Efficient Organisation of Economic Institutions: Firms and Contract Enforcement Agencies

Niko Bernd Georg Matouschek

The London School of Economics and Political Science

A thesis submitted for the PhD degree, University of London.

June 2000

Abstract

This thesis studies the efficient organisation of economic institutions. In the first chapter we analyse how foreign direct investment projects can generate spillovers through backward linkages. An investment project can generate such spillovers if local competitors in the project's own industry can benefit from the upstream efficiency improvements that were induced by the entry of the foreign firm. The existence of the spillover effect depends crucially on the supplier arrangement that is chosen by the foreign firm. The foreign firm could avoid the spillover effect by producing the input itself or by contracting with only a small number of local suppliers. We use an incomplete contract framework to study the conditions under which the foreign firm optimally chooses a supplier arrangement that generates spillovers to the local industry.

In the second chapter we study an incomplete contract model in which a buyer and a seller first agree on an efficient ownership structure and then bargain over the price of an input. We allow for asymmetric information at the ex post bargaining stage. The ownership structure that the agents agree on ex ante determines the payoff that each of them can realise before reaching agreement ex post. We show that an ownership structure that lowers the parties' joint pre-agreement payoffs accelerates ex post decision making but also makes delay in decision making more costly. We derive the ownership distribution that minimises the ex post bargaining inefficiencies.

In the third chapter we compare the efficiency of private and public provision of contract enforcement services. We show that self-interested agents with coercive power may have an incentive to use this power to enforce contracts between third parties. However, such agents also engage in extortion. We analyse how social welfare depends on the number of self-interested agents with coercive power and whether such agents face democratic elections.

Contents

Acknowledgments

I want to thank a number of people who, in many different ways, have helped me in the course of writing this thesis.

I am very grateful to my supervisor Tony Venables for his guidance, generous support, and in particular for teaching me. I was also very fortunate in having had the opportunity to learn from John Moore through his lectures, the theory seminars, and conversations. I am indebted to him for all the advice and encouragement he has given me. Further, I want to thank Kevin Roberts who has been very helpful to me ever since he taught me as an undergraduate.

On a more day to day basis I enjoyed the company of many fellow graduate students. Pawan Patil has given me a lot of much appreciated advice and also made our office a more enjoyable place to work at. I am particularly grateful to Paolo Ramezzana for the many hours we spent discussing economic and other interesting problems; graduate school would have been much less enjoyable without his company. Daniel Sturm's patience in talking economics is unsurpassed and I very much appreciate all the time he spent following, and unfortunately deflating, many of my ideas. I am also very grateful to Richard Walker from whose company I have benefited in many ways, not least through an improved, though still imperfect, understanding of the subtleties of the British way of life. Further, I want to give special thanks to Diana Khew for the help and encouragement on which I could always rely.

I am also very thankful to faculty members at the London School of Economics who took their time to discuss my research. In particular, I would like to thank Antoine Faure-Grimaud, Oliver Hart, Nobu Kiyotaki, David de Meza, François Ortalo-Magné, and Andrea Prat.

I am also grateful for the financial help I have received from the Centre for Economic Performance, the London School of Economics, and the Suntory Toyota International Centres for Economics and Related Disciplines.

Finally, my greatest debt is to my parents. I dedicate this thesis to them.

Introduction

This thesis studies the efficient organisation of economic institutions. The importance of economic institutions in market economies is evident to even the most casual observer of economic developments. Institutions, both private and public, are the outcome of the interaction between privately optimising agents, and are thus an integral part of a market economy. The analysis of the emergence, development, and functioning of institutions is therefore not only of interest in itself, but also fundamentally important for our understanding of market interactions more generally. The analysis of institutions was somewhat neglected in the early economics literature, not because of a lack of interest, but because of a lack of appropriate formal techniques. Developments in game and contract theory have since allowed economists to make vast progress in our understanding of institutions. In this thesis we aim to contribute to the existing literature on the foundations of economic institutions. In chapters 1 and 2 we analyse the efficient organisation of probably the most important private institution, namely the firm. In chapter 3 we develop a political economy model to study contract enforcement agencies.

For the remainder of this introduction we describe the broad development of the existing literature and put our analysis into its context. We start by discussing the theory of the firm and then turn to political economy.

In the standard neoclassical theory the firm is treated as a black box. In this literature the firm is simply a technology that turns inputs into outputs according to a given production function. This view of firms has proven to be a useful simplification that has allowed economists to address a large number of important economic problems. It is less useful, however, in explaining why firms exist in the first place, and in analysing the determinants of their boundaries. Given the central role that firms play in real world economies, these questions are of great importance for our understanding of the functioning of economic activity.

Coase (1937) was the first to open this black box by posing and addressing the basic questions of what defines firms, what determines their boundaries, and what are the costs and benefits of merging firms. Coase argues that the consequence of one firm buying another is that the owner-manager of the former obtains the right to give orders to the manager of the latter. When the firms are separately owned, neither manager has the right to give orders to the other. Instead, both have to agree on contracts to achieve efficient outcomes. Coase argues that, in many situations, transaction costs make contracting in the market costly. Independent firms may then decide to avoid these transaction costs by merging. While the potential reduction of transaction costs is identified as the benefit of merging firms, increased bureaucracy is seen as its primary cost. In this view it is more difficult for managers to run large rather than small firms. The optimal size of firms is then determined by the trade-off between transaction and bureaucratic costs.

Coase's analysis has been very influential. He has been credited, especially, for bringing the questions about the nature of firms to the forefront of the economic research agenda. Nevertheless, it took many years before substantial further progress was made. About 40 years after the publication of Coase's original article, Williamson (1975, 1979, 1985) and Klein, Crawford, and Alchian (1978) argued that relationship specific investments play an important role in explaining the boundaries of firms. To understand their argument, consider two firms that can only trade with each other after making initial investments. Suppose further that these investments are 'relationship specific', in the sense that their value is larger when the two firms trade with each other than when each trades with a third party. Ex post, after the investments have been sunk, the two firms then depend on each other to realise the gains from trade. This lock-in gives rise to opportunistic behaviour with each firm trying to increase its share of the gains from trade at the expense of the other. Klein, Crawford, and Alchian argue that, while the firms will always agree to trade and thereby realise the full gains from trade, ex post opportunistic behaviour can nevertheless distort the firms' ex ante investment incentives. This is the case since a firm which is in a weak bargaining position ex post can only realise a small fraction of the gains from trade and therefore only has a limited incentive to generate the surplus by investing ex ante. Klein, Crawford, and Alchian argue that by merging, firms can reduce the scope for ex post opportunistic behaviour and thus improve ex ante investment incentives. In contrast, Williamson argues that ex post haggling over the sharing of the gains from trade can itself be wasteful. In his view integration reduces opportunistic behaviour and ex post asymmetric information, thereby reducing haggling costs. While Williamson and Klein, Crawford, and Alchian elaborate on the benefits of integration, they are much less specific about its costs. Williamson follows Coase (1937) in arguing that increased bureaucracy represent the costs of integration. Klein, Crawford, and Alchian do not discuss the costs of integration in detail.

The property rights approach, which was initiated by Grossman and Hart (1986) and Hart and Moore (1990), provides a coherent theory of the firm in which both the costs and the benefits of integration are derived from the same basic principles. The central assumption in the property rights literature is that contracts are incomplete, in the sense that an initial contract between two parties cannot specify each party's obligations in all possible states of the world. Consider, for instance, two parties who contract over the use of a physical asset. It is assumed that these parties cannot costlessly anticipate all future contingencies, so that an initial contract between them is necessarily incomplete. The parties can, however, specify who has the right to decide how the asset should be used in any contingency that is not specified in the initial contract. This person is said to have the residual control rights over the asset and is identified as its owner. Hence, in the property rights literature, ownership is defined as a residual control right. Furthermore, firms are defined as a collection of productive physical assets which are jointly owned by one or more agents. When the owner-manager of one firm buys another firm she therefore simply becomes the owner of the second firm's physical assets and can decide how these assets should be used in any contingency that is not specified in a contract between the parties.

The previous paragraph outlined the property rights literature's answer to Coase's first question about the definition of a firm. We now turn to the answers that the property rights literature provides for Coase's second question about what determines the boundaries of firms. Note that, in the property rights framework, this question is equivalent to asking what determines the optimal ownership distribution of physical assets. Consider again a situation in which two parties can only trade with each other after having made relationship specific investments. Assume, furthermore, that the parties require some physical assets to engage in trade and that any ex ante contract between them is incomplete. Due to the incompleteness of the initial contract, the parties must bargain over the details of the transaction ex post. When one firm, call it firm A, buys another firm, say firm B, the owner-manager of firm A obtains the right to decide how to use firm B's physical assets in any situation that is not contractually specified. In particular, at the ex post bargaining stage, the owner-manager of the merged firm can decide how to use all the physical assets without having to agree with firm B's manager. This change in the ownership distribution of the physical assets improves the ex post bargaining position of firm A's owner-manager and worsens that of firm B's manager. When the owner-manager of firm A buys firm B she therefore obtains a larger share of the ex post surplus and has a stronger incentive to invest ex ante. This improvement in the investment incentives of firm A's owner-manager constitutes the benefit of integration. It comes, however, at the cost of reduced investment incentives for the manager of firm B who, after the merger, obtains a smaller fraction of the overall surplus. The efficient ownership distribution optimises both parties' investment incentives to maximise the overall surplus.

The property rights literature is related to the work of Williamson and Klein, Crawford, and Alchian in emphasising the importance of ex ante relationship specific investments. Importantly, however, it differs from their contributions by focusing on the role of residual control rights over physical assets and by using this concept to explain both the benefits and the costs of integration. The analysis of the property rights literature is closer to Klein, Crawford, and Alchian, who also identified potentially inefficient ex ante investments as a key determinant of the boundaries of firms, than to Williamson, who stressed the importance of ex post haggling costs in explaining the emergence of firms.

The property rights approach has been very influential and its basic framework has been extended and applied to a number of economic problems. For instance, Hart and Tirole (1990) and Bolton and Whinston (1991, 1993) use property rights models to study vertical integration and market foreclosure. Also, Rajan and Zingales (1998) extend the basic property rights framework and distinguish between residual control rights and the right to give access to an asset. They show how the right to control access can be used to influence relationship specific investments.

In chapter 1 of this thesis we develop an application of the basic property rights framework that is related to the incomplete contract literature on market foreclosure and to Rajan and Zingales (1998). In particular, we develop a property rights model to rationalise a frequently observed phenomenon of foreign direct investment projects. Multinational corporations (MNCs) which invest in less developed countries often cooperate directly with local suppliers. For instance, MNCs often transfer know-how and provide training for the work force of local suppliers. Also, some MNCs encourage upstream investments by agreeing to long term contracts or indeed by providing direct financial assistance. It is well documented (see, for instance, Matouschek and Venables (1999b)) that the upstream improvements that are induced by the entry of an MNC can spill over to local downstream firms. This can be the case if the local suppliers do not only sell to the MNC but also serve other local downstream firms. While the local industry, and local consumers, can benefit from the spillover effect, the MNC might be adversely affected by it. This will be the case if the spillover effect leads to the creation of more efficient local firms which compete with the MNC.

We argue that the creation of this kind of spillover depends on the supplier arrangement that is chosen by the MNC. The MNC could, for instance, avoid the spillover by producing the input itself. Alternatively, the MNC could contract with only a small number of suppliers and just generate enough upstream production to cover its own demand. There can only be a spillover effect if the MNC chooses a supplier arrangement that generates a *net* increase in the local production of inputs. We argue that the MNC may have an incentive to generate a net increase in upstream production, in spite of the spillover effect, since such an increase improves its bargaining position relative to the local suppliers. This improved bargaining position can then lead to a more efficient level of foreign investment. We characterise the economic environment under which the MNC finds it optimal to use a supplier arrangement that generates this kind of spillover effect.

Having applied the existing property rights literature to the specific problems of foreign direct investment projects, we proceed in chapter 2 to develop a property rights theory of the firm which focuses exclusively on ex post bargaining inefficiencies. This chapter is therefore in the spirit of Williamson who, as was noted above, stresses the importance of ex post bargaining inefficiencies in determining the boundaries of firms. The model is based on the conceptual framework of the property rights literature in that we assume that contracts are incomplete, define ownership as a residual control right, and characterise firms by the assets they own. Importantly, however, we differ from the existing literature by allowing for ex post asymmetric information. The existence of ex post private information can lead to bargaining inefficiencies which depend on the ownership distribution of physical assets.

To illustrate our basic arguments, consider a situation in which two parties, who are locked-in, can first decide on the ownership distribution of physical assets and then bargain over the gains from trade. Suppose now that there is asymmetric information ex post because, for instance, the upstream firm is better informed about the cost of producing an input or because the downstream firm knows more about the profitability of selling its product to final consumers. In the model we argue that, in the presence of private information, the parties may spend some time haggling before realising the gains from trade. The size of the resulting bargaining inefficiencies depend on the ownership distribution of physical assets. In particular, we argue that, while the two managers bargain over the input price, and before they reach an agreement with each other, each can trade with third parties on the spot market. We refer to the per period payoff that each manager can obtain during the bargaining process (and before reaching agreement with each other) as each manager's 'inside option' and argue that it depends on the ownership distribution of assets. To see how the bargaining inefficiencies depend on the ownership distribution, consider a change in the ownership distribution that lowers the managers' *joint inside option,* i.e. the sum of their individual inside options. On the one hand, for a given duration of the delay period, a reduction in the joint inside option constitutes a direct resource cost. On the other hand, however, the duration of the delay period itself depends on the ownership distribution. In particular, the managers will spend less time haggling over the input price the more dependent they are on each other, i.e. the lower their joint inside options. This *acceleration effect* constitutes an efficiency gain. The optimal ownership distribution, on which the risk neutral and financially unconstrained parties agree ex ante, minimises the bargaining inefficiencies by trading off these two effects. We believe that our model complements the existing property rights literature and can help us to understand a number of ownership patterns such as joint ventures, exclusive supplier arrangements, and the exchange of 'ugly princess hostages'.

To our knowledge the model in chapter 2 provides the first analysis of bargaining inefficiencies in a property rights framework. In a technical sense chapter 2 is related to Lockwood and de Meza (1998) and Chiu (1998) who study property rights models with strategic bargaining games. While we also apply strategic bargaining games, we differ fundamentally from their analyses by allowing for asymmetric information and by focusing on ex post rather than ex ante inefficiencies.

In chapter 3 we move away from the analysis of firms and instead turn to contract enforcement agencies. In most democracies the state uses its monopolised coercive powers to enforce private contracts. There are, however, numerous instances in which states either give up some of their monopoly powers or at least refrain from using them for contract enforcement purposes. Private organisations with coercive powers may then have an incentive to enter the market and satisfy the demand for contract enforcement. In chapter 3 we analyse the welfare implications of different contract enforcement regimes.

An understanding of contract enforcement agencies may at first seem to be of largely historical interest, applicable for instance to the development of autocracies and feudal systems or to the emergence of the Sicilian Mafia. Indeed, Gambetta (1993) argues informally that the emergence of the Sicilian Mafia was closely related to the failure of the state to provide contract enforcement services. However, recent developments in transition economies, especially in Russia, show that an understanding of private contract enforcement agencies is also of considerable current interest. A number of papers, including Greif and Kandel (1995), Hay, Shleifer, and Vishny (1996), Pistor (1996) and Hay and Shleifer (1998), have described the weaknesses of the Russian legal system in the 1990s and discussed how private agents respond to the lack of public contract enforcement. One particular response that can be observed is the emergence of private organisations with coercive powers which provide contract enforcement services. In chapter 3 we analyse this phenomenon. We start by defining 'coercive power' as the ability to influence the distribution of physical assets across agents. We then analyse the behaviour of agents who have 'some' coercive power. In the absence of democratic governments these agents have an incentive to engage in distortionary extortion. However, we show that under certain conditions the powerful agents also provide contract enforcement services. We then analyse the effect of competition between powerful agents on social welfare and study the welfare implications of controlling such agents through democratic elections. The analysis in this chapter identifies some of the trade-offs that are relevant in a discussion about the provision of contract enforcement services and that are absent from the existing literature on contract enforcement in transition economies.

Here we do not provide a comprehensive survey of the political economy literature (see Persson and Tabellini (2000) for a thorough introduction to the subject) and instead discuss those strands of the literature that are directly related to the analysis in chapter 3. Brennan and Buchanan (1977, 1978, 1980) propose a theory of the public sector in which the government aims to exploit the citizens through the maximisation of the tax revenues that it extracts from the economy. This view of the public sector as a Leviathan has received considerable attention in the literature. A key argument and empirical implication of their analysis is that, due to interjurisdictional mobility of citizens, competition between governments can limit excessive taxation. This hypothesis has been tested and rejected in a number of papers, including Oates (1985) and Forbes and Zampelli (1989). Such empirical findings, together with theoretical considerations, may lead one to be sceptical about the suitability of the Leviathan theory in describing current democratic governments. It may still, however, provide a reasonable benchmark in describing less democratic and more authotarian regimes which, across time and countries, still form the majority of governments. In our analysis we take this point of view and assume that powerful agents simply aim to maximise their life-time consumption utility. We show that, in spite of being purely self-interested, these agents have an incentive to provide socially valuable contract enforcement services.

Next to allowing for non-democratic rulers, we differ from the Leviathan literature in assuming that citizens do not have the option of leaving an economy for one with a more favourable governmental system. In our model the rulers are then 'competing' in so far as each is trying to extract as much surplus as possible from the same fixed group of private agents. Clearly, the decision whether or not to allow for interjurisdictional mobility in a formal model depends on the economic situation one aims to analyse. We believe that the assumption of limited interjurisdictional mobility is the appropriate one to make when analysing the welfare implications of the private provision of contract enforcement services. To our knowledge the only other paper studying political competition with interjurisdictional immobility is Konrad and Skaperdas (1999). In particular, they study different governmental regimes in which powerful agents use their coercive power to extort citizens but also to protect them against bandits and competing lords. In one regime "peasants are tied to their land and at the mercy of the lords who compete over how to divide them up[1]". They stress the importance of this type of competition and state: "From Mesopotamia to China, Egypt, Mesoamerica, or feudal Europe, serfs were tied to the land and free peasants typically had no outside options, with rulers coming and going but without any change in their incentives for production. Even in the past two centuries, with the rise of the rights of man, the most liberal of states have sequestered their citizens with barbed-wire borders and passport controls[2]".

As was mentioned above, we show in our analysis that self-interested private agents with coercive power have an incentive to engage in the seemingly benevolent activity of providing contract enforcement services. Similarly, Olson (1993) and McGuire and Olson (1996) argue that a revenue maximising autocrat does not only engage in extortion but also provides public goods. They argue that the autocrat has an incentive to provide public goods since it increases his tax base. While our paper is similar in spirit, the models differ in several dimensions. In particular, they discuss public goods in general, while we focus on the special public good of contract enforcement. We believe that this public good has particular characteristics which need to be model explicitly. Also, these papers do not address the reputation mechanisms which are central in our analysis. Moreover, they focus on autocracies and democracies while we allow for other government systems.

[1] Konrad and Skaperdas (1999).

[2] Konrad and Skaperdas (1999), p.17.

In our analysis we compare social welfare in an economy in which private agents can enforce contracts to two benchmarks: an anarchy, in which no agent has enough power to influence the payoff distribution, and a democracy, in which coercive power is concentrated and controlled by a democratically elected ruler. We model the democratic benchmark as a political agency problem in which citizens use elections to discipline self-interested politicians. Thus, our analysis is in the spirit of the political agency theory as initiated by Barro (1973). To our knowledge, however, the issues of contract enforcement that we focus on have so far not been addressed in the political agency literature.

Having outlined the broad strands of the literature that are relevant for the discussion in chapters 1, 2, and 3 we can now turn to the formal analysis.

Chapter 1

Foreign Direct Investment and Spillovers through Backward Linkages

The impact of foreign direct investment (FDI) projects on host economies has been a controversial issue among academics and policy makers for some time. In the 1970s the majority opinion was largely critical of the presence of multinational corporations (MNCs) in developing countries. This opinion was mainly based on the argument that the competitive advantage of an MNC can lead to the monopolisation of the local industry, thereby generating negative welfare effects in the host economy. In the last few years the prevailing view of multinational activity has become more optimistic. This shift in opinion is partly due to a large number of empirical studies that have identified various channels through which FDI can be beneficial for the host economy[1]. These empirical findings have been complemented by a number of papers[2] that provide a theoretical framework to rigorously analyse the costs and benefits of multinational activity.

The aim of this chapter is to provide microfoundations for a particular channel that has received attention in the literature, namely the generation of spillovers through backward linkages. It is well understood that the entry of an MNC can lead to quality and efficiency improvements in the local upstream industry. If there are increasing returns in the upstream industry, the efficiency improvements may simply be due to the increased demand for local inputs that is generated by the project. The MNC can also encourage upstream investments by co-operating with the local suppliers more directly, for instance, by

[1] For a comprehensive survey article see Blomström and Kokko (1996).

[2] See, for example, Rodriguez-Clare (1996), Markusen and Venables (1999), and Matouschek and Venables (1999a).

transferring know-how to local suppliers or by providing training for their work force. Also, some foreign investors encourage upstream investments by agreeing to long term contracts or, indeed, by providing direct financial assistance.

The upstream improvements that are induced by the entry of the MNC can spill over to local downstream firms, especially if local suppliers do not only sell to the MNC but also serve other local downstream firms. While the local industry, and local consumers, can benefit from the spillover effect, the MNC might be adversely affected by it since it may lead to the creation of more efficient local competitors.

In this chapter we argue that the existence of this kind of spillover depends crucially on the supplier arrangement that is chosen by the MNC. The MNC could, for instance, avoid the spillover by producing the input itself or by contracting with only a small number of suppliers which generate just enough upstream production to cover the MNC's own demand. There can only be a spillover effect if the MNC chooses a supplier arrangement that generates a *net* increase in the local production of inputs.

Given that the MNC may be adversely affected by the spillover effect, the key question is why it would choose a supplier arrangement that generates such a spillover. We argue that the MNC may itself have an interest in a net increase of local upstream production, since such an increase improves its bargaining position relative to its local suppliers. Only when the total supply of inputs is larger than the MNC's demand can it engage in what we call a 'double procurement policy'. Under such a policy the MNC establishes trading relationships with several suppliers although any single supplier would be willing, and able, to satisfy the MNC's demand. The MNC can then threaten to leave one supplier and buy more of the input from the other suppliers. Such a threat is not credible if the total supply of the input just covers the MNC's demand since the suppliers know that the MNC has no choice but to buy the input from them. The foreign firm then faces the standard hold-up problem. The observation that firms sometimes solve a hold-up problem, not by integrating with the supplier, but by engaging in a double procurement policy has been made by Matouschek and Venables (1999b). In their case study of several recent FDI projects in Eastern Europe they report cases in which the foreign investor actively encouraged the development of an independent local upstream industry. Foreign investors were certain that such a development was in their interest since it allowed them to engage in a double procurement policy which, in turn, was perceived to reduce their hold-up problem.

In this chapter we analyse different supplier arrangements and derive the conditions under which the MNC optimally chooses an arrangement that leads to a net increase in local upstream production. In this sense we provide micro-foundations for the type of spillover effect described above.

Our model is closely related to recent literature on market foreclosure in an incomplete contract setting. Hart and Tirole (1990) and Bolton and Whinston (1991, 1993) analyse how the ownership arrangement between two vertically linked firms can affect other upstream and downstream firms. In particular, they show that, as long as downstream firms compete in either input or output markets, there is "an excessive tendency towards integration"[3], since firms may use integration to engage in market foreclosure. Thus, firms use changes in the ownership distribution to influence the degree of competition in their own and related industries. While our analysis is related to this literature, it differs in a number of ways. Most importantly, we allow the downstream firm to influence the degree of competition in the upstream industry directly by determining the equilibrium number of active firms. Chemla (1996) studies how the degree of downstream competition affects the investment incentives of a monopolistic upstream supplier. In particular, he argues that an increase in downstream competition has two opposing effects on the supplier's investment incentives: on the one hand, an increase in downstream competition improves the supplier's bargaining position, on the other hand, however, it also reduces the downstream industry's profits. Our analysis differs from this in that we allow the degree of

[3] Bolton and Whinston (1991), p.214.

competition in the downstream industry to be determined endogenously and to be related to the degree of competition in the vertically linked industry.

Our analysis is also related to the literature on second sourcing (see, for instance, Farrell and Gallini (1988)) which studies a firm's private incentive to increase competition in its own industry. The key argument in this literature is that, by increasing competition in its own industry, a firm can commit to low prices in the future, which in turn improves the incentives of customers to make product-specific investments. The analysis in this chapter is also related to Rajan and Zingales (1998). They build on the basic property rights framework and show how an agent can use the power to control access to an asset to influence the (relationship specific) investments of other agents. By controlling access to an asset the agent influences the degree of (ex post) competition between 'employees'. We differ from this analysis in that we focus on the spillover effect which is absent in their set-up.

The rest of this chapter is organised as follows: in the next section we describe the basic model. In section 2 we solve the model and discuss the implications. Section 3 summarises and concludes.

1 The Model

A 'foreign' downstream firm can invest in a host country where the local downstream industry consists of a large number of potential firms. To produce locally, downstream firms must buy inputs that are produced by the local upstream industry (thus, we assume that inputs cannot be traded on the world market). This local upstream industry also consists of a large number of potential firms. The downstream industry uses the locally produced inputs to manufacture a final good that is demanded by consumers. As will become clear below, countries other than the host economy play a very passive role and merely provide an export market.

1.1 Demand

The goods that are produced by the downstream firms are perfect substitutes and their price in the local market is given by the inverse demand function

$$p(\widehat{q}+q) = \lambda P(\widehat{q}+q),$$

where $\widehat{q}$ is the total quantity sold by local firms, q is the quantity sold by the foreign firm, and λ is a measure of the size of the local market. We assume that $P_Q < 0$ and $P_{QQ} \leq 0$, $\forall Q$, where $Q \equiv q + \widehat{q}$. Throughout the analysis we use subscripts as a shorthand to indicate derivatives. We also assume that the foreign firm can sell its product in an export market at price p^e, where $p^e = P^e(q^e)$, $P^e_{q^e} < 0$, and $P^e_{q^e q^e} \leq 0, \forall q^e$. By assumption the local firms cannot sell in this market.

1.2 The Upstream Technology

Initially there is a large number of identical suppliers. Each supplier is endowed with an asset a. The foreign firm can choose which local suppliers to transfer know-how to. We assume that transferring know-how does not involve any direct costs for the foreign firm. After having received the know-how, a local supplier can make an unobservable investment $u \in \{0, \overline{u}\}$, where u indicates both the level and the cost of the investment. Suppose the supplier invests $u = \overline{u}$. Ex post he can then use the asset to produce one unit of the input at zero cost. In the absence of the supplier's human capital the asset can still be used ex post to produce the input (at zero cost), however, in this case input production is only successful with probability $\beta \in [0,1]$. With probability $1-\beta$ it is not possible to produce the input in the absence of the supplier's human capital. We interpret β as a parameter that captures the degree to which the supplier's ex ante investment is embodied in the physical asset a: the investment is entirely specific to the supplier's human capital if $\beta = 0$ and it is entirely embodied in the asset a if $\beta = 1$[4]. Inputs cannot be produced ex post if the supplier does not invest ex ante, i.e. if $u = 0$.

We follow the recent literature on incomplete contracts and property rights[5] in assuming that the upstream and downstream firms do not know ex ante what type of input is appropriate for trade ex post. Furthermore, because of the large number of potential input types it is too costly to write contingent forward contracts. Instead of writing forward contracts the firms bargain over the input price ex post. The bargaining process is described in section 1.4 below.

1.3 The Downstream Technology

To engage in production downstream firms need one unit of the input that is produced by the local upstream firms. The operational fixed cost of each local downstream firm is given by $F > 0$ and that of the foreign firm is given by $F - v$, where v indicates the level of foreign investment. Marginal costs in the downstream industry are constant and normalised to zero[6].

The foreign firm decides on the level of foreign investment v at the same time at which the suppliers choose their investment u. We assume that v is unobservable and that the foreign investment costs are given by the function $c(v) \geq 0$ which satisfies $c_v(v) \geq 0$, $c_{vv}(v) > 0$, $c_{vvv}(v) \geq 0$, $c_v(0) = 0$, and $c_v(F) = \infty$. Most of these assumptions are standard and simply ensure that the foreign firm's investment problem has a unique maximum. By assuming that $c_v(F) = \infty$ we make sure that the foreign firm's operational fixed costs are always positive. Assuming $c_v(0) = 0$ implies that the foreign firm always has strictly lower costs than the local firm. The assumption that $c_{vvv}(v) \geq 0$ may

[4] This way of capturing the asset specificity of investments is similar to Nöldeke and Schmidt (1998).

[5] See, for example, Hart (1995).

[6] This assumption implies that investments only affect fixed costs and not marginal costs. While relaxing the assumption might make the model more 'realistic' we believe that it would distract from the main points of the paper by making the analysis more cumbersome than necessary.

seem non-standard and we clarify below why we need this assumption[7].

Suppose there is at most one active local firm (in the analysis this will always be the case). Then the assumptions about the demand for the final good and the downstream production costs can be summarised in the profit functions

$$\pi(\widehat{q}, q, q^e) = p^e(q^e)q^e + p(\widehat{q} + q)q - (F - v)$$

for the foreign, and

$$\widehat{\pi}(\widehat{q}, q) = p(\widehat{q} + q)\widehat{q} - F$$

for the local firm.

We assume that the downstream firms engage in Cournot competition. Finally, we assume that it is always profitable for at least one local downstream firm to engage in production, in the sense that $\widehat{\pi}^c \geq \overline{u}$, where $\widehat{\pi}^c$ denotes the equilibrium profits of a local firm in a downstream duopoly.

1.4 The Bargaining Process

After the investments have been sunk the downstream firms need to bargain with the suppliers over the input price. In this section we specify what form the bargaining process takes when there are either one or two suppliers.

Suppose that ex post there is only one supplier who can engage in production. We then assume that the supplier Nash-bargains with the foreign firm at $t = 0$ and that both parties have equal bargaining powers. If the parties disagree and the foreign firm owns the upstream asset, then it can try to engage in upstream production at $t = 1$ after which the bargaining game ends. If the parties disagree and the supplier owns the asset, then it can engage in production and make a

[7] Note that the assumptions $c_v(v) \geq 0$, $c_{vv}(v) > 0$ and, in particular, $c_{vvv}(v) \geq 0$ are satisfied by any function $c(v) = v^\alpha$ for $\alpha > 1$.

take-it-or-leave-it offer to one of the local downstream firms at $t = 1$ after which the bargaining game ends[8].

If there are two suppliers who each own an upstream asset, bargaining takes the following form: the bargaining game has an infinite horizon and continues until the foreign firm agrees with one supplier on the price of an input. At $t = 0$ the foreign firm gets matched with one supplier and Nash-bargains with that supplier over the input price. If they disagree, the foreign firm gets matched with the other supplier at $t = 1$ and the two firms Nash-bargain over the input price. This process of alternated matching continues until the foreign firm agrees with one supplier in period $t = n \geq 0$. Once the foreign firm has agreed with a supplier, the other supplier can make a take-it-or-leave-it offer to one of the local downstream firms. During the bargaining process (i.e. prior to $t = n$) the foreign firm has to pay a real resource cost to move from one period (or, equivalently, one supplier) to the next. We assume that this resource cost is a fixed fraction $1 - \delta$, for $\delta \in [0, 1]$, of the foreign firm's net equilibrium operating profits, i.e. its equilibrium operating profits minus the input price it has to pay[9]. We interpret the parameter δ as an indicator of the degree of competition in the upstream industry: competition between the suppliers is very strong when $\delta = 1$ since, if the foreign firm disagrees with one supplier, it can costlessly move on to the other supplier. In contrast, if $\delta = 0$, it is prohibitively expensive for the foreign firm to move from one supplier to the next, so that there is essentially no competition between the suppliers.

The bargaining process which we have specified is, of course, entirely ad hoc. One could think of many other bargaining solutions that could arise in the situations which we consider. For instance, the Shapley value has been used extensively in the property rights literature (see, for example, Hart and Moore

[8] The supplier is assumed to have all the bargaining power when bargaining with the local downstream firms simply because there is a large number of such firms which all value the good equally.

[9] A bargaining game in which, instead of having to pay a real resource cost that is a fraction of its net payoff, the foreign firm discounts its payoffs by a discount rate δ gives exactly the same bargaining solution as the bargaining game presented here.

(1990)). We do not use the Shapley value for two main reasons: first, while the Shapley value would imply some competition between suppliers, it does not allow us to vary the degree of competition. Second, we conjecture that it would imply a 'collusive' outcome in which the two downstream firms and the suppliers agree to an anti-competitive arrangement. We have chosen the bargaining game specified above since it is fairly straightforward and it enables us to parameterise the degree of ex-post competition between the two suppliers.

1.5 The Game

The game is summarised in the time-line below.

$t=-2$	$t=-1$	$t=0$	$\bar{t}\geq 0$	$\bar{t}+1$
Supplier arrangement	Suppliers invest u Foreign firm invests v	Bargaining starts	Bargaining ends	Downstream competition

At $t=\bar{t}+1$ production of the final output takes place and the downstream firms compete in quantities. Between $t=0$ and $t=\bar{t}$ the suppliers and the downstream firms bargain over the price of the input. At $t=-1$ the local suppliers and the foreign firm decide on their investment levels and at $t=-2$ the foreign firm chooses the optimal supplier arrangement.

2 The Analysis

In this section we solve the game by backward induction. We first describe the equilibrium in the downstream competition subgame and then analyse the investment incentives under different supplier arrangements. In particular, we consider the following three supplier arrangements: the foreign firm contracts with only one independent supplier (indicated by 's' or single procurement), the foreign firm integrates with one local supplier ('i' or integration), and the foreign firm uses two independent suppliers ('d' or double procurement). In section 2.6

we finally derive and discuss the conditions under which the different supplier arrangements are optimal.

2.1 Downstream Competition

Given these supplier arrangements, we only need to consider the cases in which there are either one or two downstream firms. Suppose, first, that one local downstream firm competes with the foreign firm. The Nash equilibrium of the final stage subgame is then given by the simultaneous solution to the following maximisation problems:

$$\max_{q,q^e} \pi(\widehat{q}, q, q^e) \tag{1}$$

$$\max_{\widehat{q}} \widehat{\pi}(\widehat{q}, q). \tag{2}$$

Let the simultaneous solution to (1) and (2) be denoted by $\widehat{q}^*$, q^* and q^{e*} (see the appendix for details). Also, let the respective operating profits of the foreign firm and its local competitor be denoted by $\pi^c(v) = \pi(\widehat{q}^*, q^*, q^{e*})$ and $\widehat{\pi}^c = \widehat{\pi}(\widehat{q}^*, q^*)$.

Suppose, now, that only the foreign firm is active, and let q^{m*} denote the solution of (1) for $\widehat{q} = 0$. In this case the foreign firm's profits are then given by $\pi^m(v) = \pi(0, q^{m*}, q^{e*})$. Accordingly, denote by $\widehat{q}^{m*}$ the solution to (2) for $q = 0$ and the associated profits for the local downstream firm by $\widehat{\pi}^m = \widehat{\pi}(\widehat{q}^{m*}, 0)$.

2.2 First Best

Before analysing the different supplier arrangements, consider the first best investment levels that would maximise the foreign firm's profits. In the absence of a local competitor the first best investments solve

$$\max_{v,u} \frac{u}{\overline{u}} \pi^m(v) - u - c(v),$$

while, in the presence of a local competitor, they solve

$$\max_{v,u} \frac{u}{\overline{u}} \pi^c(v) - u - c(v).$$

Since $\pi^c(v) \geq \widehat{\pi}^c \geq \overline{u}$ and $c_{vv}(v) > 0$, $\forall v$, it follows that, in either case, the first best investments are given by $u = \overline{u}$ and $v = v^b$, where v^b is implicitly defined by

$$c_v(v^b) = 1 \tag{3}$$

2.3 Single Procurement

In this case the foreign firm transfers its know-how to only one independent supplier. As a result, ex post there is only one independent supplier from which the foreign firm can buy the input. Consider the bargaining game that starts at $t = 0$. The foreign firm can make an operating profit of $\pi^m(v)$ if it buys the input from the supplier. If the firms disagree, then the supplier can sell its input to one of the local downstream firms at $t = 1$ for $\widehat{\pi}^m$. The foreign firm cannot engage in production without using the input so that its outside option at $t = 0$ is zero. Let $R^s(v)$ and $W^s(v)$ denote the respective bargaining payoffs of the foreign firm and the supplier. Nash-bargaining then results in the following payoffs:

$$R^s(v) = \frac{1}{2}(\pi^m(v) - \widehat{\pi}^m)$$

$$W^s(v) = \frac{1}{2}(\pi^m(v) + \widehat{\pi}^m).$$

At $t = -1$ the foreign firm and the local supplier decide how much to invest. The equilibrium investment levels simultaneously solve the following maximisation problems:

$$\max_{v} \frac{u}{\overline{u}} R^s(v) - c(v)$$

$$\max_{u \in \{0, \overline{u}\}} \frac{u}{\overline{u}} W^s(v) - u.$$

The first order conditions are given by

$$\frac{u}{\overline{u}} R^s_v(v) - c_v(v) = 0 \tag{4}$$

$$u = \begin{cases} \overline{u} & \text{if } W^s(v) \geq \overline{u} \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

Let v^s denote the level of foreign investment that solves the foreign firm's first order condition for $u = \overline{u}$, i.e. let v^s be implicitly defined by

$$c_v(v^s) = \frac{1}{2}. \tag{6}$$

Note that $\pi^m(v) \geq \widehat{\pi}^m > \widehat{\pi}^c \geq \overline{u}$, $\forall v$. Thus, $W^s(v) \geq \overline{u}$, $\forall v$. This, in turn, implies that $u = \overline{u}$ is a best response for any v. It then follows that the unique simultaneous solution to the first order conditions (4) and (5) is given by $u = \overline{u}$ and $v = v^s$. Finally, we can confirm that the relevant second order condition for maximisation problem is satisfied since $c_{vv}(v) > 0$, $\forall v$. Hence, at $t = -1$ the firms optimally invest $u = \overline{u}$ and $v = v^s$.

Given the convexity of the cost function $c(v)$, conditions (3) and (6) imply that $v^b > v^s$. Thus, under single procurement the foreign firm underinvests relative to first best. This underinvestment is, of course, due to the standard hold-up problem: at $t = -1$ the foreign firm anticipates that it will not receive the full return from its investment and, since it bears the entire investment costs, it underinvests. Note that, in this model, the supplier does not underinvest since it can always cover its investment costs by selling the input to one of the local firms.

At $t = -2$ potential suppliers are willing to pay their entire expected surplus to learn how to produce the input. Thus, at this stage the foreign firm can extract the suppliers' entire expected surplus. At $t = -2$ the foreign firm's total expected surplus is then given by

$$T^s = \pi^m(v^s) - \overline{u} - c(v^s). \tag{7}$$

2.4 Integration

In this case, just as in the previous one, the foreign firm transfers its know-how to only one supplier. However, the foreign firm now owns the asset which the supplier needs to use to engage in production. If the supplier disagrees with the foreign firm his only option is to withdraw his human capital, thereby realising a zero payoff. In the absence of the supplier's human capital the foreign firm can use the asset to produce the input itself but it will only be successful with probability β. Its disagreement payoff is therefore given by $\beta\pi^m(v)$. Agreement between the foreign firm and the supplier results in a total payoff of $\pi^m(v)$. Thus, the respective Nash bargaining payoffs are given by

$$R^i(v) = \frac{1}{2}(1+\beta)\pi^m(v) \tag{8}$$

(9) $$W^i(v) = \frac{1}{2}(1-\beta)\pi^m(v).$$

At $t = -1$ the foreign firm and the local supplier decide how much to invest. The equilibrium investment levels simultaneously solve the following maximisation problems:

$$\max_v \frac{u}{\overline{u}} R^i(v) - c(v)$$

$$\max_{u \in \{0, \overline{u}\}} \frac{u}{\overline{u}} W^i(v) - u.$$

The first order conditions are given by

(10) $$\frac{u}{\overline{u}} R^i_v(v) - c_v(v) = 0$$

(11) $$u = \begin{cases} \overline{u} & \text{if } W^i(v) \geq \overline{u} \\ 0 & \text{otherwise.} \end{cases}$$

It immediately follows from the first order conditions that, for $W^i(0) < \overline{u}$, non-investment by both firms is an equilibrium.

There can, however, be another equilibrium in which both firms do invest. Let v^i denote the foreign investment level that solves the foreign firm's first order condition for $u = \overline{u}$, i.e. let v^i be implicitly defined by

(12) $$c_v(v^i) = \frac{1}{2}(1+\beta).$$

It then follows that $u = \overline{u}$ and $v = v^i$ solve the first order conditions if and only if $W^i(v^i) \geq \overline{u}$. The convexity of $c(v)$ again ensures that the relevant second order condition is satisfied. Hence, the investments $u = \overline{u}$ and $v = v^i$ form a Nash equilibrium if and only if $W^i(v^i) \geq \overline{u}$. Note that $W^i(v^i) < \overline{u}$ implies that $W^i(0) < \overline{u}$. Thus, $u = v = 0$ is a unique equilibrium when $W^i(v^i) < \overline{u}$. Note also that $W^i(0) < \overline{u}$ does not imply $W^i(v^i) < \overline{u}$. When $W^i(v^i) \geq \overline{u}$ there can, therefore, be two equilibria, one in which $u = v = 0$, and one in which $u = \overline{u}$ and $v = v^i$. For the remainder of this analysis we assume that, whenever both equilibria exist, the firms coordinate on the investment equilibrium in which $u = \overline{u}$ and $v = v^i$.

We now show that the condition $W^i(v^i) \geq \overline{u}$ is satisfied (i.e. the equilibrium $u = \overline{u}$ and $v = v^i$ exists) if and only if $\beta \in [\underline{\beta}, \overline{\beta}]$, where $\underline{\beta}$ and $\overline{\beta}$ are two parameters such that $0 \leq \underline{\beta} \leq \overline{\beta} < 1$. To do so we first need to show that $W^i(v^i)$ is concave in β. Differentiating $W^i(v^i(\beta))$ with respect to β gives

$$(13) \qquad W^i_\beta(v^i) = \frac{1}{2}[(1-\beta)v^i_\beta(\beta) - \pi^m(v^i)],$$

and

$$(14) \qquad W^i_{\beta\beta}(v^i) = \frac{1}{2}(1-\beta)v^i_{\beta\beta}(\beta) - v^i_\beta(\beta).$$

Applying the implicit function theorem to (12) gives

$$(15) \qquad v^i_\beta(\beta) = \frac{1}{2c_{vv}(v^i)} > 0,$$

and

$$(16) \qquad v^i_{\beta\beta}(\beta) = -\frac{c_{vvv}(v^i)}{4c_{vv}(v^i)^3} \leq 0.$$

Inspection of (13) shows that the sign of $W^i_\beta(v^i)$ is indeterminate. This is due to the two opposing effects that changes in β have on the supplier's bargaining payoff $W^i(v^i)$: on the one hand, a higher level of β increases the foreign firm's outside option, thereby reducing the gains from trade, and hence the supplier's bargaining payoff. On the other hand, however, the foreign firm's optimal investment level is increasing in β. Thus, increases in β lead to larger gains from trade and a higher bargaining payoff for the supplier. It follows from (13), (15), and (16) that $W^i_{\beta\beta}(v^i) < 0$ as long as $c_{vvv}(v) \geq 0$. This is the reason why we are making an assumption about the third derivative of the cost function $c(v)$.

Note next that $W^i_\beta(v^i) \leq 0$ for $\beta = 1$. Hence, there exist at most two values of β for which $W(v^i) = \overline{u}$. Let $\overline{\beta}$ denote the smallest value of β such that $W^i(v^i) \leq \overline{u}$, $\forall\beta \geq \overline{\beta}$, and let $\underline{\beta}$ denote the largest value of β such that $W^i(v^i) \leq \overline{u}$, $\forall\beta \leq \underline{\beta}$. The concavity of $W^i(v^i)$ in β implies that $W^i(v^i) \geq \overline{u}$ if and only if $\beta \in [\underline{\beta}, \overline{\beta}]$.

Finally, we need to show that $0 \leq \underline{\beta} \leq \overline{\beta} < 1$. Consider first $\overline{\beta}$, and note that $W(v^i) = 0$ for $\beta = 1$. Hence, $\overline{\beta} \neq 1$. Suppose now that $\max_\beta W(v^i) \geq \overline{u}$. Then $\overline{\beta}$ is uniquely defined by the β for which $W(v^i) = \overline{u}$ and $W^i_\beta(v^i) \leq 0$. Note that, in this case, $\overline{\beta}_{\overline{u}} = W^i_\beta(v^i)^{-1} \leq 0$. Clearly, $\overline{\beta} = 0$ when $\max_\beta W(v^i) < 0$. Consider now $\underline{\beta}$. Suppose that $\max_\beta W(v^i) \geq \overline{u}$, $W^i_\beta(v^i) > 0$ for $\beta = 0$, and $W^i(v^i) < \overline{u}$ for $\beta = 0$. Then $\underline{\beta}$ is uniquely defined by the β for which $W(v^i) = \overline{u}$ and $W^i_\beta(v^i) \geq 0$. Note that in this case $\underline{\beta}_{\overline{u}} = W^i_\beta(v^i)^{-1} \geq 0$. Finally, $\underline{\beta} = 0$ if either $\max_\beta W(v^i) < \overline{u}$, $W^i_\beta(v^i) < 0$ for $\beta = 0$, or $W^i(v^i) > \overline{u}$ for $\beta = 0$. The concavity of $W^i(v^i)$ then implies that $\underline{\beta} \leq \overline{\beta}$. Hence, $0 \leq \underline{\beta} \leq \overline{\beta} < 1$.

The discussion can be summarised in the following lemma:

Lemma 1 *At $t = -1$ the investment levels $u = \overline{u}$ and $v = v^i$ form a Nash equilibrium if and only if $\beta \in [\underline{\beta}, \overline{\beta}]$. For $\beta \notin [\underline{\beta}, \overline{\beta}]$ the unique equilibrium of the investment subgame is given by $u = v = 0$.*

At stage $t = -2$ the foreign firm can extract the entire surplus from the local supplier. Its total payoff is therefore given by

$$T^i = \begin{cases} \pi^m(v^i) - \overline{u} - c(v^i) & \text{if } \beta \in [\underline{\beta}, \overline{\beta}] \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

2.5 Double Procurement

Under double procurement the foreign firm transfers its know-how to two independent suppliers. Each supplier can then invest and engage in production.

Let u_i, for $i = 1, 2$, denote supplier i's investment at $t = -1$ and suppose that $u_1 = u_2 = \overline{u}$. At the bargaining stage the foreign firm then bargains with two independent suppliers in the way specified in section 1.4. To analyse this bargaining game, note first that, if one supplier has sold an input to the foreign firm, the other supplier can sell its input to a local downstream firm for $\widehat{\pi}^c$.

At $t = 0$ the foreign firm gets randomly matched with one supplier (the 'first supplier'). The two firms then Nash-bargain over the input price x. Since there are gains from trade to be realised, the firms agree on an input price x^* at $t = 0$, after which the other supplier (the 'second supplier') sells its input to a local downstream firm at $t = 1$. Thus, the agreement payoff of the foreign firm is $\pi^c(v) - x^*$, the agreement payoff of the first supplier is x^*, and their joint agreement payoff is $\pi^c(v)$. To determine x^*, we need to consider the outside options of the first supplier and the foreign firm. Suppose, therefore, that the foreign firm and the first supplier do not agree. The foreign firm then always finds it profitable to move on to the next supplier (since it only has to pay a fraction of its net payoff to do so). Except for the identity of the supplier, the foreign firm's bargaining problem at $t = 1$ is identical to its bargaining problem at $t = 0$. Thus, at $t = 1$ the foreign firm would pay x^* to the second supplier after which the first supplier would sell its input to a local downstream firm. It then follows that, if agreement is reached at $t = 0$, the gains from trade of the foreign firm and the first supplier are respectively given by $(1 - \delta)(\pi^c(v) - x^*)$ and $x^* - \widehat{\pi}^c$. Under Nash-bargaining the gains from trade of the two parties must be equalised, so that

$$(1-\delta)(\pi^c(v)-x^*) = x^* - \widehat{\pi}^c.$$

Solving for x^* gives

$$x^* = \frac{1}{2-\delta}[(1-\delta)\pi^c + \widehat{\pi}^c].$$

Thus, in equilibrium the foreign firm pays the first supplier an input price x^* at $t = 0$ and the second supplier sells its input to a local downstream firm at $t = 1$ for a price $\widehat{\pi}^c$. Note that

$$x^*_\delta = \frac{\widehat{\pi}^c - \pi^c(v)}{(2-\delta)^2} < 0, \text{ for } v > 0.$$

Hence, the stronger the degree of upstream competition, the lower the price the foreign firm has to pay for the input. We have already noted that the foreign firm's bargaining payoff is given by $\pi^c(v) - x^*$, so that

$$R^d(\delta, v) = \frac{1}{2-\delta}(\pi^c(v) - \widehat{\pi}^c).$$

Ex ante each supplier has a chance of one half to be chosen as the first supplier, in which case its bargaining payoff is given by x^*, or as the second supplier, in which case the bargaining payoff is $\widehat{\pi}^c$. Each supplier's expected bargaining payoff is, therefore, given by

$$W^d(\delta, v) = \frac{1}{2(2-\delta)}[(1-\delta)\pi^c(v) + (3-\delta)\widehat{\pi}^c].$$

When upstream competition is very strong, i.e. $\delta = 1$, the expected bargaining payoffs are given by $R^d(1, v) = \pi^c(v) - \widehat{\pi}^c$ and $W^d(1, v) = \widehat{\pi}^c$. In this case the foreign firm can extract all the gains from trade from the suppliers who only realise their 'outside option'. When there is no upstream competition, i.e. $\delta = 0$, the expected bargaining payoffs are given by $R^d(0, v) = \frac{1}{2}(\pi^c(v) - \widehat{\pi}^c)$ and $W^d(0, v) = \frac{1}{4}(\pi^c(v) + 3\widehat{\pi}^c)$.

Consider now the investment decisions at $t = -1$. At this stage the foreign firm and each supplier can decide how much to invest. Note, first, that it is always optimal for supplier i to invest $u_i = \overline{u}$, independent of v and u_j, where $j \neq i$, since $W^s(\delta, v) > \overline{u}$, $\forall v$, and $W^d(\delta, v) > \overline{u}$, $\forall v$. Thus, in any equilibrium $u_1 = u_2 = \overline{u}$. Consider now the foreign firm's best response when both suppliers invest $\overline{u}$. The best response solves

$$\max_v R^d(\delta, v) - c(v).$$

Let v^d be implicitly defined by $R^d_v(\delta, v^d) - c_v(v) = 0$, i.e.

$$\frac{1}{2 - \delta} = c_v(v^d). \tag{18}$$

The second order condition is satisfied since $c_{vv}(v) > 0$. Thus, the foreign firm's best response is given by $v = v^d$. It then follows that the unique equilibrium of the investment subgame is given by $u_1 = u_2 = \overline{u}$ and $v = v^d$.

At stage $t = -2$ the foreign firm can again extract the entire surplus from the local supplier. Its total expected surplus is, therefore, given by

$$T^d = \pi^c(v^d) + \widehat{\pi}^c - 2\overline{u} - c(v^d). \tag{19}$$

2.6 The Optimal Supplier Arrangement

We are interested in the conditions under which the entry of the foreign firm generates a spillover effect that can lead to the emergence of a local competitor.

The foreign investment will generate such a spillover effect only if its entry leads to a net increase in the local supply of inputs. Since the foreign firm uses one unit of the input itself, we therefore need to study the conditions under which the local industry is induced to produce at least two units of the critical input. The only supplier arrangement that might achieve this outcome is the double procurement policy in which the foreign firm enables two local suppliers to engage in upstream production. Single procurement does not generate the spillover effect because of the assumed capacity constraint[10]. In this section we use the above analysis to derive the conditions under which the foreign firm optimally engages in a double procurement policy.

2.6.1 Single Procurement versus Integration

It is useful to start the discussion by comparing single procurement and integration. We have already seen that a single procurement policy leads to underinvestment. It is, of course, well known that integration can reduce these kind of investment inefficiencies. In this section we study the conditions under which this is the case in the current framework.

The foreign firm prefers single procurement to integration if and only if $T^s \geq T^i$. Recall that, under single procurement, the supplier invests $u = \overline{u}$, and the foreign firm invests $v^s < v^b$. Thus, integration can only be more profitable than single procurement if it increases the level of foreign investment, without reducing the supplier's investment level.

The extent to which integration improves the foreign firm's investment incentives depends on β, the degree of asset specificity of the supplier's investment.

[10]We conjecture that we could relax the assumption of an exogenously given capacity constraint without affecting the results if we replace it with one of the following assumptions: firstly, we could allow the foreign firm to sign an exclusive dealing agreement with the single independent supplier. Even the supplier has an interest in such an agreement since it reduces downstream competition and, therefore, increases the gains from trade at the expense of the consumers. Alternatively, we could extend the game to allow suppliers to choose the capacity of their plant ex ante. We conjecture that, if the suppliers cannot commit to exclusive dealing agreements (e.g. because of anti-trust legislation), they have an incentive to choose the capacity constraint that we assume exogenously, i.e. they optimally choose to build a plant that just covers the demand of the foreign firm.

We have already noted that, for a given upstream investment, the foreign firm's bargaining payoff R^s is increasing in β. This is the case since an increase in β improves the foreign firm's ability to produce the input in the absence of the supplier's human capital which, in turn, increases the foreign firm's outside option and, thus, its bargaining payoff.

The effect of integration on the supplier's investment incentives, however, is ambiguous. This is due to the two opposing effects that integration has on the supplier's bargaining payoff W^s: on the one hand, integration improves the foreign firm's outside option and hence reduces the gains from trade between the supplier and the foreign firm. On the other hand, however, the foreign firm's improved investment incentives lead to an increase in the gains from trade. When β is very large, i.e. $\beta \in (\overline{\beta}, 1]$, the former effect dominates the latter, and the supplier cannot be induced to invest under integration. Note that, in this case, the foreign firm will also not invest since it cannot engage in production without access to locally produced inputs. Hence, for $\beta \in (\overline{\beta}, 1]$, integration leads to lower investment levels and is therefore less profitable than single procurement.

Suppose now that the supplier's investment is very human capital specific, i.e. $\beta \in [0, \underline{\beta})$. In this case the asset is not very useful for the foreign firm in the absence of the supplier's human capital. As a result, for a given upstream investment, integration only leads to a small improvement in the foreign firm's bargaining position and its investment incentives (see (12)). The less the foreign firm invests, however, the smaller are the supplier's investment incentives. When β is small enough, i.e. $\beta \in [0, \underline{\beta})$, the supplier cannot be induced to invest. This again implies $v = 0$, since the foreign firm cannot produce without access to local inputs.

Finally, suppose $\beta \in [\underline{\beta}, \overline{\beta}]$. In this parameter range the supplier invests $u = \overline{u}$ and the foreign firm invests $v = v^i$. It follows from (7) and (17) that the relative profitability of the two supplier arrangements is then given by

$$\Delta T^{i,s} = T^i - T^s = v^i - v^s - c(v^i) + c(v^s).$$

Note that $\Delta T^{i,s} = 0$ for $\beta = 0$ and

$$\frac{\partial \Delta T^{i,s}}{\partial \beta} = \frac{1}{2}(1-\beta)v^i_\beta \geq 0, \ \forall \beta.$$

Thus, integration is more profitable than single procurement if the supplier's investment is neither too human capital, nor too asset specific, i.e. $\beta \in [\underline{\beta}, \overline{\beta}]$. The analysis in this section can be summarised in the following lemma:

Lemma 2 *Integration is weakly more profitable than single procurement, i.e. $T^i \geq T^s$, if $\beta \in [\underline{\beta}, \overline{\beta}]$. Otherwise single procurement is more profitable than integration, i.e. $T^s > T^i$.*

2.6.2 Double Procurement versus Single Procurement

The foreign firm prefers double to single procurement if and only if $T^d \geq T^s$. It follows from (7) and (19) that this is the case if and only if

$$\pi^c(v^d) + \widehat{\pi}^c - 2\overline{u} - c(v^d) \geq \pi^m(v^s) - \overline{u} - c(v^s). \tag{20}$$

Condition (20) shows that the relative profitability of single and double procurement depends on three effects: first, the procurement policies differ with respect to the number of suppliers who undertake investments. Clearly, a double procurement policy requires twice as many suppliers to invest than a single procurement policy. This duplication of investment efforts constitutes a resource cost which makes double procurement relatively less profitable.

Second, the relative profitability of the two procurement policies depends on the size of the spillover effect: under double procurement suppliers produce two units of the input, one of which is then sold to the foreign firm and the other to a local downstream firm[11]. This creates competition in the downstream industry

[11] It was already mentioned earlier that, in spite of only requiring one unit of the input, the

and reduces industry profits. To analyse the size of this spillover effect consider condition (20) for $v^d = v^s = v$. Note that, in this stylized model, the foreign firm is able to extract the entire operating profit from its local downstream competitor so that the size of the spillover effect is given by $\pi^m(v) - \pi^c(v) - \widehat{\pi}^c$. There are two reasons why this spillover effect is always positive: first, in the case of perfect substitutes, the revenue of a monopolist must be larger than the joint revenue of two duopolists. Second, when there are fixed costs, production by a monopolist is more cost efficient than production by two duopolists. The size of the spillover effect depends on the extent to which the two downstream firms compete in the same market. In terms of the model this can be shown by noting that

$$\frac{\partial(\pi^m(v) - \pi^c(v) - \widehat{\pi}^c)}{\partial\lambda} = P(q^{m*})q^{m*} - P(q^{c*} + \widehat{q}^{c*})(q^{c*} + \widehat{q}^{c*}) \geq 0,$$

i.e. the larger the local market (in which the firms compete) relative to the export market (in which only the foreign firm is active), the larger the spillover effect.

Finally, the relative profitability of the two supplier arrangements depends on the difference in foreign investments. We have already seen that, under double procurement, the foreign firm's investment incentives are increasing in δ, the degree of competition in the upstream industry (see (18)). Essentially, more competition in the upstream industry improves the foreign firm's ex post bargaining position which, in turn, improves its investment incentives. It can be seen from (6) and (18) that $v^d = v^s$ for $\delta = 0$. Since $v^d_\delta > 0$ and $v^s_\delta = 0$, it follows that $v^d > v^s$ if $\delta > 0$. The first order condition (18) also shows that, even under double procurement, the foreign firm's investment is always weakly

foreign firm and the suppliers would like to agree to an exclusive dealing contract, whereby the suppliers sell only to the foreign firm and 'boycott' the local competitor. Indeed such a contract would increase the joint profit of all firms, including the local downstream firm. We rule out such a contract on anti-trust grounds since it only increases joint profits at the expense of the consumers by monopolising the downstream industry.

below the first best level. Hence, double procurement reduces the investment inefficiencies relative to single procurement.

Moving from single to double procurement therefore affects the foreign firm's profits in three ways: the duplication of upstream investment efforts and the potential creation of downstream competition have a negative and the improved investment incentives have a positive effect on the foreign firm's profits. Rearranging (20) and substituting for the profit functions gives

$$(21)\quad v^d - v^s + c(v^s) - c(v^d) \geq \overline{u} + p(q^{m*})q^{m*} - p(q^{c*} + \widehat{q}^{c*})(q^{c*} + \widehat{q}^{c*}) + F.$$

The LHS of (21) represents the gains and the RHS the costs of moving from single to double procurement. Let $\Delta B^{d,s}(\delta) = v^d - v^s + c(v^s) - c(v^d)$ and note that $\Delta B^{d,s}(0) = 0$ and

$$\frac{\partial \Delta B^{d,s}(\delta)}{\partial \delta} = \frac{1-\delta}{2-\delta} v_\delta^d \geq 0.$$

Also let $\Delta C^{d,s} = \overline{u} + p(q^{m*})q^{m*} - p(q^{c*} + \widehat{q}^{c*})(q^{c*} + \widehat{q}^{c*}) + F$ and note that it is independent of δ. Thus, there exists a critical $\widetilde{\delta}^s$ above which double procurement is more profitable than single procurement, and below which the opposite is true. If $\Delta B^{d,s}(1) \leq \Delta C^{d,s}$, then $\widetilde{\delta}^s = 1$ and if $\Delta B^{d,s}(1) > \Delta C^{d,s}$, then $\widetilde{\delta}^s$ is implicitly defined by the unique $\delta \in [0,1)$ that solves $\Delta B^{d,s}(\delta) = \Delta C^{d,s}$. Note that

$$\frac{\partial(\Delta B^{d,s} - \Delta C^{d,s})}{\partial \overline{u}} = -1$$

and

$$\frac{\partial(\Delta B^{d,s} - \Delta C^{d,s})}{\partial \lambda} = -[P(q^{m*})q^{m*} - P(q^{c*} + \widehat{q}^{c*})(q^{c*} + \widehat{q}^{c*})] < 0.$$

Thus, the parameter region in which double procurement is more profitable than single procurement is decreasing in the cost of upstream investments and in the relative size of the local market. The analysis can be summarised in the following lemma:

Lemma 3 *Double procurement is more profitable than single procurement, i.e.* $T^d > T^s$*, if* $\delta \in (\widetilde{\delta}^s, 1]$*. Otherwise single procurement is weakly more profitable than double procurement, i.e.* $T^s \geq T^d$*.*

2.6.3 Double Procurement versus Integration

Finally, we need to consider the relative profitability of double procurement and integration. It immediately follows from (7) and (17) that double procurement is more profitable than integration if the supplier's investment is either very specific to the supplier's human capital ($\beta \in [0, \underline{\beta})$) or to the asset ($\beta \in (\overline{\beta}, 1]$). If $\beta \in [\underline{\beta}, \overline{\beta}]$, double procurement is more profitable than integration if and only if

$$\pi^c(v^d) + \widehat{\pi}^c - 2\overline{u} - c(v^d) \geq \pi^m(v^i) - \overline{u} - c(v^i). \tag{22}$$

Rearranging (22) and substituting for the profit functions gives

$$v^d - v^i + c(v^i) - c(v^d) \geq \overline{u} + p(q^{m*})q^{m*} - p(q^{c*} + \widehat{q}^{c*})(q^{c*} + \widehat{q}^{c*}) + F. \tag{23}$$

The LHS of (23) represents the (potential) gains and the RHS the costs of moving from integration to a double procurement policy. Just as in the previous section the costs are due the duplication of upstream investments and the spillover effect. Double procurement can only be more profitable than integration if it leads to more efficient investment levels. Whether the level of foreign investment is more efficient under double procurement than under integration depends on the degree of upstream competition and the asset specificity of the suppliers' investments. Let $\Delta B^{d,i}(\delta,\beta) = v^d - v^i + c(v^i) - c(v^d)$ and $\Delta C^{d,i} = \overline{u} + p(q^{m*})q^{m*} - p(q^{c*} + \widehat{q}^{c*})(q^{c*} + \widehat{q}^{c*}) + F$. It then follows from (12) and (18) that $\Delta B^{d,i}(\frac{2\beta}{1+\beta},\beta) = 0$,

$$\frac{\partial \Delta B^{d,i}(\delta,\beta)}{\partial \delta} = \frac{1-\delta}{2-\delta}v_\delta^d \geq 0,$$

and

$$\frac{\partial \Delta B^{d,i}(\delta,\beta)}{\partial \beta} = -\frac{1}{2}(1-\beta)v_\beta^i \leq 0.$$

Thus, if $\beta \in [\underline{\beta},\overline{\beta}]$, the foreign firm's investment levels are more efficient under double procurement than under integration if and only if $\delta \geq \frac{2\beta}{1+\beta}$, i.e. upstream competition is strong relative to the degree of asset specificity of the supplier's investment. Since $\Delta C^{d,i}$ is independent of β and δ it then follows that $\delta \geq \frac{2\beta}{1+\beta}$ is a necessary condition for double procurement to be more profitable than integration.

Let $\widetilde{\delta}^i$ be the critical degree of upstream competition above which double procurement is more profitable than integration. Thus, $\widetilde{\delta}^i = 1$ if $\Delta B^{d,i}(1,\beta) \leq \Delta C^{d,i}$. Otherwise, $\widetilde{\delta}^i$ is implicitly defined by the unique $\delta \in [0,1)$ that solves $\Delta B^{d,i}(\delta,\beta) = \Delta C^{d,i}$. Note that $\frac{\partial(\Delta B^{d,i}-\Delta C^{d,i})}{\partial \overline{u}} < 0$, $\frac{\partial(\Delta B^{d,i}-\Delta C^{d,i})}{\partial \lambda} < 0$, and $\frac{\partial(\Delta B^{d,i}-\Delta C^{d,i})}{\partial \beta} \leq 0$. Thus, the parameter region in which double procurement is more profitable than integration is decreasing in the cost of upstream investments, in the relative size of the local market, and in the degree of asset specificity of the supplier's investment. The analysis can be summarised in the following lemma:

Lemma 4 *Double procurement is more profitable than integration, i.e.* $T^d > T^i$, *if* $\delta \in (\widetilde{\delta}^i, 1]$. *Otherwise integration is weakly more profitable than double procurement, i.e.* $T^i \geq T^d$.

2.6.4 The Optimality of Double Procurement

The following proposition follows immediately from the discussion above:

Proposition 1 *The foreign firm engages in double procurement if and only if either $\beta \in [\underline{\beta}, \overline{\beta}]$ and $\delta \in (\widetilde{\delta}^{i}, 1]$ or $\beta \notin [\underline{\beta}, \overline{\beta}]$ and $\delta \in (\widetilde{\delta}^{s}, 1]$.*

We have seen that the foreign firm is more likely to engage in double procurement, the lower the upstream investment costs (the lower $\overline{u}$) and the smaller the spillover effect (the smaller λ). Also, if $\beta \in [\underline{\beta}, \overline{\beta}]$, the foreign firm is more likely to engage in double procurement, the more specific the upstream investment is to the supplier's human capital (the smaller β).

Note also that there may be a difference between the social and the foreign firm's private return from engaging in double procurement. This is the case since the foreign firm does not take into account the gain in consumer surplus that is generated when there are more firms in the downstream industry. It follows that, whenever the foreign firm engages in double procurement, this arrangement leads to a higher level of social surplus than the alternative arrangements that we allow for. However, the reverse is not true: it can be the case that double procurement is socially desirable but the foreign firm chooses a different supplier arrangement. Again, this is possible because the foreign firm does not internalise changes in consumer surplus.

3 Conclusion

The aim of this chapter was to rationalise the observation that foreign investors sometimes use supplier arrangements that generate spillovers to local competitors. We argued that it may be problematic for the foreign firm to contract with only one independent supplier since, in this case, it anticipates to be held up ex post and therefore underinvests. Integration with the supplier does not improve the foreign firm's investment incentives if the local supplier's investment is either very human capital or very asset specific. In the former case, ownership of the asset does not improve the foreign firm's bargaining position since it still has to agree with the supplier to get access to his human capital. In the latter case, integration only improves the foreign firm's bargaining position at a very high cost, namely a strong deterioration in the supplier's own investment incentives.

The effectiveness of a double procurement policy in improving investment incentives depends on the degree of competition in the upstream industry: the more competitive the upstream industry the stronger the foreign firm's ex post bargaining position in the case of double procurement and the higher its investment level. The foreign firm only engages in double procurement if the benefit of improved investment incentives outweighs the costs of such a policy. These costs are given by the duplication of upstream investment efforts and the creation of downstream competition.

4 Appendix

The first order conditions of (1) and (2) are given by

$$p^e_{q^e}(q^e)q^e + p^e(q^e) + p_Q(\widehat{q}+q)q + p(\widehat{q}+q) = 0$$

$$p_Q(\widehat{q}+q)\widehat{q} + p(\widehat{q}+q) = 0.$$

These first order conditions solve for q^*, $\widehat{q}^*$ and q^{e*}.

Note that the second order conditions

$$\begin{aligned} 2p_Q(\widehat{q}+q) + p_{QQ}(\widehat{q}+q)q + 2p^e_{q^e}(q^e) + p^e_{q^eq^e}(q^e)q^e &\leq 0 \\ 2p_Q(\widehat{q}+q) + qp_{QQ}(\widehat{q}+q) &\leq 0, \end{aligned}$$

are satisfied, since $p_Q(Q) < 0$, $p_{QQ}(Q) < 0$, $p^e_{q^e}(q^e) < 0$, and $p^e_{q^eq^e}(q^e)$.

Also, the stability condition is satisfied, since

$$\begin{vmatrix} 2p_Q(Q) + qp_{QQ}(Q) & p_Q(Q) + qp_{QQ}(Q) \\ p_Q(Q) + \widehat{q}p_{QQ}(Q) & 2p_Q(Q) + \widehat{q}p_{QQ}(Q) \end{vmatrix} = p_Q(Q)[3p_Q(Q) + p_{QQ}(Q)Q] > 0.$$

Chapter 2

A Property Rights Theory of the Firm with Private Information

In this chapter we develop a property rights theory of the firm which focuses exclusively on the role of asset ownership in determining bargaining inefficiencies. We argue that firms at times bargain with their patrons[1] in the presence of private information. Bargaining between the parties may then be inefficient, in the sense that gains from trade are not realised or are only realised after some delay. The firm and its patrons anticipate these bargaining inefficiencies and take them into account when deciding on the ownership distribution of the physical assets. We explicitly model the interdependence between the ownership distribution and the ex post bargaining inefficiencies. This allows us to relate the economic environment to the ownership distribution that minimises the expected bargaining inefficiencies.

To illustrate the structure of the model, consider a situation in which there are two managers and two assets. One manager operates the upstream asset, for instance a plant that produces inputs, and the other manager operates the downstream asset, for instance a plant that produces a final output. Assume further that the two managers are locked-in, in the sense that trade between them is more profitable than trade between either of them and third parties. In the presence of transport costs, the lock-in may, for instance, be due to the close proximity of the two assets. The managers first contract over the ownership distribution of the assets and then bargain over the price of the input. At the ex post bargaining stage only the manager of the downstream firm knows how

[1] We adopt Hansmann's definition of a firm's patrons as "all persons who transact with a firm either as purchasers of the firm's products or as sellers to the firm of supplies, labor or other factors of production" (Hansmann (1996), p.12).

profitable it is for her to operate in the downstream industry and thus how much she values the input. Due to the presence of ex post private information the managers may spend some time haggling with each other before agreeing on the input price. While the two managers bargain over the input price, and before they reach agreement, each can trade with third parties on the spot market. We refer to the per period payoff that each manager can obtain during the bargaining process (and before reaching agreement with each other) as each manager's 'inside option' or 'pre-agreement payoff'. Each manager's pre-agreement payoff depends on the ownership distribution on which the parties agree ex ante. To see how the bargaining inefficiencies depend on the ownership distribution, consider a change in the ownership distribution that lowers the managers' *joint pre-agreement payoff*, i.e. the sum of their individual pre-agreement payoffs. On the one hand, for a given duration of the delay period, a reduction in the joint pre-agreement payoff constitutes a direct resource cost. On the other hand, however, the duration of the delay period itself depends on the ownership distribution. In particular, the managers will spend less time haggling over the input price the more dependent they are on each other, i.e. the lower their joint pre-agreement payoff. This *acceleration effect* constitutes an efficiency gain. Before the managers bargain over the input price, and before the downstream manager learns her valuation, they contract over the ownership distribution of the physical assets. We assume that both managers are risk neutral and not liquidity constraint. Thus, they agree on the efficient ownership distribution that minimises the total expected bargaining costs. In the analysis we show that the acceleration effect dominates the direct resource cost when uncertainty is small relative to the gains from trade. In this case the efficient ownership distribution *minimises* the managers' joint pre-agreement payoff. When uncertainty is large relative to the gains from trade, the direct resource cost dominates the acceleration effect and the efficient ownership distribution *maximises* the managers' joint pre-agreement payoff.

As a further illustration of the basic effects, it is useful to consider the following example. Recently, the US car manufacturer Ford was involved in negotiations about the renewal of a supply contract with its sole supplier of car locks[2]. The firms disagreed about the terms of the contract: while Ford wanted a continuation of the previous conditions, the supplier suspected that Ford's profitability had improved, and therefore insisted on more favorable terms. During the negotiations the supplier suddenly stopped delivering locks to Ford which forced the latter to suspend production. This interruption was very costly for both sides: Ford incurred a production loss of 10,000 cars, costing it turnover of DM200 million, and the supplier's stock price fell sharply. After a few days the supplier started to deliver again, and, soon after that, the companies agreed on a new contract. This contract guaranteed the lock producer a continued position as Ford's exclusive supplier of locks. At the time commentators questioned why, after having experienced such a costly dispute, the two firms did not agree on an ownership distribution that made them less dependent on each other. For instance, it was argued that, if Ford owned the supplier, it could use the upstream assets during a dispute to ensure that at least some cars can be produced. In this chapter we argue that ownership structures which maximise the parties' pre-agreement payoffs do indeed reduce the cost of disagreement for a *given disagreement period*. However, we also argue that the duration of the disagreement period itself is negatively related to the cost of disagreement. In terms of our example, this suggest that, while integration might make temporary disagreement less costly for a given disagreement period, it might also increase the duration of the disagreement period itself. It may, therefore, be optimal for Ford and its supplier to make themselves very dependent on each other, since this ensure that disputes are settled quickly.

Our analysis is related to the recently developed property rights theory of the firm. This literature was initiated by the seminal papers of Grossman and

[2]The information about this case is taken from the following articles: "Key Position," *Frankfurter Allgemeine Zeitung*, 18.6.1998, p.17; "New Contract Between Ford and Kiekert," *Frankfurter Allgemeine Zeitung*, 14.10.1998, p.8.

Hart (1986) and Hart and Moore (1990)[3], and developed further in a number of papers, including Chiu (1998), de Meza and Lockwood (1998) and Rajan and Zingales (1998). The property rights literature argues that, in the absence of comprehensive contracts, ownership of physical assets is a residual control right. That is to say that the owner of an asset has the right to take all those decisions concerning the use of the asset that have not been specified in an initial contract. The ownership of physical assets is then taken as the defining characteristic of firms. We adopt this basic conceptual framework in our analysis. We differ from the existing property rights literature, however, by focusing on ex post rather than ex ante inefficiencies. In the existing property rights literature bargaining is always assumed to be efficient. Thus, once relationship specific investments have been undertaken, the ownership distribution does not influence the size but only the sharing of the overall surplus. Ownership then affects efficiency by determining the agents' private incentives to undertake relationship specific investments prior to the bargaining stage. In contrast, in our approach ownership has efficiency implications by determining the bargaining inefficiencies that can arise in the presence of private information.

While recent work on the theory of the firm has emphasised the role of ex ante investment incentives[4], earlier, and largely informal, contributions stressed the importance of bargaining inefficiencies in determining the boundaries of firms. Coase (1937), for instance, discusses the role of imperfect information and negotiation costs in the emergence of firms and states that: "The most obvious cost of "organizing" production through the price mechanism is that of discovering what the relevant prices are. [...] The costs of negotiation and concluding a separate contract for each exchange transaction which takes place on a market must also be taken into account."[5] Also, Williamson (1975, 1985) identifies imperfect information as one reason why market transactions can be inefficient.

[3]For an overview of the literature see Hart (1995).

[4]Next to the above mentioned property rights literature see also Klein, Crawford, and Alchian (1978).

[5]Coase (1937), p.336.

In his view vertical integration allows firms to avoid costly haggling but also increases bureaucracy. Our analysis is related to this early literature in that we focus on the interdependence of ownership distributions and ex post bargaining inefficiencies.

A related literature (see, for instance, Arrow (1975) and Riordan (1990)) studies integration in the presence of private information and argues that integration reduces the degree of asymmetric information. We differ from this literature in that we assume that changes in the ownership distribution only affect the payoff agents realise during the bargaining process and have no influence on the degree of the informational asymmetry. We take this approach, not because we think that integration can never reduce informational asymmetries, but because we believe that the impact of ownership changes on agents' incentives to transmit information is ambiguous and needs to be model explicitly (see, for instance, Aghion and Tirole (1997) and Dessein (1999)). Our analysis shows that a model in which the extent of asymmetric information is not affected by changes in the ownership structure has implications that are consistent with some observed ownership patterns.

We proceed as follows: in the next section we describe a simple model with a static bargaining game. In this model the uninformed party simply makes a take-it-or-leave-it offer to the informed party. Because of the static nature of the bargaining game we have to adapt our arguments somewhat. In particular, in the static model, the ownership distribution is assumed to determine the disagreement, rather than the pre-agreement payoffs. Also, a change in the ownership distribution that lowers the disagreement payoffs increases the probability of agreement rather than accelerating agreement. In spite of these differences we can use the static model to illustrate our main arguments. We do so in section 2 by solving the static model. While the static model is straightforward to solve, it has two conceptual problems. First, it does not allow the parties to continue bargaining until the gains from trade are realised and simply assumes that the bargaining process ends after one offer. Second, by assuming that the ownership distribution that is agreed on ex ante determines the disagreement payoffs, it restricts the analysis to irreversible ownership changes. While some ownership changes are indeed irreversible, a model that allows for more general ownership changes is clearly more satisfactory. We present such a model, with a dynamic bargaining game, in section 3 and solve it in section 4. We argue that the static and the dynamic models have very similar implications for the efficient ownership distributions. In section 5 we discuss these implications and show to what extent they are consistent with observed ownership patterns. Section 6 concludes.

1 The Static Model

There are two risk neutral players, a buyer B and a seller S. Neither the buyer nor the seller is liquidity constrained. The seller can produce an input that can be used by the buyer to produce the final output. To engage in production the buyer and the seller need access to some physical assets. Let A denote the ownership distribution of these physical assets across the agents.

There are two periods, $t = -1$ and $t = 0$. At $t = -1$ (ex ante) the parties contract over the ownership distribution A. We make the important assumption that the particular type of input required by the buyer ex post is uncertain ex ante and cannot be specified in the initial contract[6].

At $t = 0$ (ex post) uncertainty about the required input is realised and production can take place. At this stage the buyer and the seller bargain over the input price. Once the parties agree on a transfer price the seller produces the relevant input and the buyer uses the input to produce the final output. We normalise the seller's production costs to zero. By producing the final product the buyer generates a payoff of π. Ex ante the value of π is uncertain and both parties only know that it is uniformly distributed on $[\pi_l, \pi_h]$, where $\pi_h = \mu + \alpha$ and $\pi_l = \mu - \alpha \geq 0$. At the beginning of the trading period ($t = 0$) the level of

[6] For microfoundations of incomplete contracts see, for instance, Hart and Moore (1999), Segal (1999), and, for the case of one-sided asymmetric information, Reiche (1999).

π is realised. We assume that only the buyer learns the realization of π.

If trade between the buyer and the seller does not take place, the players realise their respective disagreement payoffs $b(A)$ and $s(A)$. The joint disagreement payoffs are denoted by $j(A) \equiv b(A)+s(A)$. We assume that trade between the buyer and the seller is always profitable, i.e.

$$\pi_l \geq j(A), \quad \forall A. \tag{1}$$

Note that only the disagreement payoffs are functions of the ownership distribution. When the buyer and the seller do not transact with each other, and instead trade with third parties on the spot market, the return that each one realises depends on the assets he or she owns. The assumption that disagreement payoffs depend on the ownership distribution that was agreed on ex ante implies that changes in the ownership distribution are irreversible.

Note also that no other variable or parameter depends on the ownership distribution. In particular, the degree of asymmetric information α does not depend on A. This captures the idea that the degree of asymmetric information between the players is independent of the ownership distribution. Also, the buyer's valuation of the input π does not depend on the ownership distribution since, when trade between the parties takes place, both players have access to all the assets in the relationship, independent of the ownership distribution.

In this simple model bargaining takes the form of a take-it-or-leave-it offer by the (uninformed) seller. Hence, after the seller makes an offer p, the buyer decides whether to accept or to reject it. If the buyer accepts the offer, she realises a payoff of π and the seller realises a payoff of p. If, instead, the buyer rejects the offer, both parties simply realise their disagreement payoffs.

2 The Analysis of the Static Model

In this section we analyse the model that was described above. We first solve the bargaining game that takes place at $t = 0$ and then study the optimal ownership distribution on which the parties agree at $t = -1$.

2.1 The Bargaining Game

At $t = 0$ the uninformed seller makes a take-it-or-leave-it offer p. The buyer accepts p if and only if $\pi - p \geq b(A)$. Hence, $p_c(\pi_c) = \pi_c - b(A)$ denotes the price that is accepted by all buyers of type $\pi \in [\pi_c, \pi_h]$.

The seller's expected return from making an offer $p_c(\pi_c)$ is then given by

$$R(\pi_c) = p_c(\pi_c)\frac{\pi_h - \pi_c}{\pi_h - \pi_l} + s(A)\frac{\pi_c - \pi_l}{\pi_h - \pi_l}, \tag{2}$$

and his optimal offer solves

$$\max_{\pi_c \in [\pi_l, \pi_h]} R(\pi_c). \tag{3}$$

A cutoff point π_c^* is a maximum of (3) if and only if it satisfies the following first and second order conditions:

$$R'(\pi_c^*) = \begin{cases} 0 & \text{for } \pi_c^* \in (\pi_l, \pi_h) \\ \leq 0 & \text{for } \pi_c^* = \pi_l \\ \geq 0 & \text{for } \pi_c^* = \pi_h, \end{cases}$$

and

$$R''(\pi_c^*) < 0 \text{ for } \pi_c^* \in (\pi_l, \pi_h).$$

Differentiating (2) gives

$$R'(\pi_c) = \frac{1}{(\pi_h - \pi_l)}(\pi_h + j(A) - 2\pi_c)$$

and

$$R''(\pi_c) = -\frac{2}{(\pi_h - \pi_l)}.$$

Note, first, that π_c^* is unique since $R''(\pi_c) < 0, \forall \pi_c$. Note, second, that $R'(\pi_h) < 0$, so that π_h is not a maximum. Finally, note that $R'(\frac{1}{2}(\pi_h + j(A))) = 0$ and that $R'(\pi_l) < 0$ if and only if $\pi_l > \frac{1}{2}(\pi_h + j(A))$. Thus, the optimal cutoff type π_c^* that solves (3) is uniquely given by

$$\pi_c^* = \max(\frac{1}{2}(\pi_h + j(A)), \pi_l). \tag{4}$$

Since $\pi_c^* < \pi_h$, there is always a positive probability that the buyer accepts the seller's offer. Note that π_c^* is weakly increasing in $j(A)$. Thus, the higher the joint disagreement payoffs $j(A)$, the less likely it is that the buyer accepts the seller's offer. Finally, note that any type of buyer $\pi \in [\pi_l, \pi_h]$ accepts the offer if $\pi_c^* = \pi_l$.

The following lemma summarises the agents' bargaining strategies:

Lemma 1 *At $t = 0$ the seller makes a take-it-or-leave-it offer of $p_c(\pi_c^*)$. The offer is accepted if the buyer is of type $\pi \in [\pi_c^*, \pi_h]$, and rejected if the buyer is of type $\pi \in [\pi_l, \pi_c^*)$.*

2.2 The Optimal Ownership Distribution

At $t = -1$ the parties contract over the ownership distribution. Since the agents are risk neutral, forward looking, and not wealth constrained, they always agree on the jointly efficient ownership distribution. Clearly, the ex ante division of surplus depends on the relative bargaining powers of the players. However, since the division of the surplus at the ex ante stage does not affect the analysis of the optimal ownership structure, we make no explicit assumptions about the relative bargaining powers at the ex ante stage.

Trade takes place if the buyer is of type $\pi \in [\pi_c^*(A), \pi_h]$. In this case the players generate a social surplus of π. In contrast, trade does not take place if the buyer is of type $\pi \in [\pi_l, \pi_c^*(A))$. The players then realise their joint disagreement payoffs $j(A)$. Hence, at $t = -1$, expected social surplus $W(A)$ is given by

$$W(A) = \frac{1}{\pi_h - \pi_l}\left(\int_{\pi_c^*(A)}^{\pi_h} \pi d\pi + \int_{\pi_l}^{\pi_c^*(A)} j(A) d\pi\right). \tag{5}$$

Let $\overline{A}$ and $\underline{A}$ denote the ownership distributions that, respectively, maximise and minimise $j(A)$. To find the optimal ownership distribution, we first analyse the efficiency implications of moving from $\overline{A}$ to $\underline{A}$. Let $\Delta W(\underline{A}, \overline{A}) = W(\underline{A}) - W(\overline{A})$. It then follows from (4) and (5) that

$$\Delta W(\underline{A}, \overline{A}) = \frac{1}{\pi_h - \pi_l}\left(\int_{\pi_c^*(\underline{A})}^{\pi_c^*(\overline{A})} \pi - j(\overline{A}) d\pi - \int_{\pi_l}^{\pi_c^*(\underline{A})} j(\overline{A}) - j(\underline{A}) d\pi\right). \tag{6}$$

This is the key equation in the model. We know from equation (4) that the seller's offer is (weakly) increasing in $j(A)$. Hence, the lower the joint disagreement payoffs, the more likely it is that the buyer accepts the seller's offer. The first integral on the RHS of (6) represents the corresponding efficiency gain. Clearly, in the case of disagreement, a reduction in the joint disagreement payoff constitutes a direct resource cost. This effect is represented by the second integral on the RHS of (6).

By moving from $\overline{A}$ to $\underline{A}$ the parties therefore increase the probability of trade but also make disagreement more costly. We now study the conditions under which either effect dominates. To do so, we prove the following lemma:

Lemma 2 *In the permissible parameter range the function $\Delta W(\underline{A},\overline{A})$ has the following properties:*

$$\Delta W(\underline{A},\overline{A}) = \begin{cases} 0 & \text{if } \alpha \in [0, \frac{1}{3}(\mu - j(\overline{A}))] \\ > 0 & \text{if } \alpha \in (\frac{1}{3}(\mu - j(\underline{A})), \frac{3}{5}(\mu - \frac{1}{2}(j(\overline{A}) + j(\underline{A}))) \\ \leq 0 & \text{if } \alpha \in [\frac{3}{5}(\mu - \frac{1}{2}(j(\overline{A}) + j(\underline{A})), \mu - j(\overline{A})] \end{cases}$$

and

$$\frac{\partial \Delta W(\underline{A},\overline{A})}{\partial \alpha} = \begin{cases} 0 & \text{if } \alpha \in [0, \frac{1}{3}(\mu - j(\overline{A}))) \\ > 0 & \text{if } \alpha \in (\frac{1}{3}(\mu - j(\overline{A})), \frac{1}{3}(\mu - j(\underline{A}))) \\ < 0 & \text{if } \alpha \in (\frac{1}{3}(\mu - j(\underline{A})), \mu - j(\overline{A})]. \end{cases}$$

Proof: see appendix.

Lemma 2 is illustrated in figure 1 and its intuition is straightforward. When the degree of uncertainty is small relative to the gains from trade, i.e. $\alpha \in [0, \frac{1}{3}(\mu - j(\overline{A})))$, the seller's offer is always accepted by the buyer, independent of the ownership distribution. This can be seen by noting that $\pi_c^*(\underline{A}) = \pi_c^*(\overline{A}) = \pi_l$. Thus, for these parameter values, moving from $\overline{A}$ to $\underline{A}$ neither increases the probability of agreement (since the parties agree immediately even when $A = \overline{A}$) nor does it lead to any efficiency losses (since the players never realise the disagreement payoffs). As a result, changes in the ownership distribution have no efficiency implications.

Consider now the effect of increasing the degree of uncertainty relative to the gains from trade so that $\alpha \in [\frac{1}{3}(\mu - j(\overline{A})), \frac{1}{3}(\mu - j(\underline{A}))]$. In this case some buyer types reject the seller's offer when the disagreement payoffs are high ($\pi_c^*(\overline{A}) > \pi_l$), while all buyer types accept the offer when the disagreement payoffs are low ($\pi_c^*(\underline{A}) = \pi_l$). Thus, moving from $\overline{A}$ to $\underline{A}$ increases the probability of agreement. Note that, in equilibrium, the disagreement payoffs are never realised when

$A = \underline{A}$ since, in this case, the buyer always accepts the seller's offer. As a result, there is no efficiency cost to moving from $\overline{A}$ to $\underline{A}$ so that, in this parameter range, $\triangle W(\underline{A}, \overline{A})$ is positive.

Finally, consider the effect of further increasing the degree of uncertainty relative to the gains from trade so that $\alpha \in (\frac{1}{3}(\mu - j(\underline{A})), \mu - j(\overline{A})]$. In this parameter region, some buyer types reject the seller's offer even when $A = \underline{A}$ (this can be seen by noting that $\pi_c^*(\overline{A}) > \pi_c^*(\underline{A}) > \pi_l$). Moving from $\overline{A}$ to $\underline{A}$ then still increases the probability of agreement but also makes disagreement more costly. The larger α, the stronger is the former effect relative to the latter. Hence, in this region, the expected efficiency gain of moving from $\overline{A}$ to $\underline{A}$ is decreasing in α. It can be shown that, for $\alpha \in (\frac{1}{3}(\mu - j(\underline{A})), \frac{3}{5}(\mu - \frac{1}{2}(j(\overline{A}) + j(\underline{A}))]$, moving from $\overline{A}$ to $\underline{A}$ reduces expected efficiency, while, for $\alpha \in (\frac{3}{5}(\mu - \frac{1}{2}(j(\overline{A}) + j(\underline{A})), \mu - j(\overline{A})]$, it increases expected efficiency. Note that $\frac{3}{5}(\mu - \frac{1}{2}(j(\overline{A}) + j(\underline{A}))$ is increasing in μ and decreasing in $j(\overline{A})$ and $j(\underline{A})$. Thus, the parameter space in which it is (weakly) optimal to reduce the disagreement payoffs is increasing in the μ and decreasing in $j(\overline{A})$ and $j(\underline{A})$.

So far we have only considered two possible ownership allocations, namely $\overline{A}$ and $\underline{A}$. The following lemma extends the analysis to all other possible ownership allocations.

Lemma 3 *Consider any $\widehat{A}$ such that $j(\widehat{A}) \in [j(\underline{A}), j(\overline{A})]$. It can be shown that*

$$\max[W(\overline{A}), W(\underline{A})] \geq W(\widehat{A}).$$

Proof: see appendix.

Together lemmas 2 and 3 establish the following proposition.

Proposition 1 *For $\alpha \in [0, \frac{3}{5}(\mu - \frac{1}{2}(j(\overline{A}) + j(\underline{A}))]$ it is weakly optimal for the buyer and the seller to minimise their disagreement payoffs by choosing the ownership distribution $\underline{A}$. For $\alpha \in (\frac{3}{5}(\mu - \frac{1}{2}(j(\overline{A}) + j(\underline{A})), \mu - j(\overline{A})]$ it is optimal for the parties to maximise their disagreement payoffs by choosing the ownership distribution $\overline{A}$.*

Proposition 1 shows that the efficient ownership distribution maximises the joint disagreement payoffs if the gains from trade are small relative to the degree of uncertainty. If the gains from trade are large relative to the degree of uncertainty, the efficient ownership allocation minimises the joint disagreement payoffs.

3 The Dynamic Model

The static model that was presented in the previous section is very simple and tractable. There are, however, two conceptual problems with this model: first, it assumes that the seller can only make one offer. If that offer is not accepted, trade does not take place. However, it is not clear why the parties cannot continue to bargain over the price of the input after the initial offer has been rejected. After all, there are still gains from trade that the parties could realise. Second, even if the parties have to stop bargaining over the input price after only one offer, it is not evident why, in the case of disagreement over the input price, they do not renegotiate the ownership distribution (since all relevant information is common knowledge we would expect such renegotiations to be efficient). Only if changes in the ownership distribution are irreversible would the parties not be able to engage in such renegotiations. While some ownership changes may indeed be irreversible (see section 5 for an example), a model that allows for reversible ownership changes is clearly more satisfactory. In this section we present a version of the model which addresses both of these issues. In particular, we now allow for a dynamic bargaining game in which the parties continue to bargain until agreement is reached and the gains from trade are realised. We will argue that the dynamic model leads to results that are very similar to those presented above.

The bargaining game that we consider here is related to Admati and Perry (1987). They study a Rubinstein infinite-horizon, alternating-offers model with one-sided asymmetric information in which players can delay making offers. We adopt their basic framework but extend the analysis in two ways: first, while they allow for only two types of buyers, we allow for a continuum of such types[7]. Second, and more importantly, in our model the players can realise non-zero payoffs during the delay period (i.e. we allow for non-zero 'inside options' or 'pre-agreement payoffs'). In Admati and Perry (1987) the inside options are assumed to be zero.

We now turn to the formal description of the model. The ex ante period ($t = -1$) is exactly as described in section 1.

Ex post the parties bargain over the price of the input. As in the static model, the seller's cost of producing the input is zero and the buyer's return from using the input in the final good production is given by π, where π is distributed uniformly on $[\pi_l, \pi_h]$. The buyer learns the realisation of π at the beginning of the bargaining period ($t = 0$). Before they reach an agreement over the input price the buyer and the seller can each trade with third parties on the spot market. By doing so the seller can realise an inside option of $s(A)$ per instant of time and the buyer can realise an inside option of $b(A)$ per instant of time[8]. We continue to assume that trade is always profitable, i.e.

$$\pi_l \geq \frac{j(A)}{r}, \quad \forall A, \tag{7}$$

where r is the positive discount rate.

The seller can make the first offer at any time $t \geq 0$. Thereafter the parties alternate making offers. The minimum time between offers is $\bar{t} = -\frac{1}{r}\log\delta$, where $\delta \in [0,1]$ is the discount factor from one period delay. After a player has received an offer he can take either of two actions: accept the offer, in which case trade takes place and the game ends, or reject the offer and make a

[7]Crampton (1992) studies a model in which he allows for a continuum of types and two-sided asymmetric information.

[8]Note the difference between inside and outside options. Inside options are the payoffs that the parties realise while they disagree temporarily. Outside options, in contrast, are the payoffs that the parties realise if they terminate bargaining. For full information models with positive inside options see, for instance, Binmore, Rubinstein, and Wolinsky (1986) and Muthoo (1999).

counteroffer. Either action can be taken at any time after the minimum delay period $\bar{t}$. After a player has made an offer he can take no further action until the other player has either accepted the offer or made a counteroffer.

Let (t, p) denote an outcome of the game in which the parties agree at time t to trade at price p. For any (t, p) the seller's payoff is given by $e^{-rt}p + (1- e^{-rt})\frac{s(A)}{r}$ and the buyer's payoff is given by $e^{-rt}(\pi - p) + (1- e^{-rt})\frac{b(A)}{r}$. Note that these preferences imply that the players are risk neutral and that they are impatient, in the sense that they prefer agreement today to the same agreement at any later date. Note also that the buyer is more impatient the larger the gains from trade.

Let $n \in \{1, 2, ..., \infty\}$ denote the number of offers that have been made and let τ_n denote the time after the minimum delay of $\bar{t}$ after which offer p_n has been made. After n rounds have been played the history is given by $h^n = \{p_i, \tau_i\}_{i=1}^n$. Throughout the analysis, when we state that a player replies "immediately", we mean that he replies without any further delay after the minimum time $\bar{t}$. Similarly, when we say that an offer p_n has been "delayed", we mean that $\tau_n > 0$.

A pure strategy σ_S for the seller and $\sigma_B(\pi)$ for the buyer specifies, for any history h^n after which it is the player's turn to move, the delay period τ_{n+1}, whether p_n is accepted, and, if not accepted, the counteroffer p_{n+1}. Let $\lambda(\pi)$ denote the seller's belief that the buyer is of type π, i.e. the seller believes that with probability $\lambda(\pi)$ the seller is of type π.

Below we study a sequential equilibrium of this game. A sequential equilibrium specifies the strategies σ_S and σ_B and a system of beliefs such that each strategy is optimal given the other strategy and beliefs, and the beliefs are consistent with Bayes' rule (when possible).

4 The Analysis of the Dynamic Model

We proceed as in the static model: we first describe the equilibrium of the bargaining game and then turn to the optimal ownership distribution.

4.1 The Bargaining game

In this section we construct a sequential equilibrium in pure strategies for the dynamic bargaining game. We believe that the equilibrium we study is appealing because of its simplicity[9]. However, in general, there may be a large number of equilibria and we do not attempt to characterise the conditions under which the equilibrium is unique. Thus, we do not show how bargaining must proceed and only study one form it might take.

Intuitively, the equilibrium can be described as follows: the seller makes an offer at $t = 0$. The buyer then either accepts the offer or rejects it. Acceptance takes place at $t = \bar{t}$. A buyer who rejects the offer delays her counteroffer so long as to distinguish herself from the buyer with a valuation that is just higher than her own. She then makes an offer that corresponds to the complete information equilibrium offer. This offer is immediately accepted by the seller. Below we show that an equilibrium of this form does indeed exist. We first describe the complete information equilibrium offer. We then show by how long the buyer has to delay her counteroffer to credibly signal her type. Finally, we derive the seller's optimal first offer and describe the players' strategies and the beliefs that support such an equilibrium.

To economise on notation, we suppress the ownership distribution A from all the functions that are applied and derived in this section. We analyse the impact of changes of A on the described equilibrium in the section 4.2.

The Full Information Game Rubinstein (1982) has shown that a full information alternating offers game with fixed time between offers has a unique subgame perfect equilibrium. It is straightforward to extend this analysis to the case of non-zero inside options and to show that such a game also has a unique subgame perfect equilibrium. The equilibrium depends on who makes the first offer. Let $p^B(\pi)$ denote the equilibrium price if a buyer of type π makes the

[9] It is also closely related to the equilibria studied by Admati and Perry (1987) and Crampton (1992).

first offer, and let $p^S(\pi)$ be the equilibrium price if the seller makes the first offer to a buyer of type π. It is straightforward to show (see for example Muthoo (1999, pp.138-43)) that

$$\begin{aligned} p^B(\pi) &= \frac{1}{1+\delta}(\delta\pi - \delta\frac{b}{r} + \frac{s}{r}) \\ p^S(\pi) &= \frac{1}{1+\delta}(\pi - \frac{b}{r} + \delta\frac{s}{r}). \end{aligned}$$

These prices have the property that each player is indifferent between trading at the other player's offer immediately and trading at the player's own offer next period, i.e.

$$\begin{aligned} \pi - p^S(\pi) &= \delta(\pi - p^B(\pi)) + (1-\delta)\frac{b}{r} \\ p^B(\pi) &= \delta p^S(\pi) + (1-\delta)\frac{s}{r}. \end{aligned} \tag{8}$$

Delay Suppose that the seller makes an offer which he expects to be accepted by any buyer of type $\pi \in [\pi_c, \pi_h]$ and rejected by any buyer of type $\pi \in [\pi_l, \pi_c)$. Suppose further that the seller believes the buyer to be of type $\pi(\tau, \pi_c)$ if, after having rejected the seller's offer, she makes a counteroffer after a delay of τ. Finally, suppose that if the seller believes the buyer to be of type π, the buyer cannot do better than to offer $p^B(\pi)$ and the seller cannot do better than to accept. Then the present discounted value of a buyer of type $\pi(\tau, \pi_c)$ from pretending to be of type $\pi(\widetilde{\tau}, \pi_c)$ by counteroffering $p^B(\pi(\widetilde{\tau}, \pi_c))$ after a delay $\widetilde{\tau}$ is given by

$$U(\widetilde{\tau}) = \delta e^{-r\widetilde{\tau}}(\pi(\tau,\pi_c) - p^B(\pi(\widetilde{\tau},\pi_c)) + (1-\delta e^{-r\widetilde{\tau}})\frac{b}{r}. \tag{9}$$

Substituting (8) into (9) and differentiating with respect to $\widetilde{\tau}$ gives the following first order condition

$$U'(\widetilde{\tau}) = -r[\pi(\tau,\pi_c) - p^B(\pi(\widetilde{\tau},\pi_c))] - \frac{\delta}{1+\delta}\frac{\partial\pi(\widetilde{\tau},\pi_c)}{\partial\tau} + b = 0. \tag{10}$$

If (10) describes a maximum, and we check the second order condition below, then the function $\pi(\tau,\pi_c)$ has to satisfy the first order condition (10) for $\tau = \widetilde{\tau}$. Thus, we get

$$\delta\frac{\partial\pi(\tau,\pi_c)}{\partial\tau} + r\pi(\tau,\pi_c) - j = 0. \tag{11}$$

Note that the buyer does not have to signal her type if she has the highest possible valuation, so that $\pi(0,\pi_c) = \pi_c$. The differential equation (11) can then be solved:

$$\pi(\tau,\pi_c) = \pi_c e^{-\frac{r}{\delta}\tau} + (1-e^{-\frac{r}{\delta}\tau})\frac{j}{r}. \tag{12}$$

We can now use (12) to confirm that the second order condition for the maximisation problem is satisfied. Differentiating (9) twice with respect to $\widetilde{\tau}$ and setting $\widetilde{\tau} = \tau$ gives

$$U''(\tau) = -\frac{r}{\delta}e^{-\frac{r}{\delta}\tau}(r\pi_c - j) < 0,$$

so that the second order condition is indeed satisfied.

Let $\tau(\pi, \pi_c)$ denote the inverse of $\pi(\tau, \pi_c)$ with respect to τ. Hence, if the seller knows the buyer types to be distributed on $[\pi_l, \pi_c]$, a buyer of type $\pi \in [\pi_l, \pi_c]$ can credibly signal her type by delaying her offer by $\tau(\pi, \pi_c)$. Rearranging (12) gives

$$e^{-r\tau(\pi,\pi_c)} = (\frac{r\pi - j}{r\pi_c - j})^{\delta}. \tag{13}$$

It is worth to make three observations about the function $\tau(\pi, \pi_c)$. First, $\tau(\pi, \pi_c)$ is decreasing in π. Thus, the time needed by the buyer to signal her type is longer the lower her profits. This feature of the model is, of course, due to the particular preferences we described above which implicitly assumed that a buyer is more patient the lower π. Second, the time the buyer needs to credibly signal her type is increasing in the joint inside options since

$$\frac{\partial\tau(\pi, \pi_c)}{\partial j} = \frac{\delta(\pi_c - \pi)}{(r\pi - j)(r\pi_c - j)} \geq 0, \text{ for } \pi \in [\pi_l, \pi_c].$$

This feature will be important in the discussion of the optimal ownership structure below. Finally, it can be shown that

$$\frac{\partial\tau(\pi, \pi_c)}{\partial \pi_c} = \frac{\delta(r\pi - j)}{(r\pi - j)(r\pi_c - j)} > 0,$$

so that signalling takes longer the larger the support of the distribution of buyer types.

The First Offer It follows from (8) that a buyer of type $\pi = \pi_c \in [\pi_l, \pi_h]$ is indifferent between immediately accepting an offer $p^S(\pi_c)$ and buying the input for $p^B(\pi_c)$ after the minimum delay $\bar{t}$. Also, any buyer $\pi > \pi_c$ strictly prefers accepting $p^S(\pi_c)$ and any buyer $\pi < \pi_c$ strictly prefers buying the input for $p^B(\pi_c)$ after the minimum delay $\bar{t}$.

Suppose that the seller makes an offer $p = p^S(\pi_c)$, for any $\pi_c \in [\pi_l, \pi_h]$, and that the players then respond in the following way: any buyer of type $\pi \in [\pi_c, \pi_h]$ immediately accepts while any buyer of type $\pi \in [\pi_l, \pi_c)$ rejects the offer and instead counteroffers $p^B(\pi)$ after $\tau(\pi, \pi_c)$ which the seller immediately accepts. Let $R(\pi_c)$ denote the seller's expected return from making an offer of $p^S(\pi_c)$. Then

$$\begin{aligned} R(\pi_c) &= \frac{1}{\pi_h - \pi_l}[\delta p^S(\pi_c)(\pi_h - \pi_c) + \int_{\pi_l}^{\pi_c} \delta^2 e^{-r\tau(\pi,\pi_c)} p^B(\pi) \\ &\quad +(1 - \delta^2 e^{-r\tau(\pi,\pi_c)})\frac{s(A)}{r} d\pi]. \end{aligned}$$

The optimal cutoff type π_c^* is then given by

$$\pi_c^* = \arg \max_{\pi_c \in [\pi_l, \pi_h]} R(\pi_c). \tag{14}$$

The next lemma describes the solution to (14) when $\delta \to 1$ (the case for $\delta < 1$ is described in the appendix).

Lemma 4 *For $\delta \to 1$ the maximisation problem (14) has a unique solution $\pi_c^* \in (\pi_l, \pi_h)$ which solves*

$$3r(\pi_h - \pi_c^*)(r\pi_c^* - j)^2 - (r\pi_c^* - j)^3 + (r\pi_l - j)^3 = 0. \tag{15}$$

It can be shown that

$$\frac{\partial \pi_c^*}{\partial j} > 0.$$

Proof: see appendix.

It is now straightforward to prove the following proposition:

Proposition 2 *The following strategies and beliefs form a sequential equilibrium of the dynamic bargaining game:*

- *The seller's beliefs: Suppose the seller's last offer was p^s. If $p^s = p^S(\pi_c)$, for any $\pi_c \in [\pi_l, \pi_h]$, then $\lambda(\pi) = 1$ if the buyer makes a counteroffer after a delay of $\tau(\pi, \pi_c)$ and $\lambda(\pi_l) = 1$ otherwise. If $p^s > p^S(\pi_h)$, then $\lambda(\pi) = 1$ if the buyer makes a counteroffer after a delay of $\tau(\pi, \pi_h)$ and $\lambda(\pi_l) = 1$ otherwise. If $p^s < p^S(\pi_l)$ and the buyer rejects the offer, then $\lambda(\pi_l) = 1$.*

- *The seller's strategy: The seller starts by offering $p^S(\pi_c^*)$. Suppose the buyer's last offer was p^b. If $\lambda(\pi) = 1$, then the seller immediately accepts if $p^b \geq p^B(\pi)$ and immediately counteroffers $p^S(\pi)$ if $p^b < p^B(\pi)$.*

- *The buyer's strategy: Suppose the seller's last offer was p^s. If $p^s = p^S(\pi_c)$, for any $\pi_c \in [\pi_l, \pi_h]$, then any buyer $\pi \in [\pi_c, \pi_h]$ accepts the offer while any buyer $\pi \in [\pi_l, \pi_c)$ counteroffers $p^B(\pi)$ after a delay of $\tau(\pi, \pi_c)$. If $p^s > p^S(\pi_h)$ then any buyer $\pi \in [\pi_l, \pi_h]$ counteroffers $p^B(\pi)$ after a delay of $\tau(\pi, \pi_h)$. If $p^s < p^S(\pi_l)$ then all buyers accept immediately.*

Proof: see appendix.

4.2 The Optimal Ownership Distribution

In this section we analyse the optimal ownership structure when the minimum time between offers gets arbitrarily small, i.e. $\delta \to 1$. It is one of the appealing features of the Admati and Perry (1987) analysis, which continues to hold in this model, that delay occurs with positive probability even if parties can reply to offers instantaneously.

At $t = -1$ the players contract over the ownership distribution. Since the agents are risk neutral, forward looking, and not wealth constrained, they always agree on the jointly efficient ownership distribution. The players know that, ex post, the seller makes an offer which is instantly accepted by any buyer of type $\pi \in [\pi_c^*(A), \pi_h]$. If the buyer is of such a type, then the parties immediately realise a joint surplus of π. Any buyer of type $\pi \in [\pi_l, \pi_c^*(A))$ does not accept the seller's initial offer and instead she waits $\tau(\pi, \pi_c^*(A))$ before making an offer that the seller accepts immediately. Thus, for any buyer of type $\pi \in [\pi_l, \pi_c^*(A))$ the joint payoff is given by

$$\pi e^{-r\tau(\pi,\pi_c(A))} + \frac{j(A)}{r}(1 - e^{-r\tau(\pi,\pi_c(A))}).$$

Expected social surplus is then given by

$$(16) \qquad W(A) = \frac{1}{(\pi_h - \pi_l)}[\int_{\pi_c^*(A)}^{\pi_h} \pi d\pi + \int_{\pi_l}^{\pi_c^*(A)} \pi e^{-r\tau(\pi,\pi_c^*(A))} + \frac{j(A)}{r}(1 - e^{-r\tau(\pi,\pi_c^*(A))})d\pi].$$

Recall that $\overline{A}$ and $\underline{A}$ denote the ownership distributions that, respectively, maximise and minimise $j(A)$. Changing the ownership distribution from $\overline{A}$ to $\underline{A}$ affects expected social surplus in two opposing ways: on the one hand, such a change accelerates agreement but, on the other hand, it also makes temporary disagreement more costly. To understand the acceleration effect, consider how $\pi_c^*(A)$ and $\tau(\pi, \pi_c^*(A))$ depend on the joint inside options. In lemma 4 we have shown that $\pi_c^*(A)$ is increasing in $j(A)$. Hence, the higher the joint inside options, the less likely it is that the buyer will accept the seller's first offer. We have also noted above that the delay period $\tau(\pi, \pi_c^*(A))$ is increasing in the joint inside options and in the cutoff type $\pi_c^*(A)$. Together these two effects constitute the *acceleration effect*: the lower the joint inside options, the faster the parties can be expected to reach agreement. While lower inside options accelerate agreement, they also make temporary disagreement more costly. The last term on the RHS of (16) gives the joint payoff during the delay period which, for a given $\tau(\pi, \pi_c^*(A))$, is clearly decreasing in the joint inside options $j(A)$.

It follows from this discussion that the increase in expected social surplus of moving from $\overline{A}$ to $\underline{A}$ is given by

$$
\begin{aligned}
(17)\quad \Delta W(\underline{A},\overline{A}) &= \frac{1}{r(\pi_h-\pi_l)}[\int_{\pi_c^*(\underline{A})}^{\pi_c^*(\overline{A})}(r\pi - j(\overline{A}))(1-e^{-r\tau(\pi,\pi_c^*(\overline{A}))})d\pi \\
&+ \int_{\pi_l}^{\pi_c^*(\underline{A})}(r\pi - j(\overline{A}))(e^{-r\tau(\pi,\pi_c^*(\underline{A}))} - e^{-r\tau(\pi,\pi_c^*(\overline{A}))})d\pi \\
&- \int_{\pi_l}^{\pi_c^*(\underline{A})}(j(\overline{A}) - j(\underline{A}))(1-e^{-r\tau(\pi,\pi_c^*(\underline{A}))})d\pi].
\end{aligned}
$$

The acceleration effect is given by the first two integrals. The last integral represents the resource cost of reducing the inside options. Note that, since $\pi_c^*(\overline{A}) > \pi_c^*(\underline{A}) > \pi_l > j(\overline{A}) > j(\underline{A})$, both the benefit (the acceleration effect) and the resource cost of reducing inside options are always weakly positive. Note also that, for $\alpha = 0$, the parties reach agreement instantaneously, independent of the ownership distribution. Thus, $\Delta W(\underline{A},\overline{A}) = 0$ for $\alpha = 0$.

It is worth to note the similarity of the expressions for $\Delta W(\underline{A},\overline{A})$ in the static (equation (6)) and the dynamic model (equation (17)). In the static model reducing $j(A)$ increases the probability of agreement, while making permanent disagreement more costly. Here a reduction in $j(A)$ accelerates agreement, while making temporary disagreement more costly.

In contrast to the static model we have, so far, not been able to find a complete analytic solution for the dynamic model[10]. To make further progress we now rely on simple simulations of equation (17). Consider figure 2 which plots the efficiency gain of moving from $\overline{A}$ to $\underline{A}$ for several different parameter configurations[11]. As expected, $\Delta W(\underline{A},\overline{A}) = 0$ for $\alpha = 0$. For small positive

[10] The problem we face is that we cannot find a closed form solution for π_c^* (see lemma 4).

[11] Obviously, simulations do not allow us to draw definite conclusion about the shape of the $\Delta W(\underline{A},\overline{A})$ function. We have made simulations for a large number of parameter values and the general shape of the graph is always as shown in figure 2. We hope that the insights from

values of α there is an expected efficiency gain of moving from $\overline{A}$ to $\underline{A}$. In this region the acceleration effect dominates the resource cost effect. As α becomes larger, however, the resource cost effect becomes larger relative to the acceleration effect. For large enough α the resource cost effect dominates the acceleration effect. In this case a reduction in the joint inside options leads to an expected efficiency loss. Note again the similarity of figures 1 and 2 which, respectively, plot the expected efficiency gain of moving from $\overline{A}$ to $\underline{A}$ for the static and the dynamic model. In both cases moving from $\overline{A}$ to $\underline{A}$ is efficiency enhancing (reducing) if uncertainty is small (large) relative to the gains from trade. In figure 2 it can be seen that $\max(W(j = 50), W(j = 20)) \geq W(j = 30)$. This suggests that, as in the static model, only $\underline{A}$ and $\overline{A}$ can be efficient ownership distributions[12]. The players then agree on an ownership distribution that minimises their joint inside options if the degree of uncertainty is small relative to the gains from trade and they agree on an ownership distribution that maximises the joint inside options if the degree of uncertainty is large relative to the gains from trade.

5 Discussion

The analysis above shows that the ownership distribution of physical assets influences the size of ex post inefficiencies. In this section we explore how the findings of our model relate to observed ownership patterns. We start by analysing a simple example. We then move on to discuss different ownership arrangements that reduce the agents' joint inside options. We conclude this section by briefly analysing the relationship between efficiency and the size of firms.

these simulations will allow us find an analytical solution for the shape of $\Delta W(\underline{A}, \overline{A})$ in future work.

[12] We again repeated these simulations for a large number of parameter configurations which all gave the same result.

5.1 A Simple Example

Suppose there are only two assets, a_1 and a_2. Let A_B and A_S denote the assets which are owned, respectively, by the buyer and the seller given an ownership distribution A. We assume that the players are symmetric, in the sense that $s(A_S) = b(A_B), \forall A_S = A_B$. We refer to 'integration' as the ownership structure in which both assets are owned by one agent, and 'non-integration' as the one in which each agent owns one asset[13]. Furthermore, it is useful to introduce the following definitions:

Definition 1 *The assets a_1 and a_2 are 'synergetic' if*

$$b(a_1, a_2) + b(\phi) \geq b(a_1) + b(a_2).$$

Definition 2 *The assets a_1 and a_2 are 'non-synergetic' if*

$$b(a_1, a_2) + b(\phi) < b(a_1) + b(a_2).$$

Definition 3 *The assets a_1 and a_2 are 'strictly complementary' if*

$$b(a_1, a_2) > b(a_1) = b(a_2) = b(\phi).$$

Thus, assets are synergetic if, in the case of temporary disagreement, the joint payoff is higher under integration than under non-integration. Correspondingly, assets are non-synergetic if, in the case of temporary disagreement, the joint payoff is higher under non-integration than under integration. In the case of a restaurant, for instance, the room and the furniture used to equip it are likely to be synergetic assets: if the chef owns the room and the waiter owns the furniture, then neither is able to generate much surplus without the cooperation of the other. If, instead, the chef owns both assets, then, in the case of non-coorporation, he would be able to generate some surplus by serving a smaller

[13]Because of the symmetry of the example we do not need to distinguish between buyer and seller integration or between the two possible non-integration cases.

number of customers himself. In contrast, a restaurant and the plant used by the restaurant's supplier of foodstuffs are likely to be non-synergetic: in the case of disagreement more surplus can be generated if each is run independently of the other.

The analysis in the previous sections has shown that the parties optimally choose an ownership distribution which minimises their joint inside options when the degree of uncertainty is small relative to the gains from trade. When a_1 and a_2 are jointly owned by the buyer and the seller, then neither can use any assets during the delay period and their joint inside option is minimised to $2b(\phi)$. To see why joint ownership is optimal in this parameter range, consider moving from joint ownership to any other ownership distribution. Independent of whether the assets are synergetic or non-synergetic, such a change in the ownership distribution would lead to higher joint inside options. On the one hand, this implies a reduction in the cost of disagreement. On the other hand, however, it also implies a longer duration of the disagreement period. The latter effect dominates the former when the gains from trade are large relative to the degree of uncertainty. In this case moving from joint ownership to any other ownership distribution does, on average, lead to an efficiency loss.

The parties optimally agree on an ownership distribution that maximises their joint inside options when the degree of uncertainty is large relative to the gains from trade. Thus, integration is optimal when assets are synergetic, and non-integration is optimal when assets are non-synergetic. Note that this might provide an explanation for why firms own a large number of assets and do not distribute them among their work force[14]: a firm can expect to spend a lot of time haggling with its work force, even after distributing its assets among the workers, when it faces a very uncertain environment. When the firm's assets are synergetic, it is optimal for the firm to reduce the cost of haggling by becoming the owner of all the assets.

[14]Holmström (1999) argues that this asset clustering is one of the most striking ownership patterns that we observe and that it is not easily explained by the property rights literature.

Note that, in the case of non-synergetic assets, the above arguments imply a negative relationship between the average size of firms and the degree of uncertainty that they are exposed to[15]: it is optimal to have one large (jointly owned) firm when uncertainty is small, while two small, separately owned firms are optimal when uncertainty is high (relative to the gains from trade). It is well known that small firms tend to have more volatile share prices than larger firms. Taking the volatility of the share price as a proxy for the uncertainty in a firm's environment our argument reverses the causality that is usually used to explain this stylized fact: it is not because firms are small that they have a very volatile share price. Instead, it might be that, because a firm faces a very uncertain environment, it optimally chooses to be small so as to minimise bargaining inefficiencies. These arguments suggest a way of empirically testing our theory[16].

5.2 Reducing Joint Inside Options

We have already seen that the joint ownership of physical assets reduces joint inside options. Giving veto rights over certain aspects of a firm's operation to one of its patrons is another closely related arrangement[17]. In this section we briefly discuss other observed ownership arrangements that reduce the parties' joint inside options.

Agents who are locked-in sometimes increase their interdependence further by exchanging 'ugly princess hostages'[18]. The term refers to a practice in which agents exchange ownership of assets that are very valuable to themselves, but

[15]In the case of synergetic assets there is no relationship between the size of firms and the degree of uncertainty.

[16]To our knowledge the existing empirical literature on the size of firms (see, for example, Kumar, Rajan, and Zingales (1999)) does not take into account the degree of uncertainty in a firm's environment as a determinant of its size.

[17]Another similar arrangement is the German law on codetermination which entitles workers to elect half of the members of the supervisory board in all large German firms (see Hansmann (1996), p.110-112). Codetermination further increases the interdependence between firms and their work force. Note, however, that codetermination is required by law and is only observed in Germany.

[18]See, for example, Williamson (1985).

which have little value for the other party. The Japanese car industry provides an example of such an arrangement. There it can be observed that physical assets which are specific to a particular car manufacturer are often not owned by that firm but by its supplier[19].

Agents can also reduce joint inside options by abolishing assets that can be used unilaterally during a disagreement period. For instance, Holmström and Roberts (1998) report the case of an airline alliance between KLM and Northwest Airlines[20]. In this case the airlines deliberately increased their interdependence by eliminating their duplicate support operations. Interestingly, they did so after running into a costly dispute which led to the dismantling of the companies' cross-ownership structure.

Firms that are engaged in long term vertical relationships often make themselves remarkably dependent on each other by using exclusive sourcing arrangements. The Ford case that was described in the introduction is one example. Holmström and Roberts (1998) provide another example. They report the case of Nucor, which has been the most successful steel manufacturer in the US in the last 20 years. Instead of following the industry standard by integrating backward, the firm made itself very dependent on only one independent supplier. Given the size of Nucor, the supplier is also very dependent on this relationship[21]. We can interpret such an arrangement as an example of non-integration in the case of synergetic assets. By not integrating backwards, Nucor and its supplier increase their interdependence which, in turn, reduces the expected duration of potential disputes. This argument is in line with Holmström and Roberts (1998) who hypothesise that, in the case of Nucor, "one reason why the partnership has been working so well may be the high degree of mutual dependence"[22].

Finally, it can be observed that firms at times agree on the separate ownership

[19] Holmström and Roberts (1998), p.80-81.
[20] Holmström and Roberts (1998), p.84.
[21] The supplier is estimated to make about 50% of its scrap business with Nucor.
[22] Holmström and Roberts (1998), p.83.

of strictly complementary assets. For instance, Dnes (1993) observes that, in some franchise agreements, the franchisee and the franchisor agree on the separate ownership of the land on which the premise is built and the assets used to equip it. The separate ownership of strictly complementary assets is similar to joint ownership and minimises the parties' joint inside options.

The analysis in the previous section has shown that lowering joint inside options reduces ex post inefficiencies when the gains from trade are large relative to the degree of uncertainty. To some extend our model is therefore consistent with the type of ownership arrangements described in this section. Note, however, that in our model it is optimal to either maximise or minimise joint inside options. Ownership arrangements that reduce joint inside options, but do not minimise them, can therefore not be fully explained with the given framework[23]. Note, also, that our model is not the only one in the property rights literature in which joint ownership can be optimal. In particular, Rajan and Zingales (1998), de Meza and Lockwood (1998), and Chiu (1998) show that joint ownership can be optimal in models that focus on ex ante investment inefficiencies.

5.3 Decreasing Returns from Managerial Inputs

It is often argued that there are decreasing returns from managerial inputs and that this puts an upper bound on the size of firms. For instance, Coase (1937) argues that: "[...] as a firm gets larger, there may be decreasing returns to the entrepreneur function, that is, the costs of organizing additional transactions within the firm may rise"[24]. The analysis in this chapter suggest another reason for why large firms may be inefficient. To see this, consider a firm that increases its size by buying additional assets from a third party. It seems natural to assume that, at least in a weak sense, the cost of disagreeing with any patron is decreasing in the number of assets owned by the firm. This may,

[23] We conjecture that interior solutions can be optimal in a version of our model in which the gains from trade can be negative, and in which there is uncertainty about the agents' *outside* options.

[24] Coase (1937), p.340.

for instance, be due to the firm's ability to produce some of the inputs it usually obtains from a supplier. As a consequence, the size of a firm is positively related to the expected duration of disagreement, and it is negatively related to the cost of disagreement. If the gains from trade are large relative to the degree of uncertainty, then the former effect dominates the latter. In this case firms that increase in size become less efficient *because* they are in a strong bargaining position vis-a-vis their patrons and are expected to spend a long time haggling with them.

6 Conclusion

There are many situations in which agents, firstly, depend on each other to generate a surplus and, secondly, anticipate that bargaining over the sharing of the surplus may be costly. In this chapter we have shown that in such a situation agents may have an incentive to take actions prior to the bargaining stage to further increase their interdependence, in the sense of lowering the sum of the payoffs the parties realise during the bargaining process. Increased interdependence accelerates decision making but also makes disagreement more costly. In the analysis we have shown that the former effect dominates the latter if uncertainty is small relative to the agents' expected gains from trade. In this case it is efficient for the agents to increase their interdependence by minimising their joint pre-agreement payoffs. The opposite is true if uncertainty is large relative to the gains from trade. In this case it is efficient for the agents to minimise their interdependence by maximising their joint pre-agreement payoffs.

In this chapter we have applied these arguments to the theory of the firm and discussed the efficiency implications of various observed ownership patterns. We believe that the basic effects in our model can also help us understand other institutions and legal arrangements. An obvious example is the marriage contract which can be interpreted as a contract that increases the interdependence of two partners who anticipate to be locked-in in future[25]. In this interpretation

[25] Here lock-in may not only be due to exogenous factors, such a mutual attraction, but also

the marriage contract is a means of accelerating domestic decision making which, however, comes at the cost of lower disagreement payoffs[26]. While we believe that the basic effects in our model can help us understand a number of institutions, it is also evident that our formal model is quite restrictive. In particular, among other assumptions, we only allow for two players, one-sided asymmetric information, and assume a particular extensive form bargaining game. Relaxing any of these assumption may be interesting future work.

to relationship-specific investments.

[26]In so far as our model can be extended to multilateral bargaining situations, arguments similar to those presented above might also be used to explain the institution used by the Roman Catholic church to elect a new pope. A new pope is elected by an assembly of cardinals who are locked up in a part of Vatican Palace until they reach agreement. This institution, called a 'conclave', originates in the 13th century when the cardinals failed to elect a new pope for two years. A local magistrate then decided to improve the cardinals' incentives by locking them up in the episcopal palace, removing its roof, and allowing them nothing but bread and water until they elected the next pope (for more details see www.britannica.com). The observation that this institution has not been abandoned (but only somewhat adapted) suggests it might be efficient for the church as a whole (including the decision making cardinals) to accelerate the decision making process by lowering the payoff the cardinals realise during their negotiations.

7 Appendix

Proof of lemma 2:

From (4) it follows that

$$\begin{aligned}
\pi_c^*(\underline{A}) &= \pi_c^*(\overline{A}) = \pi_l \ \text{ for } \alpha \in [0, \frac{1}{3}(\mu - j(\overline{A}))] \\
\pi_c^*(\overline{A}) &> \pi_c^*(\underline{A}) = \pi_l \ \text{ for } \alpha \in (\frac{1}{3}(\mu - j(\overline{A})), \frac{1}{3}(\mu - j(\underline{A}))] \\
\pi_c^*(\overline{A}) &> \pi_c^*(\underline{A}) > \pi_l \ \text{ for } \alpha \in (\frac{1}{3}(\mu - j(\underline{A})), \mu - j(\underline{A}))].
\end{aligned}$$

Consider first $\alpha \in [0, \frac{1}{3}(\mu - j(\overline{A}))]$. It then follows from (6) that $W(\underline{A}) - W(\overline{A}) = 0$.

Consider next $\alpha \in (\frac{1}{3}(\mu - j(\overline{A})), \frac{1}{3}(\mu - j(\underline{A}))]$. From (6) we get

$$W(\underline{A}) - W(\overline{A}) = \frac{1}{\pi_h - \pi_l} \int_{\pi_l}^{\pi_c^*(\overline{A})} \pi - j(\overline{A}) d\pi.$$

Note that in this parameter range $\pi_c^*(\overline{A}) > \pi_l$ and that, by assumption (1), $\pi_l \geq j(\overline{A})$. It then follows immediately that $W(\underline{A}) - W(\overline{A}) > 0$. Substituting (4) into (6) and solving gives

$$\Delta W(\underline{A}, \overline{A}) = \frac{1}{16\alpha}[3\alpha - (\mu - j(\overline{A}))][3(\mu - j(\overline{A})) - \alpha].$$

Differentiating gives

$$\frac{\partial \Delta W(\underline{A}, \overline{A})}{\partial \alpha} = \frac{1}{16\alpha^2}((\mu - j(\overline{A}))^2 - \alpha^2).$$

Because of assumption (1) it then follows that $\frac{\partial \Delta W(\underline{A}, \overline{A})}{\partial \alpha} > 0$.

Finally, consider $\alpha \in (\frac{1}{3}(\mu - j(\underline{A})), \mu - j(\underline{A}))]$. Substituting (4) into (6) and solving gives

$$\Delta W(\underline{A}, \overline{A}) = \frac{j(\overline{A}) - j(\underline{A})}{16\alpha}(6\mu - 3j(\overline{A}) - 3j(\underline{A}) - 10\alpha)$$

Hence, $\Delta W(\underline{A}, \overline{A}) > 0$ if $\alpha \in (\frac{1}{3}(\mu - j(\underline{A})), \frac{3}{5}(\mu - \frac{1}{2}(j(\overline{A}) + j(\underline{A})))$ and $\Delta W(\underline{A}, \overline{A}) \leq 0$ if $\alpha \in [\frac{3}{5}(\mu - \frac{1}{2}(j(\overline{A}) + j(\underline{A})), \mu - j(\underline{A}))]$. Also, taking the derivative gives

$$\frac{\partial \Delta W(\underline{A}, \overline{A})}{\partial \alpha} = \frac{-3(j(\overline{A}) - j(\underline{A}))}{8\alpha^2}(\mu - \frac{1}{2}(j(\overline{A}) + j(\underline{A}))) < 0. \quad \blacksquare$$

Proof of proposition 1:

Let $j(q) = qj(\overline{A}) + (1-q)j(\underline{A})$. Furthermore, let $\widetilde{q}$ be implicitly defined by $\frac{1}{2}(\pi_h + j(\widetilde{q})) = \pi_l$. Then (5) can be rewritten as

$$W(q) = \frac{1}{\pi_h - \pi_l}\left(\int_{\pi_c^*(j(q))}^{\pi_h} \pi d\pi + \int_{\pi_l}^{\pi_c^*(j(q))} j(q) d\pi\right) \text{ for } q \in [0,1]. \tag{18}$$

Differentiating (18) with respect to q gives

$$W'(q) = \begin{cases} 0 & \text{if } q < \widetilde{q} \\ \frac{1}{8\alpha}(j(\overline{A}) - j(\underline{A}))(\pi_h + 3j(q) - 4\pi_l) & \text{if } q > \widetilde{q} \end{cases}$$

and

$$W'_-(\widetilde{q}) = 0$$
$$W'_+(\widetilde{q}) = \tfrac{1}{8\alpha}(j(\overline{A}) - j(\underline{A}))(\pi_h + 3j(\widetilde{q}) - 4\pi_l).$$

Differentiating twice gives

$$W''(q) = \begin{cases} 0 & \text{if } q < \widetilde{q} \\ \frac{3}{8\alpha}(j(\overline{A}) - j(\underline{A}))^2 & \text{if } q > \widetilde{q} \end{cases}$$

and

$$W''_-(\widetilde{q}) = 0$$
$$W''_+(\widetilde{q}) = \tfrac{3}{8\alpha}(j(\overline{A}) - j(\underline{A}))^2.$$

Hence, $W(q)$ has a kink at $q = \widetilde{q}$. It is flat for $q < \widetilde{q}$ which implies that $W(\widetilde{q}) = W(0)$ for $q \leq \widetilde{q}$. For $q > \widetilde{q}$ the function $W(q)$ is convex. Clearly, local maxima of a convex function are given by the corner solutions, i.e. $q = \widetilde{q}$ or $q = 1$. Hence, $\max[W(q=1), W(q=0)] \geq W(\widehat{q})$ for any $\widehat{q} \in [0,1]$. ■

Proof of lemma 4:

Substituting (8) and (13) into (14) and differentiating with respect to π_c gives

$$R'(\pi_c) = \frac{1}{(\pi_h - \pi_l)}[\delta p^{S\prime}(\pi_c)(\pi_h - \pi_c) - \delta p^S(\pi_c)$$
$$+\delta^2 p^B(\pi_c) + (1-\delta^2)\frac{s}{r} - \frac{\delta^4}{1+\delta}\int_{\pi_l}^{\pi_c}(\frac{r\pi - j}{r\pi_c - j})^{\delta+1} d\pi]$$

and

$$R''(\pi_c) = \frac{-1}{(2+\delta)(\pi_h - \pi_l)}[2(2-\delta)\delta + \delta^4(\frac{r\pi_l - j}{r\pi_c - j})^{\delta+2}]. \tag{19}$$

Note that $R''(\pi_c) < 0$ for any $\pi_c \in [\pi_l, \pi_h]$. For $\delta \to 1$ the first order condition reduces to

$$(20) \qquad R'(\pi_c) = \frac{1}{2(\pi_h - \pi_l)}(\pi_h - \pi_c - \int_{\pi_l}^{\pi_c} (\frac{r\pi - j}{r\pi_c - j})^2 d\pi) = 0.$$

Note that

$$R'(\pi_l) = \frac{3r(r\pi_l - j)(\pi_h - \pi_l)}{2r^3(\pi_h - \pi_l)} > 0$$

and

$$R'(\pi_h) = \frac{-[(r\pi_h - j)^3 - (r\pi_l - j)^3]}{2r^3(\pi_h - \pi_l)} < 0.$$

Thus, for $\delta \to 1$ it must be that $\pi_c^* \in (\pi_l, \pi_h)$.

Since $R''(\pi_c^*) < 0$ we can apply the implicit function theorem to (20). Totally differentiating (20) gives

$$(\int_{\pi_l}^{\pi_c} \frac{r(r\pi - j)(\pi_c^* - \pi)}{(r\pi_c^* - j)^3} d\pi)dj - \frac{1}{3}(2 + (\frac{r\pi_l - j}{r\pi_c^* - j})^3)d\pi_c^* = 0$$

Hence,

$$\frac{\partial \pi_c^*}{\partial j} = \frac{r(\pi_c^* - \pi_l)^2(r\pi_c^* + 2r\pi_l - 3j)}{4(r\pi_c^* - j)^3 + 2(r\pi_l - j)^3} > 0. \qquad \blacksquare$$

Proof of proposition 2:

To verify that these beliefs and strategies are an equilibrium we have to show that each strategy is a best response given the other strategy and that the beliefs are consistent with Bayes' rule.

Consider first the seller's strategy. Suppose the buyer's last offer was p^b and that $\lambda(\pi) = 1$. The seller can then either accept this offer, in which case he realises a payoff p^b, or he can reject it and make a counteroffer p^s. If $p^s \leq p^S(\pi)$, then the seller expects the buyer to accept immediately, while if $p^s > p^S(\pi)$ the seller expects the buyer to offer $p^B(\pi)$ after signalling her type. Thus, if the seller rejects p^b, then he cannot do better than to offer $p^S(\pi)$ and realise an expected payoff of $\delta p^S(\pi) + (1-\delta)\frac{s}{r} = p^B(\pi)$. It then follows that, if $\lambda(\pi) = 1$, the seller's best response to an offer p^b is to accept immediately if $p^b \geq p^B(\pi)$ and immediately counteroffer $p^S(\pi)$ if $p^b < p^B(\pi)$. That the first offer $p_c(\pi_c^*)$ is a best response to the buyer's strategy follows immediately from the fact that π_c^* solves the maximisation problem (14).

Consider now the best response of any buyer $\widetilde{\pi}$ to an offer of $p^s = p^S(\pi_c)$, for any $\pi_c \in [\pi_l, \pi_h]$. Note first that, given the seller's strategy and the construction of the function $\tau(\pi, \pi_c)$, the buyer $\widetilde{\pi}$ always prefers to make the counteroffer $p^B(\widetilde{\pi})$ at $\tau(\widetilde{\pi}, \pi_c)$ to counteroffering $p^B(\widehat{\pi})$ after $\tau(\widehat{\pi}, \pi_c)$, for any $\widehat{\pi} \neq \widetilde{\pi}$. Also, if the buyer $\widetilde{\pi}$ makes a counteroffer after a delay $\tau(\pi, \pi_c)$, then, given the seller's beliefs and his strategy, she cannot do better than to offer $p^B(\pi)$. Finally, it can never be optimal for the buyer to make an offer after a delay $\tau > \tau(\pi_l, \pi_c)$. Thus, if the buyer rejects p^s, then she cannot do better than to counteroffer $p^B(\widetilde{\pi})$ after $\tau(\widetilde{\pi}, \pi_c)$, in which case she makes $\delta e^{-r\tau(\widetilde{\pi},\pi_c)}(\widetilde{\pi} - p^B(\widetilde{\pi})) + (1 - \delta e^{-r\tau(\widetilde{\pi},\pi_c)})\frac{b}{r}$. If the buyer $\widetilde{\pi}$ accepts p^s, then she realises $\widetilde{\pi} - p^S(\pi_c)$. It follows that it is a best response for the buyer $\widetilde{\pi}$ to accept p^s if $\widetilde{\pi} \geq \pi_c$ and it is a best response for her to counteroffer $p^B(\widetilde{\pi})$ after $\tau(\widetilde{\pi}, \pi_c)$ if $\widetilde{\pi} < \pi_c$. Suppose now that the seller's last offer was $p^s > p^S(\pi_h)$. Since, for any $\widetilde{\pi}$, $\widetilde{\pi} - p^S(\pi_h) \leq \delta(\widetilde{\pi} - p^B(\pi_h)) + (1-\delta)\frac{b}{r}$ it follows that it is a best response for any buyer $\widetilde{\pi}$ to reject the offer and counteroffer $p^B(\widetilde{\pi})$ after $\tau(\widetilde{\pi}, \pi_h)$. Finally, suppose the seller's last offer was $p^s < p^S(\pi_l)$. If the buyer $\widetilde{\pi}$ rejects, then she cannot do better than to immediately counteroffer $p^B(\pi_l)$. Since, for any $\widetilde{\pi}$, $\widetilde{\pi} - p^S(\pi_l) \geq \delta(\widetilde{\pi} - p^B(\pi_l)) + (1-\delta)\frac{b}{r}$ it follows that the buyer's best response is to accept any $p^s < p^S(\pi_l)$

At last, we need to check that, on the equilibrium path, the seller's beliefs are consistent with Bayes' rule. Given the strategies, the seller's first offer $p_c(\pi_c^*)$ is immediately accepted by any buyer of type $\pi \in [\pi_c^*, \pi_h]$. Any buyer of type $\pi \in [\pi_l, \pi_c^*)$ rejects $p_c(\pi_c^*)$ and offers $p^B(\pi)$ at $\tau(\pi, \pi_c^*)$. Thus, on the equilibrium path an offer of $p^B(\pi)$ is made after a delay of $\tau(\pi, \pi_c^*)$ if and only if the buyer is of type π. Thus, the belief that $\lambda(\pi) = 1$ if the buyer makes an offer $p^B(\pi)$ after a delay of $\tau(\pi, \pi_c^*)$ is consistent with Bayes' rule. ■

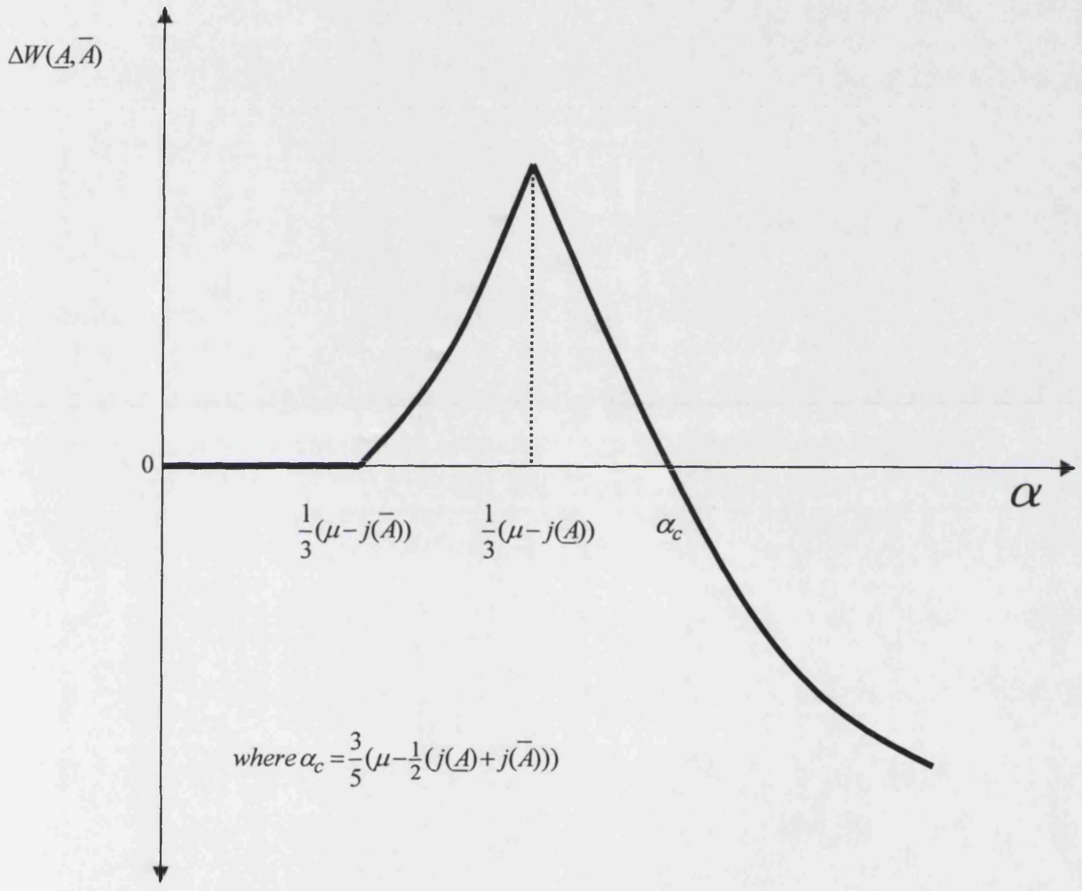

Figure 1: Efficiency gain of moving from $\overline{A}$ to $\underline{A}$ in the static model.

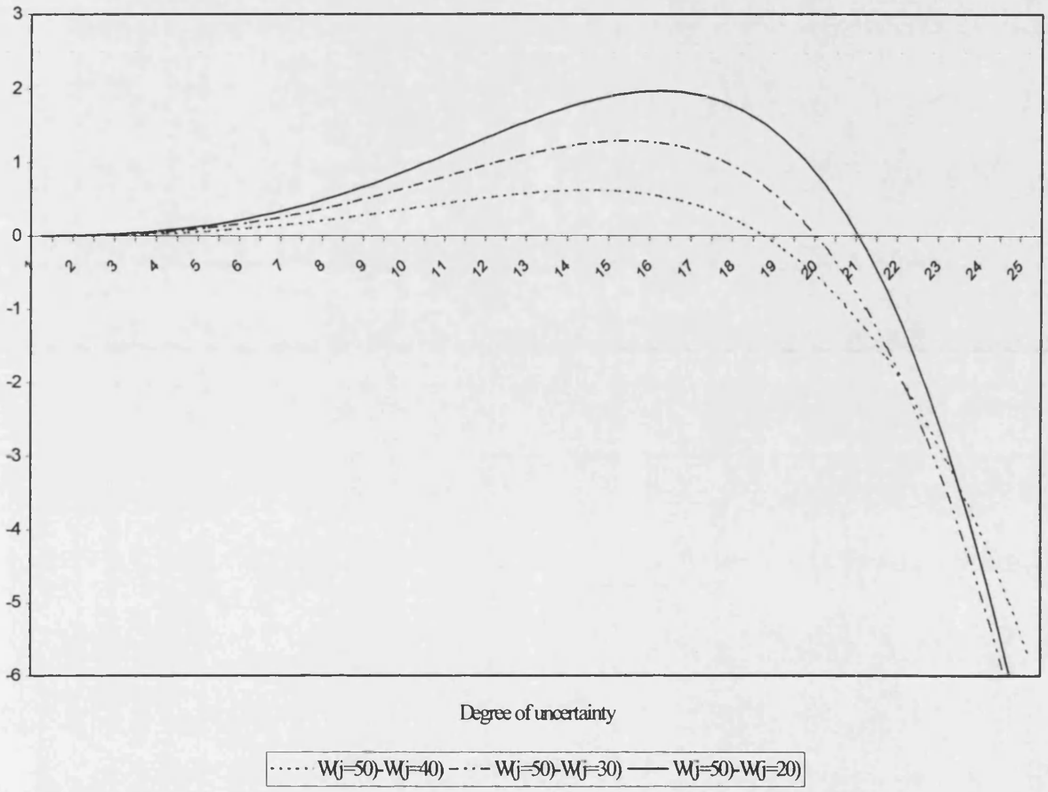

Figure 2: Simulated efficiency gain of moving from $\overline{A}$ to $\underline{A}$ in the dynamic model for $\mu = 100$, $r = 1$, $j(\overline{A}) = 50$ and $j(\underline{A}) = 40, 30, 20$.

Chapter 3

Coercive Power and Economic Exchange: The Invisible Hand versus the Grabbing Hand

The protection of property rights and the enforcement of contracts are key determinants for the development of an economy[1]. In most democracies the state uses its monopolised coercive powers to enforce private contracts. There are, however, numerous instances in which states either give up some of their monopoly powers or at least refrain from using them for contract enforcement purposes. Private organisations with coercive powers may then have an incentive to enter the market and satisfy the demand for contract enforcement. Hay, Shleifer and Vishny (1996), for instance, observe that, because of the inefficiencies in the Russian legal system, "[...] business people stay away from using the legal system, and use the services of organised crime instead"[2]. Also, Gambetta (1993) argues convincingly that, contrary to public perception, the Sicilian Mafia's primary business is that of contract enforcement.

In contrast to other markets, the provision of contract enforcement services requires the control of coercive powers by the suppliers. This suggests a particular problem with the private provision of contract enforcement services in so far as potential suppliers may find it more profitable to engage in extortion rather than to supply these services in the market place. In this chapter we take a closer look at the incentives of agents with coercive power. In particular, we study when such powerful agents have an incentive to enforce private contracts, and analyse the degree to which they engage in potentially distortionary

[1] See, among many others, Rosenthal (1992), who argues that at least a part of the difference in economic development between France and England in the 18th and 19th century can be explained by differences in their legal systems.

[2] Hay, Shleifer, and Vishny (1996), p.562. For a discussion of contract enforcement in Russia see also Greif and Kandel (1995), Pistor (1996), and Hay and Shleifer (1998).

extortion. We compare the powerful agents' optimal strategies under different governmental regimes and describe the welfare implications.

As a benchmark we first consider an anarchic economy in which nobody has enough coercive power to determine the allocation of assets and payoffs unilaterally. Because of the lack of force there is no extortion in such an economy. However, for the same reason, an anarchy depends on reputation building to ensure that agents abide by contracts. While reputation building mechanisms work well in many situations, they are less useful when agents interact infrequently or have short time horizons[3]. Thus, while an anarchy might experience little distortionary taxation, the lack of coercive power also prevents some transactions from taking place.

We then show that a patient monopoly mafia has an incentive to provide contract enforcement services. This enables private agents to transact, thereby allowing them to realise gains from trade. However, the mafia only enforces contracts to maximise the total surplus from which it extorts its own income. Since extortion distorts the private agents' investment incentives, it also reduces welfare. Thus, the welfare implications of moving from an anarchic society to one in which contract enforcement services are supplied by a monopoly mafia turn out to be ambiguous.

To analyse the effect of competition in the contract enforcement industry, we consider a regime in which coercive power is shared by two competing mafias. In this regime, as in the other regimes described in this chapter, the private agents do not have the option of leaving the economy for one with a more favourable governmental system. The mafias are therefore 'competing' in so far as each is trying to extract as much surplus as possible from the same fixed group of private agents[4]. We argue that, in this case, the two mafias have a strong incentive to collude by coordinating their contract enforcement and extortionary activities.

[3] Also, even when reputation building is possible it can be inefficient to rely on such mechanisms exclusively since they may constitute barriers to entry for new firms.

[4] Note that this is different from the more familiar type of competition where jurisdictions compete to attract mobile subjects by lowering taxes or providing other types of privileges. See section 3.

In the model we show that, as long as conflicts between mafia families are not too wasteful, competition between mafia families leads to more extortion and less contract enforcement, and therefore reduces welfare.

Finally, we compare the public and private provision of contract enforcement services. For this purpose we analyse a democratic economy in which the citizens can use elections to discipline self-interested powerful agents. We show that public contract enforcement is more efficient than the provision of such services by a private monopoly mafia.

This chapter is structured as follows: the next section presents the model. In section 2 we solve the model and compare social welfare between the various regimes. Section 3 discusses the related literature and section 4 concludes.

1 The Model

The model is an infinitely repeated game. We first describe the basic stage game when there are no elections. In section 1.2 we turn to the stage game in which the powerful agent faces elections.

1.1 The Stage Game without Elections

The basic stage game is a sequential game with four stages denoted by s_i with $i = 1, 2, 3, 4$. There is an infinitely large population of ex ante identical, risk neutral, and infinitely lived agents. We will, at times, refer to these agents as 'private agents' (as opposed to the 'powerful agents' described below). At s_1 two agents are randomly drawn from this population, one of who becomes an entrepreneur 'E' and the other a worker 'L'. They are each endowed with one discrete unit of labour and their respective labour supply is denoted by $l_E \in \{0, 1\}$ and $l_L \in \{0, 1\}$. Their utility functions are given by $U_j = y_j - l_j u_j$, for $j = E, L$, where y_j denotes agent j's consumption of the final good y and $u_j \geq 0$ denotes his disutility of working. We assume that the worker has a comparative advantage in supplying labour in the sense that $u_E \geq u_L$.

At s_2 the entrepreneur can make an unobservable investment i at cost $c(i)$. We assume that, $\forall i$, $c'(i) \geq 0$, $c''(i) > 0$, $c'''(i) \geq 0$ as well as $c'(0) = 0$ and $c'(1) = \infty$. With probability i the investment is successful and leads to the development of a new production technology while with probability $1 - i$ the entrepreneur fails in developing a new technology. The investment's success is indicated by $z \in \{0, 1\}$, where $z = 1$ if a new technology was developed and $z = 0$ otherwise. We assume that a new technology is only operational for the duration of one stage game. At times we refer to a stage game in which a new production technology has been developed as a 'successful stage game'.

At s_3 the entrepreneur and the worker choose simultaneously how much labour to supply. The labour supply by each agent is observable. The new technology, together with labour input $l = l_E + l_L$, generates output according to the following production technology:

$$y = \begin{cases} 1 & \text{if } l \geq 1 \\ 0 & \text{otherwise.} \end{cases}$$

We assume that the technology is productive even when it is operated by the entrepreneur, in the sense that $u_E < 1$. Also, while labour input is supplied at s_3, output is only generated at the next stage.

At s_4 the output is generated, distributed among the agents, and consumed. In the current stage game, without any elections, we allow for three separate regimes which only differ with respect to the specification of stage s_4. In the first regime, which we refer to as 'anarchy', any output that is produced first accrues to the entrepreneur who can then decide how to distribute it between herself and the other agents.

In the second regime, which we refer to as the 'monopoly mafia regime', there is one 'powerful agent' (denoted by 'M') who distinguishes himself from the private agents in so far as he has access to an enforcement technology that allows him to unilaterally determine the distribution of final output across agents. In particular, at s_4 the powerful agent chooses a distribution $d = (t, r, w)$, where t, r, and w denote the fraction of y that is given to the powerful agent, the entrepreneur, and the worker respectively. The distribution D that is actually implemented is then given by $D = d$. Further, we assume that the powerful agent is infinitely lived and derives utility from the consumption of the final good, so that her per stage game utility is given by $U_M = y_M$, where y_M is her consumption of the final good. In contrast to the entrepreneur and the worker, the identity of the powerful agent cannot change from one stage game to the next.

Finally, we consider a regime, the 'mafia cartel regime', in which there are two powerful agents '$\mathrm{M_m}$', for $m = 1, 2$. These agents are again infinitely lived and their per stage game utility is given by $U_{M_m} = y_{M_m}$, where y_{M_m} denotes $\mathrm{M_m}$'s consumption of the final good. At s_4 the powerful agents jointly and uncooperatively determine the distribution of payoffs across all agents. Specifically, at s_4 each powerful agent chooses a distribution $d_m = (t_1, t_2, r, w)$, where t_m denotes the fraction of y that is given to $\mathrm{M_m}$. The distribution that is actually implemented is then given by

$$D = \begin{cases} d_1 & \text{if } d_1 = d_2 \\ \frac{\alpha}{2}(d_1 + d_2) & \text{if } d_1 \neq d_2, \end{cases} \tag{1}$$

where $\alpha \in [0, 1]$ is an exogenous parameter. The enforcement technology has the following interpretation: if the two powerful agents agree on the same distribution, then this particular distribution is also implemented. In the case of disagreement, the implemented distribution is a convex combination of d_1, d_2 (each with an equal weight of $\frac{\alpha}{2}$), and the zero vector (with weight $1 - \alpha$). The parameter α therefore indicates how wasteful it is for the two powerful agents to engage in conflict. For simplicity we only consider the case in which the powerful agents are symmetric, which is captured by the equal weighting of d_1 and d_2. This implementation technology is, of course, entirely ad hoc. Nevertheless, it allows us to describe some effects that we believe to be interesting.

Note that the powerful agents only have power over physical assets (i.e. the distribution of the final good); by assumption they cannot force the entrepreneur and the worker to invest or to provide labour. As was demonstrated on the Roman galleys, physical force and intense monitoring can also be used to induce human capital investments. However, we still make this assumption since we believe that there are dimensions to human capital investments which can only be provided voluntarily. In particular, private agents cannot be forced directly to engage in the type of human capital investments that are important for the development of an economy, such as learning and conducting research.

After the payoff distribution has been determined at s_4 consumption takes place and the stage game ends. Summarising, the four stages are given by:

s_1 E and L are randomly chosen.

s_2 E invests i.

s_3 E and L decide whether to supply labour.

$$s_4 \begin{cases} E \text{ decides on payoff distribution ('anarchy')} \\ M \text{ chooses } d \text{ ('monopoly mafia regime')} \\ M_1 \text{ and } M_2 \text{ choose } d_1 \text{ and } d_2 \text{ ('mafia cartel regime').} \end{cases}$$

Finally, we assume that, while there is no discounting within each stage game, payoffs between two subsequent stage games are discounted at a rate $\delta \in [0, 1)$. Social surplus per stage game is denoted by V.

1.2 The Stage Game with Elections

We now turn to the case in which the group of private agents (the citizens) can delegate the monopoly over coercive power to any private agent for the duration of at least one stage game. We need to distinguish between the first and all subsequent stage games. At the beginning of the first stage game one agent is drawn randomly from the population of private agents and given the monopoly over coercive power. The next stages s_1 to s_4 are then the same as those described above. At s_5 there is an election in which the private agents can vote between two alternatives: either they re-elect the incumbent 'I', in which case this agent remains 'in office' until the end of the next stage game, or they vote to remove the incumbent from office. In this case the incumbent is replaced by an agent who is randomly drawn from the population of private agents[5]. From the second stage game onwards the stages are therefore given by:

s_1 E and L are randomly chosen.

s_2 E invests i.

s_3 E and L decide whether to supply labour.

s_4 I chooses d.

s_5 Election takes place.

2 The Analysis

In this section we solve the different games that were described above and compare social welfare between the regimes.

2.1 Anarchy

We first consider an anarchic economy in which there are no powerful agents. Recall that the population is assumed to be infinitely large. Hence, a worker and an entrepreneur who are drawn in one stage game do not expect to meet again in the future. Reputation building mechanisms are therefore not applicable in this environment[6]. Consider then the perfect Bayesian equilibrium of the stage game. At s_4 it is always optimal for the entrepreneur to consume the

[5]Thus, we do not allow private agents to abstain, they have to vote either for or against the incumbent.

[6]For the analysis of two historical cases in which reputation mechanisms were applied see Greif (1993) and Greif, Milgrom, and Weingast (1994).

entire amount of $y \in \{0, 1\}$ that has been produced and not to share any output with the worker. At s_3 the entrepreur and the worker have to decide how much labour to supply after observing whether or not a new machine has been developed. When a new technology has been developed the entrepreneur and the worker then face a standard time inconsistency problem: it would be profitable for the entrepreneur to 'hire' the worker, and for the worker to supply labour, since $u_L \leq u_E < 1$. However, the worker cannot commit himself to work after receiving his wage and the entrepreneur cannot commit herself to pay the worker after he supplied labour. Thus, at s_3 it is never optimal for the worker to supply any labour (independent of his beliefs about the entrepreneur's investments) and instead the entrepreneur has to engage in home production by providing labour herself. At s_2 the entrepreneur has to decide how much to invest. Anticipating a return of $1 - u_E$ if a new technology is developed, and zero otherwise, she optimally invests

$$i^a = i^*(1 - u_E) = \arg\max_i i(1 - u_E) - c(i). \tag{2}$$

In the unique perfect Bayesian equilibrium the entrepreneur then invests i^a (at s_2), plays $l_E = 1$ if the investment was successful and $l_E = 0$ otherwise (at s_3), and consumes any output that is produced (at s_4). At s_3 the worker always plays $l_L = 0$ and always believes that the entrepreneur invested i^a[7]. It immediately follows that, under anarchy, social welfare per stage game is given by V^a, where

$$V^a = i^a(1 - u_E) - c(i^a) > 0. \tag{3}$$

[7]Note that there are no information sets off the equilibrium path so that we do not have to be concerned with off the equilibrium path beliefs.

2.2 Monopoly Mafia

We now consider the monopoly mafia regime in which there is one powerful agent. It is well known that, for 'high enough' δ, an infinitely repeated game of the type described above can have a large number of trigger strategy equilibria. Here we focus on one particular trigger strategy equilibrium, namely the one in which the present discounted value of the monopoly mafia's equilibrium payoff is maximised. Most of this section is used to derive and describe this equilibrium.

The trigger strategies that maximise the mafia's equilibrium payoff must have the following four features: first, they must punish deviation by the mafia as severely as possible, in the sense that, following deviation, the mafia realises its reservation payoff of zero ever after. This has to be the case since in equilibrium deviation never occurs so that a more severe punishment merely increases the set of sustainable outcomes. Second, on the equilibrium path, the players must cooperate in every stage game. The more often the players cooperate on the equilibrium path, the higher their total equilibrium payoffs (and the lower the minimum discount rate for which their strategies form an equilibrium). Third, on the equilibrium path, it must be the workers, and not the entrepreneurs, who supply labour in every successful stage game. This has to be the case since, in every successful stage game, resources $u_E - u_L \geq 0$ can be 'saved' by having a worker rather than an entrepreneur supply labour. These resources can then be used to increase the mafia's payoff either directly, by increasing the level of extortion, or indirectly, by improving the entrepreneurs' investment incentives. Finally, the payoff distribution that is implemented on the equilibrium path must be chosen so as to optimise the trade-off between the level of extortion, the entrepreneurs' investment incentives, and the workers' equilibrium payoffs. The following strategies $\sigma_M(\overline{d})$, $\sigma_E(\overline{d})$, and $\sigma_L(\overline{d})$ satisfy the first three of these requirements:

- $\sigma_M(\overline{d})$: The mafia plays $\overline{d} = (\overline{t}, \overline{r}, \overline{w})$ in the first stage game and continues to do so as long as it always played $\overline{d}$ in the past. If ever the mafia plays any $d \neq \overline{d}$ it plays $\underline{d} = (1, 0, 0)$ at every subsequent opportunity.

- $\sigma_E(\overline{d})$: The first entrepreneur plays $i^*(\overline{r})$. Any subsequent entrepreneur plays $i^*(\overline{r})$ if and only if the mafia never played any $d \neq \overline{d}$ and she plays $i = 0$ otherwise. The entrepreneurs always play $l_E = 0$.

- $\sigma_L(\overline{d})$: The first worker plays $l_L = 1$ if the investment was successful and $l_L = 0$ otherwise. Any subsequent worker plays $l_L = 1$ if and only if the investment was successful and the mafia never played any $d \neq \overline{d}$ and he plays $l_L = 0$ otherwise.

Since there is imperfect information about the entrepreneurs' investment levels the other players have to form beliefs about them. The worker and the mafia can be in two 'types' of information sets, one in which $z = 0$ and the other in which $z = 1$. In an information set with $z \in \{0, 1\}$ the probability that the mafia and the worker attach to the entrepreneur having invested i is denoted by $b_M(i, z)$ and $b_L(i, z)$ respectively. Consider then the following beliefs:

- $b(\overline{d})$: If the mafia never played any $d \neq \overline{d}$ in the past, then $b_M(i^*(\overline{r}), z) = b_L(i^*(\overline{r}), z) = 1$, $\forall z$. If the mafia every played any $d \neq \overline{d}$, then the beliefs satisfy $\int_0^1 b_M(i, z)di = \int_0^1 b_L(i, z)di = 1$, $\forall z$.

It is now straightforward to prove the following proposition:

Proposition 1 *The strategies $\sigma_M(\overline{d})$, $\sigma_E(\overline{d})$, and $\sigma_L(\overline{d})$ and the beliefs $b(\overline{d})$ form a perfect Bayesian equilibrium if and only if the following resource (RC), no-cheating (NC), and incentive (IC) constraints are satisfied:*

$$\overline{t} + \overline{r} + \overline{w} \leq 1 \tag{RC}$$

$$\overline{t} + \frac{\delta}{1-\delta} i^*(\overline{r})\overline{t} - 1 \geq 0 \tag{NC}$$

$$\overline{w} \geq u_L. \tag{IC}$$

The constraints can be satisfied simultaneously if and only if $\delta \in [\underline{\delta}, 1)$, where $\underline{\delta} \in (0, 1)$ is defined below.

Proof: see appendix.

We now turn to the fourth requirement for profit maximisation, namely that the distribution $\overline{d}$ is chosen so as to optimise the trade-off between the level of extortion, the entrepreneurs' investment incentives, and the workers' equilibrium payoff. Suppose that $\delta \in [\underline{\delta}, 1)$ and that the players play the strategies $\sigma_M(\overline{d})$, $\sigma_E(\overline{d})$, and $\sigma_L(\overline{d})$ (and have the beliefs $b(\overline{d})$). The mafia's total equilibrium payoff is then given by

$$\pi(\overline{r}, \overline{t}) = \frac{1}{1-\delta} i^*(\overline{r})\overline{t}.$$

The profit maximising distribution $\overline{d}^* = \left(\overline{t}^*, \overline{r}^*, \overline{w}^*\right)$ maximises $\pi(\overline{r}, \overline{t})$ subject to the RC, NC, and IC constraints. It is evident that the mafia's profits can only be maximised if no resources are wasted, which is equivalent to saying that the RC constraint is satisfied with equality. Also, the mafia's profits can only be maximised if the workers always receive a zero net payoff. Thus, the IC constraint must be satisfied with equality, so that $\overline{w}^* = u_L$. Substituting into the RC constraint then gives $\overline{r}^* = 1 - u_L - \overline{t}^*$. Hence, to fully describe $\overline{d}^*$, we now only need to find $\overline{t}^*$ by solving

(4) $$\max_{t \in [0, 1-u_L]} \pi(1 - u_L - t, t)$$

$$\text{s.t.} \qquad NC(t, \delta) = t + \frac{\delta}{1-\delta} i^*(1 - u_L - t)t - 1 \geq 0.$$

In the appendix we prove the following proposition:

Proposition 2 *For $\delta \in [0, \underline{\delta})$ maximisation problem (4) does not have a solution. For $\delta \in [\underline{\delta}, 1)$ the solution is given by*

$$\overline{t}^* = \begin{cases} T & \text{if } \delta \in [\overline{\delta}, 1) \\ \widetilde{T}(\delta) & \text{if } \delta \in [\underline{\delta}, \overline{\delta}), \end{cases}$$

where $T = \arg\max_{t \in [0,1-u_L]} \pi(1 - u_L - t, t)$*,* $\widetilde{T}(\delta)$ *is the smallest* $t \in [0, 1 - u_L]$ *that satisfies* $NC(t, \delta) = 0$*, and* $\underline{\delta} \in (0, \overline{\delta})$ *is defined below.*

Furthermore, it can be shown that

i. $T \in (0, 1 - u_L)$*,* $T = \widetilde{T}(\overline{\delta})$*, and* T *is independent of* δ*.*

ii. for $\delta \in (\underline{\delta}, 1)$*,* $\widetilde{T}(\delta) \in (0, 1 - u_L)$ *and*

$$\frac{\partial \widetilde{T}(\delta)}{\partial \delta} < 0 \ .$$

Proof: see appendix.

For $\delta \in [0, \underline{\delta})$ there does not exist a distribution $\overline{d}$ that can simultaneously satisfy the RC, NC, and IC constraints. Thus, in this parameter range, there does not exist a trigger strategy equilibrium in which the workers supply labour. Since, as was argued above, labour supply by the workers is more efficient than labour supply by the entrepreneurs, this implies that there also does not exist a trigger strategy equilibrium in which the entrepreneurs supply any labour. Hence, for $\delta \in [0, \underline{\delta})$, each player simply plays the stage game equilibrium strategies in every stage game. It is straightforward to show (for a proof see proposition 12 in the appendix) that, in any perfect Bayesian equilibrium of the stage game, the mafia engages in complete extortion and, anticipating extortion, the private agents do not invest and generate no surplus. Thus, when δ is too low to sustain trigger strategy equilibria, all players realise zero payoffs. Note that each player's stage game equilibrium payoff is equal to each player's reservation payoff.

For the remainder of the analysis we assume that, in the monopoly mafia regime, the players play the strategies $\sigma_M(\overline{d}^*)$, $\sigma_E(\overline{d}^*)$, and $\sigma_L(\overline{d}^*)$ and have the beliefs $b(\overline{d}^*)$ as long as $\delta \in [\underline{\delta}, 1)$. For $\delta \in [0, \underline{\delta})$ the players use their stage game equilibrium strategies of always playing $i = l_E = l_L = 0$ and $d = (1, 0, 0)$ in every stage game.

The analysis above shows that a patient profit maximising mafia does not only engage in extortion, but also provides 'contract enforcement' services. In this chapter we say that a mafia 'enforces contracts' when it implements a distribution that compensates the workers for their labour supply. That the provision of contract enforcement services is optimal for a patient mafia can be seen by noting that $\overline{w}^* = u_L$ for $\delta \in [\underline{\delta}, 1)$. By enforcing contracts the mafia allows the transaction between the worker and the entrepreneur to take place: at s_3 the entrepreneur can 'hire' the worker since the latter is sure to receive u_L in return for his labour supply. It is optimal for the mafia to enforce contracts simply because resources can be saved by having the worker rather than the entrepreneur engage in production. As was noted above, the mafia can then use these extra resources to increase its extortionary income either directly, by increasing t, or indirectly, by improving the entrepreneurs' investment incentives.

We have already seen that a profit maximising mafia engages in extortion. Note, however, that a patient mafia does not extort the entrepreneur completely and always leaves some surplus with her (see proposition 2). Clearly, a patient mafia has an incentive to do so since the entrepreneurs only invest, and thereby create the surplus from which the mafia can extort, if they expect to receive at least some of the returns from their investments[8].

Proposition 2 also shows that the optimal level of extortion is (weakly) decreasing in the discount rate δ. While this seems to be a very intuitive result, it depends crucially on the sequential nature of the production process in our model. To see why this is the case, note first that a profit maximising mafia would ideally want to commit itself to an extortion rate T. This rate, which is independent of δ, is only credible if the mafia is very patient, i.e. $\delta \in [\overline{\delta}, 1)$. In this case $NC(T, \delta) \geq 0$, and thus $\overline{t}^* = T$. Consider now a slightly lower discount rate such as $\delta = \overline{\delta} - \varepsilon$. Then the implementation of the distribution $\overline{d} = (T, 1 - u_L - T, u_L)$ is not credible: the private agents anticipate that, whenever $y = 1$, the mafia prefers engaging in full extortion to implementing

[8]The observation that it might be optimal for an organisation like the mafia to refrain from full extortion is, of course, not new and was indeed already made by Franchetti in 1876 who, in the context of the Sicilian Mafia, stated that: "If the villains made use of their destructive abilities to an extreme degree, they would soon lack the very matter from which to steal" (Franchetti [1876] (1974), p.126, as quoted in Gambetta (1993), p.33).

$\overline{d}$. To make its tax policy credible, the mafia must then *increase* the extortion rate beyond T[9]. Since T maximises the expected tax return $i(t)t$, charging a higher extortion rate reduces the expected tax return. However, it increases the mafia's no-cheating payoff $(t + \frac{\delta}{1-\delta}i^*(1 - u_L - t)t)$ relative to its cheating payoff (1). For $\delta \in [\underline{\delta}, \overline{\delta})$ the profit maximising rate of extortion is then given by $\widetilde{T}(\delta) > T$, i.e. the smallest level of t that satisfies the no-cheating constraint. Proposition 2 shows that $\widetilde{T}(\delta)$ is decreasing in δ. This is the case since the mafia puts less weight on the negative incentive effect of deviation the less patient it is. Thus, the lower δ, the more profitable is deviation relative to non-deviation. To prevent deviation in the case of a lower discount rate, the equilibrium extortion rate must therefore be increased.

We conclude this section by deriving V^m, the expected social surplus (per stage game) that is generated in the monopoly mafia regime. For $\delta \in [0, \underline{\delta})$ the mafia always engages in full extortion, and hence no surplus is generated. For $\delta \in [\underline{\delta}, 1)$ the profit maximising mafia enforces contracts and only engages in limited extortion. Thus,

$$V^m = \begin{cases} 0 & \text{if } \delta \in [0, \underline{\delta}) \\ i^m(1 - u_L) - c(i^m) & \text{if } \delta \in [\underline{\delta}, 1), \end{cases} \tag{5}$$

where

$$i^m = i^*(1 - u_L - \overline{t}^*).$$

2.3 Anarchy versus the Monopoly Mafia

In an anarchy contracts cannot be enforced and, as a result, some transaction do not take place. A patient monopoly mafia improves efficiency by providing contract enforcement services. However, the ultimate aim of a mafia is to engage

[9] In other words, the mafia optimally chooses a tax rate on the downsward sloping part of the Laffer curve so as to make its tax policy credible.

in extortion which, in turn, reduces welfare. In this section we analyse these opposing welfare implications.

First, suppose $\delta \in [\underline{\delta}, 1)$. It follows from (3) and (5) that the difference in social surplus between the two regimes is then given by

$$(6) \qquad V^m - V^a = i^a(u_E - u_L) + (i^m - i^a)(1 - u_L) - (c(i^m) - c(i^a)).$$

The first term on the RHS of (6) gives the direct efficiency gain that is due to the enforcement of contracts: by allowing the transaction between the worker and the entrepreneur to take place, the mafia saves resources of size $u_E - u_L$ in any successful stage game. This term is always (weakly) positive since the worker has a comparative advantage in supplying labour.

The mafia only provides contract enforcement services to increase its extortionary income. Clearly, contract enforcement and extortion have opposing effects on the entrepreneurs' investment incentives: while contract enforcement makes the technology more productive and therefore improves the investment incentives, extortion reduces investment incentives by diminishing the entrepreneurs' return from developing a new technology[10]. The last two terms in (6) describe this incentive effect. It is straightforward to show that the incentive effect is positive if $u_E - u_L > \bar{t}^*$ and negative otherwise (see proof of proposition 3). Hence, a mafia improves the entrepreneurs' investment incentives if the gains from trade are larger than the level of extortion. In this case all agents, including the entrepreneurs, prefer a regime with a monopoly mafia to an anarchic society.

If the level of extortion is larger than the gains from trade, however, the mafia

[10]Note that lump sum extortion would not reduce the entrepreneurs' investment incentives. In the analysis we rule out lump sum extortion. We believe that there are a number of reasons why this is a reasonable assumption in the present context. For instance, the economies that we have in mind are very likely to have imperfect credit markets.

reduces the entrepreneurs' investment incentives. In this case the difference in social surplus between anarchy and the monopoly mafia regime is ambiguous: on the one hand, the mafia saves resources $i^m(u_E - u_L) \geq 0$ by enforcing contracts while, on the other hand, it reduces the entrepreneurs' investment incentives by engaging in extortion, thereby decreasing welfare by $(i^m - i^a)(1 - u_E) - (c(i^m) - c(i^a)) < 0$. The overall effect depends on the relative size of the gains from trade $u_E - u_L$ and the tax rate $\overline{t}^*$. In particular, we prove below (see proposition 3) that, for $\delta \in [\underline{\delta}, 1)$, the difference in expected social surplus between the monopoly mafia and the anarchic regime is increasing in gains from trade $u_E - u_L$ and the discount rate δ.

The mafia always engages in full extortion, and thus no surplus is generated, if $\delta \in [0, \underline{\delta})$. In contrast, in an anarchy the entrepreneurs engage in home production and social surplus is generated even if $\delta \in [0, \underline{\delta})$ (see (3)). Hence, in this parameter range, the presence of a monopoly mafia reduces welfare relative to an anarchy.

The discussion in this section can be summarised in the following proposition:

Proposition 3 *The difference in social surplus between the monopoly mafia regime and anarchy is given by*

$$V^m - V^a \begin{cases} < 0 & \text{if } \delta \in [0, \underline{\delta}) \\ \geq 0 & \text{if } \delta \in [\underline{\delta}, 1) \text{ and } u_E - u_L \geq \overline{t}^* \\ \gtrless 0 & \text{if } \delta \in [\underline{\delta}, 1) \text{ and } u_E - u_L < \overline{t}^*. \end{cases}$$

For $\delta \in [\underline{\delta}, 1)$, $V^m - V^a$ is weakly increasing in δ and strictly increasing in $u_E - u_L$.

Proof: see appendix.

2.4 Mafia Cartel

We have seen above that a profit maximising monopoly mafia engages in extortion and, possibly, enforces contracts. In this section we analyse how the existence of a second mafia affects the provision of contract enforcement services and the level of extortion. Consistent with the analysis in the previous sections we again focus on one particular equilibrium, namely the one in which the mafias' equilibrium payoffs are maximised. Here this implies that the mafias form a cartel and use trigger strategies to support collusion.

The informal literature on mafias and organised crime notes that collusion between separate mafias (or mafia families) is a widespread phenomenon (see Gambetta (1993)). In this model, and we believe in reality, mafias have an incentive to collude because of two main interdependencies: first, it is a feature of the enforcement technology that the implemented payoff distribution is a function of both the mafias' actions (see (1)). Thus, it is difficult, if not impossible, for one mafia to provide contract enforcement services without the implicit cooperation of the other mafias. Second, private agents cannot target their 'punishment' strategies at one particular mafia: when a mafia engages in unexpectedly high levels of extortion, the private agents respond by reducing their investment levels in future periods. This leads to a reduction in the potential extortionary income for *all* mafias. Hence, all mafias get 'punished' for excessive extortion by any one mafia.

The analysis required to describe the trigger strategy equilibrium that maximises the cartel's equilibrium payoff is very similar to the analysis in the monopoly mafia case. The trigger strategies that maximise the cartel's equilibrium payoff must have the same four features as the corresponding strategies in the monopoly mafia case: first, deviation by any mafia must be punished as severely as possible. Second, on the equilibrium path, the players must cooperate in every stage game. Third, on the equilibrium path, the workers rather than the entrepreneurs must supply labour. Finally, the payoff distribution that is implemented on the equilibrium path must be chosen so as to optimise the trade-off between the level of extortion, the entrepreneurs' investment incentives, and the workers' equilibrium payoffs. The following strategies $\sigma^c_{M_1}(\overline{d}^c)$, $\sigma^c_{M_2}(\overline{d}^c)$, $\sigma^c_E(\overline{d}^c)$, and $\sigma^c_L(\overline{d}^c)$ satisfy the first three requirements.

- $\sigma^c_{M_m}(\overline{d}^c)$, for $m = 1, 2$: Each mafia plays $\overline{d}^c = (\frac{\overline{t}^c}{2}, \frac{\overline{t}^c}{2}, \overline{r}^c, \overline{w}^c)$ in the first stage game and continues to do so as long as it was never the case that $d_1 \neq \overline{d}^c$ or $d_2 \neq \overline{d}^c$. If ever $d_1 \neq \overline{d}^c$ or $d_2 \neq \overline{d}^c$, then M_1 plays $\underline{d}^c_1 = (1, 0, 0, 0)$ and M_2 plays $\underline{d}^c_2 = (0, 1, 0, 0)$ at every subsequent opportunity.

- $\sigma^c_E(\overline{d}^c)$: The first entrepreneur plays $i^*(\overline{r}^c)$. Any subsequent entrepreneur plays $i^*(\overline{r}^c)$ if and only if it was never the case that $d_1 \neq \overline{d}^c$ or $d_2 \neq \overline{d}^c$ and she plays $i = 0$ otherwise. The entrepreneurs always play $l_E = 0$.

- $\sigma^c_L(\overline{d}^c)$: The first worker plays $l_L = 1$ if the investment was successful and $l_L = 0$ otherwise. Any subsequent worker plays $l_L = 1$ if and only if the investment was successful and it was never the case that $d_1 \neq \overline{d}^c$ or $d_2 \neq \overline{d}^c$ and he plays $l_L = 0$ otherwise.

Again we also have to specify beliefs to describe the equilibrium. Let $b_{M_m}(i, z)$, for $m = 1, 2$, denote the probability that M_m attaches to the entrepreneur having invested i given that M_m is in an information set with $z \in \{0, 1\}$. Consider then the following beliefs:

- $b^c(\overline{d}^c)$: If it was never the case that $d_1 \neq \overline{d}^c$ or $d_2 \neq \overline{d}^c$, then $b_{M_m}(i^*(\overline{r}^c), z) = b_L(i^*(\overline{r}^c), z) = 1$, $\forall z, m$. If it was ever the case that $d_1 \neq \overline{d}^c$ or $d_2 \neq \overline{d}^c$, then the beliefs satisfy $\int_0^1 b_{M_m}(i, z)di = \int_0^1 b_L(i, z)di = 1$, $\forall z, m$.

It is straightforward to prove the following proposition:

Proposition 4 *The strategies $\sigma^c_M(\overline{d}^c)$, $\sigma^c_E(\overline{d}^c)$, and $\sigma^c_L(\overline{d}^c)$ and the beliefs $b^c(\overline{d}^c)$ form a perfect Bayesian equilibrium if and only if the following resource (RC^c), no-cheating (NC^c), and incentive (IC^c) constraints are satisfied:*

$$\overline{t}^c + \overline{r}^c + \overline{w}^c \leq 1 \tag{RC^c}$$

$$\overline{t}^c + \frac{\delta}{1-\delta} i^*(\overline{r}^c)\overline{t}^c - \alpha(1 + \frac{\overline{t}^c}{2}) \geq 0 \tag{NC^c}$$

$$\overline{w}^c \geq u_L. \tag{IC^c}$$

The constraints can be satisfied simultaneously if and only if $\delta \in [\underline{\delta}^c, 1)$, where $\underline{\delta}^c \in [0, 1)$ is defined below.

Proof: see appendix.

Suppose that $\delta \in [\underline{\delta}^c, 1)$ and that the players play the strategies $\sigma_M^c(\overline{d}^c)$, $\sigma_E^c(\overline{d}^c)$, and $\sigma_L^c(\overline{d}^c)$ (and have the beliefs $b^c(\overline{d}^c)$). The cartel's joint equilibrium payoff is then given by

$$\pi(\overline{r}^c, \overline{t}^c) = \frac{1}{1-\delta} i^*(\overline{r}^c)\overline{t}^c.$$

The profit maximising distribution $\overline{d}^{c*} = (\frac{\overline{t}^{c*}}{2}, \frac{\overline{t}^{c*}}{2}, \overline{r}^{c*}, \overline{w}^{c*})$ maximises $\pi(\overline{r}^c, \overline{t}^c)$ subject to the RC^c, NC^c, and IC^c constraints. Just as in the case of a monopoly mafia, it must be that $\overline{w}^{c*} = u_L$ and $\overline{r}^{c*} = 1 - u_L - \overline{t}^{c*}$. To describe $\overline{d}^{c*}$, we therefore only need to find $\overline{t}^{c*}$ by solving

$$\max_{t \in [0, 1-u_L]} \pi(1 - u_L - t, t) \tag{7}$$

$$\text{s.t.} \quad NC^c(t, \delta) = t + \frac{\delta}{1-\delta} i^*(1 - u_L - t)t - \alpha(1 + \frac{t}{2}) \geq 0.$$

Any difference between the distribution that is implemented by a profit maximising monopolist ($\overline{d}^*$) and that implemented by a profit maximising cartel ($\overline{d}^{c*}$) is due to the difference in the no-cheating constraints ($NC(t, \delta) \geq 0$ and $NC^c(t, \delta) \geq 0$). The existence of a second mafia changes the no-cheating constraint in two ways: on the one hand, a second mafia reduces the no-cheating payoff since, in a cartel, the two mafias have to share the total exortionary income, while one mafia receives the entire income in a monopoly. In terms of the model, no-cheating gives a payoff of $t + \frac{\delta}{1-\delta} i^*(r)t$ to a monopolist and only $\frac{1}{2}(t + \frac{\delta}{1-\delta} i^*(r)t)$ to a duopolist. Ceteris paribus, a reduction in the no-cheating payoff makes cheating relatively more profitable. On the other hand, however, the existence of a second mafia also reduces the cheating payoff. In a monopoly the mafia does not have a counterpart and thus, once production has taken place, it can always extort the entire surplus $y = 1$. In a duopoly, however, deviation by any one mafia always implies a conflict with the other mafia. Since the implemented payoff distribution is a function of both mafias' strategies, a deviating mafia does not receive the entire surplus $y = 1$ even if it engages in full extortion. This effect is reinforced if disagreement between the mafias is wasteful, i.e. if $\alpha < 1$. For both these reasons the deviation payoff in a duopoly is less than in a monopoly, i.e. $\frac{\alpha}{2}(1 + \frac{t}{2}) < 1$. Whether the no-cheating constraint is 'more binding' in the duopoly than in the monopoly then depends on the relative size of these two effects.

In the appendix we prove the following proposition.

Proposition 5 *For $\delta \in [0, \overline{\delta}^c(\alpha))$ the maximisation problem (7) does not have a solution. For $\delta \in [\underline{\delta}^c(\alpha), 1)$ the solution is given by*

$$\overline{t}^{c*} = \begin{cases} T & \text{if } \delta \in [\overline{\delta}^c(\alpha), 1) \\ \widetilde{T}^c(\delta, \alpha) & \text{if } \delta \in [\underline{\delta}^c(\alpha), \overline{\delta}^c(\alpha)), \end{cases}$$

where $\widetilde{T}^c(\delta, \alpha)$ is the smallest t that satisfies $NC^c(t, \delta) = 0$ and $\underline{\delta}^c(\alpha) \in [0, \overline{\delta}^c(\alpha)]$ is defined below.

Furthermore, it can be shown that

i. $\widetilde{T}^c(\overline{\delta}^c) = T$ if $\overline{\delta}^c > 0$.

ii. for $\delta \in (\underline{\delta}^c(\alpha), 1)$, $\widetilde{T}^c(\delta) \in (0, 1 - u_L)$ and

$$\frac{\partial \widetilde{T}^c(\delta)}{\partial \delta} < 0 \, .$$

Proof: see appendix.

Proposition 4 implies that, for $\delta \in [0, \underline{\delta}^c(\alpha))$, trigger strategy equilibria cannot be sustained. In this parameter range the players then simply play their stage game equilibrium strategies in every stage game. It is straightforward to verify that the following strategies and beliefs always form a perfect Bayesian equilibrium of the stage game: both mafias always engage in full extortion (i.e.

M_1 always plays $d_1 = (1, 0, 0, 0)$ and M_2 always plays $d_2 = (0, 1, 0, 0)$), the entrepreneur does not invest and never supplies labour, the worker never supplies labour, and the worker and mafias always believe that the entrepreneur invested $i = 0$. This equilibrium corresponds to the equilibrium of the stage game in the monopoly mafia case and we assume that, for $\delta \in [0, \underline{\delta}^c(\alpha))$, the players coordinate on this equilibrium. In the appendix (see proposition 13) we show, however, that for a certain parameter range there exist perfect Bayesian equilibria of the stage game in which the mafias do not engage in full extortion and the entrepreneur makes a non-zero investment.

For $\delta \in [\underline{\delta}^c(\alpha), 1)$ the profit maximising mafia cartel enforces contracts and only engages in limited extortion. Social surplus per stage game is then given by

$$V^c = \begin{cases} 0 & \text{if } \delta \in [0, \underline{\delta}^c(\alpha)) \\ i^c(1 - u_L) - c(i^c) & \text{if } \delta \in [\underline{\delta}^c(\alpha), 1), \end{cases} \tag{8}$$

where

$$i^c = i^*(1 - u_L - \bar{t}^{c*}).$$

2.5 A Monopoly Mafia versus a Mafia Cartel

In this section we analyse the welfare implications of moving from a mafia monopoly to a mafia cartel regime. In the discussion we make use of the following results.

Lemma 1 *It can be shown that*

i. $\bar{\delta}^c(0) = 0$.

ii. $\underline{\delta}^c(1) > \underline{\delta}$, $\bar{\delta}^c(1) > \bar{\delta}$, *and, for* $\delta \in [\underline{\delta}^c(1), 1)$, $\widetilde{T}^c(\delta, 1) > \widetilde{T}(\delta)$.

iii. $\underline{\delta}^c(\alpha)$, $\bar{\delta}^c(\alpha)$, *and, for* $\delta \in [\underline{\delta}^c(\alpha), 1)$, $\widetilde{T}^c(\delta, \alpha)$ *are increasing in* α.

Proof: see appendix.

Suppose that conflicts are not wasteful, i.e. $\alpha = 1$. Lemma 1, together with propositions 2 and 5, then has the following implications. First, a cartel is 'more likely' to engage in complete extortion than a monopolist: in both regimes the mafias engage in complete extortion if $\delta \in [0, \underline{\delta})$ but only the cartel does so if $\delta \in [\underline{\delta}, \underline{\delta}^c(1))$. Thus, $V^c = V^m$ if $\delta \in [0, \underline{\delta})$ and $V^m > V^c$ if $\delta \in [\underline{\delta}, \underline{\delta}^c(1))$. Second, for $\delta \in [\underline{\delta}^c(1), \overline{\delta}^c(1))$, contracts are enforced in both regimes but the mafia cartel engages in strictly more extortion than the monopoly mafia, i.e. $\overline{t}^{c*} > \overline{t}^*$. Since extortion is socially inefficient this implies that $V^m > V^c$ if $\delta \in [\underline{\delta}^c(1), \overline{\delta}^c(1))$. Finally, if $\delta \in [\overline{\delta}^c(1), 1)$, both the mafia cartel and the monopoly mafia extort T and enforce contracts, so that $V^m = V^c$. These arguments can be summarised in the following proposition.

Proposition 6 *When conflicts are not wasteful, i.e. $\alpha = 1$, social welfare in a monopoly mafia regime is weakly higher than in a mafia cartel regime. In particular,*

$$V^m - V^c \begin{cases} = 0 & \text{if } \delta \in [0, \underline{\delta}) \\ > 0 & \text{if } \delta \in [\underline{\delta}, \overline{\delta}^c(1)) \\ = 0 & \text{if } \delta \in [\overline{\delta}^c(1), 1). \end{cases}$$

While comparing American and Sicilian mafias, Gambetta (1993, p.42) notes that: "The greater restraint of the American cartel may well depend on its relative stability. Mafia families in the United States are also fewer in number and larger - five in New York as opposed to eighteen in Palermo alone - presumably decreasing the tension between firms [...]". Our analysis is consistent with this presumption in so far as it implies that, as long as conflicts are not wasteful, collusion between mafias is more difficult to sustain when there are more mafia families.

In the context of a discussion about the welfare effects of the different types of warlords who apparently controlled large parts of China in the 1920's Olson (1993) poses the following 'puzzle': "Why should warlords, who were stationary bandits continuously stealing from a given group of victims, be preferred, by those victims, to roving bandits who soon departed? The warlords had no claim to legitimacy and their thefts were distinguished from those of roving bandits only because they took the form of continuing taxation rather than occasional plunder"[11]. In terms of our analysis stationary bandits are preferred to roving bandits because they have a higher discount rate and therefore put more weight on the negative incentive effect of extortion. This leads to less extortion and makes the provision of contract enforcement services more likely. Even if they have the same discount rate, a single stationary bandit is preferred to a group of roving bandits if the conflict technology is not wasteful. Again, the reason for this is that a single bandit is more easily deterred from engaging in complete extortion than members of a group of bandits.

Suppose now that conflicts completely destructive, i.e. $\alpha = 0$. Lemma 1 and proposition 5 show that the mafia cartel then enforces contracts and extorts T for any $\delta \in [0, 1)$. We know from proposition 2 that a monopoly mafia enforces contracts if and only if $\delta \in [\underline{\delta}, 1)$, that it extorts $\widetilde{T} > T$ if $\delta \in [\underline{\delta}, \overline{\delta})$, and that it extorts T if $\delta \in [\overline{\delta}, 1)$. Since extortion is distortionary, it immediately follows that $V^m < V^c$ if $\delta \in [0, \overline{\delta})$ and $V^m = V^c$ if $\delta \in [\overline{\delta}, 1)$. This discussion can be summarised in the following proposition.

Proposition 7 *When conflicts are completely destructive, i.e.* $\alpha = 0$, *social welfare in a mafia cartel regime is weakly higher than in a monopoly mafia regime. In particular,*

$$V^m - V^c \begin{cases} < 0 & \text{if } \delta \in [0, \overline{\delta}) \\ = 0 & \text{if } \delta \in [\overline{\delta}, 1). \end{cases}$$

Finally, consider intermediate values of α. Lemma 1 implies that the parameter range in which a cartel enforces contracts is decreasing in α and that, for any δ, the level of extortion is weakly increasing in α. It therefore follows

[11] Olson (1993), p.568.

that, for any δ, V^c is (weakly) decreasing in α. Since V^m is independent of α (see (5)) this in turn implies that $V^m - V^c$ is weakly increasing in α. Thus, relative to a monopoly mafia, a mafia cartel becomes socially more efficient the more destructive the conflict technology.

2.6 Democracy

We now consider a regime in which the private agents use democratic elections to control coercive power by solving the game that was described in section 1.2. Recall that, in this game, any agent who controls coercive power in one stage game faces an election after having implemented a tax policy. We analyse to what extent elections can be used to discipline agents with coercive power. As in the previous sections we again focus on the equilibrium in which the equilibrium payoff of the 'long-run player', in this case is the electorate, is maximised.

At the election stage all the private agents are identical. In particular, they all have the same probability of being chosen as a worker, an entrepreneur, or a ruler in any subsequent stage game. Hence, when we study the voting strategy of the private agents we only need to be concerned with the behaviour of one representative voter, who we denote by 'R'. However, it is evident that, in our model, the representative voter expects a zero return from any voting strategy. This is due to the assumption that the population of private agents is infinitely large. We, therefore, assume that if the representative voter is indifferent between voting strategies, he uses the strategy that maximises the expected payoff of the electorate[12].

In a first best world the powerful agents enforce contracts and do not engage in extortion. In terms of the model this corresponds to implementing the distribution $\widetilde{d} = (0, 1 - u_L, u_L)$. To see whether the voters can induce an incumbent to implement $\widetilde{d}$, consider the following voting strategy: vote for the incumbent if she implemented d^* in every previous stage game and do not re-elect

[12] This is equivalent to assuming that, while the population may be very large, it is not quite infinite.

her otherwise. In every stage game the incumbent must then decide whether to implement $\widetilde{d}$ and be re-elected or to engage in full extortion and be thrown out of office. The former strategy gives the incumbent a zero payoff, while the latter gives her a payoff of $y = 1$ (if production has take place). Thus, faced with such a voting strategy, the incumbent never implements $\widetilde{d}$ and engages in full extortion whenever production has taken place.

The incumbent only has an incentive to refrain from full extortion if there are some gains from being re-elected. In this simple model these gains can only take the form of extortionary income. Consider, therefore, a voting strategy in which the voters 'allow' the incumbent to extort t as long as she also enforces contracts. Hence, they re-elect the incumbent if and only if she implemented $\widehat{d} = (t, 1 - u_L - t, u_L)$ during previous terms in office. In every stage game the incumbent must then choose between either implementing $\widehat{d}$ and being re-elected or engaging in full extortion and being thrown out of office. In a stage game in which production has taken place the incumbent prefers to implement $\widehat{d}$ if and only if $t + \frac{\delta}{1-\delta} i^*(1 - u_L - t)t \geq 1$, i.e. the no-cheating constraint is satisfied[13]. Recall that the smallest level of t that satisfies this constraint is given by $\widetilde{T}(\delta)$ and that $\widetilde{T}(\delta) < 1 - u_L$ for $\delta \in [\underline{\delta}, 1)$. Since the private agents prefer low levels of t, the best policy they can induce is therefore given by $d^{v*} = (\widetilde{T}(\delta), 1 - u_L - \widetilde{T}(\delta), u_L)$[14]. It was shown above that, for $\delta \in [0, \underline{\delta})$, there does not exist a $t \in [0, 1 - u_L]$ that satisfies the no-cheating constraint. Thus, in this parameter region, any incumbent engages in full extortion, independent of the private agents' voting strategy and no surplus is generated.

Before stating the strategies and beliefs that form a perfect Bayesian equilibrium in which the incumbent implements d^{v*} it is useful to introduce the following definition that allows us to rationalise on the exposition.

[13]Clearly, the incumbent never has an incentive to cheat when production did not take place.

[14]This result, that agents 'tolerate' limited extortion by the incumbent to prevent her from engaging in full extortion, is related to Klein and Leffler (1981). In their model short run consumers pay a premium above the competitive price to induce long run firms to produce high quality goods.

Definition 1 *In any stage game 'condition A' is said to be satisfied if and only if in any prior stage game an incumbent was removed from office if and only if she ever played* $d \neq d^{v*}$.

Corresponding to the notation introduced above we denote by $b_I(i,z)$ and $b_R(i,z)$ the probability that the incumbent and the representative voter attach to the entrepreneur having invested i given that they are in an information set with $z \in \{0,1\}$. We can now state the following proposition.

Proposition 8 *For* $\delta \in [\underline{\delta}, 1)$*, the following strategies and beliefs form a perfect Bayesian equilibrium:*

- σ_R^v: *In any stage game the representative voter votes against the incumbent if and only if she ever played* $d \neq d^{v*}$. *Otherwise the representative voter votes for the incumbent.*

- σ_I^v: *In the first stage game the incumbent plays* d^{v*}. *Any subsequent incumbent plays* d^{v*} *if condition A is satisfied and she plays* $\underline{d} = (1,0,0)$ *otherwise.*

- σ_E^v: *The first entrepreneur invests* $i^v = i^*(1 - u_L - \widetilde{T}(\delta))$. *Any subsequent entrepreneur invests* i^v *if and only if condition A is satisfied and she invests* $i = 0$ *otherwise. The entrepreneurs always play* $l_L = 0$.

- σ_L^v: *The first worker plays* $l_L = 1$ *if the investment was successful and* $l_L = 0$ *otherwise. Any subsequent worker plays* $l_L = 1$ *if and only if the investment was successful and condition A is satisfied and he plays* $l_L = 0$ *otherwise.*

- b^v: *If condition A is satisfied, then* $b_I(i^v,z) = b_L(i^v,z) = b_R(i^v,z) = 1$, $\forall z$. *If condition A is not satisfied, then the beliefs satisfy* $\int_0^1 b_M(i,z)di = \int_0^1 b_L(i,z)di = \int_0^1 b_R(i,z)di = 1$, $\forall z$.

Proof: see appendix.

Finally, consider social welfare in a democracy. For $\delta \in [\underline{\delta}, 1)$ the democratic ruler enforces contracts and only engages in limited extortion. For $\delta \in [0, \underline{\delta})$ even a democratic ruler engages in complete extortion. Social surplus per stage game is, therefore, given by

$$V^v = \begin{cases} 0 & \text{if } \delta \in [0, \underline{\delta}) \\ i^v(1 - u_L) - c(i^v) & \text{if } \delta \in [\underline{\delta}, 1). \end{cases}$$

2.7 Democracy versus a Mafia Monopoly or Duopoly

Consider first the welfare implications of moving from a monopoly mafia regime to a democracy. Figure 1 shows the payoff distributions that are implemented in the two regimes as a function of the discount rate. For $\delta \in [\underline{\delta}, 1)$ both the monopoly mafia and the democratic ruler enforce contracts. The former does so because it maximises its extortionary income and the latter because she would be thrown out of office if she did otherwise. The two regimes differ, however, with respect to the level of extortion. In particular, when $\delta \in [\overline{\delta}, 1)$ the democratic ruler extorts T while the monopoly mafia extorts $\widetilde{T}$. Since it was shown in proposition 2 that $\widetilde{T} > T$ for $\delta \in (\overline{\delta}, 1)$ it follows that, in this parameter range, social welfare is strictly higher under a democratic ruler than under a monopoly mafia, i.e. $V^v > V^m$. Figure 1 also shows that the level of extortion, and thus social welfare, is the same in both regimes if $\delta \in [\underline{\delta}, \overline{\delta}]$. Finally, figure 1 shows that, in both regimes, the powerful agent engages in complete extortion if he is very impatient, i.e. $\delta \in [0, \underline{\delta})$. Thus, in this parameter region, $V^v = V^m = 0$. To see why social welfare in a democracy can be strictly higher than under a monopoly mafia, consider the 'punishment devices' which the private agents can use in either regime to penalise extortion by the powerful agent. In the case of a monopoly mafia the private agents can respond to extortion only by reducing their private investment levels and hence the mafia's potential extortionary income. In a democracy elections provide an additional punishment device which can be used to discipline the ruler. The analysis above shows that this additional punishment device 'has bite', in the sense that it leads to a lower level of extortion, if and only if the powerful agents are patient, i.e. $\delta \in (\overline{\delta}, 1)$. The analysis can be summarised in the following proposition.

Proposition 9 *Social welfare in a democracy is weakly higher than in a monopoly mafia regime. In particular,*

$$V^v - V^m \begin{cases} = 0 & \text{if } \delta \in [0, \overline{\delta}] \\ > 0 & \text{if } \delta \in (\overline{\delta}, 1). \end{cases}$$

Consider now the welfare implications of moving from a mafia cartel regime to a democracy. We have seen above that the level of social welfare under a mafia cartel depends crucially on the technology parameter α. In particular, proposition 6 shows that for $\alpha = 1$ social welfare is weakly higher under a monopoly mafia than under a mafia cartel. Since we have just shown that a democracy is weakly more efficient than a monopoly mafia it follows that for $\alpha = 1$ a democracy is also weakly more efficient than a mafia cartel. The proposition below follows immediately from propositions 6 and 9.

Proposition 10 *When conflicts are not wasteful, i.e. $\alpha = 1$, social welfare in a democracy is weakly higher than in a mafia cartel regime. In particular,*

$$V^v - V^c \begin{cases} = 0 & \text{if } \delta \in [0, \underline{\delta}) \\ > 0 & \text{if } \delta \in [\underline{\delta}, 1). \end{cases}$$

We have argued above that, in the case of a mafia cartel, the level of welfare is weakly decreasing in α. To see whether a regime with a mafia cartel can ever be more efficient than a democracy, consider the case when conflicts are completely destructive, i.e. $\alpha = 0$. Figure 2 shows the payoff distributions that are implemented in both regimes as function of the discount rate. When $\delta \in (\overline{\delta}, 1)$, the cartel extorts T, the democratic ruler extorts $\widetilde{T}$, and both enforce contracts. Since $T > \widetilde{T}$ if $\delta \in (\overline{\delta}, 1)$ it follows that in this parameter range a democratic ruler is socially more efficient than the mafia cartel, i.e. $V^v > V^c$. Figure 2 shows that, for $\delta \in [0, \overline{\delta})$, a cartel engages in less extortion, and is more likely to enforce contracts, than a democratic ruler. Thus, in this parameter range, social welfare under a mafia cartel is strictly higher than in a democracy[15]. Finally, for $\delta = \overline{\delta}$ a mafia cartel and a democratic ruler generate the same level of social welfare since they both enforce contracts and extort T. The analysis can be summarised in the following proposition.

Proposition 11 *When conflicts are not wasteful, i.e. $\alpha = 0$, the difference in social welfare between a democracy and a mafia cartel regime is given by*

$$V^v - V^c \begin{cases} < 0 & if \ \delta \in [0, \overline{\delta}) \\ = 0 & if \ \delta = \overline{\delta} \\ > 0 & if \ \delta \in (\overline{\delta}, 1). \end{cases}$$

3 Related Literature

In this section we show how the existing literature relates to our analysis. Brennan and Buchanan (1977, 1978, 1980) propose a theory of the public sector in which the government aims to exploit the citizens through the maximisation of the tax revenues that it extracts from the economy. This 'Leviathan' hypothesis has received considerable attention in the literature. A key argument in their analysis is that, due to interjurisidictional mobility of citizens, competition between governments can limit excessive taxation[16]. In contrast, in our

[15] Our model focuses exclusively on the efficiency implications of distortionary extortion and contract enforcement. It is evident that one may prefer a democratic contract enforcement to enforcement by a mafia cartel for a number of additional reasons which we do not model here. One may, for instance, have equity concerns about the income distribution that results from private enforcement activities. Also, one may worry about the association of private enforcement organisations with acts of violence.

[16] See, for instance, Brennan and Buchanan (1980), p.184: "interjurisdictional mobility of persons in pursuit of 'fiscal gains' can offer partial or possibly complete substitutes for explicit fical constraints on the taxing power". This hypothesis has been tested in a number of paper, including Oates (1985) and Forbes and Zampelli (1989).

model citizens do not have the option of leaving the economy for one with a more favourable governmental system. Instead, the mafias are 'competing' in so far as each is trying to extract as much surplus as possible from the same fixed group of private agents. To our knowledge the only other paper studying this type of competition is Konrad and Skaperdas (1999). In particular, they study different government regimes in which powerful agents use their coercive power to extort citizens but also to protect them against bandits and competing lords. In one regime "peasants are tied to their land and at the mercy of the lords who compete over how to divide them up[17]". They stress the importance of this type of competition and state: "From Mesopotamia to China, Egypt, Mesoamerica, or feudal Europe, serfs were tied to the land and free peasants typically had no outside options, with rulers coming and going but without any change in their incentives for production. Even in the past two centuries, with the rise of the rights of man, the most liberal of states have sequestered their citizens with barbed-wire borders and passport controls[18]".

Olson (1993) argues informally that a self-interested dictator might provide public goods to increase his tax base. He also outlines arguments for the difference in economic development between dictatorships and democracies. The ideas presented in Olson (1993) are formalised in McGuire and Olson (1996). While our analysis is similar in spirit, the models differ in several dimensions: first, they discuss public goods in general, while we focus on the special public good of contract enforcement. We believe that this public good has particular characteristics which need to be model explicitly. Second, McGuire and Olson (1996) use a static model and do not address the reputation mechanisms which are central in our analysis. Third, they focus on autocracies and democracies while we allow for different government systems.

In recent years there have been a number of contributions that discuss private and public contract enforcement with special reference to Russia. Drawing on

[17] Konrad and Skaperdas (1999).
[18] Konrad and Skaperdas (1999), p.17.

examples from Russia, Hay, Shleifer, and Vishny (1996) describe why private agents may prefer using private methods of dispute resolution to relying on a "dysfunctional legal system". They build a very simple model to show that basic legal rules can be used to make a public legal system more attractive to private parties than contract enforcement by a mafia. However, they do not discuss the incentives of a state or a mafia to provide contract enforcement service. Hay and Shleifer (1998) describe some shortcomings of the Russian legal system and discuss the different ways in which private agents respond to the absence of public contract enforcement services. They state that "the principal argument of this paper is that the appropriate legal-reform strategy for a country like Russia is private enforcement of public rules[19]". In this context they discuss the importance of reputation development by private enforcers. However, they do not provide a formal analysis. Greif and Kandel (1995) and Pistor (1996) provide very interesting descriptive analysis of the Russian legal system and describe the response of various private parties to the lack of public contract enforcement.

Klein and Leffler (1981) use an infinitely repeated game with short- and long-run players to analyse how reputational mechanisms can induce firms to abide by contractual obligations in the absence of third-party enforcers. Kreps (1990) develops a similar framework to model reputation building and the role of firms. While the issues we address in this chapter are very different, the methods we use are similar to those developed in these papers.

De Long and Shleifer (1993) provide empirical evidence that autocratic governments are associated with low growth. Finally, Gambetta (1993) gives a very interesting account of the Sicilian Mafia.

---

[19] Hay and Shleifer (1998), p.400.

4 Conclusion

In this chapter we analysed the incentives of agents with coercive power to supply contract enforcement services. We have shown that when these agents are patient enough they do indeed have an incentive to use their powers to enforce privately optimal contracts. However, these powerful agents are not only guided by the invisible hand to provide contract enforcement services but also use their grabbing hands to enrich themselves. Indeed, the sole purpose of providing contract enforcement is to maximise the surplus from which they can extort their income. Whether a society is better off with or without powerful agents depends on the relative size of two opposing effects: on the one hand, they allow private agents to realise gains from trade but, on the other hand, they also engage in extortion and thereby distort private investment incentives. Which effect dominates depends on the degree of competition between powerful agents and on whether or not they face democratic elections.

5 Appendix

Before proving proposition 1 it is useful to consider the following lemmas.

Lemma 2 *It can be shown that*

$$\frac{\partial^2 NC(t,\delta)}{\partial t^2}\begin{cases} =0 & \textit{if } \delta=0 \\ <0 & \textit{if } \delta\in(0,1). \end{cases}$$

Proof:

Differentiating $NC(t,\delta)$ gives

$$\begin{aligned} \frac{\partial NC(t,\delta)}{\partial t} &= 1+\frac{\delta}{1-\delta}(i^*(1-u_L-t)-t\frac{\partial i^*(1-u_L-t)}{\partial r}) \\ \frac{\partial^2 NC(t,\delta)}{\partial t^2} &= \frac{\delta}{1-\delta}(-2\frac{\partial i^*(1-u_L-t)}{\partial r}+t\frac{\partial^2 i^*(1-u_L-t)}{\partial r^2}). \end{aligned}$$

To determine the sign of these derivatives, we need to consider the optimal investment level $i^*(r)$. The optimal investment level $i^*(r)$ is implicitly defined by the first order condition $r = c'(i)$ (the second order condition $c''(i) > 0$ is satisfied for all i). Applying the implicit function theorem to the first order condition gives $dr - c''(i)di = 0$. Hence,

$$\frac{\partial i^*}{\partial r}=\frac{1}{c''(i)}>0.$$

By differentiating again we can find

$$\frac{\partial^2 i^*}{\partial r^2}=-\frac{c'''(i^*)}{c''(i^*)^3}\leq 0.$$

It then follows that

$$\frac{\partial^2 NC(t,\delta)}{\partial t^2}\begin{cases} =0 & \text{if } \delta=0 \\ <0 & \text{if } \delta\in(0,1). \end{cases} \qquad \blacksquare$$

Lemma 3 *There exists a unique δ that solves*

$$\frac{\partial NC(1-u_L,\delta)}{\partial t}=0.$$

Let this δ be denoted by $\widetilde{\delta}$. It can be shown that $\widetilde{\delta}\in(0,1)$.

Proof:

From the proof of lemma 2 we know that

$$\frac{\partial NC(1-u_L,\delta)}{\partial t} = 1 - \frac{\delta}{1-\delta}(1-u_L)\frac{\partial i^*(0)}{\partial r}.$$

Since $\frac{\partial NC(1-u_L,0)}{\partial t} = 1$, $\lim_{\delta\to 1} \frac{\partial NC(1-u_L,\delta)}{\partial t} = -\infty$, and

$$\frac{\partial^2 NC(1-u_L,\delta)}{\partial t \partial \delta} = -\frac{1}{(1-\delta)^2}(1-u_L)\frac{\partial i^*(0)}{\partial r} < 0,$$

there must be a unique $\delta \in (0,1)$ that solves

$$\frac{\partial NC(1-u_L,\delta)}{\partial t} = 0. \qquad \blacksquare$$

Lemma 4 *Consider*

$$\max_{t\in[0,1-u_L]} NC(t,\delta).$$

It can be shown that this maximisation problem has a unique solution $\lambda(\delta)$. This solution has the following properties:

i. For $\delta \in [0,\widetilde{\delta})$, $\lambda(\delta) = 1-u_L$ and $\lambda'(\delta) = 0$.

ii. $\lambda(\widetilde{\delta}) = 1-u_L$, $\lambda'_-(\widetilde{\delta}) = 0$, and $\lambda'_+(\widetilde{\delta}) < 0$.

iii. For $\delta \in (\widetilde{\delta},1)$, $\lambda(\delta) \in (0,1-u_L)$ and $\lambda'(\delta) < 0$.

Proof:

Note that, for any $\delta \in [0,1)$, $\frac{\partial NC(0,\delta)}{\partial t} > 0$. Thus, $\lambda(\delta) > 0$, for any $\delta \in [0,1)$. It follows from the proof of lemma 3 that

$$\frac{\partial NC(1-u_L,\delta)}{\partial t} \begin{cases} > 0 & \text{if } \delta \in [0,\widetilde{\delta}) \\ = 0 & \text{if } \delta = \widetilde{\delta} \\ < 0 & \text{if } \delta \in (\widetilde{\delta},1). \end{cases}$$

Since $NC(t,\delta)$ is concave in t (see lemma 2) this implies that $\lambda(\delta) = 1-u_L$ for $\delta \in [0,\widetilde{\delta}]$. Also, for $\delta \in (\widetilde{\delta},1)$, this implies that $\lambda(\delta) < 1-u_L$ is implicitly defined by

(9) $$\frac{\partial NC(\lambda,\delta)}{\partial t} = 0.$$

Suppose $\delta \in (\widetilde{\delta},1)$. Totally differentiating $\frac{\partial NC(t,\delta)}{\partial t} = 0$ gives

$$\frac{\partial^2 NC(t,\delta)}{\partial t^2}dt + \frac{\partial^2 NC(t,\delta)}{\partial t \partial \delta}d\delta = 0.$$

We know from lemma 2 that $\frac{\partial^2 NC(t,\delta)}{\partial t^2} < 0$ for $\delta \in (0,1)$. Also,

$$\frac{\partial^2 NC(t,\delta)}{\partial t \partial \delta} = \frac{1}{(1-\delta)^2}(i^*(1-u_L-t) - t\frac{\partial i^*(1-u_L-t)}{\partial r}).$$

Since, for $\delta \in (\widetilde{\delta}, 1)$,

$$i^*(1-u_L-\lambda) - \lambda\frac{\partial i^*(1-u_L-\lambda)}{\partial r} = -\frac{1-\delta}{\delta} < 0,$$

this implies that $\frac{\partial^2 NC(\lambda,\delta)}{\partial t \partial \delta} < 0$. Thus, for $\delta \in (\widetilde{\delta}, 1)$,

$$\lambda'(\delta) = -\frac{\partial^2 NC(\lambda,\delta)}{\partial t \partial \delta}(\frac{\partial^2 NC(\lambda,\delta)}{\partial t^2})^{-1} < 0.$$

Suppose now $\delta \in [0, \widetilde{\delta})$. Since $\lambda(\delta) = 1 - u_L$ for any $\delta \in [0, \widetilde{\delta})$, it follows that $\lambda'(\delta) = 0$.

Finally, consider $\delta = \widetilde{\delta}$. Then,

$$\lambda'_+(\widetilde{\delta}) = -\frac{\partial^2 NC(1-u_L,\widetilde{\delta})}{\partial t \partial \delta}(\frac{\partial^2 NC(1-u_L,\widetilde{\delta})}{\partial t^2})^{-1} < 0$$

and

$$\lambda'_-(\widetilde{\delta}) = 0. \qquad \blacksquare$$

Proof of proposition 1:

We first prove that $\sigma_M(\overline{d})$, $\sigma_E(\overline{d})$, and $\sigma_L(\overline{d})$ are best responses for each other, and that the beliefs b are consistent with Bayes' rule 'where possible', if and only if the RC, NC, and IC constraints are satisfied:

$\sigma_M(\overline{d})$: Consider any stage game. If, prior to this stage game, M always played $\overline{d}$, then M's expected return of playing $\overline{d}$ is $y(\overline{t} + \frac{\delta}{1-\delta}i^*(\overline{r})\overline{t})$ while the best alternative, playing $\underline{d}$, gives a return of y. Thus, $\overline{d}$ is a best response if and only if the NC constraint is satisfied. If, prior to this stage game, M ever played any $d \neq \overline{d}$, then playing $\underline{d}$ is a best response since doing so gives a return of y, while playing $\overline{d}$ gives $y\overline{t} \leq y$.

$\sigma_E(\overline{d})$: Consider any stage game. If, prior to this stage game, M always played $\overline{d}$, then E expects a return of $\overline{r}$ if her investment is successful and zero otherwise. Hence, $i^*(\overline{r})$ is a best response. If, prior to this stage game, M ever played any $d \neq \overline{d}$, then E expects no returns from investing so that $i = 0$ is a best response. Note that it can never be optimal for E to play $l_E = 1$ since, first, it is costly for her to do so (she incurs a disutility of u_E) and, second, given the other strategies, the payoff she realises at s_4 is independent of l_E. Thus, $l_E = 0$ is always a best response.

$\sigma_L(\overline{d})$: Consider any successful stage game. If, prior to this stage game, M always played $\overline{d}$, then $l_L = 1$ gives an expected return of $\overline{w} - u_L$ and $l_L = 0$ gives zero. Hence, $l_L = 1$ is a best response if and only if the IC constraint is satisfied. If, prior to this stage game, M ever played any $d \neq \overline{d}$, then $l_L = 1$ gives a return of $-u_L$ and $l_L = 0$ gives a zero return, so that $l_L = 0$ is a best response. Clearly, in any unsuccessful stage game $l_L = 0$ is a best response since it gives a zero return while $l_L = 1$ gives $-u_L$.

$b(\overline{d})$: Given these strategies, E invests $i^*(\overline{r})$ if M always played $\overline{d}$. Thus, the beliefs $b_M(i^*(\overline{r}), z) = b_L(i^*(\overline{r}), z) = 1$, $\forall z$, are consistent with Bayes' rule, given the equilibrium strategies. A stage game prior to which M ever played any $d \neq \overline{d}$ is off the equilibrium path and M and L can have any beliefs that satisfy $\int_0^1 b_M(i, z)di = \int_0^1 b_L(i, z)di = 1$, $\forall z$. Note that, independent of their beliefs, it is always optimal for M and L to play $\underline{d}$ and $l_L = 0$ in such a stage game.

We now prove that the RC, NC, and IC constraints can be satisfied simultaneously if and only if $\delta \geq \underline{\delta}$, where $\underline{\delta} \in (0, 1)$ is a parameter that will be defined below.

For this purpose we need to investigate the NC constraint when the RC and IC constraints are satisfied with equality. From lemma 4 it follows that $NC(\lambda(0), 0) = -u_L$, $\lim_{\delta \to 1} NC(\lambda(\delta), \delta) = \infty$, and

$$\frac{\partial NC(\lambda, \delta)}{\partial \delta} = \begin{cases} 0 & \text{if } \delta \in [0, \widetilde{\delta}] \\ \frac{1}{(1-\delta)^2} i^*(1 - u_L - \lambda(\delta))\lambda(\delta) > 0 & \text{if } \delta \in (\widetilde{\delta}, 1). \end{cases}$$

Thus, there exists a unique $\delta \in (\widetilde{\delta}, 1)$ such that $NC(\lambda(\delta), \delta) = 0$. Let this critical δ be denoted by $\underline{\delta}$. Thus, for $\delta \in [0, \underline{\delta})$, there does not exist a $t \in [0, 1 - u_L]$ for which $NC(t, \delta) \geq 0$. As a result, in this parameter range, the RC, NC, and IC constraints cannot be satisfied simultaneously. In contrast, for $\delta \in [\underline{\delta}, 1)$, the RC, NC, and IC can be satisfied simultaneously. ■

Proof of proposition 2:

We first show that $\frac{\partial^2 \pi(1-u_L-t,t)}{\partial t^2} < 0$, for any $t \in [0, 1-u_L]$. Differentiating $\pi(1-u_L-t,t)$ gives

$$\begin{aligned}
\frac{\partial \pi(1-u_L-t,t)}{\partial t} &= \frac{1}{1-\delta}(i^*(1-u_L-t) - t\frac{\partial i^*(1-u_L-t)}{\partial r}) \\
\frac{\partial^2 \pi(1-u_L-t,t)}{\partial t^2} &= \frac{1}{1-\delta}(-2\frac{\partial i^*(1-u_L-t)}{\partial r} + t\frac{\partial^2 i^*(1-u_L-t)}{\partial r^2}).
\end{aligned}$$

Thus, $\frac{\partial^2 \pi(1-u_L-t,t)}{\partial t^2} < 0$, for any $t \in [0, 1-u_L]$.

Note next that $\frac{\partial \pi(1-u_L,0)}{\partial t} > 0$ and $\frac{\partial \pi(0,1-u_L)}{\partial t} < 0$. Hence, $T \in (0, 1-u_L)$ is uniquely defined by the first order condition $\frac{\partial \pi(1-u_L-T,T)}{\partial t} = 0$ and is therefore independent of δ.

Note also that $NC(T, 0) < 0$, $\lim_{\delta \to 1} NC(T, \delta) = \infty$, and

$$\frac{\partial NC(T,\delta)}{\partial \delta} = \frac{1}{(1-\delta)^2} i^*(1-u_L-T)T > 0, \quad \text{for any } \delta \in [0,1).$$

Thus, there exists a unique $\delta \in (0,1)$ which solves $NC(T,\delta) = 0$. Let this critical δ be denoted by $\overline{\delta}$. Then, $NC(T,\delta) \geq 0$ for $\delta \in [\overline{\delta}, 1)$. Thus, in this parameter range, $\overline{t}^* = T$.

We proceed to describe $\overline{t}^*$ for $\delta \in [0, \overline{\delta})$. To do so we first show that $\underline{\delta} < \overline{\delta}$. For this purpose note that $\frac{\partial NC(T,\delta)}{\partial t} = 1$ and $NC(\lambda(\overline{\delta}), \overline{\delta}) > NC(T, \overline{\delta}) = 0$. It was shown in the proof of proposition 1 that, for $\delta \in [\underline{\delta}, 1)$, $\frac{\partial NC(\lambda(\delta),\delta)}{\partial \delta} > 0$. Thus, $\underline{\delta} < \overline{\delta}$.

We know from the proof of proposition 1 that, for $\delta \in [0, \overline{\delta})$, there does not exist a $t \in [0, 1-u_L]$ that satisfies $NC(t, \delta) \geq 0$. Thus, in this parameter range, maximisation problem (4) does not have a solution.

Consider, then, $\bar{t}^*$ for $\delta \in [\underline{\delta}, \bar{\delta})$. Note that, for $\delta \in [\underline{\delta}, \bar{\delta})$, $NC(\lambda(\delta), \delta) \geq 0$ and $NC(T, \delta) < 0$. Recall that $\frac{\partial NC(T,\delta)}{\partial t} > 0$ and $\frac{\partial^2 NC(t,\delta)}{\partial t^2} < 0$. It follows that for any $\hat{t}$ that satisfies $NC(\hat{t}, \delta) \geq 0$ it must be that $\hat{t} > T$. Thus, in the given parameter region, the optimal extortion rate must be larger than T to satisfy the constraint $NC(t, \delta) \geq 0$. Note also that $\frac{\partial \pi(1-u_L-t,t)}{\partial t} < 0$ for any $t > T$. It follows that, for $\delta \in [\underline{\delta}, \bar{\delta})$, $\bar{t}^* = \widetilde{T}(\delta)$, where $\widetilde{T}(\delta)$ is the smallest value of $t \in [0, 1-u_L]$ that satisfies $NC(t, \delta) = 0$.

Since $NC(0, \delta) = -1$ it must be that, for $\delta \in [\underline{\delta}, 1)$, $\widetilde{T}(\delta) > 0$. Also, since, for $\delta \in [\underline{\delta}, 1)$, $\frac{\partial^2 NC(t,\delta)}{\partial t^2} < 0$ and $\frac{\partial NC(1-u_L,\delta)}{\partial t} < 0$, it must be that $\widetilde{T}(\delta) < 1 - u_L$. Moreover, it follows from $NC(T, \bar{\delta}) = 0$, $\frac{\partial NC(T,\delta)}{\partial t} > 0$, and $\frac{\partial^2 NC(t,\delta)}{\partial t^2} < 0$ that $\widetilde{T}(\bar{\delta}) = T$.

Finally, consider $\frac{\partial \widetilde{T}(\delta)}{\partial \delta}$ for $\delta \in (\underline{\delta}, 1)$. Totally differentiating $NC(t, \delta) = 0$ gives

$$\frac{\partial NC(t,\delta)}{\partial t} dt + \frac{\partial NC(t,\delta)}{\partial \delta} d\delta = 0.$$

Since, for $\delta \in (\underline{\delta}, 1)$,

$$\frac{\partial NC(\widetilde{T},\delta)}{\partial t} > 0$$

and

$$\frac{\partial NC(\widetilde{T},\delta)}{\partial \delta} > 0,$$

it follows that

$$\frac{\partial \widetilde{T}(\delta)}{\partial \delta} = -\frac{\partial NC(\widetilde{T},\delta)}{\partial \delta} \left(\frac{\partial NC(\widetilde{T},\delta)}{\partial t}\right)^{-1} < 0 \text{ for any } \delta \in (\underline{\delta}, 1). \qquad \blacksquare$$

Proposition 12 *In any perfect Bayesian equilibrium of the stage game in the monopoly mafia regime E, L, and M play the following strategies:*

- *E plays $i = 0$ and $l_E = 0$.*
- *L always plays $l_L = 0$.*
- *M always plays $d = (1, 0, 0)$.*

Proof:

At s_4 it is always optimal for M to play $d = (1, 0, 0)$, whether or not $y = 0$ or $y = 1$ and independent of his beliefs about E's investment level. Thus, at s_3 E and L anticipate total extortion and therefore always play $l_E = l_L = 0$, independent of the investment's success and L's belief about E's investment. Clearly, at s_2 E then optimally plays $i = 0$. If the investment is unsuccessful, L and M (rightly) believe that E played $i = 0$. At off the equilibrium path information sets we can specify arbitrary beliefs, for any of which it is optimal for L and M to play $l_L = 0$ and $d = (1, 0, 0)$. ■

Proof of proposition 3:

$V^m - V^a \leq 0$ for $\delta \in [0, \underline{\delta})$ follows immediately from (3) and (5).

Next consider (6) for $\delta \in [\underline{\delta}, 1)$. Note that the first term on the RHS is always weakly positive. We now show that the last two terms on the RHS are weakly positive if and only if $u_E - u_L \geq \bar{t}^*$. Let $F(i) = i(1 - u_L) - c(i)$ and let $i^{**} = \arg\max_i F(i)$. We then need to show that $F(i^m) - F(i^a) \geq 0$ if and only if $u_E - u_L \geq \bar{t}^*$. Note that $i^{**} \geq \max[i^m, i^a]$. Since $F'(i) > 0$ for any $i < i^{**}$ it then follows that $F(i^m) - F(i^a) \geq 0$ if and only if $i^m \geq i^a$. Note that $i^m \geq i^a$ if and only if $u_E - u_L \geq \bar{t}^*$. Hence, $F(i^m) - F(i^a) \geq 0$ if and only if $u_E - u_L \geq \bar{t}^*$. It immediately follows that $V^m - V^a \geq 0$ if $\delta \in [\underline{\delta}, 1)$ and $u_E - u_L \geq \bar{t}^*$, and $V^m - V^a \gtrless 0$ if $\delta \in [\underline{\delta}, 1)$ and $u_E - u_L < \bar{t}^*$.

We know from proposition 2 that, for $\delta \in [\underline{\delta}, 1)$, $\bar{t}^*$ is weakly decreasing in δ. This implies that, in this parameter range, V^m is weakly increasing in δ. Since V^a is independent of δ, it then follows that, in this parameter range, $V^m - V^a$ is weakly increasing in δ.

It follows from proposition 2 that $\bar{t}^*$ is independent of u_E. We now analyse how $\bar{t}^*$ depends on u_L. For $\delta \in [\bar{\delta}, 1)$, $\bar{t}^* = T$, where T is implicitly defined by the first order condition

$$i^*(1 - u_L - T) - T\frac{\partial i^*(1 - u_L - T)}{\partial r} = 0.$$

Implicitly differentiating then gives

$$\frac{\partial T}{\partial u_L} = -\frac{-\frac{\partial i^*(1-u_L-T)}{\partial r} + T\frac{\partial^2 i^*(1-u_L-T)}{\partial r^2}}{-2\frac{\partial i^*(1-u_L-T)}{\partial r} + T\frac{\partial^2 i^*(1-u_L-T)}{\partial r^2}},$$

so that

$$-1 < \frac{\partial T}{\partial u_L} < 0. \tag{10}$$

For $\delta \in [\underline{\delta}, \overline{\delta})$, $\overline{t}^* = \widetilde{T}(\delta)$, where $\widetilde{T}(\delta)$ is defined by the unique t that solves the conditions $NC(t,\delta) = 0$ and $\frac{\partial NC(t,\delta)}{\partial t} \geq 0$. Implicitly differentiating the first condition gives

$$\frac{\partial \widetilde{T}(\delta)}{\partial u_L} = \frac{\delta}{1-\delta}\widetilde{T}\frac{\partial i^*(1-u_L-\widetilde{T})}{\partial r}\left(\frac{\partial NC(\widetilde{T},\delta)}{\partial t}\right)^{-1} \geq 0. \tag{11}$$

Differentiating (6) and substituting the first order conditions $c'(i^*(1-u_E)) = 1-u_E$ and $c'(i^*(1-u_L-\overline{t}^*)) = 1-u_L-\overline{t}^*$ gives

$$\frac{\partial(V^m - V^a)}{\partial u_E} = i^*(1-u_E) \geq 0$$

and

$$\frac{\partial(V^m - V^a)}{\partial u_L} = -(i^*(1-u_L-\overline{t}^*) + \overline{t}^*(1+\frac{\partial \overline{t}^*}{\partial u_L})\frac{\partial i^*(1-u_L-\overline{t}^*)}{\partial r}).$$

Because of (10) and (11) it follows that $\frac{\partial(V^m-V^a)}{\partial u_L} < 0$. Thus, for $\delta \in [\underline{\delta}, 1)$, $V^m - V^a$ is increasing in $u_E - u_L$. ■

Before proving proposition 4 it is useful to consider the following lemmas. In general, the proofs for the cartel and the monopoly mafia regimes are very similar.

Lemma 5 *It can be shown that*

$$\frac{\partial^2 NC^c(t,\delta)}{\partial t^2}\begin{cases} = 0 & \text{if } \delta = 0 \\ < 0 & \text{if } \delta \in (0,1). \end{cases}$$

Proof:

Differentiating $NC^c(t, \delta)$ gives

$$\begin{aligned}
\frac{\partial NC^c(t,\delta)}{\partial t} &= 1 - \frac{\alpha}{2} + \frac{\delta}{1-\delta}(i^*(1-u_L-t) - t\frac{\partial i^*(1-u_L-t)}{\partial r}) \\
\frac{\partial^2 NC^c(t,\delta)}{\partial t^2} &= \frac{\delta}{1-\delta}(-2\frac{\partial i^*(1-u_L-t)}{\partial r} + t\frac{\partial^2 i^*(1-u_L-t)}{\partial r^2}).
\end{aligned}$$

Since, as was shown in lemma 2, $\frac{\partial i^*(1-u_L-t)}{\partial r} > 0$ and $\frac{\partial^2 i^*(1-u_L-t)}{\partial r^2} \leq 0$, it follows that

$$\frac{\partial^2 NC^c(t,\delta)}{\partial t^2} \begin{cases} = 0 & \text{if } \delta = 0 \\ < 0 & \text{if } \delta \in (0,1). \end{cases} \quad \blacksquare$$

Lemma 6 *There exists a unique δ that solves*

$$\frac{\partial NC^c(1-u_L,\delta)}{\partial t} = 0.$$

Let this δ be denoted by $\widetilde{\delta}^c$. It can be shown that $\widetilde{\delta}^c \in (0,1)$.

Proof:

From the proof of lemma 5 we know that

$$\frac{\partial NC^c(1-u_L,\delta)}{\partial t} = 1 - \frac{\alpha}{2} - \frac{\delta}{1-\delta}(1-u_L)\frac{\partial i^*(0)}{\partial r}.$$

Since $\frac{\partial NC^c(1-u_L,0)}{\partial t} = 1 - \frac{\alpha}{2} > 0$, $\lim_{\delta\to1} \frac{\partial NC^c(1-u_L,\delta)}{\partial t} = -\infty$, and

$$\frac{\partial^2 NC^c(1-u_L,\delta)}{\partial t \partial \delta} = -\frac{1}{(1-\delta)^2}(1-u_L)\frac{\partial i^*(0)}{\partial r} < 0 \text{ for any } \delta \in [0,1),$$

there must be a unique $\delta \in (0,1)$ that solves

$$\frac{\partial NC^c(1-u_L,\delta)}{\partial t} = 0. \quad \blacksquare$$

Lemma 7 *Consider*

$$\max_{t\in[0,1-u_L]} NC^c(t,\delta).$$

It can be shown that this maximisation problem has a unique solution $\lambda^c(\delta)$. *This solution has the following properties:*

i. For $\delta \in [0,\widetilde{\delta}^c)$, $\lambda^c(\delta) = 1-u_L$ *and* $\lambda^{c\prime}(\delta) = 0$.

ii. $\lambda^c(\widetilde{\delta}^c) = 1-u_L$, $\lambda^{c\prime}_-(\widetilde{\delta}^c) = 0$, *and* $\lambda^{c\prime}_+(\widetilde{\delta}^c) < 0$.

iii. For $\delta \in (\widetilde{\delta}^c, 1)$, $\lambda^c(\delta) \in (0, 1-u_L)$ *and* $\lambda^{c\prime}(\delta) < 0$.

Proof:

Note that $\frac{\partial NC^c(0,\delta)}{\partial t} > 0$. Thus, $\lambda^c(\delta) > 0$, for any $\delta \in [0,1)$.

It follows from the proof of lemma 6 that

$$\frac{\partial NC^c(1-u_L,\delta)}{\partial t}\begin{cases} >0 & \text{if } \delta \in [0,\widetilde{\delta}^c) \\ =0 & \text{if } \delta = \widetilde{\delta}^c \\ <0 & \text{if } \delta \in (\widetilde{\delta}^c, 1). \end{cases}$$

Since $NC^c(t,\delta)$ is concave in t (see lemma 5) this implies that $\lambda^c(\delta) = 1-u_L$ for $\delta \in [0,\widetilde{\delta}]$. Also, for $\delta \in (\widetilde{\delta}^c, 1)$, this implies that $\lambda^c(\delta) < 1-u_L$ is implicitly defined by

(12)
$$\frac{\partial NC^c(\lambda^c,\delta)}{\partial t} = 0.$$

Suppose $\delta \in (\widetilde{\delta}^c, 1)$. Totally differentiating $\frac{\partial NC^c(t,\delta)}{\partial t} = 0$ gives

$$\frac{\partial^2 NC^c(t,\delta)}{\partial t^2}dt + \frac{\partial^2 NC^c(t,\delta)}{\partial t\partial\delta}d\delta = 0.$$

We know from lemma 5 that $\frac{\partial^2 NC(t,\delta)}{\partial t^2} < 0$ for $\delta \in (0,1)$. Also,

$$\frac{\partial^2 NC^c(t,\delta)}{\partial t\partial\delta} = \frac{1}{(1-\delta)^2}(i^*(1-u_L-t) - t\frac{\partial i^*(1-u_L-t)}{\partial r}).$$

Since, for $\delta \in (\widetilde{\delta}^c, 1)$,

$$i^*(1-u_L-\lambda^c) - \lambda^c\frac{\partial i^*(1-u_L-\lambda^c)}{\partial r} = -\frac{1-\delta}{\delta}(1-\frac{\alpha}{2}) < 0,$$

this implies that $\frac{\partial^2 NC^c(\lambda^c,\delta)}{\partial t\partial\delta} < 0$. Thus, for $\delta \in (\widetilde{\delta}^c, 1)$,

$$\lambda^{c\prime}(\delta) = -\frac{\partial^2 NC^c(\lambda^c,\delta)}{\partial t\partial\delta}(\frac{\partial^2 NC^c(\lambda^c,\delta)}{\partial t^2})^{-1} < 0.$$

Suppose now $\delta \in [0, \widetilde{\delta}^c)$. Since $\lambda^c(\delta) = 1 - u_L$ for any $\delta \in [0, \widetilde{\delta}^c)$, it follows that $\lambda^{c\prime}(\delta) = 0$.

Finally, consider $\delta = \widetilde{\delta}^c$. Then,

$$\lambda^{c\prime}_{+}(\widetilde{\delta}^c) = -\frac{\partial^2 NC^c(1-u_L, \widetilde{\delta}^c)}{\partial t \partial \delta}\left(\frac{\partial^2 NC^c(1-u_L, \widetilde{\delta}^c)}{\partial t^2}\right)^{-1} < 0$$

and

$$\lambda^{c\prime}_{-}(\widetilde{\delta}^c) = 0. \quad \blacksquare$$

Proof of proposition 4:

We first verify that the strategies $\sigma^c_M(\overline{d}^c)$, $\sigma^c_E(\overline{d}^c)$, and $\sigma^c_L(\overline{d}^c)$ are best responses for each other, and that the beliefs b^c are consistent with Bayes' rule 'where possible', if and only if the RC^c, NC^c, and IC^c constraints are satisfied.

$\sigma^c_{M_m}(\overline{d}^c)$, for $m = 1, 2$: Since the mafias are symmetric we only need to consider the best response by one mafia. Consider any stage game. If, prior to this stage game, it never happened that $d_1 \neq \overline{d}^c$ or $d_2 \neq \overline{d}^c$, then M_1 expects a return of $y(\frac{\overline{t}^c}{2} + \frac{\delta}{1-\delta} i^*(\overline{r}^c)\frac{\overline{t}^c}{2})$ by playing $\overline{d}^c$ and $\frac{\alpha}{2} y(1 + \frac{\overline{t}^c}{2})$ by playing $\underline{d}^c_1$. Hence, $\overline{d}^c$ is a best response if and only if the NC^c constraint is satisfied. If, prior to this stage game, it ever happened that $d_1 \neq \overline{d}^c$ or $d_2 \neq \overline{d}^c$, then playing $\underline{d}^c_1$ is a best response for M$_1$ since doing so gives a return of $\frac{\alpha}{2} y$ while playing $\overline{d}^c$ gives $\frac{\alpha}{4}\overline{t}^c y \leq \frac{\alpha}{2} y$.

$\sigma^c_E(\overline{d}^c)$: Consider any stage game. If, prior to this stage game, it never happened that $d_1 \neq \overline{d}^c$ or $d_2 \neq \overline{d}^c$, then E expects a return of $\overline{r}^c$ if her investment is successful and zero otherwise. Thus, $i^*(\overline{r}^c)$ is a best response. If, prior to this stage game, it was ever the case that $d_1 \neq \overline{d}^c$ or $d_2 \neq \overline{d}^c$, then E does not expect any investment returns and $i = 0$ is a best response. Note that it can never be optimal for E to play $l_E = 1$ since, first, it is costly for her to do so (she incurs a disutility of u_E) and, second, given the other strategies, the payoff she realises at s_4 is independent of l_E. Thus $l_E = 0$ is always a best response.

$\sigma^c_L(\overline{d}^c)$: Consider any successful stage game. If, prior to this stage game, it was never the case that $d_1 \neq \overline{d}^c$ or $d_2 \neq \overline{d}^c$, then $l_L = 1$ gives an expected return of $\overline{w}^c - u_L$ and $l_L = 0$ gives zero. Hence, $l_L = 1$ is a best response if and only if the IC^c constraint is satisfied. If, prior to this stage game, it was ever the case that $d_1 \neq \overline{d}^c$ or $d_2 \neq \overline{d}^c$, then $l_L = 1$ gives a return of $-u_L$ and $l_L = 0$ gives a zero return, so that $l_L = 0$ is a best response. Clearly, in any unsuccessful stage game $l_L = 0$ is a best response since it gives a zero return while $l_L = 1$ gives $-u_L$.

$b^c(\overline{d}^c)$: Given these strategies, E invests $i^*(\overline{r}^c)$ if and only if it was never the case that $d_1 \neq \overline{d}^c$ or $d_2 \neq \overline{d}^c$. Thus, the beliefs $b_{M_m}(i^*(\overline{r}^c), z) = b_L(i^*(\overline{r}^c), z) = 1$, $\forall z, m$, are consistent with Bayes' rule, given the equilibrium strategies. A stage game, prior to which it was ever the case that $d_1 \neq \overline{d}^c$ or $d_2 \neq \overline{d}^c$, is off the equilibrium path and M$_1$, M$_2$, and L can have any beliefs that satisfy $\int_0^1 b_{M_m}(i, z)di = \int_0^1 b_L(i, z)di = 1$, $\forall z, m$. Note that, independent of their beliefs, it is always optimal for M$_1$, M$_2$, and L to play $\underline{d}_1^c$, $\underline{d}_2^c$, and $l_L = 0$ in such a stage game.

We now prove that the RC^c, NC^c, and IC^c constraints can be satisfied simultaneously if and only if $\delta \geq \underline{\delta}^c$, where $\underline{\delta}^c \in [0, 1)$ is a parameter that will be defined below. For this purpose we need to investigate the NC^c constraint when the RC^c and IC^c constraints are satisfied with equality. From lemma 7 it follows that $NC^c(\lambda^c(0), 0) = -u_L$, $\lim_{\delta \to 1} NC^c(\lambda^c(\delta), \delta) = \infty$, and

$$\frac{\partial NC^c(\lambda^c, \delta)}{\partial \delta} = \begin{cases} 0 & \text{if } \delta \in [0, \widetilde{\delta}^c] \\ \frac{1}{(1-\delta)^2} i^*(1 - u_L - \lambda^c(\delta))\lambda^c(\delta) > 0 & \text{if } \delta \in (\widetilde{\delta}^c, 1). \end{cases}$$

Thus, there exists a unique $\delta \in (\widetilde{\delta}^c, 1)$ such that $NC^c(\lambda^c(\delta), \delta) = 0$. Let this critical δ be denoted by $\underline{\delta}^c$. Thus, for $\delta \in [0, \underline{\delta}^c)$, there does not exist a $t \in [0, 1 - u_L]$ for which $NC^c(t, \delta) \geq 0$. As a result, in this parameter range, the RC^c, NC^c, and IC^c constraints cannot be satisfied simultaneously. In contrast, for $\delta \in [\underline{\delta}^c, 1)$, the RC^c, NC^c, and IC^c can be satisfied simultaneously. ■

Proof of proposition 5:

Note that $\lim_{\delta \to 1} NC^c(T, \delta) = \infty$ and that

$$\frac{\partial NC^c(T, \delta)}{\partial \delta} = \frac{1}{(1-\delta)^2} i^*(1 - u_L - T)T \geq 0, \text{ for any } \delta \in [0, 1).$$

Suppose $NC^c(T,0) \leq 0$. Then there exists a unique $\delta \in [0,1)$ that solves $NC^c(T,\delta) = 0$. Let $\overline{\delta}^c(\alpha)$ be implicitly defined by $NC^c(T,\overline{\delta}^c) = 0$ if $NC^c(T,0) \leq 0$ and let $\overline{\delta}^c(\alpha) = 0$ otherwise. Then T satisfies the NC^c constraint if and only if $\delta \in [\overline{\delta}^c(\alpha), 1)$. Since it was shown above that T also satisfies the RC^c constraint it follows that $\overline{t}^{c*} = T$ for any $\delta \in [\overline{\delta}^c(\alpha), 1)$.

We proceed to describe $\overline{t}^{c*}$ for $\delta \in [0, \overline{\delta}^c)$. To do so we first show that $\underline{\delta}^c \leq \overline{\delta}^c$. For this purpose note that $\frac{\partial NC^c(T,\delta)}{\partial t} = 1 - \frac{\alpha}{2}$ and $NC^c(\lambda^c(\overline{\delta}^c), \overline{\delta}^c) > NC(T, \overline{\delta}^c) = 0$. It was shown in the proof of proposition 4 that, for $\delta \in [\underline{\delta}^c, 1)$, $\frac{\partial NC^c(\lambda^c(\delta),\delta)}{\partial \delta} > 0$. Thus, $\underline{\delta}^c \leq \overline{\delta}^c$.

We know from the proof of proposition 4 that, for $\delta \in [0, \underline{\delta}^c)$, there does not exist a $t \in [0, 1-u_L]$ that satisfies $NC^c(t,\delta) \geq 0$. Thus, in this parameter region, the maximisation problem (7) does not have a solution.

Consider, then, $\overline{t}^{c*}$ for $\delta \in [\underline{\delta}^c(\alpha), \overline{\delta}^c(\alpha))$. Note that, for $\delta \in [\underline{\delta}^c(\alpha), \overline{\delta}^c(\alpha))$, $NC^c(T,\delta) < 0$ and $NC^c(\lambda^c(\delta),\delta) \geq 0$. Recall that $\frac{\partial NC^c(T,\delta)}{t} > 0$ and $\frac{\partial^2 NC^c(T,\delta)}{t^2} < 0$. It follows that for any $\widehat{t}$ that satisfies $NC^c(\widehat{t},\delta) \geq 0$ it must be that $\widehat{t} > T$. Hence, in the given parameter region the optimal extortion rate must be larger than T to satisfy the constraint $NC^c(t,\delta) \geq 0$. Note also that $\frac{\partial \pi(1-u_L-t,t)}{\partial t} < 0$ for any $t > T$. It follows that, for $\delta \in [\underline{\delta}^c(\alpha), \overline{\delta}^c(\alpha))$, $\overline{t}^{c*} = \widetilde{T}^c(\delta)$, where $\widetilde{T}^c(\delta)$ is the smallest value of t that satisfies $NC^c(t,\delta) = 0$.

Since $NC^c(0,\delta) = -\alpha \geq 0$ it must be that, for $\delta \in [\underline{\delta}^c, 1)$, $\widetilde{T}^c(\delta) \geq 0$. Also, since, for $\delta \in [\underline{\delta}^c, 1)$, $\frac{\partial^2 NC^c(t,\delta)}{\partial t^2} < 0$ and $\frac{\partial NC^c(1-u_L,\delta)}{\partial t} < 0$, it must be that $\widetilde{T}^c(\delta) < 1 - u_L$. Moreover, since, for $\overline{\delta}^c > 0$, $NC^c(T, \overline{\delta}^c) = 0$, $\frac{\partial NC^c(T,\delta)}{\partial t} > 0$, and $\frac{\partial^2 NC^c(t,\delta)}{\partial t^2} < 0$, that $\widetilde{T}^c(\overline{\delta}^c) = T$ if $\overline{\delta}^c > 0$.

Finally, consider $\frac{\partial \widetilde{T}^c(\delta)}{\partial \delta}$ for $\delta \in (\underline{\delta}^c, 1)$. Totally differentiating $NC^c(t,\delta) = 0$ gives

$$\frac{\partial NC^c(t,\delta)}{\partial t} dt + \frac{\partial NC^c(t,\delta)}{\partial \delta} d\delta = 0.$$

Since, for $\delta \in (\underline{\delta}^c, 1)$,

$$\frac{\partial NC^c(\widetilde{T}^c,\delta)}{\partial t} > 0$$

and

$$\frac{\partial NC^c(\widetilde{T}^c, \delta)}{\partial \delta} > 0,$$

it follows that

$$\frac{\partial \widetilde{T}^c(\delta)}{\partial \delta} = -\frac{\partial NC^c(\widetilde{T}^c, \delta)}{\partial \delta}\left(\frac{\partial NC^c(\widetilde{T}^c, \delta)}{\partial t}\right)^{-1} < 0, \quad \text{for } \delta \in (\underline{\delta}^c, 1). \quad \blacksquare$$

Proposition 13 *In the mafia cartel regime there exist perfect Bayesian equilibria of the stage game in which the entrepreneur invests $i > 0$ if and only if $\alpha < \frac{2(1-u_L)}{3-u_L}$.*

Proof:

Suppose M_1 plays $d_1 = \widetilde{d} = (t, 1-t-r-u_L, r, u_L)$ when $y = 1$. It is a best response for M_2 to play $d_2 = \widetilde{d}$ if and only if

$$1 - t - r - u_L \geq \frac{\alpha}{2-\alpha}. \tag{13}$$

Now suppose $y = 1$ and M_2 plays $d_2 = \widetilde{d}$. It is a best response for M_1 to play $\widetilde{d}$ if and only if

$$t \geq \frac{\alpha}{2-\alpha}. \tag{14}$$

Conditions (13) and (14) can be satisfied simultaneously if and only if $\alpha \leq \frac{2(1-r-u_L)}{3-r-u_L}$. If this condition is not satisfied, then the mafias always engage in complete extortion so that in equilibrium it cannot be optimal for E to make non-zero investments. If this condition is satisfied, then it can be a best response for L to play $l_L = 1$ if a machine has been developed (and $l_L = 0$ otherwise). At s_2 it is then optimal for E to invest $i^*(r) > 0$ if and only if $r > 0$. Thus, for $\alpha < \frac{2(1-u_L)}{3-u_L}$ there exist perfect Bayesian equilibria for which $i > 0$. $\blacksquare$

Proof of lemma 1:

i. Note that $NC^c(T, 0) = T \geq 0$ for $\alpha = 0$. It follows from the definition of $\overline{\delta}^c(\alpha)$ that $\overline{\delta}^c(0) = 0$.

ii. Let $\alpha = 1$. Then $NC^c(T,0) < 0$, so that $\overline{\delta}^c(1)$ is implicitly defined by $NC^c(T, \overline{\delta}^c) = 0$ and $\overline{\delta}^c(1) > 0$.

Note that, for $\alpha = 1$, $NC(t,\delta) = NC^c(t,\delta) + \frac{t}{2}$. Thus, $NC^c(T, \overline{\delta}) = -\frac{T}{2}$. Since $\frac{\partial NC^c(T,\overline{\delta})}{\partial \delta} > 0$, it follows that $\overline{\delta}^c(1) > \overline{\delta}$.

Also, since $NC(t,\delta) = NC^c(t,\delta) + \frac{t}{2}$, $NC(\lambda,\delta) > NC^c(\lambda^c,\delta)$ for any $\delta \in [0,1)$. In particular, $NC(\lambda, \underline{\delta}^c) > NC^c(\lambda^c, \underline{\delta}^c) = 0$. Since, for $\delta \geq \underline{\delta}$, $\frac{\partial NC(\lambda,\delta)}{\partial \delta} > 0$, it follows that $\underline{\delta} < \underline{\delta}^c(1)$. Also, for $\delta \geq \underline{\delta}^c$, $NC(\widetilde{T}^c(\delta),\delta) > NC^c(\widetilde{T}^c(\delta),\delta) = 0$. Since $\frac{\partial NC(\widetilde{T}^c(\delta),\delta)}{\partial t} > 0$ it must be that, for $\delta \in [\underline{\delta}^c(1), 1)$, $\widetilde{T}(\delta) < \widetilde{T}^c(\delta)$.

iii. Recall the definition of $\overline{\delta}^c$: if $NC^c(T,0) \leq 0$, then $\overline{\delta}^c$ is implicitly defined by $NC^c(T,\overline{\delta}^c) = 0$. If $NC^c(T,0) > 0$, then $\overline{\delta}^c = 0$.

Note now that
$$\frac{\partial NC^c(T,0)}{\partial \alpha} = -(1 + \frac{T}{2}).$$
Totally differentiating $NC^c(T,\delta) = 0$ gives
$$\frac{\partial \delta}{\partial \alpha} = (1-\delta)^2(1+\frac{T}{2})(i^*(1-u_L-T)T)^{-1} > 0.$$
Hence, $\overline{\delta}^c$ is weakly increasing in α.

Next recall the definition of $\underline{\delta}^c$: if $NC^c(\lambda^c(0),0) \leq 0$, then $\underline{\delta}^c$ is implicitly defined by $NC^c(\lambda^c(\underline{\delta}^c), \underline{\delta}^c) = 0$. If $NC^c(\lambda^c(0),0) > 0$, then $\underline{\delta}^c = 0$.

Note that
$$\frac{\partial NC^c(\lambda^c(0),0)}{\partial \alpha} = -(1 + \frac{\lambda^c(0)}{2}) < 0.$$
Totally differentiating $NC^c(\lambda^c(\delta),\delta) = 0$ gives
$$\frac{\partial \overline{\delta}^c}{\partial \alpha} = (1-\delta)^2(1+\frac{\lambda^c}{2})(i^*(1-u_L-\lambda^c)\lambda^c)^{-1} \geq 0.$$
Hence, $\underline{\delta}^c$ is weakly increasing in α.

Finally, recall that, for $\delta \in [\underline{\delta}^c, 1)$, $\widetilde{T}^c$ is defined as the smallest t for which $NC^c(t,\delta) = 0$. Totally differentiating $NC^c(t,\delta) = 0$ gives
$$\frac{\partial NC^c(t,\delta)}{\partial t}dt + \frac{\partial NC^c(t,\delta)}{\partial \alpha}d\alpha = 0$$
Thus,
$$\frac{\partial \widetilde{T}^c}{\partial \alpha} = (1 + \frac{\widetilde{T}^c}{2})(\frac{\partial NC(\widetilde{T}^c,\delta)}{\partial t})^{-1} > 0. \qquad \blacksquare$$

Proof of proposition 8:

We first verify that the strategies σ_L^v, σ_E^v, σ_I^v, and σ_R^v are best responses for each other.

σ_L^v: Consider any successful stage game. If condition A is satisfied, then $l_L = 1$ and $l_L = 0$ both give zero returns so that $l_L = 1$ is a weak best response. If condition A is not satisfied, then $l_L = 1$ gives a return of $-u_L$ and $l_L = 0$ gives a zero return, so that $l_L = 0$ is a best response. Clearly, in any unsuccessful stage game $l_L = 0$ is a best response since it gives a zero return while $l_L = 1$ gives $-u_L$.

σ_E^v : Consider any stage game. If condition A is satisfied, then E expects a return of $1 - u_L - \widetilde{T}(\delta)$ if her investment is successful and zero otherwise. Hence, $i^*(1 - u_L - \widetilde{T}(\delta))$ is a best response. If condition A is not satisfied, then E expects no returns from investing so that $i = 0$ is a best response. Note that it can never be optimal for E to play $l_E = 1$ since, first, it is costly for her to do so (she incurs a disutility of u_E) and, second, given the other strategies, the payoff she realises at s_4 is independent of l_E. Thus $l_E = 0$ is always a best response.

σ_I^v : Consider any stage game. If condition A is satisfied, then d^{v*} and the best alternative, playing $\underline{d}$, both give an expected return of y, so that d^{v*} is a best response. If condition A is not satisfied, then playing d^{v*} gives a return of $\widetilde{T}(\delta)y < y$ while playing $\underline{d}$ gives a return of y so that $\underline{d}$ is a best response.

σ_R^v : Consider any stage game. If condition A is satisfied, then re-electing an I who always played d^{v*} gives an expected return of $\frac{1}{1-\delta}(i^v(1 - u_L - \widetilde{T}(\delta)) - c(i^v)) \geq 0$ while voting against her gives a zero return. Also, re-electing an I who ever played any $d \neq d^{v*}$ gives a zero return while voting against her gives $\frac{1}{1-\delta}(i^v(1 - u_L - \widetilde{T}(\delta)) - c(i^v)) \geq 0$. Thus, if condition A is satisfied, it is optimal to re-elect I if and only if she never played any $d \neq d^{v*}$. Finally, note that if condition A is not satisfied, then the private agents are indifferent between all voting strategies since they all lead to a zero return.

b^v : Given these strategies, E invests i^v in any stage game in which condition A is satisfied. Thus, $b_I(i^v, z) = b_L(i^v, z) = b_R(i^v, z) = 1$, $\forall z$, are consistent with Bayes' rule, given the equilibrium strategies. A stage game in which condition A is not satisfied is off the equilibrium path and I, L, and R can have any beliefs that satisfy $\int_0^1 b_M(i, z)di = \int_0^1 b_L(i, z)di = \int_0^1 b_R(i, z)di = 1$, $\forall z$. Note that, independent of their beliefs, it is always optimal for I and L to play $\underline{d}$ and $l_L = 0$ and for R to vote against I in such a stage game.

The claim that the given strategies do not form a perfect Bayesian equilibrium if $\delta \in [0, \underline{\delta})$ follows immediately from proposition 1 where we have shown that the RC, NC, and IC constraints can be satisfied simultaneously if and only if $\delta \in [\underline{\delta}, 1)$. ■

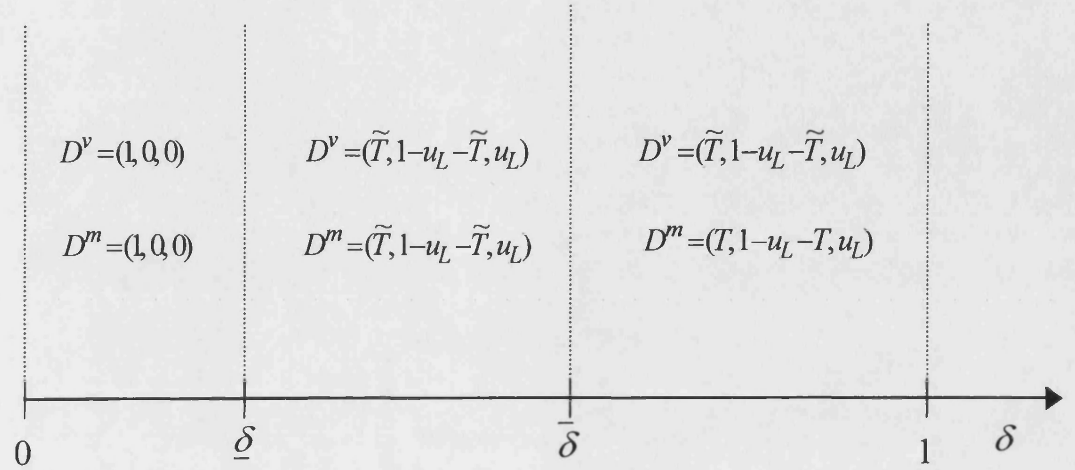

Figure 1: Distribution implemented by a democratic ruler and a monopoly mafia.

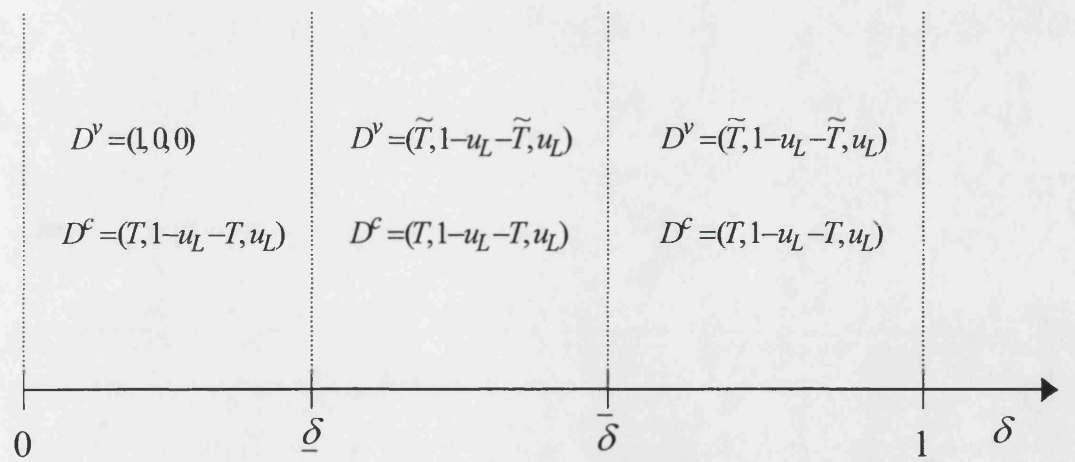

Figure 2: Distribution implemented by a democratic ruler and a mafia cartel when $\alpha = 0$.

Conclusion

In this thesis we studied the efficient organisation of economic institutions. Chapter 1 applied the well established property rights framework to foreign direct investment projects. The recent developments in the theory of the firm provide very useful tools which can be used to improve our understanding of foreign direct investment projects and multinational corporations. So far this application of the property rights approach has received surprisingly little attention and seems to provide a promising area for future research.

In chapter 2 we developed a new property rights theory of the firm that focuses exclusively on the role of bargaining inefficiencies in determining the boundaries of firms. This theory has clear empirical implications which are, at least in principle, testable. We intend to investigate the empirical evidence in future work. Also, the role of bargaining inefficiencies in corporate finance is an interesting and somewhat neglected research area. We are currently working on a model of debt and equity financing that uses the general framework introduced in chapter 2.

Finally, in chapter 3 we analysed the private and public provision of contract enforcement services. An understanding of these issues is not only of historical interest but also relevant for current policy discussions in some transition countries. The limited amount of work that has been done in this area so far does not do justice to the importance of the subject and much interesting work remains to be done.

On the whole, we hope that this thesis makes a modest contribution to our understanding of economic institutions and also offers some ideas for future work.

Bibliography

Admati, A. and Perry, M. (1987), "Strategic Delay in Bargaining," Review of Economic Studies, LIV: 345-64.

Aghion, P. and Tirole, J. (1997), "Formal and Real Authority in Organizations," Journal of Political Economy, 105(1): 1-29.

Arrow, K. (1975), "Vertical Integration and Communication," Bell Journal of Economics, 6(1): 173-83.

Barro, R. (1973), "The Control of Politicians: An Economic Model," Public Choice, 14, 19-42.

Binmore, K., Rubinstein, A., and Wolinsky, A. (1986), "The Nash Bargaining Solution in Economic Modelling," Rand Journal of Economics, 17: 176-88.

Blomström, M. and Kokko, A. (1996), "Multinational Corporations and Spillovers," CEPR Discussion Paper 1365.

Bolton, P. and Whinston, M. (1991), "The Foreclosure Effects of Vertical Mergers," Journal of Institutional and Theoretical Economics, 147: 207-26.

— — (1993), "Incomplete Contracts, Vertical Integration, and Supply Assurance," Review of Economic Studies, 60: 121-48.

— **and Xu, C.** (1998), "Ownership and Managerial Competition: Employee, Customer, or Outside Ownership," mimeo.

Brennan, G. and Buchanan, J. (1977), "Towards a Tax Constitution for Leviathan," Journal of Public Economics, 8(3), 255 - 73.

— — (1978), "Tax Instruments as Constraints on the Disposition of Public Revenues," Journal of Public Economics, 9(3), 310 - 18.

— — (1980), *The Power to Tax: Analytical Foundations of a Fiscal Constitution*, Cambridge University Press.

Chemla, G. (1996), *Essays on the Theory of the Firm: Interactions Between Capital, Product and Labour Markets*, PhD Thesis, LSE.

Chiu, S. (1998), "Noncooperative Bargaining, Hostages, and Optimal Asset Ownership," American Economic Review, 88(4): 882-901.

Coase, R. (1937), "The Nature of the Firm," Economica, 4: 386-405.

Crampton, P. (1992), "Strategic Delay in Bargaining with Two-Sided Uncertainty," Review of Economic Studies, 59: 205-25.

De Long, B. and Shleifer, A.(1993), "Princes and Merchants: European City Growth before the Industrial Revolution," Journal of Law and Economics, 26, 671 - 702.

De Meza, D. and Lockwood, B. (1998), "Does Asset Ownership Always Motivate Managers? Outside Options and the Property Rights Theory of the Firm," Quarterly Journal of Economics, 113(2): 361-86.

Dessein, W. (1999), "Authority and Communication in Organizations," mimeo.

Dnes, W. (1993), "A Case-Study Analysis of Franchise Contracts," Journal of Legal Studies, 22, 367-93.

Farrell, J. and Gallini, N. (1988), "Second-Sourcing as a Commitment: Monopoly Incentives to Attract Competition," Quarterly Journal of Economics, 103(4), 673-94.

Ferejohn, J. (1986), "Incumbent Performance and Electoral Control," Public Choice, 50, 5-25.

Fiorentini, G. and Peltzman S. eds. (1995), *The Economics of Organised Crime*, Cambridge University Press.

Forbes, K. and Zampelli, E. (1989), "Is Leviathan a Mythical Beast?," American Economic Review, 79(3), 568 - 77.

Franchetti, L. [1874] (1985), "Condizioni politiche ed amministrative della Sicilia," in L. Franchetti and S. Sonnino, eds., *Inchiesta in Sicilia*, Florence: Vallecchi.

Fudenberg, D. and Tirole, J. (1991), *Game Theory*, MIT Press.

— **and Levine, D.** (1994), "Efficiency and Observability with Long-Run and Short-Run Players," Journal of Economic Theory, 62, 103-35.

Gambetta, D. (1993), *The Sicilian Mafia: The Business of Private Protection*, Harvard University Press.

Greif, A. (1993), "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition," American Economic Review, 83(3), 525 - 548.

— **and Kandel, E.** (1995), "Contract Enforcement Institutions: Historical Perspective and Current Status in Russia," in E. Lazear, eds., *Economic transition in Eastern Europe and Russia: Realities of reform.* Stanford, CA: Hoover Institution Press, 1995, 295 - 321.

— **Milgrom, P., and Weingast, B.** (1994), "Coordination, Commitment, and Enforcement: the Case of the Merchant Guild," Journal of Political Economy, 102(4), 745 - 776.

Grossman, S. and Hart, O. (1986), "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," Journal of Political Economy, 94: 691-719.

Hansmann, H. (1996), *The Ownership of Enterprise*, Harvard University Press.

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

Hart, O. and Moore, J. (1990), "Property Rights and the Nature of the Firm," Journal of Political Economy, 98: 1119-58.

— — (1998), "Cooperatives versus Outside Ownership," NBER Working Paper No.6421.

— — (1999), "Foundations of Incomplete Contracts," Review of Economic Studies, 66(1): 115-138.

— **and Tirole, J.** (1990), "Vertical Integration and Market Foreclosure," Brookings Papers on Economic Activity, Microeconomics: 205-76.

Hay, R. and Shleifer, A. (1998), "Private Enforcement of Public Laws: A Theory of Legal Reform," American Economic Review, 88(2), 398 - 403.

— — **and Vishny, R.** (1996), "Toward a Theory of Legal Reform," European Economic Review, 40, 559 - 567.

Holmström, B. (1999), "The Firm as a Subeconomy," mimeo.

— **and Tirole, J.** (1989), "Transfer Pricing and Organizational Form," Journal of Law, Economics and Organization, 7(2): 201-28.

— **and Roberts, J.** (1998), "The Boundaries of the Firm Revisited," Journal of Economic Perspectives, 12(4): 73-94.

Klein, B. and Leffler, K. (1981), "The Role of Market Forces in Assuring Contractual Performance," Journal of Political Economy, 89(4), 615 - 41.

— **Crawford, R., and Alchian, A.** (1978), "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process," Journal of Law and Economics, 21: 297-326.

Konrad, K. and Skaperdas, S. (1999), "The Market for Protection and the Origin of the State," CEPR Discussion Paper No. 2173.

Kreps, D. (1990), "Corporate Culture and Economic Theory," in J. Alt and K. Shepsle, eds., *Perspectives on Positive Political Economy*, Cambridge University Press.

Kumar, K., Rajan, R., and Zingales L. (1999), "What Determines Firm Size?," mimeo.

Markusen, J. and Venables, A. (1999), "Foreign Direct Investment as a Catalyst for Industrial Development," European Economic Review, 43, 335-56.

Matouschek, N. and Venables, A. (1999a), "Evaluating Investment Projects in the Presence of Sectoral Linkages," mimeo.

— — (1999b), "Evaluating Investment Projects in the Presence of Sectoral Linkages - Case Studies," mimeo.

Muthoo, A. (1999), *Bargaining Theory with Applications*, Cambridge University Press.

Nöldeke, G. and Schmidt, K. (1998), "Sequential Investments and Options to Own," Rand Journal of Economics, 29(4), 633-53.

Oates, W. (1985), "Searching for Leviathan," American Economic Review, 75(4), 748-57.

Olson, M. (1993), "Dictatorship, Democracy, and Development," American Political Science Review, 83(3), 567 - 576.

Osborne, M. and Rubinstein, A. (1990), *Bargaining and Markets*, Academic Press.

Persson, T. and Tabellini, G. (2000), *Political Economics: Explaining Economic Policy,* MIT Press.

Pistor, K. (1996), "Supply and Demand for Contract Enforcement in Russia: Courts, Arbitration, and Private Enforcement," Review of Central and East European Law, 22(1), 55 - 87.

Rajan, R. and Zingales, L. (1998), "Power in a Theory of the Firm," Quarterly Journal of Economics, 386-432.

Reiche, S. (1999), "Foundation of Incomplete Contracts: Asymmetric Information and Renegotiation," mimeo.

Riordan, M. (1990), "What is Vertical Integration?," in M. Aoki, B. Gustafsson, and O.Williamson (eds.), The Firm as a Nexus of Treaties, London, Sage: 94-111.

Rodriguez-Clare, A. (1996), "Multinationals, Linkages, and Economic Development," American Economic Review, 86, 852-873.

Rosenthal, J. (1992), *The Fruits of Revolution*, Cambridge University Press.

Rubinstein, A. (1982), "Perfect Equilibrium in a Bargaining Model," Econometrica, 50: 97-109.

— **and Yaari, M.** (1983), "Repeated Insurance Contacts and Moral Hazard," Journal of Economic Theory, 30, 74-97.

Segal, I. (1999), "Complexity and Renegotiation: A Foundation for Incomplete Contracts," Review of Economic Studies, 66(1): 39-56.

Williamson, O. (1975), *Markets and Hierarchies*, The Free Press.

— (1985), *The Institutions of Capitalism*, The Free Press.