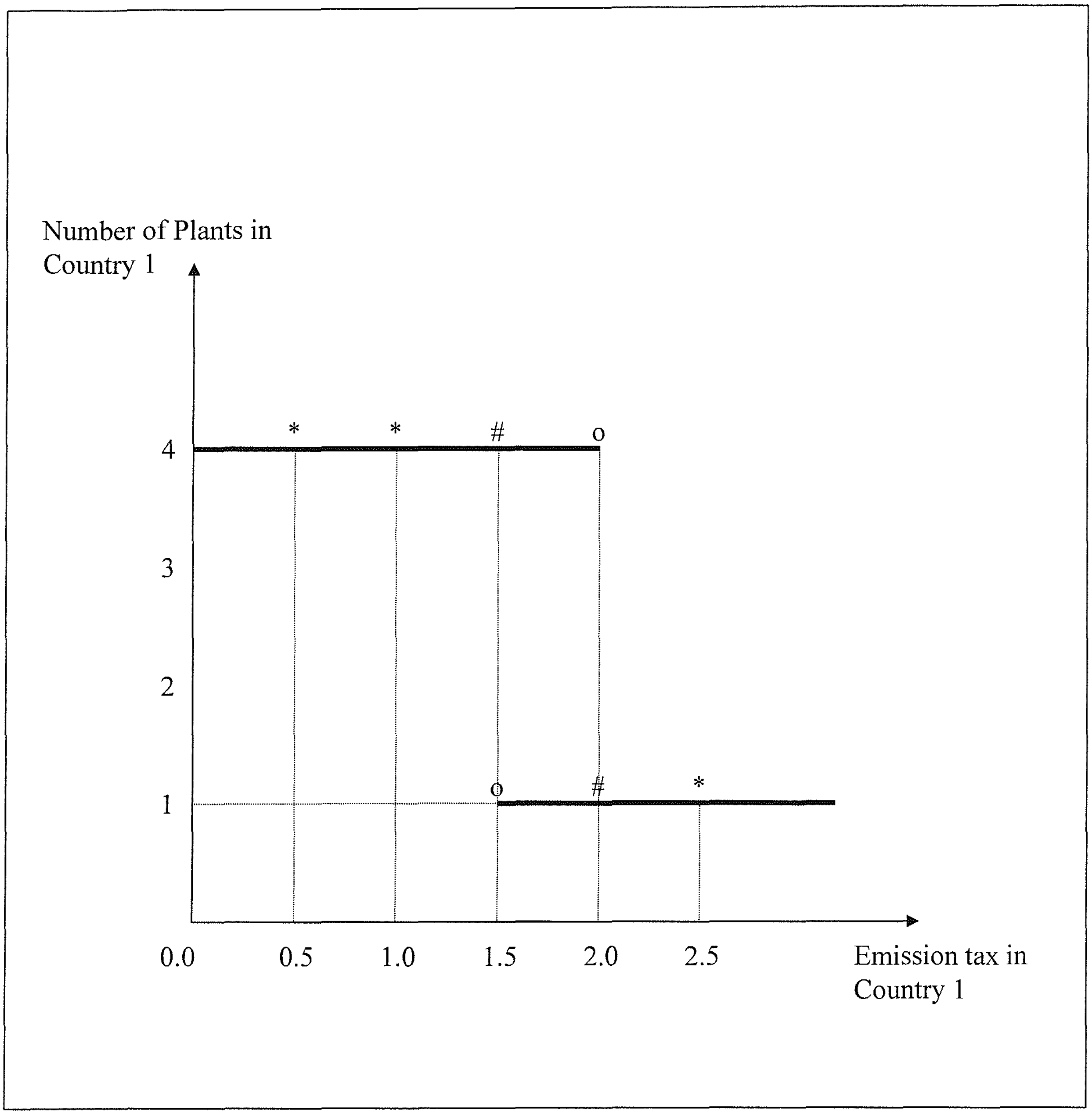

Fig. 3.1

Chapter 4

Is environmental dumping greater when firms are footloose?

Abstract

Concerns have been expressed that in a global market place with mobile capital, national governments will have incentives to set weak environmental policies ("environmental dumping") to protect the international competitiveness of their domestic firms and that these incentives are particularly strong in industries where plants may be relatively footloose so that governments are concerned to prevent "capital flight". In this paper we investigate whether the mobility of firms does indeed increase incentives for environmental dumping. We do this by taking a simple model of imperfect competition and comparing the environmental policies that would be set by non-cooperative governments for two different move structures - where governments set environmental policies after firms decide where to locate (the exogenous location case or Market Share Game) and where governments set environmental policies before firms decide where to locate (the endogenous location case or Location Game).

This raises an important modelling issue for it is natural in the Market Share Game that governments would set different environmental policies depending on the number of firms that locate in their countries, and if we are to compare just the effect of different move structures then in the Location Game we should also allow governments to condition their environmental policies on the number of firms that locate in them, contrary to previous models of the Location Game where governments set a single instrument independent of the number of firms that locate in their countries.

We show that the extent of environmental dumping in the Market Share Game may be greater or less than in the Location Game, depending in particular on the degree of substitution between products of the firms and hence the intensity of market competition. We also show that there is more environmental dumping in the Location Game when governments use a single instrument than when they condition their instruments on the number of firms that locate in their countries.

4.1. Introduction

There is considerable public concern that globalisation - the liberalisation of trade and capital flows - leads to policy competition between governments, which in the context of environmental policy may induce governments to set too lax environmental policies (ecological dumping) in order to protect the international competitiveness of their domestic firms, and that this incentive is particularly strong when firms are footloose so that governments may worry about 'capital flight'; the possibility of capital flight is often associated with multinational firms who must decide where to locate production activities. In the NAFTA debate the possibility of firms/plants relocating from the US to Mexico was an issue and a similar concern is expressed about the enlargement of the EU to include Eastern European countries. There has been a considerable theoretical literature on this topic (see Ulph (1997a) for a recent survey) which has included models where firm/plant locations are taken as fixed and others where firm/plant locations are endogenous ('footloose'). In both cases an important conclusion of the literature is that there can be no general presumption that globalisation will lead to environmental dumping - it is possible to construct models in both cases where competition between non-cooperative governments leads them to set policies which are tougher than would be the case if the governments cooperated and so eliminated strategic policy competition.

However, with a few exceptions (e.g. D. Ulph (1995) Markusen (1996)) there has been little attempt to compare the extent of strategic environmental policy competition when firm/plant location is fixed or endogenous. As noted by D. Ulph (1995) the distinction between models where location is fixed or endogenous can be thought of in terms of the relative ability of governments to commit themselves to environmental policies and firms to commit themselves to locations. Models of strategic environmental policy with endogenous location assume that governments commit to environmental policies prior to firms choosing where to locate while models with fixed location assume that firms commit to locations prior to governments setting their environmental policies. In this paper we shall compare the extent of environmental dumping with exogenous and endogenous firm locations by using a particular two-stage game model of imperfect competition in which we vary the move structure - with exogenous location (what we call the Market Share Game) firms first choose locations and governments then set policies while with endogenous location (what we call the Location Game) governments set their policies and firms then choose where to locate[1].

As already noted, the usual presumption is that the incentives for environmental dumping will be stronger when firms are footloose - the intuition being that there is more to compete for, or, using the commitment argument, that if firms have already committed to their location then this gives more scope for governments to raise taxes. What lies behind this presumption can be illustrated by the following very simple model, a variant of an argument used by Hoel (1997). Assume that there is only one firm and two countries, with no transport costs or capacity constraints, so that the firm will only use one plant, and will locate in the country that sets the weakest environmental policy. For simplicity assume that all consumption occurs in a third country and that profits accrue to the country of production. Governments use a simple emission limit (e) to regulate pollution. Welfare of the country where the firm locates is thus given by profits of the firm less environmental damage, shown as $V(e)$ in Figure 4.1, while welfare of the other country is zero.

If the firm chooses its location before the government sets its policy, then the government in which the firms locates will simply set the emission limit, e^*, which maximises welfare. This will be the usual first-best emission limit such that marginal abatement costs equal marginal damage costs, so there is no strategic element to environmental policy. If governments choose their policies first, then if governments set the same emission limit, e, we assume that the firm randomises where it locates so that each government gets expected welfare $U(e) = 0.5V(e)$,

[1] Of course a richer model would be a multi-stage game in which governments and firms where one models more explicitly what leads to different degrees of commitment; Feenstra (1998) conducts such analysis for the case where firms have to choose investment in capital, and shows that a multi-stage game can lead to different conclusions from a two-stage game but does not consider location decisions. D. Ulph (1995) also notes that rather than starting from the assumption that firms choose locations *de novo* in a world of liberalised trade and capital flows it may make more sense to assume that globalisation is a process so that one starts from a position of autarky in which there are already established firms located in different countries.

shown in Figure 4.1. If governments set different emission limits the firm locates in the country with the higher limit, and that country gets welfare $V(e)$. The equilibrium of the game in emission limits will be a simple 'race-to-the-bottom' in which both governments set emission limit $\bar{e}$ where welfare gets driven to zero. It cannot be an equilibrium for the governments to set the same emission limit below $\bar{e}$ because by setting a marginally higher emission limit a government could get welfare $V(e) > U(e)$, and it cannot be an equilibrium for governments to set different emission limits if the government with the higher emission limit gets strictly positive welfare, since the other government could set a marginally higher emission limit and get positive rather than zero welfare. Of course this argument depends on the assumption that $V(e^*) \geq 0$. If this is not true than while the outcome with fixed location will be as before (the government in the country with the firm sets emission limit e^*) with endogenous location both governments will set such tough environmental policies that no firm would want to locate in their countries - the NIMBY outcome[2]. If we ignore this NIMBY outcome, the conclusion is that there will be greater environmental dumping when the firm is footloose than when its location is fixed.

However, as we show in this paper, even within the simple framework set out in the example above this conclusion is not robust if we simply extend the number of firms from one to two, where the firms engage in Cournot competition. There are four reasons why this changes the conclusions reached above:

(i) With fixed locations there is now another possibility - that the two firms locate in separate countries. In this case, given the assumptions we have made, there will be strategic incentives for governments to engage in environmental dumping and set weaker environmental policies than they would set if they cooperated[3]. To see why this might change the conclusion, it is possible that, with cooperation and fixed location, governments can achieve positive welfare while with non-cooperation and fixed location governments get negative welfare. When

[2] Markusen, Morey and Olewiler (1995) reach a similar conclusion in a somewhat richer model in which there are trade costs so the single firm also has to decide whether to build one or two plants.
[3] In particular the assumptions of Cournot competition and the fact all profits accrue to the country in which the firm is located lead to environmental dumping (see Ulph (1997a) for more discussion.

governments set their policies before firms locate they can never end up with negative welfare, so there could be cases where even if there was a race-to-the-bottom with endogenous locations there would be even laxer policies if firms had located in different countries.

(ii) Precisely because there is now competition between the firms, it may no longer be the case that relaxing environmental policies always leads to higher profits for firms; in the absence of environmental policies total output would be higher than that which maximises profits, so there will be cases (with low environmental damage costs) where toughening environmental policies may raise profits and this may limit the race-to-the-bottom.

(iii) If, as we shall assume, damage costs are strictly convex, then, for a given level of output and pollution by each firm having both firms locate in one country will double the profits that country can earn it will more than double the environmental damage costs it gets. Now if we think of the two-firm analogue of the race-to-the bottom argument set out in Figure 4.1, we want to compare the welfare a country gets if it sets a higher emission limit than its rival and hence attracts both firms, $V(e)$, with the expected welfare it gets if it sets the same emission limit as its rival and both firms randomise where to locate, $U(e)$; in calculating this expected welfare we now include the possibility that the two firms locate in different countries. Because of the convex damage cost argument it may no longer be the case that $V(e) > U(e)$; indeed we will show that there must always be some values of e below $\bar{e}$ for which $V(e) < U(e)$ and indeed this may be true for all e. As we shall show this introduces the possibility of multiple equilibria for the game where firms are footloose, and while these will include the race-to-the-bottom equilibrium they may also include equilibria which have tougher emission limits than when firm locations are fixed. Hoel (1997) makes a similar argument when extending his model from one to many firms.

(iv) A further implication of the convex damage cost argument is that when we consider the environmental policies governments will set when firms' locations are fixed, it is quite natural to assume that they will set different environmental policies depending on whether one or two firms locate in their countries, in other words we can think of environmental policies being conditioned on the number of firms that locate in those countries. However the argument sketched out in (iii), and which is used by other authors when considering models of endogenous firm location with more than one firm[4], assumes that governments set a single environmental policy independent of the number of firms, so that if governments set different emission limits then all firms locate in the same country. But if we want to compare the difference in policies when firms have fixed or endogenous locations, and we do this by varying the move structure of the game, then if we allow a government to condition its policy instrument on the number of firms located in its country under one move structure we should do so under both. This means that when governments set policies before firms locate, governments can separate the emission limits they set to attract one firm to locate from that which they set to attract both firms to locate. Speaking loosely this means that governments are not driven into an all-or-nothing race-to-the-bottom but can settle for sharing the firms between them. Perhaps not surprisingly, if governments know they can secure an equilibrium where one firm will locate in each country they will set the same level of emission limits as if these locations were fixed, so we get exactly the same outcome irrespective of whether firms locations are fixed or endogenous. This occurs for a wide class of parameter values.

In this paper we use the simple model of an industry with more than one firm as set out above to compare the extent of environmental dumping where firms have fixed locations (Market Share Game) or endogenous locations (Location Game) where these two games represent two different move structures of a two-stage game. In the Location Game we compare the case where governments set only a single instrument irrespective of the number of firms with the case where governments condition policies on the number of firms located in their countries. We show that the extent of environmental dumping in the Market Share Game may be greater or less than in the Location Game, depending in particular on the degree of substitution between products of the firms and hence the intensity of market

---

[4] See Hoel (1997), Markusen, Morey and Olewiler (1993) although in the latter case the authors do not study policy competition between governments.

competition. We also show that there is more environmental dumping in the Location Game when governments use a single instrument than when they condition their instruments on the number of firms that locate in their countries.

In the next section we set out the model; sections 4.3 and 4.4 set out the equilibria of the Market Share Game and the Location Game; section 4.5 uses some numerical results to compare the equilibria of the two games, and section 4.6 concludes.

4.2. The Model

We use the simplest possible model to make our point as sketched in the Introduction. We consider a partial equilibrium model of a single industry in which, because of fixed costs and market size, there are only two firms, i = 1, 2, producing products which may be imperfect substitutes for each other. There are only two countries in which they can locate their production activities c = a, b. Firms and countries are identical. In neither country are there any consumers of the products they produce. There are sunk costs, F, of setting up a plant, but there are no transport costs, so each firm will only ever use one plant. Production involves the generation of a pollutant, which can be abated, at a cost. To control emissions of the pollutant the government of the country in which a plant is located can set plant-specific emission limits.

We denote firm i's output and emission limit by x_i, e_i and its profit function (net of abatement costs but excluding fixed costs) by $\pi(x_i, x_j, e_i)$ where $\pi_2 < 0, \pi_3 > 0$. In words, profits are decreasing in other firm's output, and an increase in emission limits for given output raises profits by an amount equal to marginal abatement costs. Emission limits are set prior to firms choosing their outputs, so in the output game each firm takes as given its emission limit and the output of its rival and chooses its own output to maximise profits. Firm i's *equilibrium* output is denoted $X^i(e_i, e_j)$ where $X_1^i > 0, X_2^j < 0\ X_1^i + X_2^j > 0$, i.e. an increase in firm i's emission limit increases its own output, reduces the rival firm's output but increases overall industry output (i.e. the own-effect is stronger than the cross-effect). Firm i's equilibrium profits are then denoted by $\Pi(e_i, e_j) \equiv \pi[X_i(e_i, e_j), X_j(e_j, e_i), e_i]$. This has the following properties:

$$\Pi_1(e_i, e_j) = \pi_2 X_2^j + \pi_3 \geq 0;$$
$$\Pi_2(e_i, e_j) = \pi_2 X_1^j \leq 0;$$

and we assume that $\Pi_{11} < 0, \Pi_{12} < 0, \Pi_{22} > 0$.

Thus an increase in firm i's emission limit will increase (or at least not decrease) its profits for two reasons: it reduces the output of the rival firm and it saves the firm some abatement cost. An increase in the rival firm's emission limit lowers firm i's profits because it expands the rival firm's output. For future purposes note that if both firms face the same emission limit then the effect of an increase in firm i's emission limit on total industry profits is given by: $\Pi_1(e_i, e_j) + \Pi_2(e_j, e_i) = \pi_2(X_2^j + X_1^j) + \pi_3$ where the first term is non-positive; this is because Cournot behaviour leads to industry output being above the profit-maximising level, so an increase in emission limit for one firm which causes total industry output to rise will reduce total industry profits. This discussion leads naturally to the definition of two critical levels of emission limits:

$$\bar{e}: \ \Pi_1(\bar{e}, \bar{e}) = 0; \ \Pi_1(e,e) > 0 \ \Leftrightarrow \ e < \bar{e};$$
$$\hat{e} < \bar{e}: \ \Pi_1(\hat{e}, \hat{e}) + \Pi_2(\hat{e}, \hat{e}) = 0; \ \Pi_1(e,e) + \Pi_2(e,e) > 0 \ \Leftrightarrow \ e < \hat{e}$$

In words, if emission limits are equal to e for both firms, then an increase in own emission limits will increase own profits provided $e < \bar{e}$, and will increase total industry profits provided $e < \hat{e} < \bar{e}$. Another way of describing $\bar{e}$ is that it corresponds to the emissions that would be emitted by each firm in a (symmetric) Cournot equilibrium with no environmental policy in place, so that any further increase in emission limits would have no effect since firms are doing no abatement and would not wish to increase their output. To make the model interesting we assume that $\Pi(\bar{e}, \bar{e}) - F \geq 0$, i.e. in the absence of environmental policy the Cournot equilibrium yields non-negative profits to the two firms. Finally we assume that there exists a non-negative emission limit, ε, which is the smallest non-negative value of e such that $\Pi(\varepsilon, \varepsilon) = 0$; we shall interpret ε as the NIMBY level of emission limit, i.e. if a country sets its emission level at or below ε it can guarantee that any firm which locates in its territory cannot make positive profits and will in general make losses if the rival country sets its emission level above ε.

Unabated emissions cause environmental damage to the country in which the plant is located (i.e. we ignore transboundary pollution). If total emissions in a country are E, then total damage costs from emissions are given by the damage cost function $\delta(E) \equiv \theta d(E)$, where $d(E)$ is a strictly convex function with $d(0) = 0$ and

θ is a parameter which allows us to compare the impact of varying the intensity of environmental damage costs. In addition we allow for the possibility that establishing a plant in a country is itself polluting (e.g. visual disamenity), so that there may be a fixed environmental cost $D \geq 0$ per plant.

Finally, we consider welfare. We assume that countries *a, b* impose a 100% profit tax on all firms located within their territories, so that all profits accrue to the country in which plants are located[5]. Given our assumptions[6], we have only two components of welfare to consider in a country - profits and environmental damage caused by firms located within that country. We distinguish two cases:

Firms locate in separate countries.

The welfare of, say, country *a* is given by:

$$W(e_a, e_b) \equiv \Pi(e_a, e_b) - F - \delta(e_a) - D$$

where e_c, $c = a, b$, is the emission limit imposed by country *c* on the firm located in its territory.

Both firms locate in one country.

The welfare of the country with no firms located in it is just 0. The welfare of the country where the plants locate is:

$$V(e_1, e_2) \equiv \Pi(e_1, e_2) + \Pi(e_2, e_1) - 2F - \delta(e_1 + e_2) - 2D$$

[5] Strictly speaking we need to give firms some incentive to care about the impact of their decisions, such as location, on profits. We could assume that managers are delegated by the owners to maximise profits, or that a fixed proportion of profits is retained by owners (who reside outside countries *a* and *b*), and that this proportion is sufficiently small that it can be ignored in calculating country welfare levels.

[6] The two key assumptions are that there are no consumers located in countries *a* and *b* and that these countries levy 100% profits tax. The former assumption means that countries can ignore consumer welfare in setting environmental policy; if there were consumers located in these countries this would reinforce incentives to relax environmental policies. The second assumption means that we can ignore the location of the two firms' owners (but see previous footnote); without this assumption governments may have incentives to set tougher environmental policies (as a form of implicit tax on foreign shareholders). See Ulph and Valentini (1997) for further discussion of these points.

where $e_i, i = 1,2$ is the emission limit that country imposes on firm *i*.

It will simplify notation in future if we define $\phi = F + D$ as the *sunk costs* both environmental and non-environmental associated with having a firm locate in a country.

This completes the description of the model. We now discuss two alternative games depending on whether firms choose location before governments set their environmental policies or alternatively governments set their policies before firms choose their locations.

4.3. Firms locate before governments set policies

In this section we model a two-stage game, which we refer to as the "Market-Share Game" in which in the first stage the two firms choose in which of the two countries to locate their plants and in the second stage governments of the two countries set their environmental policies. This game captures the situation where governments are unable to commit to environmental policies for the length of time in which firms will operate their plants within their territories. If the firms locate in different countries the outcome will correspond to the standard strategic environmental policy model of rent-shifting, which we have referred to as competition for market share. As usual we solve the two stages of the game backwards.

4.3.1 Second-Stage Game.

We focus only on equilibria from the first stage in which both firms are in the market.

<u>(i) Two Firms Locate in Different Countries.</u>

This is the standard Brander-Spencer model of strategic trade adapted to environmental policy (see Barrett (1994), Conrad (1993), Kennedy (1994), Ulph (1996)). For ease of notation we shall refer to the emission limits set by the two countries as $f_{a,}f_b$. Then country a, say, takes as given f_b and chooses f_a to maximise $W(f_a, f_b)$. The resulting symmetric Nash equilibrium will involve both countries setting the emission limit f^N which satisfies:

$$W_1(f^N, f^N) = \Pi_1(f^N, f^N) - \delta'(f^N) = \pi_2 X_2 + \pi_3 - \delta'(f^N) = 0 \qquad (1)$$

We noted in section 4.2 that $\pi_2 X_2 \geq 0$, so, as is now well know, this Nash equilibrium will involve "environmental dumping" in the sense that marginal abatement cost is below marginal damage cost. Governments relax environmental policies to try to reduce the output of the firm located in the other country. It is also straightforward to show if the two countries cooperated to maximise the sum of welfare in the two countries they would select a common emission limit f^C which satisfied the condition:

$$\begin{aligned} W_1(f^C,f^C)+W_2(f^C,f^C) &= \Pi_1(f^C,f^C)+\Pi_2(f^C,f^C)-\delta'(f^C) \\ &= \pi_2(X_1+X_2)+\pi_3-\delta'(f^C)=0 \end{aligned} \quad (2)$$

We noted in section 4.2 that $\pi_2(X_1+X_2)\le 0$ so that the cooperative emission limits will be such that marginal damage costs are less than marginal abatement costs; the reason is that governments use emission limits not only to deal with the environmental externality but also with the fact that Cournot output is too high relative to the industry-profit maximising level. It is clear that $f^C < f^N$, so we also get "environmental dumping" in the sense that environmental policies set by the governments when they do not cooperate are laxer than when they do cooperate.

To summarise, if the two firms locate in different countries then the payoffs to the countries and the firms respectively are: $W^N \equiv W(f^N,f^N)$; $\Pi^N \equiv \Pi(f^N,f^N)-F$.

(ii) Both Firms Locate in the Same Country.

Obviously the country in which neither firm locates gets a payoff of zero. Again for ease of notation we shall denote the emission limits set for firms 1 and 2 by the country in which they do locate by g_1, g_2. These are chosen to maximise $V(g_1,g_2)$. The solution must involve the country setting the same emission limit, g^*, for each firm, where g^* satisfies the first-order condition:

$$\begin{aligned} V_1(g^*,g^*) &= \Pi_1(g^*,g^*)+\Pi_2(g^*,g^*)-\delta'(2g^*) \\ &= \pi_2(X_1+X_2)+\pi_3-\delta'(2g^*)=0 \end{aligned} \quad (3)$$

Comparing (1), (2) and (3), and given our earlier discussion, it is clear that $g^* < f^C < f^N$. Having both firms locate in the same countries achieves all the benefits of cooperation - so the government in that country no longer has any strategic incentive to weaken environmental standards to reduce the market share of a foreign competitor, and indeed has an incentive to set tougher emission limits than in the simple first-best rule of equating marginal abatement cost and marginal damage cost in order to restrict total industry output below the level that Cournot firms would set. However, if both firms are located in the same country emission limits per firm will be lower than in the cooperative equilibrium given our assumption of strictly convex damage costs, so that marginal damage costs per firm are higher than when the two firms locate in different countries.

In summary, when both firms locate in a single country then the profits to each firm and welfare to the country respectively are: $\Pi^* \equiv \Pi(g^*, g^*) - F; \quad V^* \equiv V(g^*, g^*)$.

4.3.2 First-Stage Game.

The two firms can predict perfectly the policies that governments will implement in the second stage once they have chosen their location. So they will locate in a single country if $\Pi^* > \Pi^N$, locate in separate countries if $\Pi^N > \Pi^*$, and if profits are equal each firm will just randomise where to locate, choosing each country with equal probability. Now achieving this outcome involves a degree of co-ordination between firms, which we shall solve as follows. We assume that firm 1 first chooses its location randomly, so that with 50% probability it locates in country a and with 50% probability in country b. Firm 2 then chooses its location, and will locate in the same country as country 1 if $\Pi^* > \Pi^N$, will locate in the country other than that in which firm 1 is located if $\Pi^N > \Pi^*$ and otherwise it too will randomise where it locates.

Recall from section 4.2 that marginal damage costs are strictly increasing in a parameter θ and that $\hat{e}$ maximises $\Pi(e,e)$. If $\theta \cong 0$ then $g^* \cong \hat{e}, \; f^N \cong \bar{e} > \hat{e}$, so that $\Pi^* > \Pi^N$ i.e. for small values of θ profits will be greatest if both firms locate in the same country. The reason is that with low environmental damage costs, firms benefit from the fact that a single government tries to restrict output below the Cournot level. As θ increases, both g^* and f^N decrease and certainly if $g^* < f^N \leq \hat{e}$ then $\Pi^* < \Pi^N$, so that when θ is high firms are better off locating in separate countries. We suppose there exists a unique $\bar{\theta} \; s.t. \; \Pi^* > \Pi^N \Leftrightarrow \theta < \bar{\theta}$.

Thus firms will locate in the same country if $\theta < \bar{\theta}$, will locate in separate countries if $\theta > \bar{\theta}$ and will randomise their location if $\theta = \bar{\theta}$.

4.3.3 Conclusion of Market Share Game.

This completes the analysis of the two stages of the Market Share Game. The overall equilibrium can be summarised as follows. If $\theta < \bar{\theta}$, then both firms locate in the same country, and it is equally likely that this will be country *a* or *b*; the country in which they locate sets emission limit g^* for each firm; each firm makes profit Π^*, and each country gets expected welfare $0.5V^*$. If $\theta > \bar{\theta}$ both firms locate in different countries; each country sets emission limit f^N, obtaining welfare W^N while each firm makes profits Π^N. Finally, if $\theta = \bar{\theta}$ each firm randomises where it locates. Each firm makes profits $\Pi^* = \Pi^N$ while each country gets expected welfare $0.25V^* + 0.5W^N$.

We have ignored the possibility that only one firm may be able to survive in the market, and to justify this we need to assume that $\max[\Pi^*, \Pi^N] \geq 0$. Note that this implies that for one of the two equilibria profits may be negative. Note also that there is nothing which guarantees that a country is better off having one or more firm locate in its territory than having no firm locate in its territory.

4.4. Firms locate after governments set policies.

We now turn to an alternative game, the "Location Game" in which governments first set their environmental policies and then firms choose where to locate. This corresponds to a situation where governments can commit themselves to environmental policies for the duration of a plant's life.

It was quite natural in the previous section to suppose that the environmental policy set by a country would depend on whether one or two firms located within its territory. If the only thing that is being changed is the order of moves, then we should also allow that governments can condition their environmental policies on the number of firms that locate in their territories. However, as we noted in the introduction, many of the existing models of location games, e.g. Hoel (1997), Markusen, Morey and Olewiler (1993) assume that governments set environmental policies which are independent of the number of firms that locate within their boundaries. We argue that if we want to explore the implications of changing the move structure then we should not compound this with a change in the set of available instruments. In order to see what impact it makes when we change assumptions about the number of instruments available we shall analyse the location under two sets of assumptions; the first is that the government of each country c must use a single instrument, i.e. it must set $f_c = g_c = h_c$; the second assumption is that the government of each country can use two instruments, f_c and g_c which *need* not be the same, though they may choose to set them equal.

4.4.1 Second-Stage Game - Single Instrument Case.

The analysis here is straightforward. We shall suppose that emission limits that are relevant for this analysis lie below level $\bar{e}$ (defined in section 4.2 by $\Pi_1(\bar{e},\bar{e}) = 0$) so that profits are increasing in own emission limit. Then whatever location decision the other firm makes, and hence whatever emission limit h_c the other firm faces, $\Pi(h,h_c)$ is increasing in h and so it is a dominant strategy for a firm to always locate in the country which offers the higher emission limit. Thus we have a simple outcome of this game: if emission limits differ, both firms locate in the country with the higher limit; if emission limits are the same in both countries then firms are indifferent where they locate and we suppose that each firm just randomises where it locates, so that with probability 0.25 both firms locate in country a, with probability 0.25 both locate in country b and with probability 0.5 the firms locate in separate countries.

4.4.2 Second-Stage Game - Two Instrument Case.

This game is more complicated than the previous analysis, because the ability of the countries to condition their instruments on the number of firms that locate in their territories means that the simple argument that own profits are always higher when locating in the country with a higher emission limit, irrespective of the location of the other firm, breaks down. A full analysis of all possibilities is given in Appendix A. For the purposes of the later analysis we can focus on a subset of cases. By symmetry any equilibrium will involve the two countries setting the same emission limits, i.e. $f_a = f_b = f, g_a = g_b = g, f \neq g$ and we denote the associated profits by Π_f and Π_g. The location decision is then simple: firms will locate in separate countries if $\Pi_f > \Pi_g$ and in the same country if $\Pi_g > \Pi_f$. As in the Market Share Game we assume they achieve this outcome by firm 1 first choosing with equal probability which country it locates in and firm 2 then locating in the same or a different country as required by the above profit conditions.

We need to also consider what happens if a country deviates from such an equilibrium. If the countries set different values of f then we denote by Π'_f and Π''_f respectively the minimum and maximum values of profits; we define Π'_g and Π''_g in a similar fashion. Then we consider deviations from equilibrium in two stages.

<u>$\Pi_f > \Pi_g$</u> The equilibrium will involve both firms locating in separate countries. If a country deviates in its f policy instrument to try to earn higher profits for its firm, that will have no impact on the location decision of firms unless this results in $\Pi'_f < \Pi_g < \Pi''_f$, in which case the location equilibrium switches to one in which both firms locate in the same country. In particular this means that if $f \cong g$ so that

the initial difference in profits $\Pi_f - \Pi_g$ is positive but very small, then the deviation being discussed will lead to a switch to an equilibrium with both firms locating in the same country. However if there is a significant difference between *f* and *g* then deviations in *f* will leave the location decisions unaffected. If a country deviates from the initial equilibrium in its *g* policy instrument then that will have no influence on the location decision unless this results in $\Pi'_g = \Pi_g < \Pi_f < \Pi''_g$ when the location equilibrium will switch to having both firms locate in the country that offers the higher profits.

$\underline{\Pi_g > \Pi_f}$ The equilibrium will involve both firms locating in the same country, but randomising which one that is. Suppose one country deviates in its *f* policy instrument. Even if it achieved higher profits $\Pi''_f > \Pi_f$ such that $\Pi''_f > \Pi_g$, since this would also imply that, with unchanged policy instrument *f* in the other country, the firm which located in that country would have profits $\Pi'_f < \Pi_f < \Pi_g < \Pi''_f$, then, as we have seen above, the resulting equilibrium would still be one in which both firms located in the same country. It would require $\Pi_g < \Pi'_f < \Pi''_f$ for the equilibrium to switch to one in which the firms locate in separate countries, and no single deviation by a country on its own can achieve this outcome. Now suppose one country deviates in its *g* policy instrument. The result would then be that $\Pi_f < \Pi_g = \Pi'_g < \Pi''_g$ and the location equilibrium will change from one in which both firms locate in the same country but randomise which country will be to one in which they both locate in the country offering the higher profits.

This completes the description of the second-stage of the Location Game with one and two instruments and we now turn to the first stage.

4.4.3 First-Stage Game - Some Properties of Welfare Functions.

Before setting out the equilibria of the first-stage game it will be useful to establish some properties of the relevant welfare functions and also establish some notation. For a start, where we consider symmetric situations, we shall write the profit function and welfare functions with a single argument. From the description of the second-stage of the location game, it is clear that there are four possibilities for equilibria whose welfare we will need to evaluate:

(i) both firms locate in a single country with emission limits g which offers the higher profits; that country will have welfare: $V(g) = 2\Pi(g) - \delta(2g) - 2\phi$; the other country will have welfare 0.

(ii) both firms locate in a single country, but because both countries set the same emission level g they randomise which that country will be; each country will then have expected welfare: $\overline{V}(g) = 0.5V(g) = \Pi(g) - 0.5\delta(2g) - \phi$.

(iii) the two firms locate in separate countries which set the same emission limit f and so each country earns welfare: $W(f) = \Pi(f) - \delta(f) - \phi$.

(iv) both countries use a single instrument and set the same value h; both firms randomise where they locate and so each country gets expected welfare: $U(h) \equiv 0.25V(h) + 0.5W(h) = \Pi(h) - [0.5\delta(h) + 0.25\delta(2h)] - \phi$.

We begin by noting that since the damage cost function δ (.) is strictly convex, then for any positive emission limit g, say, $0.5\delta(2g) > [0.5\delta(g) + 0.25\delta(2g)] > \delta(g)$, and hence $W(g) > U(g) > \overline{V}(g)$. By the same argument, the slopes of these welfare functions satisfy the inequalities: $W'(g) > U'(g) > \overline{V}'(g)$. Three things follows:

(a) Define f^*, h^*, g^* as the values of the emission limits which maximise, respectively, $W(f)$, $U(h)$, $\overline{V}(g)$ [equivalently $V(g)$]; then $g^* < h^* < f^*$.

(b) Define $\bar{f}, \bar{h}, \bar{g}$ as the maximum values of emission limits such that $W(\bar{f}) = U(\bar{h}) = \overline{V}(\bar{g}) = V(\bar{g}) = 0$ then $\bar{g} < \bar{h} < \bar{f}$.

(c) We can define values of ϕ, $\overline{\phi}_1 > \overline{\phi}_2 > \overline{\phi}_3$, such that:

$$\phi \geq \overline{\phi}_1 \Leftrightarrow W(f) \leq 0 \quad \forall f \geq 0$$
$$\phi \geq \overline{\phi}_2 \Leftrightarrow U(h) \leq 0 \quad \forall h \geq 0$$
$$\phi \geq \overline{\phi}_3 \Leftrightarrow V(g) \leq \overline{V}(g) \leq 0 \quad \forall g \geq 0$$

For the single instrument case we shall need to know when $V > U$ while for the two instrument case we shall need to know when $V > W$. It is straightforward to see that there will be values of ϕ, $\bar{\phi}_4$ and $\bar{\phi}_5$ where $\bar{\phi}_3 > \bar{\phi}_4 > \bar{\phi}_5$ such that:

$$\phi \geq \bar{\phi}_4 \Leftrightarrow V(g) \leq U(g) < W(g) \; \forall g \; s.t. \; V(g) \geq 0$$
$$\phi \geq \bar{\phi}_5 \Leftrightarrow V(g) \leq W(g) \; \forall g \; s.t. \; V(g) \geq 0$$

Then it is readily shown that:

$$\phi < \bar{\phi}_4 \Rightarrow \exists \, \tilde{h}, \; h^* < \tilde{h} < \bar{g} \quad s.t. \; V(\tilde{h}) = U(\tilde{h}), h < \tilde{h} \Rightarrow V(h) > U(h), h > \tilde{h} \Rightarrow V(h) < U(h)$$
$$\phi < \bar{\phi}_5 \Rightarrow \exists \, \tilde{f}, \; f^* < \tilde{f} < \bar{g} \; s.t. \; V(\tilde{f}) = W(\tilde{f}), f < \tilde{f} \Rightarrow V(f) > W(f), f > \tilde{f} \Rightarrow V(f) < W(f)$$

Since $V(\tilde{h}) = U(\tilde{h}) < W(\tilde{h}), \tilde{h} > \tilde{f}$.

These properties of the welfare functions are illustrated in Figure 4.2.

4.4.4 First-Stage Game - Single Instrument.

There are only two possible outcomes: either one country sets a higher emission limit, *g*, say, than the other country and has both firms locating in that country, yielding welfare *V(g)*, or both countries set the same emission limit, *h*, say and the firms randomise where to locate and both countries get expected welfare *U(h)*. Before discussing possible equilibria of the first-stage game, we provide an argument which rules out certain possibilities. There cannot be an equilibrium with country *c* setting a higher emission limit than the other, thereby attracting both firms to locate in country *c* and yielding country *c* positive welfare, since that could be overturned by the other country setting an even higher emission limit. This is the essence of the race-to-the-bottom argument. Similarly there cannot be an equilibrium with both countries setting the same emission limit *h* such that $V(h) > U(h)$ since one country could set an emission limit slightly higher than *h*, attract both firms to locate in its territory and earn welfare $\approx V(h) > U(h)$. We now consider possible equilibria in four ranges.

(i) $\phi \geq \overline{\phi}_2$ Since $V < U \leq 0 \; \forall h$ countries are at least as well off, and in general better off, having no firms locate in them, and so we get the NIMBY equilibrium in which both countries set emission limits at or below ε.

(ii) $\overline{\phi}_3 \leq \phi < \overline{\phi}_2$ In this range, $V \leq 0 \; \forall h$ but $U(h^*) > 0$. Clearly no country will wish to have both firms locate within its territory. We begin by arguing that any h such that $U(h) \geq 0$ would be an equilibrium. The argument is that if the other country has set such an emission level h, then if country c also sets emission limit h it will get welfare $U(h) \geq 0$, which in general will be positive; if it sets its emission limit below h then it will certainly get zero welfare while if it sets its emission limit above h it will have both firms locate in it and earn non-positive (in general negative) welfare.

(iii) $\overline{\phi}_4 \leq \phi < \overline{\phi}_3$ In this range $V(g^*) > 0, U(h^*) > 0, V(h) \leq U(h)$. We begin by defining $\hat{h}$ as the minimum value of h such that $U(\hat{h}) = V(g^*)$ and arguing that any h in the range $[\hat{h}, \overline{h}]$ is an equilibrium. The argument is essentially the same as in (ii); if the other country has set h then country c can get a payoff $U(h) \geq 0$ by also setting emission limit h. Setting its emission limit below h would give country c zero welfare; setting its emission limit above h, say h' will give it a payoff $V(h') \leq U(h)$. Note that the latter argument rules out emission limits below $\hat{h}$ as possible equilibria defecting to g^* would make a country better off. However we shall argue that countries may restrict themselves to the narrower range of equilibria $[h^*, \overline{h}]$. Our argument is that for any potential equilibrium h in the range $[\hat{h}, h^*)$ there is another equilibrium $> h^*$ which will yield the same payoff; countries will have a natural bias in setting the higher equilibrium because of an asymmetry of risks; if they set an emission limit lower than their rival they end up with a zero payoff; if they set an emission limit above their rival they obtain a lower, but not necessarily zero, payoff; so there is a natural bias to setting higher than lower emission limits if they earn the same payoffs; so we shall argue that only emission limits at least as great as h^* would be ones on which countries focus. We shall assume that this argument applies to the previous case (ii) as well.

(iv) $\phi < \overline{\phi}_4$ In this range *V(g*)* > *U(g*)* > 0, *U(h*)* > 0. By the general argument made at the beginning of this subsection we can rule out any emission limits below $\tilde{h}$ as possible equilibria. Otherwise the arguments in (ii) and (iii) apply. Thus the range of possible equilibria is $[h^*, \overline{h}]$ if $\tilde{h} \leq h^*$ and $[\tilde{h}, \overline{h}]$ if $\tilde{h} > h^*$.

These are illustrated in Figure 4.3.

To summarise the single instrument case, there are three possible kinds of equilibria: the NIMBY equilibrium, any h in the range $[h^*, \overline{h}]$, and any h in the range $[\tilde{h}, \overline{h}]$. However, in trying to compare the equilibrium when governments use a single instrument with that when they use two instruments, it will be useful to select a particular equilibrium from the range. We shall use three possible rules for selecting such an equilibrium. First countries may choose the equilibrium which gives the *maximum welfare*, which could be either what we shall call the Maximum Expected Welfare (MEW) equilibrium, h^*, or the Intermediate Equilibrium, $\tilde{h}$. Second, the asymmetric risks argument could be used to select the highest emission limit in the range, which leads to the selection of the *race-to-the-bottom* equilibrium, $\overline{h}$. These give us the two extremes of the range of equilibria, so a natural third candidate would be an average of the maximum welfare (either MEW or Intermediate equilibria) and the race-to-the-bottom equilibrium, which we shall take to be just the mid-point of the appropriate range.

This completes the description of the first-stage of the Location Game when countries use a single instrument and we now turn to the case where countries use two instruments.

4.4.5 First-Stage Game - Two Instruments.

We begin by noting that any equilibrium of the model with a single instrument remains an equilibrium when governments can use two instruments. To see this, suppose that there is an equilibrium in which $f_a = g_a = f_b = g_b = \hat{h}$, say. Suppose that country a, say, deviates by setting $f_a > \hat{h}$ so that $\Pi(f_a, \hat{h}) > \Pi(\hat{h}, \hat{h}) > \Pi(\hat{h}, f_a)$. Then from the analysis of the second stage game we know that the equilibrium location decision will be that both firms will locate in the same country, but randomise which one that will be. So both countries will now get welfare $\overline{V}(\hat{h}) < U(\hat{h})$ so country a (and b) will be worse off than in the original equilibrium. Now suppose that country a deviates by setting $g_a \neq \hat{h}$ so that $\Pi(g_a) > \Pi(\hat{h})$ and having both firms locate in country a. But that cannot improve welfare for country a since such a deviation would have been available to country a in the single instrument model by setting $h_a = g_a$ and if that improved welfare for country a then $\hat{h}$ could not have been an equilibrium of the single instrument model.

We now consider possible equilibria where the governments set different instruments depending on the number of firms that locate in their territories. Note first that the general argument we used to rule out possible equilibria for the single instrument model applies here too, but with one qualification. There cannot be an equilibrium involving both firms locating in one country which yields that country strictly positive welfare and where the policy instrument contingent on two firms locating in one country lies strictly below $\hat{e}$, the level at which profits are maximised; in other words the simple race-to-the-bottom argument that applies with a single instrument may be thwarted in the two instrument case by one government not being able to offer higher profits to attract both firms while there is still positive welfare. While there will be some other similarities with the analysis of the single instrument case, there will be two important differences which arise from the ability of governments to set instruments contingent on the number of firms. First, where the equilibrium involves firms locating in separate countries we can use the separation of this instrument from the instrument when two firms locate in a country to specify which equilibrium governments will choose and hence eliminate the multiple equilibria that arose in the single instrument model. Second, when the equilibrium involves firms locating in the same country we need to take account of the constraint that the profit function $\Pi(g,g)$ reaches a maximum at $g = \hat{e}$, so if a government is trying to attract both firms to locate in its territory it will never pay to set its emission limit contingent on two firms locating in its country above $g = \hat{e}$. This latter constraint makes the analysis of the possible equilibria a bit messier than in the single instrument model. We now consider four ranges of parameter values.

(i) $\phi \geq \overline{\phi}_1$ In this range $W \leq 0\ \forall f \geq 0,\quad V \leq 0\ \forall g \geq 0$ so no government wants any firms locating in its country. Thus both governments will pursue the NIMBY policies of setting f and g less than or equal to ε.

(ii) $\overline{\phi}_3 \leq \phi < \overline{\phi}_1$ In this range $W(f^*) > 0, V(g) \leq 0\ \forall g \geq 0$, so no government would want two firms locating in its territory and hence both governments will set g less than or equal to ε. However having a single firm locate in a country can yield positive welfare for some emission limits. Given that governments are blocking both firms locating in their territories, the obvious candidate for an equilibrium choice of policy instrument contingent on a single firm locating in a country is the Nash equilibrium f^N of the Market Share Game and this will be the chosen policy instrument provided $W(f^N) \geq 0$, i.e. $f^N \leq \bar{f}$. If this condition is not satisfied then governments will set policy instruments as close as possible to the Nash equilibrium while having non-negative welfare, so they will set the Constrained Nash equilibrium emission limits $\bar{f}$. Thus the equilibrium in this range involves governments setting the Nash or Constrained Nash equilibrium instruments contingent on a single firm locating in their territories and this will indeed be the equilibrium choice of location by firms.

(iii) $\overline{\phi}_5 \leq \phi < \overline{\phi}_3$ In this range $W(f^*) > 0, V(g^*) > 0, V(f) \leq W(f)\ \forall f \geq 0$. Again the natural candidate for an equilibrium in this range is the Nash equilibrium of the Market Share game, and this will be true provided $f^N \leq \hat{e}, f^N \leq \bar{f}$. It would not pay any government to deviate from this equilibrium in order to attract both firms to locate in its country since that would require setting $g > f^N$ to yield welfare $V(g) < W(f^N)$, and given that countries are not interested in having both firms locate in their territories it is natural to select the Nash equilibrium among policies involving only a single firm in each country. We need to address what happens when the two conditions fail to apply.

(a) Suppose first that $\hat{e} < f^N < \bar{f}$, then the argument just given fails because to attract both firms to one country a government would not set $g > f^N$ but rather $g \geq \hat{e}$; the usual race-to-the-bottom argument rules out any equilibrium with $g < \hat{e}$ In effect then there may be two possible equilibria: both governments set $g = \hat{e}, f = f^N$, both firms locate in a single country, but randomise which one that is, so that each country gets expected welfare $\overline{V}(\hat{e})$; or both governments set policies $f = f^N, g = g'\ s.t.\ \Pi(g') < \Pi(f^N)$, the firms locate in separate countries and each country gets welfare $W(f^N)$. The first of these equilibria, which we call the Profit-Constrained Race-to-the-Bottom, will be selected if $\overline{V}(\hat{e}) > W(f^N)$, otherwise the Nash Equilibrium will be selected.

(b) Finally we consider what happens if $f^N > \bar{f}$. There are three possibilities. We get the Constrained Nash Equilibrium $(f = \bar{f}, g = \overline{g})$ with both firms locating in separate countries and both countries getting welfare $W(\bar{f}) = 0$ if $\Pi(\bar{f}) > \Pi(\overline{g})$. We get the Race-to-the-Bottom Equilibrium $(f = \bar{f}, g = \overline{g})$ with both firms locating in the same country but randomising which one that is, so that each country gets expected welfare $\overline{V}(\overline{g}) = 0$ if $\Pi(\bar{f}) < \Pi(\overline{g})$ and $\overline{g} < \hat{e}$. Finally we get the Profit-Constrained Race-to-the-Bottom $(f = \bar{f}, g = \hat{e})$ with both firms locating in the same country but randomising which one that is, and both countries getting expected welfare $\overline{V}(\hat{e})$ if $\hat{e} \leq \overline{g} < \bar{f} < f^N$.

To summarise, the prime candidate for an equilibrium in this range is the Nash Equilibrium of the Market Share Game, but constraints on either profits or non-negativity of welfare mean that we may also get various "corner solutions" - the Constrained Nash Equilibrium, the Race-to-the-Bottom Equilibrium and the Profit-Constrained Race-to-the-Bottom Equilibrium.

(iv) $\phi < \overline{\phi}_5$ In this range $V(g^*) > 0$, $W(f^*) > 0$, $V(g^*) > W(g^*)$. All of the possible equilibria identified in (iii) remain possible equilibria in this range, but there is an additional condition required for the Nash equilibrium to be an equilibrium, namely $W(f^N) \geq V(f^N)$ i.e. $f^N \geq \tilde{f}$. This condition was automatically satisfied in (iii) but may not be satisfied in this range of parameters, and when it is not this introduces the possibility of another equilibrium, which we shall call the Intermediate Equilibrium (intermediate between the Nash and the Race-to-the-Bottom Equilibria). It is defined by the two governments setting policy instruments $f = \tilde{f}, g = g''\ s.t.\ \Pi(g'') < \Pi(\tilde{f})$, the two firms locating in separate countries and each country earning welfare $W(\tilde{f})$. The conditions for the Intermediate Equilibrium to be the equilibrium are: $f^N < \tilde{f} \leq \hat{e}$ or $f^N < \tilde{f}, \tilde{f} > \hat{e}$ but $W(\tilde{f}) \geq \overline{V}(\hat{e})$; failure of the last inequality means that the Profit-Constrained Race-to-the-Bottom Equilibrium will be the equilibrium.

We illustrate the possible equilibria of the two instrument model with different instruments in Figure 4.4.

This completes the analysis of the possible equilibria when governments set different policy instruments contingent on the number of firms locating in their countries. There is one final step to complete the analysis of the two instrument model. We have noted that for any parameter values, an equilibrium of the single instrument model remains an equilibrium of the two instrument model. We have also just analysed for any parameter value the equilibrium that arises when governments choose to set the two instruments different from each other. So for any set of parameter values we have two candidate equilibria and we assume that governments select the one which yields higher welfare.

4.4.6 Conclusions of Location Game

This completes the analysis of the Location Game. The questions we wish to address are how does the difference in move structure affect the extent to which environmental policies are manipulated for strategic reasons of competition for market share or location of plants and what are the welfare consequences of such strategic competition? How does the ability of governments to condition their policy instruments on the number of firms locating in a country affect these indicators of strategic competition? To answer such questions we need to use a benchmark of what the outcome would be in the absence of such strategic competition. The benchmark we use is the cooperative equilibrium where the two firms locate in separate countries, the governments set policy instruments $f = f^*$ and each country earns welfare $W(f^*)$. We choose this for the obvious reasons that co-operative behaviour eliminates the incentive for strategic competition, and given the strict convexity of the damage cost function, for any given total production and emissions by the two firms it will always be better in terms of environmental damage costs and hence welfare if the firms are located in different countries.

It is difficult to provide general answers to the questions we have just posed because for the Market Share Game, the Location Game with a Single Instrument and the Location Game with Two Instruments we get different equilibria depending on the particular parameter values. However a few general comments should be noted.

(i) We have shown that when governments use two instruments in the Location Game, then an equilibrium of the Market Share Game - the Nash Equilibrium with firms located in separate countries - may be an equilibrium of the Location Game, so that where this occurs the order of moves is irrelevant.

(ii) When governments use a single instrument in the Location Game and select the equilibrium according to the maximum welfare criterion, then when the equilibrium involves governments setting h^*, since $h^* < f^*$, governments will be setting tougher environmental policies than in our co-operative benchmark; if the Market Share Game equilibrium was g^* then this would be an even tougher environmental policy (recall that $g^* < h^*$) but if the Market Share Game equilibrium was $f^N > f^* > h^*$ then this is an example where competition for market share leads to weaker environmental policies than competition for location.

(iii) Another example of how competition for market share may lead to weaker environmental policies than competition for location arises when the equilibrium of the Market Share Game is the Nash Equilibrium f^N but $W(f^N) < 0$, i.e. $f^N > \bar{f}$; in the Location Game governments need never set policies that lead to negative welfare, and we have seen that in these circumstances the equilibrium of the Location Game with Two Instruments would be the Constrained Nash Equilibrium, the Race-to-the-Bottom Equilibrium or the Profit- Constrained Race-to-the-Bottom, with equilibrium policy instruments $f^N > \bar{f} > \overline{g} > \hat{e}$.

(iv) It is also worth commenting on the comparison between the outcome of the Location Game with one or two instruments. In general we would believe that using a single instrument would be likely to encourage governments to engage in more environmental dumping, because with two instruments governments can control separately their competition to get both firms to locate in their countries from competition for market share contingent on a single firm in each country. But this need not always be the case. One example of this is that there is a range of parameter values ,$\overline{\phi}_2 \leq \phi \leq \overline{\phi}_1$, where the outcome with a single instrument will be NIMBY while the outcome with two instruments could be the Nash equilibrium or the Constrained Nash equilibrium. Further, the example in point (ii) above could equally be used to argue that there is an equilibrium of the single instrument game (h*) which is tougher than an equilibrium of the two instrument model - the Nash equilibrium. Moreover the example in (iii) can also be interpreted to say that even if the outcome of the single instrument model is the race-to-the-bottom equilibrium, $\overline{h}$, this can also be tougher than an equilibrium of the two instrument model, namely the Constrained Nash equilibrium, $\bar{f}$.

More examples could be given, but to get an appreciation of the relative importance of different types of equilibria we turn to the numerical simulations in the next section.

4.5. Simulation results.

In the previous two sections we have analysed the equilibria of the Market Share Game, the Location Game with a Single Instrument and the Location Game with Two Instruments. By comparing the equilibrium of the Market Share Game and the Location Game we want to assess how the difference in move structure affects the incentives for governments to engage in "environmental dumping" and by comparing the Location Game with one and two instruments we want to assess how the difference in the number of instruments affects incentives for "environmental dumping". We are interested also in how differences in policies are reflected in differences in welfare. However, as we argued at the end of the last section, because each game has a range of possible equilibria depending on precise parameter values, comparing the equilibria of these games is not straightforward. We provided some examples to show that there can be no general presumption that competition for location leads to laxer environmental policies than competition for market share, and indeed that for some parameter values differences in move structure may have no effect on policy instruments. Equally there can be no general presumption that using only one policy instrument causes more of a race-to-the-bottom than using two policy instruments.

To get beyond these examples and assess the relative occurrence of different equilibria for the three games and hence the relative prevalence of strategic competition in environmental policy we have taken a special case of the model set out in section 2 which allows us to conduct some numerical experiments. Firm i has an inverse demand function $P = A - x_i - \sigma x_j$ where σ, $0 < \sigma \leq 1$ is a parameter capturing the degree of substitutability (and hence intensity of market competition) between the two products. Firm i has an abatement cost function $(x_i - e_i)^2 / 2$.

Firm i's equilibrium output is $X_i(e_i, e_j) = \dfrac{(3-\sigma)A + 3e_i - \sigma e_j}{(9-\sigma^2)}$.

The *symmetric* profit function is:

$$\Pi(e) = be - ae^2 - (F - c) \quad \text{where } a = \frac{[(3+\sigma)^2 - 3]}{2(3+\sigma)^2}, b = \frac{3A}{(3+\sigma)^2}, c = \frac{3A^2}{2(3+\sigma)^2}$$

The damage cost function is given by $\delta(E) = \theta E^2$, so that the relevant *symmetric* welfare functions are:

$$W(f) = bf - (a+\theta)f^2 - \phi$$
$$V(g) = 2bg - 2(a+2\theta)g^2 - 2\phi$$
$$\overline{V}(g) = bg - (a+2\theta)g^2 - \phi$$
$$U(h) = bh - (a+1.5\theta)h^2 - \phi$$

where $\phi = F + D - c$ and, to reduce the number of parameters, we shall assume that $F = D$. It is readily shown that the five critical values of ϕ are:

$$\overline{\phi}_1 = \frac{b^2}{4(a+\theta)}, \overline{\phi}_2 = \frac{b^2}{4(a+1.5\theta)}, \overline{\phi}_3 = \frac{b^2}{4(a+2\theta)}, \overline{\phi}_4 = \frac{b^2}{4(a+2.5\theta)}, \overline{\phi}_5 = \frac{b^2}{4(a+3\theta)}$$

and we can also readily calculate the key emission limits:

$$\hat{e} = \frac{b}{2a}; f^C = f^* = \frac{b}{2(a+\theta)}; h^* = \frac{b}{2(a+1.5\theta)}; g^* = \frac{b}{2(a+2\theta)};$$

$$\overline{f} = \frac{b+\sqrt{b^2 - 4(a+\theta)\phi}}{2(a+\theta)}; \overline{h} = \frac{b+\sqrt{b^2 - 4(a+1.5\theta)\phi}}{2(a+1.5\theta)}; \overline{g} = \frac{b+\sqrt{b^2 - 4(a+2\theta)\phi}}{2(a+2\theta)};$$

$$f^N = \frac{9A}{[(3-\sigma)(3+\sigma)^2(2\theta+1) - 9]}; \quad \varepsilon = \max\{0, \frac{b - \sqrt{b^2 - 2a(\phi - c)}}{2a}\}$$

$$\tilde{f} = \frac{b+\sqrt{b^2 - 4(a+3\theta)\phi}}{2(a+3\theta)}; \quad \tilde{h} = \frac{b+\sqrt{b^2 - 4(a+2.5\theta)\phi}}{2(a+2.5\theta)}$$

There are three key parameters in this model: σ, θ, and ϕ. We have used values of

σ = 0.1, ...1.0 in steps of 0.1. The more significant parameters for determining which equilibrium occurs are the intensity of damage cost parameter θ, and the 'sunk cost' parameter ϕ. We have set an upper bound on θ of $\overline{\theta} = a$, and then taken values of θ between 0 and $\overline{\theta}$ in steps of $0.01\overline{\theta}$. For ϕ we have set an upper bound $\overline{\phi} = c + \frac{b^2}{2a}$ which ensures that $\Pi(\hat{e}) \geq 0$ and searched over values of ϕ between 0 and $\overline{\phi}$ in steps of $0.01\overline{\phi}$. However, since we know that for $\phi \geq \overline{\phi}_1$ all (expected) welfare functions are non-positive, and hence our benchmark level of welfare *W(f*)* would be non-positive, for any given θ we have confined attention to values of $\phi < \overline{\phi}_1$.

Thus for any given θ, $0 \leq \theta \leq \overline{\theta}$ and $\phi, 0 \leq \phi \leq \overline{\phi}_1$ we first calculate the equilibrium of the Market Share Game (MSG), the equilibrium of the Location Game with One Instrument (LGOI), and the equilibrium of the Location Game with Two Instruments (LGTI), where for the LGOI (and hence the LGTI) we use the three different selection criteria - Maximum Welfare (MW), Race-to-the-Bottom (RttB) and the Average of MW and RttB (Ave). If we denote the equilibrium emission limit and welfare by e''' and W''' we then calculate two measures of the deviation of this equilibrium from the benchmark emission limit and welfare (*f** and *W(f*)*): the proportionate increase in emission limit and the proportionate reduction in welfare, defined as:

$$\mu \equiv (f^* - e''')/f^*; \omega \equiv (W(f^*) - W''')/W(f^*).$$

For each game we can then define the average values of μ and ω over all values of θ and ϕ, which we denote by $\overline{\mu}$ and $\overline{\omega}$ respectively.

We begin by discussing the relative frequency of the different equilibria as parameters vary. We begin with the Market Share Game. As discussed in Section 4.3, there are two possible equilibria - where firms locate in a single country and the government of that country sets emission limits g^*, and where firms locate in different countries and the two governments set the Nash Equilibrium emission limits f^N; the first equilibrium occurs for low values of θ, the latter for high values of θ. Table 4.1 shows the proportion of (θ, ϕ) parameter space for which the two equilibria occur for different values of the substitution parameter σ. The Nash Equilibrium is by far the dominant equilibrium, and this increases as σ falls.

The reason is that while the Nash equilibrium gives higher emission limits than in the single country case, it also encourages firms to produce too much output, which reduces profits; this 'overcompetition' gets reduced as σ falls and the products of the two firms become more independent of each other.

We now turn to the Location Game. When governments use two distinct instruments, then there are five possible equilibria - the Nash Equilibrium, f^N, the Constrained Nash Equilibrium, $\bar{f}$, the Race-to-the-Bottom Equilibrium, $\bar{g}$, the Profit-Constrained Race-to-the-Bottom Equilibrium, $\hat{e}$, and the Intermediate Equilibrium, $\tilde{f}$. When the government uses a single instrument, then the equilibria depend on the equilibrium selection criterion; with Maximum Welfare the possible equilibria are NIMBY, ε, the MEW Equilibrium, h^*, and the Intermediate Equilibrium, $\tilde{h}$. The equilibria of the single instrument model are also possible equilibria when the governments use two instruments. When the equilibrium selection is Race-to-the-Bottom then the only equilibria with a single instrument are the NIMBY and the Race-to-the-Bottom Equilibrium, $\bar{h}$. Finally when the selection criterion is the Average, then the possible equilibria are the NIMBY, the average of MEW and RttB, $(h^* + \bar{h})/2$, and the average of Intermediate and RttB, $(\tilde{h} + \bar{h})/2$. Tables 4.2, 4.3 and 4.4 present the proportions of (θ, ϕ) parameter space in which the different equilibria occur for the Location Game with one and two instruments using, respectively, the MW, Ave. and RttB selection criteria.

We note the following points:

(a) With a single instrument the relative proportions of parameter space for the different equilibria are independent of the value of σ and of the selection criterion. It is not surprising that the NIMBY equilibrium is the same fraction of all parameter space, given the way we constructed $\bar{\phi}_1$ and $\bar{\phi}_2$, and obviously the proportions of parameter space involving the MEW and Intermediate equilibria will be same using the MW and Ave criteria. However, we have no immediate intuition why the proportions of MEW and Intermediate equilibria should be independent of σ.

(b) Since the MW criterion gives the highest welfare for the single instrument case and the RttB criterion the least, it is not surprising that when governments have two instruments a higher proportion of the single instrument equilibria remain equilibria in the two instrument case under the MW criterion for the single instrument than with either the Ave or RttB criteria. This is most marked for σ =1, and by the time σ gets below 0.7 the proportion of single instrument equilibria which remain as equilibria with two instruments is less than 20% even with MW criterion. When σ gets below 0.5 then single instrument equilibria become negligible with two instruments and the proportion of equilibria with two instruments is independent of how the single instrument equilibria are selected.

(c) A striking result is the predominance of the Nash equilibrium with two instruments - accounting for between 50% and 70% of equilibria for almost all values of σ and all criteria. Correspondingly with two instruments we almost never get the race-to-the-bottom equilibrium with both firms locating in the same country. With the single instrument model and RttB criterion we get RttB equilibria in about 90% of cases. This confirms the importance we have given to separating the issue of the move structure from the issue of the number of instruments. When we have the same number of instruments as in the Market Share Game what we have seen is that for the majority of cases the Move Structure is irrelevant, and the prevalence of race-to-the-bottom outcomes is significantly reduced. Even if we treat the Constrained Nash Equilibrium as a form of race-to-the-bottom equilibrium our conclusion that race-to-the-bottom equilibria are negligible when two instruments are used remains valid.

Finally we turn to Table 4.5 which presents the results for average proportionate increase in emission limits (proportionate weakening of environmental policies), $\overline{\mu}$, and the proportionate welfare loss, $\overline{\omega}$, for the different games and different values of σ. We discuss first the variations within each game with respect to σ and then compare the results across games. For the MSG, the striking result is that both $\overline{\mu}$ and $\overline{\omega}$ decline sharply as σ declines; this is what we would expect since a reduction in the degree of substitutability between products reduces the incentive for strategic competition between governments in the Nash Equilibrium. For LGSI, again the striking result is that the percentage increase in emission limits is independent of σ ; to achieve this requires two factors: for a given kind of equilibrium (e.g. Intermediate, the proportionate increase in emission limits must be independent of σ, which reflects the simple functional forms we have used; and the proportion of different kinds of equilibria must be independent of σ, which we have already seen is correct. However, despite the fact that the proportionate increase in emission limits is independent of σ , the proportionate reduction in welfare increases as σ falls. This reflects the particular welfare functions we have employed[7]. For LGTI, in general as σ falls both $\bar{\mu}$ and $\bar{\omega}$ first increase and then fall. The initial increase reflects the reduction in the proportion of equilibria from the LGSI such as NIMBY and MEW which involve toughening environmental policies relative to f^* ; the subsequent fall reflects the growing dominance of the Nash equilibrium as σ declines, and hence mirrors the results for the MSG.

Turning to a comparison of the different games, we begin by comparing the MSG and LG to assess how move structure affects $\bar{\mu}$ and $\bar{\omega}$. The key result is that for high values of σ, MSG involves a greater proportionate increase in emission limits than does LG; i.e. when products are close substitutes competition for market share is more intense than competition for location in terms of its impact on environmental dumping. For low values of σ this result is reversed. Thus averaging over all parameter values, there is no presumption that competition for location is more intense than competition for market share. Similar results apply to comparisons of welfare losses between the MSG and LG.

---

[7] It can be argued that making welfare comparisons across different values of σ is not sensible since we are using different welfare functions. However, given that we wish to construct this measure, there is a technical problem that because we have calculated the proportionate reduction in welfare from $W(f^*)$, and because for some parameter values $W(f^*)$ becomes close to zero, this caused problems for computing $\bar{\omega}$. To overcome this we added a constant to the welfare function, which of course does not affect any substantive result. However it does mean that while in the absence of the constant a proportionate reduction in f^* which was independent of σ would have led to a proportionate reduction in $W(f^*)$ which was independent of σ the addition of the constant gives the result we see in Table 4.5.

Comparing different forms of LG, not surprisingly the proportionate increase in emission limits and proportionate reduction in welfare gets greater as we move from the MW to the Ave to the RttB criteria for the selection of the equilibrium with a single instrument.

Finally we compare the results of using one and two instruments in the Location Game. In terms of the proportionate increase in emission limits, not surprisingly the result depends crucially on which equilibrium selection criterion we use for the single instrument model. With the MW criterion for almost all values of σ there is more environmental dumping with two instruments than with one; while with the RttB there is more environmental dumping with one instrument than two for all values of σ ; the Ave criterion yields more environmental dumping with one instrument than two except for two values of σ . If we take the Ave criterion as being more representative of what might happen with a single instrument, then on balance a single instrument leads to more environmental dumping than two instruments. This accords with our general rationale for using two instruments - that it separates the competition for location of both firms from competition for market share when firms locate in a single country and the latter provides a floor to the extent of race-to-the-bottom that might otherwise occur. Whatever the degree of environmental dumping, the use of a single instrument always leads to greater welfare losses than the use of two instruments. The rationale for this relates to our previous comment. For the majority of parameter values there is positive welfare to be earned from having a firm locate in each country, and for almost all equilibria with two instruments this is the outcome that prevails. However when a single instrument is used for all equilibria other than NIMBY there is only a 50% chance of having one firm locate in each country and a 25% chance that no firm at all will locate in a country. As we have seen from the properties of the welfare functions for all parameter values the *expected* welfare with a single instrument *(U(f))* is lower than the certain welfare (*W(f))* each country would get when each firm locates in a separate country for all values of *f*, which is the outcome with two instruments for most parameter values; so even if the use of a single instrument or two instruments leads to the same equilibrium emission limits, welfare would be lower with one instrument than two. This again reflects the fact with two instruments one can ensure that both firms will locate in separate countries.

4.6. Conclusions.

The results reported in the last section confirm the two key points we have been making in this paper - that there is no general presumption that competition for location leads to greater environmental dumping than does competition for market share with fixed locations and that in comparing the two forms of competition it is important not to compound a change in move structure with a change in the number of instruments available to governments. Allowing governments to condition their policies on the number of firms that locate in their countries will in general lead to a reduction in environmental dumping for location games and will certainly improve welfare.

Obviously the model we have used is extremely simple - we have omitted considerations of consumer welfare, of more sophisticated treatment of distribution of profits, and the use of other instruments such as emission taxes which generate government revenue. These would all change the welfare functions we have used and hence might change the detailed results but we do not believe that they would change some of the key rankings of welfare functions which derived simply from the assumption of increasing marginal damage costs and hence the extra welfare cost of having more firms emitting pollution located in one country. We have also assumed symmetry and that the number of firms equals the number of countries but again we do not believe these assumptions change the basic message about the importance of modelling carefully the instruments available to governments.

There are two directions in which we think more work is needed. The first is to move to a proper dynamic analysis where location and environmental policy decisions are taken over an infinite time horizon and one models differences in degrees of commitment by governments and firms. Work by Feenstra (1998) in related but not identical contexts shows that the results from dynamic games can be significantly different from those in simple multi-stage games that we have employed here. Second, we are not aware of governments precommitting environmental policies conditional on the number of firms that might locate in their countries and we would like to explore what constraints there might be which would prevent governments using such policies, given that we have shown the benefits of such an approach.

APPENDIX A - ANALYSIS OF SECOND-STAGE LOCATION GAME WITH TWO INSTRUMENTS

The two firms take as given the emission limits set by the governments of the two countries at stage one. These emission limits are contingent on the number of firms that locate in a country, so we denote by $f_{c,}$ g_c the emission limits set in country c when respectively one and two firms locate in country c, $c = a$, b. The two firms now play a location game in which, for example, if firm 1 locates in country a while firm 2 locates in country b the payoff to firm 1, ignoring fixed costs, is $\Pi(f_a, f_b)$, while if both firms locate in country a each firm would get a payoff $\Pi(g_a, g_a)$. We can construct the rest of the payoffs in a similar fashion. To analyse the Nash equilibria of this game, it will ease notation to define the following:

$$\Pi'_f \equiv \min[\Pi(f_a, f_b), \Pi(f_b, f_a)] \quad \Pi''_f \equiv \max[\Pi(f_a, f_b), \Pi(f_b, f_a)]$$
$$\Pi'_g \equiv \min[\Pi(g_a, g_a), \Pi(g_b, g_b)] \quad \Pi''_g \equiv \max[\Pi(g_a, g_a), \Pi(g_b, g_b)]$$

We assess the possible equilibria in four parts.

Part A. We begin by assuming that the two countries set different emission limits, so the above maxima and minima are distinct. It turns out that there are three cases to distinguish:

I. $\Pi''_g \leq \Pi'_f < \Pi''_f$ In this case the equilibrium is that both firms locate in different countries. There are two Nash equilibria depending on which firm locates in which country; to resolve this we shall retain our assumption from the Market Share Game that firm 1 has first mover advantage, so the equilibrium will be that both firms locate in different countries with firm 1 locating in the country with higher profits.

II. $\Pi'_f < \Pi''_f \leq \Pi''_g$ The equilibrium is that both firms locate in the same country. There is either a unique Nash equilibrium with both firms locating in the country with higher profits, or two Nash equilibria with both firms locating in either country a or country b. If we resolve the latter by allowing firm 1 to move first, then there is a unique equilibrium in which both firms locate in the country with the higher profits.

III. $\Pi'_f < \Pi''_g < \Pi''_f$ There are a number of possible equilibria. First, there could be two Nash equilibria where the two firms locate in different countries; again we can resolve this by allowing firm 1 to move first, so that this equilibrium involves both firms locating in different countries, with firm 1 locating in the country with higher profits. A second possibility is that there is a unique equilibrium in which both firms locate in the same country, that country being the one which offers higher profits. The third possibility is a unique equilibrium of a Prisoner's Dilemma nature in which both firms locate in the same country, but the one which offers *lower* profits.

Part B. We now suppose that $f_a \neq f_b$ but $g_a = g_b \Rightarrow \Pi'_g = \Pi''_g = \Pi_g$. Then the equilibria can again be classified using the three cases set out above, except that $\Pi''_g = \Pi_g$. In case I, we have the same outcome as above; in case II there are again two equilbria with both firms locating in the same country, although the firms are indifferent about which country that is; we solve the coordination problem by allowing firm 1 to locate first (choosing whether to locate in a or b with equal probability); finally in case III we now get a unique equilibrium with both firms locating in the same country.

Part C. We now suppose that $g_a \neq g_b$ but $f_a = f_b \Rightarrow \Pi'_f = \Pi''_f = \Pi_f$. There are three cases:

I. $\Pi'_g < \Pi''_g \leq \Pi_f$ In this case there are two equilibria with both firms locating in different countries, but the firms being indifferent where they locate; we solve the coordination problem by again having firm 1 locate first, randomising where to locate, with firm 2 locating in the other country.

II. $\Pi'_g < \Pi_f < \Pi''_g$ There is a unique equilibrium with both firms locating in the same country, the one which offers higher profits.

III. $\Pi_f \leq \Pi'_g < \Pi''_g$ There are two equilibria in which both firms locate in the same country; assuming that firm1 locates first reduces this to the same outcome as case II.

Part D. Finally we suppose that $f_a = f_b$, $g_a = g_b$. There are now just two cases:

I. $\Pi_f > \Pi_g$ There is a single equilibrium in which both firms locate in separate countries.

II. $\Pi_f < \Pi_g$ There is a single equilibrium in which both firms locate in the same country. Again we assume that firm 1 moves first, randomising in which country it locates.

Table 4.1 - Equilibria in market share game

σ	g^*	f^N
1.0	0.32	0.68
0.9	0.27	0.73
0.8	0.24	0.76
0.7	0.20	0.80
0.6	0.27	0.83
0.5	0.14	0.86
0.4	0.11	0.89
0.3	0.08	0.92
0.2	0.06	0.94
0.1	0.03	0.97

Table 4.2 -Equilibria for location game -max. welfare

	SINGLE INST.			TWO INSTRUMENTS							
σ	ε	h^*	$\tilde{h}$	ε	h^*	$\tilde{h}$	f^N	$\bar{f}$	$\bar{g}$	$\hat{e}$	$\tilde{f}$
1.0	0.11	0.20	0.69	0.08	0.19	0.43	0.04	0.04	0.00	0.23	0.00
0.9	0.11	0.20	0.69	0.07	0.19	0.30	0.19	0.04	0.00	0.21	0.00
0.8	0.11	0.20	0.69	0.10	0.17	0.15	0.38	0.01	0.00	0.19	0.00
0.7	0.11	0.20	0.69	0.05	0.08	0.07	0.56	0.05	0.00	0.17	0.02
0.6	0.11	0.20	0.69	0.06	0.04	0.03	0.65	0.02	0.00	0.16	0.05
0.5	0.11	0.20	0.69	0.04	0.02	0.01	0.68	0.01	0.00	0.15	0.08
0.4	0.11	0.20	0.69	0.02	0.01	0.00	0.68	0.02	0.00	0.15	0.12
0.3	0.11	0.20	0.69	0.01	0.00	0.00	0.66	0.01	0.00	0.15	0.17
0.2	0.11	0.20	0.69	0.01	0.00	0.00	0.63	0.00	0.00	0.15	0.21
0.1	0.11	0.20	0.69	0.00	0.00	0.00	0.59	0.00	0.00	0.15	0.26

Table 4.3 - Equilibria of location game - average

	SINGLE INST.			TWO INSTRUMENTS							
σ	ε	h^*	$\tilde{h}$	ε	h^*	$\tilde{h}$	f^N	$\bar{f}$	$\bar{g}$	$\hat{e}$	$\tilde{f}$
1.0	0.11	0.20	0.69	0.08	0.19	0.18	0.26	0.04	0.00	0.26	0.00
0.9	0.11	0.20	0.69	0.07	0.19	0.09	0.37	0.04	0.00	0.24	0.00
0.8	0.11	0.20	0.69	0.10	0.11	0.04	0.53	0.01	0.00	0.20	0.00
0.7	0.11	0.20	0.69	0.05	0.06	0.02	0.63	0.05	0.00	0.18	0.02
0.6	0.11	0.20	0.69	0.06	0.03	0.01	0.68	0.02	0.00	0.16	0.05
0.5	0.11	0.20	0.69	0.04	0.01	0.00	0.69	0.01	0.00	0.16	0.08
0.4	0.11	0.20	0.69	0.02	0.00	0.00	0.68	0.02	0.00	0.15	0.12
0.3	0.11	0.20	0.69	0.01	0.00	0.00	0.66	0.01	0.00	0.15	0.17
0.2	0.11	0.20	0.69	0.01	0.00	0.00	0.63	0.00	0.00	0.15	0.21
0.1	0.11	0.20	0.69	0.00	0.00	0.00	0.59	0.00	0.00	0.15	0.26

Table 4.4 - Equilibria of location game - race-to-the-bottom

	ONE INST.		TWO INSTRUMENTS						
σ	ε	$\bar{h}$	ε	$\bar{h}$	f^N	$\bar{f}$	$\bar{g}$	$\hat{e}$	$\tilde{f}$
1.0	0.11	0.89	0.08	0.10	0.41	0.10	0.01	0.29	0.00
0.9	0.11	0.89	0.07	0.07	0.52	0.08	0.01	0.25	0.00
0.8	0.11	0.89	0.10	0.04	0.61	0.04	0.00	0.21	0.00
0.7	0.11	0.89	0.05	0.02	0.67	0.06	0.00	0.19	0.02
0.6	0.11	0.89	0.06	0.01	0.70	0.02	0.00	0.17	0.05
0.5	0.11	0.89	0.04	0.00	0.70	0.02	0.00	0.16	0.08
0.4	0.11	0.89	0.02	0.00	0.69	0.02	0.00	0.15	0.12
0.3	0.11	0.89	0.01	0.00	0.66	0.01	0.00	0.15	0.17
0.2	0.11	0.89	0.01	0.00	0.63	0.00	0.00	0.15	0.21
0.1	0.11	0.89	0.00	0.00	0.59	0.00	0.00	0.15	0.26

Table 4.5 Average increase in emission limits and proportionate welfare losses

	Market Share Game		Location Game - Max. Welfare				Location Game - Average				Location Game - Race-to-the-Bottom			
			One Inst.		Two Insts.		One Inst.		Two Insts.		One Inst.		Two Insts.	
σ	$\bar{\mu}$	$\bar{\omega}$	$\bar{\mu}$	$\bar{\omega}$	$\bar{\mu}$	$\bar{\omega}$	$\bar{\mu}$	$\bar{\omega}$	$\bar{\mu}$	$\bar{\omega}$	$\bar{\mu}$	$\bar{\omega}$	$\bar{\mu}$	$\bar{\omega}$
1.0	0.393	0.256	0.053	0.205	0.007	0.144	0.187	0.282	0.163	0.179	0.321	0.403	0.276	0.208
0.9	0.367	0.220	0.052	0.217	0.055	0.145	0.186	0.299	0.181	0.169	0.320	0.427	0.270	0.186
0.8	0.331	0.182	0.052	0.230	0.080	0.140	0.186	0.316	0.174	0.152	0.320	0.451	0.215	0.160
0.7	0.295	0.147	0.051	0.243	0.182	0.125	0.185	0.334	0.220	0.130	0.319	0.477	0.236	0.134
0.6	0.253	0.113	0.050	0.257	0.171	0.105	0.184	0.353	0.185	0.107	0.318	0.504	0.191	0.108
0.5	0.216	0.083	0.052	0.271	0.173	0.084	0.186	0.372	0.177	0.085	0.319	0.531	0.179	0.085
0.4	0.173	0.056	0.048	0.286	0.169	0.065	0.182	0.391	0.170	0.065	0.316	0.559	0.170	0.066
0.3	0.131	0.033	0.051	0.302	0.139	0.050	0.185	0.412	0.139	0.050	0.319	0.588	0.139	0.050
0.2	0.088	0.016	0.050	0.317	0.112	0.039	0.183	0.432	0.112	0.039	0.317	0.618	0.112	0.039
0.1	0.045	0.004	0.047	0.333	0.088	0.033	0.180	0.453	0.088	0.033	0.313	0.646	0.088	0.033

Welfare Functions: One Firm

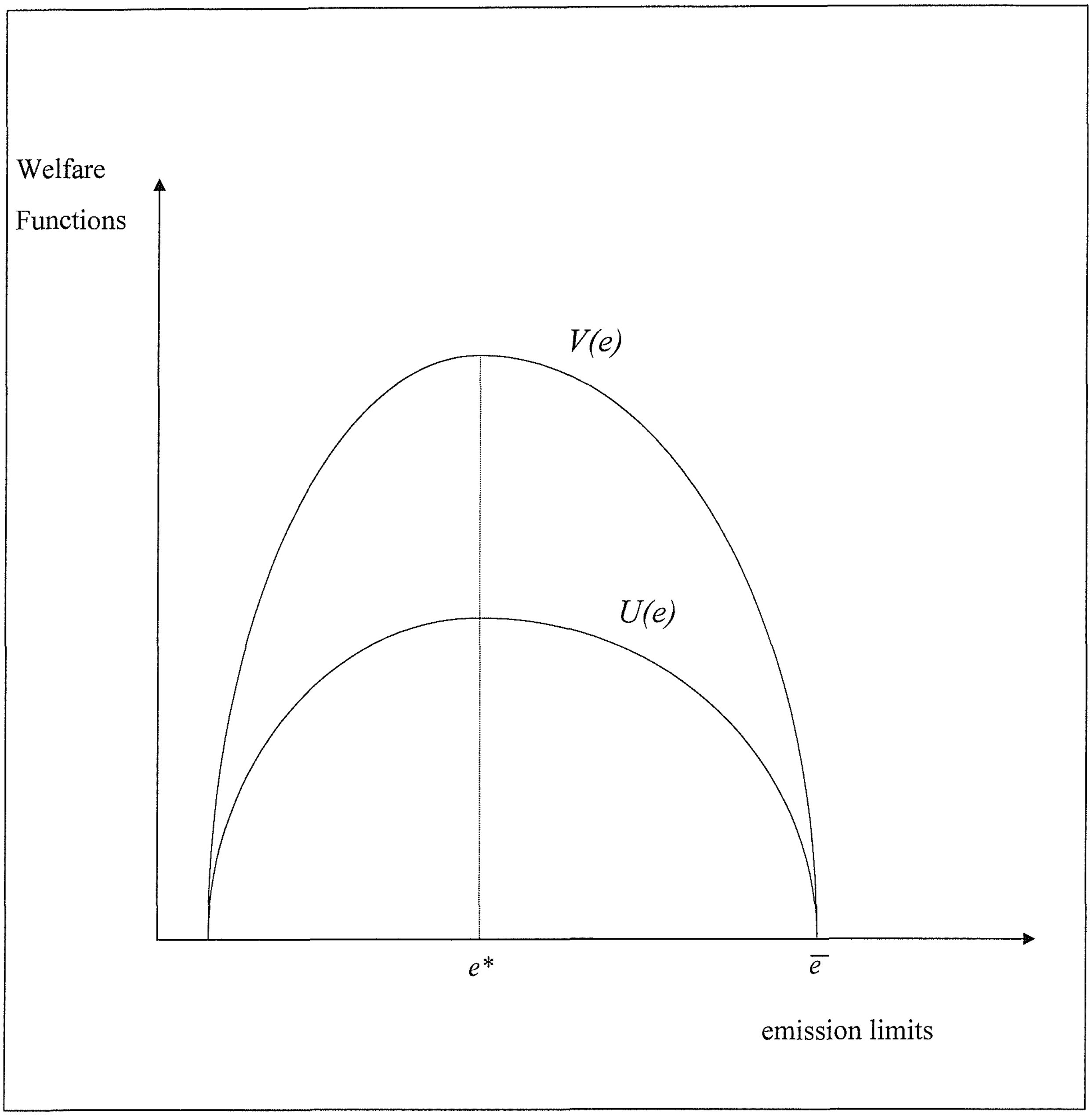

Fig. 4.1

Welfare Functions: More than One Firm

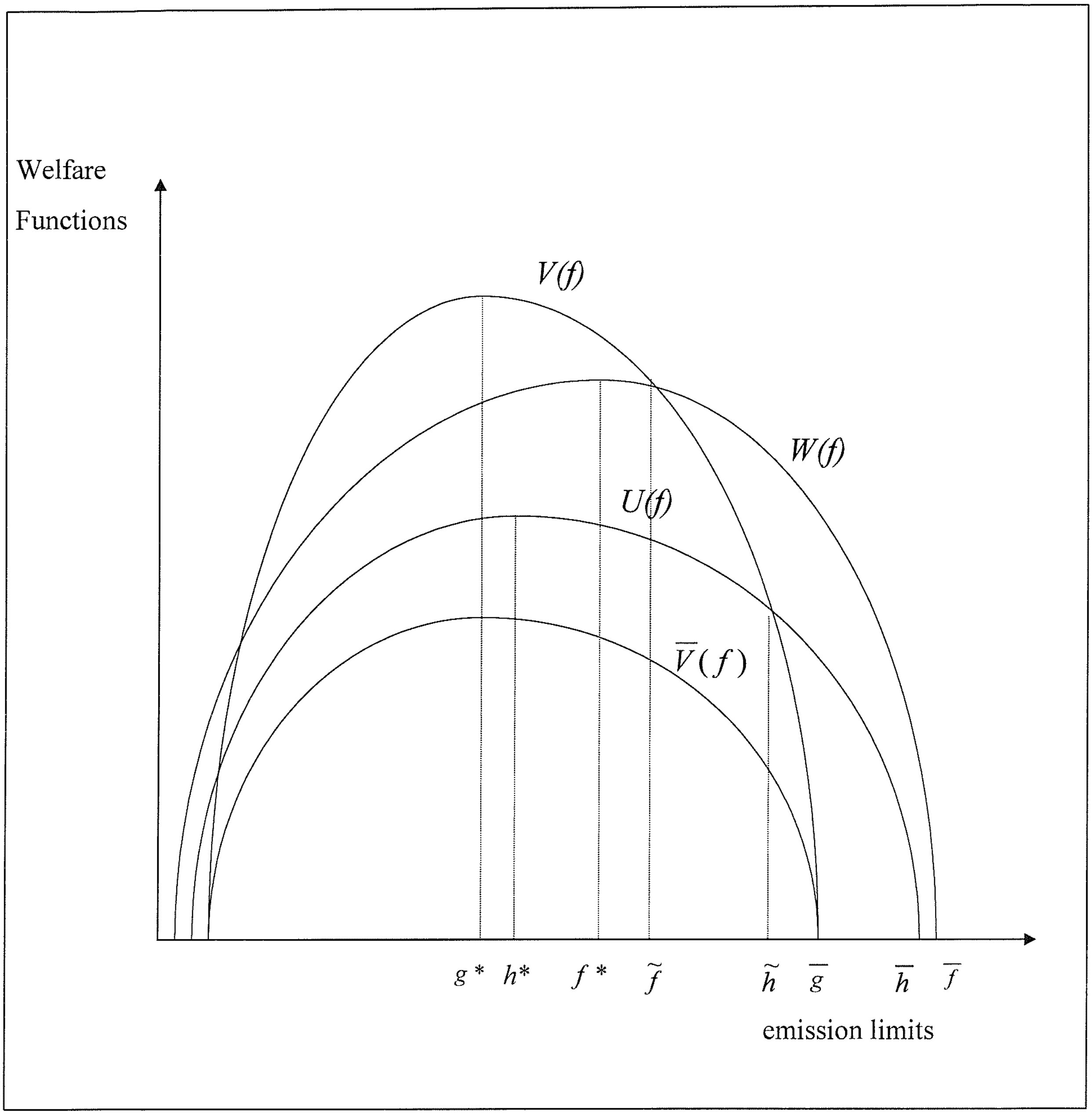

Fig. 4.2

Equilibria of Location Game with a Single Instrument

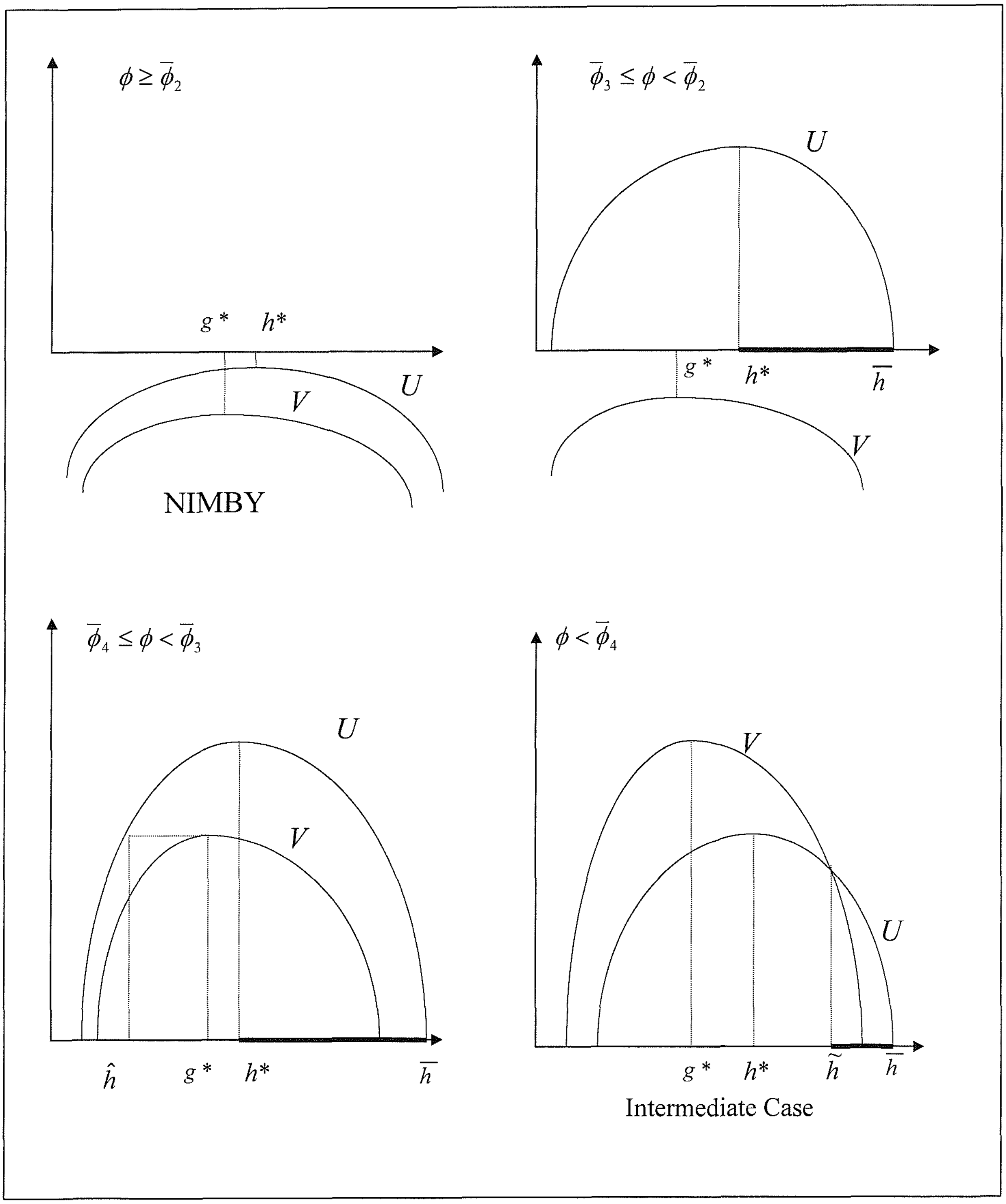

Fig. 4.3

Equilibria of the Location Game for Two Intruments

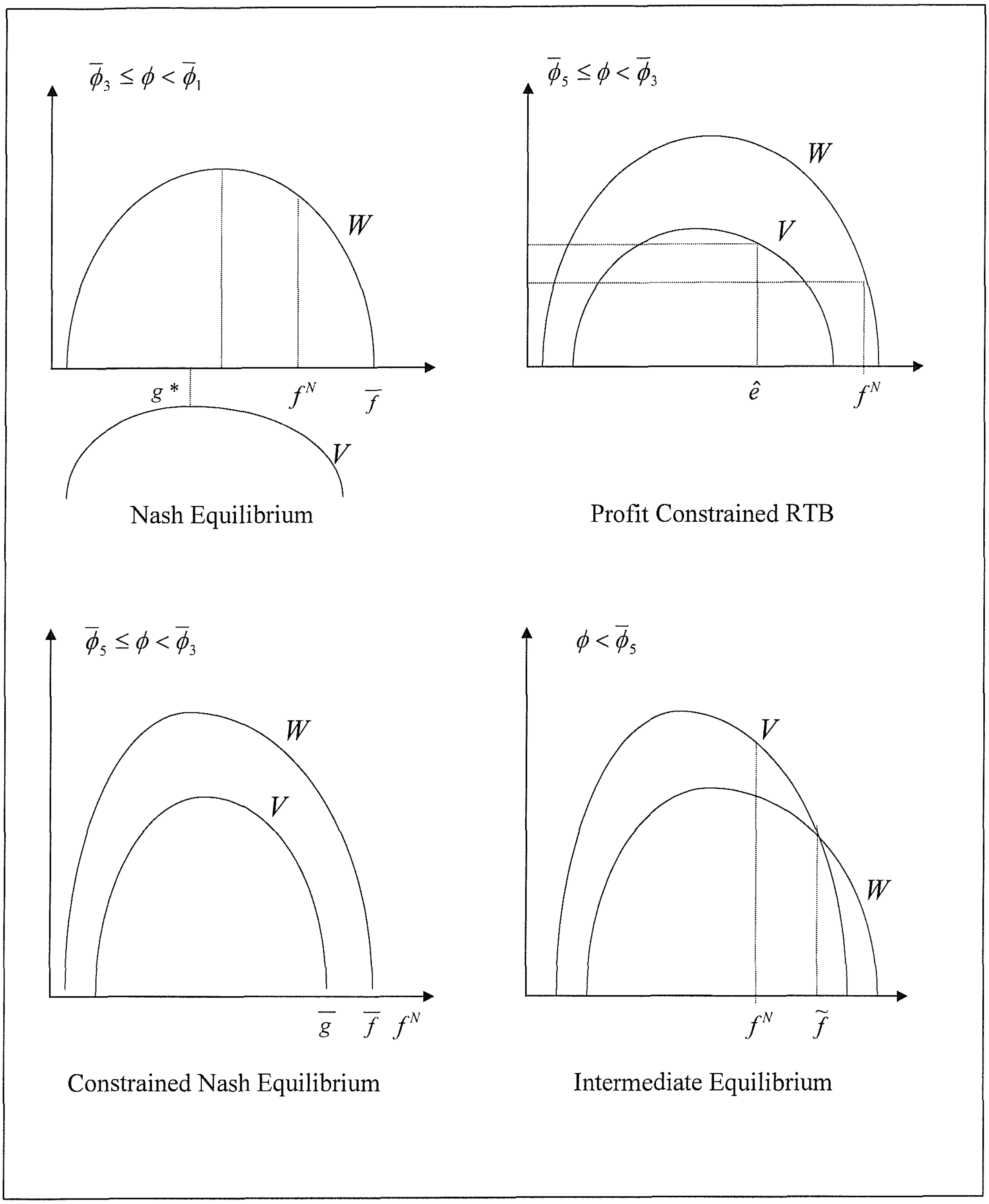

Fig. 4.4

Chapter 5

Credit rationing and capital structure: a survey of empirical studies

Abstract

This survey is part of a project on The Financial Implications of Environmental Legislation. The purpose of it is to see how the problem of credit rationing, according to borrowers' characteristics, has been empirically estimated in some literature. The suvey considers first allocation of credit from the lenders' side, and second demand of credit from the borrowers' side. Both commercial and consumers' loan markets have been considered. For commercial loan the availability of data made it possible to build up more appropriate models, providing a positive answer to the question whether borrowers are discriminated according to their characteristics. The same has not been possible for the commercial loan market, mainly because data on this are scarce. This made modelling of firms' and banks' behaviour more unsatisfactory

5.1. Introduction

This survey is part of a research project on the financial consequences of environmental legislation. The survey will thus be oriented with this ultimate goal in mind, that is to see if and how credit is allocated or rationed according to firms' environmental riskiness.

The quantity of commercial loan is the equilibrium result of the profit-maximising behaviour of firms and banks. For the supply there are two aspects to consider: the first is the total supply of credit and the second is the allocation of credit among firms. The total amount of credit is determined by banks characteristics like their marginal costs of deposit, the interest they are going to charge, by macroeconomic variables or conditions in general and by monetary policies more specifically. Most of the studies are related to this aspect (Sealey (1979)). The second aspect is more related to firms' characteristics, according to which banks discriminate among them in order to decide to whom they should lend. For the banks it is in general quite difficult to observe firms' characteristics as they are usually not observable. Melnik and Plaut (1986) however manage to model how banks discriminate firms according to their characteristics, with particular regard to the allocation of loan commitments. The role of customers' characteristics in the allocation of credit will be analysed mainly in the consumers' loan market, because there is more availability of data.

On the other side the demand for credit depends on firms' preferences which are also strongly related to their characteristics. More specifically, decisions concerning capital structure are based on firms' value maximisation, and usually the decisions on borrowing or internal financing or other sources will be determined, along with other factors, by the costs associated with the different financing forms. In turn these costs depend very much on firms' characteristics, for example it is very likely that a bank will charge a lower interest rate to an old customer rather than to a new one. Or the decision to borrow privately or publicly will depend on issuance costs: the larger is the firm the higher is the amount of the debt that can be issued publicly and thus the fixed issuance costs can be spread over a large volume. This matter has recently been analysed by some authors (Titman-Wessels (1988), Johnson (1997), Easterwood-Kadapakkam (1991), James-Wier (1990), Houston-James (1996) and Petersen-Rajan (1994)).

It is possible, however, that the credit market is in fact not in equilibrium and the quantity of credit observed at any time is the lowest between the demand and the supply of credit at the ongoing interest rate (and usually the reason for not adjusting the interest rate is institutional). Jaffee (1971)suggests the use of disequilibrium models for this. Sealey (1979) uses this approach to study the commercial loan market, while Avery (1981), Cox-Jappelli (1993), Jappelli (1990) and Perraudin-Sorensen (1992) analyse the consumers' loan market. They are able to report also consumers' characteristics in their analysis., while Sealey (1979) takes into account more macro-economic variables.

The fact that interest rates do not adjust instantaneously is usually given to explain rationing, Berger and Udell (1992) doubt instead that credit rationing is always associated with rates stickiness, while they find evidence that the latter is more related to contractual terms.

With regard to interest rates stickiness it must also be observed that usually this works in the sense that it is not possible to increase the interest rate above a given ceiling, while it is always possible to decrease it (along the same supply curve). This means that if there is a disequilibrium (always given the same supply and demand curves) this is in the sense that there is excess demand (and thus credit rationing). If there is a decrease in the interest rate it usually means that there has been a change in the monetary policy, and this is going to affect the supply (and demand) of credit by shifting the supply curve. This matter is relevant for estimation purposes because rather than a problem of attribution of the observed quantity to the demand or to the supply schedule there is a problem of identifying the changes in the two curves.

All these papers fail in taking simultaneously into account both demand and supply in the determination of the equilibrium quantity of credit. It must be observed however that there are objective difficulties for this. Also, as most of the studies are cross section analysis it must be reported that the ordering of the data is not indifferent for the results.

The survey is organised in the following way:

In the first two sections the supply of credit and its allocation according to firms' characteristics is considered. The first part describes the usual disequilibrium models framework, while the second reviews some empirical models on this subject. The third section considers empirical studies in the consumers' loan market. The last two sections will be devoted to analyse how firms choose the sources of their borrowing according to their own preferences, relating the proportions of debt to their own attributes. In the first of these two I describe firms' characteristics, how they can be measured and what are the relationships with debt sources. In the second I consider again some empirical studies.

5.2 Credit supply and firms' characteristics

The literature on the supply of credit is not too homogeneous, so it is difficult to build a common framework on it. I will subsequently present the papers which in my view are more related to the topic we are directly interested in. Even though our main concern is credit rationing some of these papers are more about commitments or collaterals rather than credit rationing. Their inclusion is motivated by the fact that they are interesting either for the methodology used or for the relationships with firms' characteristics. The selection of papers is also restricted by the fact that we are not interested in analysing whether there is credit rationing or not, but rather on the relationships between bank lending and firms' characteristics.

It is traditional to approach credit rationing by using disequilibrium models (as by definition credit rationing implies excess demand at the prevailing interest rate) of the kind:

$$D_t = \alpha_1 X_t + \beta_1 P_t + u^D \tag{1}$$

$$S_t = \alpha_2 X_t + \beta_2 P_t + u^S \tag{2}$$

and

$$Q_t = \min\{S_t, D_t\} \tag{3}$$

where Q_t is the quantity really observed at any time t.

In equilibrium $Q_t=D_t=S_t$ and equations (1) and (2) form a system of simultaneous equations in Q and P, which can be estimated by the usual methods. This is however rarely the case, and it creates a major empirical problem since at any time only the effective quantity of granted loans, Q_t, is actually observed, while there are no observations on S and D. Or better, at any time Q_t will represent either D or S, whichever is the smallest, and simultaneous equations methods cannot be used. As Maddala (1983) points out, even knowing which observations belong to the demand function and which to the supply function, OLS cannot be used because the residuals of the truncated series will not have zero mean and, more important, they are also correlated with the explanatory variables (in particular the price).

In case of disequilibrium Jaffee (1971) suggests some alternatives of which I summarise two:

i) use the change in price as an indicator of excess demand or supply:

$\Delta P > 0$ as D>S (4.a)

$\Delta P < 0$ as D<S (4.b)

the sign of ΔP can then be used to identify two groups: one in which Q belongs to the demand schedule and one in which Q belongs to the supply schedule. When ΔP=0, Q can be attributed to both groups for estimation. As said previously OLS methods applied separately to each group give inconsistent estimates of the parameters.

i) assume that the change in price is directly proportional to the excess demand:

$$\Delta P = \gamma(D - S) + \varepsilon, \qquad 0 \leq \gamma \leq \infty \qquad (5)^{1}$$

and use the change in price as a quantitative surrogate for excess demand. This allows both D and S to be estimated over the entire sample.

Equations (1), (2) and (5) form a system of simultaneous equations in D, S and P which can be estimated using Full Information Maximum Likelihood. Later some studies that use this approach will be examined. It must be noticed however that in the case of a credit rationing there are not uncertainties in assigning the observed value of loans to the supply function as it is sure that this is the smallest (relative to the demand). Moreover it is extremely rare that there is excess supply, since there are no reasons why interest rates shouldn't adjust (decrease), in particular as these models usually assume that banks are perfectly competitive. The argument about non-zero-mean errors still holds.

Being on the supply side, reasons other than borrowers characteristics are also important in determining the allocation of credit and in particular the possibility of a credit rationing. Variables strictly related to the supply could be marginal costs (of

---

[1] In fact ΔP is not really proportional to excess demand because the error term has been introduced. It is also possible to have ΔP_{t+1} instead of ΔP_t, and to have two equations for the cases in which D>S and D<S.

deposits, for example), institutional constraints (capital requirements), macroeconomic policy instruments or conditions.

5.3. Some empirical studies

Sealey (1979) uses the approach described above to estimate a disequilibrium model of the commercial loan market. He specifies a demand and a supply equations where desired demand depends on the commercial loan interest rate and on the rate of alternative sources of external financing (corporate bond rate), lagged unpaid profits (which stands for internal financing) and the lagged index of industrial production (for the expectations on future activity[2]). Similarly, the supply function depends on the commercial loan rate and on the treasury bill rate (as alternative assets), total bank deposits, costs for a dollar of deposits and the lagged index of industrial production. These two equations are estimated together with the switching regression (5) by FIML. The sample goes from 1952 to 1977, and the interest rates are not seasonally adjusted, while the other variables are. Sealey (1979) carries out estimations of two versions of the model: one that corresponds to the equilibrium case and one that corresponds to the disequilibrium. In the equilibrium model demand responds negatively to the differential interest rate (loan rate minus rate of alternative sources) and positively to industrial production index and unpaid profits. Supply responds positively to all variables. For the disequilibrium model the signs of the estimated parameters are as in the previous case, moreover the parameter γ is also estimated and it is significantly different from 0 (though less than infinite) meaning that the loan rate does adjust partially (but not instantaneously) to excess demand. A comparison between the equilibrium and disequilibrium versions shows also that rates respond more to supply than to demand (meaning that when there is excess supply interest rates adjust quite quickly).

---

[2] Sealey (1979) also includes a dummy variable for the period 1973-75 to take account of the special economic conditions of that moment and which may have affected the loan market.

As we saw credit rationing is normally associated with interest rates stickiness. However Berger and Udell (1992) show that very often this stickiness is not because of credit rationing but because of the presence of commitment loans (which by definition exclude rationing), or more generally it may be due to contractual terms rather than credit constraints. B-U use individual loan data and develop a number of empirical tests that enables them to distinguish between different reasons for rates stickiness. They examine how rate stickiness changes according to different contractual terms of credit that may be related to rationing.

To test for rate stickiness they regress the loan rate premium on real and nominal rates, some contractual variables and some macro variables. In some regressions dummy contractual variables are included to test if the different contractual terms matter for the degree of stickiness. The dummy variables are: i) one dummy that takes value 1 if the contract is a commitment one; this implies that observed stickiness on commitment loans cannot reflect credit rationing. In fact it is quite the opposite since commitment borrowers cannot be rationed; ii) a second dummy refers to the situation of floating or non floating rates; iii) and a last dummy takes value 1 when the loan is secured (collateral).

A second test examines the proportion of new loans with different contractual characteristics and how this varies with credit rationing. Specifically they form logit models of probabilities that newly issued loans are commitment, floating or collateral. For example in periods of rationing the proportion of commitment loans should increase because non commitment borrowers will be rationed while commitment ones cannot. B-U conclude that rates stickiness is not related to credit rationing. A large portion of rate stickiness is related to commitment loans which in fact preclude rationing.

From a theoretical point of view there is not a clear relationship between bank risk and loan commitments. On the one hand loan commitments are riskier because there is a lack of information at the moment of signing the contract that might be solved in a later period. This lack of information may induce borrowers to switch to riskier projects once they have obtained the loan. However, banks may offer the loan commitment to borrowers that do not have these informational problems (for example to old traditional customers that also want to keep this relationship) and thus commitments could be safer than other forms of credit. Berger and Udell (1992) find evidence that on average loan commitments perform better than other forms of credit, but whether this is due to less riskiness of commitments or to a better borrower discrimination by the banks they are not able to tell.

Melnik and Plaut (1986) develop and test a model that explains the quantitative allocation of loan commitments . The focus is on the supply side of the contract: the behaviour of the borrower is taken as given and is modelled as strictly as necessary, in order to derive the relationships among the variables. However, differently from the previous studies, now also the firms' characteristics are taken into account in the empirical analysis, because the purpose here is to determine how banks allocate loans among their customers, and thus how they are able to discriminate firms' riskiness. Under a loan commitment a borrower is entitled to have a credit up to a certain amount (the commitment) for a certain period. In exchange for this right the borrower pays the bank a commitment fee (k) at the moment of signing the contract. These contracts usually specify also a set of different terms, such as the time length of the contract, the interest paid, the amount of collateral placed with the bank, and so on. The terms of the contract are strongly related to each other and the determination of the loan commitment size cannot be carried out without taking into account these relationships. An implication of this is that commitment contracts are very much borrower specific, that is they are the result of a bargaining between borrower and lender and are mainly designed according to the borrower's specific characteristics , and in this sense each contract is of a unique type. The loan contract is seen as a 'package' of terms and is described by a vector of the components of the contract:

$$B[L^*, m, k, T, C] \qquad (6)$$

where:

- L^* is the amount of commitment (this will be the endogenous variable in the regression) and L is the amount actually borrowed each period.
- T is the time length of the contract
- k is the commitment fee paid ex-ante, at the beginning of the contract and is then independent on the other variables. The fee is specific to the borrower.
- m is the mark-up placed on the borrower and is also specific to his characteristics[3]
- C is the amount of collateral required and placed with the lender.

The objective is to analyse how the size of the loan commitment varies with the other components of the contract. As far as the borrower is concerned he will select the amount L_t of credit to borrow at any time, under the constraint that $0 \leq L_t \leq L^*$, and will have to pay an interest equal to $r_t + m$. L_t will depend, among others, on L^*: the greater L^*, the larger L_t. The amount L_t will depend also on some factors unknown at the signing of the contract, and in particular on some state of the world θ_t which has density distribution $f(\theta_t)$. This creates a second source of uncertainty, more relevant to the lender, which is the probability of default, that is, under some states of the world the borrower will be unable to repay back the debt. The probability of repayment is denoted by π and is assumed to be a function of the elements of the contract (apart from k which is paid in advance), according to the following functional relationships:

---

[3] The interest rate paid on the loan commitment is composed of two parts: a prime interest rate r, which is floating and is determined by the market, plus the mark-up m, or risk premium, which is customer specific. r is not really relevant in this analysis because is determined by the market and thus it is exogenous to the contract. Moreover in M-P formulation r is assumed also equal to the lender's discount rate (or the rate on 'safe' assets) , and in the analytical formulation the two cancel out.

$$\frac{\partial \pi}{\partial L^*}, \frac{\partial \pi}{\partial T}, \frac{\partial \pi}{\partial m} < 0; \frac{\partial \pi}{\partial C} > 0 \qquad (7)^4$$

The lender will then maximise the expected utility from the loan contract. He receives in any case k at the beginning of the period. If he gets repaid (with probability π) he also receives the interest on the loan, otherwise he keeps the collateral C and looses the amount of loan (in the last period only). Taking partial derivatives of expected utility w.r.t. the contract terms gives the relationships among these variables and the conclusions are that the size of the loan commitment is larger the higher are m, k or C, while the relationship with T is ambiguous.

The model by M-P can be safely summarised in the following way: suppose T=1 and r=0 and that there is no uncertainty on the state of the world, at least as far as the lender is concerned (we could think of there being an exogenous source of uncertainty which affects the probability of repayment). Moreover π is a function of the characteristics of the borrower but not of the other terms of the contract. In this case the expected profits for the lender are:

$$E(\Pi) = k + \pi \cdot L(1+m) + (1-\pi) \cdot (C - L) \qquad (8)$$

If banks are perfectly competitive then E(Π)=0 (which in turn implies that C<L). This gives the amount of loans the bank is willing to lend:

$$L^* > L = \frac{k + (1-\pi)C}{(1-\pi) - \pi(1+m)} > 0 \qquad (9)^5$$

with $0<L_t<L^*$, and since $k+(1-\pi)C>0$, then it must be that also $(1-\pi)-\pi(1+m)>0$

It can be shown that not surprisingly:

[4] There are two things to note: the relationship between π and θ is not specified; and it is assumed that default is a possibility only once and at the end of the contract, that is at time T, while the first type of uncertainty occurs at any time t.

[5] I have determined here the size of actual loan rather than of commitment.

$$\frac{\partial L}{\partial C} \geq 0,\ \frac{\partial L}{\partial m} \geq 0,\ \frac{\partial L}{\partial \pi} > 0,\ \frac{\partial L}{\partial k} > 0 \qquad (10)^{6}$$

Next M-P regress the commitment loan on the terms of the contract and some other variables, namely some characteristics of the lenders. Data are collected from a survey among borrowers that provide information on the commitment size, risk premium, contract length, collateral requirements, sales, assets and liability structure. The sample was over the period 1980-81. The regression is:

$$L^{*} = \beta_0 + \beta_1 m + \beta_2 m^2 + \beta_3 T + \beta_4 k + \beta_5 C + \beta_6 S + \beta_7 CR + \varepsilon \qquad (11)$$

The theory predicts that β_1, β_4, β_5 are greater than zero, and β_3 is ambiguous: if it is less than zero then longer contracts are seen as riskier and they decrease the size of the commitment, while if it is positive then a longer continuous relationship is seen as more favourable by the lender.

Observations on the amount of collateral were not available, but only on whether a collateral was placed or not. So C is in fact a dummy variable which takes value 1 if there is collateral and 0 if there isn't. S and CR are included to proxy π. S is a dummy variable that tells whether or not a borrower had sales greater than 200 million dollars (the idea is that larger firms are safer), and thus β_6 is expected to be positive. CR is the ratio of current assets to current liabilities (used as a measure of cash flow and credit worthiness of the borrower) and thus its coefficient should be greater than zero.

The empirical results show that $\beta_1 > 0$ and significant, while $\beta_2 < 0$: the relationship with the risk premium is concave. The time length of the contract does not have a very significant impact (it is positive but not statistically significant), and the commitment fee is not significant at all. β_5, β_6 and β_7 are instead all positive and significant; in

[6] In this simplification it is not possible to derive the relationship between T and L. However this is not a big problem, since also in M-P the relationship between these two is not uniquely determined: if it is positive it means that a continuous relationship with the customer is awarded, if it is negative it means that a longer contract is seen as more risky.

particular β_7 has a very large impact, indicating that borrower-specific characteristics are relevant in determining the allocation of credit.

As already said there may be other reasons for having credit rationing, more related to the supply side or to macroeconomic reasons. Peek and Rosengren (1995) test whether there has been a capital crunch in the credit slowdown of 1990 in New England. The reason why there had been a slowdown was that banks had problems in meeting their capital requirements. P-R derive a theoretical model that allows them to distinguish between loan demand shocks and loan supply slowdown. In the model the capital/asset ratio requirement is explicitly taken into account. When the constraint becomes binding banks behave differently. The first order conditions for banks' profit maximisation (derived w.r.t. deposits) give the relationships between deposits and capital and betweem loans and capital, for the case in which the capital requirement is not binding and for the case in which it is. In particular when the capital requirement constraint is not binding a reduction in capital increases deposits and reduces loans (but less than one by one). At the same time if there is a fall in demand both deposits and loans decrease. On the other hand when the constraint is binding, a reduction in capital makes banks decrease both deposits and loans (moreover deposits decrease proportionately and loans must decline more than one by one, as an effect of a reduction in both capital and deposits), while a reduction in demand leaves both deposits and loans unchanged. It is clear then that shocks in demand or supply have different effects on deposits and loans (whether the constraint is binding or not)

This is supposing that it is possible to detect whether the constraint is binding or not: a shock in supply when the capital-asset ratio is binding has the same effect on deposits and loans of a shock in demand when the constraint is not binding. Therefore in order to be able to detect if there has been a capital crunch on demand it is necessary first to observe if the capital-asset requirement is binding. P-R partially solve this observability problem by sampling banks in the same geographical area , where all banks were subject to the same economic conditions, even though they were not completely subject to equal demand shocks. The remaining demand factors that cannot be removed and that differ across banks are explicitly taken into account by including in the regression some variable that refer to banks activities.

The empirical analysis confirms that banks' behaviour was indeed affected by the capital requirements constraint[7].

The estimated equation is the following:

$$\frac{\Delta D}{A} = \alpha_0 + (\alpha_1 + \alpha_2 \frac{K}{A})\frac{\Delta K}{A} + \alpha_3 Log(A) + \varepsilon \qquad (12)^8$$

ΔD is the change in deposits from 1990-1 to 1991-I, and similarly for K, A is total Assets. Data are cross sectional. If there has been a capital crunch α_1 is expected to be positive and α_2 is expected to be negative.

Estimates are carried out through 2SLS. The model predicts that banks with lower capital that face negative shocks will shrink their liabilities more than banks with larger capital. This is why the effects of ΔK/A depends also on K/A. Log(A) is a variable reflecting demand factors, which may influence some banks but not all (for example if these banks specialise in some kind of credit, like saving banks, or if a bank has a special relationship with a particular customer). Estimates are consistent with expectations and the parameter of ΔK/A is positive, while that for (ΔK/A)K/A and α_3 are negative. The results support the idea that there has been a capital crunch.

5.4. Credit rationing in the consumers' loan market

Here I consider few studies on credit rationing in the consumers' loan market. These studies have in common the main modelling structure: data are based on surveys and

[7] The constraint may become binding because of a loss of capital, this is often then the ultimate reason for a credit slowdown, thus the term capital crunch. This may result in a credit crunch if there is asymmetric information and large costs of acquiring and monitoring it, because the capital crunch may cause a rationing which is not offset by other lenders.
[8] I included only the variables which have significant coefficients

the main purpose is to determine whether banks discriminate between households according to their socio-economic characteristics when deciding to whom to grant credit. And this is also the main difference with the previous studies: before the total supply of credit was analysed, and if there was excess demand over supply; here the allocation of credit among consumers is considered instead, and how banks are able to discriminate among consumers according to their characteristics.

Avery (1981) used an approach similar to Sealey (1979) to estimate credit constraints in the consumers' loan market. A main difference is the introduction of a demand function for durable goods which is closely related to the demand for loans (in fact the goods demand function includes the quantity of granted loans as explanatory variable). This demand in turn affects both the demand and the supply of loans. This may pose a problem of identification. The model becomes:

$$Y_t = \alpha_1 X_{1t} + \beta_1 Q_t + u \tag{13}$$

$$D_t = \alpha_2 X_{2t} + \beta_2 Y_t + u^D \tag{14}$$

$$S_t = \alpha_3 X_{3t} + \beta_3 Y_t + u^S \tag{15}$$

$$Q_t = \min\{D_t, S_t\} \tag{16}$$

We have then a set of two systems of regressions depending on whether Q=S or Q=D[9]

If S≠D restrictions on the parameters must be imposed in order to have a unique equilibrium (with either demand-constrained regime or supply-excess regime prevailing for every observation)[10].

The reduced form gives rise to two switching regressions models: one regression for the demand of durable (depending on whether D and S is the smallest) and one

[9] The introduction of another demand function which is simultaneously determined with Q makes things more difficult and necessitates a different estimation procedure.

[10] Maddala observes that the imposition of restrictions on parameters is often necessary in switching regression models and it is one of the major drawbacks.

regression for the quantity of loans. These equations can then be written in reduced form. Avery (1980) proposes the following procedure : i) estimate the switching regression of the demand for durables and of the quantity of loans; ii) build instrumental variables for the unknown variables (Y and Q); iii) estimate the structural form parameters. For Y the parameters are estimated using OLS. Estimates of D and S are computed using MLE. Avery states that these estimates are consistent. This is not a FIML and it is possible that there is a loss of efficiency. The advantage is that this procedure is computationally more simple. Estimates show a positive coefficient for the quantity of loans in the durable demand function, and a positive significant coefficient of the demand for durables in the supply function of loans.

In the papers by Cox-Jappelli (1993), Jappelli (1990) and Perraudin-Sorensen (1992) data are taken from a Federal Reserve Survey of 1983 in which Households were asked, among other things, if they had been refused a loan or if they hadn't applied for fear of being refused. From this they are able to infer whether households wanted a loan or not in the first instance, and second if banks discriminate between borrowers according to their characteristics, which in this case are of demographic type.

J. and P-S model both sides of the market, that is both banks' and consumers' behaviour, while C-J concentrate on the demand side only. A main difference between the first two is that J. does not estimate the disutility of applying for credit, and does not distinguish between demand and supply effects. The idea is that the existence of a credit constraint cannot be detected by analysing only one side of the loan market : first the amount of debt consumers ask for must be determined, and second if banks are willing to offer that amount or not.

First I describe consumers' and second banks' behaviour.

Consumers maximise expected utility from lifetime consumption subject to a budget constraint . The optimisation problem gives rise to demand for debt (which can be either positive or negative), to which is added a disutility-cost of applying for debt (to account for those who do not apply for fear of being rejected). From this consumers can be classified as belonging to one of the following groups (which are determined by the demand for loans being either positive or negative):

1. those who do not want credit (the derived demand function is negative)

2. those who want credit but do not apply (the application costs are too large)

3. those who apply for credit

Banks' behaviour, whether an application is accepted or rejected, determines the further partition of category 3 into two subcategories:

3.a those who apply for credit and are not constrained

3.b those who apply for credit but are rationed

J. does not consider what happens in the first two cases, but goes directly to determine the positive demand for a loan, and a consumer is credit constrained if

$$C^* - Y - A(1+r) > D \tag{17}$$

where C^* is the optimal consumption level in the absence of liquidity constraints (so the left hand side is the desired amount of debt D^*), and D is the amount of debt issued to the borrower, Y is income, A non-human wealth and r is the exogenous real interest rate. The probability of being constrained is then determined by both demand and supply functions. The probability is related to a set of observable variables and is explained through a logit model. The characteristics of liquidity constrained consumers can then be analysed.

The following assumptions are made: the reduced forms for consumption and debt can be written as:

$$C^* = X'\alpha + \varepsilon \tag{18}$$

$$D = X'\delta + \eta \tag{19}$$

D is the debt ceiling, X is the set of demographic variables, α (common to all consumers) and δ are the parameters to estimate, ε is a vector specific to each consumer and η is an error term.

Substituting these two equations into the demand for debt gives:

$$Z = X'\beta + \mu \tag{20}$$

where: $\beta=\alpha-\delta$ and $\mu=\varepsilon-\eta$. The following dummy variable is then defined:

$$\begin{cases} Z = 0 \quad X'\beta + \mu \geq 0 & \Rightarrow \quad \textit{consumer constrained} \\ Z = 1 \quad X'\beta + \mu < 0 & \Rightarrow \quad \textit{consumer unconstrained} \end{cases} \tag{21}$$

In the first dummy J. includes both rejected and discouraged consumers, while P-S keep them distinguished. Thus the probability that a consumer is credit constrained is a function of the variable X:

$$P = F(X'\beta) \tag{22}$$

where *F(.)* has logistic distribution and $X'\beta$ is the reduced form for excess demand of loans.

One difference between J and P-S is that the reduced form equation for the probability function does not allow the first to distinguish between demand and supply effects, in fact the two parameters α and δ are not identified, but only β can be estimated, and thus only the net effect can be detected. One problem is also that when demand and supply work in opposite directions it is possible to see at least which one of the two prevails, while nothing can be said when demand and supply go in the same direction. The results are that the probability of being constrained depends significantly on both consumers' and banks' behaviour and on some demographic variables. In particular the probability P is a negative function of current income (though the relationship is not completely linear) and wealth for two reasons: because the demand for credit falls and because the debt ceiling increase. Age has also a negative effect: on the supply side banks tend to lend to older people because they have a longer credit history and are considered less risky. On the demand side the effect is ambiguous. The estimated net effect is negative: the older the consumer the lower the probability of him being constrained. Relevant demographic variables are: marital status (couples are less likely to being constrained), race (not surprisingly being white has a negative effect on the probability of being rationed). while the size of household has a positive effect on the probability

On the contrary, the study by P-S seems more general, even though to have an explicit demand function they have to assume a quadratic utility function. The theoretical framework is the following: consumers maximise a utility function from consumption subject to some budget constraint. Whenever the optimal level of consumption is higher than current income they would apply for a credit, however there is also a disutility of application, d_α that makes some consumers with positive loan demand not apply. This disutility depends on consumers' characteristics. Last, those consumers that apply face probability P_α of not being accepted by the bank. P_α is determined by the bank and is a function of consumers' characteristics.

So consumers' behaviour identifies the demand for debt, while banks' behaviour determines the probability of acceptance. In summary P-S estimate the three following equations:

$$\lambda_{\alpha} = \beta_{1i}\alpha_i \tag{23}$$

$$d_{\alpha} = \beta_{2i}\alpha_i \tag{24}$$

$$\gamma_{\alpha} = \beta_{3i}\alpha_i \tag{25}$$

where the first equation refers to the demand for loans, the second to application costs and the third to loan supply. The subscript α denotes which group consumers fall into. The α are the socio-economic characteristics, and β are the parameters to be estimated. Errors are assumed to have logistic distribution and so the model is an ordered logit.

The estimates show that for both demand and supply the demographic characteristics of the borrower are indeed relevant. In particular banks discriminate between consumers according to some particular characteristics. Also the application costs differ among consumers. Estimates show that the probability of being given a credit is an increasing function of: good health conditions, household head being white, being young and unmarried with no children and it increases with income. On the contrary, this probability decreases when the household head is male and when living in an urban area. Consumers have lower loan demand when household and spouse are both unemployed or is one of the two, when the household head is male and when the size of the family is relatively small. Demand increases when household head is white and married. The application cost is lower when household head is male, when family size is lower and when living in suburbs; it increases when household head is white and has lower education.

Also C-J find that consumers are constrained according to their socio-economic characteristics . Similar to the previous papers they classify consumers according to their having/not debt, being/not constrained. They estimate the desired amount of debt (supposedly derived from a utility maximisation problem) in a three equation Tobit scheme. In general the desired amount of debt can be expressed as:

$$D^* = X_1\beta_1 + \varepsilon_1 \tag{26}$$

It must be noted however that D^* (which is the desired loan demand) can actually be observed only when the demand of debt is positive and the consumer is not (fully) rationed. If the consumer is not constrained D^* denotes the optimal amount of debt demanded. In order to estimate D^* C-J define two latent variables:

$$b = X_2\beta_2 + \varepsilon_2 \tag{27}$$

$$n = X_3\beta_3 + \varepsilon_3 \tag{28}$$

where $b>0$ when the consumer holds a positive debt and $n>0$ when the consumer is not constrained. X_2 contains X_1 and other variables that may affect consumers' convenience of holding debt. X_3 contains proxies for bank characteristics and some variables affecting the demand plus the debt ceiling denoted by D_0. The consumer is constrained whenever $D^*>D_0$. Thus observing positive debt doesn't imply that a consumer is not rationed; it may well be that the amount of debt granted is less than the optimal demanded.

The two latent variables define four regimes w.r.t. consumers: unconstrained consumers with positive debt ($b>0$ and $n>0$), unconstrained consumer with zero debt, constrained consumer with positive debt (but still less than D^*) and constrained consumers with zero debt. As said before D^* is not directly observed so what C-J do is to use the first category of consumers (unconstrained consumers with positive debt for which D^* is the actual desired level of debt demanded and obtained) to have consistent estimates of the reduced form of desired debt. Thus the regression for debt is conditional on the consumer holding positive debt and being unconstrained:

$$E(D^* | b > 0, n > 0) = X_1\beta_1 + E(\varepsilon_1 | b > 0, n > 0) \tag{29}$$

X_1 contains measures of consumers' resources (current income, permanent income and net worth) and some demographic variables (marital status, family size, race, gender of household head). X_2 has also years of education, occupation, area income, employment status and rural/urban status (this is the explanatory variable for having positive debt). While in X_3 they use variables to proxy the scoring function that banks use to evaluate consumers' riskiness (which affects the debt ceiling): regional dummies to capture different regulations and characteristics of the credit market, and employment dummies to proxy job stability.

For the probability of being unconstrained estimates give a positive coefficient for permanent income, net worth, age and concentration ratio (C-J) explain this as the fact that banks in areas with higher concentration might be willing to accept higher risk); while the coefficients for race (being black), family size and living in an urban area are negative (all others are not significant). With regard to. the probability of holding positive debt, permanent earnings, age female household head and family size have a positive coefficient; living in an urban area has instead a negative coefficient.

5.5. Capital structure and firms' own characteristics.

So far I have considered mainly studies that model the supply of credit from a bank perspective, particularly for commercial loans. Especially in this market it has been seen that there are really few studies that manage to distinguish the allocation of credit with regard to firms' characteristics. This is mainly for lack of data. On the contrary, the availability of a survey in the consumers' loan market made it possible to distinguish between demand and supply and consumers' characteristics, rendering these studies more complete. In what follows I will analyse again the commercial loan market but on the demand side instead. First I will describe firms' characteristics and how they relate to firms' preferences over debt structure and to banks' discrimination criteria over borrowers.

Some of these characteristics are not directly observable and thus proxies or indicators are used instead; a description of these is also outlined, as well as the links between characteristics and indicators (since as we shall see later the same indicator can be used to proxy different characteristics, or different indicators can be used for the same characteristic)

First the common framework to all studies is described, while the differences will be outlined in the empirical survey. Here I start with a description of firms' attributes and their indicators, together with the implications in terms of significance and signs of regression coefficients. It is clear that even though characteristics and indicators are described separately one by one, they are strictly connected to each other and at the end it is difficult to say that a characteristic belongs to one attribute rather than to another, and the same holds for indicators.

1. Historical relationship between the firm and the bank: for the bank there are two types of customers, 'old' and 'new'. Old customers are less likely to be rationed, as they are considered less risky and less costly to screen and to administer for the bank.

2. Reputation: this is usually proxied by age and size and volatility of earnings growth (with a negative relationship with this last indicator)

3. Age: (this is usually measured as number of years since incorporation) older firms have matured a higher reliability reputation and thus can rely more on public debt . On the contrary younger firms do not really have to rely more on bank debt. From what was said before it follows that older firms have a higher proportion of public debt, while very young firms have a higher proportion of private non bank debt. 'Middle age' firms should have a higher proportion of bank debt.

4. Size: the leverage ratio may be related to the firm's size as well as the maturity debt structure. Small firms rely more heavily on bank debt than larger firms, and usually prefer short term debt rather than issuing larger term debt because of the lower fixed costs associated with this alternative. Indicators of size can be sales or total assets (the two measures are highly correlated) or quit rates (the idea is that larger firms usually provide better career opportunities and thus have lower quit rates).

5. Collateral value of assets: the type of physical assets owned by a firm may affect its capital structure. Indicators are the ratio of intangible assets to total assets (negative relation) and the ratio of inventory, gross plant and equipment (fixed assets ratio) to total assets (positive relation).

6. Non-debt tax shields: tax deductions for depreciation and investment tax credit are substitutes for the tax benefits of debt financing.

7. Growth of investment opportunities: firms with higher expected growth (in investment opportunities) prefer short term debt (for which there is thus a positive relationship) to long term debt (for which the relationship is negative). Firms must decide quite simultaneously the amount of debt and the maturity of the debt as well; and not only the capital structure but also debt maturity depends on firms' characteristics. In general firms have an incentive to increase the proportion of debt over capital in order to take advantage of a higher leverage, and to issue more long term debt (for example to minimise the costs associated with rolling short term debt). However in both cases a disincentive to invest may arise. This occurs because the benefits from profitable projects are divided between stockholders and bondholders, and in some cases bondholders may capture most of them, leaving stockholders with less than normal returns. As a result projects that hold positive net present value may be rejected. The disinvestment incentives can be controlled by decreasing the proportion of debt. However also in this case the problem is not completely solved, in fact if most of the debt is long term debt, and firms want to ricapitalise to undergo the investment, they have to buy back the outstanding debt at a price that does not reflect the value of the new project, and if information about the project becomes public then the price of the long term debt increases and the disincentive remains. If debt is more short term this problem is solved. Indicators include capital expenditures over total assets and the growth of total assets measured by the percentage change in total assets. R&D over sales is also an indicator of growth of investment opportunities.

8. Investment opportunities and project quality (intending by this the likelihood that a future project will be successful): these are measured by the ratio between the market value of firm's assets to its book value (market/book ratio from now on). This can be used as a negative indicator of project liquidation values (early liquidation is less valuable for firms which have projects with low liquidation value, for which public funds are then preferred)

9. Uniqueness: the more unique is the product a firm sells, the more specific are workers skills and capital and suppliers characteristics, and the more difficult it is for the customers to find alternatives to this product. This means that in case of crisis or even bankruptcy this firm becomes more risky (substitutability becomes relevant in case of bankruptcy) for a lender. The debt ratio is thus negatively related to uniqueness. Possible indicators for this attribute are: R&D expenditures as a percentage of sales (firms with close substitutes are less likely to invest in R&D) or selling expenses over sales (the claim here is that firms with more unique products are expected to advertise and promote more than other firms. In my view this is not always true, even if selling expenses are weighted for sales. In fact the argument that is usually used doesn't take into account the use of advertising as a strategic variable for competition between firms, to gain market shares or to make consumers less price elastic. Another problem with the use of selling expenses is that these models make use of cross section data, while it may well be that it is past advertising rather than contemporaneous which is more related to uniqueness.)[11]. Quit rates (percentage of industry's total work force that voluntarily left the job in the sample years) are also used

10. Industry classification: the debt may vary according to type of industry (e.g. firms manufacturing machines and equipment should be financed with less debt). The reason is related to collateral value of assets. The collateral value is important when liquidation value is considered: the more specialised the asset is, the less

[11] All these models in general cannot take into account time lagged effects, which may be quite important in determining credit allocation or in affecting firms' choices of credit sources.

liquidation value the asset has and the less suitable as a collateral it is. It follows that firms with more specialised assets are, relatively, financed with less debt.

11. Profitability: the past profitability and thus the amount of earnings that could be retained to refinance future projects should also be an important determinant of capital structure. Indicators are operating income over sales and operating income over total assets. It must be observed that there may be problems of endogeneity with this variable.

12. Credit riskiness: firms with highest and lowest credit risk issue short term debt, while firms with intermediate credit risk issue long term debt (this follows from possibility of monitoring and reputation concerns). A key result is that borrowers with credit rating in the middle of the spectrum borrow from banks (the reason is that at the two extremes there are higher rated borrowers, who have to lose a lot of reputation from default and thus do not need monitoring, and lower rated borrowers, who do not have much to lose in terms of reputation from default, and for whom monitoring will not give the right incentives for non default. Monitoring is worthwhile and successful only for those borrowers who have enough to lose from defaulting in terms of reputation, and who want to build up a higher reputation. These borrowers thus have a preference for banks' debt)[12]

13. Volatility of earnings growth: the optimal debt should be a decreasing function of volatility of earnings. This variable is used as a proxy of credit risk and likelihood of financial distress: more volatile firms are more credit risky because they may experience more states in which cash flows are too low to repay the debt. More volatile firms would prefer to have the option to renegotiate the debt and thus prefer private debt to public debt.

14. Leverage: defined sometimes as the book value of long term debt over the book value of total assets and some other times as short and long term debt over total

[12] Demand for monitoring is also higher in periods of higher interest rates or negative expectations about future profitability.

assets. Johnson (1997) suggests introducing this variable even though it is almost the dependent variable, and in fact appropriate techniques must be used if it is to be used as explanatory variable. Leverage is in fact related to the same determinants as the debt structure, and being the numerator of the dependent variable it is certainly correlated to the error term. Simply introducing it as an explanatory variable would make estimates biased and inconsistent. Johnson (1997) uses an auxiliary regression to tackle the problem: he partitions the leverage into an endogenous part related to the other explanatory variables, and an exogenous part. Thus he first regresses leverage on the other explanatory variables (this is the auxiliary regression) and then uses the residuals as an instrumental variable for leverage in the debt ownership regression. This residual has the advantage of being highly correlated with leverage and orthogonal to the other explanatory variables. Leverage is related to investment opportunities, collateral value, firm's size and volatility.

5.6. Some empirical studies

In order to analyse the determinants of capital structure Titman and Wessels (1988) regress (separately) six measures of leverage on firm's attributes. These measures are long term, short term and convertible debt divided respectively by market and book values of equity (debt is measured in terms of book value).

For the attributes T-W adopt a procedure quite different from usual, which may be interesting but I think inefficient. Firms' attributes, as I have said, are quite often not observable and for this reason proxies are used instead. However, there are some problems in so doing. For example, one attribute can usually be represented by different proxies, or conversely the same proxy can be used as an indicator for different attributes. In particular this last poses serious problems for interpreting the statistical results, as we do not know how to distinguish the net effects of different attributes. Another problem (not mentioned by T-W, but still quite important) is that the indicators may be (and often are) multicollinear. The approach they use can indeed solve this problem. Last, because measurement errors of the proxies might be correlated with measurement errors of the dependent variable, there could be spurious correlation even when the attribute is not related at all with the dependent variable.

For all these reasons they adopt a factor analysis approach, which works in the following way:

there is a variable Y we want to explain and which depends on some unobservable attributes ξ. These attributes, however, may be measured by a vector of observable indicators. For the particular technique they use the relationship between attributes and indicators need be linear:

$$X=\Lambda\xi+\delta \tag{30}$$

where X is the vector of indicators, ξ is the vector of unobservable attributes, Λ is a matrix of coefficients and δ is the error term. This part is called by the authors the measurement model. If X and ξ were of the same dimension, the relationship would be exact and we wouldn't need an error term. However, since in this model X has dimension 15, while ξ has dimension 8, the error term must also be included[13].

The main idea underlying this approach is the following: we can observe the vector X which has a given variance Σ. The issue is to find the 'components' ξ of X that can explain the most variance of X[14], and then use these estimated components as explanatory variables in a regression for Y:

$$Y=\Gamma\xi+\varepsilon \tag{31}$$

where Γ is the matrix of coefficients, and this part of the system is called structural model.

[13] There are two other reasons to use this approach: one is to solve multicollinearity if present, the other is to reduce the dimension of the number of regressors.

[14] As a side effect these components should also be orthogonal to each other, which is desirable in econometrics. However this is not the case in T-W.

Johnston (1984) observes that estimates obtained through this method may be biased, but the variance of the estimators could be lower than that of an OLS estimator. However, one of the disadvantages is that they must impose a linear relationship between attributes and indicators and, more important, in order to be able to identify the measurement model they need to impose quite many restrictions (105) on the matrix of coefficients Λ. On the other hand they do not need to impose restrictions on the structural part of the model.

The attributes considered by T-W are: growth, uniqueness, non-debt tax shields, asset structure, size, profitability, volatility, industry classification. The indicators used are: Non-Debt Tax/ Total Assets, Tax Credit/TA, Depreciation/TA, R&D/Sales, Selling Expenses/Sales, Capital Expenditure/TA, Intangible Assets/TA, Fixed Assets Ratio (Inventory+Plants+Equipment)/TA, Log of Sales, Percentage change in Total Assets, Quit Rates, Operating Income/TA, Operating Income/S (it is not clear why they include both these variables), Standard Deviation of Percentage Change in Operating Income, Industry dummy (1 for firms that produce machinery and equipment, 0 otherwise).

Their sample period goes from 1974 to 1982, which is then divided into three subperiods over which they take averages of the variables. The reason for doing this is to have a sort of past and future (expected) values of some variables (to take account of lagged and lead effects). Thus the observations for debt refer to the average through 1977-79; the growth rate of total assets and capital expenditures over total assets were measured over the period 1980-82 (thus the lead realised value was taken as a proxy for expectations). Uniqueness, Non-debt tax shield, asset structure and industry dummy are instead measured contemporaneously, while for size and profitability the lagged observations (1974-76) were used (this is to see if there is more of a long run effect rather than a short run; lagged size is also to avoid a spurious correlation with debt size). The standard deviation of change in operating income was measured using all nine years.

The estimation procedure is a method derived from MLE, and they estimate the two equations simultaneously.

The estimated coefficients for the measurement model are significant and of the predicted signs. In particular, growth is measured by R&D as a percentage of sales, capital expenditures over total assets and percentage variation in total assets: all coefficients are positive. Uniqueness is measured by R&D over sales, selling expenses (both with positive coefficients) and quit rates (negative coefficient). Non Debt Tax Shield is measured by tax credit over total assets, depreciation over total assets and a direct measure which is the tax shield itself (all coefficients are positive). The collateral value is proxied by the ratio of intangible assets over total assets and Inventory+Plants and Equipment over total assets; the coefficient of the first is negative while the coefficient of the second is positive. Size is measured by the log of sales (with a positive coefficient) and quit rates (negative coefficient). Profitability is measured by operating income over total assets and operating income over sales (both have positive coefficients). Volatility is identically defined by the standard deviation of percentage variation in operating income; its coefficient is constrained to be equal to 1 (implying that there is no error of measurement). Industry classification is also measured without error by the industry dummy.

For the structural model the 6 regressions are:

a) Long term debt divided by the market value of equity. For this regression the coefficients of uniqueness, profitability and industry dummy are all negative, while all the other parameters are not significant.

b) Short term debt divided by the market value of equity: the coefficients of uniqueness, size and profitability are negative and the others not significant.

c) Convertible debt divided by the market value of equity: the cross sectional variation seems not to be explained by any of these attributes.

d) Long term debt divided by the book value of equity: the coefficient of growth is positive, while uniqueness and size are negative.

e) Short term debt divided by the book value of equities: uniqueness and size are negative.

f) Convertible debt is again left unexplained

The fact that the coefficient of uniqueness is always negative implies (according to T-W) that firms that can impose high costs (in case of liquidation) on customers, workers etc., tend to have lower debt ratios. The fact that size is negative and significant only for the proportion of short term debt over market value of equity, but not for the proportion of long term debt seems to suggest that smaller firms tend to prefer shorter term debt compared to larger firms. However, when debt is divided by the book value, size becomes negative and significant for both short and long term debt, implying that smaller firms prefer any kind of (private) debt maturity compared to larger firms. A possible interpretation for this difference in results is that many firms are guided by the market value of their equity, rather than their size, when selecting their long term debts.

Profitability is significant only for the market value measures, which suggests that an increase in the market value of equity, due to an increase in operating income is not completely offset by an increase in firm's borrowing. A positive coefficients of growth in the long term debt/ book value regression implies that since growth adds value to a firm, it increases its debt capacity and hence the ratio of debt to book value.

It must be said that overall they are not really able to explain too much of the cross-section variability with the attributes included in the regression. Convertible debt moreover is left completely unexplained, while for the other two kinds of debt most of the coefficients are not significant.

Johnson (1997) also examines firms' choice of debt sources, but distinguishing between public debt, private-bank debt and private non-bank debt (all refer to long term debt)[15]. Although a firm holds different kinds of debt at the same time, their proportions will vary among firms according to their characteristics. In particular J. considers the following characteristics (attributes): age, Book Value of Total Assets as a proxy for firm's size. Size in turn is used as a proxy of monitoring costs which decrease for larger firms. Age and Assets are a negative indicator for not observable risk; Volatility of earnings growth[16,17]; Market-to-Book Ratio, Leverage, Fixed Asset Ratio is a proxy for assets collateral value, however the collateral value of an asset is lower the more specialised this asset is. To take account of this J. includes also a variable which is the fixed asset ratio multiplied by a dummy that stands for a SIC code (1 for manufacturers of machinery tools and instruments that have very specialised assets and 0 otherwise). the sum of this coefficient with the fixed asset ratio coefficient reflects the response to collateral value of very specialised assets; while the simple coefficient of the fixed assets ratio reflects the collateral value of non-specialised assets.

J. estimates three Tobit regressions, for the proportion of long term public debt, the proportion of long term bank debt and the proportion of long term private non bank debt, which give the following results:

Public debt depends positively on age, total assets, fixed asset ratio and leverage, and negatively on a dummy variable for fixed assets for the machinery industry and on volatility. It does not depend on the market-to-book ratio

Private non-bank debt depends positively on market-to-book ratio and negatively on the fixed assets ratio and leverage; it does not depend on the other variables.

[15] The distinction between bank and private non bank lending is justified by problems of asymmetric information and monitoring. there is counter-evidence about the significance of this distinction, in the sense that some studies find it relevant while some others do not.

[16]In J. this variable is defined as the standard deviation of first differences in earnings before interest, taxes and depreciation, for 5 years preceding the sample year, scaled by the average asset of that period

[17] The idea is that firms with more volatile earnings are more credit risky because they face more states in which they may not be able to repay the loan.

Private bank debt depends negatively on age, total assets, market-to-book ratio, and positively on fixed assets ratio and on leverage.

In summary J's results seem to confirm that the hypothesis that older and larger firms (with higher reputation to defend) prefer to borrow publicly. They find it more convenient to incur fixed costs of issuance and to produce all the necessary information for a debt which is otherwise cheaper than the private one, and to build up or protect their reputation. Also, less credit risky firms prefer public debt, while more risky firms tend to borrow more from banks. Project quality seems relevant only in determining a preference towards private non bank debt rather than bank debt. The positive coefficient for fixed assets ratio also suggests that concerns for ineffiecient liquidation or asset substitution are important, while it goes contrary to the idea that public debt should decrease as project liquidation values increase (however being a proxy of all these characteristics it only measures the net effect, so at the end it is only possible to say that one effect dominates over the other, but not that one hypothesis must be rejected. In this sense the approach of T-W seems preferable and more appropriate).

Since public and bank debt are at the two extremes of credit quality spectrum, J. analyses also the mix of public and bank debt. More precisely J. regresses the proportion of bank debt over firms' characteristics, for only those firms that have access to public debt. This is interesting because -apart from fixed costs- bank debt is more expensive than public debt, and so firms do borrow from banks only if they need some monitoring -for example to build up a reputation-or want to have renegotiability. Thus for this subsample of firms, bank debt is negatively related to age, total assets, market-to-book ratio, and positively related to the fixed asset ratio and leverage. The signs and significance of the coefficients are the same as in the full-sample regression.

Usually debt maturity is related to private and public debt in the following way: bank loans are in general assumed to have shortest maturities, followed by private non bank loans, while public loans have longer maturities. As Barclay and Smith (1995) show debt maturity is also associated to firms' characteristics. It could be possible thus that the preferences for private or public debt are indeed preferences for different maturities rather than for different sources. J. finds evidence that the choice between debt sources is not only due to preferences for different debt maturities. However, for public debt the relationship with fixed assets and volatility may also reflect maturity effects.

Easterwood and Kadapakkam (1991) also examine the determinants of the choice between public and private debt. Their sample contains medium to large firms, but not small firms; for the amount of long term debt, however, the variability is large enough[18]. E-K infer the proportion of private debt from the figures of total long term and public long term debt which are available. Differently from J. they do not distinguish between bank and private non bank debt.

E-K regress the proportion of long term private debt on the following variables: log of book value of long term debt; fixed assets ratio; expenditures on advertising as a percentage of sales; R&D as a percentage of sales; expenditures as a percentage of total assets; Ratio of earnings before depreciation, interest and taxes to total assets (profitability) this variable is mainly included as a control variable for leverage; volatility of stock returns adjusted for leverage: private debt should increase with volatility (this represents the risk of financial distress, and private debt should increase for leverage problems). The coefficient is thus expected to be positive.

Variables 2-5 are proxies for leverage-related costs. The first 6 variables are averaged over the sample period (1980-88), while the 7th is calculated as the proportion of market value of equity to firm value times the standard deviation of daily stock returns for each year, averaged over the sample period.

[18] Long term debt is measured as the book value of long-term debt and leverage is measured as the book value of long term debt to the sum of this value and the market value of equity.
Public long term debt is any publicly traded debt with maturity greater than one year at the time of issuance.

For comparison E-K run two regressions: one for private debt and another one for leverage (defined as long term debt over total assets). Both regressions include the same explanatory variables, but the coefficients in the two regressions should have opposite signs[19].

Regressions are cross section. For leverage the significant coefficients are those of R&D (negative), profitability (negative) and volatility (negative), while for the regression for private debt the only significant coefficient is the one for long term debt (negative). This last result seems to indicate that the smaller the level of long term debt, the more firms prefer private to public debt, and this is consistent with the hypothesis that issue costs and secondary markets are important in determining the choice between private and public debt. The explanatory power of these results is not very large, and so it is quite difficult to draw conclusions.

E-K give the following possible explanations for the poor results:

- Non linear relationships may cause parameters to be not significant even though a relationship does exist. They try to deal with this problem by grouping variables in the upper and lower quartiles, but still private debt is related only to long term debt.
- Also the averaging over such a long period might smooth some variables. Very likely firms' characteristics and debt vary through time, and so averaging the data eliminates some important information. E-K handle this problem by using a pooled data set (139 firms for five years 1980-82-84-86-88. And some firms had to be eliminated because they did not provide information for the whole period). Now also the coefficients for capital expenditure and profitability are significant

[19] Here again a problem is that one of these variables is long term debt, which contains bank debt in the first regression and corresponds to the numerator of leverage in the second regression, and thus must be in some way correlated to the errors.

(negative and positive respectively). The fact that capital expenditure is negatively related to private debt is inconsistent with the hypothesis that private debt reduces leverage-related costs. E-K conclude that averaging is not the reason for the poor explanatory power of the pooled regressions.

- Last, including variables that differ largely for the amount of private debt may cause problems[20]. Here they take into account the possibility that larger firms may require private funding only occasionally. For these firms then the relationships between private debt and the other variables may not be significant, and because of their size they may overshadow also the relationships for smaller firms. In practice they partition the sample according to firms' sizes, however also this hypothesis seems not to work, and in fact results do not improve particularly (and for smaller firms still only long term debt is significant and negative, while for larger firms long term debt is not significant, but only R&D is -and it is negative-).

One of the reasons why private debt may be 'cheaper' for the firm is that the costs necessary to provide information are lower for the firm even if it does not have to incur such costs, since the private intermediary will face them. On the contrary, when the firm wants to borrow publicly it has to incur some costs to provide information to the public. These costs are usually quite large, and so only larger firms find it convenient to give such information. One fact that is also worth noticing is that initial private monitoring may also give information and provide credibility for later public borrowing applications.

Information costs are also larger when a firm first wishes to borrow publicly. The higher uncertainty associated with firms 'unknown' to the public may make the initial public offering (IPO) of their equity lower, and consequently these firms can benefit from reputation and credibility established through private debt. James-Wier's (1990) purpose is precisely to examine the effects of private debt on IPO costs. In particular

[20] It must also be remembered that differently from time series, in cross sections there is not a uniquely determined order, and different data ordering may also give different results.

they want to test a model for which the existence of a borrowing relationship (between a firm and a private lender) can reduce the ex-ante uncertainty of the value of equities and thus reduce the problem of IPO underpricing. For our purposes I will illustrate only the empirical results, in which they show the relationships between borrowings and firms' characteristics.

J-W estimate a logit model for the borrowing relationship with firms' characteristics that may be associated with IPO underpricing. For firms' characteristics they use age, fixed assets ratio and sales in the 12 months prior to the IPO. Only age and sales are significantly positive, meaning again that credit history is longer the more reputation the firm has and the larger it is. J-W conclude also that debt is associated with characteristics that may affect the costs of borrowing .

Houston and James (1996) also aim at analysing the structure of firms' debt relative to their characteristics. The sample includes some firms observed at three points of time: 1980, 85, 90. They regress the proportion of bank debt (over total debt) on some firms' characteristics. They are interested mainly in multibank relationships and on the relationships between private loan and growth opportunities. For this reason they use two measures of growth opportunities: the market-to-book ratio and R&D as a percentage of sales. They run two types of regressions including each of these variables per time. One problem here is that both variables are significant in each separate regression and there are no reasons not to include them jointly in the same regression, as they are indicators also of other attributes.

H-J run six regressions: 4 OLS estimates and two Tobit regressions since the dependent variable is constrained to lie between 0 and 1. Common explanatory variables to all regressions are: coverage, leverage, the log of assets and two time dummies (one for 1985 and one for 1990). These last two are included to test whether there is a time pattern in the decrease of bank debt. However, both variables are always not significantly different from zero, meaning that if there is a decline in the proportion of bank debt this is more related to firms' characteristics (but perhaps this should have been better tested by including some dynamics).

As said previously in the first three regressions H-J include the market to book ratio and this variable multiplied by a dummy for multibank relationships. In the second three regressions they include instead R&D and R&D multiplied by the multibank relationship dummy. Two regressions in each group differ also for the inclusion, in the second ones, of the multibank dummy. Again, this variable is significantly positive and there is no reason for its exclusion. The third regressions of each group are Tobit estimates.

Results from estimates are the following: coverage has a negative coefficient but it is almost always not significant. Leverage is always negative, and it is not significant in the two Tobit regressions, meaning that highly levered firms rely less on bank debt, and it is consistent with the idea that there are economies of scale in issuing public debt. The log of assets is negative: larger firms have lower proportion of banks' debt, and in fact usually these firms have higher convenience to go public. The market to book ratio is negative: this variable has been included as a proxy for growth opportunities and this result seems a bit counterintuitive (J-H explain it with potential hold-up problems). Similarly also the coefficient of R&D/Sales is negative. On the contrary, the interaction between market to book ratio and multiple banks relationship is positive (and again the use of R&D gives the same result): this result seems consistent with the idea that multiple banks relationships mitigate hold up problems.

This argument however is on the bank's side decision process, and the models described so far do not take it into account. This leads to the conclusion that the two, firms' and banks' behaviour (that is demand and supply) should indeed be estimated together, since it is extremely difficult to distinguish whether a firm decreases its proportion of bank debt because of its preferences or because of a restriction by the bank.

Petersen and Rajan (1994) regress instead the interest rate charged on some variables which include firms' characteristics and economy and industry conditions. The aim is to estimate the impact of lending relationships on availability and price of loans. However in their sample they consider only small firms (less than 500 employees) and relatively young firms (the median age is 10 years). They run 4 OLS regressions; the first two differ only for the functional relationship between the dependent variable (interest rate charged on the most recent loan) and the age of the firm and the length of the relationship: that is, in the first regression this relationship is assumed linear, in the second is logarithmic. In these two regressions, the prime rate, default spread and the number of banks have a positive effect on the interest rate charged. With regard particularly to the multiple bank relationship it seems here that borrowing from more banks is considered as a sign of riskiness by the banks itself, which will then tend to charge a higher rate. Conversely, firms' size (book value of assets) and age have a negative impact on interest rates. They also include a dummy variable to take into account the fact that the firm may be a corporation, and this has also a negative effect on the interest rate, similarly for other two dummies related to the loan characteristics (a floating rate loan and a nonfinancial firm loan). In a third regression P-R include also two other variables that are significantly negative: sales growth and profits/interest. And similarly in the last regression they also include two industry related dummies, one of which is significant. As before the same critique applies: if these variables are significant they should have been included in the first two regressions as well; not doing so biases the results. Other variables such as the term structure spread, debt book assets, a dummy for a bank loan (distinguishes whether the loan is from a bank or from another financial intermediary), the length of the relationship between the firm and the bank and others are non significant.

P-R's ultimate goal is to estimate the effects of relationships on credit availability. Since the credit market is often in disequilibrium it is usually difficult to estimate credit availability directly. In fact it is difficult to determine whether a firm has low debt ratio because it is credit rationed or because its demand is low. All the empirical studies analysed before use the debt ratio as dependent variable for single regression models, while the credit allocation is determined simultaneously by both demand and supply. And in fact P-R show as an example a regression where debt/assets is regressed on firm's size (book value of assets), profitability (profits/assets), a dummy to account for the fact that the firm is a corporation or not, age, length of relationship, Herfindhal index for bank deposits and two industry dummies (one for the mean profits/assets for 1987, and one the mean-variance of the same variable for the period 1983-87). And as they report, size, profitability and age are proxies for firm's reputation and credit quality: the larger and older the firm is, the less risky it is for the lender, and thus the higher its debt ratio should be. Empirical estimates, however, give the opposite sign for profitability and age (which are negative). And the point is that in this relation they cannot distinguish if there is credit rationing or if these firms simply rely on other sources of credit. Other results are that: the coefficients of size, corporation dummy, Herfindhal index and mean industry dummy are positive, while all the others are negative

So they suggest an alternative measure for credit availability. The idea is that if a firm is credit rationed it should subsequently borrow from more expensive lenders. They define two variables (determined by a survey): late payments (percentage of trade credit that is paid after the due date) and discounts taken (percentage of discounts for early payment). Each is a proxy for the amount borrowed from alternative sources. These two variables seem to depend on age rather than on size. Thus late payments and discounts taken are regressed on measures of investment opportunities (age, size and industry dummy which are included later) industry dummies, measures of cash flow (internal cash flow normalised by book assets and the ratio of outstanding institutional debt -total loan less family and owner loans- to book assets) and measures of relationships: length of longest relationship with a financial institution (a measure of how informed a lender is),-fraction of borrowing that comes from institutions that provide at least one significant financial service to the firm-, measure of how concentrated the firm's borrowing is, and the Herfindhal index for financial institutions in the neighbourhood of the firm. The dependent variable is a percentage and thus a Tobit regression is estimated.

For the first regression, that is the percentage of trade credit that were paid late, results are: the coefficients of debt from institutions/assets, dummy for corporate or non-corporate firms, number of firm's lenders are positive, while those for age, length of relationship, debt from financial provide and Herfindhal index are negative. Book value of assets and profits/assets are non significant. In other regressions they also include other variables which are significant.

The second regression is in some sense the mirror image, in that the dependent variable is the percentage of early payment discounts taken by the firm, and the coefficients should have thus opposite signs: now book value of assets, profits/assets, age, length of relationship, Herfindhal index have positive coefficients; debt from institutions/assets, dummy for corporate firms and the number of institutions from which the firm borrows are positive. Again in alternative regression specification they include also other variables which are significant.

Table 5.1: Summary of empirical studies

	T-W		J		E-K		J-W		H-J		
Dependent variable	LTD	STD	Pub D	BD	NBD	Lev.	D	BL/ Td	BL/ TD	Tobit	i
Explicative variables											
	Growth (or Market/Book ratio)		Growth (M/B)						Growth M/B* multibank dummy		
	Uniqueness										
	Non Debt Tax Shield										
					Advert./S (unique.)						
					R&D				R&D		
									R&D* multi-bank relationship		
									Multi-bank dummy		
	Fixed Asset Ratio		Fixed Assets		Fixed Assets		Fixed Assets				
									Coverage		
	Size		Total Assets				Sales				
	Profitability										
			Leverage						Leverage		
					Log long term debt						
	Volatility of earnings growth		Volatility								
	SIC dummy										
			Fixed Assets*Sic dummy								
			Age				Age				

5.7. Conclusions.

In this survey I have analysed the determinants of credit allocation and/or rationing. Credit allocation is the result of lenders' and borrowers' behaviour and thus it is related to their characteristics, implying that even if there is a disequilibrium, demand and supply of credit should be estimated together, since demand shocks have supply effects and vice-versa.

Lenders' characteristics are for example their capital requirement constraints. Borrowers' characteristics are of socio-demographic type for consumers' loans, or firms' attributes for commercial loans.

In section 5.4 of this survey I have considered the consumers' loan market, and in particular the question whether banks discriminate between consumers according to their characteristics. Empirical estimates confirm this hypothesis. The presence of surveys made data available and this made it possible to estimate relatively rich models of both demand and supply of credit, and thus the presence of credit rationing determined by borrowers characteristics.

A similar question has been asked also for the commercial loan market and some studies have been surveyed in section 5.6. However the answer in this case is less conclusive for two main reasons. The first is the availability of only a few studies, mainly because of lack of data. The second is that it is more difficult to discern the reason for a credit rationing in the commercial loan market, as this is also strongly related to macroeconomic variables.

As I have said, for the commercial loan market there are few studies that estimate the supply of loans according to firms' characteristics. Most of the studies described in this survey modelled instead only the demand for credit.

Demand for credit is one of the components of a firm's capital structure, and along with the other components is determined by the firm's preferences and its characteristics. These characteristics are not always directly observable and so proxies are used instead, with the problem that it is very difficult to state clearly and uniquely which indicator is proxing which characteristics. So, together with the problem in determining which characteristics are indeed relevant in determining the demand of credit, there is the additional problem of choosing the appropriate proxies, and then to uniquely relate them to the characteristics, and thus to the demand of credit.

One of the implications of this is that relationships may not be linear, affecting in particular the significance of the estimates.

From these studies it is thus difficult to derive a unique conclusion with regard to which variables affect credit allocation in the commercial loan market, for several reasons.

First the dependent variable (the definition of debt used) varies largely across studies; second, also the inclusion of explanatory variables changes widely. What is mostly affected is the significance of some characteristics. In some cases results are very poor, in that estimates do not explain any of the variability of data. As it was said before, one reason could be the non linearity of the relationships. Also omissions of significant variables is crucial. A further problem with these studies is that some variables (such as profitability or leverage) included amongst the explicative are in fact endogenous

Chapter 6

Environmental liability and the capital structure of firms

Abstract.

A number of countries have recently introduced legislation which holds polluters liable for the costs of cleaning up environmental damage they have caused. While in principle this can give polluters appropriate incentives to reduce the risk of environmental damage, these incentives are weakened if polluters enjoy limited liability and can avoid paying large damages through bankruptcy (the 'judgement proofness' problem). This has led to suggestions that liability be extended to lenders such as banks, who will have incentives to condition loans on the effort made by firms to reduce environmental damage. However, this in turn can be confounded by problems of asymmetric information, and the costs imposed on banks in monitoring the environmental risks being incurred by firms. This in turn leads to fears that holding banks liable for environmental risks could substantially reduce the use of bank debt by firms.

In this paper we analyse the impact of different environmental liability regimes on the capital structure of firms, and in particular with how much bank debt they will use. We use a simple theoretical model to show that introducing environmental liability only on firms who have limited liability will cause them to increase their use of bank debt (essentially to protect their shareholders); extending liability also to banks has an ambiguous effect on bank borrowing, but under plausible assumptions will lead to lower bank borrowing than with liability only on firms, but higher or lower bank borrowing than with no liability at all.

The US has had the longest history of environmental liability legislation (CERCLA, 1980), and there have been differences over time in the extent to which banks have been held to be liable for environmental damages of insolvent firms. We use US industry-level data to estimate a reduced-form model of bank borrowing by polluters and show that the empirical model supports the theoretical findings. For industries which are heavily exposed to environmental liabilities, such as chemicals, we show that the introduction of environmental liability only on firms caused bank borrowing to increase by 15 - 20%, but when liability was extended to banks, borrowings returned to a level only slightly higher than with no liability. Our findings suggest that extending environmental liability to banks does not have drastic consequences for bank lending to firms.

6.1. Introduction.

Environmental policy has traditionally relied on *ex ante* regulation (emission standards, technology standards, economic instruments) to induce polluters to reduce their emissions. But recently some countries have moved towards *ex post* regulation, whereby polluters are made strictly liable for the costs of the environmental damage they have already caused. In principle, the anticipation of these liabilities will provide polluters with appropriate incentives to reduce the risk of environmental damage (this can be thought of as an application of the Coase Theorem). However, this incentive to reduce pollution is weakened if polluters enjoy limited financial liability and can avoid paying large damages by becoming insolvent (they become 'judgement proof').

One possible solution to this problem that has been proposed is to extend environmental liability also to lenders such as banks. Provided they are not likely to become bankrupt (they have 'deep pockets') they will have incentives to condition their loans on the efforts firms make to reduce the risk of environmental liabilities, and, again in principle this will ensure that efficient levels of pollution abatement are undertaken (Segerson (1993)). However, this solution is in turn problematic if there are significant asymmetries of information between banks and firms so that it is either impossible or expensive for banks to properly monitor environmental risks being run by firms[1]. If there are significant problems of asymmetric information then this may result in both inefficiently low levels of environmental care being taken by polluters[2] and also distortions to the capital structure of firms. If only firms are held liable for environmental damages, and they have limited liability, then firms may take on too much bank debt to protect shareholders against environmental risks. On the other hand if banks are also held liable, then firms may take on too little bank debt if either banks impose significant

[1] On the other hand, an argument sometimes advanced for imposing environmental liability on banks is that if banks are already monitoring conventional credit risks, then there may be economies of scope in extending this monitoring role to include environmental risks, and this could reduce the costs of monitoring using traditional *ex ante* regulation.

[2] See Pitchford (1995), Boyer and Laffont (1997), Balkenborg (1997) Lewis and Sappington (1997).

charges to cover monitoring costs or banks use credit rationing in response to the asymmetry of information.

The best known use of environmental liability legislation is in the USA which introduced the *Comprehensive Environmental Response Compensation and Liability Act* (CERCLA) in 1980. This Act holds 'potentially responsible parties' (PRP's) jointly and severally liable to pay for the clean-up of contaminated properties. PRP's are owners and operators of sites. The Act contained a 'Secured Creditor Exemption' whereby lenders were exempt from liability as long as the property was kept mainly as security for a loan. There were a number of tests of this exemption, and in 1992 the EPA issued a ruling which said that lenders could only be liable as owners if they foreclosed on a loan (which meant that lenders avoided foreclosure) or as operators if they were "involved in the day-to-day management of the insolvent firm". There were a few cases where this was invoked - e.g. where an employee of a bank was on the board of directors. But the case which seemed to open up banks more widely to liability was the *Fleet Factors* case (1990) where the judge ruled that banks could be held to have a "capacity to influence" the environmental risks run by operators. The argument was that lenders could not gain exemption from liability by pleading ignorance, since this would reduce the incentives for banks to carry out environmental audits as part of the normal prudential credit approval process. This seemed to imply that the normal monitoring of credit risks by banks would have to include environmental risks and hence give banks a knowledge of and therefore capacity to influence the steps taken by firms to affect environmental risks. Thus for a period of time in the US there was a view that banks as well as firms could be held liable for environmental damages.

In the UK the Environment Act (1995) introduced strict liability on PRP's for cleaning up contaminated land. The draft guidance notes for implementation of the Act apply six sequential exemption tests which appear to ensure that banks are exempt from liability (see Jewell (1997 a, b) for a legal analysis of the EA). However, these have still to be tested in court, and there is evidence that UK banks still believe there is a small risk they could be held liable for environmental damages of insolvent borrowers (McKenzie and Wolfe (1998)). Elsewhere in Europe there are variations in liability regimes, and while in principle lenders could be held liable on the same two grounds as in the US, no lender has yet been held liable. The EU has been considering proposals for harmonisation of liability regimes.

In this chapter we are concerned primarily with an empirical analysis of how the environmental liability regimes in the US have affected the capital structure of firms, and in particular the extent to which firms have relied on long-term bank borrowing. In section 6.2 we shall set out a simple theoretical model of firms' choice of capital structure and show how different environmental liability regimes affect bank borrowing. We shall show that if only firms are held liable, and they have limited financial liability, then firms will take on an inefficiently high level of bank debt, essentially to protect their shareholders from environmental liabilities. If firms and banks are held jointly liable, then we show that the effect on bank borrowing is ambiguous. However, on plausible assumptions the level of bank borrowing will be lower with joint liability than with liability only on firms; bank borrowing may be higher or lower with joint liability than when there is no liability.

In section 6.3 we use US industry-level data to estimate a reduced-form model of long-term bank debt over three time phases: (i) pre-1980 (i.e. pre-CERCLA): no liability; (ii) 1981-1990 (post-CERCLA, pre-*Fleet Factors*): liability only on firms; (iii) post-1991 (i.e. post-*Fleet Factors*): joint liability on firms and banks. We show that the empirical analysis strongly supports the theoretical predictions: bank borrowing rises when liability is imposed only on firms; when liability is extended to banks, then bank borrowing is lower than when liability is imposed only on firms, but is slightly higher than with no liability at all. For industries, such as Chemicals, which are heavily exposed to CERCLA liabilities, the model predicts that imposing environmental liabilities on firms would cause bank borrowing to rise by 15-20%, *ceteris paribus*, while if liability is imposed jointly on firms and banks then bank borrowing in such heavily exposed industries would only be about 1-2% above the level without any liability. Our results do not suggest that extending environmental liabilities to banks, which would be desirable in terms of improving the level of effort firms make to reduce environmental risks, would have a drastic effect on bank borrowing by firms.

6.2. The Theoretical Model

In this section we set out a simple theoretical model to show how different liability regimes affect the choice of capital structure by firms. The model extends the work of Feess and Hege (1997), which, to our knowledge, is the first paper which focuses explicitly on the impact of liability regimes on capital structure.

We analyse a firm which is deciding whether to undertake a project with investment cost I, and if so, how to finance it. There are three possible ways of funding the project: by equity, supplied by the owner-manager (hereafter owner) of the firm who has initial wealth W; or by debt which can take the form of either bank (private) debt or by public debt (corporate bonds). Both forms of debt are supplied on competitive markets[3], and we denote by D the amount of debt used. The two forms of debt are equivalent except for a lump sum benefit or cost of using bank debt rather than public debt, denoted by m which may reflect factors such as better monitoring or negotiation capacity of the bank or the higher cost of bank debt. Thus m may be positive (if the benefit of bank debt outweighs the cost) or negative (the reverse occurs). The fixed cost nature of this benefit/cost means that the firm will only choose one form of debt. Denote by Φ a variable which takes the value 1 if the firm chooses bank debt and the value 0 if it chooses public debt. Then $Z(\Phi, m) \equiv (2\Phi-1)m$ is the gain (or loss) to the owner from the choice of public or private debt. Finally $I - D$ is the amount of funds supplied by the owner to finance the project. The owner is assumed to be risk neutral.

Rather than explicitly model why the capital structure might matter to the firm, we simply assume that the return to the project, $R(D)$, is a strictly concave function of D, where $D^* = argmax(R(D))$, $0 < D^* < I$, is the optimal level of debt. This formulation is consistent with a wide range of theories of optimal capital structure, such as tax shields, signalling, costs of financial distress. Finally we assume that

[3] The assumption of competitive markets is significant. Heyes (1996) and Balkenborg (1997) show that the conclusions drawn by Pitchford (1995) about the implications of environmental liability in models of competitive banking with asymmetric information do not all carry through when the banking market is imperfectly competitive.

$I\text{-}D^* < W < I$ so that the owner can finance the equity required for the optimal capital structure, but not the whole project. Indeed we shall make the stronger assumption that for all relevant values of D, $W > I - D$; this assumption is made just to avoid having to worry about cases where the owner is wealth constrained.

We assume that the project carries a risk of causing environmental damage X, with probability $p(e)$, where $e \geq 0$ is the level of care expenditure made by the firm. For all $e \geq 0$, $1 \geq p(e) \geq 0$ and $p'(e) < 0$, $p''(e) > 0$. We suppose that the firm chooses e at the start of the project and that it is possible to condition debt or insurance contracts on e (though, as noted below, there may be costs to monitoring such effort).

We denote by $L \leq X$ the level of liability imposed by a regulator. The aim of the regulator is to maximise welfare, which is defined as profits earned by the owner minus expected environmental damage net of any liability payments. We shall consider in turn five possible regimes for how liability may be imposed: *Regime 0:* No liability; *Regime I:* Liability only on the firm with compulsory insurance with costless monitoring; *Regime II:* Liability only on the firm with compulsory insurance with costly monitoring; *Regime III:* Liability only on the firm with no requirement for insurance; *Regime IV:* Liability jointly on the firm and any bank which has lent to the firm, but no requirement for insurance.

6.2.1 *Regime 0:* No Liability

In this regime the firm is not held liable for any damages in the event of any environmental accident occurring, so clearly it will make no investment in care expenditure, i.e. $e = 0$. The owner has then to choose the capital structure (i.e. how much of the investment cost I is financed with debt D or with equity) and the form of debt, bank or public, Φ. D and Φ are chosen to maximise:

$$\Pi(D,\Phi) = [R(D) - D] - [I - D] + (2\Phi - 1)m \qquad (1)$$

where the first term is the value of the equity-holding of the owner, the second term is the equity invested by the owner, and the third term is the net gain/cost from the choice of private or public debt. We can rearrange (1) as:

$$\Pi(D,\Phi) = R(D) - I + (2\Phi - 1)m \qquad (2)$$

Given the separable nature of Π, the choice of D and Φ can be simply summarised by:

i) <u>choice of D</u>: D is chosen so as to maximise $\Pi(D, \Phi)$ or equivalently $R(D)$ so $D=D^*$;

ii) <u>choice of Φ</u>: $\Phi = 1$ if $m \geq 0$ and $\Phi = 0$ if $m < 0$. Call this assignment rule Φ^*. Then $Z(\Phi^*, m) = | m |$.

Thus without liability the firm earns profits: $\Pi_0 \equiv R(D^*) - I + |m|$ and society gets welfare $V_0 \equiv \Pi_0 - P(0)X$.

6.2.2 *Regime I:* Liability on Firm with Compulsory Insurance with Costless Monitoring

We assume that a regulator imposes a liability L on the firm, and requires that the firm can pay this liability through appropriate insurance. Before allowing the project to proceed the regulator requires that the firm demonstrates how its liability is to be insured. This can take the form of self-insurance, denoted by S, where the owner pledges to cover the amount S of the liability; the balance $L - S$ must then be insured externally, and we assume that there is a competitive market in insurance contracts, offered by either insurance companies or banks. For this section we assume that both banks and insurance companies can costlessly monitor the level of care expenditure, e, undertaken by the firm, and hence can costlessly observe $P(e)$, the probability of environmental liability being incurred. The level of self-insurance, S, is subject to an upper bound given by the available wealth of the owner, defined as:

$$\bar{S}(D) \equiv W - (I - D) + (R(D) - D) = W + R(D) - I \qquad (3)$$

There are three components of the owner's maximum available wealth: initial wealth W, less the costs of the funds put up to fund the project $(I - D)$, plus the equity returns on the project $(R(D) - D)$.

With our assumption of costless monitoring and competitive insurance markets, the cost to the owner of external insurance is just the actuarial premium $p(e)(L-S)$, while the cost of self-insurance is simply $p(e)S$.

The problem facing the owner can now be summarised as choosing the capital structure of the firm, D, the form of debt, Φ, the level of care expenditure, e, and the level of self insurance, S, to maximise:

$$\Pi(D,\Phi,e,S|L) \equiv [R(D)-D]-[I-D]-[p(e)S+p(e)(L-S)+e]+(2\Phi-1)m$$

$$\begin{aligned} &= R(D)-I-[p(e)S+p(e)(L-S)+e]+(2\Phi-1)m \\ &s.t. \quad S \leq \bar{S}(D) \end{aligned} \qquad (4)$$

It is clear from the formulation of (4) that the solution decomposes into three stages:

(i) Choice of e and S: the owner-manager chooses e and S to minimise:

$$\begin{aligned} &p(e)S+p(e)(L-S)+e = p(e)L+e \\ &s.t. \quad S \leq \bar{S}(D) \end{aligned}$$

The objective function does not depend on S since the owner is completely indifferent between self-insurance and external insurance. For simplicity we assume $S = 0$. e is then chosen to minimise $p(e)L + e$. Denote the solution by $E(L)$, where $E(L)$ satisfies the first order condition:

$$p'[E(L)]L+1=0 \qquad (5)$$

This is just the usual condition that the marginal cost of effort equals the marginal reduction in the expected cost of liability. Totally differentiating (5) yields:

$$E'(L) = -\frac{p'}{p''L} > 0 \qquad (6)$$

so that the optimal level of care is an increasing function of liability.

Finally denote by $K(L) = p(E(L))L + E(L)$ the minimum cost to the owner of facing liability L under costless mandatory insurance. Note that $K'(L) = p(E(L))>0$.

(ii) Choice of D. While the level of D affects the upper bound of self-insurance, we have seen that this plays no role, so D is chosen simply to maximise $R(D)$, i.e. $D = D^*$, the same as with no liability.

(iii) Choice of Φ. As with no liability the owner will choose Φ by the rule Φ^* i.e. $\Phi = 1$ if $m \geq 0$, $\Phi = 0$ if $m < 0$.

Thus for any given level of liability L with mandatory costless insurance the firm chooses $D = D^*$, $e = e^* = E(L)$, $S = 0$, $\Phi = \Phi^*$ and earns profits

$$\Pi^*(L) \equiv R(D^*) - I - K(L) + |m|$$

We now consider the problem of the regulator in setting L. This is chosen to maximise social welfare, defined as profits less the expected costs of damage not covered by liability payments, i.e. L is chosen to maximise:

$V(L) \equiv \Pi^*(L) - p(E(L))[X - L]$ for which the first order condition is:

$$\frac{d\Pi^*(L)}{dL} - p'E'[X - L] + p[E(L)] = 0 \qquad (7)$$

Since $\frac{d\Pi^*(L)}{dL} = -K'(L) = -p[E(L)]$ (7) reduces to

$$-p'E'[X - L] = 0 \quad i.e. \quad L = L^* = X$$

This is just the obvious condition that with costless mandatory insurance it is optimal to impose full liability on the firm.

Thus with costless mandatory insurance we get the **first-best** outcome:

(i) the externality is fully internalised as a liability to the firm ($L = L^* = X$);

(ii) the firm chooses effort so that $p'(e)X + 1 = 0$, i.e. the marginal cost of a bit more effort equals the marginal benefit of reduced expected environmental damage cost;

(iii) the firm makes the optimal choice of capital structure, $D = D^*$;

(iv) the firm makes the optimal choice of private or public debt, $\Phi = \Phi^*$

This is the first-best solution that would be obtained if the regulator had directly chosen D, e, Φ to maximise

$$R(D) - I - e - p(e)X + (2\Phi - 1)m$$

i.e. profits minus expected environmental damage costs. Indeed, the regulator only has to set liability $L = L^* = X$ and leave it to the profit-maximising owner to choose the first-best values for D, e and Φ.

Finally note that the project should only proceed if $V(L^*) \geq 0$, i.e.$\Pi^*(X) \geq 0$.

6.2.3 *Regime II:* Liability on Firm with Compulsory Insurance with Costly Monitoring

As in the previous section we assume that the regulator imposes a liability L on the owner in the event of environmental damage being incurred and requires that this risk is fully insured, either by self insurance or through external insurance contracts. But now we suppose that there is a moral hazard problem: banks and insurance companies cannot costlessly monitor e. If the owner has announced care expenditure $E(L)$, of which an amount S, $0 \leq S \leq L$ is self-insured, and $E(L)$ cannot be observed, the firm would have an incentive to carry out only expenditure level $E(S) \leq E(L)$. If the owner manager wants to commit to a level of effort $e > E(S)$, then banks and insurance companies will have to carry out monitoring activities which are costly[4].

We assume that banks and insurance companies incur monitoring costs

$$C(e,S) \equiv \gamma[e - E(S)] \qquad e \geq E(S)$$

where γ is an increasing strictly concave function ($\gamma' \geq 0,\ \gamma'' > 0$) with $\gamma(0)=\gamma'(0)=0$, i.e. the marginal cost of doing the first bit of monitoring is zero. While the cost function $C(e, S)$ is common to both banks and insurance companies, we also assume that there is a fixed cost difference between banks and insurance companies in monitoring costs denoted by δ, where δ > (<) 0 means that banks incur higher (lower) monitoring costs than insurance companies. Thus the cost of monitoring by a bank is: $\gamma(e - E(S)) + \delta$. Finally we assume that if a firm decides to borrow from a bank, then it has a choice of whether to insure from a bank or from an insurance company; while if the firm uses public debt, then it must use insurance companies to monitor its activities. This assumption is meant to capture the possibility that there may be economies of scope in a bank monitoring both the financial and environmental risks of a loan and the bank is only willing to incur the costs of monitoring the insurance risk if it is also monitoring the loan. Thus there is an additional benefit from using bank debt, that the firm may be able to reduce monitoring costs by an amount $b = min(0, \delta)$.

We continue to assume that the market for insurance by either banks or insurance companies is competitive, so that the firm has to pay the full cost of both the actuarially fair insurance premium and the cost of monitoring for any liability that is not self-insured. The problem for the owner faced with liability L and mandatory costly insurance is to choose D, Φ, e, S to maximise:

[4] Note that the assumption that e can be monitored, albeit at a cost, and therefore debt and insurance contracts can be conditioned on e, is an important one. Pitchford (1995), Balkenborg(1997), Heyes (1996), Boyer and Laffont (1997) analyse the case when there is a moral hazard problem but without the possibility of monitoring e, so any debt or insurance contract can condition payments only on whether or not an environmental accident occurs.

$$\hat{\Pi}(D,\Phi,e,S|L) \equiv R(D) - I - [p(e)L + e + \gamma[e - E(S)] + (2\Phi - 1)m - \Phi b \quad (8)$$

$$s.t. \qquad S \leq \bar{S}(D)$$

(8) differs from (4) by the inclusion of the cost of monitoring. Again it is possible to decompose the decision into three stages.

(i) Choice of e and S. For any given D, and hence $\bar{S}(D) = W - I + R(D)$, the firm chooses e and S to minimise:

$$p(e)L + e + \gamma[e - E(S)] \qquad s.t. \ S \leq \bar{S}(D)$$

It is obvious that to minimise these costs, the firm should choose the maximum level of self-insurance, since this minimises monitoring costs. This leaves the problem of choosing e to minimise: $p(e)L + e + \gamma[e - E(\bar{S}(D))]$ for which the first order condition is:

$$p'L + 1 + \gamma' = 0 \qquad (9)$$

Define the solution as $\hat{E}(L,D)$. Total differentiation of (9) shows that:

$$\frac{\partial \hat{E}}{\partial L} = \frac{-p'}{[p''L + \gamma'']} > 0, \qquad \frac{\partial \hat{E}}{\partial D} = \frac{\gamma''E'R'}{[p''L + \gamma'']}$$

So effort will be an increasing function of liability[5], but the effect of an increase in debt on care expenditure depends on the sign of R': if increasing debt increases R, and hence the maximum amount of self insurance, then care expenditure will increase. Note also that as long as $L > \bar{S}(D)$ then $\hat{E}(L,D) < E(L)$, i.e. the optimal effort level with costly insurance is less than with costless insurance, for the simple reason that the firm has to pay the monitoring costs of committing to higher levels of effort.

[5] Note that this result depends crucially on the assumption that effort can be monitored. As noted in footnote 1, in other models where effort cannot be monitored, payments can depend only on whether or not an environmental accident has occurred, and, as shown by Pitchford (1995), in such models increasing liability may reduce the care effort made by firms.

Define $\hat{K}(L,D) \equiv p[\hat{E}(L,D)]L + \hat{E}(L,D) + \gamma[\hat{E}(L,D) - E(\bar{S}(D))]$ as the minimum cost of insuring against liability L when the debt level of the firm is D. It is straightforward to see that:

$$\frac{\partial \hat{K}}{\partial D} = -\gamma' E' R', \qquad \frac{\partial \hat{K}}{\partial L} = p[\hat{E}(L,D)]$$

(ii) Choice of D. We can now solve the problem of the firm as choosing D and Φ to maximise:

$$R(D) - I - \hat{K}(L,D) + (2\Phi - 1)m - \Phi b$$

The first order condition is:

$$R'(D) - \frac{\partial \hat{K}}{\partial D} = R'(D) + \gamma' E' R'(D) = 0 \quad \Rightarrow R'(D) = 0 \;\Rightarrow\; D = D^*$$

Thus, since the firm wishes to maximise self insurance, and the limit on self insurance is increasing in the returns on the project, this simply reinforces the firm's wish to choose the capital structure which maximises the returns to the project. Because of mandatory insurance cover, there is no incentive to increase debt as a means of passing liability on to lenders. This would still need to be covered by mandatory insurance and increasing debt beyond D^* just reduces the scope for self-insurance and so increases the monitoring cost the firm has to pay.

(iii): Choice of Φ. If the firm chooses to use bank debt then it gets a net benefit of m-b where $b = min\{0, \delta\} \leq 0$; if it chooses to use public debt then it receives a net benefit of $-m$. So the firm will choose bank debt ($\Phi = 1$) if $m \geq b/2$, and public debt ($\Phi = 0$) if $m < b/2$. Call this assignment rule $\hat{\Phi}$. Since $b \leq 0$, then in general there will be a wider range of values of m for which bank debt will be chosen compared to the case of costless monitoring, because the choice of bank debt now offers a choice of monitoring agency which is not available if public debt is chosen. Using this rule the value of the net benefit of choosing the optimal type of debt is $\left|m - \frac{b}{2}\right|$.

To summarise, for any given liability level, L, with mandatory but costly insurance the firm chooses $D = D^*$, $e = \hat{E}(L, D^*), S = \bar{S}(D^*), \Phi = \hat{\Phi}$, and earns profits:

$$\hat{\Pi}(L) \equiv R(D^*) - I - \hat{K}(L, D^*) + \left| m - \frac{b}{2} \right|$$

The regulator will choose L to maximise $\hat{V}(L) = \hat{\Pi}(L) - p(\hat{E}(L, D^*))(X - L)$ for which the first order condition is:

$$\frac{\partial \hat{\Pi}}{\partial L} - p' \frac{\partial \hat{E}}{\partial L}(X - L) + p(\hat{E}(L, D^*)) = 0 \qquad (10)$$

Since $\frac{\partial \hat{\Pi}}{\partial L} = -\frac{\partial \hat{K}}{\partial L} = -p[\hat{E}(L, D^*)]$ (10) again gives the optimal rule for liability: $L = \hat{L} = X$. This is the optimal second best outcome in which with mandatory coverage and costly monitoring: there is full internalisation of environmental damage costs, the firm chooses the first best capital structure, as in the first-best with mandatory insurance and costless monitoring; but there is a lower level of care expenditure, $\hat{E}(X, D^*) < E(X)$, and there will be a wider range of values m for which bank debt is preferred to public debt, compared to the case of costless monitoring. Again this outcome is exactly the same as if the regulator chose all the variables e, S, D and Φ; so the second-best outcome can be achieved by just imposing mandatory insurance cover with full liability for environmental damage, and leave it to the firm to choose the second-best level of care expenditure and financial structure.

6.2.4 *Regime III*: Liability on Firm with No Mandatory Insurance

We now assume that in the event of environmental damage being incurred the firm is required to pay liability L, but is no longer required to insure this liability, either by itself (self-insurance) or by external parties. We continue to assume that an external insurance contract will involve paying costly monitoring services to overcome the moral hazard problem. Because of limited liability, the maximum amount for which the owner may be held directly liable is just the equity invested in the firm:

$$R(D) - D \leq \bar{S}(D) = W + R(D) - I$$

(since $I\text{-}D \leq W$). Since there is no requirement to take out insurance to cover any liability greater than $R(D)\text{-}D$, it is clear that the firm can simply ignore any liability greater than $R(D)\text{-}D$ and, since monitoring by external agencies is costly, has no incentive to take out any external insurance. Define $\Psi(L, D) = min\ (L, R(D)\text{-}D)$. Then, $S = \Psi(L, D)$, and the problem for the firm is to choose D, e and Φ to maximise:

$$\Pi(D, e, \Phi | L) \equiv R(D) - I - [p(e)\Psi(L, D) + e] + (2\Phi - 1)m \tag{11}$$

Again the solution can be decomposed as follows:

(i) <u>Choice of e</u>. Since (11) is similar to (4), it is clear that the optimal choice of e is given by $e = \tilde{E}(L, D) = E(\Psi(L, D))$. If $L < R(D) - D$, $\frac{\partial \tilde{E}}{\partial L} = E' > 0; \frac{\partial \tilde{E}}{\partial D} = 0$. While if $L \geq R(D) - D$ $\frac{\partial \tilde{E}}{\partial L} = 0$, $\frac{\partial \tilde{E}}{\partial D} = E'(R' - 1)$ which says that the effect of an increase in debt on care expenditure depends on the sign of $R' - 1$, i.e. if an increase in debt increases the equity value of the firm (because it increases returns more than it increases debt) than this will increase care, because more of the owner's equity is at risk. However if increasing debt reduces the equity value of the firm then that will reduce the care taken by the firm.

It is straightforward to see that $\tilde{E}(L, D) \leq \hat{E}(L, D) < E(L)$, where the first inequality arises because $\tilde{E}(L, D) = E(\Psi(L, D)) \leq E(\bar{S}(D)) \leq \hat{E}(L, D)$.

Finally define $\tilde{K}(L, D) = p[\tilde{E}(L, D)]\Psi(L, D) + \tilde{E}(L, D)$ as the minimum cost of facing environmental liability L when the firm has debt D and limited liability. If $L<R(D)\text{-}D$, $\frac{\partial \tilde{K}}{\partial L} = p(\tilde{E})$, $\frac{\partial \tilde{K}}{\partial D} = 0$, while if $L \geq R(D)\text{-}D$ $\frac{\partial \tilde{K}}{\partial L} = 0$, $\frac{\partial \tilde{K}}{\partial D} = p(\tilde{E})[R' - 1]$.

In the rest of this section we shall concentrate on the more interesting case where $L \geq R(D) - D$.

(ii) Choice of D. The firm chooses D to maximise: $R(D) - I - \widetilde{K}(L, D) + (2\Phi - 1)m$

for which the first order condition is: $R'(D) - \frac{\partial \widetilde{K}}{\partial D} = 0 \Rightarrow R'(D) = -\frac{p}{(1-p)} < 0$. So the optimal level of debt is $D = \widetilde{D} > D^*$. Thus if the liability is imposed only on the firm, with no requirement for mandatory insurance cover, and the firm can limit its liability to the equity holding, then there is a clear incentive to distort the capital structure in the direction of too much debt.

(iii) Choice of Φ. As in the first-best case, the firm will choose $\Phi = \Phi^*$, i.e. use bank debt if $m > 0$ and public debt if $m < 0$.

In summary, assuming that $L \geq R(\widetilde{D}) - \widetilde{D}$, the firm sets $e = \widetilde{E}(L, \widetilde{D})$, $D = \widetilde{D}$, $\Phi = \Phi^*$ and earns profits: $\widetilde{\Pi}(L) \equiv R(\widetilde{D}) - I - \widetilde{K}(L, \widetilde{D}) + |m|$ (12).

Since (12) does not depend on L, $\widetilde{\Pi}'(L) = 0$.

For $L \geq R(\widetilde{D}) - \widetilde{D}$, social welfare is $\widetilde{V}(L) \equiv \widetilde{\Pi}(L) - p[\widetilde{E}(L, \widetilde{D})][X - (R(\widetilde{D}) - \widetilde{D})]$. Thus any liability in the range $R(\widetilde{D}) - \widetilde{D} \leq L \leq X$ will lead to exactly the same outcome, and so the best the regulator can do is to achieve care level $\widetilde{E}(L, \widetilde{D})$. So if the only instrument available to the regulator is the level of liability, and if the regulator cannot impose mandatory insurance cover, then the level of care expenditure will be below the second best level, while the capital structure of the firm will be distorted to give too high a level of borrowing. The choice of bank or public debt is the same as the simple rule Φ^* which the firm followed with either no liability or in the first best case with liability and mandatory insurance with costless monitoring.

6.2.5 *Regime IV:* Joint Liability on Firm and Bank, with no mandatory insurance.

Finally we analyse the case where liability L is imposed jointly on the firm and bank lender, so that in the event that the liability cannot be paid by the owner, then it will be paid by the bank, if the firm has borrowed from a bank. Note that if the firm has used public debt, then it is assumed that it is not possible for individual holders of corporate debt to undertake any monitoring of the firm's behaviour, and so they would be exempt from any liability.

We follow Feess and Hege (1997) in noting that this case can be viewed as a combination of the two previous cases. If the firm decides to borrow from a bank, then the bank will insist that any liability to which it is exposed must be fully covered by insurance and the level of care expenditure appropriately monitored. The outcome will thus be as if we were in the regime of mandatory insurance with costly monitoring, except that the firm has chosen to use bank financing, i.e. $\Phi=1$. The solution, as in *Regime II,* will result in the firm setting $D=D^*$, $S=\overline{S}(D^*)$, $e=\hat{E}(L,D^*)$ and will yield the firm profits:

$$\overline{\Pi}^1(L) \equiv R(D^*) - I - \hat{K}(L,D^*) + m - b \tag{13}$$

(13) differs from (10) only in the last term because $\Phi = 1$, necessarily, rather than being set optimally.

On the other hand if the firm chooses public debt, then holders of public debt cannot be held liable, and then, as in *Regime III*, the owner can only be held liable to the extent of the equity invested in the firm. So if $\Phi = 0$ the firm will set $D=\widetilde{D}$, $S=R(\widetilde{D})-\widetilde{D}$, $e=\widetilde{E}(L,\widetilde{D})$ and earn profits:

$$\overline{\Pi}^0(L) \equiv R(\widetilde{D}) - I - \widetilde{K}(L,\widetilde{D}) - m \tag{14}$$

which again differs from (12) because Φ is necessarily 0 rather than being chosen optimally.

The only decision left now is for the firm to choose which form of debt it will use, and it will choose bank debt ($\Phi = 1$) iff:

$$R(D^*) - I - \hat{K}(L, D^*) + m - b \geq R(\tilde{D}) - I - \tilde{K}(L, \tilde{D}) - m$$
$$\text{i.e. iff} \quad m \geq \frac{b+\theta}{2} \quad \text{where} \quad \theta \equiv [R(\tilde{D}) - K(L, \tilde{D})] - [R(D^*) - K(L, D^*)] \tag{15}$$

Now it is straightforward to see that $\theta \geq 0$, since the second term in brackets is the result of choosing e and D to maximise:

$$G^1(e, D) \equiv R(D) - [p(e)L + \gamma(e - E(\bar{S}(D))) + e]$$

while the first term in brackets is the result of choosing e and D to maximise:

$$G^0(e, D) \equiv R(D) - [p(e)(\min\{L, (R(D) - D)\}) + e]$$

$\theta \geq 0$ because $G^0(e, D) \geq G^1(e, D)$, for two reasons: the limitation of potential liability the firm is exposed to, and the fact that there is no need to pay monitoring costs.

Now in the case of no liability (*Regime 0*), liability with mandatory insurance and costless monitoring (*Regime I*) and liability of the firm with no mandatory insurance (*Regime III*), the firm will choose bank debt iff $m \geq 0$; in the case of liability and mandatory insurance with costly monitoring (*Regime II*) the firm will choose bank debt iff $m \geq b/2$; so (15) tells us that with joint liability on the firm and the bank (*Regime IV*) the firm will use bank debt for a lower range of parameter values than *Regime II*, and will use it less (more) than in *Regimes 0, I* and *III* if $\theta \geq -b/2$ ($\theta < -b/2$); recall that $b = min(0, \delta)$ is non-positive.

Thus if lender liability is imposed on the firm and bank jointly, then, if the firm chooses bank debt we will get the second best level of care expenditure and the first-best capital structure, but if it chooses public debt, we will get inefficiently high level of debt and less than second-best level of care expenditure. The firm is more likely to choose public debt than in the second-best assignment rule.

6.2.6 Summary

This completes the analysis of the outcomes of the five liability regimes. We now summarise the implications of these results for the empirical analysis which follows. We shall model bank borrowing over three time periods: (i) pre-CERCLA, which corresponds to *Regime 0*: No Liability; (ii) post-CERCLA but pre-*Fleet-Factors*, which we shall take to correspond to *Regime III*: Liability only on Firm, with No Mandatory Insurance; and (iii) post-CERCLA and post-*Fleet-Factors* which we shall take to correspond to *Regime IV*: Joint Liability on Firm and Bank, with No Mandatory Insurance. We are interested in how the extent of bank borrowing by different industries varies across the three phases, or equivalently Liability Regimes, and the analysis shows that this depends on two factors: the level of debt, *D*, and whether borrowing was from banks or from the public, Φ.

The analysis shows that moving from phase (i) to phase (ii), i.e. when CERCLA introduced liability on firms only, but without mandatory insurance, should unambiguously increase bank borrowing: the firm will choose debt level $\widetilde{D} > D^*$ and will choose to use bank rather than public borrowing for the same parameter values as with no liability ($\Phi = 1$ iff $m \geq 0$). Moving from phase (ii) to phase (iii), by extending liability fully to banks, will have an ambiguous effect on bank borrowing; first, the level of borrowing will revert to the pre-liability (phase (i)) level of debt (D^*), so bank lending would fall; but, compared to the no liability case, the firm will make more or less use of bank borrowing depending on whether $\theta < -b/2$ or $\theta \geq -b/2$. In general one would expect that the unambiguously negative effect of the first factor would mean that bank lending would be less when liability is extended from the firm to include the bank, but the ambiguity of the second factor means that it is not possible to predict whether bank borrowing with joint liability would be greater or less than bank borrowing with no liability. If the second effect is weak, then one would expect bank borrowing to return to close to the pre-CERCLA level.

6.3. Empirical Analysis of the Effects of Environmental Liability on Bank Borrowing.

In this section we use US data to estimate the effect of both CERCLA and the *Fleet Factors* case on the level of industrial bank borrowing. We estimate a reduced-form model of bank borrowing over the period 1973-97 using data for 14 industries, and use dummies for both CERCLA and *Fleet Factors* to divide the period into three phases: (i) pre-CERCLA: No Liability; (ii) post-CERCLA, pre-*Fleet-Factors*: Liability only on Firms; (iii) post-CERCLA and post-*Fleet-Factors*: Joint Liability on Firms and Banks. To motivate our choice of variables for the reduced-form model of bank borrowing, we begin by briefly summarising some previous studies of the determinants of firms' capital structure.

6.3.1 Survey of Previous Studies of Firms' Capital Structure and of the Effect of Environmental Liability on Cost of Capital.

We focus on the relationship between firms' characteristics and firms' capital structure. The quantity of commercial loans taken out by firms is the equilibrium result of the profit-maximising behaviour of firms and banks, and in summary it is determined by three types of variables: *macroeconomic variables*, such as interest rates or GNP (affecting both demand and supply); *supply variables,* reflecting credit suppliers' characteristics and their preferences over customers' characteristics; *demand variables* reflecting firms' choice of capital structure. When estimating a reduced-form function to explain bank borrowing by firms these three sets of variables interact, sometimes with opposite effects on supply or demand. This may leave undetermined the net effect and thus the sign and the statistical significance of a reduced form regression.

Firms' capital structure concerns the proportion of equities over debt, the proportion of debt that is publicly issued or privately borrowed and the term structure of debt. The main decision about capital structure studied in the literature concerns the proportion of debt to borrow privately and the proportion to borrow publicly. Bank financing is considered less expensive than public debt, only larger firms thus rely on public debt, because they can spread fixed issuance costs over a larger amount. On the other hand monitoring costs from banks may make bank lending more expensive.

The demand for bank loans depends on interest rates, the costs of alternative funds and firms' characteristics. The supply of bank loans depends again on interest rates, deposits and capital requirements, banks' costs, and firms' characteristics, since banks will prefer to allocate to less risky firms than to more risky ones. We survey three studies: Johnson (1997), Houston and James (1996), Titman and Wessels (1988), which for brevity will be referred to as J, HJ and TW; these studies use reduced form regressions of the proportion of long-term private debt on firms' characteristics. Some of these characteristics are not directly observable and thus proxies or indicators are used instead. The firms' characteristics considered are the following:

Age: this is usually measured by the numbers of years since incorporation. Age is used as a proxy of reputation: older firms have developed a reputation for reliability and thus can rely more on public debt. On the contrary younger firms may be considered too costly to administer and monitor by the banks, and do not necessarily have to rely on bank debt. The relationship is thus that older firms have a higher proportion of public debt, while very young firms have a higher proportion of private non-bank debt. Middle aged firms should have a higher proportion of bank debt. J finds a positive coefficients of age in estimating long run public debt, while for bank debt the coefficient was not significant and for private non-bank debt it is negative.

Size: The leverage ratio may also be related to size as well as the term structure of debt. Small firms rely more heavily on bank debt than larger firms, and usually prefer short term debt rather than issuing longer term debt because of the lower fixed costs associated with this alternative. Size is usually measured by sales or total assets or quit rates (the idea is that larger firms provide better career opportunities and thus have lower quit rates). Size is also used as another proxy for reputation. TW first determine the relationship between logs of sales and quit rates and size with a factor analysis. They find that sales have a positive coefficient and quit rates have a negative coefficient in determining size. Second they estimate the impact of size (as estimated through factor analysis) on long term and short term debt, and find that it has a positive impact on long term debt and a negative one on short term debt. J measures size with total assets and age (older firms are usually larger) and finds that the coefficient of size is positive for public long term debt and bank long term debt. HJ measure size with the logs of assets and find a negative relationship with bank debt (confirming one of the explanations given above that larger firms rely more on public debt).

Collateral value of assets: firms need funds to finance their fixed assets, on the other hand fixed assets may be used as a collateral for default debt. The main variable used to proxy the collateral value of assets is the ratio of fixed assets (inventory, plant and equipment) over total assets. TW also use the ratio of intangible assets over total assets as collateral value but find a non significant coefficient for both long and short term debt. J finds a positive coefficient for public debt and for bank debt, and a negative one for non bank private debt. J also measures the collateral value of assets by leverage and market to book ratio. Leverage has a positive impact on public debt and on bank debt, and a negative impact on private non bank debt. It must be said however that leverage is not a predetermined variable, as it contains the amount of debt itself, and it is thus correlated with the residuals. The market to book ratio is non significant for public debt, positive for private non bank debt and negative for bank debt.

Non debt tax shield: tax deductions for depreciation and investment tax credit are substitutes for the tax benefits of debt financing. TW use tax credit over total assets or depreciation over total assets but they find a non significant coefficient for both cases in explaining long and short term debt.

Growth: TW measure growth by R&D as a percentage of sales or capital expenditure over total assets or with percentage variation of total assets; in all cases the coefficient for growth is positive for long term debt. By contrast, HJ measure growth by the market to book ratio and find a negative coefficient for growth (they also use R&D/S and still find a negative coefficient).

Volatility: this variable may have different impacts depending whether a demand or a supply function is estimated. On the one hand, firms with higher volatility of earnings would prefer to have the option to renegotiate the debt and thus prefer private debt to public debt, and would prefer a more secured long term debt than a short term one. On the other side lenders view more volatile firms as more credit risky because they may experience more states in which cash flows are too low to repay the debt and so they are less willing to lend to such firms. The net effect on the amount of bank loans thus is ambiguous. TW use the standard deviation of percentage changes in operating income but this variable is not significant for either long term or short term debt. J instead finds a negative coefficient for public debt (consistent with a demand function) and a non significant coefficient for private debt.

Uniqueness: the more unique the firm is the more specific are the workers skills and the capital characteristics. In the event of bankruptcy the more specialised the asset is the less its liquidation value and thus the lower the collateral value of fixed assets. These firms are considered more risky to lenders and thus they should find it more difficult to find loans. TW in fact find a negative coefficient of uniqueness for long and short term debts, measuring uniqueness by R&D/S, selling expenses, quit rates and a dummy variable which is the SIC classification. J instead uses a dummy variable which is the product of fixed asset ratio by the SIC classification, and he finds a negative coefficient of this variable in explaining public debt, while he finds it not significant for private debt.

Profitability: the past profitability and thus the amount of earnings that could be retained to refinance future projects should also be an important determinant of capital structure. This variable should have opposite effects on demand and supply. For demand, profitability is an alternative source of financing and thus should be negatively related to debt; for supply, profitability should be a positive indicator of good financial health of the firm and should be positively related to credit allocation. Indicators are operating income over sales and operating income over total assets. TW use both measures and find a negative relationship for long term debt and short term debt.

The studies surveyed in this section have used firm-level data, but as discussed below we were able to use only industry-level data. So some of the variables described above, such as age, size, or uniqueness will be difficult to capture with industry-level data, while for some others, such as volatility, there may be significant aggregation problems.

We now summarise some of the existing empirical studies of the impact of environmental liability on capital structure.

The impact of environmental liability on capital structure has recently been analysed by Garber and Hammit (1998) (GH), specifically, they investigate whether CERCLA has increased the cost of capital of firms belonging to the chemical industry. Because Superfund liability adds an element of risk to the securities of a firm, investors will require higher returns as compensation for the increased risk. Superfund should thus increase the cost of capital. GH analyse whether firms' equity betas (the sensitivity of stocks' returns to fluctuations in returns on market portfolios) vary with the level of exposure to Superfund liability. For this they use six alternative exposure measures: 1) the number of NPL (National Priority List) sites at which a company is named, 2) the number of sites proposed for NPL at which a company is named, 3) measure 1 divided by real market equity, 4) measure 2 divided by real market equity, 5) the sum over sites in measure 1 of firm's shares of market equities of large PRPs at site, 6) sum over sites in measure 2 of firm's shares of market equities of large PRPs at site. Estimates from three sets of regressions are reported: in a first set all firms are included, but results about the impact of Superfund on return rates are ambiguous; in fact Superfund increases significantly the cost of capital only for the first two exposure measures used . In the other two sets of regressions GH divide the sample of firms between High market equity firms and Low market equity firms. For the first subsample Superfund increases significantly the cost of capital always, for all six exposure measures, while for the second subsample the coefficient relative to Superfund liability is not significant for any of the six exposure measures. It seems thus that investors of firms belonging to the first subsample are aware of

Superfund liability and require higher returns for this, while investors of firms in the second subsample are not aware of Superfund liability.

An obvious difference with our paper is the dependent variable: GH want to explain the impact that CERCLA had on the cost of capital, and for this they analyse the rates of returns on securities. Conversely we want to explain the impact of CERCLA on the allocation of credit, and our dependent variable is the proportion of bank lending. Moreover, GH do not want to explain the cost of capital in terms of firms' characteristics; their only explanatory variables are market rates of returns and Superfund exposure measure. We instead try to explain how bank loans depend on the characteristics of the industries. The last difference is the level of data aggregation and the sample considered: GH's sample comprises firms that belong to the chemical industry only, but they have monthly observations at firm level. We couldn't rely on this level of data but we had quarterly observations for 14 industries.

In a paper related to the previous one Barth and McNichols (1994) (BM) investigate the relationship between firms' share prices and estimates of Superfund liabilities. BM regress the firms' market value of equities over firms' assets, firms' liabilities and a proxy for environmental liability. With no disturbances or omissions the market value of equities should be equal to assets minus total liabilities, this is not the case in the analysis reported. However, there is a problem because they still estmate an accounting identity and do not explain what determines the market value of assets and the role that environmental liability plays in it. However, in their analysis environmental liability is measured with alternative proxies: number of sites at which a firm is PRP, the sum of estimated clean up costs across all sites at which a firm is PRP (treating each PRP as 100% liable), partial allocation of remediation costs for each site at which a firm is PRP summed across all sites (PRP partially liable). All environmental proxies are significantly negative for cross-section time series analysis, while large percentages of them are significant for year by year cross section regressions. The liability measure with highest explanantory power (highest t-statistics and R-squared) is the number of sites at which a firm is nominated PRP.

6.3.2 Description of Data and Variables

We attempted to obtain US firm-level data on capital structure, but it is not possible to obtain published data on firms' bank borrowings. Such data has to be collected from the financial accounts of individual firms, and we did not have the resources to obtain such data. Nor were we able to gain access to the data sets of researchers who had collected such data. Thus we were forced to rely on industry-level data on bank borrowing; data on bank borrowing and other non-environmental variables was made available from Standard and Poor's. These consisted of quarterly observations from 1973 last quarter to 1997 second quarter at industrial level. The 14 industries, classified according to SIC code, for which data were available are Food and Tobacco, Textiles, Paper, Printing and Publishing, Chemicals, Petroleum and Coal, Rubber and Plastics, Stone Clay and Glass, Primary Metals, Fabricated Metals, Non-Electrical Machinery, Electric and Electronic Equipment, Transportation Equipment, Instruments.

We next capture two different aspects of the impact of environmental liabilities on bank borrowing. The first aspect is a measure of the exposure of the different industries to clean-up costs under CERCLA. The only observation we had were clean-up costs directly related to CERCLA for 1996(2) for the following seven industries: Food and Tobacco, Chemicals, Petroleum and Coal, Primary Metals, Fabricated Metals, Non-Electrical Machinery, Electric and Electronic Equipment. This was available from Resources for the Future. This variable was scaled to the size of the industry by dividing it by the value of total industry assets in 1996 (2).

The second aspect we capture is the different liability regimes, as set out in the three phases noted at the start of this section. We captured this by the use of dummy variables. The first is a dummy variable (called CERCLA1) that takes the value zero from 1973 (4) to 1980 (4), the value one from 1981(1) to 1991(4), and the value zero thereafter. This variable is designed to capture the impact of the introduction only of CERCLA with environmental liability imposed only on firms. The second is a dummy variable (called CERCLA2) that takes the value zero from 1973(4) to 1991(4) and one thereafter, and is designed to capture the impact of *Fleet Factors*, making both firms and banks jointly liable.

These two sets of variables are then multiplied together. So for industries for which there is no data on clean-up costs, the environmental dummies are zero throughout; for industries for which we had data on clean-up costs, the dummy variables are either zero or positive, with the size of the positive variable reflecting the exposure of the industry to CERCLA liabilities.

We now describe the remaining non-environmental variables. The dependent variable is long term bank debt as a proportion of total liabilities. The (non-environmental) independent variables are:

- *interest rate*: long term real interest rates (three years maturity bonds); this would be expected to have a negative effect on demand for bank borrowing, a positive effect on supply, and so an ambiguous effect in a reduced-form equation; clearly it is also an endogenous variable;

- *monetary base*: this is expected to have a positive effect on the supply of bank lending and hence on the reduced-form;

- *profitability*: retained income over sales (net operating revenue). As discussed in 6.3.1, this represents an alternative source of funds to firms and so should reduce the demand for bank loans; on the other hand it is a measure of the financial health of the sector and so may make banks more willing to lend. So the overall effect in a reduced-form equation is ambiguous. Moreover, as the theoretical model makes clear, differences in environmental liability regime will affect firms' profits, so profitability may not be considered to be fully exogenous.

- *equities*: measured as stockholders equities over total assets. While this is not exactly the complement of bank debt, as in the capital structure of industries there is bank debt, non-bank debt and stockholders equities, it represents an alternative source of funds and so will reduce the demand for bank debt. However, it will be determined simultaneously with the level of bank debt.

- *fixed assets ratio*: measured as the value of property, plant and equipment as a proportion of total assets. This represents the collateral value of assets and as noted in 6.3.1 this should increase the supply of bank loans. It could also increase the demand for loans as to finance the capital investment. However, as noted above, the collateral value of assets diminishes the more highly specialised they are. It is difficult to capture this aspect with industry-level data.

- *current assets ratio*: measured by cash and securities divided by total current assets. This variable represents the liquidity and we would expect a negative impact on the demand for bank debt. However this is more correlated with short term debt than with long term debt.

- *working capital ratio*: measured as working capital divided by total assets; plays a similar role to current assets ratio.

- *growth*: this is designed to capture growth of investment opportunities. We used three different definitions: growth of total assets, growth of capital stock and depreciation over total assets; the results reported use the first definition; we obtained similar results with other definitions. Firms with high growth of investment opportunities will need more long-term debt in order to invest, so this will increase the demand for debt. Lenders could also consider firms (or industries) with high growth as firms (industries) with high future profitability potential and may be more willing to lend.

- *volatility*: measured by the standard deviation of profitability over 12 periods. As noted in 6.3.1, this variable is a proxy of credit riskiness and likelihood of financial distress: more volatile firms are more credit risky because they may experience more states in which cash flows are too low to repay the debt. The supply of bank debt is thus a decreasing function of volatility. On the demand side, more volatile firms would prefer to have the option to renegotiate the debt and thus prefer private to public debt. However there is an issue whether this variable is sensibly captured at industry level, since a lot of firm-specific volatility could be smoothed out by aggregation.

Before conducting the regressions, we carried out integration tests on all the variables and found that they are not integrated.

The relationships between the explanatory variables and demand and supply of bank loans are summarised in Table 6.1.

6.3.3 Regression results

Data are cross section-time series so fixed-factors panel regressions are estimated. Moreover, because the dependent variable lies between 0 and 1 a logit model is also estimated. This implies that coefficients do not really measure the direct effect of the explanatory variables on bank debt but rather on a transformed variable. The sign of the coefficient has the same meaning of a negative or positive impact of the explanatory variables on the dependent variables. While the coefficients of a non transformed regression would still be unbiased[6], the transformation was necessary to test the significance of the coefficients. Because, as noted in 6.3.2, there may be problems of endogeneity of some variables, in particular interest rates, profitability and equities, instrumental variable estimation has been used; instruments are lagged values of these variables. This does not change the size and significance of the coefficients significantly. Finally, consistent GLS estimation[7] was employed on the logit model.

We employed two variants of the model. The first included interest rates as a macroeconomic variable. However, as noted above, it may be considered inappropriate to include interest rates as an independent variable in a reduced-form regression (even if instrumented), and as an alternative we used a variant with the monetary base as a macroeconomic variable. For each variant we ran three models: Model I is a base case without any environmental liability variables; Model I introduces just the environmental variable CERCLA1; Model III introduces both environmental variables CERCLA1 and CERCLA2. Tables 6.2 and 6.3 present the results of the consistent GLS logit estimations with instrumental variables for Model I - III using the interest-rate and monetary-base variants respectively.

[6] It is possible then to use the coefficient of the untransformed regression to tell us something more about the impact of the independent variables on the dependent one. For this reason, simple OLS panels were also performed. The OLS results are not reported here.

[7] These are performed by first estimating the variance-covariance matrix using OLS and then estimating the coefficients using GLS; the process is then iterated to get better results.

The underlying reduced-form model of bank borrowing is well-behaved: where we have been able to predict the sign of a variable, that variable has the correct sign (monetary base, equities, fixed asset ratio, current asset ratio, working capital ratio, growth); in all cases almost all variables are significant, at least at 10% and mostly at 5%; and the coefficients are stable across the different models. The empirical results also strongly confirm the theory about the impact of environmental liability on bank borrowing. The introduction of liability only on the firm has a strong positive effect on bank borrowing. When joint liability on firms and banks was introduced following *Fleet Factors* bank borrowing fell; in the interest rate model the coefficient on CERCLA2 remains positive and significant, though smaller than on CERCLA2, while in the monetary base model, the coefficient on CERCLA2 is positive but insignificant, suggesting that borrowing fell back close to its pre-CERCLA level.

6.3.4 The Impact of CERCLA on US Bank Borrowing

While the regressions confirm the theory set out in section 2, as presented they give no indication of the scale of the impact of environmental liability on bank borrowing in different industries. To assess this we have used the model with monetary base to predict the effect of the different liability regimes on bank borrowing in different industries, *ceteris paribus*. What we have done is to set all the independent financial variables for each industry at their mean values; denote the vector of the variables for industry i by $\overline{X}_i$. We have then used the estimated coefficients (denoted $\hat{\beta}$ of the monetary-base variant of Model III to calculate the following three estimates of expected bank borrowing in industry i with (i) no liability; (ii) liability only on firms; (iii) joint liability on firms and banks:

$$y_i^1 = \hat{\beta}_{0i} + \hat{\beta}_1\overline{X}_i$$
$$y_i^2 = \hat{\beta}_{0i} + \hat{\beta}_1\overline{X}_i + \hat{\beta}_2 CERCLA1_i$$
$$y_i^3 = \hat{\beta}_{0i} + \hat{\beta}_1\overline{X}_i + \hat{\beta}_3 CERCA2_i$$

and

$$z_i = \frac{\exp(y_i)}{1+\exp(y_i)}$$

where y is the estimated logit transformation of the proportion of bank debt and z is the proportion of bank debt.

Note that it is only the constant-term which has an industry-specific estimated coefficient. The resulting three estimates of expected bank lending for the seven industries which have non-zero environmental liability variables are shown in columns (1) to (3) of Table 6.4. Columns (4) and (5) then show the percentage difference in expected bank borrowing from the no liability base case that would result from imposing environmental liability only on firms, and jointly on firms and banks respectively. What this shows is that in industries such as chemicals and primary metals which have high CERCLA liabilities, the introduction of liability on firms caused a significant increase in bank borrowing (above 15%) as firms tried to protect equity-holders from these liabilities. However, if liability is imposed jointly on firms and banks, then the impact on bank borrowing relative to no liability is trivial.

6.4. Conclusions

In this paper we have analysed the consequences for the capital structure of firms, and particularly their use of bank borrowing, of imposing legal environmental liability either on firms alone or jointly on firms and banks. We used a simple theoretical model to show that the imposition of liability only on firms would lead to an inefficiently high use of bank borrowing, as well as an inefficiently low level of effort to reduce environmental damage, while imposing liability jointly on firms and banks had a more ambiguous effect on bank borrowing, but under plausible assumptions would lead to lower bank borrowings than when liability is imposed only on firms, and could lead to higher or lower bank borrowings than with no liability at all. We then used US industry-level data to estimate a reduced-form model of bank borrowing which confirmed the theoretical findings and showed that for industries heavily exposed to environmental liabilities the imposition of liabilities on firms only had caused bank borrowing to rise by 15 - 20%, but then when liability was extended to banks, borrowings fell back almost to the level expected without any environmental liabilities.

The research was constrained by the difficulty of getting access to firm-level data, and it would be interesting to check if the results are upheld using such data. Further work could also be done on the dynamic aspects of the econometric model using industry-level data. Nevertheless, we believe that our work provides a useful empirical analysis of the impact of environmental liabilities on financial structures of firms. Moreover our findings suggest that the extension of environmental liability from firms to banks, which would be desirable on the grounds of improving the incentives for firms to take adequate steps to reduce environmental damage, may not have the drastic consequences for bank lending to firms that some commentators had feared.

Table 6.1: Relationship between bank debt and financial variables

Variable	Demand	Supply	Net effect
Interest Rate	-	+	
Monetary Base		+	+
Profitability	-	+	
Equities	-		-
Fixed Asset Ratio	+	+	+
Current Asset Ratio	-		-
Working Capital	-		-
Growth	+	+	+
Volatility	+	-	
CERCLA1	+		+
CERCLA2			

Table 6.2: Estimates of Models I - III Using Interest Rate

Variable	Model I	Model II	Model II
Interest Rate	-0.008*	-0.016*	-0.016*
Profitability	0.546*	0.726*	0.716*
Equities	-2.059*	-1.967*	-1.838*
Fixed Asset Ratio	2.915*	2.831*	2.863*
Current Asset Ratio	-0.070	-0.056	-0.062
Working Capital	-1.957*	-1.816*	-1.823*
Growth	0.144	0.169°	0.174°
Volatility	-0.172	-0.168°	-0.172°
CERCLA1	-	4.139*	5.362*
CERCLA2	-	-	2.130*
Adjusted R^2	0.824	0.834	0.835

Table 6.3: Estimates of Models I - III Using Monetary Base

Variable	Model I	Model II	Model II
Monetary Base	1.016*	1.443*	1.414*
Profitability	0.353	0.476°	0.479°
Equities	-1.700*	-1.513*	-1.499*
Fixed Asset Ratio	3.154*	3.010*	3.005*
Current Asset Ratio	-0.125*	-0.120*	-0.120*
Working Capital	-1.651*	-1.307*	-1.311*
Growth	0.165°	0.195*	0.196*
Volatility	-0.276*	-0.303*	-0.300*
CERCLA1	-	4.673*	5.043*
CERCLA2	-	-	0.701
Adjusted R^2	0.829	0.841	0.841

* Significant at 5%
° Significant at 10%

Table 6.4 Impact of CERCLA on Bank Borrowing

	(1)	(2)	(3)	(4)	(5)
Industry	**Model I**	**Model II**	**Model III**	**(2)/(1) %**	**(3)/(1) %**
Food and Tobacco	**10.577**	**11.061**	**10.619**	**4.6**	**0.8**
Chemicals	**6.776**	**7.965**	**6.873**	**17.5**	**1.4**
Petroleum & Coal	**5.387**	**5.638**	**5.409**	**4.7**	**0.4**
Primary Metals	**8.103**	**9.367**	**8.207**	**15.6**	**1.3**
Fabricated Metals	**14.865**	**16.139**	**14.973**	**8.6**	**0.7**
Non-Elec. Equip.	**10.206**	**10.211**	**10.207**	**0.5**	**0.0**
Electr. Equip.	**6.224**	**6.305**	**6.231**	**1.3**	**0.1**

Chapter 7

Conclusions and directions for further research

This thesis has discussed two main issues: first the impacts of *ex-ante* environmental regulation on firms' location decisions and vice-versa, and second the impacts of *ex-post* environmental regulation on firms' capital structure.

In the second chapter some of the concerns of the effects of trade and capital movements liberalisation, often expressed by policy makers and by environmentalists, have been formally addressed. More precisely the second chapter reviews the literature on the impacts of free trade on firms' location and on environmental policies with particular reference to multinational companies. It has been shown how in general economic theory does not provide a substantial ground to support those concerns. In particular there can be no general presumption that governments will engage into a 'race to the bottom' in order to prevent capital flight. However a single model that can incorporate all these concerns does not exist yet and most of the results and conclusions reached depend very much on the assumptions and are not robust to changes.

In the third chapter we analysed the impacts of imposing environmental taxes on plant location decisions when there are positive transport costs, more than one industry and more than one firm per industry. We intended to focus mainly on the impacts of environmental regulation when there are incentives to agglomerate. A particular feature is that welfare functions are not continuous. Two important conclusions are reached: first a toughening of the environmental policy by one country does not mean that that country is necessarily going to lose all its production, and second that in any case there exists a threshold level on the policy instrument above which a country may instead lose all its industrial base. Even in this case it does not follow that a country is worse off by not having production in its territory. Last, while there are cases in which governments will engage in some environmental dumping, still there are other cases in which governments will set too tough environmental policies (a typical example is the NIMBY case).

In the fourth chapter we have shown with a theoretical model that the policy governments implement when they want to affect location is not necessarily different from the one they choose when location is fixed. This result was also supported by numerical simulations. In that paper we showed that it is important to distinguish between a difference in the move structure and a difference in the number of instruments available. And indeed it was a difference in the number of instruments that mostly generated the difference in the policies set by governments rather than the timing of governments setting policies relative to firms locating plants. We have also been able to show that the extent to which government engage in environmental dumping is not necessarily greater when firms are footlose than when location is fixed. Last, it would be natural to think that using one instrument would encourage governments to weaken environmental policies more than using two instruments. Again, we have shown that this is not necessarily the case. An example is when the outcome with one instrument is the NIMBY case while the outcome with two instruments is the Nash equilibrium or the constrained Nash equilibrium. With the help of numerical simulations we assessed the probability of occurrence of the different outcomes for the two types of games. For the Market Share game (when firms decide location before governments set environmental policies) the predominant outcome is firms locating in two countries and governments setting Nash equilibria emission limits. This outcome increases as the degree of substitutability between products falls. The reason is that as competition becomes weaker the more independent the two products are. For the Location game (when governments set environmental policies before firms locate), when governments use two instruments the Nash equilibrium with firms locating in separate countries is also the dominant equilibrium. When governments use a single instrument then the Race-to-the-Bottom occurs in most of the cases.

In the second chapter some empirical literature was reviewed that estimated the impacts of environmental policies on plant location. This literature in general found no significant impact. It has been observed that a possible reason for the discrepancy between theoretical conclusions and empirical observations may rely on the fact that the empirical models did not really capture firms' footlooseness. Moreover, it has also been observed that firms base their decisions on expectations about future policies and the existing models do not take this into account. A possible extension to the literature is to test the impacts of environmental policies on plant location taking into account expectations, proper footlooseness of firms, agglomeration incentives and political dimension. If the concern is that there has been a capital flight towards pollution havens, this may imply to test whether firms have moved production from more developed countries to less developed countries where environmental regulation is less stringent.

The analysis briefly outlined in this section can be extended in other different ways. One possible extension is to move to a proper dynamic analysis, where it is possible to design different degrees of commitments between firms and governments decisions by using different concepts of strategies. For example, if locating a plant implies an initial irreversible investment, then it seems reasonable to model firms' location decisions with open loop strategies. If on the other hand a situation is modelled where investment decisions are flexible and can be changed at any time, then feedback strategies could be used instead. Similarly, if governments are committed to adopt particular environmental policies, for example because of the existence of international agreements, then again open loop strategies could suit as well. And again, if on the contrary governments are not able to commit themselves, then it seems more appropriate to model environmental policy decisions with feedback strategies. Clearly the outcome depends on what strategy concept is adopted. However a problem that we would expect is that this type of dynamic formulation, and in particular the adoption of feedback strategies, is likely to generate a relatively complicated model to solve and the need to rely on numerical simulations. The main advantage of using feedback strategies is that these are both credible and time consistent.

With regard to the model presented in the fourth chapter, in order to keep our formulation as simple as possible we abstracted from problems like consumers' surplus, or profits distributions. It is known that, if countries have also to worry about consumers' surplus, they will set laxer policies when firms compete in a Cournot fashion. This creates a further incentive for environmental dumping. Also, we considered environmental policies in form of emission limits. If taxes are used instead, then the incentive for environmental dumping increases since taxes generate government revenue.

The second part of the thesis focused on the analysis of a particular type of *ex-post* regulation, which is the financial liability of firms causing environmental damage and what the effects are of possibly extending this financial liability also to firms' creditors. We first derived a theoretical model of firms' capital structure distinguishing between own equities, private debt and public debt. In this model we distinguished between different regimes of liability to capture the various stages of legislation as it occurred in the USA. The implications of the model are that when only firms are held liable for the environmental damage then the recurrence to bank borrowing increases. When liability is fully extended also to banks then the effect on bank borrowing is ambiguous. The first effect is that the level of total borrowing will go back to the level before any liability was imposed. However the composition of debt may change compared to the no-liability situation, and private debt may increase or decrease relative to public debt (always comparing the second liability regime with the no-liability regime) depending on the values of the parameters. In fact it may happen that the composition of debt returns to the 'old' composition as in the pre-liability situation. This theoretical model has been tested empirically. We used US data at industry level to estimate the coefficients of a reduced form regression, where the proportion of long run private debt was explained as a function of the characteristics of the industries and of two environmental dummies, capturing the effect of different periods of liability. The coefficient of the first dummy, representing firms' liability, was significantly positive as predicted by the model. The coefficient of the second environmental dummy, representing joint liability of firms and banks, was positive but lower than the coefficient of the first environmental dummy, meaning that bank debt had decreased compared to the firms' liability case. For some specifications of the regression this coefficient was non significant, meaning that bank lending had gone back to the pre-CERCLA levels.

In our theoretical model we focussed our attention on the behaviour of the firm only; bank's monitoring solved the problems generated by asymmetric information. A possible extension is to consider a game between banks and firms and model more explicitly asymmetric information and monitoring costs. Banks and firms could now choose different degrees of monitoring rather than just a 0-1 type of monitoring choice. A further extension is a dynamic analysis where one could think of a situation in which the firm decides its optimal capital structure in a first stage and thus establishes a credit relationship with a bank. Environmental liability is subsequently introduced and banks and firms renegotiate the credit relationship. The question is to see how joint environmental liability changes demand and supply of private debt over a long time horizon.

In the sixth chapter it was assumed that markets for both private and public debt are perfectly competitive. The impacts of imposing financial liability on firms and their lenders have been tested under this assumption. Another possible extension is to assume that the banking system is imperfectly competitive and test the impacts of the different liability regimes under this assumption. This would allow cross-countries comparisons, for countries which have different banks' market structures. This would particularly be the case in the European Union.

With regard to imperfect competition, a further analysis is to consider the possibility for the firm to establish multiple banks relationships. In the presence of asymmetric information the bank with which the firm has already established a credit relationship may exercise a form of monopoly power generated by the better information it has compared to rival banks. If the bank becomes jointly and severally liable, on the one hand it may find it convenient if the firm establishes multiple banks relationships in order to share the environmental risk also with the other banks. On the other hand the existing bank could find it convenient if the firm keeps a single bank relationship first to keep the monopoly power generated by the better information, and second to have better control on the firm's operations, since now the bank is responsible for them as well.

The empirical research has been highly constrained by data availability. In our analysis we have seen if imposing environmental liability has had any impact on the observed quantity of credit at industry level. Results are quite promising, and for this it would be very interesting to test the model using firm-level data. It would be ideal to have information on both demand and supply of bank loans, to see if and how the allocation of credit has changed according to environmental riskiness of firms. This would be a more appropriate analysis, and therefore test, of the possibility of credit rationing to due the environmental legislation.

REFERENCES

Anderson, K. and R. Blackhurst (eds.), (1992): *The Greening of World Trade*, Harvester Wheasheaf, Hemel Hempstead.

Avery, R.B., (1981): "Estimating Credit Constraints by Switchng Regressions", in C. F. Manski, D. McFadden *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press: 435-472

Avery, R.B. and A. N. Berger, (1991): "Loan exposure and bank risk exposure" in *Journal of Banking and Finance*, vol 15: 173-192

Balkenborg, D., (1997): "Bargaining Power, Judgement-Proofness and the Impact of Vicarious Liability", *University of Southampton Discussion Paper in Economics and Econometrics 9709.*

Baltagi, B., (1995): *Econometric Analysis of Panel Data,* J. Wiley and Sons.

Barclay, M.J. and C.W. Smith JR., (1995): "The Maturity Structure of Corporate Debt", in *The Journal of Finance*, col. L, n. 2: 609-631

Barrett, S., (1994): "Strategic environmental policy and international trade", in *Journal of Public Economics*, 54, 325-338.

Barth, M. E. and McNichols M. F., (1994): "Estimation and Market Valuation of Environmental Liabilities Relating to Superfund Sites", in *Journal of Accounting Research,* 32, 177-209.

Berger, A. N. and G.F. Udell (1992): "Some Evidence on the Empirical Significance of Credit Rationing", in *Journal of Political Economy,* vol.100, n.5:1047-1077.

Bhagwati, J. and R. Hudec (eds.) (1996): *Fair Trade and Harmonisation*, MIT Press, Cambridge, Mass.

Boyer, M. and J-J. Laffont (1996): "Toward a Political Theory of Environmental Policy", *Nota do Lavoro* 56.96, FEEM, Milan.

Boyer, M. and J-J Laffont (1997): "Environmental Risks and Bank Liability" *Europeqn Economic Review*, 41(8) 1427 - 1460

Cox, D. and T. Jappelli, (1993): "The Effect of Borrowing Constraints on Consumer Liability", in *Journal of Money, Credit and Banking*, 25(2), 197-213.

Cropper, M. and W. Oates (1992): "Environmental Economics: A Survey" *Journal of Economic Literature* , **30**, 675 - 740.

Easterwood, J. C. and P.R. Kadapakkam, (1991): "The Role of Private and Public Debt in Corporate Capital Structures", in *Financial Management*: 49-57

Feenstra, T. (1998): *Environmental Policy Instruments and International Rivalry - A Dynamic Analysis*, Tilburg University.

Feess, E. and U. Hege (1997): "Lender Liability, Capital Structure and Mandatory Coverage", *mimeo*, University of Frankfurt.

Garber, S. and J.K Hammitt (1998): "Risk premiums for Environmental Liability: Does Superfund Increase the Cost of Capital?" *Journal of Environmental Economics and Management*, 36 (3), 267 - 294

Frederikson, P. (1997a): "The Political Economy of Pollution Taxes in a Small Open Economy", *Journal of Environmental Economics and Management*, 33, 44-58.

Frederikson, P. (1997b): "Why Don't Environmental Regulations Influence Trade Patterns?", *mimeo*, World Bank.

Grossman, G. and A. Krueger (1993): "Environmental Impacts of a North American Free Trade Agreement", in P. Garber (ed.) *The Mexico-US Free Trade Agreement*, MIT Press, Cambridge, Mass.

Grossman, G. and G. Maggi (1998): "Free Trade versus Strategic Trade: A Peek into Pandora's Box", *CEPR Discussion Paper 1784*, London.

Heyes, A. (1996): "Lender Penalty for Environmental Damage and Equilibrium Cost of Capital", *Economica*, 63, 311 -323

Hillman, A. and H. Ursprung (1992): "The Influence of Environmental Concerns on the Political Determination of Trade Policy", in K. Anderson and R. Blackhurst (eds.) *The Greening of World Trade Issues*, Harvester-Wheatsheaf, HemelHempstead, 195-220.

Hillman, A. and H. Ursprung (1994): "Greens, Supergreens and International Trade Policy: Environmental Concerns and Protectionism", in C. Carraro (ed.) *Trade, Innovation, Environment*, Kluwer, Dordrecht, 75-108.

Hoel, M. (1997): "Environmental Policy with Endogenous Plant Locations", *Scandinavian Journal of Economics*, 99(2), 241-259.

Houston J., C. James, (1996): "Bank Information Monopolies and the Mix of Private and Public Debt Claims", in *The Journal of Finance*, vol. LI, n. 5: 1863-1889

Jaffee D. M. (1971): *Credit Rationing and the Commercial Loan Market,* Wiley & Sons

Jaffe, A. B. and Palmer K., (1997): "Environmental Regulation and Innovation", *The Review of Economics and Statistics*, 79(4), 610-19.

James, C. and P Wier, (1990): "Borrowing relationships, intermediation, and the cost of issuing public securities", in *Journal of Financial Economics*, vol. 28: 149-171.

Jappelli, T. (1990): "Who Is Credit Constrained in the U.S. Economy?", in *Quarterly Journal of Economics*, 105, 219-234.

Jewell, T (1997a): "Contaminated Land" in A. Waite and T. Jewell (eds.) *Environmental Law in Property Transactions*, Butterworths, London, 161 - 191.

Jewell, T. (1997b): "Causing or Knowingly Permitting" in A. Waite and T. Jewell (eds.) *Environmental Law in Property Transactions*, Butterworths, London, 238 - 247.

Johal, S. and A. Ulph (1998): "Tying Governments' Hands: Why Harmonisation of Environmental Policies May Be Desirable" , *Southampton Discussion Paper in Economics and Econometrics 9802.*

Johnson, A. (1997): "An Empirical Analysis of the Determinants of Corporate Debt Ownership Structure" in *Journal of Financial and Quantitative Analysis*, vol. 32, n. 1: 47-69

Johnston, J., (1984): *Econometric Methods*, New York, Mc.Graw-Hill

Kanbur, R., M. Keen and S. van Wijnbergen (1995): "Industrial Competitiveness, Environmental Regulation and Direct Foreign Investment", in I. Goldin and A. Winters (eds.) *The Economics of Sustainable Development*, Paris, OECD, 289-301.

Kennedy, P., (1994): "Equilibrium Pollution Taxes in Open Economies with Imperfect Copetition", in *Journal of Environmental Economics and Management*, 27, 49-63.

Knogden, G. (1979): "Environment and Industrial Siting" , *Zeitschrift fur Umweltpolitik*

Krugman, P. R. (1991a): "Increasing Returns to Scale and Economic Geography", in *Journal of Political Economy* 99, 483-499.

Krugman, P. R. (1991b): *Geography and Trade*, MIT Press, Cambridge, Mass.

Krugman, P. R. and A. Venables (1994): "Globalization and the Inequality of Nations", *CEPR Discussion Paper No. 1015.*

Leebron, D. (1996): "Lying Down with Procrustes: an Analysis of Harmonisation Claims", in J. Bhagwati and R. Hudec (eds.) Ch. 2.

Levinson, A. (1996a): "Environmental Regulations and Industry Location: International and Domestic Evidence", in J. Bhagwati and R. Hudec (eds.) Ch. 11.

Levinson, A. (1996b): "Environmental Regulations and Manufacturers' Location Choices: Evidence from the Census of Manufactures", *Journal of Public Economics*, **62**, 5 - 29.

Lewis, T. and D. Sappington (1997): "Using Decoupling and Deep Pockets to Resolve Judgement-Proofness Problems", *American Economic Review.*

Low, P. (ed.) (1992): *International Trade and the Environment*, World Bank, Washington.

McConnell, V. and R. Schwab (1990): "The Impact of Environmental regulation on Industry Location Decisions: The Motor Vehicle Industry", *Land Economics*, **66**, 67 - 81.

Maddala G. S., (1983): *Limited Dependent and Quantitative Variabels in Econometrics*, Cambridge University Press

McKenzie, G. and S. Wolfe (1998): "Environmental Risk Management by UK Financial Institutions" *mimeo, Department of Management, University of Southampton.*

Mani, M., S. Pargal and M. Huq (1997): "Is There an Environmental "Race to the Bottom"? Evidence on the Role of Environmental Regulation in Plant Location Decisions in India" *mimeo*, World Bank, Washington.

Markusen, J., E. Morey. and N. Olewiler, (1993): "Environmental Policy when Market Structure and Plant Location are Endogenous", *Journal of Environmental Economics and Management* 24, 69-86.

Markusen, J. (1995): "Costly Pollution Abatement, Competitiviness and Plant Location Decisions", *Paper presented to Workshop on Environmental Capital Flight*, Wageningen University.

Markusen, J., E. Morey and N. Olewiler, (1995) "Noncooperative Equilibria in Regional Environmental Policies when Plant Locations are Endogenous", *Journal of Public Economics* 56, 55-77.

Markusen, J. (1996): "Location Choice and Environmental Quality and Policy" in J. van der Bergh (Ed.) *The Handbook of Environmental and Resource Economics*, Edward Elgar, Cheltenham, UK.

Melnik, A. and S. Plaut (1986): "Loan Commitment Contracts, Terms of Lending and Credit Allocation", in *The Journal of Finance*, vol.XLI, n.2: 425-435.

Motta, M. and J-F. Thisse (1994): "Does Environmental Dumping Lead to Delocation?" , *European Economic Review*, **38**, 563-576.

Oates, W. and R. M. Schwab (1998): "Economic Competition Among Jurisdictions: Efficiency Enhancing or Distortion Inducing?" *Journal of Public Economics,* 35: 333-354.

Peek, J. and E. Rosengren, (1995): "The Capital Crunch: Neither a Borrower nor a Lender be", in *Journal of Money, Credit and Banking*, vol.27, n.3: 625-638.

Perraudin, W. R. and B. E. Sorensen, (1992): "The Credit-Constrained Consumer: an Empirical Study of Demand and Supply in the Loan Market", in *Journal of Business & Economic Statistics*, vol. 10, n.2: 179-192.

Petersen, M. A. and R. G. Rajan (1994): "The Benefits of Lending Relationships: Evidence from Small Business Data", in *The Journal of Finance*, vol. XLIX, n.1: 3-37.

Pitchford, R. (1995): "How Liable Should a Lender Be? The Case of Judgement-Proof Firms and Environmental Risk", *American Economic Review*, 85, 1171 - 1185

Rauscher, M. (1994): "On Ecological Dumping", in *Oxford Economic Papers,* 46, 822-840.

Rauscher, M. (1995): "Environmental Policy and International Capital Allocation", *University of Kiel Working Paper.*

Rauscher, M. (1997): *International Trade, Factor Movements and the Environment*, Clarendon Press, Oxford.

Sealey, C.W. Jr. (1979): "Credit Rationing in the Commercial Loan Market: estimates of a Structural Model Under Conditions of Disequilibrium", in *The Journal of Finance*, vol.XXXIV, n.3: 689-702

Segerson, K. (1993): "Liability Transfers: An Economic Analysis of Buyer and Lender Liability", *Journal of Environmental Economics and Management*, 25, 46 - 63.

Titman, S. and R. Wessels (1988): "The Determinants of Capital Structure Choice", in *The Journal of Finance*, vol. XLIII, n.1: 1-19.

Tobey, J. (1990): "The Impact of Domestic Environmental Policies on Patterns of World Trade: an Empirical Test", *Kyklos*, **43**, 191-209.

Ulph, A. (1994): "Environmental Policy, Plant Location and Government Protection" in C. Carraro (ed.) *Trade, Innovation, Environment*, Kluwer, Dordrecht, 123-163.

Ulph, A. (1997a): "International Trade and the Environment: A Survey of Recent Economic Analysis", in H. Folmer and T. Tietenberg (eds.) *The International Yearbook of Environmental and Resource Economics 1997/8*, Edward Elgar, Cheltenham, 205-242.

Ulph, A. (1997b): "Harmonisation, Minimum Standards and Optimal International Environmental Policy under Asymmetric Information", *Discussion Paper in Economics and Econometrics* No. 9701, University of Southampton.

Ulph, A. (1997c): "Environmental Policy and International Trade" in C. Carraro and D. Siniscalco (eds.) *New Directions in the Economic Theory of the Environment*, Cambridge University Press.

Ulph, A. (1998): "Political Institutions and the Design of Environmental Policy in a Federal System with Asymmetric Information", *European Economic Review* **42**, 583-592.

Ulph, A. and. L. Valentini (1997): "Plant Location and Strategic Environmental Policy with Inter-sectoral Linkages", *Resource and Energy Economics* 19, 363-383.

Ulph, A. and L. Valentini (1998): "Is Environmental Dumping Greatet When Firms Are Footloose?" *mimeo*, University of Southampton.

Ulph, D., (1995): "Globalisation and Environmental Dumping: Firm Location and Environmental Policy", *Paper presented to Workshop on Environmental Capital Flight, Wageningen University.*

van Beers, C. and J. van den Bergh (1997): "An Empirical Multi-Country Analysis of Environmental Policy on Foreign Trade", *Kyklos*, **50**, 29 - 46.

Venables, A. (1993): "Equilibrium Locations of Vertically Linked Industries, *CEPR Discussion Paper No. 802.*

Venables, A. (1994): "Economic Policy and the Manufacturing Base - Hysteresis in Location" in J. Francois and R. Baldwin (eds.) *Dynamic Issues in Commercial Policy Analysis*, Cambridge University Press.

Wilson, J. (1996): "Capital Mobility and Environmental Standards: Is There a Theoretical Basis for a Race to the Bottom?", in J. Bhagwati and R. Hudec (eds.) Ch. 10.

Xing, Y. and C. Kolstad (1994): "Environment and Trade: A Review of Theory and Issues" (*mimeo*) University of California, Santa Barbara.