The London School of Economics and Political Science

An Economic Theory of Collusion, Blackmail and Whistle-Blowing in Organisations

Thesis submitted by

Mikko Leppämäki

to the

University of London

for the degree of

Doctor of Philosophy

July 1997

Tuijalle

Abstract

This thesis examines informal and corruptive activities agents may pursue within organisations. Chapter 1 is a brief introduction to the general theme and the related literature.

Chapter 2 develops a simple theory of non-monetary collusion, where agents collude by exchanging favours. It examines the optimal use of supervisory information in a simple hierarchy under potential collusion. It is shown that when only the supervisor's information about the agent is used, collusion does not arise, since favours can not be exchanged. Secondly, it is analysed whether the agent's information about his superior should be used. In this case collusion is possible, and there is an interesting trade-off between the benefits of using additional information and the costs of collusion. It is then shown that sometimes the principal may be better off when using less than all available information.

Chapter 3 considers task assignment and whistle-blowing as measures a principal may use to break collusion. The principal's response to potential collusion is to allocate less time to monitoring, and he breaks collusion with money. It is shown that the principal may also break collusion by hiring a third worker, and the decision how to break collusion optimally is endogenously determined. Breaking collusion by task assignment is costly, and therefore we consider whistle-blowing as a collusion breaking device. It provides the principal strictly higher welfare than the collusion-proof solution. It is also shown that under reasonable conditions, the collusion-free outcome will be achieved with no further cost.

Chapter 4 develops a model of blackmail, where a piece of information an agent prefers to keep private may facilitate blackmail when another agent, namely a blackmailer, threatens to reveal that information. The crucial feature of the blackmail game is the commitment problem from the blackmailer's side. The blackmailer can not commit not to come back in future to demand more despite the payments received in the past. The chapter outlines conditions under which successful extortion may arise, and shows that there is a unique Markov Perfect Equilibrium, which gives a precise prediction how much money the blackmailer is able to extort from the victim. It is also shown that the blackmailer receives a blackmail premium that compensates the blackmailer for not taking money from the victim and revealing information anyway.

Contents

List of Figures

Acknowledgements

My greatest debt, of course, is due to Professor Patrick Bolton. I could not have been luckier than to have a supervisor like him. Patrick has been an excellent supervisor and he has always been open and enthusiastic to my preliminary research ideas. Under his careful, friendly and patient supervision these ideas have developed into the chapters of this thesis. During the last years I have learned tremendously from him, and his way of thinking about economic problems has profoundly effected my view of economics. For all this I am very grateful for him.

I should also thank Dr. Leonardo Felli with whom I had an opportunity to talk when I stayed at The London School of Economics (LSE). Leonardo's expertise and advice helped me to improve the paper, which is Chapter 2 of the present thesis. I also got some useful comments for a preliminary version of Chapter 3. During my Ph.D studies I have learned a great deal from my fellow students on both sides of the English Channel. I thank them from the bottom of my heart.

The paper which is here Chapter 2 has benefitted from discussions with Gilles Chemla, Guido Friebel, Jean-Jacques Laffont, Anna-Lisa Luporini, and Iestyn Williams. I have also benefitted from the comments of seminar participants at the LSE, 1994 EARIE conference in Crete and the 1995 Econometric Society World Congress in Tokyo. Chapter 3 has benefitted from the comments of Micael Castanheira, Gilles Chemla, Mathias Dewatripont, John Fingleton, and Guido Friebel. Chapter 3 has also benefitted from the comments of seminar participants at the Helsinki School of Economics, LSE, ECARE, and the University of Helsinki. Chapter 4 has benefitted from the comments of Mathias Dewatripont, Guido Friebel, Michael Raith, and Martin Ruckes.

At the LSE I was fortunate to be a member of the Economics of Industry Group headed by Professor John Sutton. His Industry Group, "self help" meetings and numerous workshops offered a challenging and inspiring environment in which to do research work. I am also grateful for the Finnish Institute in London, which provided me office space during the first few months of my studies in London. I was fortunate to spend the most important year of my Ph.D. studies at the European Centre for Advanced Research in Economics (ECARE) and at the Centre d'Economie Mathématique et d'Econométrie (CEME), both at the Free University of Brussels. I remember well the excellent research facilities and friendly atmosphere there. When I was finishing my thesis in Finland, the Department of Economics at the University of Lapland provided me excellent facilities and a "home base".

I wish to thank my first teacher in economics, Professor Osmo Forssell, for encouraging me to start my graduate studies. I wish also to thank Professor Mikko Puhakka, who, over the years, has been my teacher, co-author, and now also a colleague, for continuously supporting me in my studies. During my stay in London I made two very special friends: with Olli Haltia and Espen Moen I had some of the most enjoyable moments of my life, which I will never forget.

I have benefitted from financial support from the Academy of Finland, the Yrjö Jahnsson Foundation, and the Foundation of Lapland. Due to these, I have been able to concentrate fully on my studies.

Finally comes Tuija. She has been unselfish and made sacrifices in her own career, and without them it would have been impossible for me to study full-time. Tuija has been a loving and supportive wife and an excellent life companion during these years. I can hardly remember the time when I did not use my Ph.D studies as an excuse for so many things. At least now I can promise that I will not do it anymore.

Chapter 1

Collusion, Blackmail and Whistle-Blowing in Organisations: An Introductory Essay

1.1 Introduction

This thesis examines informal and corruptive activities agents may pursue within organisations. The very origin of these activities arises due to asymmetric information between members of an organisation and an organisation designer (a principal). Within the class of models which consider a so-called hidden information (adverse selection) problem, an agent has more information than the principal, and due to this private information the agent is able to earn informational rents. Hence, the principal offers a so-called second-best contract, which optimally distorts the effort level required from an inefficient type to reduce an informational rent the principal has to pay for the efficient type.

However real world organisations are not isolated principal-agent relationships, since the members of organisations interact as well. For example, a role for a so-called middle manager or a supervisor arises endogenously if the supervisor has access to more detailed information about the agent's private information than the principal does. In that case the supervisor is able to help the principal in controlling the agent, and thus reducing the informational rents the principal otherwise has to pay.

This immediately introduces the possibility of collusion (side contracting) between the supervisor and the agent. Hence organisations become a nexus of official incentive contracts between the principal and many agents, and a nexus of illegal side contracts between the members of organisations. As a consequence some new and quite important questions arise. Namely, how is the principal able to control the agents when they prefer to pursue activities apart from those which the principal prefers them to do? What are the consequences of potential collusion for the organisation design, optimal use of information, mode of incentive schemes, optimal task assignment, and time allocation between different activities the agents are expected to perform? Furthermore, how can the principal prevent collusion or at least make it less profitable? And, more generally, how can individual incentive contracts be designed so that they minimize the costs of potential collusion.

Roughly speaking, one can distinguish four types of corruptive activities agents may pursue in organisations. The broadest class is so-called influence activities, which may take various forms. The influence activities refer to the often observed behaviour of members of an organisation who spend a great deal of time and effort attempting to influence decision makers. An important economic point is that time and effort are lost from productive activities: i.e., the activities for which the workers were hired in the first place. In addition, influence activities usually have a negative effect on the quality of decision making.[1]

The second activity, and the one which gets the most attention in the present thesis, is collusion. This refers to a situation where two agents cooperate against a third party. In our context this refers to a supervisor and an agent (or simply two agents) who together manipulate information against a principal for their own benefit. This, of course, is costly for the principal, who has to pay informational rents which would be saved within a collusion-free world. A third type of activity agents may pursue is called framing. Framing means a situation where one agent alone is able to frame another agent, making him, for instance, appear to be responsible for some misbehaviour: e.g., accepting bribes, neglecting organisation rules, and so on. An important point is that the agent who frames the other agent does not need help from anyone. Framing is a serious possibility that an organisation designer has to take into consideration especially in a situation where the agent is able to benefit himself if he frames the other agent and makes him appear to be responsible for some misbehaviour. Finally comes blackmail or extortion. The concept of blackmail refers to payments to avoid revelation of information. Blackmail arises when there exists a piece of information which an agent prefers to keep private, and another agent, namely a blackmailer, threatens to reveal that information. The crucial feature of the blackmail situation is the commitment problem from the blackmailer's side. The blackmailer can not commit not to come back in future to demand more despite the payments he has received in the past.

It is important to notice that in practice it is not always possible to separate informal activities as clearly as above, because sometimes they overlap; and in some cases they even coexist. Furthermore, it is perhaps not wrong to argue that in the broadest sense the first activity, namely influence activities, includes all the remaining ones, and that collusion, framing and blackmail are merely different forms of influence activities. Nevertheless, each of them deserves closer analysis.

The present thesis examines collusion and blackmail in greater detail. Chapter 2 develops the idea that within organisations collusion appears more often in a non-monetary

---

[1] For more about influence activities, see Milgrom (1988) and Milgrom and Roberts (1988).

form rather than in monetary payments and examines the optimal use of supervisory information. Chapter 3 in turn analyses possible actions a principal may take in breaking collusion.The principal is able to prevent collusion either by task assignment (time allocation) or by creating a whistle-blowing mechanism, which is an additional incentive device that makes collusion less profitable and ensures that workers perform well and honestly the tasks for which they were hired in the first place. Chapter 4 puts aside collusion and examines the economics behind the phenomenon of blackmail in a dynamic framework. There a blackmailer extorts money from a victim in return for suppressing a piece of information the agent prefers to keep private.

In the next section we briefly discuss the development and the current state of the related theoretical literature.

1.2 The Related Literature

Corruption, bribery and extortion are unfortunate and well-documented phenomena as old as humankind itself. Already the Old Testament warned of the harmfulness of bribery. The Lord warned Moses: "Do not accept bribes, since it blinds even the most careful man, and it destroys the businesses of those who are right“.[2]

Rent seeking, bribery and corruptive activities have long been among the interests of economists.[3] Generally speaking, the harmful effects of these activities are reasonably well known. That is, resources are allocated to those who have money, but who are not necessarily those in need. And, more generally, decision making and resource allocation are inefficient, public funds are diverted from public use to private pockets, and so on. Furthermore, bribery has an adverse effect on growth, especially in developing countries.

Interestingly enough, it took quite a long time before corruptive activities within

---

[2] The quotation is from Newsweek (1994).

[3] See, for example, Krueger (1974) for a seminal article on rent-seeking and Shleifer-Visny (1993) on the economic organisation of corruption, and a recent work by Carrillo (1995a) on the optimizing model of corruption.

organisations received the full attention of economists. Presumably, one of the main reasons was that the main tools for analysing such issues, namely incentive theory and informational economics in general, have been developed during the last twenty years.

Before economists started to analyse collusion, sociologists - e.g., Crozier (1964) and Dalton (1959) - had already realised the potential problems of collusion and capture in organisations and within agency relationships. The sociological and also organisation theoretical approach emphasize that behaviour of organisation members should not be understood and considered only from the point of view of an individual organisation members's incentive, but one should also realise interaction between group and individual incentives. This is precisely what modern economic research does by introducing coalition incentive constraints in addition to individual incentive constraints. Political scientists have also raised the question of potential collusion. Their worry is that politicians collude with interest groups and pursue policies which are favourable to them but not to the public in general. In the early 1970s some economists - e.g., Stigler (1971) and later Peltzam (1976) - paid attention to the capture problem within the regulatory framework. Stigler in particular emphasized that the power of an interest group depends on the (benefits) stakes of collusion on one hand and a group's organisation costs on the other. Even though economists in the early 1970s raised important questions concerning capture in the regulatory process, they did not really answer those questions thoroughly. For example, the effects of potential collusion on optimal organisation design as well as potential solutions to collusion problems were left unanswered for some 10 to 15 years.

In this respect a big change in the depth of the analysis occurred when economists adapted incentive theory in the analysis of collusion in organisations. In contrast to the earlier literature, now the very source of collusion is considered to arise from the fundamental asymmetry of information between a principal and other members of the organisation. Ten years ago Tirole's (1986) seminal paper started a whole new branch of literature, which modelled and analysed the effects of potential collusion in organisations by using incentive theory.

Tirole's seminal paper had a profound effect on the way collusion is modelled. The basic set-up in his paper is actually very simple. An organisation is modelled as a three-tier hierarchy: a principal, a supervisor, and an agent. The principal hires the agent to perform some production task and the supervisor to control the agent. The supervisor has access to more precise information about the agent's type than the principal, and thus he is able to help the principal control the agent. In short, the principal is able to save in informational rents he otherwise would have to pay. The problem is that the agent has an incentive to bribe the supervisor to conceal information about his type by promising to pay the supervisor an ex-post bribe. This side contract is by assumption enforceable. To prevent collusion the principal has to match the benefits of collusion: i.e., he has to satisfy a so-called coalition incentive constraint. The optimal solution is that the principal offers the supervisor and the agent a collusion-proof contract which does not leave any room for profitable collusion. This collusion-proof contract in fact replicates the potential bribes between the supervisor and the agent, and thus of course it is costly for the principal. Note that this is the simplest possible set-up, which facilitates the analysis of tension between group and individual incentives, which is the very heart of the collusion problem.

Laffont and Tirole (1991) elegantly formalise the idea of collusion in a regulatory framework. In their model a firm may bribe a regulator who has access to more detailed information about the firm's costs than the government does. They show, among other things, that the optimal response to potential collusion is to offer a low-powered incentive scheme to the firm.[4]

Parallel to this evolving literature which focuses on the adverse effects of collusion, a closely-related, but still different literature developed by considering beneficial collusion. The seminal article is Holmström and Milgrom (1990), where the authors make the point that cooperation (effort coordination) between agents may actually benefit a principal in a case where agents share some private information which is not available to the principal. The benefits of cooperation are totally driven by risk-sharing motives. Some other authors have further elaborated these ideas along the lines of Holmström and Milgrom.[5]

Tirole's (1992) article from the 1990 Econometric Society's World Congress reviews the early theoretical literature and summarises the most important results that were discovered by that time. In particular, the roots of collusion and the potential measures taken

---

[4] Laffont (1990) and Varian (1990) are other early contributors to this literature.

[5] See for example Itoh (1993) and Felli and Villas-Boas (1996).

to prevent collusion in terms of organisational design were reasonably well known. In his article Tirole also raises some interesting and new research problems and develops a simple model of dynamic collusion. We will come back to this later.

In recent years the number of studies analysing collusion in organisations has increased rapidly. In order to understand what has been done within the last few years and in order to judge what the main theoretical achievements in the theory are, it may be useful to categorise the studies in some way. The simplest and also perhaps the most fruitful way to do this is to categorise them along the lines of how collusion is modelled and what type of collusion these studies consider.

In principle, the members of an organisation are able to collude in three alternative ways. First, most of the papers model collusion as ex-post monetary transfers which are based on fully enforceable side contracts. The second approach takes the problem of the enforceability of illegal side contracts more seriously and considers dynamic collusion, where side deals have to be self-enforcing. Thirdly - the approach which is developed here - collusion may also take a purely non-monetary form where members of an organisation collude by exchanging favours and no money is needed.

The overwhelming majority of studies that consider collusion borrow the basic "machinery" from Tirole's (1986) seminal article: that is, a static adverse model with a three-tier hierarchy with a principal, supervisor and agent. Monetary side contracts between the supervisor and the agent are by assumption fully enforceable. In addition an assumption of “hard” supervisory information is adopted. That is, the supervisor can not lie when reporting to the principal, but the supervisor can however conceal information.

Perhaps the most important and widely known among these papers is an article by Koffman and Lawarrée (1993), who consider collusion within an auditing context. There a manager, who is hired by shareholders, is willing to bribe an internal free corrupt auditor to report him favourably to a principal. To prevent collusion the principal has to offer a collusion-proof contract, which leaves no room for profitable collusion. Then the authors introduce a costly external uncorrupt auditor who learns the same information as the internal auditor does. Among other interesting results, they show that costs of collusion can be constrained by hiring the external auditor with a positive probability in a case where the internal auditor and the manager may have colluded.

In Koffman and Lawarrée (1995), the authors show that under some circumstances it may actually be profitable for a principal to let collusion occur in an equilibrium, because they consider a case where the auditor probably is corruptible. Therefore, when the probability that the auditor is honest is high enough, it may be profitable to take a chance and offer a normal second-best contract.

Laffont-Martimort (1994) is an interesting paper that considers the optimal organisational response to potential collusion. They compare two regulatory structures, a unified regulator or two separate regulators in cases where there is a firm which has two pieces of private information concerning its costs and a regulator knows only the prior distributions of them. They show that separation of regulators works as a commitment against the threat of regulation capture, since separation reduces the amount of asymmetric information a single regulator holds, and thus also reduces the potential benefits of collusion.

Felli (1996) considers delegation as an organisation measure taken by a principal to prevent collusion between a middle manager and a worker. The idea in his paper is strikingly simple. A crucial assumption is that collusion takes place under conditions of asymmetric information. When a supervisor has the authority to exploit an agent's private information, the latter has no incentive to collude, because collusion will reveal information to the supervisor, who will exploit him by reporting what he has learned to the principal. Thus, in an equilibrium, collusion does not take place. One must note that a delegation works against the supervisor, who is worse off under an optimal collusion-proof contract. Felli also demonstrates that his collusion prevention scheme coincides with one where the supervisor chooses a contract for the agent.[6]

What, then, are the main contributions of these studies that have adopted the traditional approach to collusion with monetary side payments? Even though the approach has its limitations, it seems to have achieved many goals. Firstly, the literature has introduced

---

[6] Other interesting work includes Olsen and Torsvik (1994) and Mogilianski (1995), both of whom consider collusion when a principal is unable to commit an incentive contract he proposes. Brundin (1995) in turn considers collusion in a debt contract. Che (1995) considers the phenomenon of revolving doors. Koffman and Lawarrée (1996) consider the old question, namely "who polices the police" within their "standard model". Laffont-Martimort (1995a) and (1995b) and Baliga (1995) have considered collusion when agents collude by manipulating "soft" information. Prendergast and Topel (1996) is an interesting work in which the authors analyse favouritism in organisations.

many real-life applications and extended the analysis of collusion in several important directions. Generally speaking, it has increased our understanding of potentially corruptive organisation structures that may facilitate collusion. Thirdly, it has provided a large number of rather simple and realistic organisational measures to prevent collusion that an organisation designer may use in practice. This literature has also shown the precise conditions under which it may be optimal not to fight against potential collusion.

A starting point for the analysis of side contracts within a dynamic framework is that, due to their illegality, they must be self-enforceable. The most natural way to proceed in this direction is to look at the repeated relationship between the colluding parties and analyse reputation as a mechanism that enforces side deals. Tirole (1992) develops a simple dynamic model of collusion using reputation as such a mechanism that makes side contracts self-enforcing. Interestingly enough, the insights of the main results that are derived from a static model seem to carry over to self-enforcing side contracts as well.

Acemoglu's (1995) paper is an interesting attempt to model collusion in a dynamic framework where enforceable monetary side contracts are not feasible. There shareholders hire an empire builder manager to run a firm, and a manager in turn hires an auditor. The possibility for collusion between the manager and the auditor arises since the auditor can hide some information to the benefit of the manager and the manager prefers to continue a project whatever the state of nature is. The auditor has an incentive to hide information due to his own career concerns, since the manager has the authority to fire him. Whenever this threat to fire is credible, collusion may arise even though no enforceable side deals can be written. In Acemoglu's model, collusion is a serious problem when the manager has the authority to hire and fire the auditor, which clearly is an inefficient arrangement. This brings us to the interesting question of why we see this type of arrangement in practice, since it facilitates collusion.

Recently Martimort (1996a) has written an interesting paper on self-enforceable collusion. He considers an infinitely repeated game where a principal hires two agents to produce. In his model collusion between agents arises due to the possibilities for agents to realise future benefits from an ongoing relationship. That is, even though from a static point of view, truth-telling is a dominant strategy for a single agent, it may still be profitable to lie when reporting his type to the principal. This is because of the existence of future benefits that will be achieved from reciprocal behaviour by a fellow agent. This type of collusion introduces dynamic coalition incentive constraints that the principal has to satisfy in order to prevent collusion.[7]

In contrast to static monetary collusion, a dynamic approach has at least purely technical merits, since it has dropped an arguably strong assumption of fully enforceable monetary side contracts. Of course, at this point it is very difficult to fully acknowledge the main contributions, since there exist only a very few studies. However, the value of studies seems to be at least two-hold. The existing studies have greatly increased our understanding of the organisational structures and features that may lead to self-enforcing collusion. They have also clarified the role of such organisational measures as rotation and task assignment in collusion prevention. On the other hand, the complexity of models has increased enormously, and, at the moment, not very many new insights have been achieved. So it seems to me that a dynamic approach has its limitations as well.

The present thesis takes a slightly different approach to collusion than the "mainstream" literature. We consider collusion that takes place only between the members of an organisation, and the main idea is that collusion in organisations appears more often in a non-monetary form than in monetary payments. A third possible way to collude is developed by proposing a simple theory of non-monetary collusion, where members of organisations collude not by transferring money, but by exchanging favours among themselves. That is, an agent bribes another agent to do something illegal by offering a counter favour. This type of collusion is fairly common in most real-world organisations and everyday examples are numerous.

In the present thesis, collusion is modelled as mutually beneficial information manipulation where agents conceal private information they have learned about each other from a principal. In that way, the agents are able to earn information rents that would not be possible if they reported to the principal truthfully. Of course, this is not the only possible way to model non-monetary collusion, and arguably it is quite mechanical as well. It is, however, probably the simplest possible way to incorporate the idea of non-monetary

---

[7] See also Carrillo (1995b), who develops a dynamic optimization model of corruption and analyses different measures in fighting bribery. Martimort (1996b) constructs a theory of the life cycle of regulatory agencies.

collusion, which takes the form of a simultaneous exchange of favours. In addition, it is very convenient and tractable since the modeller is then able to use a normal adverse selection framework to analyse many interesting questions.

In practice, of course, there are many alternative ways agents can collude without any actual monetary transfers. We can easily relabel and reinterpret variables we have chosen to use in modelling non-monetary collusion. For example, workers may agree not to reveal some unfavourable information to their superiors just to save their own jobs. A foreman and a worker may jointly cover up each other's mistakes to avoid interference by their superiors and possible firings. An auditor and a manager may agree on information concealment from shareholders for their own benefit. A civil servant may look the other way when his superior is involved in bribery if, for example, the superior civil servant performs a counter favour by promoting the civil servant for being quiet, and so on. Note, by the way, that this type of collusion motivates Tirole's (1986) seminal paper that, however, considers only monetary bribes.

In principle both monetary and non-monetary types of collusion may coexist in organisations. In the theoretical literature, it has been assumed that monetary bribes involve transaction costs. That is, if an agent bribes a supervisor with a bribe b, then the supervisor's utility from such a bribe is: $b/(1+\lambda)$, where $\lambda > 0$ stands for the transaction costs of this illegal side deal.

One way to justify our choice to consider only non-monetary collusion is to assume that the transaction costs of monetary bribes are infinitely high ($\lambda=\infty$), because, for example, a principal has a perfect monitoring technology over monetary transfers within an organisation. In that case, members of organisations may collude only by exchanging favours. Note that this type of collusion has an extremely efficient transaction technology, since transaction costs are zero ($\lambda=0$). Collusion by exchanging favours does not involve actual transfers at all. Rather, agents simply exchange favours, and the benefits flow directly to them, for example, in the form of higher wages.

Non-monetary collusion as it is modelled here extends hierarchial collusion between a supervisor and an agent to lateral collusion, where collusion takes place between agents who work at the same organisational level. For example, in Chapter 3 workers on the "factory floor" jointly conceal information about each other for their own benefit. Collusion as it is modelled here does not require ex-post monetary transfers as is the case with the traditional approach. Consequently, collusion is much easier for agents, and for the principal it is more difficult to observe and more costly to prevent.

The second point where the present thesis departs from the current theoretical literature is our analysis of blackmail. The literature has almost totally ignored the phenomenon of blackmail and concentrated purely on collusion. Tirole (1992), however, mentions the phenomenon of blackmail, where a supervisor is able to hurt an agent by threatening not to report information that is favourable to the agent. Mogiljanski's (1994) study then tries to model this type of blackmail within a static adverse selection framework. However, that approach is not very satisfactory, since it overlooks the two fundamental features that we believe are crucial for blackmail. The first is the credibility problem: that is, whether the blackmailer will carry out the threat if a victim declines to pay. Secondly, there is a fundamental commitment problem: the blackmailer may come back to ask for more money in future. Clearly, the phenomenon of blackmail has a very strong dynamic feature, and any satisfactory model of blackmail should be dynamic. At the moment there is no economic theory of blackmail. Chapter 4 is a first attempt in this direction. There we construct a dynamic model of blackmail, where a piece of information an agent prefers to keep private may facilitate blackmail, when the blackmailer is able to reveal it. There also the credibility and the commitment problems are carefully analysed.

This is a good point at which to reflect on the value that economists have added to the understanding of organisational collusion. What have economists said about collusion that has not been analysed by sociologists and political scientists? In general, the rigorous economic modelling of collusion has greatly helped in discovering the roots of collusion in organisations. Economists have chosen to emphasize asymmetric information as a crucial feature that facilitates collusion. How realistic or relevant is this in terms of practise? Interestingly, asymmetric information seems to play a major role in most real-life collusion problems, and in that respect the informational economics approach is well justified.

The theoretical literature has derived a large number of organisational measures with which to combat collusion. In practice, many of these measures are already in use and organisations' response to potential and existing collusion in terms of organisation design is also very often within the lines of the theory. If one thinks about the development of the theoretical literature, one comes easily to the conclusion that the main achievements are not the development of an elegant and unified theory of collusion. The most important contributions lie in the numerous applications, as well as in a better understanding of the features that facilitate collusion and the ways collusive activities can be prevented. In terms of the methodology of modelling collusion, the recent contributions of dynamic and non-monetary collusion will be further elaborated. Since Chapter 4 is to the best of my knowledge, the first economic model of blackmail, it is clear that there are still many open questions. On the whole economic analysis of corruptive activities in organisations is still quite a young field, and there are still many open questions, and much work remains to be done.

Chapter 2

Non-Monetary Collusion and Optimal Use of Information

2.1 Introduction

Why in real-world organisations is all available information typically not used? Furthermore, when only a part of the available information is used, why is it the case that only superiors' information about their subordinates is used, and workers' information about their superiors is ignored? This chapter proposes a simple model to explain the above-mentioned real-life facts in terms of organisation members' possibilities to engage in harmful side-contracting. We model non-monetary collusion as mutually beneficial information manipulation, where agents exchange favours by jointly concealing information from a principal.

Within the model collusion arises only in the case where all available information is used. There the principal is trading off between the benefits of using all available information and the costs of collusion. It is shown that under certain conditions it is optimal for the principal not to use all available information, since then the principal deters collusion and gains more valuable information without additional cost. The principal may simply be better off when using less information.

In practice, members of real-world organisations often learn information which would be valuable for an organisation designer. We refer to information an agent has learned about another agent as "supervisory information". It is not so uncommon that the principal prefers and expects such information to be reported to him. For example, middle managers are supposed to report on their subordinates, line managers should report on the workers, and so on. In any organisation in which workers and managers work closely, workers also gain information about their superiors, and sometimes they also have the possibility to communicate that information to the top of the hierarchy.

In general, an agent's incentive to manipulate supervisory information depends on the benefits it brings to him. Compensation from engaging in side-trading and information manipulation may take various forms such as money, help, favours, promotions, and so on. In contrast to most of the existing theoretical literature, this chapter considers the case where compensation takes a purely non-monetary form. More precisely, collusion between agents is modelled as a simultaneous exchange of favours, which takes the form of supervisory information concealment.This type of collusion is fairly typical and, in contrast to monetary transfers, a type of collusion one certainly expects to find in real-world organisations.

Consider, for example, co-workers who agree not to reveal unfavourable information to their superiors, a foreman and worker who jointly cover up each other's mistakes, and so on.

This chapter analyses whether the principal should use all available supervisory information or simply a part of it, and if only part of it is used, whose information should be used and whose information should be ignored. The chapter uses a simple, three-tier Principal-Supervisor-Agent (P-S-A) model to analyse the optimal use of supervisory information when a supervisor and an agent are able to exchange non-monetary side transfers.[1] The novelty in our version of the P-S-A model is the self-interested supervisor who also has a production role and the agent who potentially also has a monitoring role. The supervisor is regarded as a middle manager who monitors the agent and informs the principal by means of reports, but he also has a productive function of his own. This productive function may take various forms such as cost minimization, coordination of production, coordination of management, and so on. We regard the supervisor's production task as a fine-tuning of production.

The agent is a worker at the bottom of the hierarchy, and the agent learns information about his superior. Therefore, the principal may also ask the agent to monitor his superior and report back his observations. In practice, this monitoring by workers may be in the form of questionnaires distributed to the workers in which they are asked to assess the manager's management and coordination activities, advisory activities, management style, the manager's "type", and so on. The mode of collusion we choose to consider can be understood, for example, in terms of practice as a case where the supervisor and the agent together fill out questionnaires, seal them and send them away.

We analyse two organisational structures: Organisation I: principal-productive supervisor-agent: We have a situation in which the supervisor has the dual task of supervision and production, but the worker's information about his superior is not used. Organisation II: principal-productive "supervisors": This structure allows for the dual role of production and monitoring by the agent as well as the supervisor. The principal may or may not use the agent's information about his superior. Organisation I is, therefore, a particular case of Organisation II.

---

[1] Tirole (1986) is the seminal paper which introduced the P-S-A model in the analysis of collusion in organisations.

We derive an interesting trade-off which is related to the use of supervisory information. We show that when only one supervisory information source (Organisation I) is used, collusion problems do not arise, because favours cannot be exchanged. A collusion-proof equilibrium is achieved without further costs. When two supervisory information sources (Organisation II) are used, collusion problems do arise. In general, the more information, the better; but in our model it makes non-monetary collusion possible, and this creates additional costs which are borne by the principal. Thus, a trade-off arises with respect to whether it is more beneficial to use only one source of supervisory information and avoid the costs of collusion or to use both supervisory information sources and bear the costs of collusion.

We describe precise conditions under which additional supervisory information the agent may provide about his superior should and should not be used. This decision is related to the value which supervisory information provides the principal. The value of supervisory information is directly related to the size of the supports of the random variables a member's production task includes. Then, by considering asymmetric supports, the value of supervisory information differs.

It is shown that when collusion is not an issue, all available information should indeed be used. When collusion is a problem, the principal has to calculate the benefits which new detailed information provides and compare them to the costs of collusion. After that, the decision concerning the optimal organisational mode is straightforward. In the case in which an agent and a supervisor are symmetric, we show that all available supervisory information should be used. By considering asymmetry we derive interesting results, one of the main ones being that when the asymmetry between the supervisor and the agent is great enough, it is optimal to use only one supervisory information source. Namely, it is optimal to use only the supervisor's information about the agent. On the whole, we show that the optimal use of supervisory information is endogenously determined. The model demonstrates that in certain situations the principal is better off using less information; if the principal commits not to use all supervisory information, non-monetary collusion is deterred and he will gain some additional important information without further cost.

The main contribution of this chapter lies in the development of a simple theory of non-monetary collusion that we expect to find in many real-world organisations. It is shown that the decision concerning the use of supervisory information is endogenously determined. In particular, we show that the choice of organisation mode itself effects whether collusion arises or not. Interestingly, we show that sometimes it is not optimal to use all available supervisory information, and, moreover, we show whose supervisory information should be used and whose information should be ignored. Another aspect of this model is that it tells the principal that the supervisor should have on his possession the most valuable information. One more interesting fact that arises is that the supervisor will perform his tasks better if he is not monitored by a third party. The above results are very much those that real-world organisation use in practice, too. Typically, tasks in organisations are assigned in such a way that only the superiors monitor and the subordinates perform production tasks. Our model provides a rational explanation for this separation of monitoring and production tasks. On the whole the chapter provides a basis why organisations typically do not use all available information.

This chapter is related to a small and relatively new literature of collusion and the theory of organisations.[2] The studies closest to the present one are Tirole (1986), Kofman and Lawarrée (1993) and Laffont and Tirole (1991).[3] They consider monetary collusion between the monitor and the monitored agent. The main issue in these studies lies in the analysis of how a principal can reduce the costs of collusion. For example, Kofman and Lawarrée (1993) show how, by hiring an external auditor, the principal reduces the costs of collusion. Laffont and Tirole (1991) in turn show how the principal optimally modifies the incentive contract to reduce the costs of collusion. In contrast to these approaches, this chapter introduces a model of reciprocal monitoring between the members of an organisation and considers only non-monetary collusion. In particular, we choose to emphasize the optimal use of supervisory information in organisations. Recently, Laffont and Meleu (1996) have independently of my work written a paper that considers reciprocal supervision and collusion. They consider both monetary and non-monetary collusion, and in this sense the model presented here is a special case of their model, since here the transaction costs related to monetary bribes are assumed

---

[2] See Tirole (1992) for a survey of this literature.

[3] Another related literature has adopted a principal-multi agent framework. Papers such as Holmström and Milgrom (1990) and Itoh (1992) analyse agents' cooperation (beneficial collusion) when the agents share information about their effort choices.

to be infinite. Laffont and Meleu also conclude that sometimes an asymmetric monitoring structure (Organisation I in this study) is better, but the way how they derive the result is different than here. Even though Laffont and Meleu's model is quite similar to the one presented here, there are differences as well. For example, the moral hazard problem where the agent reveals his own type to the other agent so that they will both be rewarded more generously by the principal does not arise here. Consequently, in the present model it is never optimal to let collusion occur in an equilibrium.

The remainder of this chapter is organized as follows. In section 2.2 we present the model. It is analysed in section 2.3. Section 2.4 discusses the economics and interpretations of Organisations I and II, and section 2.5 concludes.

2.2 The Model

The Parties

The model we construct borrows basic features from Tirole (1986) and Laffont and Tirole (1991). We consider a simple, three-tier hierarchy with three players: a principal (P), a supervisor (S), and an agent (A). The principal is a risk-neutral residual claimant of an organisation who hires a supervisor and an agent to perform production and monitoring tasks on an indivisible project. The supervisor and the agent are risk neutral in income and risk averse in effort. In other words, they have quasi-linear preferences. P does not have time to supervise either the agent or the manager because, by assumption, his attention is limited. A is a productive agent at the bottom of the hierarchy. S is a middle manager who has a dual role: he contributes to production, but, in addition, he monitors A and reports his information to P (Organisation I). In Organisation II A's information about his superior may also be used.

The agent, A has private information about a random cost parameter θ_1, and he can reduce costs by exerting effort e_1, which only he knows. This effort causes him a disutility, which in monetary terms is $\psi(e_1)$ with $\psi'(e_1)>0$ and $\psi''(e_1)>0$. θ can be interpreted, for instance, as the agent's type or technological variable related to production. Later, we refer to $\underline{\theta}_1$ as a good state of nature and to $\bar{\theta}_1$ as a bad state of nature. The production cost of the process in which A is involved can be written as follows: $c_1=(\theta_1-e_1)$. θ_1 has a binary support

$\{\underline{\theta}_1,\overline{\theta}_1\}$ with probability p and (1-p) respectively, $\underline{\theta}_1<\overline{\theta}_1$ and $\Delta\theta_1=\overline{\theta}_1-\underline{\theta}_1$.

The supervisor, S also has private information about random cost parameter θ_2, and he can also reduce costs by exerting effort e_2, which only he knows. This effort presents him with a disutility, which in monetary terms is $\psi(e_2)$. We also assume that $\psi'(e_2)>0$ and $\psi''(e_2)>0$. The cost of the supervisor's fine-tuning production task can then be written as follows: $c_2=(\theta_2-e_2)$. θ_2 also has a binary support $\{\underline{\theta}_2,\overline{\theta}_2\}$, and without loss of generality we assume that with probability p: $\theta_2 = \underline{\theta}_2$, and with probability (1- p): $\theta_2 = \overline{\theta}_2$. In addition $\underline{\theta}_2<\overline{\theta}_2$ and $\Delta\theta_2=\overline{\theta}_2-\underline{\theta}_2$.

When hiring the agent and the supervisor, P must offer contracts which guarantee them at least their reservation utility, which is normalized to zero. We also assume that both the agent and the supervisor are protected by limited liability; that is, individual rationality constraints must hold ex post in all states of nature. The principal pays the wage w_1 to the agent and w_2 as a function of realized costs. The agent's utility is $U_1 = w_1 - \psi(\theta_1-c_1)$, and the supervisor's utility is $U_2 = w_2 - \psi(\theta_2-c_2)$. The ex post individual rationality (IR) constraints are as follows:

$$\underline{U}_i=\underline{w}_i-\psi(\underline{\theta}_i-\underline{c}_i)\geq 0 \quad i=1,2, \tag{2.1}$$

$$\overline{U}_i=\overline{w}_i-\psi(\overline{\theta}_i-\overline{c}_i)\geq 0 \quad i=1,2. \tag{2.2}$$

To induce the agent and the supervisor to exert effort, the contract must satisfy incentive compatibility (IC) constraints:

$$\underline{w}_i-\psi(\underline{\theta}_i-\underline{c}_i)\geq\overline{w}_i-\psi(\underline{\theta}_i-\overline{c}_i), \quad i=1,2, \tag{2.3}$$

$$\overline{w}_i-\psi(\overline{\theta}_i-\overline{c}_i)\geq\underline{w}_i-\psi(\overline{\theta}_i-\underline{c}_i), \quad i=1,2. \tag{2.4}$$

In addition to cost-reducing activities, the supervisor and the agent sometimes learn each other's type. Thus, they are able to help the principal reduce informational rents by reporting their observations of each other. We assume that observing and reporting do not require any effort. The important point is that the principal does not have to pay for that supervisory information. To see why, note that once the agent and the supervisor have been offered the incentive compatible contracts, they have no incentive to lie in their reports on each other. The supervisor, for example, once offered a second-best contract, is indifferent between reporting the agent's type or concealing it if he has learned it. More precisely, he has no incentive to conceal that information, because a report does not affect his own welfare. In short, the principal gets supervisory information with no additional cost: i.e., the wage for supervisory information can be set as equal to zero.

It is assumed that the random cost parameters θ_1 and θ_2 are independently distributed, and therefore there are no gains to be achieved by conditioning the agent's and the supervisor's compensation on each other's types.[4] Note that we assume that θ_1 and θ_2 are not necessarily identically distributed. In particular, we later consider $\Delta\theta_2 \leq \Delta\theta_1$, which captures the idea that the supervisor's fine-tuning activity involves less uncertainty than the agent's main production task. As usual with the combined moral hazard and adverse selection, it is also assumed that even if P observes realized costs c_1 and c_2, he cannot disentangle their components. For example, P does not know whether the realized high cost $\bar{c}_1$ is due to A's laziness or a bad state of nature.

The principal, P hires S and A to realize an indivisible project which has exogenously set gross value R. The principal's expected utility is:

$$E[R-c_1-c_2-w_1-w_2]\,. \tag{2.5}$$

In the above, expectations are formed for all possible states of nature.

Information structure

One of the key elements of the information structure is that P observes neither random cost parameters θ_1 and θ_2 nor signals of them. However, the principal has priors over θ_1 and θ_2. Slightly abusing the notation, we write σ_1 as the signal the supervisor has of the agent's type θ_1. The supervisor can, therefore, help the principal control A by reporting any information he learns to the principal. Similarly, the agent gets signal σ_2 about the supervisor's type θ_2, and thereby he can help the principal control the supervisor. The agent learns signals with probability 1, and without loss of generality we assume that the supervisor

---

[4] We use IC constraints which state that truth-telling is a dominant strategy. We could have used Bayesian IC constraints without changing the results. This follows from the assumption of independent production and the fact that we use ex post IR constraints.

also learns the signal with the same probability. The probability of learning signals is exogenously determined.

We follow the same line as Tirole (1986) and most of the existing literature in assuming that signals are hard information: that is, they are verifiable. Having observed the other's type, one can report it to P, and by assumption that report is verifiable. Given that a random cost parameter has value θ, a signal can take two values, $\sigma=\{\theta,\varnothing\}$. In other words, the true state of nature is either observable or not. If signal $\sigma=\{\theta\}$, one can report $r=\theta$ or $r=\varnothing$. That is, one can report the true state of nature or remain silent and claim that one has not observed anything. However, one cannot report a "wrong" state of nature; that claim is by assumption unverifiable. Given that the true state of nature is θ, and signal $\sigma=\theta$, one can then report $r=\{\theta,\varnothing\}$. However, if signal $\sigma=\varnothing$ then one can only report $r=\varnothing$. The main point here is that the principal cannot distinguish whether the agent or the supervisor has actually observed a true state of nature or not.

We also assume that A (S) learns whatever signal S (A) learns. In other words, both A and S know a state of nature. Bearing in mind that random variables θ_1 and θ_2 are independently distributed, we must consider all sixteen cases. Fortunately, most cases behave symmetrically, which simplifies the analysis. For further details concerning states of nature, see Appendix A.

Collusion

In principle, collusion (side transfers) may take either a monetary or non-monetary form. In this paper we concentrate on non-monetary side transfers. There are three basic reasons for this. First, monetary transfers (bribes) are typically illegal in most societies. Secondly, monetary transfers between members of any organisation are typically strictly forbidden. Thirdly, in some cases it is quite easy for the principal to monitor monetary transfers by examining accounting records. In contrast, non-monetary side transfers are rather difficult to monitor. On the whole, non-monetary collusion is a rather more realistic and frequently-observed type of collusion in real-world organisations. This idea is modelled by assuming that transaction costs related to monetary transfers are infinite. The members of an organisation are able to collude only by exchanging favours, and transaction costs of non-monetary transfers are assumed to be zero.

The possibility of collusion is, of course, related to the supervisor's and the agent's monitoring roles. The idea is that the monitor has discretion over the monitored party, because he may release or withhold supervisory information he has learned. Within our framework, A and S may agree not to release information that could hurt each other. The agent and supervisor prefer the principal to be uninformed about the true state of nature when it is good, since they can earn rents because of the asymmetry of information between themselves and the principal.

In this chapter collusion appears purely in the form of non-monetary side transfers, and these side transfers are unobserved by P. In short, the monitored party "buys" the monitor's silence with a favour in return, and he also remains quiet about the true state of nature. Within Organisation II, both A and S have discretion over each other's activities, and the side contract takes the form of "a favour and a counter favour".[5] In our static model these favours and counter favours are exchanged simultaneously.

We assume, as does most of the literature, that side contracts are enforceable (enforceability approach). Of course, side contracts are unenforceable in a court of law due to the very nature of their illegality. However we assume that they are enforced by the agent's and the supervisor's willingness to report according to the side contract. In particular, the parties' willingness to adhere to the side contract is based on the threat that the benefits of the side contracts will be lost if one party breaches the agreement. There is strong reciprocity between the agent and the supervisor. When they both face a "good" state of nature, and when they have observed each other's type, reporting honestly would make both of them worse off. In contrast, by remaining silent, they both end up better off. Note that in contrast to monetary side payments that are paid ex post, here agents need to coordinate at the reporting stage, and after that there are no enforceability problems.[6] It is important to note that non-monetary side transfers are invisible. The transaction technology of the non-monetary transfers is extremely efficient; in fact, there are no actual transfers between the agent and the supervisor. What happens is a mutual explicit agreement of silence, in which

---

[5] Tirole (1988) reviews four possible categories of non-monetary transfers: human relations, acts of cooperation, supervision, and authority.

[6] For more about enforceable and self-enforcing side contracts, see Tirole (1992).

the benefits flow directly to the colluding parties.

Timing

Finally, the timing of the model is summarized as follows:

P designs a main contract   A reports r_1 S reports r_2 c_1 and c_2 realised   Payments w_1 and w_2

0-------------------1---------2-------------------3-----------4----------------5-------------------6------------------> time

A learns θ_1, S learns σ_1, S learns θ_2, A learns σ_2 side contracts   A chooses e_1 S chooses e_2

At stage 0, all parties learn relevant information. A and S learn their types and signals, and P learns the prior distributions of types. At stage 1, P designs a main contract which determines wages. At stage 2, A and S may secretly draw up an enforceable side contract. At stage, 3 A and S report r_1 and r_2 respectively to the principal. At stage 4, A and S choose effort levels. At stage 5, costs are realized; and finally, at stage 6 compensations are paid to A and S according to the terms of the main contract.

2.3 The Analysis

2.3.1 Organisation I: Principal-Productive Supervisor-Agent

In this section we consider a situation often observed in real-world organisations. The middle manager monitors the worker and reports to the principal, but the worker's information about his superior is not used. We begin our analysis with a case of perfect information (first best). The principal has full information about θ_1 and θ_2, and therefore he can always implement the first-best and optimal efforts e^* from both A and S. There is no need for monitoring.

The principal's problem is simply:

$$Max_{e_1,e_2}\ [R-(\theta_1-e_1)-w_1)-(\theta_2-e_2)-w_2)],\qquad (2.6)$$
$$s.t.\ w_i-\psi(e_i)\geq 0, i=1,2.$$

In the above, i=1 is the agent and i=2, the supervisor. A solution to (12) is $\psi'(e^*)=1$, $w_1=\psi(e^*)$ and $U_1=0$ for the agent and $\psi'(e^*)=1$, $w_2=\psi(e^*)$ and $U_2=0$ for the supervisor. We state the solution as a Lemma 2.1:

Lemma 2.1 Under perfect information the principal keeps both the agent and the supervisor at their reservation utility levels and induces efficient levels of effort e^* in all states of nature. The wage levels are independent of the state of nature.

Proof. The first-best results follow directly from the principal's maximization problem subject to IR constraints. □

We next analyse a benchmark case - *that of a collusion-free equilibrium* - in which collusion does not exist, and the supervisor behaves honestly. Once the principal offers the supervisor a normal second-best contract, he does not have any incentive to lie in his reports about the agent's type, because the reports do not effect his own utility at all. If the supervisor has learned the agent's type, he reports it truthfully to the principal. P then has perfect information about A's type, and there the solution coincides with the first best. If the supervisor has not learned the agent's type, then of course P also remains uninformed about it. The principal's problem is:

$$Max\ E\ [R-(\theta_1-e_1)-w_1-(\theta_2-e_2)-w_2]\qquad (2.7)$$
$$s.t.\ (2.1)-(2.4).$$

It is a standard result in contract theory that the individual rationality constraint is binding for the inefficient type and the incentive compatibility constraint is binding for the efficient type. Thus we can simplify the principal's problem and solve the relaxed problem with binding constraints. The binding IR and IC constraints are:

$$\overline{w_i}-\psi(\overline{\theta_i}-\overline{c_i})=0, \quad i=1,2. \tag{2.8}$$

$$\underline{w_i}-\psi(\underline{\theta_i}-\underline{c_i})=\overline{w_i}-\psi(\underline{\theta_i}-\overline{c_i})=\Phi(\overline{e_i}), \quad i=1,2. \tag{2.9}$$

In equation (2.9), $\Phi(\overline{e}_i)= \psi(\overline{e}_i) - \psi(\overline{e}_i - \Delta\theta)$ is a rent function with $\Phi'(\overline{e}_i) > 0$. It indicates that a rent enjoyed by the efficient type is an increasing function of the effort level required from the inefficient type. This function demonstrates effectively that there is an important trade-off between incentives and rent extraction.

The principal's problem is to maximize (2.7) subject to (2.8) and (2.9). We derive a solution which involves a combination of the first-best and second-best contracts for the agent in Appendix A. The supervisor's contract is a normal second-best contract. There is no need to compensate the supervisor for his reports about the agent's type, because the supervisor is hired and compensated for production anyway, and his utility is independent of reports. With the help of the supervisor the principal is able to reduce the asymmetric information between himself and the agent. The solution for (2.7) forms a collusion-free contract, which we refer to as Lemma 2.2:

Lemma 2.2 The collusion-free contract within Organisation I provides for the agent a combination of the first-best and second-best contracts. The supervisor's contract is a normal second-best contract.

Proof. See Appendix A.

Assume now that *non-monetary transfers between A and S can take place.* After the principal offers a main contract, A and S may secretly sign a side contract. When offered a collusion-free contract, A and S must figure out whether they can do better by colluding.

It follows immediately that the supervisor cannot do any better than with the collusion-free contract. The agent, however, has an incentive to collude; the supervisor's reports reduce his own utility. The agent prefers that the principal be uninformed about his type in good states of nature so that he can earn informational rent $\Phi(\overline{e}_1)$. However, the agent cannot provide any favours for the supervisor as compensation for favourable reports.

Therefore non-monetary collusion does not arise. The collusion-free contract as defined in Lemma 2.2 is also (trivially) a collusion-proof contract. Thus we have proposition 2.1:

Proposition 2.1 In Organisation I, the agent is unable to do any favours for the supervisor, and therefore collusion problems do not arise. The collusion-free contract coincides with a collusion-proof contract.

Proof. In Organisation I, the agent does not have discretion over the supervisor and thus, by definition, non-monetary side-transfers cannot be exchanged. The second part of the proposition follows immediately. ☐

Within Organisation I, the principal always asks for reports from the supervisor. This follows from the fact that he is not obliged to reward the supervisor separately for his reports because he is compensated for production anyway. The principal's expected profits within Organisation I are higher than in the usual second-best case without the supervisor's reports. To see this, note that the principal can always ignore the supervisor's information, and therefore he can do at least as well as in the second-best case. To see that he is strictly better off, note that within Organisation I, the principal acquires more detailed supervisory information about the agent's type without any additional costs. This implies that the principal must be better off within Organisation I. The conclusion is clear-cut. It is optimal for the principal to use the supervisor's information about the agent.

It is interesting to contrast this theoretical result with the actual behaviour of real-world organisations. As a general rule in most organisations, a superior monitors subordinates. Think, for example, of the hierarchial organisational behaviour in the armed forces, the civil service, and so on. Roughly speaking, they all have in common the fact that only the superiors' information about their subordinates is used, and no direct communication from bottom to the top takes place. As far as only non-monetary transfers between a middle manager and a worker are possible, that sort of monitoring and reporting pattern does not give rise to the possibility of collusion by exchanging favours.

2.3.2 Organisation II: Principal-Productive "Supervisors"

From now on, we allow for the possibility that a workers's information about his superior can also be used. It is interesting to compare this type of reporting behaviour with what happens in the real world, and, in particular to analyse, under which conditions the principal should also use the worker's information about his superior. Without the possibility of collusion, new, detailed information which the worker can provide about his superior would be used without hesitation to tighten the middle managers's incentive scheme. However, the possibility of collusion creates new costs which the principal must take into consideration when deciding whether or not to use two supervisory information sources. In short, the principal has to balance the benefits of using all supervisory information with the costs of collusion. Furthermore, the value of supervisory information provided by different members of an organisation may differ. In the next section we analyse how this will affect the optimal use of information under potential collusion.

The perfect-information case within Organisation II naturally coincides with the first-best of Organisation I. Supervisory information is not needed, and wages are independent of the state of nature. (See Lemma 2.1)

Consider next the *collusion-free case.* Suppose that the principal offers both A and S normal second-best contracts. Thus A and S earn informational rents when the state of nature is a good one, and they are kept on their reservation utility levels when the state of nature is a bad one. The key point concerning supervisory information is that, for example, the agent cannot increase his own welfare by concealing information he has learned about the supervisor. Thus, both the agent and the supervisor behave honestly, and P does not have to motivate them to report the supervisory information they hold. The principal is, however, ignorant of A's and S's types in the remaining states of nature. The principal's problem is:

$$\begin{aligned}&Max\ E\ [R-(\theta_1-e_1)-w_1-(\theta_2-e_2)-w_2)\\&\quad s.t.\ (2.8)-(2.9).\end{aligned} \tag{2.10}$$

With the binding constraints the above problem is easy to solve. The solution to the principal's problem in the collusion-free case is simply a combination of the first- and second-best contracts for both the supervisor and the agent. A detailed solution is stated in Appendix A. The collusion-free contract within Organisation II is merely a replica of the collusion-free contract within Organisation I, except that here the supervisor's wage schedule is also tailored to the supervisory information the agent provides. The collusion-free contract within Organisation II is stated as Lemma 2.3:

Lemma 2.3 The Collusion-free contract within Organisation II provides for both the agent and the supervisor a combination of the first-best and second-best contracts.

Proof. See Appendix A.

Assume now that *A and S can collude,* and the collusion-free contract defined in Lemma 2.3 is offered to them. The question then becomes, can they do any better by colluding and manipulating the supervisory information they have? Clearly for both of them there are strictly positive gains to be realized by coordinating their reports. Namely, they both would like to keep the principal uninformed when they face a good state of nature, and thus earn informational rents instead of being kept on their reservation utility level. Recall that in Organisation I, A could not compensate S for his favourable reports in any way, and in particular he could not do any favours for him. Now within Organisation II, both A and S can compensate each other's favours with counter favours.

From the principal's point of view giving discretion also to A creates a new way to acquire more detailed supervisory information. It also introduces a way to form a side contract, which takes the form of "a favour and counter favour" . To put it more simply, an explicit, non-monetary side contract is simply the following agreement:" When we both face a good state of nature, and if you do not report my type, neither will I report your type". This side contract has a monetary equivalent, which of course equals the rents the agent and the supervisor are able to earn in a good state of nature. It is important to realize here that the rents just flow to A and S, and there are no actual transfers between A and S.

In contrast to monetary side-contracts, non-monetary collusion can occur only in one state of nature: namely when both A and S are efficient and perfectly observe each other's type.[7] Note that this is the only state of nature when favours and counter favours take place simultaneously, which facilitates profitable non-monetary collusion.[8]

The principal's problem is that he cannot distinguish when A and S are colluding and when they actually have not learned each other's types, because they can send the same messages in both cases. In particular, in a state of collusion, they are able to jointly conceal information from the principal.

To prevent side contracts the principal has to pay S and A as much as they would gain from not releasing supervisory information about each other - the principal must match the gains that result from collusion. In short, in order to prevent collusion the principal must provide both A and S informational rents in the state of nature when they are able to collude. The principal has to respect the following coalition incentive constraint

$$[w_1-\psi(e^*)+w_2-\psi(e^*)]\geq[\underline{w_1}-\psi(\underline{e_1})+\underline{w_2}-\psi(\underline{e_2})]. \tag{2.11}$$

The left hand side of equation (2.11) states A's and S's utility when they truthfully reveal supervisory information ($r_1= \underline{\theta}_1$, $r_2= \underline{\theta}_2$), and the right hand side states their utility when they conceal it ($r_1=r_2=0$). The principal has to respect the above coalition incentive constraint by rewarding the agent and the supervisor such that they are as well off as when colluding. When $w_i = \psi(e^*) + \Phi(\hat{e}_i)$, i=1,2 A and S have no reason to collude at all, and then collusion does not arise in equilibrium.

It is important to note here that in all other states of nature, the agent and the supervisor will report honestly because non-monetary side transfers cannot be exchanged. And, more importantly, there the agent and the supervisor can not increase their own welfare by concealing the supervisory information they hold. The principal's problem under non-monetary collusion is:

---

[7] Monetary side-contracts would also arise when only one of the colluding parties faces a good state of nature, because there the monitored party would buy the monitor's silence with money.

[8] Collusion in other cases would require repetition (dynamic collusion). Recently Martimort (1996a) has developed a model of dynamic collusion along this line of thought.

$$\begin{aligned}&Max\ E\ (R-(\theta_1-e_1)-w_1)-(\theta_2-e_2)-w_2)\\&s.t.\ (2.8),(2.9),(2.11).\end{aligned} \qquad (2.12)$$

The detailed solution is derived in Appendix A. The solution for the principal's problem forms a collusion-proof contract, which we state as Lemma 2.4:

Lemma 2.4 The collusion-proof contract within Organisation II provides for both the agent and the supervisor a combination of first-best and second-best contracts. Furthermore $\hat{e}_i < \bar{e}_i$, *and* $w_i = \psi(e^*) + \Phi(\hat{e}_i)$ *in the case when* $\{\underline{\theta}_1 \underline{\theta}_2\}$ *have been reported,* $i=1,2$, $j=1,2$, *and* $i \neq j$. *Otherwise the optimal collusion-proof contract is similar as the independent collusion-free contract characterized in Lemma 2.3.*

Proof. See Appendix A.

The optimal collusion-proof contract involves further distortion ($\hat{e}_i$) in the inefficient type's effort as compared to the collusion-free case. This further distortion in the inefficient type's effort is introduced due to the possibility of collusion in order to reduce the informational rents the principal has to leave for A and S. Of course, this further distortion increases the ex post inefficiency of the collusion-proof contract.

The general properties of the collusion-proof contract are similar to those in Laffont and Tirole (1991), where the authors derive a collusion-proof contract when a regulator and a firm collude by transferring money. Here collusion technology is more efficient, since collusion by exchanging favours does involve any transaction costs. The principal optimally rewards the supervisor and the agent for their supervisory information only in the case when they could have colluded. This, however, occurs here less often than in Laffont and Tirole (1991), since monetary bribes are not feasible here.[9] Therefore, the principal pays less often for supervisory information provided by the agent and the supervisor. The difference in the optimal collusion-proof contract compared to that of Laffont-Tirole (1991) is purely due to the mode of collusion we choose to consider.

---

[9] In fact, this occurs with probability $p^2 l^2$.

When deciding whether to use one or two supervisory information sources, the principal has to balance with a trade-off, which arises from the benefits and costs of using two information sources. To judge whether it is profitable to use one or two supervisory information sources, one has to compare the principal's expected profits under Organisations I and II. The question then is, should all available supervisory information be used?

2.3.3 What is the Optimal Organisation?

In the above we come to the conclusion that the decision as to whether or not to use all supervisory information available depends on both the benefits one can accrue by using that information and on the costs of collusion which arise immediately when two information sources are combined. So far, we have assumed that the value of supervisory information the agent and the supervisor hold is equal. From here on the value of their information may differ. When the value of supervisory information differs, some comparative statics are required to see how this affects the optimal use of information within organisations. We start with the case where the supervisor and the agent are symmetric, and in the second part we allow asymmetry. When we analyse the effects of asymmetry on the optimal organisation, we fix p=l=1/2. Given this restriction, we carry out our analysis and consider how the possible asymmetry $(\Delta\theta_1 - \Delta\theta_2) \geq$ affects which organisation ends up being the most profitable one.[10]

Symmetric Supervisor and Agent

This section analyses the symmetry between the supervisor and the agent. By symmetry we refer to random variables θ_1 and θ_2, which are identically distributed: that is, they have equal supports, $\Delta\theta_1 = \Delta\theta_2$. In terms of the value of supervisory information, this means that the monetary value of the supervisor's and the agent's reports is equal. Consider the benchmark case of no collusion. Then the following result follows immediately:

---

[10] Alternatively, we could have fixed the size of the supports and done comparative statics with respect to p and l.

Proposition 2.2 Without the problem of collusion, and when the agent and the supervisor are symmetric and report honestly, Organisation II will always dominate Organisation I. The principal's expected profits with a collusion-free (cf) contract are always higher with Organisation II: $E\Pi_{IIcf} > E\Pi_{Icf}$

Proof. To prove this we need only two steps. Note first that in Organisation II, the principal can ignore the agent's information about the supervisor (i.e., not ask the agent to report supervisory information). This means that in OII the principal can do at least as well as he can in OI. To see that in fact the principal does strictly better, note that in OII the principal can acquire more detailed supervisory information at no cost. Therefore, the principal is strictly better off under Organisation II. ☐

The economics behind the above result is that with the aid of honest reports, the principal eliminates the rents and ex post inefficiency which the supervisor's second-best contract without A's reports would include. Does the possibility of collusion change which organisation ends up being the dominant one? It can be shown that Organisation II dominates also under collusion, when symmetry between the agent and the supervisor exists. Thus we have Proposition 2.3:

Proposition 2.3 Under symmetry, it is always optimal to use every single piece of available supervisory information even if the agent and the supervisor are able to collude. The principal's expected profits with a collusion-proof (cp) contract are higher in Organisation II: $E\Pi_{IIcp}>E\Pi_{Icp}$

Proof. See Appendix A.

Under symmetric agent and supervisor, costs related to collusion are always lower than the benefits of new, detailed information in OII. Organisation II dominates Organisation I with or without collusion in the case where the agent and the supervisor are symmetric. Therefore, it is optimal for the principal to ask for reports from both the agent and the supervisor.

Asymmetry between Supervisor and Agent

From now on we shall focus on the case in which the agent and the supervisor are no longer symmetric. In particular, the supervisor's fine-tuning production task includes smaller variation amongst random variables than does the agent's main production task, that is, $\Delta\theta_1 > \Delta\theta_2$. Basically we are curious to know how this asymmetry, typically observed in real-world organisations and firms, affects the optimal use of supervisory information. Another interpretation of this difference is that in general the quality of the supervisor's information is higher. In other words, it is important to keep in mind that the monetary value of supervisory information may differ. Note that asymmetry does not change anything in the case of no collusion, and Organisation II is still the dominant one.

Now it is interesting to look at whether the possibility of non-monetary collusion changes which organisation ends up being the most profitable one. To this end it is necessary to compare the principal's expected profits under both Organisations I and II over all possible values of $(\Delta\theta_1 - \Delta\theta_2)$. Comparing the principal's expected profits leads to Proposition 2.4:

Proposition 2.4 In the case of asymmetry between the supervisor and the agent, which organisation ends up being dominant depends on the difference $(\Delta\theta_1 - \Delta\theta_2)$. *In particular,*

(i) If $0 < (\Delta\theta_1 - \Delta\theta_2) < k^*$, *then it is always better to use every single piece of available supervisory information even if the agent and the supervisor are able to collude. The principal's profits under the collusion-proof contract are higher in OII:* $E\Pi_{IIcp} > E\Pi_{Icp}$.

(ii) If $(\Delta\theta_1 - \Delta\theta_2) = k^*$, *then the principal's expected profits under collusion-proof contracts are equal in Organisations I and II.* $E\Pi_{IIcp} = E\Pi_{Icp}$.

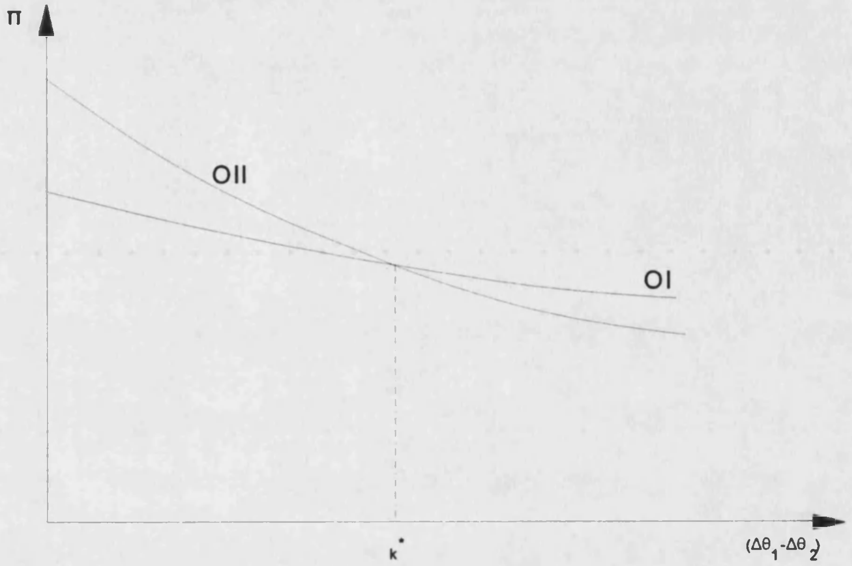
(iii) If $(\Delta\theta_1 - \Delta\theta_2) > k^*$, *it is optimal to use only the supervisory information the supervisor hold about the agent. Therefore Organisation I will dominate Organisation II. With a collusion-proof contract the principal's expected profits with a collusion proof contract are higher in Organisation I.* $E\Pi_{IIcp} < E\Pi_{Icp}$

Proof. See Appendix A.

The economics behind Proposition 2.4, which is further clarified in Figure 2.1, is quite straightforward. The proposition demonstrates that the decision as to whether or not to use all supervisory information is endogenously determined. This, in turn, determines the organisation mode firms adopt in different environments. When making the decision whether to use the additional supervisory information a worker can provide, an organisation designer has to balance the benefits of that information and the costs of collusion which arise immediately when the second information source is used. This trade-off defines which organisational mode is the optimal one.

The first part of proposition 2.4 shows that when the agent and the supervisor are not "too asymmetric" it is optimal to adopt the mode of Organisation II. This follows simply from the fact that the benefits of new detailed information are greater than the costs of potential collusion in OII. However, when asymmetry increases, Organisation I may provide greater profits. When the asymmetry is greater than threshold value k^*, the expected profits under OI are higher, and this is because the costs of collusion are high enough to offset any gain from new, detailed information in Organisation II. Therefore under this regime Organisation I dominates Organisation II. The intuition behind the domination result is straightforward. When the value of the agent's supervisory information about the supervisor decreases (i.e., when $(\Delta\theta_1 - \Delta\theta_2)$ increases), the gains from new additional information decrease, and evidently the costs of collusion becomes higher than the benefits. Therefore, it is optimal to break collusion by ignoring the agent's supervisory information by adopting Organisation II.

Figure 2.1 Expected profits in Organisations I and II. (p=l=1/2)



The results indicate that reporting patterns in all organisations are endogenously determined when the members of those organisations are able to exchange non-monetary side-transfers. When collusion is not an issue, all available supervisory information should be used. However, when collusion may arise, it is shown that there are instances when it is optimal for the principal not to listen to the worker at the bottom of the hierarchy: not only is his supervisory information less valuable, but also listening to the worker raises the possibility that harmful side-trade will take place. In addition, the model gives a precise prediction who should be a middle-manager. The supervisor should be an agent who has on his possession the most valuable supervisory information, and under some cases he should not be monitored. This is a nice result, since it says that the supervisor will perform his task better if he is not monitored by anyone. In sum, not only the decision whether or not it is optimal to use all available supervisory information is endogenous, but so is the decision who should be a supervisor in the first place.[11]

The main result of this chapter is that in some environments it is optimal for the principal to ignore intentionally some additional supervisory information, because by doing

---

[11] See also Aghion's and Tirole's (1996) discussion of the delegation of authority in an incomplete contract framework. They conclude that authority should be delegated to the agent whose preferences are most congruent with the principal's.

so he can get some other and more important supervisory information without further cost. Thus, it may be optimal for the principal to commit not to use the workers' information about their superiors so as to deter non-monetary collusion and guarantee that the superiors will truthfully report their observations of their subordinates' types. This will be achieved only if Organisation I is adopted.This result may also explain the empirical fact that if subordinates' information about their superiors is used at all, it is typically carried out in such a manner that their reports are kept secret, ensuring the reporting subordinate's anonymity. This in turn precludes the exchange of non-monetary side transfers. In particular, the colluding parties cannot condition side contracts on observable and verifiable reports.

2.4 A Discussion: The Economics and Some Interpretations of Organisations I and II

In this section we offer some economic interpretations of the organisational forms and analyses carried out earlier. We concern ourselves here with a discussion of task assignments within organisations and vertical integration versus non-integration. Our focus on these situations relates to the optimal use of information when members of an organisation are able to exchange non-monetary side transfers.

Task Assignment in Organisations

Since the days of Adam Smith, the gains from specialization in general have been put forth as a main reason for the division of labour in economies. This applies to the organisation as well as to the firm. The workers typically perform specific tasks, and task sharing (or task overlapping) is not implemented as often as is technically possible. In this section we hope to demonstrate convincingly that there may be some other reasons for specialization as well that are related to the agents' opportunities to exchange non-monetary side transfers.

Suppose that, in addition to paying attention to the normal incentive considerations, the principal has the option to design agents' jobs as well. In the analysis above we have shown that non-monetary collusion problems do indeed arise when the supervisor and the agent are able to exchange favours. In this context an interesting question that presents itself is whether or not the principal can design the agents' jobs in such a way that non-monetary side transfers cannot be exchanged.

The solution takes the form of a specialization mode, where both production tasks are carried out by the agent and both monitoring tasks are carried out by the supervisor.[12] This solution automatically avoids any non-monetary collusion problems. Alternatively, a collusion-free outcome is achieved when the principal hires four agents: two to produce and two to monitor them. The worker who produces cannot do any favours for the monitor; thus, by definition, non-monetary side-transfers are unexchangeable. The monitoring party needs no motivation to report truthfully; he does so willingly when he is paid at least his reservation wage. It is important to note, however, that compensation for supervisory tasks can not necessarily be set equal to zero as in our analysis. This is due to the fact that the supervising agent's only task is to monitor, and the wage depends on his reservation wage. Therefore, when the supervisory wage can be set relatively low, we know that the specialization mode is called for as the optimal task-assignment mode.

Specifically, through the specialization the principal can reap the full benefits of additional supervisory information in addition to being able to avoid any of the costs of collusion. In other words, the principal is able at last to reach a performance level equivalent to that of Organisation II in the collusion-free case. Here task assignment also works as an effective incentive device against collusion.[13] By separating production and supervisory tasks, the principal makes the existence of non-monetary collusion impossible.

It is interesting to expand this last observation to a broader context. These results may help us to understand how it is that the often observed hierarchial organisation mode can be the optimal solution to the problem of non-monetary collusion between the members of that

---

[12] Here we assume that the effort required to perform one production task is independent of the effort required to perform the other production task. In addition this requires that the aggregate informational rents equal to the sum of informational rents in a case in which production tasks are carried out by the separate agents.

[13] The above observation goes along with Holmström and Milgrom (1991)'s task-grouping result. The reason behind their results is not, however, connected to collusion problems.

hierarchy. Namely, designing production and supervisory tasks in an optimal way eliminates all collusion problems as long as monetary side-transfers are excluded, as was assumed here. Taken together, this implies a hierarchy with a principal at the top, a supervisor in the middle, and a worker at the bottom. Hierarchial structures are often characterized as inefficient. Here, however, a standard hierarchial solution turns out to be the optimal response to the chronic problem of non-monetary collusion. The possibility that members of an organisation can exchange non-monetary favours may be an important factor in determining how different tasks should be assigned between workers.

Vertical Integration vs. Non-Integration

Suppose the supervisor and the agent are two different units of the firm. These units can be under unified control - an integrated firm - or under separate control - a non-integrated firm. Both units have some private information which is unavailable to headquarters (a principal).

The non-integrated firm is such that it is comprised of a principal and one unit. Then the principal contracts with a separate supplier as well. These units may be able to learn each other's type even if they are not under unified control, due to the existence of long-term relationships among themselves. The units, however, may be unable to provide verifiable reports of each other's types under non-integration, because verifiability requires access to the other firm's records. Basically, the principal of the non-integrated firm has no way to delegate supervisory authority to his own unit manager such that he would have access to the separate supplier's accounting records. In other words, the separate supplier is not obliged to provide access to its accounting records, guaranteeing thereby informational rents due to the asymmetric information between the separate supplier itself and the non-integrated firm.[14]

The principal of a non-integrated firm can delegate supervisory authority to the separate supplier in such a way that monitoring the non-integrated firm's own unit manager is possible. The separate supplier can then be regarded as an external auditor. The separate

---

[14] Note that here we analyse the case in which the separate supplier is an owner-managed firm. If we had also assumed that the separate supplier was also comprised of a principal and an agent, then the principals could always write a contract which would allow cross-checking between the firms. However, then the whole distinction between the firms would become irrelevant.

supplier has no incentive to lie if it learns something about the non-integrated firm's unit manager; the supplier's own utility is independent of the non-integrated firm's unit manager's utility. Furthermore, the is compensated for supplying, no matter what.

Under non-integration, non-monetary collusion problems do not arise. The principal hires a unit manager and separate supplier of his own choice as specified in Organisation I. Here the separate supplier is also a supervisor. Under non-integration, the best the principal can achieve is the performance of Organisation I.

Suppose now that the non-integrated firm buys up the separate supplier. The integrated firm is then comprised of a principal and two units with which the principal contracts. It is important to note that under integration, the principal can delegate supervisory authority to both of his unit managers such that they are able to compile verifiable reports. This is because the principal now has authority over the decision as to whether or not access to the firm's accounting records is permitted. Note that in this present interpretation both Organisation I and II's forms characterize an integrated firm, and the only distinction between them is the use of supervisory information. If in the integrated firm the principal asks his unit managers to monitor each other, they may exchange non-monetary side transfers, and thus collusion problems do arise as shown earlier.

In general, under integration more verifiable supervisory information is available to the principal. However, the actual value of this information depends on whether collusion between the unit managers is possible. Within the present framework, the integrated firm adopts the organisation form of either I or II. The vertically-integrated firm can do at least as well as the non-integrated firm, because it can always ignore any additional supervisory information, thereby mimicking the non-integrated firm.

Our discussion here is quite closely related to what is known in the literature as influence activities. Milgrom and Roberts (1988) discuss the types of influence activities which can arise in organisations under centralized authority. There, members of the firm can try to influence decision makers in order to gain personally. These influence activities are costly to firms in at least two ways: they skew the decision-making process, because the information decision makers use may be biased. Secondly, members who try to influence decision makers use up lots of working time and effort advancing their own agendas. Milgrom and Roberts go on to argue that the existence of influence activities may be an important reason why some firms' vertical integration decisions end up being the wrong ones. Their theory explains why firms are quite often observed selling their loss-making additional units. They conclude that if these loss-making units were part of the integrated firm, they would spend by far too much time and effort lobbying the decision makers to save them.

Our model points out some factors which may affect firms' integration decisions because it allows for the costs and benefits of bringing together two separate units under unified control to be considered simultaneously.[15] The main relation our model has to the vertical-integration literature is the existence of integration costs which are incurred when a non-integrated firm becomes integrated. In vertical integration literature these costs are not explicitly modelled, but are treated as exogenous with a brief mention that they are caused by some sort of incentive problems. In the present paper, we regard the costs of collusion as endogenous integration costs which occur when a firm integrates, and the firm's unit managers can exchange non-monetary side transfers.

2.5 Concluding Remarks

In this chapter we have analysed the optimal organisation modes and the optimal use of supervisory information under circumstances of potential collusion. The chapter proposes a simple theory of non-monetary collusion, where the members of an organisation collude by simultaneously exchanging favours. Collusion is possible only when all available information is used, and the principal has to decide whether to use only one information source and avoid collusion altogether, or use two information sources and bear the costs of collusion.

It has been shown that under circumstances of collusion it may be optimal not to use all available information. In particular, when the agent and the supervisor are asymmetric enough, it may be optimal not to listen to the worker at the bottom of a hierarchy, since that prevents collusion and the supervisor provides more valuable supervisory information at no further cost. The main result of the chapter is not only that the use of information is

---

[15] Note that the present model cannot, however, answer the ultimate question, which determines a firm's boundaries.

endogenously determined, but so is the decision to whom the supervision task should be assigned. The supervisor should be an agent who has on his possession the most valuable information. An interesting result is that the supervisor performs his supervision task better when he is not monitored by some third party. On the whole, the chapter presents us with general guidelines about whose information in a real-world organisation should and should not be used, and the simple reason why all available supervisory information is typically not used.

We embrace the concept of hard supervisory information, which is clearly theoretically questionable. Undoubtedly, it would be better to use soft information, which is completely manipulable. An interesting point, but one which lies outside the scope of this paper, would be to analyse whether the soft information the agent and the supervisor have about one another has any value to the principal. Intuitively, it is less useful; and the question then becomes, does it have any value to the principal at all?

The issues of task assignment deserve closer analysis than they receive here. It seems that it is possible to design tasks in such a way that non-monetary collusion simply does not arise. The challenge for an economist is to figure out whether it is possible to assign tasks in such a way that monetary side transfers can be deterred as well.

Appendix A

All possible "states of nature"

1. A and S observe $\underline{\theta}_1$, $r_2=\{\underline{\theta}_1,\varnothing\}$ S and A observe $\underline{\theta}_2$, $r_1=\{\underline{\theta}_2,\varnothing\}$
2. A and S observe $\underline{\theta}_1$, $r_2=\{\underline{\theta}_1,\varnothing\}$ S observes $\underline{\theta}_2$, A observes $\varnothing$, $r_1=\{\varnothing\}$
3. A and S observe $\underline{\theta}_1$, $r_2=\{\underline{\theta}_1,\varnothing\}$ S observes $\bar{\theta}_2$, A observes $\varnothing$, $r_1=\{\varnothing\}$
4. A and S observe $\underline{\theta}_1$, $r_2=\{\underline{\theta}_1,\varnothing\}$ S and A observe $\bar{\theta}_2$, $r_1=\{\bar{\theta}_2,\varnothing\}$
5. A observes $\underline{\theta}_1$, S observes $\varnothing$, $r_2=\{\varnothing\}$ S and A observe $\underline{\theta}_2$, $r_1=\{\underline{\theta}_2,\varnothing\}$
6. A observes $\underline{\theta}_1$, S observes $\varnothing$, $r_2=\{\varnothing\}$ S observes $\underline{\theta}_2$, A observes $\varnothing$, $r_1=\{\varnothing\}$
7. A observes $\underline{\theta}_1$, S observes $\varnothing$, $r_2=\{\varnothing\}$ S observes $\bar{\theta}_2$, A observes $\varnothing$, $r_1=\{\varnothing\}$
8. A observes $\underline{\theta}_1$, S observes $\varnothing$, $r_2=\{\varnothing\}$ S and A observe $\bar{\theta}_2$, $r_1=\{\bar{\theta}_2,\varnothing\}$
9. A observes $\bar{\theta}_1$, S observes $\varnothing$, $r_2=\{\varnothing\}$ S and A observe $\underline{\theta}_2$, $r_1=\{\underline{\theta}_2,\varnothing\}$
10. A observes $\bar{\theta}_1$, S observes $\varnothing$, $r_2=\{\varnothing\}$ S observes $\underline{\theta}_2$, A observes $\varnothing$, $r_1=\{\varnothing\}$
11. A observes $\bar{\theta}_1$, S observes $\varnothing$, $r_2=\{\varnothing\}$ S observes $\bar{\theta}_2$, A observes $\varnothing$, $r_1=\{\varnothing\}$
12. A observes $\bar{\theta}_1$, S observes $\varnothing$, $r_2=\{\varnothing\}$ S and A observe $\bar{\theta}_2$, $r_1=\{\bar{\theta}_2,\varnothing\}$
13. A and S observe $\bar{\theta}_1$, $r_2=\{\bar{\theta}_1,\varnothing\}$ S and A observe $\underline{\theta}_2$, $r_1=\{\underline{\theta}_2,\varnothing\}$
14. A and S observe $\bar{\theta}_1$, $r_2=\{\bar{\theta}_1,\varnothing\}$ S observes $\underline{\theta}_2$, A observes $\varnothing$, $r_1=\{\varnothing\}$
15. A and S observe $\bar{\theta}_1$, $r_2=\{\bar{\theta}_1,\varnothing\}$ S observes $\bar{\theta}_2$, A observes $\varnothing$, $r_1=\{\varnothing\}$
16. A and S observe $\bar{\theta}_1$, $r_2=\{\bar{\theta}_1,\varnothing\}$ S and A observe $\bar{\theta}_2$, $r_1=\{\bar{\theta}_2,\varnothing\}$

Proof of Lemma 2.2

Under a collusion-free regime the principal offers a normal second-best contract to the supervisor, and thus the supervisor has no incentive to lie when reporting about the agent's type. Therefore, the principal acquires the supervisor's information about the agent's type in states 1-4 and 13-16, and there the first-best solution $\psi'(e^*)=1$ and $w=\psi(e^*)$ will apply. In the remaining states of nature, the principal knows only the priors of the agent's type. Since the principal does not learn any additional information about the supervisor, he knows only the priors of the supervisor's type. The principal's problem is thus reduced to two independent programs: one for the agent and one for the supervisor.

We know from the basic result of adverse-selection models that at the optimum the IR constraint of the inefficient type and the IC constraint of the efficient type are binding: $\bar{w}_1$-$\psi(\bar{e}_1)$=0 and $\underline{w}_1$-$\psi(\underline{e}_1)$=$\bar{w}_1$-$\psi(\bar{e}$-$\Delta\theta)$. If they were not binding, the principal could increase his expected profits by offering a lower wage to the inefficient type without violating his IR constraint, and thus the original solution would not be the optimum one. We use this result to solve the principal's problem. This relaxed problem is solved by taking binding constraints into account. Since the two programs are independent, we only consider the principal's problem concerning the agent:

$$\begin{aligned} Max_{\underline{e}_1, \bar{e}_1} \; & p[R-(\underline{\theta}_1-\underline{e}_1)-\psi(\underline{e}_1)-\psi(\overline{e_1})+\psi(\overline{e_1}-\Delta\theta_1) \\ & +(1-p)\;[R-(\overline{\theta_1}-\overline{e_1})-\psi(\overline{e_1})] \end{aligned} \tag{A.1}$$

The first-order conditions with respect to $\underline{e}_1$ and $\bar{e}_1$ are as follows:

$$=p(1-\psi'(\underline{e}_1))=0 \tag{A.2}$$

$$\begin{aligned} & =p(-\psi'(\overline{e_1})+\psi'(\overline{e_1}-\Delta\theta_1)) \\ & \;+(1-p)\;(1-\psi'(\overline{e_1}))=0. \end{aligned} \tag{A.3}$$

The first-order conditions for the supervisor are calculated in a similar fashion, but for the sake of brevity we do not include them here. With the help of the first-order conditions we can now solve for the optimal effort levels and wages in all states of nature. First, the principal has perfect information about the agent's type in states 1-4 and 13-16, and then the first-best solution will apply : $\psi'(e^*)$= 1, w_1=$\psi(e^*)$, and U_1=0. In stating the results, we use the following notation: $\Phi(\bar{e})$=$\psi(\bar{e})$-$\psi(\bar{e}$-$\Delta\theta)$. In the good states of nature (5-8), the solution is $\psi'(\underline{e}_1)$= 1, $\underline{w}_1$=$\psi(\underline{e}_1)$+$\Phi(\bar{e}_1)$, $\underline{U}_1$=$\Phi(\bar{e}_1)$; and in the bad states of nature (9-12), the solution is $\psi'(\bar{e}_1)$=1- (p_1 $\Phi'(\bar{e}_1)$)/(1-p_1), $\bar{w}_1$=$\psi(\bar{e}_1)$, $\bar{U}_1$=0.

The solution for the supervisor is a normal, second-best contract. In the good states of nature (1, 2, 5, 6, 9, 10, 13, 14) the solution is $\psi'(\underline{e}_2)$= 1, $\underline{w}_2$=$\psi(\underline{e}_2)$+$\Phi(\bar{e}_2)$, $\underline{U}_2$=$\Phi(\bar{e}_2)$, and in the bad states of nature (3, 4, 7, 8, 11, 12, 15, 16) the solution is $\psi'(\bar{e}_2)$=1- (p_2 $\Phi'(\bar{e}_2)$)/(1-p_2), $\bar{w}_2$=$\psi(\bar{e}_2)$, $\bar{U}_2$=0.

It follows from the convexity of disutility function $\psi(e_1)$ that $\bar{e}_1 < e^* = \underline{e}_1$, and so $\bar{w}_1 < w_1 < \underline{w}_1$. Similarly $\bar{e}_2 < \underline{e}_2 = e^*$, and $\bar{w}_2 < \underline{w}_2$. These wage and effort levels form the collusion-free contract defined in Lemma 2.2.

We still must make sure that our solution does not violate the ignored efficient type's IR and inefficient type's IC constraints. To check this we must simply substitute the optimal effort and wage levels to the ignored constraints. It follows immediately that the efficient type's IR constraint is never binding. The inefficient type will never mimic an efficient type, because this would push his utility below zero. The inefficient type's IC constraint never binds. □

Proof of Lemma 2.3

Under a collusion-free regime, the principal offers the supervisor and the agent normal, second-best contracts, and he inherits the supervisory information which the supervisor and the agent hold concerning one another. The principal has perfect information about the agent's type in states 1-4 and 13-16 and the supervisor's type in states 1, 4, 5, 8, 9, 12, 13, and 16. Thus, in those states of nature, the first-best solution will apply. In the remaining states of nature, the principal's problem is reduced to the normal adverse-selection problem. The solution is identical to that presented above, with the distinction that now the principal is also able to tighten the supervisor's incentive scheme with the supervisory information provided by the agent.

As a space-saving measure we do not state the first-order conditions here, but skip ahead directly to the characterization of the collusion free-contract within Organization II. For an agent in Organization II, a collusion-free contract includes $\psi'(e^*)= 1$, $w_1=\psi(e^*)$, and $U_1=0$ in the perfect-information states (1-4 and 13-16), $\psi'(\underline{e}_1)= 1$, $\underline{w}_1=\psi(\underline{e}_1)+\Phi(\overline{e}_1)$, $\underline{U}_1=\Phi(\overline{e}_1)$ in good states of nature (5-8), $\psi'(\overline{e}_1)=1- (p_1\ \Phi'(\overline{e}_1))/(1-p_1)$, $\overline{w}_1=\psi(\overline{e}_1)$, $\overline{U}_1=0$ in bad states of nature (9-12).

For the supervisor the solution is $\psi'(e^*)= 1$, $w_1=\psi(e^*)$, $U_1=0$ and $\psi'(\underline{e}_2)= 1$ in perfect information states (1, 4, 5, 8, 9, 12, 13, 16), $\underline{w}_2=\psi(\underline{e}_2)+\Phi(\overline{e}_2)$, $\underline{U}_2=\Phi(\overline{e}_2)$ in good states of nature (2, 6, 10, 14), and $\psi'(\overline{e}_2)=1- (p_2\ \Phi'(\overline{e}_2))/(1-p_2)$, $\overline{w}_2=\psi(\overline{e}_2)$, $\overline{U}_2=0$ in the bad states of nature (3, 7, 11, 15).

Similarly as above $\overline{e}_1 < e^* = \underline{e}_1$, and $\overline{w}_1 < w_1 < \underline{w}_1$. And of course also $\overline{e}_2 < \underline{e}^* = \underline{e}_2$ and $\overline{w}_2 < w_2 < \underline{w}_2$. These results give rise to Lemma 2.3. □

Proof of Lemma 2.4

If the principal offers the collusion-free contract derived in Lemma 2.3, the supervisor and the agent can do better by colluding and manipulating supervisory information. They can profitably collude when they perfectly observe each other's type. In state 1 when both the agent and the supervisor are efficient and have observed each others' types, they can coordinate and send the same messages as in state 6. By concealing supervisory information, they have learned they are able to enjoy informational rents. The principal's problem is that he cannot distinguish between cases 1 and 6.

In all other cases, when the agent and the supervisor observe each other's type, they will report it honestly, because the agent or the supervisor cannot increase his own utility by concealing supervisory information. Thus, collusion is an issue only in case 1, when the supervisor and the agent are able to exchange favours, and that occurs with probability p^2l^2. Due to potential collusion, the principal must motivate the supervisor and the agent to report honestly their supervisory information in state 1. The principal must compensate the supervisor and the agent for their supervisory information such that they are equally well off as when colluding.The compensation for supervisory information is determined by the binding coalition incentive constraint:

$$U_1+U_2 \geq \underline{U_1}+\underline{U_2}. \tag{A.4}$$

In other words,

$$[w_1-\psi(e^*)+w_2-\psi(e^*)] \geq [\underline{w_1}-\psi(\underline{e_1})+\underline{w_2}-\psi(\underline{e_2})]. \tag{A.5}$$

In the equation above, the left-hand side shows the agent's and the supervisor's utility in state 1, when (r_1= $\underline{\theta}_1$, r_2= $\underline{\theta}_2$). The right-hand side shows their utility in state 6, when they have not actually learned (or they have concealed) supervisory information, (r_1=r_2=0). Note that both U_1 and U_2 go to U^*=0 when the supervisor and the agent report honestly. To make the supervisor and the agent report truthfully and not collude, it has to be the case that

$$\begin{gathered} U_i=\underline{U_i}=\Phi(\overline{e_i})=\psi(\overline{e_i})-\psi(\overline{e_i}-\Delta\theta_i),\ i=1,2, \\ when[r_1=\underline{\theta_2},r_2=\underline{\theta_1}]\ reported. \end{gathered} \tag{A.6}$$

When the principal satisfies (A.6), he has perfect information about the agent's type in cases 1-4, 13-16, and the supervisor's type in cases 1, 4, 5, 8, 9, 12, 13 and 16. In other cases asymmetric information exists, and there the efficient type's IC constraint and the inefficient type's IR constraint are binding. Since non-monetary collusion can take place only in one case, the coalition incentive constraint has to be satisfied with probability p^2l^2. The principal's problem can be considered in two parts: one for the agent and one for the supervisor, which are identical; and thus we calculate explicitly only the agent's optimal contract. The principal's problem is:

$$\begin{aligned}
Max_{\underline{e}_1,\overline{e}_1}\quad & p^2l^2[R-(\underline{\theta}_1-e^*)-\psi(e^*)-\Phi(\overline{e}_1)]\\
&+p^2l(1-l)\,[R-(\underline{\theta}_1-e^*)-\psi(e^*)]\\
&+p(1-p)\,l(1-l)\,[R-(\underline{\theta}_1-e^*)-\psi(e^*)]\\
&+p(1-p)\,l^2[R-(\underline{\theta}_1-e^*)-\psi(e^*)]\\
&+p^2(1-l)\,l[R-(\theta_1-\underline{e}_1)-\psi(\underline{e}_1)-\Phi(\overline{e}_1]\\
&+p^2(1-l)^2[R-(\theta_1-\underline{e}_1)-\psi(\underline{e}_1)-\Phi(\overline{e}_1]\\
&+p(1-p)\;(1-l)^2[R-(\theta_1-\underline{e}_1)-\psi(\underline{e}_1)-\Phi(\overline{e}_1]\\
&+p(1-p)\;(1-l)\,l[R-(\theta_1-\underline{e}_1)-\psi(\underline{e}_1)-\Phi(\overline{e}_1]\\
&+(1-p)p(1-l)\,l[R-(\overline{\theta}_1-\overline{e}_1)-\psi(\overline{e}_1)]\\
&+(1-p)p(1-l)^2[R-(\overline{\theta}_1-\overline{e}_1)-\psi(\overline{e}_1)]\\
&+(1-p)^2(1-l)^2[R-(\overline{\theta}_1-\overline{e}_1)-\psi(\overline{e}_1)]\\
&+(1-p)^2(1-l)\,l[R-(\overline{\theta}_1-\overline{e}_1)-\psi(\overline{e}_1)]\\
&+(1-p)pl^2[R-(\overline{\theta}_1-e^*)-\psi(e^*)]\\
&+(1-p)pl(1-l)\,[R-(\overline{\theta}_1-e^*)-\psi(e^*)]\\
&+(1-p)^2l(1-l)\,[R-(\overline{\theta}_1-e^*)-\psi(e^*)]\\
&+(1-p)^2l^2[R-(\overline{\theta}_1-e^*)-\psi(e^*)]\,.
\end{aligned}\tag{A.7}$$

The first order conditions with respect to $\underline{e}_1$:

$$\begin{gathered}
=(1-l)p[1-\psi'(\underline{e}_1)]=0\\
\psi'(\underline{e}_1)=1.
\end{gathered}\tag{A.8}$$

For the efficient type we get the same solution as in the collusion-free case - no distortion on the top: $\underline{e}_1 = e^*=1$. The first order conditions with respect to $\overline{e}_1$:

$$\begin{aligned} &=(1-l)(1-p)[1-\psi'(\overline{e_1})]- \\ &[p^2l^2+(1-l)p]\Phi'(\overline{e_1})=0 \\ \psi'(\overline{e_1}) &=1-\frac{p}{(1-p)}[1+\frac{pl^2}{1-l}]\Phi'(\overline{e_1}) \end{aligned} \tag{A.9}$$

The inefficient type's effort $\bar{e}_1$ is further distorted due to the possibility of collusion, which costs the principal one more informational rent with probability p^2l^2. We label the further distorted effort level $\hat{e}_1 < \bar{e}_1$. The first-order conditions with respect to $\underline{e}_2$ and $\bar{e}_2$ follow in a similar fashion, and they are not stated here.

With the above first-order conditions we can solve the optimal contract. For the agent, the optimal collusion-proof contract offers $\psi'(e^*)= 1$, $w_1=\psi(e^{*})$, $U_1=0$ under perfect information, $\psi'(\underline{e}_1)= 1$, $\underline{w}_1=\psi(\underline{e}_1)+\Phi(\hat{e}_1)$,, $\underline{U}_1=\Phi(\overline{e}_1)$ in cases 5-8, and $\psi'(\hat{e}_1)=1-p/(1-p)$ $[1+pl^2/(1-l)]\ \Phi'(\hat{e}_1)$, $\bar{w}_1=\psi(\hat{e}_1)$, $\bar{U}_1=0$ in cases 9-12. The principal has to provide informational rent the agent in case 1: $w_1 =\Phi(\hat{e}1)$ and $U_1 = \Phi(\hat{e}_1)$, and this happens with probability p^2l^2.

For the supervisor, the solution is similar: $\psi'(e^*)= 1$, $w_2=\psi(e^*)$, and $U_2=0$ in the perfect information cases, $\psi'(\underline{e}_2)= 1$, $\underline{w}_2=\psi(\underline{e}_2)+\Phi(\hat{e}_2)$, $\underline{U}_2=\Phi(\hat{e}_2)$ when S is efficient, and $\psi'(\hat{e}_2)=1-p/(1-p)\ [1+pl^2/(1-l)]\Phi'(\hat{e}_2)$, $\bar{w}_2=\psi(\hat{e}_2)$, $\bar{U}_2=0$ when S is inefficient. The principal has to provide also the supervisor informational rent in case 1: $w_2=\Phi(\hat{e}_2)$, $U_2=\Phi(\hat{e}_2)$.

The solution above involves further distortion in the inefficient type's effort compared to the collusion-free case. The agent's and the supervisor's effort levels are further distorted to reduce the gains from collusion. From the strict convexity of disutility function, it follows immediately that $\hat{e}_1 < \bar{e}_1 < \underline{e}_1$. Similarly for the supervisor. These results give rise to Lemma 2.4. The ignored constraints are automatically satisfied with the same arguments as earlier.

To show that this contract is collusion-proof is straightforward. Assume that the principal offers the contract described above. Can the agent and the supervisor do any better by colluding when the collusion-proof contract is offered? When the principal offers this contract, which satisfies the coalition incentive constraint, IR and IC constraints, the agent and the supervisor do not have any incentive to collude. In particular, consider case 1, which is the only possibility for profitable collusion. If the agent and the supervisor colluded, they would gain the same informational rents that they get in case 6 ($r_1=r_2=\emptyset$). The collusion-proof contract, however, offers them in case 1 as much when they truthfully report supervisory information ($r_1 = \underline{\theta}_1$, $r_2 = \underline{\theta}_2$). There are no gains to be achieved by colluding. The contract derived above is, indeed, a collusion-proof contract. ☐

Proof of Proposition 2.3

Since we have fixed the gross value of the project exogenously as R, all that matters is the expected total costs under both organisation modes. We divide the proof into two parts, and analyse first the expected costs related to the agent, and then we consider the expected costs related to the supervisor. Then it is shown that, indeed, Organization II will always dominate when the agent and the supervisor are symmetric.

Note first that Organisation I is a special case of Organisation II where A's supervisory information about S is not used. Since we consider here a symmetric relationship between S and A, we label $\Delta\theta_1 = \Delta\theta_2 = \Delta\theta$. Since we want to look at whether the value of supervisory information has an effect on which is the most profitable organisation, we fix $p=l=1/2$.

When the principal decides to use A's information, the possibility of collusion arises, and the principal has to satisfy the coalition incentive constraint. The principal is certainly worse off since he has to provide rents the agent to get him to report his supervisory information about the supervisor. There are no other changes when A's information is used. The cost increase is purely due to the informational rent $\Phi(\bar{e}_1)$ that has to be provided the agent with probability p^2l^2.

The difference of the expected profits between OII and OI boils down to the following: $\Delta\pi = - p^2l^2\, \Phi(\bar{e}_1)$. The total expected costs related to the agent are indeed always higher in Organisation II.

When the principal decides to use A's information, the expected total costs related to the supervisor will decrease. To see this, note that within Organisation II, with the help of the agent's supervisory information about the supervisor, the principal avoids paying the informational rents $\Phi(\bar{e}_2)$ in cases 5, 9, and 13. These savings accrue with probability $pl(1-pl)$. In addition, with the help of the new information, the principal avoids inefficiency in four cases (4, 8, 12, 16), and there the solution is at the first-best level. These savings occur with probability (1-p)l. However, the principal has to provide the supervisor the informational rent in case 1 to motivate the supervisor to report information about the agent. That is, in case 1 the principal breaks collusion with money.

The expected costs related to the supervisor will then definitely decrease, and the profits increase when OII is adopted:

$$\Delta\pi_s = pl(1\text{-}pl)\Phi(\bar{e}_2) + (1\text{-}p)l\ [\ (\bar{\theta}_2 - \bar{e}_2) + \psi(\bar{e}_2) - (\bar{\theta}_2\text{-}e^*) - \psi(e^*)\ ]$$

Given that we consider a symmetric case, and have fixed p=l=0.5: $\Phi(\bar{e}_1) = \Phi(\bar{e}_2) = \Phi(\bar{e})$. Then we can combine the losses and the gains, $\Delta\pi = \Delta\pi_A + \Delta\pi_s$. Now we see that Organisation II is dominant always under symmetry, since: $\Delta\pi = pl(1\text{-}2pl)\Phi(\bar{e}) + (1\text{-}p)l\ [\ (\bar{\theta}_2 - \bar{e}_2) + \psi(\bar{e}_2) - (\bar{\theta}_2\text{-}e^*) - \psi(e^*)\ ] = 1/8\ [\Phi(\bar{e})] + 1/4\ [\ (\bar{\theta}_2 - \bar{e}_2) + \psi(\bar{e}_2) - (\bar{\theta}_2\text{-}e^*) - \psi(e^*)\ ] > 0$ □

Proof of Proposition 2.4

Now we allow asymmetry $\Delta\theta_1 \geq \Delta\theta_2$ and in proving Proposition 2.4 we use the expressions derived in the earlier proof: $\Delta\pi = -1/16\ \Phi(\bar{e}_1)$, and $\Delta\pi_s = 3/16\ \Phi(\bar{e}_2) + 1/4[\ (\bar{\theta}_2 - \bar{e}_2) - \psi(\bar{e}_2) - (\bar{\theta}_2\text{-}e^*) - \psi(e^*)\ ]$. The increase in costs, $-\ 1/16\ \Phi(\bar{e}_1)$ can now be written as:
$-\ 1/16\ [\psi(\bar{e}_1) - \psi(\bar{e}_1 - \Delta\theta_1)] = -\ 1/16[\Delta\theta_1 - 3/2\ \Delta\theta_1^2]$. The decrease in costs, $3/16\ \Phi(\bar{e}_2) + 1/4\ [\ (\bar{\theta}_2 - \bar{e}_2) - \psi(\bar{e}_2) - (\bar{\theta}_2\text{-}e^*) - \psi(e^*)\ ]$ can be written as :
$3/16\ [\Delta\theta_2 - 3/2\ \Delta\theta_2^2] + 1/4[\Delta\theta_2^2/2]$.
Now the decision concerning the organisation mode will depend on the sign of expression
$-\ 1/16[\Delta\theta_1 - 3/2\ \Delta\theta_1^2] + 3/16\ [\Delta\theta_2 - 3/2\ \Delta\theta_2^2] + 1/4[\Delta\theta_2^2/2]$
$= -1/16[\Delta\theta_1 - 3/2\ \Delta\theta_1^2] + 3/16\ \Delta\theta_2 - 1/32\ \Delta\theta_2^2$.
When this is positive Organisation II will dominate, and when it is negative Organisation I will dominate. Somewhere in between, Organisations I and II will yield equal expected profits. In deriving the above expression, we used a disutility function which takes the form $\psi(e) = e^2/2$, and the second-best effort level for the inefficient type: $\bar{e} = 1 - \Delta\theta$.

Clearly the sign of $-2/32\ \Delta\theta_1 + 3/32\ \Delta\theta_1^2 + 6/32\ \Delta\theta_2 - 1/32\ \Delta\theta_2^2$ depends on $(\Delta\theta_1 - \Delta\theta_2)$. When deriving characterizations we must start with the case where OI and OII will yield equal expected profits. We multiply the above expression by 32 and write it as follows:

$-\Delta\theta_2^2 + 6\Delta\theta_2 + 3\Delta\theta_1^2 - 2\Delta\theta_1 = 0$. When the above holds, the expected profits are equal between Organization I and II. The above relation can be solved by analysing implicit functions. We have a relation which takes a form $F(\Delta\theta_2, \Delta\theta_1) = 0$, and we ask whether there is an implicit function $\Delta\theta_2 = g(\Delta\theta_1)$ such that it is a solution to the above relation. Using the implicit function theorem, the derivative of $(d\Delta\theta_2/d\Delta\theta_1) = -F_{\Delta\theta 1}/F_{\Delta\theta 2} = -(6\Delta\theta_1 - 2)/(-2\Delta\theta_2 + 6)$. The condition for the existence of a local solution is $F_{\Delta\theta 2} \neq 0$, thus $\Delta\theta_2 \neq 3$. This is satisfied due to our restriction of the size of supports: $\Delta\theta_2 \in (0,1/2)$. Now we can solve $\Delta\theta_2$ as a function of $\Delta\theta_1$, and by using the general quadratic equation's solution method, we get the following roots:

$$\Delta\theta_2 = \frac{-6 \pm \sqrt{36-4*(-1)*(3\Delta\theta_1^2-2\Delta\theta_1)}}{-2}$$

Only one root, namely

$$\Delta\theta_2 = \frac{-6+\sqrt{36+12\Delta\theta_1^2-8\Delta\theta_1}}{-2}$$

satisfies the restriction we imposed on the support.

Now for all $\Delta\theta_1 \in (0,1/2)$ the above root will give $\Delta\theta_2^*$ such that $-\Delta\theta_2^{2+}6\Delta\theta_2 + 3\Delta\theta_1^2 - 2\Delta\theta_1 = 0$. For any $\Delta\theta_1$ define $\Delta\theta_1 - \Delta\theta_2^* = k^*$, which gives us the measure of asymmetry under which Organisation I and II are equally profitable. From now on we consider actual support $\Delta\theta_2$. Now it is evident that when $(\Delta\theta_1 - \Delta\theta_2) = k^*$, then the principal is indifferent between OI and OII. This concludes part (ii) of Proposition 2.4

To prove part (i) of Proposition 2.4 - i.e., that OII will dominate - we must prove that $-\Delta\theta_2^{2+}6\Delta\theta_2 + 3\Delta\theta_1^2 - 2\Delta\theta_1 > 0$. This is indeed the case if $(\Delta\theta_1 - \Delta\theta_2) < k^*$. Now the relationship between the agent and the supervisor is not too asymmetric, and Organisation II will remain the most profitable one. This concludes the first part of Proposition 2.4.

To prove that OI will dominate we must show that $\Delta\theta_2^{2+}6\Delta\theta_2 + 3\Delta\theta_1^2 - 2\Delta\theta_1 < 0$. The above is true if $(\Delta\theta_1 - \Delta\theta_2) > k^*$. Then the relationship between the agent and the supervisor is asymmetric enough, and the costs of collusion are higher then the benefits of new information. Thus, Organisation I will dominate. This concludes the proof. ☐

Chapter 3

Breaking Collusion by Task Assignment and Whistle-Blowing

3.1 Introduction

One of the most important and difficult decisions an organisation designer has to deal with is how to assign various tasks between workers. In this chapter our particular interest lies in the analysis of interaction between task assignment and collusion. We consider an organisation of two agents and a principal where collusion arises due to reciprocal monitoring. More precisely, the agents collude by exchanging favours by jointly concealing information from the principal. We do not consider monetary bribes; we simply assume that the transaction costs related to monetary bribes are infinite, and thus the agents may collude only by exchanging favours.

The chapter analyses how the principal is able to break collusion either by task assignment or whistle-blowing. With both of these instruments the principal is able to manipulate side-contracting opportunities. Task assignment and whistle-blowing are two different ways to break collusion. Through task assignment the principal creates a situation where the agents cannot exchange any favours. The idea behind using whistle-blowing is to create a situation where, if collusion has taken place, the benefits of it will be "taxed" away with some probability by imposing penalties for wrongdoers. When the penalties are high enough, the agents do not collude. In this chapter we are interested in whistle-blowing as an incentive device that blocks wrongdoing before it has even started. It is important to realize that here whistle-blowing is not part of the whistle-blowers' everyday tasks in their job. It simply ensures that those tasks are well executed.

In general, whistle-blowing means the disclosure by organisation members (former or current) of illegal, immoral or illegitimate practices to persons or organisations that may be able to affect the action.[1] Most of the real life whistle-blowing examples are found in situations where an organisation member exposes that the particular organisation is doing something illegal: e.g., a worker discloses to a newspaper that cars produced in the factory in which he is working are unsafe, a policeman discloses that his boss is taking drug money from criminals, and so on.

It is interesting to note that whistle-blowing arrangements are also adopted in some

---

[1] See more about whistle-blowing in Near and Miceli (1995).

big firms. For example, a telephone "hot line" has been opened to encourage workers to expose wrongdoing.[2] In the United Kingdom the government have even opened a telephone hot line where people are able to call anonymously and expose citizens who are claiming welfare benefits on wrong grounds.

At this point, the reader may wonder about the relevance and effectiveness of whistle-blowing where the agents who are colluding (the wrongdoers) are expected to blow the whistle themselves. In other words, they are practically turning in themselves when blowing a whistle. In that respect, it is interesting to note that the type of mechanism developed here has been proposed in a fight against firms' cartelization within the EU. In fact, the main idea there seems to be almost exactly the same as in the present chapter. "Under the proposal, firms that blow a whistle on a cartel before the commission gets wind of it either would not be fined or would see their fine greatly reduced "[3]

The first part of this chapter considers task assignment and time allocation in general as the measures the principal may use in preventing collusion. We consider a firm with two agents and a principal where collusion arises due to reciprocal monitoring. The principal allocates fixed working hours between monitoring and production tasks, and working time is by assumption contractible, but there are fixed adverse selection parameters for each agent. Since the agents learn each others' types, they are able to collude by concealing that information from the principal. Due to collusion, monitoring becomes less effective, and therefore less time will be allocated to monitoring activities. Moreover, time allocation now becomes interdependent, and the principal has to solve both agents' time allocation problems simultaneously.

In the collusion-proof equilibrium, the principal breaks collusion with money. We show that there is a simple way the principal can break collusion, namely by hiring a third worker and assigning one of the monitoring tasks to him. Under that task assignment structure no favours can be exchanged. Then with the fixed cost of the third worker the principal implements the collusion-free outcome. However, both ways of breaking collusion

---

[2] In the United Kingdom, for example, National Westminster Bank, Esso, and Lucas have installed telephone hotlines for whistle-blowers. Financial Times, 29 June 1995.

[3] The Wall Street Journal Dec. 7, 1995.

are costly.

The second part of the present chapter proposes a whistle-blowing mechanism as a collusion-breaking device. There the agents blow the whistle after collusion with positive probability, and, depending on the rewards and penalties, we show that the principal is strictly better off than by breaking the collusion with money (the collusion-proof solution.) We also show that if tougher penalties are allowed, the principal breaks collusion altogether. The agents do not collude in the reporting stage, out of fear of being exposed by their fellow agent. Interestingly, the principal may also break collusion by adopting non-monetary penalties. For example, the principal may fire a worker who has participated in corruptive activities, and if that is costly for the worker - which is quite plausible - that may prevent collusion.

This chapter is related to the literature of collusion in hierarchies. The current theoretical literature of collusion has assumed it as given so far that production and monitoring tasks are separated and exogenously assigned to workers.[4] This chapter builds on Chapter 2, where the idea that often many agents hold valuable information and that collusion appears more often as a form of non-monetary side transfers was introduced. That chapter showed that optimal use of information is endogenously determined, and some organisational forms are trivially collusion-proof. Sometimes the best response is to ignore one source of information, since doing so blocks collusion. This chapter in turn examines whether the principal can do any better than ignore information.

The papers closest to the present one are Felli (1996), Kofman and Lawarrée (1996) and Acemoglu (1995).[5] Acemoglu analyses how implicit collusion may arise in agency relations between a manager and an auditor. The manager has the authority to hire the auditor, and because of this the auditor may withhold information from the shareholders not

---

[4] See Tirole (1992) for a survey of this literature. The only decision concerning task allocation which is somehow endogenous is a decision to hire a supervisor in the first place. However, if the supervisor were not hired, one would be back in a normal principal-agent model, and collusion would not be a problem. Therefore, to make a hierarchy model interesting and collusion, possible the supervisor is typically always hired, and his services are used.

[5] See also Holmström and Milgrom (1990), who analyse side-trading among agents in their multi-task model.

in order to receive monetary bribes, but rather due to his own career concerns. In Acemoglu's study the task of the auditor is to find out the type of a project which an empire-builder manager is running, and to report what he finds to the shareholders. The type of the project is not observed by the shareholders, and thus there is asymmetric information between the shareholders and the manager. When the auditor announces that the project is a bad one, Acemoglu defines that as whistle-blowing, and discusses how an auditor should be rewarded for his task. However, there the auditor is simply doing what he is supposed to do. Whether we should call this whistle-blowing or not is not clear. The type of whistle-blowing Acemoglu considers is known in sociological and management literature as role-prescribed whistle-blowing. Whistle-blowing is just part of the auditor's job. He is expected to report any wrongdoing.

Felli (1996) analyses monetary collusion problems in a three-tier hierarchy model where an agent who is the only productive unit in an organisation may collude with a supervisor. He shows that by using an augmented revelation mechanism, a principal is able to achieve a collusion-free solution. In particular, the principal is able to achieve a collusion-free equilibrium because the supervisor has the discretion to exploit the agent's private information, which he learns during collusion negotiations; therefore an agent never accepts any collusion offers from the supervisor. Felli (1996) shows that this corresponds to the situation where the supervisor has the authority to choose contracts for the agent.

Kofman and Lawarrée (1996) analyse how a principal deters monetary collusion between an agent and an auditor by hiring a separate auditor who may also collude with the agent. They show that the principal is able to achieve a collusion-free solution by doubling the number of auditors. In particular, independent auditors play a prisoner's dilemma type of game. They also show that by sending auditors with probability less than one and sometimes informing the second auditor of his position, the principal prevents collusion between the agent and the auditors. In their model the first and second auditor are not able to collude between themselves. In contrast to Kofman and Lawarrée, in our case the agents who are involved in production and monitoring (working thus as an "auditor") themselves are able to collude.

The present chapter shares some lines of thought with all the above papers. However, this work is, to the best of my knowledge, the first one to analyse whistle-blowing as a collusion-breaking device. The emphasis here is to consider non-role-prescribed whistle-blowing as an incentive device.[6] That means that whistle-blowing is not the primary purpose for which the agents are hired. Rather, they are hired to perform production and monitoring tasks. In addition to those activities the agents can, if they observe some wrongdoing, blow the whistle and collect a reward. Here effective whistle-blowing ensures that the agents do not collude at the reporting stage.

The rest of the chapter is organized as follows. In section 3.2 the model is presented, and the main analysis is carried out in section 3.3. Section 3.4 discusses some interpretations of the present model. Section 3.5 concludes.

3.2 The Model

The Parties

Consider a principal (P) who is the risk-neutral owner of a firm which comprises two production processes ("machines 1 and 2"). The principal hires two risk-neutral agents to work with the machines. The production processes can be characterized simply by production functions, $f_i(t)$, i=1,2 and where t=time. Thus, output depends only on the hours that the agents work. The two production functions, $f_i(t)$, i=1,2 are identical, and have the following properties: $f'() > 0$, $f''() < 0$, $f(0) = 0$, $f'() = \infty$, when t=0, and $f'() = 0$, when $t \rightarrow 1$.

Assumption 3.1 Time, t is observable, verifiable and thus contractible.

Both agents have some private information concerning their types that is not observed by the principal. There is a fixed adverse selection parameter, θ_i associated with agent i. θ_i is a worker-specific fixed cost of working. We assume that θ_i belongs to the binary support $\{\underline{\theta}_i < \bar{\theta}\}$, $\underline{\theta}_i < \bar{\theta}_i$, and $\Delta\theta_i = \bar{\theta}_i - \underline{\theta}_i$, i=A,B. It is also assumed that p=probability$\{\theta_i = \underline{\theta}_i\}$ and (1-p)=probability$\{\theta_i = \bar{\theta}_i\}$, and that the types are identically and independently

---

[6] See Near and Miceli (1995), who discuss the difference between role-prescribed and non-role-prescribed whistle-blowing.

distributed. Later, we refer to $\overline{\theta}_i$ as a bad (inefficient, high-cost) type and $\underline{\theta}_i$ as a good (efficient, low-cost) type.

When the agent diverts some time away from production to monitoring, he observes with probability P(t) the fellow worker's type. The agent can then report it to the principal. Hence, the informational rents the principal has to pay are reduced. The monitoring task involves a monitoring function, P(t), that also has "normal properties": $P'() > 0$, $P''() < 0$, $P(0) = 0$, $P(1)=1$, and $P_i'() = \infty$, when $t=0$, $i= 1,2$.

Assumption 3.2 The total working time available for production and monitoring activities is two units (t=2).

Each worker's working hours are given, and the hours are normalized to 1; the working hours can be allocated to production and monitoring tasks. Agent i's time constraint is then simply $1 = t_i^m + (1-t_i^m)$, where t_i^m stands for time allocated to monitoring activities. The idea behind this assumption is to constrain our analysis here to the principal's problem of how to divide fixed working hours between different activities. Many modern organisations share this feature. Workers are expected to work, say, 8 hours per day, and these 8 hours are divided between different tasks.

The principal controls the agents simply by allocating the working hours between different activities. In effect, the principal chooses output levels and the expected cost savings by time-allocation decisions. In this model the wages are purely determined by monitoring reports. Within the model there are no moral hazard problems, and the only problem is related to the fixed adverse selection parameters that allow the efficient agents to earn the informational rent. Note that in the present model production levels are not contingent on the agents' types as in a normal adverse selection model. Here the agents do not report their own types to the principal. We come back to this later in this section.

The contract the principal offers A and B has to satisfy individual rationality (IR) constraints ex-post, say, due to limited liability:

$$U_i(\theta_i)=w_i(\theta_i)-\theta_i\geq 0 \quad \forall\ \theta_i\in(\underline{\theta}_i,\overline{\theta}_i),\ i=A,B. \tag{3.1}$$

Whatever the agent's type is, ex-post utility can not be negative. To make our analysis more tractable, the following assumption is made:

Assumption 3.3 The principal wants to produce even with the bad type, and does not want to separate the agents' types without further information. The principal offers only one wage $w_i(\theta_i)= \bar{\theta}_i$.

When offering $w_i(\theta_i)= \bar{\theta}_i$, the principal wants to produce even with the high-cost type and does not want to run the risk that the high-cost type will decline to work. This is precisely captured by a condition: $f_i(t) - \bar{\theta} > p\ (f_i(t) - \underline{\theta})$, which we assume always holds. It is therefore optimal to offer $w_i = \bar{\theta}_i$, i= A,B.[7] Consequently, ex post $\bar{U}_i = 0$ and $\underline{U}_i = \Delta\theta_i$. Assumption 3.3 is made without loss of generality to make the analysis of time allocation more tractable. Thanks to it we can concentrate purely on the principal's time allocation under potential collusion.

Information Structure

If an agent's working hours are allocated to monitoring activities, he learns his fellow agent's type with positive probability. If, for example, A spends t_A^m hours monitoring, he learns B's type with probability $P(t_A^m)$. And similarly for B. When the agent has learned the other agent's type, he can report it to the principal. That report is hard information, meaning that he can report it to the principal in a verifiable matter. It is further assumed that monitoring information is hard in the sense that the agent can not modify the signal he has learned, but can only conceal it.

To reduce the cases we have to consider, we assume that agent i may learn new information about agent j only if agent j is of an efficient (low-cost) type. Then, given that he has learned a signal $\sigma=\{\underline{\theta}\}$, he can report, $r=\{\underline{\theta},\varnothing.\}$; and if he has not learned a signal, $\sigma=\varnothing$, then he can report only $r=\varnothing$. If agent j is of an inefficient (high-cost) type, agent i learns nothing. In other words, learning that agent j is inefficient does not have any value, and in

---

[7] The simplifying of Assumption 3.3 and the way information structure is modelled are both borrowed from Tirole's (1992) "bare bones" model.

particular the principal does not learn anything new.

Since the principal is unable to observe types and signals, it is clear that with the aid of new additional information provided by the agents, the principal's welfare increases. The principal cannot distinguish whether an agent has observed a signal or not when he receives a report $r=\varnothing$. This, of course, facilitates information manipulation within a model.

For later purposes it is important to notice that an agent learns whether another agent has learned a signal. This, in turn, implies that collusion occurs under symmetric information, and then the agents observe the signals, which greatly simplifies the analysis of side-contracting. Even though the agents collude under symmetric information, we make the following assumption:

Assumption 3.4 The agent is not able to prove to the principal in a verifiable manner whether the other agent was informed or not.

This means that when A has learned a signal $\sigma_A=\{\underline{\theta}_B\}$, he can prove it to the principal by a verifiable report $_{\text{Ra..}}=\underline{\theta}_{B.}$ When collusion takes place under conditions of symmetric information, A also learns that B has learned A's type. However, it is assumed that A is unable to prove to the principal whether B was informed or not without some further evidence. We come back to this in section 3.3.2.

The reason why monitoring is profitable is the following. Since the wages are purely determined by monitoring reports, the efficient (low-cost) type's informational rent can be reduced by monitoring. If there is no monitoring or if monitoring does not provide new information, the principal pays $w_i=\overline{\theta}_i$.[8] If new information is received, $w_i=\underline{\theta}_i$, i=A,B. Due to this wage contract, the agents' expected utility is:

$$\begin{aligned}&\underline{U}_i=[1-P(t_j^m)]\overline{w}_i+P(t_j^m)\underline{w}_i-\underline{\theta}_i=\Delta\theta_i-P(t_j^m)\Delta\theta_i,\\&\overline{U}_i=\overline{w}_i-\overline{\theta}_i=0,\quad i=A,B,\ i\neq j.\end{aligned}\tag{3.2}$$

---

[8] The reason why $w_i=\overline{\theta}_i$ is still optimal is that once the principal receives a report $r=\varnothing$ he updates the probability for the efficient type according to Bayes' rule: $\text{prob}\{\theta=\underline{\theta},\ r=\varnothing\}=$ p (1-P(t)) / (p (1-P(t)) + (1-p)), which is smaller than p. Now recall the condition: $f_i(t)-\theta > p\,(f_i(t)-\underline{\theta})$. Therefore, it still is optimal to offer $w_i=\theta_i$.

The principal benefits from monitoring due to the reduced informational rent the efficient type earns. Without monitoring, the principal pays an informational rent equal to $\Delta\theta_i$, and thanks to monitoring the principal pays : $\Delta\theta_i$ - $P(t_j^m)\Delta\theta_i$, i=A,B, and $i \neq j$. The inefficient type's (IR) constraint binds always, and he does not earn any informational rents.

It is important to realize that after an agent has been offered and has accepted a contract where wages are determined as in (3.2), agent i has no incentive to lie when reporting a signal he has learned about the fellow worker. In particular, since a signal is hard information, an agent can not increase his own welfare by individually concealing the signal.

In the present model the agents do not report their own types at all. Alternatively, we could have considered a different model, where agents report their types at stage 1, and where the principal's time allocation decision would then be based on these announcements. However, that model would, in effect, be almost identical to the present one, because there also the problem of collusion would arise. Moreover, the principal faces there a commitment problem, because monitoring is costly. By committing to monitor ex post after high cost type has been announced the principal can deter the efficient type from not claiming to be the inefficient one. But then monitoring ex post is costly, because it takes working hours away from production. Moreover, the principal knows that he will not learn anything new, since only the inefficient types announce high cost. The principal would be better off by not allocating time to monitoring after high-cost announcements. After reasoning this, the agents would therefore always prefer to claim to be inefficient types. The best the principal can do is to monitor always (irrespective of what has been announced) and pay $w_i=\theta_i$ if he does not get any new information.

The principal is the risk-neutral owner of an organisation and maximizes expected profits:

$$E\Pi=E[f_1(t)+f_2(t)-w_A-w_B]=f_1(t)+f_2(t)-E[w_A+w_B]. \qquad (3.3)$$

The principal's problem is to allocate fixed working hours between production and monitoring tasks in a profit-maximizing way.[9] When making this decision, he faces a simple,

---

[9] The price of output has been normalized to one.

yet fundamental trade-off: working hours allocated to monitoring are lost from production. The organisation comprises two production and two monitoring tasks and all of them are delegated to A and B.[10] For later purposes it is useful to recall that monitoring has a positive effect on the principal's welfare, but a negative effect on the workers' welfare. This latter is precisely the reason why collusion may arise when the agents cross-monitor one another.

Collusion

Before turning to the analysis, a few words about collusion are in order. The mode of collusion we consider here is based on the simultaneous exchange of favours within a coalition as developed in Chapter 2. The agents collude to conceal private information they have learned about each other from the principal. By jointly concealing information from the principal, they are able to earn higher wages due to the informational rents.

Since we consider non-monetary side-transfers, collusion can take place only when both A and B are efficient and have learned one another's types. Only in that case can favours be exchanged simultaneously. In contrast to monetary collusion, this type of collusion does appear on the factory floor. For example, workers may cover each other's mistakes or jointly conceal some unfavourable information from their immediate superiors, and so on.

The crucial question concerning side contracts is their enforceability. Of course, due to the informal nature and illegality of side contracts, they cannot be enforced in a court. Therefore, in the literature, enforceability is taken to be the case by assumption, and this is also the starting point here. Note, however, that with the type of collusion we consider enforceability is a less serious problem, since the collusion takes place in the reporting stage, and no further actions are needed. Moreover, there is no need for ex-post monetary transfers from one agent to the other.

In the present chapter our interest lies in the question why, apart from their being illegal and thus unenforceable, agents may be unable to strike profitable side contracts. We show that there are some fairly simple ways the principal can affect side-contracting opportunities. So as a first step we take it for granted that if agents agree on a side contract,

---

[10] We are implicitly assuming here that hiring more agents is too expensive compared to the expected benefits the principal would gain. In the next section we show that the principal is able to break collusion easily by hiring a third worker.

they will follow it (ex-ante information manipulation). As a second step we analyse how the principal is able to reduce profitable side-contracting opportunities by task assignment and by introducing whistle-blowing as a collusion-breaking device.

Timing

The timing of the model is as follows:

The Main Contract   Production Monitoring   A reports r_A B reports r_B Wage Payments

0-------------------1-------------------2-----------------3-------------------4-------------------5----------> t

A learns θ_A
B learns θ_B
Side Deals

At stage 0 the agents learn their own types. At stage 1 the principal offers the main contract. At stage 2 both agents accomplish production and monitoring tasks as specified in the main contract. At stage 3 agents are able to agree on side deals, which are based on information they learned at stage 2. At stage 4 the agents report. Finally, at stage 5 the principal rewards A and B according to the main contract.

3.3 The Analysis

3.3.1 Breaking Collusion by Task Assignment

This section considers the principal's time allocation problem when agents are able to collude, and possible ways the principal can break collusion. We start by stating the first-best result under perfect information. After that we consider time allocation under a collusion-free regime, which will be our reference point for the later analysis. Next we consider time allocation under potential collusion and derive the collusion-proof equilibrium, where the principal breaks collusion with money. At the end of this section we consider how the principal can break collusion by task assignment. It is shown that collusion can be broken by assigning tasks in such a way that no favours can be exchanged. In effect, this means that at least one agent has to be unmonitored.

Under conditions of perfect information there is no need for monitoring, and all available time will be allocated to production. There is nothing to optimise: all available working hours are allocated to production, and both agents work full-time with the "machines". It is not optimal to leave any time unused. Thus, we have:

Lemma 3.1. Under perfect information all available time is allocated for production:$t_i^{m*}=0$*.* *The wage* $w_i = \theta_i$, $\theta_i = \{\underline{\theta}_i, \bar{\theta}_i\}$, *and* $U_i = w_i - \theta_i = 0$, $\forall\, \theta_i \in \{\underline{\theta}_i, \bar{\theta}_i\}$.

Next we turn to the case of asymmetric information. The principal's problem is to minimize the informational rents the agents are able to earn due to their private information. To reduce informational rents (wages), the principal allocates working hours also to monitoring. Note, however, that the opportunity cost of monitoring is reduced output due to the fact that now less time is available for production. However, the benefits of additional information more than outweigh the reduced output. When all available working hours are allocated to production, the last unit of time in the production task is very inefficient, $f'(1) = 0$. Now those units can be allocated to very effective monitoring activity, since $P'(0)=\infty$.

The principal chooses the working hours the agent spends in each activity in such a way that at the optimum the marginal productivity of each task is the same. If this were not the case, the principal could increase his expected profits by transferring time from the less efficient activity to the more efficient one.

Under a collusion-free regime, the agents are rewarded according to the main contract, which gives them: $w_i=\bar{\theta}_i$, if $r_j = \varnothing$, and $w_i=\underline{\theta}_i$, if $r_j = \underline{\theta}_i$, i=A,B, j = A,B and $i \neq j$. Once the agent has been offered such a contract, he has no incentive to lie when reporting a signal he has learned about the fellow worker's type, because he can not increase his own welfare by individually concealing the signal. In the collusion-free case, the principal gets additional information about the agents at no further cost. The time allocation problem is:

$$Max_{t_A^m, t_B^m}[f_1(1-t_A^m)+f_2(1-t_B^m)-E(w_A+w_B)] \qquad (3.4)$$

$$s.t. \quad (3.2).$$

The full solution is derived in Appendix B. The optimal contract is a combination of the first- and second-best solutions. The time allocation between production and monitoring tasks is determined by the first-order conditions, which boil down to the two following condition:

$$\begin{aligned} A&: \quad P'(t_A^m)p\Delta\theta_B = f_1'(1-t_A^m), \\ B&: \quad P'(t_B^m)p\Delta\theta_A = f_2'(1-t_B^m). \end{aligned} \tag{3.5}$$

The condition (3.5) has a very intuitive interpretation. The left-hand side stands for the expected savings in informational rents. The right-hand side is the marginal value of production. Due to the concavity of $P(t_i^m)$ and $f(1-t_i^m)$, we see that the time allocated to monitoring increases when the probability of a good state, p, and the rent, $\Delta\theta_i$ increase, since then time spent in the monitoring task becomes more valuable. The results of the collusion-free case can be stated as follows:

Proposition 3.1 Under a collusion-free regime, the optimal unique allocation of working hours between production and monitoring tasks is symmetric: $t_A^{m*} = t_B^{m*}$*, and most of the working hours are allocated to production,* $1-t_i^{m*} > t_i^{m*}$*,* $i=A,B$*. The optimal contract is a combination of the first- and second-best solutions.*

Proof: See Appendix B.

Under the collusion-free case, the agents' time allocation between the monitoring and the production tasks is independently determined. Thus we have:

Corollary 3.1 When collusion is not a problem, the workers' time allocation problems can be solved independently.

Proof: To prove Corollary 3.1 it is enough to consider the first-order conditions. As is evident from (3.5), optimal t_i^m does not depend on t_j^m, $i=j=A,B$, and $i\neq j$. □

In the collusion-free regime the principal can concentrate purely on time allocation, and each agent's working hours are allocated to production and monitoring activities independently. This is due to a simple economic reason. Here an agent's utility is independent of the other agent's type. The principal is, of course, better off due to the lower wages he has to pay.

Now we are in a position to analyse the effects of potential collusion on the optimal time allocation. When the principal has offered the collusion-free contract, it is obvious that the agents can do better by jointly concealing information they have learned from the principal. However, collusion can take place only when both A and B are efficient and have learned one another's types, since only then can they exchange favours simultaneously by jointly concealing the signals they have learned from the principal.

The principal can break collusion only by making the agents as well off as they would be when colluding. This means that the principal has to provide the agents the informational rents they would earn by colluding. When the principal satisfies a so-called coalition incentive constraint, the agents have no reason to collude. In other words, the principal breaks collusion with money. In the collusion-proof equilibrium, the coalition incentive constraint (CIC) has to be satisfied:

$$(w_A-\underline{\theta}_A)+(w_B-\underline{\theta}_B)\geq(\overline{\theta}_A-\underline{\theta}_A)+(\overline{\theta}_B-\underline{\theta}_B). \tag{3.6}$$

In the above, the left-hand side represents the agents' joint utility when they report truthfully the information they have learned ($r_A=\underline{\theta}_{B,}$ $r_B=\underline{\theta}_A$); it has to equal the right-hand side, which is what the agents could get by concealing information ($r_A=\varnothing_,$ $r_B=\varnothing$). The time allocation problem under potential collusion is:

$$\begin{aligned}&Max_{t_A^m,t_B^m}[f_1(1-t_A^m)+f_2(1-t_B^m)-E(w_A+w_B)]\\&\qquad s.t.\quad(3.2),(3.6).\end{aligned} \tag{3.7}$$

The full solution can be found in Appendix B. It is a combination of the first- and the second-best solutions as in the collusion-free case. The principal's welfare is, however, reduced due to the binding CIC constraint, and thus also the optimal time allocation differs. At the optimum, the following time allocation conditions have to be satisfied simultaneously:

$$\begin{aligned} A: &\quad P'(t_A^m)p\Delta\theta_B(1-pP(t_B^m))=f_1'(1-t_A^m) \\ B: &\quad P'(t_B^m)p\Delta\theta_A(1-pP(t_A^m))=f_2'(1-t_B^m). \end{aligned} \tag{3.8}$$

The left-hand side of (3.8) implies that the expected savings are now lower than in the collusion-free case (recall (3.6)). Therefore, at the optimum, less time is allocated to monitoring than in the collusion-free case: $t_i^{m**} < t_i^{m*}, i = A, B$. We can summarise the results of the collusion-proof case as follows:

Proposition 3.2 In the collusion-proof equilibrium, the optimal unique time allocation between the agents is also a symmetric one, but less time is allocated to monitoring and more to production than in the collusion-free solution: $t_i^{m**} < t_i^{m*}, i = A, B$. *The optimal contract provides for both A and B a combination of the first and second best solutions.*

Proof: See Appendix B.

Due to the possibility of collusion, the principal has to provide the informational rents the agents as a "monitoring wage". Hence, monitoring becomes less effective, and therefore less time is allocated to monitoring activities.The following important corollary follows directly from the first-order conditions:

Corollary 3.2 Under potential collusion, the optimal time allocation between the agents is interdependent.

Proof: To see this, consider, for example, A's time allocation condition in (3.8). The optimal time A spends in monitoring (t_A^{m**}) is negatively related to the time B spends in monitoring activities.Taking the cross derivative with respect to t_B^m A's time allocation condition becomes: $- P'(t_A^m)\ p\ \Delta\theta\ P'(t_B^m) < 0$ ☐

The interdependency under collusion is an important and very interesting result. Indeed, it shows that the possibility of collusion changes the optimal time allocation, and, moreover, it makes totally independent production processes and time allocation problems interdependent. The principal can no longer allocate the agents' working hours separately.

In the above we have seen that breaking collusion through time allocation and by satisfying the coalition incentive constraint is costly, since the principal breaks collusion with money. It is evident that collusion is an issue as long as reciprocal monitoring exists. But then a simple question arises: would it be optimal to break collusion by assigning tasks in such a way that the agents can not exchange any favours and the coalition incentive constraint does not have to be satisfied?

Clearly, the way to break collusion is to create asymmetric time allocation, which means that with two agents, at least one of them must be left unmonitored. But this is the solution we already know from Chapter 2. Rather than elaborating on that further, we want to show that the principal may also break collusion by assigning one of the monitoring tasks to an uninterested third party.

Recall the principal's time constraint of two units of time. So far we have explicitly ruled out the possibility of hiring more workers. Assume now, for a moment, that the principal is able to hire a third worker with a fixed cost F. Now it is easy to show that with the help of a third worker the principal prevents collusion altogether; moreover, this is done within the time constraint. Indeed, we can show that the principal implements a collusion-free outcome.

The idea is simple. The principal hires a third agent to do one of the monitoring tasks. Then the principal is able to assign tasks in such a way that no favours can be exchanged. This can be implemented in the following way. A produces for $(1- t_A^{m*})$ hours and monitors B for t_A^{m*} hours. B only produces for $(1- t_C^{m*})$ hours, and a third agent, C, monitors A for t_C^{m*} hours. Under this task assignment structure, B can not do any favours for A, and A in turn can not do favours for C; therefore, there is no room for collusion, and the principal breaks collusion by hiring an extra worker and reallocating the working hours. Note that the principal's time allocation under the above task assignments coincides with the time allocation of the collusion-free solution. Since the extra worker is costly, the decision how to break collusion becomes endogenous:

Proposition 3.3 If the fixed cost of hiring a third agent is not too high, it is better for the principal to break collusion by hiring an external monitor. If the fixed cost are high, it is better to break collusion with money and adjust the time allocation accordingly.

Proof: See Appendix B.

This is a very interesting and important result, because it implies that one should be able to build a simple theory which tells us when a new worker should be hired. Here that decision is driven purely by the principal's concern for potential collusion, rather than for any technological reasons.

The more general implication is that, given that the agents collude by exchanging favours due to reciprocal monitoring, it is always possible to break collusion by hiring an extra worker whose only task is to monitor and who himself is not monitored. Then all remaining monitoring tasks can be assigned in such a way that no reciprocal monitoring exists, and thus collusion is not a problem.

Since, however, our main interest in the present chapter is to consider collusion-breaking in the case where there are only two workers, we introduce in the next section a whistle-blowing mechanism as a collusion-breaking device. Indeed, we are able to show that under some conditions, the principal is able to break collusion with lower costs than by simply offering the collusion-proof contract with adjusted time allocation.

3.3.2 Whistle-Blowing as a Collusion-Breaking Device

This section considers a simple whistle-blowing mechanism as a collusion-breaking device. In particular, we consider whether the principal is able to achieve the collusion-free outcome of Proposition 3.1 by adopting an organisation policy which promotes whistle-blowing. It is shown that with appropriate rewards and penalties, the whistle-blowing mechanism guarantees strictly higher profits than does the collusion-proof solution. And if more severe penalties are allowed, the principal achieves the performance of the collusion-free solution at no further cost.

Before going any further, it would be useful to state precisely what whistle-blowing means in the present context. In our model whistle-blowing refers to a situation where, after collusion has taken place, agent i reports to the principal that agent j lied when he reported that he did not learn anything. However, whistle-blowing is possible only when agent i receives hard information (evidence) which facilitates his task of proving that agent j was actually informed about the signal even though he reported otherwise. In the following we assume:

Assumption 3.5 Given that collusion has taken place, then with probability γ_A (γ_B) A (B) receives hard information with which he is able to prove whether B (A) was informed about the signal $\sigma_B=\{\underline{\theta}_A\}$ ($\sigma_A=\{\underline{\theta}_B\}$) or not.

It is precisely this new additional hard information, which we assume arrives exogenously, that facilitates whistle-blowing. Agent i may now turn in agent j to the principal: "Agent j was actually informed about the signal even though he reported otherwise, and here is the evidence". An alternative and maybe more intuitive way to express the idea of new additional hard information is to assume that there is a third agent, and the hard information about collusion (wrongdoing) reaches him. This allows him to blow the whistle and inform the principal that collusion has taken place.[11]

As pointed out in the introduction, the type of whistle-blowing we examine is a so-called non-prescribed whistle-blowing. The agents are hired to perform production and monitoring tasks, and whistle-blowing is not their priority task. It works as an additional incentive device to ensure that the primary tasks are well executed. Here, that means the agents do not collude at the reporting stage.

Note that within the time line there is now a new stage for a whistle-blower, who can act after the official reporting stage, given that he has learned some additional information.

---

[11] Whistle-blowing should be encouraged only when hard information about wrongdoing is available. If no evidence is required from the whistle-blower, it leads to situations where innocent persons will be punished. This is especially the case where an agent can not prove that he was uninformed when it has been claimed otherwise. See Okuno-Fujiwara, Postlewaite and Suzumura (1990) for a similar situation where an agent can not prove his own ignorance.

We assume the absence of any coordination at the whistle-blowing stage due to the fact that there is no official point in the time line when whistle-blowing should take place. There is only an interval between the reports and the payments, and the agents are unable to control one another.

The above is, of course, an assumption, but we believe it captures well the whistle-blowing arrangements adapted in practice. Consider, for example, a case where a firm has installed a telephone hot-line that workers can use to announce corruptive activities in an organisation. In that case coordination among the whistle-blowers is very difficult if not impossible to achieve or to sustain.

We are now ready to describe the whistle-blowing mechanism. Given that collusion has taken place at the reporting stage, then with probabilities γ_A and γ_B the agents learn information which facilitates whistle-blowing. The whistle-blowing device is as follows. The agent who blows the whistle first will be rewarded, and the second agent will be penalized. By manipulating rewards and penalties, the principal creates a situation where whistle-blowing becomes a dominant strategy when an agent has learned additional information which facilitates whistle-blowing. More generally, the introduction of a whistle-blowing mechanism makes an agent unable to commit not to blow the whistle ex post, since that is a profit-maximizing strategy. This, in turn, destroys the possibility for collusion at the reporting stage, when the penalties are high enough. And if the penalties are very mild, the agents are unable to fully realize the benefits of collusion, since whistle-blowing makes collusion less profitable for the agents. The principal is better off even though he may be unable to prevent collusion with probability one.

We consider next different combinations of rewards and penalties the principal may use in preventing collusion. In the following we assume that the toughest penalty for being found participating in collusion is that the agent's information rent is "taxed away". That is, the agent is held at his reservation utility level. In the following we show that within the model of non-prescribed whistle-blowing, a reward for the exposer has to be related to the gains from collusion.

Suppose that the principal offers the collusion-free contract. As we have shown earlier, collusion may now arise. Assume further that the main contract includes the following clause: "If I, the principal, receive information about collusion, then all agents will be penalised by taxing their informational rents away."

Consider, for example, A, who is considering whether to collude or not. If there is no collusion, the agents report truthfully and both of them get their reservation utility: $\underline{U}_i = 0$, i=A, B. If, instead, A and B collude, then ex post after the reporting stage, an agent i, who has on his possession a piece of verifiable information showing that agent j was actually informed, has to decide whether to blow the whistle or not. Given that the principal does not provide any reward for a whistle-blower, but imposes penalties for both agents, we have:

Lemma 3.3 Without a reward, and given that the penalties are as defined above, no whistle-blowing is observed, and we have a collusive equilibrium with a probability $p^2P(t_A^{m})P(t_A^{m})$.

Proof: To prove this needs only two steps. Firstly, given that collusion has taken place and if A has learned additional information he must decide whether to blow the whistle or remain quiet, by blowing the whistle he gets:$(1-\gamma_B)(0) + \gamma_B(0) = 0$, and by remaining quiet he gets: $(1-\gamma_B)(\Delta\theta_i) + \gamma_B(0) = (1-\gamma_B)(\Delta\theta_i)$. Clearly, it is a dominant strategy to remain quiet, and similarly for B. Therefore, the expected profits from collusion are $(1-\gamma_A)(1-\gamma_B)(\Delta\theta_i)+(1-\gamma_A)(\gamma_B)(\Delta\theta_i)+\gamma_A(1-\gamma_B)(\Delta\theta_i)+\gamma_A\gamma_B(\Delta\theta_i) = \Delta\theta_i > 0$, i= A, B, because nobody blows the whistle. Thus, when possible, the agents always collude and nobody blows the whistle. ☐

Here we have collusion in an equilibrium with a positive probability, and, furthermore, nobody blows the whistle. The clear implication is that the principal has to provide the informational rent the agent who blows the whistle, because otherwise no agent would ever choose to do so. The reward for a whistle-blower has to be related to the gains from collusion.

Assume now that the principal rewards the agent who blows the whistle first by allowing him to keep his informational rent and giving him a small additional reward, ϵ. In this case the collusion-free contract is supplemented by the following clause: "If I, the principal, receive information about collusion, then the first agent who blows the whistle can keep his informational rent and he will get an additional reward. The other agent will be penalised by taxing his informational rent away."

For an agent who has a piece of verifiable information that facilitates whistle-blowing, it becomes a dominant strategy to blow the whistle. This is because only the agent who blows the whistle first collects a reward, and the other agent will be punished. We have:

Proposition 3.4. With the above rewards and penalties, we have a whistle-blowing equilibrium with positive probability where either A or B blows, or both blow, the whistle, and a collusive equilibrium with probability $p^2P(t_A^m)P(t_A^m)(1-\gamma_A)(1-\gamma_B)$.

Proof: See Appendix B

In contrast to the earlier result, here we have a case where in an equilibrium we have collusion and whistle-blowing with positive probability. The economics behind Proposition 3.3 is fairly straightforward. The main point is that the introduction of a whistle-blowing mechanism reduces the costs of potential collusion, and the principal is better off. Consequently, more time is allocated to monitoring, because whistle-blowing has made it more effective. Now, given that the probability of hard information about wrongdoing increases, more time can be allocated to monitoring activities, and the optimal time allocation approaches that of the collusion-free case. The principal's welfare increases and approaches the one of the collusion-free equilibrium. Indeed, under the whistle-blowing mechanism, the principal is better off than by satisfying the coalition incentive constraint by offering the collusion-proof contract with a modified time allocation. Note that this is true since the savings in informational rents ($\Delta\theta$) are higher than the small reward (ϵ) the principal has to pay the first whistle-blower, when γ_A and γ_B increase. Interestingly, under the whistle-blowing scheme it may also be the case that in an equilibrium we have collusion.

In the above we have restricted penalties in such a way that ex post an agent's IR constraint has to be satisfied ex-post even in that when he is caught for collusion. If this is relaxed, it is easy to show that collusion can be broken altogether. Interestingly, an organisation designer may not have to use tougher monetary penalties to achieve this. Consider the following penalty scheme, which may be more realistic. Let the rewards be the same as above, but let the penalty structure be the following. If both agents blow the whistle, then both agents' informational rent is taxed away, and they are allowed to stay in the organisation. But if only one agent blows the whistle, then he will be rewarded as above, and the other agent's informational rent will be taxed away; and in addition to this he will be fired. Assume also that an agent attaches a big negative utility, say - D, to being fired from an organisation. One may think of D as a loss of future benefits, pension, and so on. Then, the following corollary follows:

Corollary 3.4 If more severe penalties are allowed, which, of course, are not executed in an equilibrium, a symmetric collusion-free outcome can be supported with no additional cost.

Proof: The first step is to show that, given that if an agent - say A - has learned new hard information it is a dominant strategy to blow a whistle, by blowing the whistle he gets: $(1-\gamma_B)(\Delta\theta_i+\alpha) + \gamma_B(0) = (1-\gamma_B)(\Delta\theta_i+\alpha)$, and by remaining quiet: $(1-\gamma_B)(\Delta\theta_i) + \gamma_B(-D)$, where D is large. Clearly then it is a dominant strategy for A to blow the whistle, and that is true for B as well. This, in turn, implies that the expected utility from collusion:$(1-\gamma_A)(1-\gamma_B)(\Delta\theta_i)+(1-\gamma_A)(\gamma_B)(-D)+\gamma_A(1-\gamma_B)(\Delta\theta_i+\epsilon)+\gamma_A\gamma_B(0) < 0$, i= A, B. Therefore, nobody tries to collude, and thus we have no whistle-blowing either. □

The important point, thus, is that the more severe penalties do not necessarily have to be monetary penalties. For example, here the penalty is a combination of monetary and non-monetary punishments: rents are taxed away as above, and a worker is fired. In practice, this type of punishment may be very effective, and if a worker has been fired because he has participated in corruptive activities it may be a very efficient punishment. The worker may, for example, be excluded from any vacancies in civil service. Similarly, an employer may refuse to give references, and make it explicitly known why a worker was fired. These all may work as very efficient punishments, and more importantly the agents' wealth is not a constraint here in imposing punishments. No doubt, this type of punishment is a far more important control mechanism in civil service than the monetary penalties.

3.4 A Discussion: Some Interpretations

Separation of Monitoring and Production Tasks

In general, all organisations where favours can be exchanged between members of an organisation are potentially corruptible. Therefore, the monitoring and production tasks are often separated and carried out by different actors. One recent real-life example of this is related to the monitoring of MPs' outside interests in the United Kingdom.[12] The main point in this discussion has been that earlier internal control by fellow MPs (cross monitoring) has not worked well enough. Therefore, an independent monitoring committee has been established. This is a clear response from "the principal" to break potential collusion, mutually beneficial information manipulation by separating policy-making and monitoring tasks.[13]

Similar kinds of arrangements where an agent's or a unit's main task and the monitoring task have been separated from one another can be found in many real-life situations. For example, in big multinational firms, accounting departments of different subsidiaries do not monitor (audit) one another. Typically, auditing is performed by separate auditing units or independent auditing firms. That is, book-keeping and auditing tasks are separated and carried out by different units. Similarly, in politics, decision-making and monitoring tasks are separated from one another. In the UK, for example, Government offices do not monitor one another, but a particular office is monitored by a respective shadow minister.

In the main analysis we have, however, shown that it may actually be beneficial to have non-specialisation task structure; yet, as we have seen above, there are plenty of real-life examples where specialisation has been chosen. An interesting question arises: is there too much specialisation in practice?

---

[12] The Economist 1995, 25 Feb and The Financial Times 1 Aug 1995.

[13] Among the police forces similar accusations about internal investigation procedures have been made public.The problem indeed is that "a policeman is policing himself".

The Structure of Accounting and Documenting Systems

It is important to note that an organisation designer may actually increase the probability that hard information about wrongdoing in an organisation becomes available by certain organisational arrangements. For one thing, he may smooth the path for a whistle-blower by providing rewards and protection to persons who choose to blow the whistle. For example, in the US there are laws that provide protection for whistle-blowers.

Secondly, an organisation designer may reduce the number of corporate structures that facilitate cover-ups by increasing the number of people who are likely to know about possible wrongdoing and therefore capable of blowing the whistle. In practice, a measure taken by many organisations is to make the decision-making process more open, which creates information. Also, the ISO-9000 standard, as part of Total Quality Management, aims precisely at that by making all transactions in an organisation transparent to third parties. This type of measure improves the effectiveness of external auditing, and making it easier for an external auditor to discover wrongdoing, mistakes, etc.

In practice, accounting and documenting systems serve an important role which facilitates whistle-blowing. In many situations these documents reveal who in an organisation is informed about something, and in practice signatures in these documents work for this purpose. Imagine, for example, an auditor who has to sign documents which show what sort of material he has gone through in an auditing process. Later, these signatures may serve as evidence that the auditor and perhaps also a manager were informed about, say, the bad state of the company, etc. In a similar way, all signatures in contracts, documents, etc. are, in fact, evidence ex post that the persons who approved and signed those contracts were informed about them. The signature procedure in general can be interpreted in such a way that it ensures that different parties are informed enough to blow the whistle.

This type of situations arise also in politics. For example, when President Bush was running for the presidency, there was a heated debate whether he was informed about the Iran-Contra deal or not. His defence was precisely: "I was uninformed, I was out of the loop". Recently, a General Secretary of Nato, Willy Claes, had to resign, because he was involved in a bribery scandal. He also defended himself by claiming that he was uniformed about the bribes. However, some new evidence came up which proved that he was actually informed, and yet he had not blown the whistle.[14]

In the model presented here, an interpretation is that after collusion, an agent learns with some probability new hard information (e.g., documents) which facilitate whistle-blowing. Of course, within the abstract model, this looks artificial; but in practice whistle-blowing is generally facilitated by some verifiable information concerning wrongdoing.

3.5 Concluding Remarks

In this chapter we have analysed different measures the principal may adopt in breaking collusion. In the first part of the chapter, we considered task assignment and time allocation in general as measures the principal may use. It was shown that, due to collusion, less time is allocated to monitoring than in the collusion-free case. More importantly, the time allocation is then interdependent, and the principal has to solve both agents' time allocation problems simultaneously. Under the collusion-proof solution, the principal breaks collusion with money. It was also shown that the principal may implement the collusion-free situation with three agents under the time constraint by delegating one of the monitoring tasks to the hired third agent. On the whole, breaking collusion through task assignment is costly in one way or another.

In the second part we introduced a whistle-blowing mechanism that aims at discouraging the agents from colluding at the reporting stage. Depending on the penalties and rewards, we were able to show that the principal does strictly better than in the collusion-proof equilibrium. Moreover, if penalties for participating in collusion are high enough, the principal is able to prevent collusion with no further cost and reach a symmetric collusion-free solution. Interestingly, the principal is also able to break collusion by adopting a non-monetary penalties.

In future, it might be interesting to look at how effective the measures adopted here are in breaking monetary collusion. Most obviously, the whistle-blowing mechanism will

---

[14] See more in Elliot and Willingham (1980), who discuss various aspects of management fraud, and especially its detection and deterrence.

break monetary collusion as well. Also, it would be interesting to elaborate further the idea that potential collusion may be a factor that determines when new workers should be hired. One potential extension to the present model is to develop a model that also captures framing, an issue that was ignored here.

Appendix B

The Proof of Proposition 3.1

The principal's problem of allocating working hours between production and monitoring tasks is as follows:

$$\begin{aligned} & Max_{t_A^m, t_B^m} \ [f_1(1-t_A^m)+f_2(1-t_B^m)-E(w_A+w_B)] \\ & \quad s.t. \ \ \underline{U_i}=\Delta\theta_i-P(t_j^m)\Delta\theta_i, \\ & \quad \quad \overline{U_i}=0, \ i=j=A,B, \ i\neq j. \end{aligned} \tag{B.1}$$

After substituting IR constraints into (B.1) and noticing that the principal pays w$_i$ = $\overline{\theta}_i$ - unless he receives a report r$_j$ = $\underline{\theta}_i$, in which case he pays w$_i$ = $\underline{\theta}_i$ - the principal's problem becomes:

$$\begin{aligned} & Max_{t_A^m, t_B^m} \ [f_1(1-t_A^m)+f_2(1-t_B^m)-[p^2P(t_A^m)P(t_B^m)(\underline{\theta}_A+\underline{\theta}_B) \\ & +p^2(1-P(t_A^m))P(t_B^m)(\underline{\theta}_A+\overline{\theta}_B)+p(1-p)P(t_B^m)(\underline{\theta}_A+\overline{\theta}_B) \\ & +p^2P(t_A^m)(1-P(t_B^m))(\overline{\theta}_A+\underline{\theta}_B)+ \\ & p^2(1-P(t_a^m))(1-P(t_B^m))(\overline{\theta}_A+\overline{\theta}_B) \\ & +p(1-p)(1-P(t_B^m))(\overline{\theta}_A+\overline{\theta}_B)+(1-p)pP(t_A^m)(\overline{\theta}_A+\underline{\theta}_B) \\ & +(1-p)p(1-P(t_A^m))(\overline{\theta}_A+\overline{\theta}_B)+(1-p)(1-p)(\overline{\theta}_A+\overline{\theta}_B)]. \end{aligned} \tag{B.2}$$

After optimizing (B.2) with respect to t$_A$^{m} and t$_B$^{m}, the first-order conditions boils down to:

$$\begin{aligned} & \frac{\partial\pi}{\partial t_A^m}=-f_1(1-t_A^m)+P'(t_A^m)p\Delta\theta_B=0 \\ & \frac{\partial\pi}{\partial t_B^m}=-f_2(1-t_b^m)+P'(t_B^m)p\Delta\theta_A=0. \end{aligned} \tag{B.3}$$

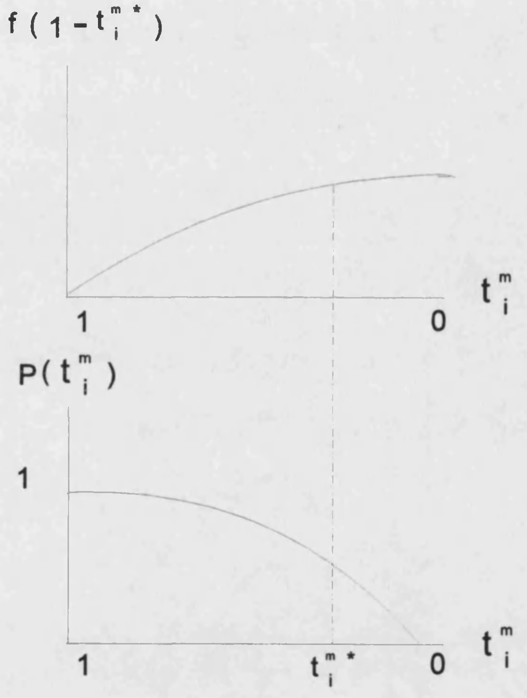
Rewriting (B.3) we get the condition which determines the optimal time allocation under the collusion-free regime:

$$P'(t_i^m)p\Delta\theta_j=f'(1-t_i^m), \quad i=j=A,B, \quad i\neq j. \tag{B.4}$$

At the optimum, the marginal productivity of each task has to be the same. If this were not the case, the principal could increase his profits by allocating working hours from the less productive task to the more productive one.

Let t_A^{m*} and t_B^{m*} be the solutions to the principal's problem. Since the problem is concave in t_i^m, a pair of solution (t_A^{m*}, t_B^{m*}) is a unique solution given any p and $\Delta\theta_i$, i=A,B. The optimal unique allocation of working hours is symmetric between the agents: $t_A^{m*} = t_B^{m*}$. Figure 3.1 further characterizes time allocation:

Figure 3.1 Optimal Time Allocation under the Collusion-Free Regime.



In the figure a vertical dash line indicates the number of working hours allocated to monitoring activities: t_i^{m*}, and at the optimum $P'(t_i^{m*})\Delta\theta_j = f'(1-t_i^{m*})$.

The optimal contract is a combination of the first- and second-best solutions: $w_i = \bar{\theta}_i$ if $r_j = \varnothing$, and $w_i = \underline{\theta}$ if $r_j = \underline{\theta}_i$. The efficient type earns an informational rent, $\Delta\theta_i$ if another agent does not provide new information to the principal. The inefficient type's IR constraint binds: $U_i = 0$. The optimal time allocation is: $t_i^{m*} < (1-t_i^{m*})$, i= A, B. □

The Proof of Proposition 3.2

The principal's problem is to allocate fixed working hours between the production and monitoring tasks under potential collusion. To break collusion the principal has to compensate A and B for revealing information they have learned: i.e., the principal has to match the gains from collusion. The principal does so by satisfying the coalition incentive constraint in a case where both A and B are efficient types and have learned the signals. This occurs with probability $p^2P(t_A{}^m)P(t_B{}^m)$. When the coalition incentive constraint is satisfied, the agents are as well off as they would be by colluding. Then they have no reason to conceal the signals they have learned. Since the rents are costly the coalition incentive constraint binds:

$$(w_A-\underline{\theta}_A)+(w_B-\underline{\theta}_B)=(\overline{\theta}_A-\underline{\theta}_A)+(\overline{\theta}_B-\underline{\theta}_B) \qquad \text{(B.4)}$$

The left-hand side represents the agents' joint utility when they report truthfully information they have learned ($r_A=\underline{\theta}_B,\ r_B=\underline{\theta}_A$); it has to match the right-hand side, which is the welfare the agents could get by concealing information ($r_A=\varnothing,\ r_B=\varnothing$).

Note that in all other cases the principal does not have to pay the agents for the information they provide. This is due to the fact that there collusion is not a problem, since favours can not be exchanged and an agent can not increase his own welfare by individually concealing the signals. After substituting (CIC) and (IR) constraints for the principal's time, the allocation problem becomes:

$$\begin{aligned}
&Max_{t_A^m,t_B^m}\ [f_1(1-t_A{}^m)+f_2(1-t_B{}^m)-[p^2P(t_A{}^m)P(t_B{}^m)(\overline{\theta}_A+\overline{\theta}_B)+\\
&p^2(1-P(t_A{}^m))P(t_B{}^m)(\underline{\theta}_A+\overline{\theta}_B)+p(1-p)P(t_B{}^m)(\underline{\theta}_A+\overline{\theta}_2)+\\
&p^2P(t_A{}^m)(1-P(t_B{}^m))(\overline{\theta}_A+\underline{\theta}_B)+p^2(1-P(t_a{}^m))(1-P(t_B{}^m))(\overline{\theta}_A+\overline{\theta}_l\\
&+p(1-p)(1-P(t_B{}^m))(\overline{\theta}_A+\overline{\theta}_B)+\\
&(1-p)pP(t_A{}^m)(\overline{\theta}_A+\underline{\theta}_B)+\\
&(1-p)p(1-P(t_A{}^m))(\overline{\theta}_A+\overline{\theta}_b)+(1-p)(1-p)(\overline{\theta}_A+\overline{\theta}_B)].
\end{aligned} \qquad \text{(B.6)}$$

After optimizing (B.6) with respect to t_A^m and t_B^m and simplifying the first-order conditions, we have a pair of equations:

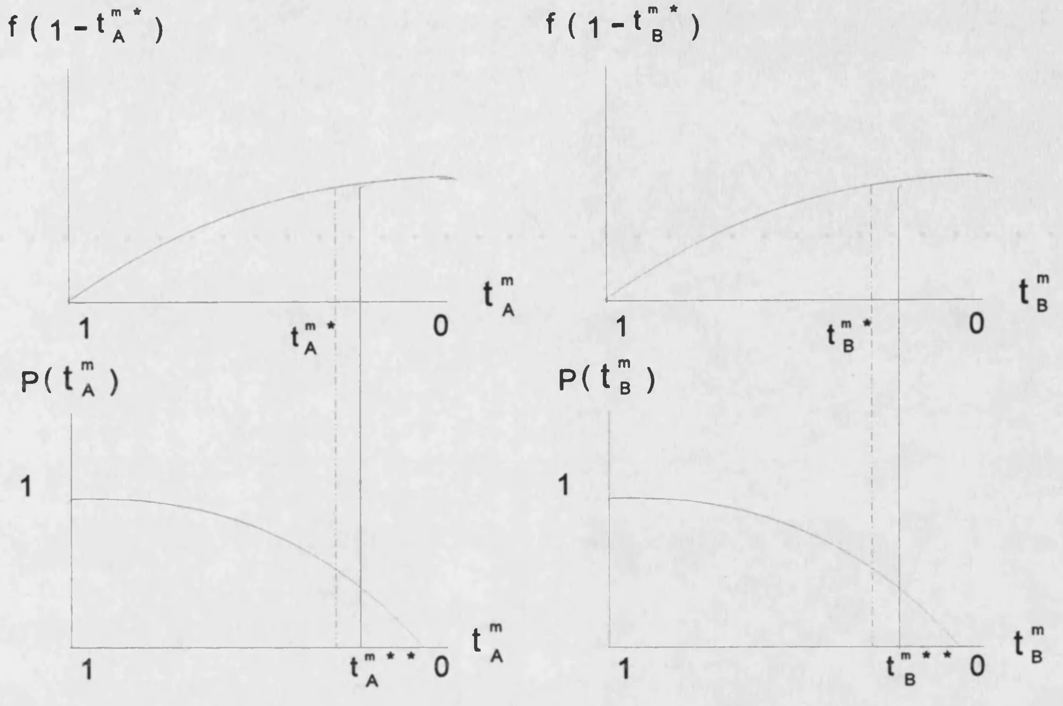
$$\begin{aligned}\frac{\partial\pi}{\partial t_A^m}&=-f_1^{\prime}(1-t_A^m)+P^{\prime}(t_A^m)p\Delta\theta_B(1-pP(t_B^m)=0\\ \frac{\partial\pi}{\partial t_B^m}&=-f_2^{\prime}(1-t_b^m)+P^{\prime}(t_B^m)p\Delta\theta_A(1-pP(t_A^m)=0.\end{aligned} \tag{B.7}$$

Let t_A^{m**} and t_B^{m**} be the solutions to the principal's problem under potential collusion. Rewriting (B.7) gives us the optimal time allocation conditions:

$$\begin{aligned}&P^{\prime}(t_A^m)p\Delta\theta_B(1-pP(t_B^m)=f_1^{\prime}(1-t_A^m)\\ &P^{\prime}(t_B^m)p\Delta\theta_A(1-pP(t_A^m)=f_2^{\prime}(1-t_B^m).\end{aligned} \tag{B.8}$$

When we compare (B.8) to (B.4), we see that due to the possibility of collusion the expected savings are smaller, since the possibility of collusion has reduced the effectiveness of the monitoring activity. At the optimum, less time is allocated to monitoring activities and more to production: $t_i^{m**} < t_i^{m*}$. In contrast to the collusion-free regime, the optimal time allocation is now interdependent, and, of course, the principal's welfare is lower. In Figure 3.2 this time allocation problem is further clarified.

Figure 3.2 Optimal Time Allocation under the Collusion-Proof Solution.



In Figure 3.2 the vertical continuous lines correspond to the optimal time allocation under the collusion-proof solution (t_i^{m**}), and the dash line shows the optimal time allocation under the collusion-free regime (t_i^{m*}), i=A, B. The optimal contract offers: $w_i = \bar{\theta}_i$ if $\{r_A = \underline{\theta}_B, r_B = \underline{\theta}_A\}$, $w_i = \bar{\theta}_i$ if $r_j = \varnothing$, $w_i = \underline{\theta}_i$ if $r_j = \underline{\theta}_i$, i=j=A, B, i≠j. And $t_i^{m**} < (1-t_i^{m**})$, i= A, B. □

Proof of Proposition 3. 3

Assume that the principal is able to hire a third agent with a fixed cost F, whose only task is to monitor. With the help of the third worker, the principal is able to implement the collusion-free outcome. The optimal time allocation corresponds to the time allocation of the collusion-free regime. A produces for (1- t_A^{m*}) hours and monitors B for t_A^{m*} hours. B only produces for (1- t_C^{m*}) hours and a third agent, say C, monitors A for t_C^{m*} hours.

Given that tasks are assigned in this way, A and B are unable to collude, because they can not exchange any favours. Note also that A can not do any favours for C, and thus collusion is not an issue. The benefits of hiring a third agent are clear, since it breaks collusion; but the cost side is the fixed cost of hiring the third worker. The principal's welfare with three agents is:

$$\pi_3 = f_1(1 - t_A^{m*}) + f_2(1 - t_C^{m*}) - \overline{\theta}_A - \overline{\theta}_B + P(t_A^{m*}) p\Delta\theta_B + P(t_C^{m*}) p\Delta\theta_A - F. \tag{B.14}$$

Recall the principal's welfare under the collusion-proof solution:

$$\pi_{cp} = f_1(1 - t_A^{m**}) + f_2(1 - t_A^{m**}) - \overline{\theta}_A - \overline{\theta}_B + P(t_B^{m**}) p\Delta\theta_A + P(t_A^{m**}) p\Delta\theta_B \tag{B.15}$$

Now, depending on $\pi_3 - \pi_{cp}$, the principal breaks collusion either with money by satisfying the coalition incentive constraint or by hiring a third agent. When (i)$\pi_3 - \pi_{cp} > 0$, the principal prefers to hire an extra worker to monitor one of the agents, and when (ii) $\pi_3 - \pi_{cp} < 0$, the principal satisfies the coalition incentive constraint. Whenever the costs of collusion are higher than the fixed cost of hiring a third worker, it is optimal for the principal to break collusion by hiring the third worker. ☐

The Proof of Proposition 3. 4

The first step is to show that blowing the whistle is a dominant strategy. The expected utility from blowing the whistle is $EU_i = \gamma_j\,(0) + (1-\gamma_j)(\Delta\theta_i+\epsilon)$, which is higher than the expected utility from not blowing the whistle: $EU_i = \gamma_j\,(0) + (1-\gamma_j)(\Delta\theta_i)$, $i=j= A,B$, $i\neq j$. The agent will always blow the whistle if possible.

Secondly, note that the expected utility from collusion is higher than the utility from not colluding. Even though whistle-blowing may occur, the agents still willing to take a gamble, because they can not be worse off by colluding than by not colluding. The expected utility from collusion is: $\gamma_i[\,\gamma_j\,(0) + (1-\gamma_j)\,(\Delta\theta_i + \epsilon)\,] + (1-\gamma_i)[\,\gamma_j\,(0) + (1-\gamma_j)\,\Delta\theta_i\,]$ $= (1-\gamma_j)\,\Delta\theta_i + (1-\gamma_j)\,\gamma_i\,\epsilon > 0$, $i=j=A,B$, $i\neq j$. Now we can also write the expected wages in the case where collusion and whistle-blowing take place: $\gamma_A\,\gamma_b(\,\underline{\theta}_A + \underline{\theta}_B) + \gamma_A(1-\gamma_b)\,(\,\overline{\theta}_A + \epsilon + \underline{\theta}_B) + (1-\gamma_A)\gamma_b\,(\,\underline{\theta}_A + \overline{\theta}_B + \epsilon) + (1-\gamma_A)\,(1-\gamma_b)\,(\,\overline{\theta}_A + \underline{\theta}_B\,)$ $= \overline{\theta}_B - \gamma_A\,\Delta\theta_B + \overline{\theta}_A - \gamma_B\,\Delta\theta_A + \epsilon(\gamma_B - \gamma_A\gamma_B) + \epsilon(\gamma_A - \gamma_A\gamma_B)$. Notice that this stands for the principal's expected wages under potential collusion when the whistle-blowing mechanism is adopted. The principal's problem is:

$$\begin{gathered}
Max_{t_A^m, t_B^m}\ [f_1(1-t_A^m)+f_2(1-t_B^m)- \\
[p^2P(t_A^m)P(t_B^m)(\overline{\theta_B}-\gamma_A\Delta\theta_B+\overline{\theta_A}-\gamma_B\Delta\theta_A \\
+\epsilon(\gamma_B-\gamma_A\gamma_B)+\epsilon(\gamma_A-\gamma_A\gamma_B)) \\
p^2(1-P(t_A^m))P(t_B^m)(\underline{\theta_A}+\overline{\theta_B})+ \\
p(1-p)P(t_B^m)(\underline{\theta_A}+\overline{\theta_2}) \\
+p^2P(_A^m)(1-P(t_B^m))(\overline{\theta_A}+\underline{\theta_B})+ \\
p^2(1-P(t_a^m))(1-P(t_B^m))(\overline{\theta_A}+\overline{\theta_B})+ \\
p(1-p)(1-P(t_B^m))(\overline{\theta_A}+\overline{\theta_B})+ \\
(1-p)pP(t_A^m)(\overline{\theta_A}+\underline{\theta_B})+ \\
(1-p)p(1-P(t_A^m))(\overline{\theta_A}+\overline{\theta_b})+(1-p)(1-p)(\overline{\theta_A}+\overline{\theta_B})]
\end{gathered} \tag{B.9}$$

Note that the term:$p^2P(t_A{}^m)P(t_A{}^m)[\bar{\theta}_B - \gamma_A\Delta\theta_B + \bar{\theta}_A - \gamma_B\Delta\theta_A+\epsilon(\gamma_A - \gamma_A\gamma_b) + \epsilon(\gamma_A - \gamma_A\gamma_b)$] stands for the expected wages under potential collusion. We see that if the probability of hard information is zero - i.e., $\gamma_i = 0$, i= A, B - then the principal's problem is equivalent to that of the collusion-proof case. And if $\gamma_i = 1$, i= A, B, the problem is identical to that of the collusion-free case. After optimising the principal's problem with respect to $t_A{}^m$ and $t_A{}^m$ and some manipulation, the first-order conditions are:

$$\begin{gathered}
\frac{\partial\pi}{\partial t_A^m}=-f_1'(1-t_A^m)+P'(t_A^m)p\Delta\theta_B[1-pP(t_B^m)((1-\gamma_B)+ \\
\epsilon(\gamma_A-\gamma_B\gamma_A))]=0 \\
\frac{\partial\pi}{\partial t_B^m}=-f_2'(1-t_b^m)+P'(t_B^m)p\Delta\theta_A[1-pP(t_A^m)((1-\gamma_A)+ \\
\epsilon(\gamma_B-\gamma_B\gamma_A))]=0.
\end{gathered} \tag{B.10}$$

And by rewriting (B.10) we get the optimal time allocation under the whistle-blowing mechanism:

$$\begin{gathered}
P'(t_A^m)p\Delta\theta_B[1-pP(t_B^m)((1-\gamma_B)+\epsilon(\gamma_A-\gamma_A\gamma_B))]]=f_1'(1-t_A^m) \\
P'(t_B^m)p\Delta\theta_A[1-pP(t_A^m)((1-\gamma_A)+\epsilon(\gamma_B-\gamma_A\gamma_B))]=f_2'(1-t_B^m).
\end{gathered} \tag{B.11}$$

When we compare (B.11) to B(8) and (B.4), we immediately observe the obvious relationship. We see that if $\gamma_A = \gamma_B = 0$, the optimal time allocation is equivalent to that of the collusion-proof case. And if $\gamma_A = \gamma_B = 1$, (B.11) boils down to (B.4), and the time allocation is identical to that of the collusion-free case.

Chapter 4

"Money or Reputation" - A Rational Theory of Blackmail

4.1 Introduction

"Even if you pay the blackmail, that is no guarantee that he won't demand more later."
(The Oxford Thesaurus. An A-Z Dictionary of Synonyms.)

"The blackmailed person to Madson: "How can I trust the boy that he won't come back to ask more money?" Madson: Don't worry, he is just a kid, he is not a professional blackmailer."
(Madson, BBC 1, 1996.)

What is the economics behind the phenomenon of blackmail? Should a victim pay a blackmailer or not? How much, if anything, should the victim pay the blackmailer? When is the blackmailer's threat to carry out his action credible? What if the blackmailer comes back to ask for more money? These are among the questions the present chapter tackles.

In recent years relatively much has been written about collusion (bribery) in organisations. Very little thought has been given to blackmail, which, in general, forms a complementary part of corruption.[1] The concept of blackmail in legal and sociological literature has originally been used to refer to payments to avoid physical harm; today it primarily refers to payments to avoid revelation of discreditable information.[2] This is precisely how blackmail is modelled in the present chapter.

Even though collusion and blackmail are closely related, there is, however, an important difference. Under collusion two parties (e.g., a manager and an auditor) act together and collude against a principal (e.g., the shareholders), for example, by agreeing on information manipulation. They enter the collusive relationship voluntarily, and after successful collusion they are both better off than by not colluding. In the case of blackmail, a blackmailer operating alone, is able to hurt a victim. The blackmailer extorts the victim by threatening to reveal a piece of information which the victim prefers to keep private. The relationship between the blackmailer and the victim is involuntary and takes the form of pure

---

[1] See, for instance, Tirole (1992) on collusion.

[2] See Hepworth (1980), who discusses several aspects of the phenomenon of blackmail from a sociological point of view.

extortion. Furthermore, after successful blackmail the blackmailer is better off, but the victim worse off compared to the case of no blackmail.

This chapter develops a dynamic model of blackmail. It takes into consideration the blackmail game, where a piece of information which an agent prefers to keep private may facilitate blackmail when the other agent, namely the blackmailer, is able to reveal that information. The crucial feature of the blackmail game is the commitment problem from the blackmailer's side. The blackmailer cannot commit not to come back in future to demand more despite the payments received in the past.[3]

The infinite horizon blackmail game has a very simple structure. Two players move sequentially in every period. First the victim decides how much to hand over as a blackmail payment, and then the blackmailer decides whether to reveal or suppress a piece of information about the victim. The equilibrium concept we use is a Markov Perfect equilibrium. That is, we consider strategies that are conditioned only on the payoff-relevant variables, and not on the entire history of the game.

We analyse first a situation where there exist no rewards for information revelation. This includes the cases where, for example, the tabloid press do not pay rewards for scandal stories, an organisation designer does not pay rewards to organisation members who turn in fellow workers, and so on. In this case it is shown that there is a unique Markov Perfect equilibrium where blackmail does not arise, and the victim pays nothing, and the blackmailer suppresses information.

Next we introduce rewards for information revelation, and consequently then the blackmailer has two potential buyers for a piece of information: the victim or the tabloid press. When a piece of information is revealed to the tabloid press, the game ends. On the contrary, if the information is suppressed, the game continues, and the blackmailer comes back in the next period to demand more money. In this case we show that in an equilibrium blackmail is an issue. The victim pays the blackmailer and information is not revealed. This Markov Perfect equilibrium is unique, and it gives us a precise prediction how much money the blackmailer will get by extorting the victim.

---

[3] Imagine, for example, a case where the blackmailer's evidence is a videotape. The victim can not be sure when buying the tape that the blackmailer does not have a copy of the tape. If the blackmailer has a copy, he can come back and ask for more money.

Our results remain the same in case where the blackmailer announces a blackmail demand in a beginning of each period. What is important for the outcome is that the blackmailer always moves after the potential blackmail payment, and the blackmailer has an option to end the game by revealing the information. This, by the way, works partially against the blackmailer, who is unable to fully exploit the victim. Interestingly, the blackmail payment does not depend on the victim's valuation of the information, but only on the external net reward.

Due to the blackmailer's commitment problem, blackmail appears in the form of an infinite stream of small blackmail payments, rather than in the form of a large one-off payment in the beginning of the relationship. Since there is nothing which forces the blackmailer to suppress the information, even though he has received money from the victim, the optimal blackmail payment turns out to be a combination of the offered external net reward and a blackmail premium. The premium compensates the blackmailer for not taking money from the victim and revealing information anyway. The case when potential rewards are the blackmailer's private information is also considered. It is shown that information about the victim will be revealed with positive probability.

Before going any further, a few words about the related literature are in order. This paper naturally relates to the recent theoretical literature on corruption in organisations. As said above not very much has been done so far. For example, Koffman and Lawarrée (1993) explicitly rule out the possibility of blackmail. Tirole (1992) also mentions the phenomenon of blackmail, but does not model it.

To the best of our knowledge, the only serious attempt to incorporate blackmail in the economic analysis of organisations is done by Mogiljanski (1994). Mogiljanski develops a concept of regulatory blackmail in a static adverse selection model à la Laffont and Tirole (1991). She defines blackmail as a situation where a regulator threatens a firm that it will conceal some favourable information from a principal, unless the firm pays the regulator ex post. The regulator has all the bargaining power when proposing a "take it or leave it offer" to the firm. Note that if the firm declines to pay, the regulator is indifferent between reporting the information or concealing it. Then the question becomes, whether the blackmailer will release or conceal the information if the firm declines to pay. She assumes that the firm's threat to conceal information is a subgame perfect. However, if one follows the paradigm of the principal-agent models strictly in this case, the regulator should behave as the principal prefers by reporting a piece of information to the principal. Then, of course, the threat is not credible, and blackmail is not an issue. Note also that if the principal commits to an infinitesimal reward for information revelation, the blackmail there is not an issue anymore, since then the regulator will certainly report. Mogiljanski's study ignores the dynamic aspect of the phenomenon of blackmail, which we choose to emphasize. We put aside collusion and concentrate purely on blackmail in a dynamic framework. In particular, we focus on the way the potential surplus will be shared between the victim and the blackmailer.

Outside of organisation theory Konrad and Skaperdas (1995) consider credible threats in the extortion business between a gang and shopkeepers. They emphasize a fundamental credibility problem: i.e., whether the gang will actually carry out a threat if a shopkeeper declines to pay. There the gang has to make a costly up-front investment that is unobserved by the shopkeepers, and these investments facilitate punishments in a case when the shopkeepers decline to pay. They show that only a no-extortion equilibrium exists if there is only one shopkeeper. This equilibrium disappears when the number of shopkeepers increases, and then the only subgame perfect equilibrium that will remain is an extortion equilibrium. They do not consider the blackmailer's commitment problem: i.e., that the blackmailer may come back. In order to derive their main results, they have to make two very strong assumptions. They assume that if the gang has invested in punishment technology, it has to always use it if a shopkeeper declines to pay. They also assume that the gang can not punish if the shopkeeper pays the gang. In the present chapter both of these assumptions are relaxed.

Recently some authors have also considered collusion in a dynamic framework. Acemoglu (1995) develops a dynamic model of implicit collusion between an auditor and a manager. Martimort (1996) in turn has proposed a model of self-enforcing collusion by modelling a static adverse selection problem as an infinitely repeated game. However, none of those papers examines the phenomenon of blackmail and the question of how the benefits of information suppression will be shared.

The outline for the rest of this chapter is as follows. Section 4.2 describes the model. The analysis is carried out and the main results are provided in section 4.3. In section 4.4 some interpretations and potential extensions are discussed. Section 4.5 concludes.

4.2. The Model

The Players

Consider a model with two agents: a blackmailer (B) and a blackmailed person (V); we refer to V also as the victim. V has on his possession a piece of information (I), which he prefers to keep private: i.e., not to share with others. The monetary value of V's privacy (or, say, reputation) is equal to $v(I) > 0$, and $v(I) = 0$ if information is revealed. In particular, we assume that the value of privacy is a per-period benefit. Later, we use v as a short-hand notation for v(I). The time horizon we consider is infinite. We also assume that information revelation causes the victim a permanent damage. That is, the victim never get back his privacy.

The blackmailer, B also has access to a piece of information I, and is able to reveal it to a third party. The cost of information revelation for B is a fixed cost, $c \geq 0$. We assume, for the sake of simplicity, that B also learns v.[4] The blackmailer does not derive any direct utility himself from releasing discreditable information about V, and B hopes that the victim will pay him to suppress the information. In his article Hepworth (1990) puts this nicely: "At the heart of reputational blackmail lies the willingness of the blackmailer to exploit the victim's desire to prevent others sharing a secret".

The model we consider has many potential interpretations. Within an organisation, V may be a civil servant who has taken bribes from a contractor, and B is another civil servant who has observed this and blackmails V. Alternatively, V may be a politician who has had an "affair", and B is a person who blackmails the politician by threatening to reveal that information to the tabloid press. This latter example is more closely related to the very idea of reputational blackmail ("newspaper blackmail") on which we want to concentrate here.

So far the model has a simple feature where one agent prefers his privacy, and the other agent is able to provide that by remaining silent. The blackmail game has a very intuitive and familiar interpretation, namely that of a seller and a buyer. Here the blackmailer

---

[4] This is not a restricting assumption. Later, it is shown that the optimal blackmail payment does not depend on the victim's valuation at all.

(the seller) is able to provide the victim (the buyer) a service by remaining silent. The question is how much the buyer has to pay for this service. Or, to put it otherwise, we can even state that trade has already occurred, because the victim has on his possession an infinite surplus stream from privacy: $v(I)/(1-\delta)$, where δ is a common discount factor. Now the question becomes, how will the benefits from privacy be shared?

Note that in contrast to a normal seller-buyer model, the blackmailer is not needed here to create a surplus, since it already exists. However, the blackmailer is able to destroy the surplus permanently simply by revealing the information. Note that if the blackmailer reveals the information, he basically then also destroys his only asset, and certainly will not get anything out of the relationship in future.

The trade here is not a one-off event, since the victim's payment to the blackmailer does not end the game. In every period, given that no information has been revealed in the past, a new surplus arrives and the blackmailer is around demanding a share of it. Before describing the blackmail game, we make the following assumptions.

Assumption 4.1: No contracts can be conditioned on whether the blackmailer comes back in future to demand more money.

Assumption 4.1 is a crucial one, and states the fundamental commitment problem we believe captures the very essential feature of blackmail. The blackmailer can not commit not to come back in future to demand more money. That is, the blackmailer is unable to commit not to exercise a profitable action. For example, this is the case of unfortunate and well-documented practices among small businesses who have to pay "protection money" to criminals just to be able to run their businesses. In those cases blackmail exists exactly in the form of "small" blackmail payments criminals collect from small businesses, say every week, rather than in a form of a large one-off payment. The model presented here describes therefore also racketeering.[5]

---

[5] Interestingly, Konrad and Skaperdas (1995) do not consider this feature in their extortion analysis.

Assumption 4.2: No contracts can be conditioned on information suppression.

According to assumption 4.2 such contracts are not enforceable. The only reason why the blackmailer may suppress information is that it is individually rational to do so. That is, even if the victim hands over money, it is no guarantee that the blackmailer will suppress the information. This last point further emphasizes the fact that the relationship between the victim and the blackmailer is a game and not an enforceable contract.

How much money the blackmailer is able to extort from the victim is what we call blackmail, and a blackmail payment is labelled as m. There are presumably many alternative ways to model blackmail, and we have chosen to stick with that of dynamic games.[6]

Timing, Strategies, and the Solution Concept

As stated above, the time horizon we consider is infinite ($T=\infty$). Within every period of the blackmail game, and given that information has not been revealed, there are two sequential stages (see figure 4.1). First, the victim decides how much to hand over as a blackmail payment (m). In stage 2, the blackmailer decides whether to reveal (r) or suppress (s) the information.[7] The blackmailer either accepts V's offer and suppresses the information, or rejects it and reveals the information. Since the blackmailer pockets m in any case, we can further simplify the blackmailer's action in each period. He either reveals or suppresses information. If B suppresses information, the game continues and in a next period stages 1 and 2 are repeated, and so on. Information revelation ends the blackmail game immediately in that very period, because after that there is no valuable information left anymore. It is in the blackmailer's hands whether the game continues or not.

The introduction of an external reward (R) has a crucial effect on the blackmail game, since then the blackmailer is able to sell a piece of information also to the third party. We do

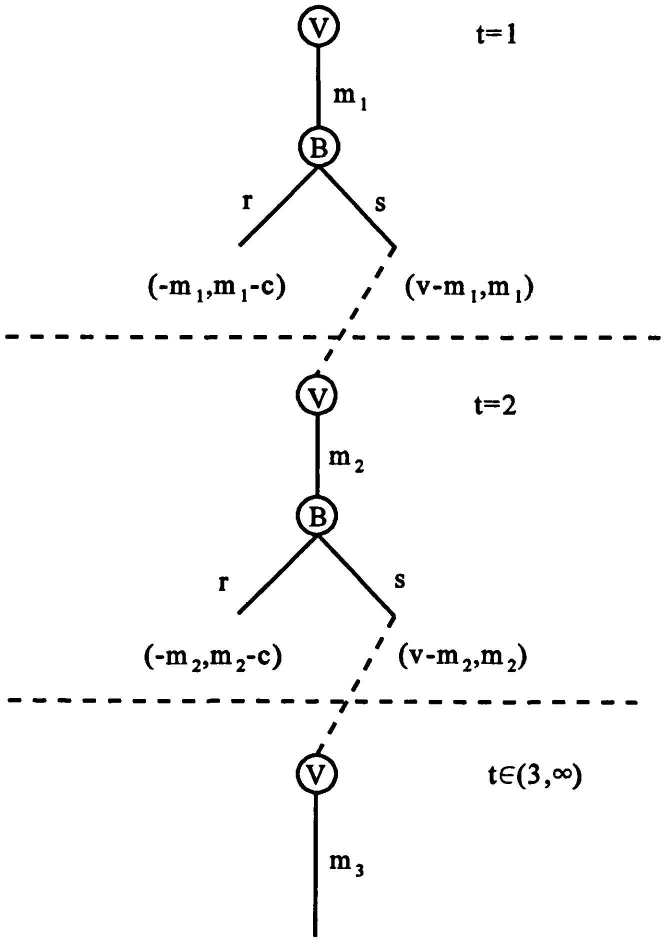
---

[6] The Rubinstein-Ståhl model could provide another way to proceed. However, there an acceptance by one party ends the game, which is not valid here. There a contract between players is an enforceable contract; here the relationship is a non-cooperative game.

[7] Later, it is shown that this order of moves is preferred by both the victim and the blackmailer. The blackmailer does not want to move first, but he prefers to move after the victim has handed over the blackmail payment.

not incorporate the third party explicitly in the model, and it is assumed that when offered, the exogenously given R is available in every period to the blackmailer, who exposes discreditable information about the victim.

Figure 4.1 The Blackmail Game. The victim moves first.



Rather than finding all the perfect equilibria of the blackmail game, we restrict ourselves to the Markov Perfect Equilibrium (MPE). That is, we consider only Markov strategies, where player i's strategy in period t does not depend on the whole history of the game, but only on the variables that affect its present period's payoffs.[8] Once we have constrained the set of strategies to Markov, the structure of the game becomes stationary, and on the condition that information has not been revealed earlier, the blackmail game looks similar in every period.

Consider next the players' strategies. The victim moves first by handing out the blackmail payment m_t. m_t is chosen from a finite set (0,v). The blackmailer's strategy space is discrete: he either suppresses (s) information or reveals (r) it: $b_t \in \{s,r\}$. Define the available net reward for the blackmailer in the beginning of period t as $\theta_t = (R_t\text{-}c)$. We say that θ_t describes "a state of the system" in period t.

Define the history of the game in period t as $h_t=\{(m_1,b_1,\theta_1), (m_2,b_2,\theta_2), (m_3,b_3,\theta_3), ...(m_{t-1},b_{t-1},\theta_{t-1})\}$. Note that the game has a history in period t only when no information has been revealed earlier. In period t the only aspect of the history that directly affects the victim's action in the present period is $\theta_{t.}$ Interestingly, since the external net reward is exogenously given, θ_t depends only on what B did in period t-1. Thus if $b_{t-1} = r$, then the game would have ended, and if $b_{t-1} = s$, then $\theta_t = (R_t\text{-}c)$. Therefore, the victim's strategy depends only on θ_t: $m_t\ (\theta_t)$. Due to the sequential timing of moves in each period, the aspect of the history that directly affects the blackmailer's payoff is θ_t, but the blackmailer's payoff relevant history also includes m_t, the blackmail payment handed over by the victim in that very period: $b_t\,(m_t,\theta_t)$. Moreover, the reactions by the victim and the blackmailer do not depend on the calendar time, but only on the state variable and the state of the system. The victim's strategy is then a reaction function: $m_t = R(\theta_t)$, and the blackmailer's strategy is: $b_t = R(m_t,\theta_t)$.

Both the victim's and the blackmailer's objective is to maximize the present discounted value of their payoffs:

$$\sum_{t=1}^{\infty} \delta^{t-1} \Pi_t^V(m_t\ ,b_t) \tag{4.1}$$

---

[8] See Maskin and Tirole (1997) for the Markov Perfect Equilibrium.

$$\sum_{t=1}^{\infty} \delta^{t-1}\Pi_t^B(m_t\ ,b_t) \qquad (4.2)$$

Define V^V and V^B respectively as net present values of being the victim and the blackmailer at the beginning of a period. Now we can use dynamic programming and write both the victim's and the blackmailer's payoffs (intertemporal profits from period t onwards) as valuation functions:

$$V_t^V = Max_{m_t}\Pi_t^V(m_t, b_t) + \delta V_{t+1}^V. \qquad (4.3)$$

$$V_t^B = Max_{b_t}\Pi_t^B(b_t, m_t) + \delta V_{t+1}^B. \qquad (4.4)$$

Now we are in a position to write down the payoffs in terms of the strategies. The victim's payoffs in the case of information suppression and revelation are respectively:

$$V_t^V = \Pi_t^V(m_t, s) + \delta V_{t+1}^V = (v - m_t) + \delta V_{t+1}^V. \qquad (4.5)$$

$$V_t^V = \Pi_t^V(m_t, r) = -m_t. \qquad (4.6)$$

If the blackmailer suppresses the information in period t, the victim's payoff is the net benefit $(v\text{-}m_t)$ plus the continuation value. In the case of information revelation, the victim loses his reputation immediately as well as the payment m_t, which the victim handed over. Note that this means that the victim pays a blackmail payment out of his pocket, and that v is realised at the end of each period only in the case of information suppression. The blackmailer's payoffs are:

$$V_t^B=\Pi_t^B(m_t,s)+\delta V_{t+1}^B=m_t+\delta V_{t+1}^B. \tag{4.7}$$

$$V_t^B=\Pi_t^B(r,m_t)=(m_t-c). \tag{4.8}$$

In the case of information suppression, the blackmailer pockets m_t immediately, and he gets the continuation value as well. If the blackmailer decides to reveal the information he bears a cost of revelation c as well, and receives (m_t - c), and the blackmail game ends immediately in that very period.

4.3 The Analysis

4.3.1 No Rewards for Information Revelation

This section considers a case where no rewards exist for an agent who reveals some discreditable information in public. In practice this includes cases where the tabloid press pays no rewards for scandal stories, an organisation designer does not reward workers who disclose information about fellow workers' activities, and so on. The case of "no rewards" further clarifies the idea that the blackmailer does not enjoy any direct utility from releasing discreditable information about the victim.

The blackmail game proceeds as follows. In every period t, given that information has not been released earlier, the victim hands over m_t to the blackmailer. The blackmailer's strategies are then simply: to suppress (s) or reveal (r) the information. That strategy corresponds to the decision of ending the game or letting it to continue.

In the following we argue and prove that in the case of "no rewards", blackmail will not arise. The victim pays nothing and the blackmailer suppresses the information. In order to derive an optimal blackmail payment, we have to find out the blackmailer's optimal responses to an arbitrary blackmail payment. It is shown that B's best response to any blackmail payment is to suppress the information. And, of course, then the victim optimally pays nothing for information suppression.

Consider period t and the blackmailer to whom the victim has handed over m_t. If the blackmailer suppresses the information, he receives m_t and the game continues. If B reveals the information, the game ends immediately in period t, and the blackmailer receives (m_t - c). Note that once the information is revealed the blackmailer can not come back in future. In fact, a piece of information is like an asset for the blackmailer, and revelation of the information destroys this asset. Thus we have Lemma 4.1:

Lemma 4.1 Given that the victim has handed over a blackmail payment, the blackmailer's best response is to suppress information.

Proof: Suppose not, and assume that B reveals the information, thus ending the game immediately. In that case B gets $(m_t\text{-}c) < m_t + \delta V^B$, which he would get by suppressing information. Clearly, it is optimal for the blackmailer to suppress the information, and we have a contradiction, and information suppression is the blackmailer's optimal response. □

Consider next in turn the victim's problem of choosing m, when information has not been released earlier. Note that if the information had been released in period t-1, the game would have ended and V's action would be irrelevant. In the following we show that an optimal offer $m^*=0$, which makes the blackmailer suppress the information and the game continues. Therefore we have Lemma 4.2:

Lemma 4.2 Given that the blackmailer suppresses information after any arbitrary blackmail payment, the victim's optimal response is $m^* = 0$.

Proof: Suppose not, and assume that the victim hands over $m>0$, given that the blackmailer always suppresses the information. In this case the victim receives $(v\text{-}m) + \delta V^V < v + \delta V^V$, which the victim receives when he pays $m=0$. Thus, we have a contradiction, and the victim's optimal and unique best response is $m^* = 0$. □

The main result in the case when there are no rewards is stated as Proposition 4.1:

Proposition 4.1 In the case of "no rewards" there is a unique MPE of no blackmail. The victim pays nothing and the information is suppressed; $m^*=0$, $V^V = v/(1-\delta)$, $V^B = 0$.

Proof: We want to show that a pair of strategies, ($m^*=0$ and s) forms an equilibrium. We refer here to "one-period deviation - principle". By playing his equilibrium strategy, $m^* = 0$, the victim gets: $v/(1-\delta)$, which is greater than $(v-\epsilon) + \delta v/(1-\delta)$, which is what he would get by deviating and handing over $m=\epsilon$, and then conforming to the equilibrium strategy. By playing the equilibrium strategy, the blackmailer receives: $0 + \delta V^B > 0 - c$, which he would get by deviating and revealing the information. In sum, neither the victim nor the blackmailer prefers to deviate from the equilibrium path. ☐

The economics behind Proposition 4.1 is strikingly simple and yet it confirms the intuition Hepworth (1980) provides: "If the blackmailer is unable to persuade the victim he has access to a receptive audience he is powerless to commercialise or gain any other kind of advantage from the information which has fallen in his hands." In short, when a piece of information is not valuable to anyone apart from the victim, the blackmailer is unable to extort money from the victim.

Note that it is not important for the above result that there is a positive cost of information revelation. Even if c went to zero, information suppression would remain the blackmailer's optimal response, since by suppressing information he would get $0 + \delta V^B \geq 0$. More precisely, information revelation is then weakly dominated by information suppression. And after the elimination of dominated strategies, information suppression remains a weakly dominant strategy.The victim would choose $m^*=0$, the blackmailer would suppress information, and the "no-blackmail equilibrium" would survive.

So far we have assumed that the victim moves first, and the blackmailer is passive, and only reacts to the blackmail payment. Would the blackmailer ever prefer to move first? It is clear that this is never the case, since by moving first the blackmailer can not receive anything which he does not get by waiting for the victim's blackmail payment. The blackmailer does not want to move first, but waits until the victim has handed over the payment, and then moves. But here also the blackmailer is a "passive player". Since we consider blackmail or extortion in terms of how much money the blackmailer is able to extort from the victim, a far more interesting and relevant case for us is a situation where the blackmailer is able to announce a blackmail demand, d_t in the beginning of period t. See figure 4.2 in Appendix C. This is actually a very interesting case, since it is a kind of robustness test of Proposition 4.1.

Suppose now that B moves first, and announces a blackmail demand, $d_t > m^*$. How does the victim react to the blackmailer's demand? He can either accept it and pay d_t or he can reject it. Even if he rejects the blackmailer's demand, the victim still has to decide how much to hand over as a blackmail payment. But now the set-up is exactly identical to the case where the victim moved first, and of course it is optimal to pay the same amount as before, $m^* = 0$. And for the blackmailer it is optimal to suppress the information. Consequently, no matter what the blackmailer's demand is, the victim pays only $m^*=0$.

The logic behind the result remains the same even though the blackmailer moves first. What is crucial for the result is that he also moves last. He moves after the blackmail payment has been handed over. As the last mover the blackmailer has an option to end the game by revealing information. Here the last mover has a disadvantage, since he will always prefer that the game continues, and this destroys the credibility of his threat of terminating the game by revealing information if his blackmail demand is not matched. Consequently, we have Corollary 4.1:

Corollary 4.1. If the blackmailer moves first by making a blackmail demand ($d_t > m^*$), the result of no blackmail holds. The victim pays nothing, and the information will be suppressed.

Proof: The logic is exactly similar to that of the proof of Proposition 4.1. Assume that B has announced a blackmail demand $d_t > m^*$. V can either pay this, in which case he gets $(v - d_t) + \delta V^V$. Alternatively, he can reject B's demand and pay $m^* = 0$, which is the optimal payment he would choose when moving first. From earlier we know that B's optimal response to that is to suppress the information, which gives V: $v + \delta V^V > (v - d_t) + \delta V^V$. Therefore, no matter what B asks, the victim hands over m*, and B suppresses the information. □

This an important and interesting result. Indeed, the "no-blackmail" result will hold even if the blackmailer moves first by announcing a blackmail demand. The reason behind this is that the blackmailer is unable to make use of his threat of information revelation, since he is the agent who moves last. In particular, he moves after the victim has handed over m*. In effect, this resembles a situation where the victim makes a "take it or leave it" offer, and where the last mover has a disadvantage, since it is optimal for him to accept the payment and suppress the information.

4.3.2 Rewards for Information Revelation

In this section we consider whether the "no-blackmail" result will hold if there is a reward for information revelation. In practice, for example, the tabloid press do pay rewards for "scandal stories"; similarly an organisation designer may reward a worker who exposes wrongdoing in the workplace, and so on. Here we analyse whether the victim now can decline to pay anything, and how much money the blackmailer is able to extort from the victim.

Note that now the blackmailer can sell a piece of information also to a third party, which values information as well. However, there is an important difference whether a piece of information is sold to the tabloid press or to the victim. In the former case, revelation of the information will end the blackmail game, since the blackmailer no longer has valuable information on his possession. That is, the value of his asset has disappeared. In the latter case, there is nothing to keep the blackmailer from returning back to demand more money in future despite the payments he received earlier. From the victims's point of view, it is not optimal to hand over a blackmail payment which matches the reward offered by the third party, since the blackmailer would be back asking for more in the very next period.

Recall that the relationship between the victim and the blackmailer is a game, and not an enforceable contract. Therefore, there is nothing to keep the blackmailer from taking money from the victim and revealing the information anyway. The reward for information revelation represents here the blackmailer's opportunity cost of not releasing information.[9] Before proceeding any further, we make the following assumption:

Assumption 4.3 An external reward, R is small compared to the damage of information revelation to the victim and higher than the cost of information revelation to the blackmailer: $c < R < v/(1-\delta)$.

If assumption 4.3 were violated, and (R-c) was greater than the damage, the blackmailer would not bother to blackmail, but would end the blackmail game by revealing the information and collecting the net reward in the very first period.[10]

We start again by considering B's best responses to V's offer. Now the victim has to take into consideration that if he does not pay the blackmailer, the blackmailer may reveal the information, in which case the victim loses his reputation immediately. In the next lemma we derive an optimal payment m* which makes the blackmailer indifferent between releasing information and suppressing it.

Lemma 4.3 There exists $m^*=(R-c)(1-\delta)/\delta$ *such that if V's offer* $m < m^*$, *then B's optimal response is to reveal the information. Alternatively if V's offer* $m \geq m^*$, *then B's optimal response is to suppress the information.*

Proof: First we derive an optimal payment m*. Assume that V offers m, and the blackmailer suppresses the information. In this case B gets: $m + \delta V^B$. If B reveals the information, he keeps m that V handed over, but he also receives a net reward (R-c). The payoff for information revelation becomes: $m + (R-c)$. The victim's problem is to choose m he hands over to the blackmailer in every period so that B is just indifferent between suppressing and

---

[9] The role of an external reward is quite similar to that of the outside option in the Rubinstein-Ståhl Bargaining model. Here, however, the external reward is available to the blackmailer in every period, even after the victim's payment.

[10] Interestingly, in this case it might also happen that the victim sells his own story, since that would be the profit-maximizing strategy. Presumably, this is why we see in practice that occasionally celebrities sell their own scandal stories to the tabloid press, and indeed this is rational behaviour.

revealing the information: $m + \delta V^B = m + (R\text{-}c)$. That is, $m + \delta m + \delta m^2 + ... = m + (R\text{-}c)$; and after some manipulation, this gives $m^* = (R\text{-}c)(1\text{-}\delta)/\delta$. The claim is that if V's offer $m < m^*$, then B's best response is to reveal the information. Assume that V hands over $(m^*\text{-}\epsilon) < m^*$. If B suppresses the information, it yields a payoff : $(m^*\text{-}\epsilon) + \delta V^B = (m^*\text{-} \epsilon)/(1\text{-}\delta)$. Alternatively, if B reveals the information, the payoff is: $(m^*\text{-}\epsilon) + (R\text{-}c)$. Now it is clear that B will reveal the information, since $(m^*\text{-}\epsilon) + (R\text{-}c) > (m^*\text{-} \epsilon) /(1\text{-}\delta)$, where $m^*= (R\text{-}c)(1\text{-}\delta)/\delta$. Alternatively, if the victim offers $m \geq m^*$ - say, for example, $m = m^*$ - the blackmailer's best response is to suppress the information. ☐

An important economic insight follows immediately from Lemma 4.3: the blackmailer will remain silent as long as the victim hands over $m^* = (R\text{-}c)(1\text{-}\delta)/\delta$. If the victim deviates from this by handing less than m^*, the blackmailer will reveal the information; and this ensures him a payoff equal to that of the continuation value. In an equilibrium path, V is willing to pay B for information suppression, and the blackmailer is willing to suppress the information. We have a case of successful blackmail.

The victim's problem of choosing an optimal offer m^* becomes now quite straightforward, since he knows that the blackmailer will reveal the information and end the game if $m < m^*$. Thus, we have Lemma 4.4:

Lemma 4.4 Given that the blackmailer will reveal the information if $m < m^*$ and suppress it if $m \geq m^*$, then the victim's optimal response is to hand over m^*.

Proof: The proof of Lemma 4.4 is included in the proof of Proposition 4.2 below, and is thus omitted here.

The main result in the case when there is an external reward for information revelation is stated as Proposition 4.2:

Proposition 4.2 When there is a reward for information revelation, there is a unique MPE of blackmail. The victim pays the blackmailer m in every period, and the blackmailer suppresses the information;* $m^* = (R-c)(1-\delta)/\delta$, $V^V = (v-m^*)/(1-\delta)$, $V^B = m^*/(1-\delta)$.

Proof: We want to show that a pair of strategies ($m^*-\epsilon$, r and m^*, s) forms an equilibrium. Once again we refer to the "one deviation only-principle". Given that V and B play equilibrium strategies, we show that neither player prefers to deviate. The victim's equilibrium payoff is: $(v-m^*) + \delta(v-m^*) + \delta^2(v-m^*) + \ldots. = (v-m^*)/(1-\delta)$, where $m^* = (R-c)(1-\delta)/\delta$. Suppose that he deviates from his equilibrium strategy by handing over $(m^*-\epsilon)$. Then, for a moment, the victim potentially has on his possession $v-(m^*-\epsilon)$ from that period. However, we know that in this case B will reveal the information immediately. The victim loses his reputation, and the blackmail game ends; and the victim's payoff is: $-(m^*-\epsilon)$. The victim does not want to deviate. The blackmailer's equilibrium payoff is: $(m^*) + \delta(m^*) + \delta^2(m^*) + \ldots. = (m^*)/(1-\delta)$. If B deviates and reveals the information, he pockets $m^* + (R-c)$ and the game ends. However, by playing his equilibrium strategy, he receives: $m^* + \delta V^B \geq m^* + (R-c)$, and $V^B \geq (R-c)/\delta$. The blackmailer has no incentive to deviate from equilibrium. ☐

Note that a blackmail payment $m^* = (R-c)(1-\delta)/\delta$ does not depend on the victim's valuation (v) at all. It only depends on the external net reward and a discount factor δ. Now we can see that m^* increases in (R-c) and decreases in δ, the latter meaning that when the blackmailer becomes less patient (δ ->0), the optimal blackmail payment increases. Interestingly, the blackmailer is able to get more money by extorting the victim than by selling the information directly to a third party. The intuition behind this is that the victim has to compensate the blackmailer for not taking money (m^*) and revealing the information anyway. Thus we have Proposition 4.3:

Proposition 4.3 In an equilibrium path the blackmailer gets more by extorting the victim than the potential net reward (R.-c) he would get by selling a piece of information directly to the third party.

Proof: We know that $m^* = (R-c)(1-\delta)/\delta$. The flow payment that would match the external net reward is $(R-c)(1-\delta)$, which we label as M. Now $m^*/M = 1/\delta$. And $1/\delta > 1$, since $\delta < 1$. ☐

We see from Proposition 4.3 that the more impatient the blackmailer is, the bigger is the premium which the victim has to pay over the external net reward.

As in the earlier section, we see immediately that the blackmailer prefers for the victim to move first. By moving after the victim, the blackmailer is able to get $(R-c)/\delta$. By moving first the blackmailer is able to get at most (R-c). Then it is clear that the blackmailer, as well as the victim, prefers that the victim moves first and the blackmailer last.

What if the blackmailer is able to announce a blackmail demand before the victim moves? Intuitively, now the blackmailer should be in a stronger position, since he can make profits by revealing information to the third party if the victim declines to pay. One would expect that the blackmailer would get more out of the victim than in the case considered so far. Interestingly, and against intuition, this is not the case. In arguments which are very similar to those in section 4.2, it can be shown that the equilibrium payoffs do not change. Irrespective of what the blackmailer demands, the victim hands over $m^*= (R-c)(1-\delta)/\delta$, and the blackmailer will suppress the information. He does not get any more out of the victim even though he moves first by announcing a blackmail demand. And thus we have:

Corollary 4.2 When there exists a reward and the blackmailer demands $d_t > m^*$, then in an equilibrium path the result of Proposition 4.2 holds: the victim pays m^* in every period and the information is suppressed.

Proof: Assume that B moves first by announcing a blackmail demand $d_t > m^*$. The victim can either pay it, and then his payoff is : $(v- d_t) + \delta V^V$, since B suppresses the information. Alternatively, the victim can reject the demand, and hand over less than the demanded d_t. Suppose that the victim hands over $(d_t-\epsilon) > m^*$. How will the blackmailer react? If B reveals the information, he gets: $(d_t-\epsilon) + (R-c)$. However, if he suppresses the information he receives:$(d_t-\epsilon) + \delta V^B$. Now, we know that since $m^*= (R-c)(1-\delta)/\delta$, we can write $(R-c)= m^*\delta/(1-\delta)$. Then, by revealing the information, B gets: $(d_t-\epsilon) + m^*\delta/(1-\delta)$. And now since the victim is ready to pay $(d_t-\epsilon)$ in every period for information suppression, the blackmailer gets $(d_t-\epsilon) + \delta(d_t-\epsilon)/(1-\delta)$ by suppressing the information. Clearly, it is a dominant strategy for the blackmailer to suppress the information, since $(d_t-\epsilon) + \delta(d_t-\epsilon)/(1-\delta) > (d_t-\epsilon) + m^*\delta/(1-\delta)$. But now it is evident that V can lower his payment all the way down to m* - i.e., the optimal blackmail payment which the victim chose when moving first. And from earlier we know that the blackmailer will suppress the information. This concludes the proof. □

So far we have seen that in an equilibrium, information is not revealed, and the blackmailer is able to extort money from the victim only in the case when there exists a third party which values the information as well. In the next section we consider whether the fact that information about the potential reward is the blackmailer's private information will change this result, and, in particular, whether a blackmailer who does not have access to an external reward is able to extort money from the victim.

4.3.3 If a Potential Reward Is Private Information

Will information be revealed if a potential reward is the blackmailer's private information? In the following we consider a case where the victim of "newspaper blackmail" does not know whether the blackmailer has been promised a reward or not. Indeed, now we are able to show that it may well happen that the information about the victim is going to be revealed in an equilibrium.

Perhaps the simplest way to model this idea is to assume that the victim does not know whether or not the blackmailer will get a reward, R if he reveals the information. He only knows a prior probability $\beta=\text{prob}\{R=R\}$ and that with probability $(1-\beta)=\text{prob}\{R=0\}$. Recall that we use the Markov Perfect equilibrium as a solution concept, and note that now these prior beliefs are part of the “state of the system”. Consider next the victim's (the uninformed party's) problem of choosing a blackmail payment to hand over to the blackmailer.

The victim knows that if he hands over m=0 he faces the probability β that the blackmailer will reveal the information immediately and the probability $(1-\beta)$ that the blackmailer will suppress the information.The problem from the victim's point of view is that both blackmailer types prefer to behave ex ante as if they had been offered R. Consequently, if the victim wants to be sure that no information is revealed, he has to pay $m^*=(R-c)(1-\delta)/\delta$. The intuition for this is straightforward. If the victim hands over nothing, the information will be revealed in the very first period with probability β or with probability $(1-\beta)$. If the victim instead hands over $m = m^*$, both blackmailer types will suppress the information. We have a pooling equilibrium in which the victim is unable to distinguish the blackmailer's type without taking a risk by paying nothing. Thus we have:

Proposition 4.4 When R is the blackmailer's private information, and the victim moves first by offering m=0, then the information about the victim will be revealed with positive probability, and if m = m the information will be suppressed with probability one.*

Proof: Consider first the case where the victim hands over $m = 0$. We know from earlier that the blackmailer who has been offered R will reveal the information, and the blackmailer who has not been offered R will suppress the information.V's expected utility is :$\beta(0) + (1-\beta)(v/(1-\delta))$. If the victim hands over m^*, the expected utility is: $(v-m^*)/(1-\delta)$, since both types will suppress the information. Given that $\beta > m^*/v$, it is optimal to hand over m^*. □

Now it is clear that, depending on probabilities and a reward, we may end up with the case that in an equilibrium the information is revealed in the very first period. This may happen, for instance, if the external reward is large, but the probability of it is low, since the victim may take a gamble and hand over $m = 0$. In this case, the blackmailer who has been promised a reward will reveal the information. Note that if the information has not been revealed when $m=0$ has been handed over, the victim faces a blackmailer who has certainly not been promised a reward. After this, the victim optimally pays $m=0$ from that period onwards.

In sum, by paying m^*, the victim is unable to distinguish between the blackmailers' types, since both types will suppress the information; but it is certain that the information will not be revealed. And by paying $m=0$, the victim is able to distinguish between the types after the very first period. The blackmailer either reveals or suppresses the information, and that behaviour reveals the blackmailer's type completely. Of course, in practice the victim's behaviour is very much related to risk aversion. If the victim is very risk-averse, he will pay money even if the probability of facing a blackmailer who has been promised a reward is very low.

What if the blackmailer moves first by announcing a blackmail demand? The problem is that both blackmailer types announce a demand, since it does not cost anything. The blackmailer who has not been offered an external reward also tries also to convince the victim that he has on his possession a potential reward, R.

Interestingly, the victim's problem remains the same as above. If the victim wants to be sure that no information will be revealed, then he has to pay $m^*=(R-c)(1-\delta)/\delta$ as in the case where the victim moved first. The victim does not learn anything new about the blackmailer's type from the blackmail demand, since both types make the demand. It is only after the victim has not paid anything, that he learns something new, since the blackmailer's action reveals his type completely. Therefore, we have:

Corollary 4.3 When a reward is the blackmailer's private information, and the blackmailer moves first by demanding $d_t > m^*$*, then the information about the victim will be revealed with probability* β *if the victim declines to pay* $(m=0)$*; and if the victim hands over* $m=m^*$ *the information will be suppressed with probability one.*

Proof: The logic of proof is exactly the same as above, and in Corollary 4.2, and is thus omitted here.

Interestingly also, now information about the blackmailer's type is revealed only in the case where the victim declines to pay (m=0), since then the blackmailer who has been promised a reward will reveal the information, and the blackmailer who does not have access to a reward will suppress the information.

The case when only the blackmailer knows whether he has access to R changes the possibility of information revelation in an equilibrium. Earlier we have shown that in an equilibrium information will not be revealed. Now we have the possibility that information will be revealed, and of course the victim is worse off. To ensure that the information is not released, the victim has to pay m* even though it may well be collected by the blackmailer who has no access to an external reward. The blackmail demand does not reveal any new information, since both blackmailer types make a blackmail demand; and it is only the blackmailer's action after the payment of (m=0) that may reveal new information to the victim.

4.4 A Discussion, Some Interpretations, and Potential Extensions

I A Finite Horizon Blackmail Game

The analysis of this chapter has been done within an infinite horizon blackmail game. How realistic is this, since almost every relationship between economic agents is finite? How sensitive are our results in respect to the assumption of an infinite time horizon? The finite horizon blackmail game provides a robustness test for the main results derived earlier.

Consider now a finite horizon blackmail game without any uncertainty and asymmetric information. Then it is obvious that in the last period, the blackmailer will take any money the victim hands over and reveal the information if that is his profit-maximising strategy. In particular, when the net reward is positive, the blackmailer will reveal the information in the very last period, since the blackmailer can not commit to an unprofitable action (information suppression) in the last period.

The crucial feature that separates the infinite horizon model from the finite horizon model is the following. When the net reward is positive in the infinite horizon model, the blackmailer can not commit not to come back to ask for more money in future; but he can commit not to reveal the information, since there is no last period. In contrast, in the finite time horizon model, the blackmailer is able to commit not to come back to ask for more money, but he is not able to commit not to reveal the information in the last period. When the net reward is negative, the unique equilibrium of no-blackmail derived earlier in the infinite horizon model coincides with the equilibrium of the finite time horizon model.

Consider next the finite horizon model and assume that T $< \infty$ is large. Assume that there is $\alpha < 1$, which is the probability that the blackmailer will reveal the information in the last period even if the victim has handed over the blackmail payment.Alternatively one may think that α is the probability that T is the last period. Does the equilibrium of the finite horizon model converge now with the one of the infinite horizon model: blackmail and information suppression, or something else? Indeed, we can show that we get qualitatively the same equilibrium: the information is suppressed and the victim pays the blackmailer. However, the optimal blackmail payment differs, since now the victim has to compensate the blackmailer for the possibility that in the last period he may be unable to reveal the information. To see this, note that the equilibrium condition that determines the optimal blackmail payment is as follows:

$$\begin{aligned} m+(R-c)=m+\delta m+\delta^2 m+\ldots+\delta^{T-1}m+ \\ (1-\alpha)m+(\alpha)(m+(R-c))\delta^T. \end{aligned} \tag{4.9}$$

From equation (4.9) we can solve the optimal blackmail payment as :

$$\begin{aligned} m*&=\frac{(R-c)(1-\delta)(1-\alpha\delta^T)}{\delta(1-\delta^{T-1})+\delta^T(1-\delta)(1-\delta)} \\ &=\frac{(R-c)(1-\delta)(1-\alpha\delta^T)}{\delta-\delta^{T+1}}. \end{aligned} \tag{4.10}$$

Now we can see the effect of α on m* given that T is large. If $\alpha=1$, we get exactly the expected result. The optimal blackmail payment coincides with the one of the infinite horizon model. This result is obvious, since now the blackmailer is able to reveal the information after any blackmail payment. Since T is large, the equilibrium converges with the one of the infinite horizon model. When α decreases, the equilibrium remains the same - i.e., information is not revealed - but the victim has to pay more (m* increases). This compensates the blackmailer for the possibility that in period T he may be unable to reveal the information after the blackmail payment. In an equilibrium, the victim pays more, and the blackmailer suppresses the information. We can conclude that the main result of information suppression and blackmail carries over also to the finite case. The only difference is that now with probability α the information may be revealed in the last period. However, the victim will be willing to pay the blackmailer if the blackmailer suppresses the information in all periods up to T-1.

Escalation of Corruption in Organisations

It is a well-known fact that corruptive activities in organisations often escalate from one level to another.[11] Here we want to point out how bribery in the first place may lead blackmail. In the following we assume that bribe-taking is against the rules in the civil service, and that the penalty for being caught is that a civil servant is fired.

Imagine a simple organisation, say a government agency with three members and an outside contractor. The members of the organisation are two civil servants (A and B) and their superior civil servant (P), whose preferences are the same as the government's. A's task is to choose an outside contractor who will supply material to the government agency and to approve the quality of delivered materials. Suppose that the contractor bribes A in one way or another in return for accepting low-quality material. That is, the government agency pays the contractor according to the standard price, the contractor's profit margins are higher due to the less costly low-quality materials it supplies, and A is bribed by the contractor. On the whole, the government is losing money, which is going into the contractor's and A's pockets. Note that bribery here is not just a redistribution of wealth, but it has serious economic consequences as well, since the government is worse off due to the low-quality materials. Assume now that B, the second civil servant, observes the bribery with probability one. And after observing it, B may start to blackmail A. Will he reveal information about the bribery to their superior or nor? We keep the assumption of a small cost, c due to information revelation.

We consider first the case where the organisation designer has not offered a reward for a member of the organisation who exposes any wrongdoing. We know from earlier that in this case B will not reveal the information, since the net reward is negative. A pays nothing to B, who, however, suppresses the information about the bribery. The bribery between the

---

[11] See, for instance, Basu et al (1992), who consider the controlling of corruption when corruption may escalate. See also Carrillo (1995b), who also considers corruption in a case where there exist potentially dishonest agents at several levels of a hierarchy.

contractor and A takes place, but here blackmail is not an issue. The contractor and A end up being better off. The economic consequences are severe, since the government loses money and receives low-quality materials.

Assume now that there is a reward R > c for a whistle-blower who exposes bribery. Now it is clear from our earlier analysis that blackmail may arise, since B is able to reveal information about bribery, which will also bring him a positive net reward. Hence, bribery in the first place facilitates blackmail, and nobody blows the whistle. In short, the corrupted civil servant buys silence from the initially honest civil servant, who becomes a blackmailer. The contractor, A and now also B end up being better off, and again the government is worse off. Compared to the earlier case of no reward, here the second civil servant, B also gets his share of bribes. At the aggregate level the economic consequences of bribery are the same as above.

Note that the possibility for blackmail would arise even if the first civil servant had been corrupted only once. What is required for the successful blackmail is that A has been corrupted at least once, and that there is a positive net reward for the civil servant who blows the whistle. In that case blackmail arises, redistributing the wealth between A and B. Note that if the organisation designer wants to prevent corruptive activities altogether, he has to destroy the roots of corruption in the first place.

Implications for the Organisation Design

From the theoretical literature we have learned a reasonable number of measures organisations may adopt in preventing collusion. Among them is rotation of workers. The idea is that rotation prevents collusion, since it blocks long-term relationships that are essential for collusion to be feasible.

Interestingly, rotation alone does not necessarily prevent blackmail; rather it helps the blackmailer, since now the blackmailer can credibly commit not to come back to ask for more in the very next period. Assume that a worker is able to blackmail only when being an employee in a position to expose a wrongdoer. In other words, a reward for whistle-blowing is available only in the period when corruption has taken place.

We introduce here rotation by assuming that B, the second civil servant, is rotated in every period. Therefore, A meets a different civil servant in each period. In effect, rotation changes the relationship between A and B to a one-shot game.

What happens when B is rotated and there is no reward for whistle-blowing? The only change is that A and B play a one-shot game. Since information revelation is costly, we know that B will not reveal his information, and thus A pays nothing to B. Therefore, A will be bribed by the contractor, and blackmail is not an issue. The economic consequences are the same as in the case of no rotation. That is, the government loses money and receives low-quality materials, and the contractor and A are better off.

Consider now the case where the organisation designer rewards a whistle-blower. In this case rotation has an effect, and corruptive activities will be prevented altogether. To see this, note that if A does not pay B, B will certainly blows the whistle. And if A pays B, the blackmailer will reveal the information in any case. In short, B can not commit not to blow the whistle after any bribe paid by A. Now B's threat to reveal information about the bribery is credible. Therefore, A knows that B will expose him in any case, and thus A won't get involved in bribery in the first place.

Here the possibility of blackmail is beneficial from the organisation's point of view, since it prevents all corruptive activities. Rotation and a positive net reward are a powerful combination in preventing corruption. The economic consequences are as follows. A does not get involved in bribery, the government receives high-quality materials, and blackmail is not an issue. The organisation designer does not have to pay rewards, since nobody blows the whistle. The simple economics behind this result is the very powerful last period effect due to rotation. B can commit not to come back, but he can not commit not to reveal even after A has paid him.

A Creditor and a Firm

Consider a firm that has borrowed D from a creditor to realise a research and development project. When raising the debt, the firm has to disclose and share valuable information about the project with the creditor. The debt contract defines the repayment schedule, where the firm agrees to make fixed payments p in each period. If the firm does not make the payment, the creditor is able to end the project by liquidating the assets. Liquidation of a R&D project means the same as selling or revealing information to an interested third party - for example, a competitor. That is, the creditor is able to destroy the firm's potential profit stream in the case of default.[12]

Now two interesting questions arise. First, how much will the creditor get back from the firm? Secondly, what is the smallest amount the firm has to pay the creditor so that the creditor does not end the relationship by liquidating the assets? Assume that in period t the firm, for one reason or another, makes a payment $m_t < p$. What does the creditor do? In principle, he has two alternatives. The creditor can either end the game by liquidating the firm with a small cost c. In this case, the creditor gets m_t+(L-c) immediately and nothing in future. Alternatively, the creditor can accept the smaller payment, and let the project go ahead. In this case the creditor gets: $m_t + \delta V^C$, where V^C is the creditor's continuation value. Clearly, what is optimal for the creditor depends on (L-c).

Suppose first that the liquidation value is low. Then, of course, the firm's position is now stronger and it has to pay only the project's termination value, and the creditor lets the project go ahead. Assume an extreme case, L=0, where no third party values the intangible assets of the R&D project. In this extreme case, the firm pays nothing, and the creditor can not do anything but let the project go ahead.

Here the last mover, the creditor, has an option to end the game by revealing information to the third party - i.e., by liquidating the R&D project. This option works against him, since he will always prefer to continue the game and wait one more period for the payments. For the creditor it is better to accept lower payments than to terminate the relationship by liquidation, since the continuation value of the project is always greater or equal to the termination value. The model above seems to belong to a class of models that have in common a so-called last-mover disadvantage.

Gromb (1994) considers repeated lending between a creditor and a borrower. In his analysis a creditor, the last mover whose decision is whether or not to refinance a project, gets zero surplus, which is a return for the termination of a project. This is due to the fact that a creditor cannot fully commit to terminate the relationship if the borrower does not meet his repayment, since it is mutually beneficial for them to write a new contract under which both are better off.

---

[12] We do not consider the possibility that the firm is excluded from the credit market in future, and we also rule out enforceable credit contracts.

In the present model the creditor gets a termination value as well, which is (L-c). That is the value of the R&D project to the interested third party. But, in addition to that, here the last mover (the creditor) receives a premium, which compensates him for not taking money from the firm and selling information about the R&D project to the third party in spite of the payment. To see this consider next a case where the liquidation value is higher. As earlier, the amount the firm has to pay the creditor depends on (L-c). Interestingly, it can be shown that here the firm may have to pay more than the agreed fixed payment p. Recall that no enforceable debt contracts are in place, and the firm has to hand over the payment that makes the creditor indifferent between liquidating the firm (revealing information) and letting the project go ahead (suppressing information).[13] That is, it has to be the case that: $m_t + \delta V^C = m_t + (L\text{-}c)$, and thus $m^* = (L\text{-}c)(1\text{-}\delta)/\delta$. Therefore, it may well happen that the flow payment $m^* > p$. The firm has to pay more than the agreed fixed payment p.

In this latter case, the creditor is able to blackmail the firm, since the creditor is able to sell information about the R&D project to the interested third party. The firm may end up paying more than it has borrowed from the creditor. Of course, the possibility of blackmail here is quite an extreme case, but presumably not totally unrealistic in the world of project financing. Perhaps this is one of the reasons why R&D projects are rarely financed by debt.

4.5. Concluding Remarks

This chapter has considered the phenomenon of blackmail in a simple dynamic framework. In particular, we have considered the question of how the potential surplus due to the victim's privacy is going to be shared between the victim and the blackmailer. We have shown that there is a unique Markov Perfect equilibrium which gives a precise prediction how much money the blackmailer will get by extorting the victim. Interestingly, and against intuition, it depends only on the external reward, and not the victim's valuation of his privacy. Furthermore, we have shown that the blackmailer gets more money by extorting the victim than by selling his piece of information directly to the interested third party. This follows

---

[13] Alternatively, if we assume that the creditor is able to liquidate only if $p < d$, then p is the upper limit of the payment the firm has to pay.

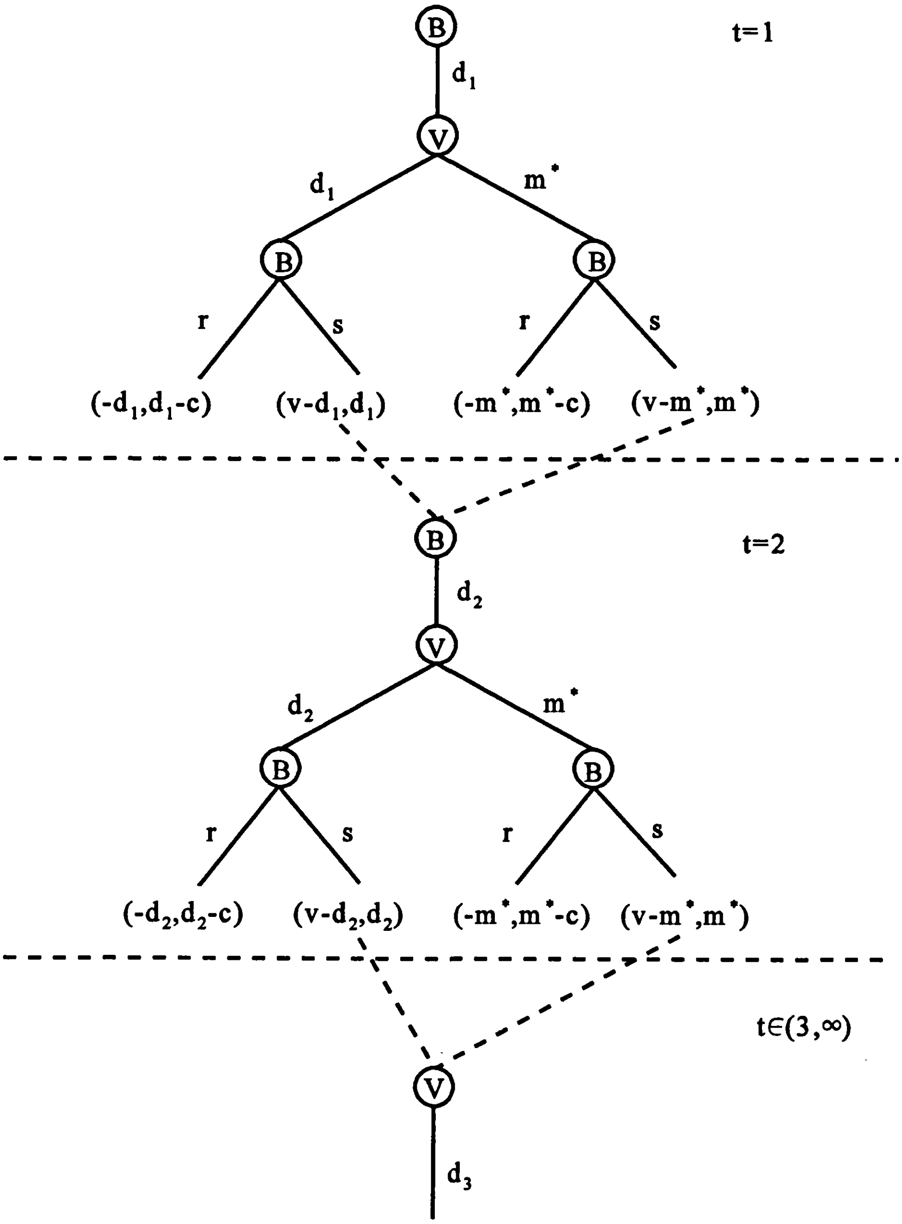
from the fact that in the former case the victim pays a blackmail premium to the blackmailer for not taking the victim's money and revealing the information anyway. It was also shown that even if the blackmailer is able to move first by announcing a blackmail demand, the blackmailer is unable to get more money from the victim than he does in the case when the victim moves first. This counter-intuitive result follows from the fact that in each period, irrespective of who moves first, the blackmailer always moves last - i.e., after the victim has handed over the blackmail payment. Despite the blackmailer's demand, the victim pays only as much as he would pay when moving first, and after this optimal payment the blackmailer prefers to suppress the information. In this sense, the model has one feature of a last mover disadvantage, and the blackmailer is unable to fully exploit the victim.

In future it might be worthwhile to examine thoroughly how an introduction of the coexistence of a rational and an insane blackmailer would change the results of the present paper. Here we have considered only a "rational blackmailer" who does not reveal information if it is unprofitable for him. In practice, of course, there may also exist "insane blackmailers" who will reveal information even if it is costly. However, it is a well-known fact in game theory literature that the introduction of an insane player may change the results greatly. Most obviously, here the victim is worse off, and both blackmailer types are better off. However, this potential extension does not add much to the analysis of reputation and imperfect information by Kreps and Wilson (1982). The main difference with Krep-Wilson is that here the blackmailer who reveals information in the very first period will end the game immediately, which is not the case in Kreps and Wilson (1982).

A more interesting case for further study would be to develop a model that fully integrates collusion and blackmail. Also, it would be interesting to look at whether the results presented here will carry on into the bargaining literature. For example, it would be interesting to try to incorporate the idea of blackmail with the present assumptions into the alternative offers' bargaining model. How much would the blackmailer get from the victim there? What would determine the shares the bargainers get? Is there a blackmail premium? These are among the open questions left for future studies.

Appendix C

Picture 4.2 The Blackmail Game. The Blackmailer Moves First



References

Acemoglu, D. 1995 Career Concerns and Cover-ups: A Dynamic Model of Collusion, mimeo. MIT, Department of Economics.

Aghion, P. and J. Tirole 1996: Formal and Real Authority in Organisations, *Journal of Political Economy 105.* 1 - 29.

Baliga, S. 1995: Universal Collusion and renegotiation, Dictators and Contracts, *University of Cambridge, DAE Working Papers No. 9511.*

Basu, K., Bhattacharya, S. And Mishra, A. 1992: Notes on Bribery and the control of corruption, *Journal of Public Economics 48.* 349 - 359.

Brundin, I. 1995: Collusion in A Debt Contract, mimeo. Cremaq, Toulouse.

Carrillo, H. 1995a: Grafts, Bribes and the Practice of Corruption, mimeo. IDEI, Toulouse.

Carrillo, H. 1995b: Corruption in Hierarchies, mimeo. IDEI, Toulouse.

Che, Y.- K. 1995: Revolving Doors and the Optimal Tolerance for Agency Collusion, *Rand Journal of Economics 26.* 378 - 397.

Crozier, M. 1964: *The Bureaucratic Phenomenon*, The University of Chicago Press, Chicago.

Dalton, M. 1959: *Men Who Manage.* John Wiley & Sons, Inc. New York.

The Economist 1995: The Nolan Committee. Spring Cleaning. February 25th. 30 - 31.

Elliot, R. - Willingham, J. 1980: *Management Fraud: Detection and Deterrance.* Petrocelli Books. New York.

Felli, L. 1996: Preventing Collusion Through Discretion, *London School of Economics, Theoretical Economics, Discussion Paper, TE/96/303.*

Felli, L and J. Villas-Boas 1996:"Friendships" in Vertical Relations,*London School of Economics, Theoretical Economics, Discussion Paper, TE/96/314..*

The Financial Times 1995: Nolan starts inquiry into public bodies: Education providers under standards spotlight. August 1st.

The Financial Times 1995: June 29th.

Gromb, D. (1994): Renegotiation of Debt contracts, mimeo. MIT, Sloan School of Management.

Hepworth, 1980: Modern Blackmail, *Scottish Journal of Sociology 4.* 291 - 299.

Holmström, B. and P. Milgrom 1990: Regulating Trade Among Agents, *Journal of Institutional and Theoretical Economics 146.* 85 - 105.

Holmström, B. and Milgrom, P. 1991: Multitask Principal-Agent Analysis: Incentive Contracts, Asset Ownership, and Job Design, *Journal of Law, Economics, and Organisation, Special Issue.* 24 - 52.

Itoh, H. 1992: Cooperation in Hierarchial Organisations: An Incentive Perspective, *Journal of Law, Economics, and Organisation 8. 321 - 345.*

Itoh, H. 1993: Coalitions, Incentives, and Risk Sharing, *Journal of Economic Theory 60.* 410 - 427.

Kofman, F. and J. Lawarrée 1993: Collusion in Hierarchial Agency, *Econometrica 61. 629 - 656.*

Kofman, F. and J. Lawarrée 1995: On the Optimality of Allowing Collusion, *Journal of Public Economics* (forthcoming).

Koffman, F, and J. Lawarree 1996: A Prisoner's Dilemma Model of Collusion Deterrence, *Journal of Public Economics* 59. 117 - 136.

Konrad, K. and S. Skaperdas 1995: Credible Threats in Extortion, *Free University of Berlin, Department of economics, Discussion Paper Series, No. 1/1995.*

Kreps, D and R. Wilson 1982: Reputation and Imperfect Information, *Journal of Economic Theory 27.* 253 - 279.

Krueger, A. 1974: The Political Economy of Rent Seeking Society, *American Economic Review 64.* 291 - 303.

Laffont, J.J. 1990: Analysis of Hidden Gaming in A Three Level Hierarchy, *Journal of Law, Economics and Organization 6.* 301 - 324.

Laffont, J.J. and M. Meleu 1996: Reciprocal Sepervision, Collusion and Organizational Design, mimeo. IDEI-CREMAQ, Universite de Toulouse.

Laffont, J.J. and D. Martimort, D. 1994: Separation of Regulators against Collusive Behaviour, *Institut d'Economie Industrielle. Toulouse, Document de Travail, No. 44.*

Laffont, J.J. and D. Martimort 1995a: Collusion Under Asymmetric Information,mimeo. IDEI, Toulouse and ECARE.

Laffont.J.J and D.Martimort 1995b: Collusion and Delegation, mimeo. IDEI, Toulouse and ECARE.

Laffont, J.-J. and J. Tirole, J. 1991: The Politics of Government Decision-Making: A Theory of Regulatory Capture, *Quarterly Journal of Economics 106. 1089 - 1127.*

Madson 1996: TV Drama serie presented at the BBC1.

Martimort, D. 1996a: A Theory of Self-Enforceable Collusion in Organizations, mimeo. IDEI, Toulouse.

Martimort, D. 1996b: The Life Cycle of regulatory Agencies:Dynamic Capture and Transaction Costs, mimeo. IDEI, Toulouse.

Maskin, E. And J.Tirole 1997: Markov Perfect Equilibrium, I:Observable Actions, mimeo. Harvard University and IDEI, Toulouse.

Milgrom, P. 1988: Employment Contracts, Influence Activities, and Efficient Organization Design, *Journal of Political Economy 96.* 42 - 60.

Milgrom, P. - Roberts, J. 1988: An Economic Approach to Influence Activities in Organisations, *American Journal of Sociology, Supplement*. 154 - 179.

Mogiljanski, A. 1994: Corruption in Procurement- The Economics of Regulatory Blackmail. *Research Papers In Economics 1994:4. University of Stockholm.*

Mogilianski, A. 1995: Indirect Monitoring and Optimal Corruption, mimeo. Stockholm University.

Near, J. and M. Miceli 1995: Effective Whistle-Blowing, *Academy of Management Review 20.* 679 - 708.

Newsweek 1994: Money Talks, November 14th, 10 - 15.

Okuno-Fujiwara, M., Postlewaite, A. and K. Suzumura 1990: Strategic Information Revelation, *Review of Economic Studies* 57. 25 - 47.

Olsen, T. and G. Torsvik 1994: Collusion and Renegotiation in Hierarchies: The Case of Beneficial Corruption, mimeo. University of Bergen.

The Oxford Thesarus. An A-Z Dictionary of Synonyms.1994.The Oxford University Press. Oxford.

Peltzman, S. 1976: Toward A More General Theory of Regulation, *Journal of Law and Economics 19.* 211 - 240.

Prendergast, C. and R. Topel 1996: Favoritism in Organizations, *NBER Working Paper No. 4427.*

Shleifer, A. and R.Visny 1993: Corruption, *Quarterly Journal of Economics 108.* 599 - 617.

Stigler,G. 1971: The Economic Theory of Regulation, *Bell Journal of Economics 2.* 3 -21.

Tirole, J. 1986: Hierarchies and Bureaucracies: On the Role of Collusion in Organisations, *Journal of Law, Economics and Organisation* 2. 181 - 214.

Tirole, J. 1988: The Multicontract Organisation, *Canadian Journal of Economics 21.* 459 - 466.

Tirole, J. 1992: Collusion and The Theory of Organisations. *In Advance in Economic Theory. eds. J.-J. Laffont, Proceedings of the 6th World Congress of the Econometric Society.* Cambridge University Press.

Varian, H. 1990: Monitoring Agents With Other Agents, *Journal of Instituional and Theoretical Economics 146*. 153 - 174.

The Wall Street Journal 1995: Dec. 7th.