Essays on Incentive Contracting, R&D, and Growth

Inauguraldissertation

zur Erlangung des akademischen Grades

eines Doktors der

Wirtschaftswissenschaften

der Universität Mannheim

Dezsö Szalay

vorgelegt im Sommersemester 2001

Referent: Professor Martin Hellwig, Ph.D.

Korreferent: Professor Dr. Benny Moldovanu

Dekan: Professor Dr. Paul Gans

Tag der mündlichen Prüfung: 18.7.2001

Acknowledgements

I owe special thanks to my thesis supervisor Martin Hellwig. Without the intellectual challenge his teaching at University of Basle provided I would not have continued my studying of economics. After my graduation he took lots of time for and devoted much effort to discussions and comments on my doctoral thesis. I am very greatful for his teaching and supervision.

I thank Benny Moldovanu for his many useful comments and for agreeing to be my second examiner. Many thanks also to Mathias Dewatripont for being my host when I visited ULB in Brussels within ENTER. His views on contract theory clearly reshaped my thinking about the subject. Chapter two benefitted from an excellent discussion with Jacques Crémer. This discussion ultimately laid the seed to chapter three.

I'm grateful for useful comments and discussions to the participants of various sessions at ULB, ESEM99 at Santiago de Compostela, ES in Seattle, EEA2k in Bolzano, and especially to the participants of the Mannheim workshop in economic theory.

Last but not least, I'd like to thank my friends and colleagues for excellent discussions and comments: Claudia Aebersold Szalay, Malte Cherdron, Christian Groh, Sascha Haller, Markus Jaggi, Christian Laux and Nicolas Melissas.

Financial Support of the Swiss Science National Foundation and the Deutsche Forschungsgemeinschaft is greatfully acknowledged.

Contents

1 Introduction **3**

2 Optimal Delegation **12**

2.1 Introduction 12

2.2 The Model 16

2.2.1 First Best 18

2.2.2 Freedom of Action as a Contracting Problem 19

2.2.3 The Degree of Freedom as an Incentive Device 22

2.3 Perfectly Aligned Interests 24

2.4 Conflicting Interests 29

2.5 Robustness Considerations 32

2.5.1 The Optimality of $\Gamma_{\epsilon,\lambda}$ 32

2.5.2 Monetary contracts 33

2.6 Conclusion 39

2.7 Appendix A 40

2.8 Appendix B 57

3 Procurement with an Endogenous Type Distribution **69**

3.1 Introduction 69

3.2 The Model 72

3.3 Analysis 74

3.3.1 The Principal's Problem 74

3.4 Implementation 78
3.5 Investments 80
3.6 Learning 86
3.7 Conclusion 92
3.8 Appendix A 93
3.9 Appendix B 95
3.9.1 Implementation 95
3.9.2 Investments 96
3.9.3 Learning 99

4 Financial Contracting, R&D and Growth **101**

4.1 Introduction 101
4.2 The model 103
4.3 Properties of the Nash-equilibrium 106
4.4 Moral Hazard: The Impact of Outside Financing Needs of the Challenger . 113
4.4.1 First best 114
4.4.2 Unobservable effort: Second best contracts 116
4.5 Leapfrogging reconsidered 122
4.6 Implications for growth 125
4.6.1 The model 125
4.6.2 Optimal research policies 127
4.6.3 The steady state 129
4.6.4 Welfare 131
4.7 Conclusions 132
4.8 Appendix A 133
4.9 Appendix B 133
4.10 Appendix C 134

Chapter 1

Introduction

This dissertation studies incentive issues in bilateral contracting relations. The literature has termed such relations principal-agent relations: one party - the principal - seeks to "employ" another party - the agent - to perform a particular task on his behalf. When the two parties do not have access to the same amount of information their relation is plagued by incentive problems. Two types of informational problems may arise: those resulting from hidden actions and those resulting from hidden information. In the hidden action case, known as "moral hazard" the principal cannot observe directly how hard the agent works. In contrast, hidden information refers to the case where the agent has superior information on the opportunities of the venture. Many real world situations contain an element of both types of informational problems: at the same time the principal cannot observe how hard the agent works *and* the agent possesses superior information on relevant characteristics of the venture.

The principal's problem amounts to foreseeing the consequences of his informational disadvantage and to design the contract governing the relation with the agent the way that best mitigates problems of asymmetric information. In the context of moral hazard problems (see, e.g., Grossman and Hart (1983)) this amounts to an indirect control problem for the principal. Although the input of the agent is not directly observable, there may still be measures available that provide some information about the agent's input, e.g. profits, so that contracts can be based on these measures. In general - that is, to the extent that the principal cannot punish the agent extremely hard - lack of observability entails welfare losses. In order to provide the agent with incentives to work hard, the agent has to bear more than the first best amount of risk. The intuition for this more general statement can be captured with a simple example: suppose a risk neutral principal and a risk averse agent contract with each other. Clearly, an efficient allocation of risks provides perfect insurance to the agent. However, for a wage scheme that is independent of profits, the agent has no incentive to work at all. Consequently, due to the moral hazard problem, a trade-off between insurance and incentive arises. Even between risk neutral parties, lack of adequate punishing schemes - that is inability to pay on the part of the agent - may entail welfare losses because it prevents the principal from holding the agent completely responsible for his actions.

In the case of hidden information the principal faces the problem of how to make use of the agent's information. To provide a particular example, Baron and Myerson (1982) have considered how to design optimal procurement contracts with a seller of unknown costs. Ideally the buyer would like to buy an amount that reflects costs of production. However, the seller always has an incentive to exaggerate these costs. Consequently, if unbounded punishments are infeasible, an optimal solution entails a trade-off between buying an efficient amount and clamping down on the supplier's overstating of costs and involves a welfare loss.

The question in each particular application is how to best structure contracts that govern the relation of the principal with the agent. The essays to follow in the next chapters discuss such contracting arrangements in three particular contexts.

Chapter two studies how to delegate a choice to an agent who has no superior information to begin with but may acquire such information at a personal cost. The principal cannot observe how hard the agent tries to acquire information nor can he tell what information the agent has in the end. In chapter three we reconsider the procurement problem of Baron and Myerson assuming that the seller's characteristic is endogenous and determined by an unobservable prior choice of the agent. Two canonical examples are discussed:

productivity enhancing investment and - similarly to chapter two - learning. Within the classification given above, the problems we address in these chapters belong to the third class of "hybrid" problems. Chapter four studies the consequences of a pure moral hazard problem. In "Financial Contracting, R&D, and Growth" an entrepreneur contracts with an investor for the financing of an R&D outlay. It requires finance as well as an unobservable effort input to be provided by the entrepreneur to make a research project a success. Anticipating his inability to monitor the entrepreneur's effort choice the investor guards himself by making finance more costly to obtain for the entrepreneur. Moral hazard on the part of the agent gives rise to an agency cost of outside finance (Jensen and Meckling (1976)). We study its consequences on industry structure and growth in the context of a patent race.

The remainder of this introduction provides an overview of the results of each of the chapters. More in-depth overviews of each chapter can be found at the beginning of each essay.

Optimal Delegation

Chapter two studies how to use freedom of action of an agent optimally as a means of motivating his "creativity on the job". The existing literature views the problem of delegation in terms of the following trade-off. An agent is given freedom of action to pursue some of his ideas because he has better information on the subject than the principal has. However, the problem is that the principal and the agent may have differing preferences concerning actions to be taken for some given information. Thus, the trade-off involved is one of improved quality of decision making versus loss of control.

We treat the case where the agent has no superior information ex ante but is capable of acquiring such information as part of the job. Starting from the existing literature it would be most natural to think of the problem in terms of a trade-off between incentives for information acquisition ex ante and conflicts of interest ex post. However, it turns out that by proceeding this way an interesting trade-off remains hidden to the eye. So, suppose there is no conflict of interest ex post: all ideas the agent comes up with are in the best interest of the principal. Does that mean, that the principal will simply have the agent have his way all the time? Somewhat surprisingly, this is not necessarily optimal. Instead, the principal may find it optimal to exclude some middling choices from the agent's action set. That is, he may find it optimal to force the agent to take a clear stance on a matter. To provide a specific example: suppose an economic advisor is asked by a politician for his advice on the desirability of more or less regulation of the telecommunications industry. It may be better to force the advisor to say either, "add more rules" or "deregulate" rather than allow him also to remain undecided and say "don't know, keep the status quo". In other words, it may be optimal to demand clear statements from experts even if they don't have (yet) clear answers on their mind. The reason is that experts' incentive to think hard about the matter and gather evidence in favour or against certain changes are enhanced. To be certain, the expert is forced to make two kinds of mistakes when he has to take a clear stance. The first kind of mistake occurs when the expert really has no information at hand. The second kind of error occurs when the expert has found out that the evidence in favour and against certain changes of policy really hold the balance. However, seen from the ex ante perspective, the first kind of error involves a higher utility loss to the expert. Therefore forcing him to take a clear stance increases his incentive to avoid having to give clear answers with a clouded mind. Consequently, the agent will come up with information more often if his freedom of action is cut this way. However, the drawback of cutting the agent's freedom of choice is that ex post the quality of decision making can only be decreased. In other words there is a trade-off between incentives for information acquisition and too much control.

It is straightforward to extend our framework to encompass conflicts of interest ex post. The resulting trade-off is one of increased creativity versus a loss of control. On the one hand the agent comes up with ideas more often when he has more freedom of action to pursue his ideas. On the other hand, these ideas will not always conform to the principal's taste.

Procurement with an Endogenous Type Distribution

The framework we've developed in chapter two is well suited to address problems of optimal regulation as well. In "Procurement with an Endogenous Type Distribution" we reconsider the problem of optimal procurement as studied by Baron and Myerson (1982) and allow for endogeneity of the seller's hidden characteristic. In Baron and Myerson's problem the principal faces a trade-off between making use of the agent's information in order to buy an efficient amount reflecting true costs of production and limiting the agent's informational rent. The optimal contractual arrangement entails less than optimal production for all but the most efficient type of producers. However, there is no distortion at the top.

This view of the problem of procurement essentially presumes that the provision of the service is in some sense "standard". The technology of production is given and well known to the producer. In practice, at least one of these presumptions might be unwarranted. If the producer provides the service for the very first time, it may have no superior information of production costs than the buyer has. But the seller may have means to acquire such information prior to production. Or the seller may have means to improve his technology through cost reducing innovations. However, the diligence he devotes to such tasks is in many cases hard to monitor. How are standard contracts best adjusted to provide incentives for investment and research?

We find that incentives for investments are provided through premia for the most efficient types of sellers. These premia may be so large as to generate overproduction and distortions at the top level of efficiency. The intuition behind these results is straightforward. An increase of investments prior to production increases the likelihood of low costs of production. In order to provide incentives for investments ex ante the optimal contract entails a reward for the case where the seller is able to produce at very low costs. Therefore production is set over and above the level that would be optimal ex post. The precise height of the reward depends on how informative low costs of production are about the amount invested ex ante. The more informative low costs the higher the premium for low costs.

In the limiting case where the informativeness is arbitrarily large, very simple contracts can provide very good incentives for investments. Optimal contracts can be approximated arbitrarily closely by standard procurement contracts adjusted for a premium to be paid whenever a minimum level of performance is reached.

Equilibrium investments of the seller will always be suboptimally low from the buyer's perspective. Higher investments increase the likelihood of low costs of production and the buyer benefits ex post when the seller lower costs. In other words, the investment decision entails a positive externality on the buyer.

Consider now research on costs of production. In this context we explain the emergence of extreme reward schemes. Optimal contracts involve a reward for very efficient producers and a punishment for very inefficient producers. The reason is straightforward. A contract that provides incentives for information acquisition must make more use of information than a contract that would be optimal for the same but given information structure. Making more use of information in turn means that the level of production reacts more strongly to the private information of the seller. If the buyer finds it optimal to give incentives for research the optimal contract may entail premia at the top and overproduction at the top as for the case of investments. Similarly, very simple contracts may emerge as an optimal solution to the incentive problem if the lowest costs of production is very informative about research efforts.

However, there is a decisive difference between investments that reduce costs and research that merely results in better information about costs. Here, more research on the part of the seller is not unambiguously a blessing to the buyer. The externality exerted on the seller's payoff may also be negative. More information about production costs results sometimes in a better deal for the buyer ex post and sometimes in a worse deal ex post. Roughly speaking, when the perceived costs of the seller adjust upwards the buyer is worse off than if it had contracted with the seller for an average amount of production at an average level of costs. He is better off if perceived costs adjust towards lower levels of costs. Thus, more research on the part of the seller is akin to a lottery. Whether the buyer benefits from such a lottery depends on his induced preferences towards risk. If he does not like to incur such a lottery he will respond with a more equalized contractual solution that makes less use of information than standard procurement contracts and consequently provides less of an incentive to acquire information in the first place.

Financial Contracting, R&D and Growth

"Financial Contracting, R&D and Growth" takes up a simple line of thought. Young firms are usually full of ideas but short of cash. In contrast, mature firms often have financial means in abundance. Yet, their behaviour does not always display too much creativity. Is there a mismatch between ideas and finance?

While we argue that this is not precisely the case, we show that there is nevertheless a grain of truth in the argument: more precisely, we establish a causal relationship between financing needs of start-ups and lack of creativity of mature firms. Mature firms can afford to innovate more slowly, thereby capturing more profits from the sale of their products, *because* they do not have to fear too much competition from start-ups. Lack of competition from start-ups in turn arises because problems of asymmetric information make finance more costly to obtain for young firms.

More precisely, we envision the following constellation. Imagine a monopolistic market for a patented product currently sold by the inventor of that product, the incumbent henceforth. The quality of the product can be improved through costly innovations. We consider the patent race for such a quality improving innovation between the incumbent and (a) potential entrant(s).

In such a race the initial starting position of the incumbent is a natural source of a comparative advantage vis a vis outsiders. The profits currently obtained can be used to finance R&D outlays. In contrast, outsiders with no such basis of profits have to contract with investors to finance their expenditures for research and development. Consequently, any gains from the venture entrants get have to be shared with the financiers. As a result, start-up's incentives to conduct research in the first place are reduced. Foreseeing these effects outside finance will be more costly to obtain for start-ups, putting them c.p. at a comparative disadvantage relative to the incumbent who can selffinance.

However, while financial standing power is an advantage of the incumbent, there are also drawbacks that stem from his starting position. In particular, innovating is a mixed blessing to the incumbent. For one thing, an innovation creates a new market for a new product with associated profits to the innovator. However, the sale of a higher quality product renders the current product obsolete, destroying the profits that could have been made through the sale of the old product. The incumbent replaces his monopoly with a new one so that only the marginal increase in profits is valuable to him. Entrants on the other hand have no such basis of profits to destroy. As a consequence total future profits count as an incentive for innovations, putting them c.p. at a strategic advantage vis a vis the incumbent.

The relative strength of these two forces determines who will spend more on research and who is more likely to win the race. In particular, the chances of outside entrants to win against established firms will be modest if it requires a large increase in hard-to-measure (and hence not contractable) inputs to increase the success probability in the lab by a given amount.

Thus, in a partial equilibrium framework, our model delivers a prediction about relative research outlays and relative research successes. When extended to a general equilibrium framework, differences in costs give rise to extreme outcomes: only the most efficient researcher will be active in equilibrium. When financing problems of young firms are severe, this means that only the incumbent will be an active researcher. Moreover, his research policy will involve less than socially optimal research. The financing needs of start-up firms create the leeway the incumbent needs to follow his most preferred research policy. Since new products render his old products obsolete the incumbent wants to rest a little bit on his laurels so that his privately optimal research policy entails too little innovations from the social perspective.

Our arguments make clear that venture capital is important, not only because it fuels the creation of new ideas and products through the creation of new firms replacing old ones but also because of its disciplining effect on the behaviour of established firms.

Chapter 2

Optimal Delegation

2.1 Introduction

Standard agency theory assumes that the agent's action space is given. In some contexts, however, it is of interest to study how much freedom of action an agent *should* be given. This is a key issue for the theory of delegation.

Delegation is usually studied in terms of a trade-off between the benefits of exploiting the agent's superior information and the costs arising from conflicting interests with the result that the principal may want to constrain the agent's freedom of action to mitigate the consequences of conflicts. The present paper shows that there may be another reason to reduce the agent's freedom of action. The paper develops a model where there is no conflict of interest but the principal forbids the agent to choose actions in some intermediate set forcing him to take a clear choice in one direction or the other. This is costly for both the principal and the agent whenever middling actions are indeed optimal. However, it has the advantage that the agent will try harder to find out what the best choice is.

The situation we have in mind is one where a principal hires an agent to implement a specific project for him. Both parties have well defined preferences over all possible choices the agent might take. However, the identity of the best choice depends on circumstances unknown to both the principal and the agent at the outset. By spending privately costly effort the agent can try to become informed about these circumstances. The agent's effort choice, however, is unobservable to the principal. The question then is how the contract affects the agent's effort choice and the problem is to provide the agent with proper incentives for information acquisition. Prohibiting the agent from doing what would be optimal under ignorance provides such an incentive by reducing his payoff in the event where he is not informed. To be sure, such a prohibition also reduces his payoff in the event where he is informed and still his information has no effect on his optimal choice. However, from an ex ante perspective the former effect dominates the latter. By prohibiting the agent form taking choices in the neighbourhood of the optimal choice under ignorance the difference between the expected payoff from being informed and the expected payoff without information is increased. Consequently the agent's marginal benefit of information, i.e. his incentive to become informed, is increased when the principal limits the agent's ability to choose middling actions. The resulting improvement in the quality of decision making may outweigh the losses the principal inflicts on himself through the fact that he punishes himself (ex post) just as he punishes his agent. Therefore, limiting the agent's freedom of action can be optimal in the very case where objectives are shared. In a straightforward extension of the model we allow for conflicting interests with respect to project choices. When conflicting interests are an issue the principal wants to mitigate the extent of conflicts in worst case scenarios. This is achieved by forbidding extreme choices a priori. Limiting the agent's discretion this way comes, however, at the cost of stifling the agent's incentives for information acquisition.

The expected extent of conflicts is the larger the more severe the conflict of interest between principal and agent. Therefore the region of prohibited extreme choices will be the *larger* the *more severe* the conflict. The region of prohibited intermediate choices will be the *larger* the *better interests are aligned*, because the marginal value of information to the principal is the higher the closer interests are aligned. Taken together this implies a non-monotonicity in the reaction of the agent's optimal degree of freedom to changes in the severity of the conflict of interests.

Our theory might help explain some real world phenomena: politicians, when they approach economists for their advice, want to hear clear statements about whether a policy is good or bad[1]. Thus, the answer, "it depends" is excluded from the agent's choice set, although it is often the best answer that can be given (and it is certainly the best answer for an economist, if he has no specific information). As a consequence the economist - provided that he has a preference for correct answers - will spend more time thinking about the correct solution.[2]

Whinston Churchill complained that he could not get one-handed economists. If he asked them for advice they would keep telling him: "Well, on the one hand..., but on the other hand also..." In contrast, according to our model, a one handed economist has to make up his mind for either one side and take a clear stance.[3]

Judges or members of a jury can effectively only choose between guilty and non guilty. The judgement "don't know" is excluded from the choice set. Consequently the judge and the jurors have a strong incentive to think hard about the guilt of the accused. As a last example, consider a corporate culture with a strong bias for innovations. The finding in the paper might give some rationale for the "we do it differently" doctrine. If headquarters follows the policy "whatever we do we will not stick to the status quo" division managers will have a strong incentive to become informed about what change exactly should be implemented.

The analysis of the paper bears on two distinct strands of the literature on delegation. On the one hand Holmström (1984) and Armstrong (1994) have investigated the optimal degree of freedom a principal should grant his agent in situations of conflicting interests with given information structures. In their models there was no need to motivate the agent to acquire information. On the other hand Aghion and Tirole (1997) have developed a theory of the allocation of authority within organizations. In a situation where the contracting parties' information structures are endogenous they show that a principal may gain from delegating authority to an agent even if principal and agent do not share the

[1]I thank Martin Hellwig for pointing out this interpretation to me.

[2]However, there are other factors which also come to mind in this particular situation. Politicians also like to have someone to blame if things go wrong.

[3]I thank Nobuhiro Kiyotaki for this example.

same objective and the latter sometimes abuses the freedom of action this provides. Giving the agent more freedom (or authority) increases his incentive to acquire information. The resulting improvement in the quality of decision making may outweigh the loss of control for the principal.

The present paper studies a hybrid of these two approaches to the problem of delegation. As in Aghion and Tirole (1997) the agent's incentives to become informed are determined by the allocation of authority, i.e. the agent's freedom of action. In contrast to their model only the agent can gather information in the present model, so it is optimal to delegate decision making to him. We are concerned with the optimal freedom of action, the agent should be given, or equivalently, how much authority he should have: as in Holmström (1984) and Armstrong (1994) the principal is given the right to "veto" certain projects or choices a priori.[4] Thus, the question we address is: "How should a principal optimally use his veto power in a situation of decentralized decision making?"

In terms of results the difference to Aghion and Tirole (1997) is that in their model the agent's *ability to choose* provides incentives for information acquisition while here it is the agent's *inability to choose* that is favorable for incentives.

At a more general level, the present paper raises the question: "How are contracts best structured when there is a need to provide incentives for both the acquisition of information and the subsequent use of the information?" This question has also been addressed in a different line of research.[5] Crémer, Khalil and Rochet (1998a) have investigated how the properties of optimal contracts are altered in the standard procurement problem when the agent's decision to gather productive information must be incentive compatible. Lewis

[4]This choice of "how to delegate" is not considered in Aghion and Tirole (1997). Tirole (1999) develops the "complete contracting" version of the Aghion and Tirole (1997) paper. He discusses the institution of giving authority to the agent with the principal having "gatekeeping counterpower". In this institutional setting the principal *is* given the right to exclude certain alternatives a priori. However, his point is to show that the concept of authority is not an artefact of modelling the problem as an incomplete contracting problem. To illustrate this idea, he shows that we can find an institution in an incomplete contracting context that implements the same outcome as the optimal complete contract. Giving authority to the agent with the principal having gatekeeping counterpower is an example of such an institution. This point is distinct from our question of how to use gatekeeping counterpower optimally.

[5]I was not aware of this link to the literature on procurement when I did this research. I'm greatful to Jacques Crémer, who pointed it out to me as a comment on an earlier draft of this paper.

and Sappington (1997) consider a related question in an organizational design problem to argue for the separation of information acquisition and its subsequent employment. A common finding of this literature is that the problem of inducing information acquisition and its revelation must be solved simultaneously. Optimal contracts therefore involve additional distortions relative to the standard procurement situation. In our model with perfectly aligned interests, efficient production could ex post be achieved costlessly. Yet, the principal may voluntarily choose to introduce distortions ex ante to provide incentives for information acquisition.

The remainder of the paper is structured as follows. Section 2.2 introduces the model. Section 2.3 treats the case of complete alignment of interests. Section 2.4 offers a more general treatment of the problem, allowing for conflicts of interests. Section 2.5 discusses the robustness of our approach and section 2.6 concludes. All proofs are relegated to the appendix.

2.2 The Model

We consider an agency problem in which the payoffs of both the principal and the agent depend on an action x chosen by the agent as well as on two parameters, η and φ, according to the specification

$$U(x, \eta) = k - \frac{A}{2}(x - \eta)^2 \tag{2.1}$$

and

$$\pi(x, \varphi) = K - \frac{1}{2}(x - \varphi)^2 \tag{2.2}$$

where U is the payoff of the agent and π the payoff of the principal. A is a measure of the agent's relative distaste for risk (read: relative to the principal). We will henceforth interpret the action x as the choice of a particular "project". The parameters η and φ are assumed to be realizations of random variables $\tilde{\eta}$ and $\tilde{\varphi}$ with joint distribution function

$F_{\eta,\varphi}$. More precisely

Assumption 2.1 (η, φ) have identical, symmetric marginals $f_\eta = f_\varphi$ with mean μ and Variance σ^2.

At the time of contracting both agent and principal know f but neither of them knows the realization of $\tilde{\eta}$ and $\tilde{\varphi}$.

Between the time the contract is written and the choice of the action x the agent tries to get informed. By exerting effort e the agent is informed about the realizations of both $\tilde{\eta}$ and $\tilde{\varphi}$ with probability e - uninformed with probability $1-e$ - and bears costs of effort $g(e)$, where $g(e)$ is convex and satisfies the Inada conditions: $g'(e) > 0$, $g''(e) > 0$; $\lim_{e\to 0} g'(e) = 0$; $\lim_{e\to 1} g'(e) = \infty$. The agent's choice of e is not observable to the principal. If the agent became informed, the true realizations of $\tilde{\eta}$ and $\tilde{\varphi}$ are his private knowledge. Moreover, the principal does not observe whether the agent successfully learned the true realizations of $\tilde{\eta}$ and $\tilde{\varphi}$ or not.

The agent chooses the action x according to his information[6].

Contractibility: In section 2.5, we will analyze the optimal choice of monetary contracts. But for the moment, we study a situation where the principal cannot use such contracting schemes:

Assumption 2.2 The agent does not react to monetary incentives.

In light of assumption 2.2 there is no possibility to make the reward of the agent dependent on the *realized* payoff of the principal, π, even if π is perfectly observable and verifiable. Therefore, the principal can only contract on the observable decision x in a nonmonetary sense:

Contracts: the principal restricts the choice of x to an admissable set Γ.

[6] As long as the principal can commit to a decision rule in advance it is inessential whether the agent directly takes actions or whether the agent tells the principal his information and the principal subsequently takes the action. If, however, commitment were not possible and no payments were allowed one would face a communication game in the spirit of Crawford and Sobel (1982). Hence our assumption is, that *the principal can commit to a decision rule.*

The situation is thus one where the agent cares "much more" for his private benefit than he cares for other sources of income. The assumption serves to emphasize the impact of the agent's discretion on his incentives to acquire information. We believe that there are interesting situations that fit into this simplified world: the introductory example of the judges' discretion is a case in point where simple rules rather than complex financial arrangements are used in practice. Moreover we will show that *the assumption is not crucial for the main results*: the main result will be shown to be robust with respect to the introduction of monetary contracts as long as payoffs are non-contractible. (see section 2.5)

The nature of conflicts: Under assumption 2.1 there is no conflict of interest *concerning project choice* ex ante: both principal and agent agree that $x = \mu$ is the best choice. However, ex post - i.e. conditional a specific realization (η, φ) of $(\tilde{\eta}, \tilde{\varphi})$ - the principal and his agent do not necessarily agree what the best project is: the agent's preferred alternative is $x = \eta$ while the choice $x = \varphi$ is in the principal's best interest. For the analysis of optimal contracts when interests are conflicting ex post in section 2.4 it will be necessary to impose more structure on the distribution $F_{\eta,\varphi}$. In this case we will analyze a special case of the distributions satisfying Assumption 2.1:[7]

Assumption 2.1': (η, φ) are bivariate normal with joint density $f_{\eta,\varphi}$ and identical marginals $f_\eta = f_\varphi$, i.e. $\eta \sim N(\mu, \sigma^2)$; $\varphi \sim N(\mu, \sigma^2)$.

Since marginals are identical, the correlation coefficient ρ between η and φ summarizes all relevant information on the expected extent of this *conflict of interest concerning project choice ex post.*

2.2.1 First Best

The first best corresponds to a situation where all variables can be written into an enforceable contract, i.e. everything is observable and verifiable. Such a contract will directly specify a particular choice of e and a choice of x contingent on the agent's information.

[7]Note that the analysis of section2. 3 will not make use of Assumption 1.1'.

To provide a benchmark we will characterize this first best for the case of perfectly aligned interests concerning project choice, i.e $\eta \equiv \varphi$. The contract will thus specify $x = \eta$ for the case where the agent knows the realization of $\tilde{\eta}$ and $x = \mu$ for the case where the agent does not know the true realization of $\tilde{\eta}$.

The first best effort level is defined as the effort level that maximizes joint surplus.[8] To derive this we must first calculate the value of information. From (2.2) the principal's payoff conditional on the agent knowing η is K, while the agent will enjoy utility k (from (2.1)). Conditional on the agent being ignorant the principal's payoff is: $K - \frac{1}{2}\int(\mu - \eta)^2 dF_\eta \; = K - \frac{1}{2}\sigma^2$, while the agent will then have utility $k - \frac{A}{2}\sigma^2$. The optimal choice of e solves

$$\max_e K + k - \frac{(1+A)\sigma^2}{2} + e\frac{(1+A)\sigma^2}{2} - g(e)$$

and therefore satisfies

$$\frac{(1+A)\sigma^2}{2} \overset{!}{=} g'(e)$$

I.e. the agent spends effort to learn until the marginal cost of effort is equal to the marginally avoided disutility of risk for the principal and himself.

2.2.2 Freedom of Action as a Contracting Problem

There are two problems of asymmetric information present. The agent's information and his effort choice are not observable to the principal. *Ex post*, i.e. given the agent's information structure, there will be hidden knowledge, both with respect to whether the agent has learned the true realizations of $\tilde{\eta}, \tilde{\varphi}$ and with respect to the realizations themselves if the agent has learned them. *Ex ante*, i.e. before information is acquired, the incentive problem is to motivate the agent to become informed. Both problems must be solved

[8] The present model is one of private benefits, which makes the interpretation of first best a bit difficult. It should be noted that our results do not depend on this definition of first best. The only important thing is that there is some underinvestment at all.

simultaneously by choice of the contract Γ. In the absence of monetary compensation schemes, the principal will choose the admissable set Γ in such a way as to maximize his expected payoff. More formally the principal's maximization problem is:

$$\max_{x'',x'(.),\Gamma,e} eE\pi(x'(\tilde{\eta}),\tilde{\varphi})+(1-e)E\pi(x'',\tilde{\varphi}) \tag{2.3}$$

s.t.

$$\begin{aligned} x'(\eta) &\in \arg\max_{x\in\Gamma} U(x,\eta) \quad \forall\eta \\ x'' &\in \arg\max_{x\in\Gamma} EU(x,\tilde{\eta}) \end{aligned} \tag{2.4}$$

$$EU(x'(\eta),\tilde{\eta})-EU(x'',\tilde{\eta})=g'(e) \tag{2.5}$$

$$\begin{aligned} &eEU(x'(\eta),\tilde{\eta})+ \\ &(1-e)EU(x'',\tilde{\eta})-g(e)\geq 0; \end{aligned} \tag{2.6}$$

The first term in (2.3) represents the principal's expected payoff when he knows that an agent who knows the realizations of $\tilde{\eta}$ and $\tilde{\varphi}$ (henceforth an *informed* agent) chooses his most preferred alternative (given the restriction Γ), weighted by the probability of reaching this state, e. The second term represents the analogue for an *ignorant* agent's choice. (2.4) are the incentive compatibility conditions on the choice of alternative of the informed agent (henceforth $x'(.)$) and on the choice of an ignorant agent (x'' henceforth), respectively. (2.5) is the IC condition for the agent's effort choice. Finally (2.6) is the agent's individual rationality constraint, *which is assumed to be nonbinding for all* Γ, i.e. k is relatively large.[9]

[9] Without monetary transfers this assumption is obviously needed for a meaningful analysis. We will discuss the impact of this assumption on our results below in detail. See section 2.5.2. and Appendix B.

The problem is not completely standard. The difference stems from the fact that the principal controls the degree of freedom the agent has when he takes his decision: the principal maximizes with respect to sets which in turn provide the constraints on the agent's choices. At first sight this problem is quite unstructured. However one can prove:

Proposition 2.1 *An optimal contract takes the form*

$$\Gamma_{\epsilon,\lambda} = \Gamma_\epsilon \cap \Gamma_\lambda$$

where[10]

$$\Gamma_\epsilon = dom\ f_\eta \backslash (\mu - \epsilon, \mu + \epsilon)$$

and

$$\Gamma_\lambda = dom\ f_\eta \backslash [-\infty, \mu - \lambda)(\mu + \lambda, \infty]$$

for some $\epsilon, \lambda \geq 0; \lambda \geq \epsilon$ *whenever (i) interests are perfectly aligned or (ii) interests are conflicting and* σ^2 *is large.*

The optimal contract can be understood in terms of two parameters, ϵ and λ. The agent is prohibited from taking extreme actions x in the intervals $[-\infty, \mu - \lambda)$ and $(\mu + \lambda, \infty]$ and he is prohibited from taking middling actions in the interval $(\mu - \epsilon, \mu + \epsilon)$. The smaller is λ the more severe is the prohibition of extreme actions; the larger is ϵ the more severe is the prohibition of middling actions. In the case of a worker in charge of buying goods for his firm, Γ_ϵ means "buy either less than $\mu - \epsilon$ or more than $\mu + \epsilon$ goods but don't buy any quantity in between." Γ_λ means "buy at least $\mu - \lambda$ but not more than $\mu + \lambda$ items". Setting $\epsilon = 0$ and $\lambda = \infty$ means "do whatever you deem right".

The intuition for Proposition 2.1 is explained in more detail in section 2.5. The explanation builds on heuristic arguments developed in the analysis of the optimal choice of ϵ, λ. The

[10] $dom\ f_\eta$ stands for the domain of the density.

formal proof of Proposition 2.1 is given in appendix A.[11]

Notation: In light of Proposition 2.1 we will henceforth use the following shorthand for the agent's and the principal's expected payoffs under contract $\Gamma_{\epsilon,\lambda}$ (anticipating incentive compatible choices according to the agent's information and (2.4)):

$$\begin{aligned} E\pi'_{\epsilon,\lambda} \quad &: \quad = E\pi(x'(\tilde{\eta}),\tilde{\varphi}); \qquad E\pi''_{\epsilon,\lambda} := E\pi(x'',\tilde{\varphi}) \\ EU'_{\epsilon,\lambda} \quad &: \quad = EU(x'(\eta),\tilde{\eta}); \qquad EU''_{\epsilon,\lambda} := EU(x'',\tilde{\eta}) \end{aligned}$$

2.2.3 The Degree of Freedom as an Incentive Device

The agent's incentive to acquire information is determined by the marginal value of information for him, $EU'_{\epsilon,\lambda} - EU''_{\epsilon,\lambda}$(see (2.5)). This in turn depends on the contract, because the agent's actions x' and x'' are constrained by Γ (see (2.5)). To determine the agent's effort choice, we solve backwards. For any contract $\Gamma_{\epsilon,\lambda}$ the agent's best response with respect of the choice of x is:

$$\begin{aligned} x'(\Gamma_{\epsilon,\lambda}) \quad &= \quad \begin{cases} \mu-\lambda & for\ \eta\in \quad [-\infty,\mu-\lambda) \\ \eta & \text{"} \quad [\mu-\lambda,\mu-\epsilon] \\ \mu-\epsilon & \text{"} \quad (\mu-\epsilon,\mu] \\ \mu+\epsilon & \text{"} \quad (\mu,\mu+\epsilon) \\ \eta & \text{"} \quad [\mu+\epsilon,\mu+\lambda] \\ \mu+\lambda & \text{"} \quad (\mu+\lambda,\infty] \end{cases} \\ x'' \quad &= \quad \{\mu+\epsilon \end{aligned} \tag{2.4'}$$

i.e. he minimizes the (expected) deviation from the unrestricted best choice, which would be $x=\eta$ if the agent is informed and $x=\mu$ if the agent is ignorant.[12]

[11]Since both the intuition and the formal proof build on arguments developed below, the reader interested in the formal proof is invited to go through the analysis of sections 2.3 and 2.4 before going through the proof of Proposition 2.1.

[12]We can pick $x''=\mu+\epsilon$ without loss of generality since both principal and agent are indifferent between

The agent's expected utility levels under $\Gamma_{\epsilon,\lambda}$, conditional on being informed ($EU'_{\epsilon,\lambda}$) and, respectively, on being ignorant ($EU''_{\epsilon,\lambda}$) are then (by symmetry)

$$\begin{aligned} EU'_{\epsilon,\lambda} &= k - A\int_{\mu}^{\mu+\epsilon}(\mu+\epsilon-\eta)^2 dF_\eta - A\int_{\mu+\lambda}^{\infty}(\mu+\lambda-\eta)^2 dF_\eta \\ EU''_{\epsilon,\lambda} &= k - \frac{A}{2}\int(\mu+\epsilon-\eta)^2 dF_\eta = k - \frac{A}{2}(\sigma^2+\epsilon^2) \end{aligned} \tag{2.7}$$

From (2.5) and (2.7) the agent's incentive compatible effort choice is

$$e = h\left[\ A\frac{(\sigma^2+\epsilon^2)-2\int_{\mu}^{\mu+\epsilon}(\mu+\epsilon-\eta)^2 dF_\eta - 2\int_{\mu+\lambda}^{\infty}(\mu+\lambda-\eta)^2 dF_\eta}{2}\ \right] \tag{2.5'}$$

where $h := g'^{-1}(.)$ (which exists by $g''(e) > 0$). It turns out that the parameters ϵ, λ have very clear-cut effects on the agent's effort choice:

Lemma 2.1 *e is increasing in λ and increasing in ϵ.*

The agent's incentive to acquire information is determined by the wedge between utility levels across the state where the agent is informed and the state where he is not. To foster the agent's effort choice, this wedge must be increased: either *knowledge* must be *rewarded* or *ignorance* must be *punished*. The first result corresponds to the intuitive notion that "limiting the agent's freedom of action stifles his incentives for information acquisition". From (2.5') it is clear that the maximal possible reward is achieved by setting $\lambda = \infty$. For any given choice of λ, there is, however, a second way to foster the agent's initiative. Prohibiting the agent from choosing actions which would be optimal in case he remains ignorant effectively serves as a punishment device. The principal thereby interferes with the optimal choice under ignorance as well as with the optimal choices whenever the agent knows that middling actions are optimal. The punishment is effective because its impact on the agent's utility is more severe when the subordinate is ignorant: conditional on the agent being informed, the optimal choice will only by coincidence, i.e. small probability, lie in the excluded interval, while conditional on being ignorant the agent *always* wants

$\mu - \epsilon$ and $\mu + \epsilon$.

to choose the excluded alternative $x = \mu$. In short, both levels $EU'_{\epsilon,\lambda}$ and $EU''_{\epsilon,\lambda}$ *decrease*, but their difference, $EU'_{\epsilon,\lambda} - EU''_{\epsilon,\lambda}$, *increases.* The *inability* to choose can be used to foster the agent's initiative. It should be noted that this result does not depend on any distributional assumptions at all. It will hold as long as there are at least three possible choices and there is some uncertainty at all.

2.3 Perfectly Aligned Interests

At first sight the case of perfectly aligned interests with respect to project choices seems totally uninteresting. By definition the agent's selfishly motivated project choices are in the principal's best interest. The principal could simply tell the agent to do whatever he deems right and thereby make costless use of the agent's superior information. However, only the ex post incentive problem is solved. From the ex ante perspective even when granted full freedom of action the agent still exerts too little effort relative to the social optimum because he does not value the improvement in the principal's utility from an informed decision. The principal may therefore want to restrict the agent's choice set, so as to increase his effort choice. In light of lemma 2.1 this can be achieved by setting $\epsilon > 0$. We now proceed to show that this might indeed be optimal.

Analogously to (2.7) one can derive the principal's expected utility conditional on the fact that his agent has learned the realizations of $\tilde{\eta}$ and $\tilde{\varphi}$ (that the agent is ignorant of the realizations, respectively)

$$\begin{aligned} E\pi'_{\epsilon,\lambda} &= K - \int_{\mu}^{\mu+\epsilon} (\mu + \epsilon - \eta)^2 dF_\eta - \int_{\mu+\lambda}^{\infty} (\mu + \lambda - \eta)^2 dF_\eta \qquad (2.8) \\ E\pi''_{\epsilon,\lambda} &= K - \frac{1}{2}(\sigma^2 + \epsilon^2) \end{aligned}$$

Plugging (2.5') and (2.8) into (2.3), the principal's maximization problem can be restated as an unconstrained problem:

$$\max_{\epsilon,\lambda} eE\pi'_{\epsilon,\lambda} + (1-e)E\pi''_{\epsilon,\lambda} \qquad (2.9)$$

Solutions satisfy the first order necessary conditions:

$$\frac{\partial E\pi''_{\epsilon,\lambda}}{\partial\epsilon} + e\frac{\partial}{\partial\epsilon}\left[E(\pi'_{\epsilon,\lambda} - \pi''_{\epsilon,\lambda})\right] + \frac{\partial e}{\partial\epsilon}E(\pi'_{\epsilon,\lambda} - \pi''_{\epsilon,\lambda}) \stackrel{!}{=} 0 \tag{2.10}$$

and

$$e\frac{\partial}{\partial\lambda}\left[E(\pi'_{\epsilon,\lambda} - \pi''_{\epsilon,\lambda})\right] + \frac{\partial e}{\partial\lambda}E(\pi'_{\epsilon,\lambda} - \pi''_{\epsilon,\lambda}) \stackrel{!}{=} 0. \tag{2.11}$$

Define $e_f^{sb} := e_{|\epsilon=0,\lambda=\infty}$, i.e. the agent's effort choice in the second best situation when he is left to choose what he deems right (where the f stands for "free"). Then

Proposition 2.2 (*I*) *An optimal solution to program (2.9) exists.*
(*II*) *The optimal contract takes the form* $\Gamma_{\epsilon^*} = dom\ f_\eta \backslash (\mu - \epsilon^*, \mu + \epsilon^*)$, *i.e.* $\Gamma_{\epsilon^*,\lambda^*}$ *with* $\lambda^* = \infty$ *and* ϵ^* *finite.*
(*III*) *The principal optimally sets* $\epsilon^* > 0$ *if* $(1 - e_f^{sb}) < \frac{A\sigma^2}{2g''(e_f^{sb})}$.
(*IV*) *If* $\frac{g'''(e)}{g''(e)^2}\frac{A\sigma^2}{2} > 2; \forall e$ *then* (*a*) *the optimal contract is unique and* (*b*) $\epsilon^* > 0$ *only if* $(1 - e_f^{sb}) < \frac{A\sigma^2}{2g''(e_f^{sb}).}$.

Obviously the principal has no reason to prohibit the agent from choosing extreme alternatives when interests are perfectly aligned: both (2.8) and (2.5') are increasing in λ. Therefore $\lambda = \infty$ is optimal. To ease the discussion in this section we will skip the subscript λ on expected payoffs. Figure 2.1 shows the shape of the optimal contract if $\epsilon > 0$ is optimal. Since the problem is symmetric we consider the upper half of the support of η only and the origin of the diagram corresponds to $\mu = E\eta$.

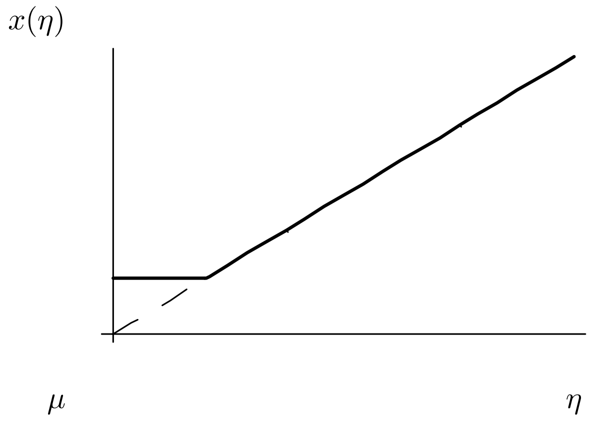

Figure 2.1

The dashed line corresponds to the 45° degree line and depicts first best optimal project choices $x = \eta$. If the optimal contract is such that the agent's effort choice is fostered above the level that he supplies voluntarily, then a departure from optimal choices is forced for realizations of $\tilde{\eta}$ close to the mean in that the agent is prohibited completely from choosing actions which are too close to the mean. Note in particular that these are the only distortions introduced at the optimum.

Whether the principal finds it worthwhile to set $\epsilon > 0$ depends on the *level* of effort supply when the agent is left to choose what he deems right, e_f^{sb}, on the agent's *responsiveness* to incentives, $\frac{A}{g''(e_f^{sb})}$, and on the marginal value of avoided risk to the principal $\frac{\sigma^2}{2}$. The increase in the agent's effort supply must be large enough to justify the ex post utility losses the principal inflicts on himself by interfering with the agent's choices. A sufficient condition that guarantees this is $(1 - e_f^{sb}) < \frac{A\sigma^2}{2g''(e_f^{sb})}$. To understand this result , consider ϵ very small but positive: with probability (approximately) $1 - e_f^{sb}$ the agent is ignorant and the principal just shoots himself in the foot, by setting $\epsilon > 0$. The marginal impact of increasing ϵ on $E\pi''_\epsilon$ is then equal to $-\epsilon$. With probability e_f^{sb} the agent is informed. In this event there is also a negative impact of ϵ on $E\pi'_\epsilon$, but this effect vanishes as ϵ is chosen small enough. These costs must be compared with the *marginal risk avoidance*, i.e. additional effort supplied due to a marginal increase in the wedge $E(U'_\epsilon - U''_\epsilon)$ times the impact of avoided risk on the principal's utility: $\frac{A}{g''(e_f^{sb})}\frac{\sigma^2}{2}$.[13] Under the condition in

[13] It is shown in the appendix that the impact of increasing ϵ on $E(U'_\epsilon - U''_\epsilon)$ becomes indistinguishable

the proposition the latter, beneficial effect will dominate the costs and the principal will set $\epsilon > 0$.

Precisely for the same reason, however, the principal's maximization problem cannot be globally concave. But if the product of the marginally avoided risk, $\frac{A}{g''(e_f^{sb})}\frac{\sigma^2}{2}$, and the curvature of the marginal cost function, $\frac{g'''(e)}{g''(e)}$, is bounded below by a suitable constant, we are able to prove uniqueness of the solution of the principal's problem.[14] Note, however, that existence of a solution does not depend on this condition.

It is straightforward to show that the argument underlying proposition 2.2 does not depend on the form of the utility function: one could as well take functions $\hat{\pi}(x,\eta)$, $\hat{U}(x,\eta)$ and assume that, for each η, $\frac{\partial}{\partial x}\hat{\pi}(x,\eta) = 0$ for some x and $\frac{\partial^2}{\partial x^2}\hat{\pi}(x,\eta) < 0$. Likewise for $\hat{U}(x,\eta)$. Furthermore assume that $\hat{\pi}(x,\eta)$ and $\hat{U}(x,\eta)$ are such that their unique maximizer is identical for each η, hence that there are no conflicts of interests. The proofs are essentially the same.

One may wonder about the robustness of the analysis to individual rationality considerations. In the absence of monetary transfers the costs of prohibiting middling actions make the agent worse off. However, when monetary contracts are admitted, it is straightforward to extend the analysis to binding IR-constraints. This will be discussed in detail in section 2.5 and in Appendix B.

There is an issue of commitment and renegotiation. Once effort is sunk, both principal and agent would be better off when the agent is given full freedom. But if the principal cannot commit not to renegotiate the contract, the incentive effect of the initial contract would be undermined.

A natural question is how the size of the prohibited interval is affected by the parameters of the model.

Proposition 2.3 *If, given σ, $\frac{g'''(e)}{(g''(e))^2}\frac{A\sigma^2}{2} > 2, \forall e, A$ the optimal value of ϵ^* is nonincreasing in the risk aversion parameter A. Likewise, given A, ϵ^* is nonincreasing in σ if*

from ϵ for ϵ very small. (see (2.17) in the Appendix.)

[14] Unfortunately I have no (good) intuition for the third derivative of the cost function.

$\frac{g'''(e)}{(g''(e))^2}\frac{A\sigma^2}{2} > 2, \forall e, \sigma$. If in addition $(1 - e_f^{sb}) < \frac{A\sigma^2}{2g''(e_f^{sb})}$, ϵ^* is, c.p., strictly decreasing in A and σ.

Whether the prohibited region will be smaller or larger when the parameters of the model are varied again boils down to the question how the level of effort and the responsiveness of the agent to incentives are affected by the change in the parameters. In general there are countervailing forces at work. A smaller A causes a lower level of effort supply of the agent for any given contract, because the agent gets relatively indifferent with respect to the risk of choosing the wrong projects. Therefore the risk of shooting into oneself's foot by setting $\epsilon > 0$ increases and the principal has less of an incentive to use a prohibition. On the other hand, a decrease in A also changes the way the agent reacts to incentives: both the nominator and the denominator of $\frac{A\sigma^2}{2g''(e_f^{sb})}$ are affected by a change in A. The nominator is decreased by a decrease in A. The denominator is affected by the fact that e is the smaller the smaller A is. If the cost function is convex enough, i.e. if $g'''(e) > 0, \forall e$, then this implies that the denominator is decreased by a decrease in A. Moreover, if this convexity of $g(e)$ is so pronounced in relation to the other parameters of the model such that the condition $\frac{g'''(e)}{(g''(e))^2}\frac{A\sigma^2}{2} > 2, \forall e, A$ holds then $e + \frac{A\sigma^2}{2g''(e)}$ will be increased for all values of e where it is evaluated. Hence we can conclude that the principal has more of an incentive to use a prohibition of intermediate choices, or use a larger prohibition of such choices, respectively, when the agent is relatively indifferent with respect to the risk of choosing wrong actions.

The reader might be worried about the fact that we made reference to the endogenous quantity e_f^{sb} in the conditions stated in proposition 2.2 as being sufficient for $\epsilon^* > 0$. To see that these conditions are not vacuous it is instructive to consider the following **example**: Let[15] $g(e) = ce + \frac{1}{3}e^3$ and take $\sigma = 1$. Hence $e_f^{sb} = \sqrt{\frac{A}{2} - c}$ and $e^{fb} = \sqrt{\frac{(A+1)}{2} - c}$. One

[15]Obviously, the Inada conditions are violated. Interior solutions must be guaranteed by a suitable choice of c, or, for given c by a suitable restriction on the range of A. For this example $\frac{g'''(e)}{(g''(e))^2}\frac{A\sigma^2}{2} = \frac{1}{2}\left[\frac{\frac{A\sigma^2}{2}}{\frac{A\sigma^2}{2} - c}\right]$ is a constant and larger than 2 if $c > \frac{3}{8}A\sigma^2$. If $c < \frac{3}{8}A\sigma^2$, uniqueness of the equilibrium is not guaranteed and the local comparative statics predictions are reversed. Furthermore, we know from proposition 2 that $e_f^{sb} + \frac{A\sigma^2}{2g''(e_f^{sb})} > 1$ is sufficient in this case for $\epsilon^* > 0$, not necessary.

has $e_f^{sb} + \frac{A}{2g''(e_f^{sb})} > 1 \; \forall A$ if $c > \frac{1}{8}$. If $c = \frac{1}{16}$, $e_f^{sb} + \frac{A}{2g''(e_f^{sb})} > 1$ for $\frac{1}{8} \leq A < 0.135$ or $A > 0.49$. If $c = 0$, $e_f^{sb} + \frac{A}{2g''(e_f^{sb})} > 1$ for all $A > 0.5$.[16]

Henceforth I will assume that the conditions in proposition 2.2 are satisfied. Formally I will impose:

Assumption 2.3: Let $(1 - e_f^{sb}) < \frac{A\sigma^2}{2g''(e_f^{sb})}$ and $\frac{g'''(e)}{g''(e)^2} \frac{A\sigma^2}{2} > 2; \forall e, A$.

2.4 Conflicting Interests

So far only the extreme case of perfect alignment of interests has been discussed. In this special case there is no need for the principal to protect himself from ex-post-opportunistic behavior of his agent, since the only deviation from first best was the agent's effort choice. When $\tilde{\eta}$ and $\tilde{\varphi}$ are not identical this is no longer true and the contract must provide the agent with proper incentives for his effort choice *and* the project choice. Concerning the project choice, the worst case scenario for the principal is given by a realization-tuple η, φ of $\tilde{\eta}, \tilde{\varphi}$ where η and φ are extremely far apart. As a consequence, the agent when left unrestricted to choose whatever he deems right, might choose an alternative which gives the principal a very low private benefit. The principal can majorize the expected impact on his private benefits of this divergence of interests ex post if he forbids extreme alternatives a priori, i.e. sets $\lambda < \infty$.

The expected utility of the principal from informed decisions of his agent now takes account of the joint distribution of $\tilde{\eta}, \tilde{\varphi}$. Hence:

$$E\pi(\rho)'_{\lambda,\epsilon} := K - \left\{ \begin{array}{c} \int \int_{\mu}^{\mu+\epsilon} (\mu + \epsilon - \varphi)^2 dF_\eta dF_{\varphi|\eta} \\ + \int \int_{\mu+\epsilon}^{\mu+\lambda} (\eta - \varphi)^2 dF_\eta dF_{\varphi|\eta} \\ + \int \int_{\mu+\lambda}^{\infty} (\mu + \lambda - \varphi)^2 dF_\eta dF_{\varphi|\eta} \end{array} \right\} \tag{2.12}$$

[16]The reader who finds it counterintuitive that there is also a range for A such that $e_f^{sb} + \frac{A}{2g''(e_f^{sb})} > 1$ if A is *large* enough, recall that for $c < \frac{3}{8}A\sigma^2$, $e_f^{sb} + \frac{A}{2g''(e_f^{sb})}$ is *increasing* in A.

while $E\pi''_{\lambda,\epsilon}$ is not affected relative to the case before, because there is no information we could condition on. The benefit of forbidding extreme alternatives a priori for the principal is the protection against extreme conflicts. The flipside of the coin is of course that the principal might end up forbidding a choice that he would actually want to have implemented ex post and that the agent will respond with less effort.

Formally, the principal now solves

$$\max_{\epsilon,\lambda} eE\pi(\rho)'_{\epsilon,\lambda} + (1-e)E\pi''_{\epsilon,\lambda} \tag{2.13}$$

where $E\pi(\rho)'_{\epsilon,\lambda}$ is given by (2.12). Solutions of this program satisfy

$$\frac{\partial}{\partial\epsilon}E\pi''_{\epsilon,\lambda} + e\frac{\partial}{\partial\epsilon}\left[E(\pi(\rho)'_{\epsilon,\lambda} - \pi''_{\epsilon,\lambda})\right] + \frac{\partial e}{\partial\epsilon}E(\pi(\rho)'_{\epsilon,\lambda} - \pi''_{\epsilon,\lambda}) \stackrel{!}{=} 0 \tag{FOC$_\epsilon$}$$

$$e\frac{\partial}{\partial\lambda}E\pi(\rho)'_{\epsilon,\lambda} + \frac{\partial e}{\partial\lambda}E(\pi(\rho)'_{\epsilon,\lambda} - \pi''_{\epsilon,\lambda}) \leq 0 \tag{FOC$_\lambda$}$$

where FOC$_\lambda$ holds with equality if $\lambda^* > 0$.

Proposition 2.4 *(ia)An optimal solution to problem (2.13) exists. (ib)The solution is unique (a.e.).*

(ii)The principal is better off if he contracts with agents whose interests are more in line with his own, i.e. $\frac{\partial}{\partial\rho}E\pi(\rho)'_{\epsilon,\lambda} > 0$.

(iii)Under Assumption 2.3 it is optimal to set $\epsilon > 0$ *if* ρ *is sufficiently close to* 1; $\frac{\partial\epsilon^*}{\partial\rho} \geq 0$.

(iv)If $\rho \leq 0$ *it is optimal to set* $\lambda^* = 0$.

(v)If $0 < \rho < 1$ *the principal will choose the boundary of the interval control such that* $\lambda^* > \bar{\lambda}$, *where* $\bar{\lambda} = \arg\max E\pi(\rho)'_{\epsilon,\lambda}$. *If* $\rho = 1$, $\lambda^* = \bar{\lambda}$.

(vi) λ^* *increases in* ρ.

A choice of $\epsilon^* > 0$ serves to motivate the agent to supply high effort. Prohibiting the agent from choosing extreme actions protects the principal from opportunistic behavior of his agent. Forbidding the choice which is best under ignorance is a very costly incentive instrument. It will only be used if the benefit of doing so is sufficiently large. Since the value of delegation increases with the correlation between interests this will be the case when the principal's and the agent's interests are closely aligned. Ultimately, when interests are negatively correlated, the marginal value of additional effort for the principal is negative, how ever he delegates the task to the agent. This logic suggests that the principal will stop fostering the agent's effort supply by prohibiting the agent from choosing middling actions when the measure of the severity of conflicts of interests, ρ, falls below a critical value and will contend himself with using safeguards only.

As long as the principal does not find it optimal to give his agent full discretion, he will give a freedom of action, which is larger than would be optimal from an ex post perspective only. By sacrificing some of the expected benefits from an informed decision of his agent, he induces a higher effort choice of him. This results reflects the analogue to the result in Aghion and Tirole (1997): the principal trades off increasing the agent's initiative versus losing control.

Consider a range of ρ such that $\epsilon = 0$ is optimal in this range. Then, agents whose interests are closer aligned with those of the principal get more freedom. This result hinges crucially on the assumption of identical marginals. If the prior means of principal and agent differ, agents with higher ρ need not be better for the principal nor need they be given a larger degree of freedom.

Note however that the result that a principal can induce higher effort by prohibiting the agent from taking intermediate choices - i.e. by forbidding the agent's most preferred alternative when ignorant - applies in this case too. Note also that the cost of using such an incentive device decreases as the prior means move apart. In contrast to the present situation the principal does not punish himself ex post when he makes the ignorant life for his agent unpleasant. Apart from these considerations it should be possible to eliminate differences in opinions ex ante trough some kind of screening procedure.

2.5 Robustness Considerations

2.5.1 The Optimality of $\Gamma_{\epsilon,\lambda}$

Proposition 2.1 has been stated without providing the idea behind the result, because the analysis of the past two sections was needed to shape the intuition for the result. The contract $\Gamma_{\epsilon,\lambda}$ is optimal if and only if it can never be optimal to prohibit the agent from choosing actions in intervals of the type $(\mu+\gamma-t, \mu+\gamma+t)$ with $\gamma - t > 0$. This is how we prove proposition 1 formally in the appendix.[17] This result implies a number of things: it implies that the principal cannot improve upon the contract $\Gamma_{\epsilon,\lambda}$ by prohibiting the agent from choosing actions that are allowed under contract $\Gamma_{\epsilon,\lambda}$. It also implies that the principal does not want to allow choices in the interior of an interval that he has forbidden with $\Gamma_{\epsilon,\lambda}$.[18] Combined with the fact that the boundaries of the interval control and the interval prohibition were already chosen optimally, these arguments imply that $\Gamma_{\epsilon,\lambda}$ will be the optimal contract, when the principal can choose to forbid any collection of open sets.

The following intuition underlies the result: prohibiting the agent from choosing actions in an interval $(\mu+\gamma-t, \mu+\gamma+t)$ must decrease the agent's incentives for information acquisition. It does not interfere with the agent's choice under ignorance but limits his discretion in case he is informed. Thus, the agent's marginal value of information must decrease, stifling his incentives for information acquisition. A sufficient condition that effectively rules out $\Gamma_{\gamma t} := dom \quad f_\eta \backslash (\mu+\gamma-t, \mu+\gamma+t)$ as part of an optimal contract is then that the principal's expected payoff from decisions of an informed agent under contract $\Gamma_{\gamma t}$, $E\pi_t'$, is weakly decreasing in t. In case interests are perfectly aligned this is trivially true. In case interests are conflicting it requires that the variance of the distribution be large enough. Intuitively, it can never be optimal to leave the agent the possibility of reacting

[17] Any interval other than the ones considered in $\Gamma_{\epsilon,\lambda}$ can be written as such an interval, located symmetrically around some point $\mu+\gamma$. Note, however, that the concept can only apply to finite values, since otherwise we cannot find the "middle" of a prohibited interval. Thus interval controls are not the limit of these policies as $t \to \infty$!

[18] Reallowing some selected alternatives within the already forbidden intervals is effectively nothing else than setting ϵ and λ differently and prohibit the agent in addition from choosing actions in an interval of the type discussed here.

in the wrong direction.

Consequently, if it is optimal to exclude some specific alternatives beyond a certain threshold $\mu + \lambda$, then it must be optimal to exclude all alternatives beyond the threshold. Furthermore, if the principal chooses to elicit more effort from the agent than e_f^{sb}, there is no better way to achieve this goal than through the contract Γ_{ϵ^*}. Intuitively, the principal wants to increase the agent's effort supply at the lowest possible costs. If an alternative is allowed, that is closer to the mean than $\mu + \epsilon^*$, this alternative will be chosen when the agent is ignorant. Let this alternative be $\mu + \epsilon^a$. But then prohibiting the agent from choosing actions in the interval $(\mu + \epsilon^a, \mu + \epsilon^*)$ both decreases the agent's effort supply and the principal's expected payoff ex post. Hence this cannot be optimal.

2.5.2 Monetary contracts

So far we have not allowed for monetary contracts. Consequently we cannot be sure whether the results are artifacts of our focus on nonmonetary contracting schemes. We now show that this is not the case. Of particular interest (in our view) is the robustness of the result that the principal finds it optimal to interfere with the agent's choices even when there are no conflicts of interests. Therefore we shall reimpose Assumption 2.1: interests are perfectly aligned.

Let the agent and the principal each have utility functions that are additive separable in financial income and private benefits. The principal is risk neutral with respect to income shocks. The agent's utility function is now

$$U(x, \eta) + V(\tau(.))$$

where τ stands for the transfer from the principal to the agent.

Observability and Verifiability: absent monetary payments a contract was simply a restriction on the agent's action space. Of course this presumes that the agent's action is verifiable. In the following we maintain, for reasons of consistency with the previous analysis, the same assumptions about observability and verifiability as before: the action

x is observable and verifiable. In contrast, the true realizations of the underlying random variable are neither observable nor verifiable to the principal. Moreover, he can neither observe nor verify whether the agent has learned the realization of $\tilde{\eta}$ at all.

The possibility of monetary contracting opens various possibilities depending on what the contract can condition on. In particular one has to distinguish two situations depending on whether the contract can condition on payoffs or not. We consider these cases in turn.

Contractible Payoffs

If payoffs are contractible and the agent is risk neutral it is easy to see that first best can be implemented. In this case one has $V(\tau) \equiv \tau(x, \eta)$. The principal can implement first best with the contract $\tau(x, \eta) = \frac{-1}{2}(x - \eta)^2$ and by giving the agent full freedom of action to take any decision x that he finds optimal. The resulting choices will be $x = \eta$ if the agent is informed and $x = \mu$ if he is not. The agent's effort choice satisfies $g'(e) = \frac{1+A}{2}\sigma^2$. Hence, first best is implemented. This corresponds to the well known solution of selling the store to the agent.

It is also well known that the first best can in general not be attained if the agent is risk averse or/and if the agent faces limited liability constraints. Since these issues are well known by now we shall not pursue them here any further but rather turn to the case of noncontractible payoffs. Hence we assume:

Assumption 2.4: the agent is risk neutral with respect to income shocks, $V(\tau) \equiv \tau$ and has unlimited wealth.

Noncontractible Payoffs

Contracts and the Revelation Principle. In case payoffs are not contractible the final outcome will depend on how the contract affects the agent's choice of what value of η to announce to the principal (under a direct mechanism) or equivalently what pair of x, τ to "choose" when the true value of the information variable is η. By the revelation principle we can think of a contract as specifying a direct, incentive compatible mechanism. Such a mechanism prescribes (i) an effort level e, (ii) an action $x_\emptyset$ and a transfer payment $\tau_\emptyset$ for the case where the agent is ignorant, and (iii) and for each realization of the random variable $\tilde{\eta}$ an action $x(\eta)$ and a transfer payment $\tau(\eta)$ for the case where the agent has learned the realization of $\tilde{\eta}$ and announces a specific η. Such a mechanism is incentive compatible concerning project choices if (a) the agent prefers to reveal his ignorance rather than announce any other specific value of the information variable η in case he is ignorant, (b) the agent prefers to announce the true realization of $\tilde{\eta}$ in case he knows it, rather than claim to be ignorant, and (c) the agent prefers to announce the true realization, η, rather than any other value of the information variable, $\hat{\eta}$. The mechanism is incentive compatible with respect to the agent's effort choice if the marginal value of information to the agent just covers the agent's marginal cost of effort. Finally, the mechanism is ex ante individually rational if the agent's expected utility is at least as high as his cost of effort.[19]

The solution to this problem is somewhat less than straightforward. The problem does not fit into the standard framework usually studied in mechanism design theory because of the constraints (a), (b) and the fact that the agent's effort choice is endogenous and must be guaranteed to be incentive compatible.[20] However, we can solve the problem nevertheless. Moreover, as in the previous analysis, the interesting questions and trade-offs involved can be studied without determining the precise level of the second best optimal effort choice being implemented. More precisely, we raise the following sequence of questions: (i) under what condition does the principal not contend himself with the effort level that the agent is willing to supply voluntarily by just granting him full freedom of choice? (ii) Suppose then that the principal does want to foster the agent's effort choice beyond the level of the agent's voluntary supply. What is the cheapest, incentive compatible way to increase the agent's effort supply to the desired level?

[19] To the best of our knowledge this is the first analysis of delegation in a framework of optimal mechanism design allowing for monetary transfers. Holmström (1984) has stated his general problem in a mechanism design framework but his analysis is carried through and explicit solutions are derived in a situation where monetary contracts are not admitted. Every paper on delegation we know of has followed this approach.

[20] At the end of this section, we will be more precise with what we mean by "standard" and "non-standard" as well as explain the relation of the present analysis to the existing literature in this field.

Since the problem is symmetric the solution is characterized for the upper half of the support only. Letting θ denote the Lagrange multiplier attached to the incentive compatibility constraint on effort choice, we prove in appendix B that the optimal mechanism has the following features:

Proposition 2.5 *If* $e_f^{sb} + \frac{A\sigma^2}{g''(e_f^{sb})2}(1-A) > 1$ *it is not optimal to implement the same project choices as in the first best situation. An optimal contract satisfies*

$$\begin{aligned} x_\emptyset &= x(\mu)\, and \\ \frac{\partial x}{\partial \eta} &\geq 0. \end{aligned}$$

If $\frac{\partial}{\partial \eta}\left[\eta + \frac{\theta}{e}\frac{A}{1+A}\frac{1-F_\eta}{f_\eta}\right] \geq 0\, \forall \eta$, *the optimally implemented decisions satisfy*

$$x^*(\eta) = \eta + \frac{\theta}{e}\frac{A}{1+A}\frac{1-F_\eta}{f_\eta}.$$

Figure 2.2 gives a graphical illustration of the main properties of the contract.

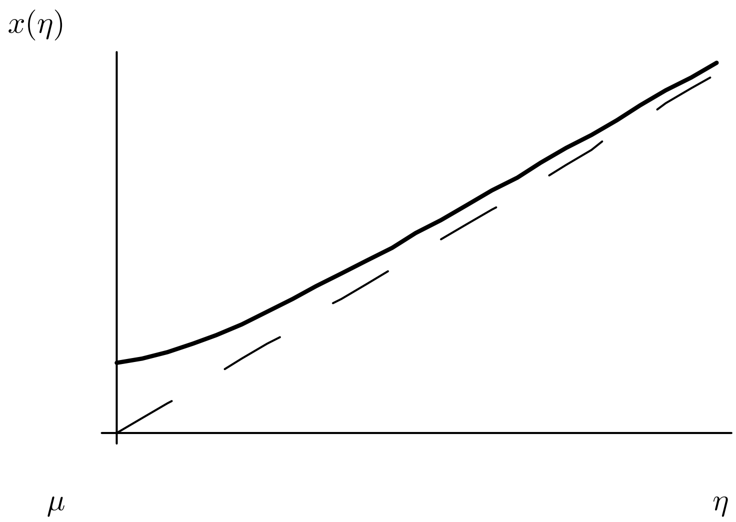

Figure 2.2

The dashed line in the figure corresponds to the 45° line and represents first best optimal project choices. The solid curve represents the second best optimally implemented decisions and is drawn assuming an increasing hazard rate.

The reason why it is not optimal to implement $x = \eta$ is the same as in proposition 2.2.

The agent supplies too little effort from a social perspective if first best optimal project choices are implemented. In order to increase the agent's effort choice, the principal rewards the agent's knowledge, or equivalently, punishes the agent's ignorance. This, in turn, cannot be done without a departure from first best project choices: as before the agent's incentive to become informed is determined by the marginal value of information to him. This consists of (i) the avoided disutility of risk, $\frac{A\sigma^2}{2}$, and (ii) the expected excess rent of an informed as opposed to an ignorant agent under the contract. The second term can be understood by an inspection of the figure. An informed agent's indirect utility is the higher the smaller the difference between the implemented decision and the first best decision. The expected indirect utility of an agent that he expects to receive under the contract when he ends up being informed is therefore proportional to a weighted average of these deviations from first best project choices. Hence, the expected rent of an informed agent is proportional to the area between the solid and the dashed curve. But if first best choices are implemented, the solid and the dashed curve coincide and the agent's incentive to become informed is exactly $\frac{A\sigma^2}{2}$. A sufficient condition such that this cannot be optimal is that $\frac{A\sigma^2}{2g''(e_f^{sb})}(1-A) > 1 - e_f^{sb}$. Essentially the same argument as in proposition 2.2 underlies this comparison (for details see appendix B). Prohibiting the agent from choosing the optimal decision x is bad if he is not informed, which happens with probability $1 - e_f^{sb}$. This must be compared with the marginally avoided risk, i.e. the increase in the agent's effort supply times the principal's disutility of risk, $\frac{A\sigma^2}{2g''(e_f^{sb})}$, *net of the compensation for the agent's increased effort cost,* $A \cdot \frac{A\sigma^2}{2g''(e_f^{sb})}$.

The difference between the present and the preceding analysis is that the presence of monetary contracts combined with the unboundedness of the agent's utility function implies that his individual rationality constraint is binding at the optimum. This shows in particular that the essence of the main result does not depend on the absence of a quit option for the agent.

Turning to the details of the contract: there is no way to distinguish between the case where the agent claims to be ignorant and the case where the agent knows that the realization of $\tilde{\eta}$ is μ, because in the former case his best guess of $\tilde{\eta}$ is μ as well. Therefore the same decision has to be implemented for both cases, hence the constraint $x_{\emptyset} = x(\mu)$. $\frac{\partial x}{\partial \eta} \geq 0$ is the standard condition, stating that only monotone decision schedules are implementable. For purely suggestive reasons we have drawn the solid curve assuming an increasing hazard rate[21], satisfying $\lim_{\eta\to\infty} \frac{f_\eta}{1-F_\eta} = \infty$. In this case distortions are greatest around the mean and decrease steadily the more one moves outwards. At the tails of the distribution, implemented decisions coincide approximately with the first best optimal decisions. This, however, depends on the distribution F and is not a general property of the optimal contract. Some, in our view, interesting features of the optimal contract *are* independent of the specific distributional assumptions. The most interesting of which has been discussed at length already: the contract does not implement first best optimal project decisions even though there are no conflicts of interest. In particular, some projects within an interval around the mean will never be implemented and the agent is again forced to take a clear choice in either one direction. Second, and also irrespective of the specific distribution F: the principal always wants to spread all distortions over the whole support of η. Hence we conclude:

Whether monetary contracts are admitted or not, it may pay for the principal to demand a clear statement from the agent even if the latter does not necessarily have a clear answer on his mind.

We end with a technical remark on the relation of the model presented in this last section to the literature on optimal mechanism design. The development of our solution concept has built on Lewis and Sappington (1993b) and Crémer, Khalil and Rochet (1998a). Both papers study modified versions of the procurement problem originally due to Baron and Myerson. In Lewis and Sappington the adverse selection problem is complicated by the possibility of ignorance by the agent. However, they do not allow for endogeneity of information acquisition. Hence the probability of ignorance is exogenous in this model. Crémer et.al allow for the endogeneity of the information structure but work with a learning tech-

[21] Increasing, but not too fast, since otherwise $\eta + \frac{\theta}{e}\frac{A}{1+A}\frac{1-F_\eta}{f_\eta}$ might not be monotonic in η.

nology eliminating, in equilibrium, all uncertainty about the agent being ignorant or not. In the present analysis the adverse selection problem is complicated by the simultaneous presence of both problems: endogeneity of information structure as well as the possibility that the agent is as ignorant as the principal himself.

2.6 Conclusion

The principal uses two types of instruments to control the agent's behavior. Agents which are known to have interests very much out of line with those of the principal have to be controlled with instruments restricting their misbehavior. Their choice set will be a narrow set around the optimal choice under ignorance. Obviously those agents will also deliver little effort, because they will hardly be able to use their improved knowledge in their own best interest. Agents with interests which parallel those of the Principal may be subject to another kind of restriction on their choice set. Since the principal benefits very much from their increased efforts he could try to make life unpleasant for them if they have to decide under ignorance. Such an effort enhancing prohibition must therefore forbid the choice which is optimal under ignorance. The surprising thing is that a principal might find it optimal to use such instruments even if the agent is a perfect clone of himself.

The bottom line of our argument is that curtailing the discretion or limiting the authority of agents may have as favorable incentive effects as increasing their discretion or giving them more authority. Moreover, the impact of such rules on the expected payoff of the principal may be positive. While Aghion and Tirole (1997) have found that delegating authority -thus increasing discretion- may be a good thing, even if there are conflicting interests, the present paper points at the mirror image of this result: curtailing discretion might be a good thing even if there are no conflicts of interests. As a consequence, there need not be a monotonic relationship between the severity of the conflict of interest and the restrictiveness of the optimal decision space. An agent with better aligned interests need not be given a higher degree of freedom - even if he would always (on expectation) use this additional freedom in the principal's interest.

2.7 Appendix A

Definitions and Notation:

- $dF_\eta = f(\eta)d\eta; \qquad dF_{\varphi|\eta} = f(\varphi\,|\eta)d\varphi$
- $\Delta F_\epsilon := F(\mu) - F(\mu - \epsilon) = F(\mu + \epsilon) - F(\mu)$
- $dG_\epsilon := \frac{dF_\eta}{\Delta F_\epsilon}$
- $\delta := \mu - \int_{\mu-\epsilon}^{\mu} \eta dG_\epsilon = \int_{\mu}^{\mu+\epsilon} \eta dG_\epsilon - \mu$
- $F_{\mu-\lambda} := 1 - F(\mu + \lambda) = F(\mu - \lambda)$
- $dG_\lambda := \frac{f(\eta)}{F_{\mu-\lambda}} d\eta$
- $\nu := \int_{\mu+\lambda}^{\infty} \eta dG_\lambda - \mu = \mu - \int_{-\infty}^{\mu-\lambda} \eta dG_\lambda$

$\frac{dG_\epsilon}{d\eta}$ is the density of η, conditional on the fact that $\eta \in [\mu - \epsilon, \mu]$. $\mu - \delta$ is the mean of η, conditional on $\eta \in [\mu - \epsilon, \mu]$. Likewise, $\frac{dG_\lambda}{d\eta}$ is the probability density of η, conditional on the fact that $\eta \in [\mu + \lambda, \infty]$. $\nu + \mu$ is the conditional mean of η, conditional on the fact that $\eta \in [\mu + \lambda, \infty]$.

Proof of Proposition 2.1. It is suggested that the reader go through this proof at the end of the analysis. Consider conflicting interests and the candidate policy
$\Gamma_t = dom \quad f(\eta) \backslash (\mu + \gamma - t, \mu + \gamma + t)$. By the same arguments as in Lemma 2.2 below
$E\pi_t' = K - (1 - \rho)\sigma^2 -$

$$\left(\begin{array}{c} \int_{\mu+\gamma-t}^{\mu+\gamma} (\mu + \gamma - t - \eta)^2 dF_\eta + \int_{\mu+\gamma}^{\mu+\gamma+t} (\mu + \gamma + t - \eta)^2 dF_\eta \\ +2(1-\rho) \int_{\mu+\gamma-t}^{\mu+\gamma} (\mu + \gamma - t - \eta)(\eta - \mu) dF_\eta \\ +2(1-\rho) \int_{\mu+\gamma}^{\mu+\gamma+t} (\mu + \gamma + t - \eta)(\eta - \mu) dF_\eta \end{array} \right)$$

Taking derivatives with respect to t : $\frac{\partial}{\partial t} E\pi_t' =$

$$(-1) \left(\begin{array}{c} (-2) \int_{\mu+\gamma-t}^{\mu+\gamma} (\mu + \gamma - t - \eta) dF_\eta - 2(1-\rho) \int_{\mu+\gamma-t}^{\mu+\gamma} (\eta - \mu) dF_\eta \\ +2 \int_{\mu+\gamma}^{\mu+\gamma+t} (\mu + \gamma + t - \eta) dF_\eta + 2(1-\rho) \int_{\mu+\gamma}^{\mu+\gamma+t} (\eta - \mu) dF_\eta \end{array} \right) \tag{2.14}$$

We must show that the term inside the brackets is nonnegative. If ρ is equal to 1, the term inside the brackets becomes $(-2)\int_{\mu+\gamma-t}^{\mu+\gamma}(\mu+\gamma-t-\eta)dF_\eta + 2\int_{\mu+\gamma}^{\mu+\gamma+t}(\mu+\gamma+t-\eta)dF_\eta$, which can never be negative: in the first integrand $(\mu+\gamma-t-\eta)$, all realizations of $\tilde{\eta}$ are larger than $\mu+\gamma-t$, in the second all realizations are smaller than $\mu+\gamma+t$. (Since marginals are identical, this is also a proof for the case where distributions are not normal. If $\rho<1$, this is no longer true, since we have used the fact that conditional means are linear.) Making use of the fact that the distribution is normal, to calculate the conditional means of the truncated distribution,[22] the term inside the brackets is equal to

$$\begin{aligned} &-2(\gamma-t)\left(F_{\mu+\gamma}-F_{\mu+\gamma-t}\right)+2\rho\sigma\left(\phi\left(\frac{\gamma-t}{\sigma}\right)-\phi\left(\frac{\gamma}{\sigma}\right)\right) \\ &+2(\gamma+t)\left(F_{\mu+\gamma+t}-F_{\mu+\gamma}\right)-2\rho\sigma\left(\phi\left(\frac{\gamma}{\sigma}\right)-\phi\left(\frac{\gamma+t}{\sigma}\right)\right) \end{aligned} \tag{2.15}$$

where ϕ is the density of the standard normal. We must show that this is positive. The first line in (2.15) is positive iff

$$(\gamma-t)<\rho\sigma\frac{\left(\phi\left(\frac{\gamma-t}{\sigma}\right)-\phi\left(\frac{\gamma}{\sigma}\right)\right)}{\left(F_{\mu+\gamma}-F_{\mu+\gamma-t}\right)}=\rho E\left(\eta\middle|\eta\in[\frac{\gamma-t}{\sigma},\frac{\gamma}{\sigma}]\right)$$

hence iff ρ is large enough. The second line in (2.15) is positive iff

$$(\gamma+t)>\rho E\left(\eta\middle|\eta\in[\frac{\gamma}{\sigma},\frac{\gamma+t}{\sigma}]\right)$$

hence for all values of ρ. In particular these two things imply that $\frac{\partial}{\partial t}E\pi_t'<0$ for ρ close to 1.

If ρ is equal to zero, we must only consider

$$-2(\gamma-t)\left(F_{\mu+\gamma}-F_{\mu+\gamma-t}\right)+2(\gamma+t)\left(F_{\mu+\gamma+t}-F_{\mu+\gamma}\right)$$

[22] See Johnson and Kotz (1970).

This term will be positive if σ is large enough: $(\gamma + t) - (\gamma - t) = 2t$.

However $(F_{\mu+\gamma} - F_{\mu+\gamma-t}) - (F_{\mu+\gamma+t} - F_{\mu+\gamma}) > 0$ since an interval of size t contains more mass the lower the location of the interval. Therefore we must make sure that $(F_{\mu+\gamma+t} - F_{\mu+\gamma})$ is not "much" smaller than

$(F_{\mu+\gamma} - F_{\mu+\gamma-t})$. This can be done by choosing σ large enough, since

$$\lim_{\sigma\to\infty} (F_{\mu+\gamma} - F_{\mu+\gamma-t}) - (F_{\mu+\gamma+t} - F_{\mu+\gamma}) = 0.$$

To sum up: if σ is large, we cannot find a γ such that F is "very steep" in the interval $[\mu + \gamma - t, \mu + \gamma]$ and "very flat" in the interval $[\mu + \gamma, \mu + \gamma + t]$.

Since (2.15) is linear in ρ, it is either minimized at $\rho = 1$ or at $\rho = 0$, depending on

$$sign\left(\phi\left(\frac{\gamma - t}{\sigma}\right) + \phi\left(\frac{\gamma + t}{\sigma}\right) - 2\phi\left(\frac{\gamma}{\sigma}\right)\right)$$

But then, provided that σ is large enough, the above arguments imply that $\frac{\partial}{\partial t} E\pi'_t < 0$ for all values of ρ and $t \geq 0$. This completes the proof. ■

Proof of Lemma 2.1 .

$$e = h\left[\frac{A}{2}(\sigma^2 + \epsilon^2) - A\int_{\mu}^{\mu+\epsilon} (\mu + \epsilon - \eta)^2 dF_\eta - \int_{\mu+\lambda}^{\infty} (\mu + \lambda - \eta)^2 dF_\eta\right]$$

where $h := g'^{-1}$ exists by strict convexity of g. By Leibniz's rule and the inverse function theorem

$$\frac{\partial e}{\partial \lambda} = \frac{-A2}{g''(e)}\left\{\int_{\mu+\lambda}^{\infty} (\mu + \lambda - \eta) dF_\eta\right\} > 0$$

since $\tilde{\eta} \geq \mu + \lambda$.

$$\begin{aligned}\frac{\partial e}{\partial \epsilon} &= \frac{A}{g''(e)}\left\{\epsilon - \int_{\mu}^{\mu+\epsilon} 2(\mu + \epsilon - \eta) dF_\eta\right\} \\ &= \frac{A}{g''(e)}\left[\epsilon - 2\Delta F_\epsilon \{\epsilon - \delta\}\right]\end{aligned}$$

Then

$$\frac{\partial e}{\partial \epsilon} > 0 \Longleftrightarrow \epsilon - 2\Delta F_{\epsilon}\left\{\epsilon - \delta\right\} > 0$$

Observe now that this will always be true except in pathological cases: We have $\delta \leq \epsilon$ because a mean is a convex combination. $2\Delta F_{\epsilon} \leq 1$ because at most half of the mass can lie in the upper half of the distribution. Thus only in the case where $\delta = 0$ and ϵ is set equal to the upper bound of the distribution effort will not increase. Observe however that $\delta = 0$ means that the distribution has point mass around μ, i.e. is degenerate on μ. We can thus safely ignore this case. ■

Proof of Proposition 2.2. (I) A continuous function on a compact domain must attain a maximum by Weierstrass' Theorem.
(II) $\lambda^* = \infty$ is obvious from Lemma 2.1: $\frac{\partial E\pi'_{\epsilon,\lambda}}{\partial\lambda} > 0 \Leftrightarrow \frac{\partial e}{\partial\lambda} > 0$. Lemma 2.3 below will establish that $\frac{\partial E\pi'_{\epsilon,\lambda}}{\partial\lambda} \geq 0$ for all $\lambda \geq 0$ and that $\lim_{\lambda\to\infty}\frac{\partial E\pi'_{\epsilon,\lambda}}{\partial\lambda} = \lim_{\lambda\to\infty}\frac{\partial e}{\partial\lambda} = 0$. Since $\lim_{\lambda\to\infty} E(\pi'_{\epsilon,\lambda} - \pi''_{\epsilon,\lambda})$ and $\lim_{\lambda\to\infty} e$ are both constant, this implies that:

$$\lim_{\lambda\to\infty}[e\frac{\partial E\pi'_{\epsilon,\lambda}}{\partial\lambda} + \frac{\partial e}{\partial\lambda}E(\pi'_{\epsilon,\lambda} - \pi''_{\epsilon,\lambda})] = 0.$$

To prove that ϵ^* must be finite, observe that:

$$\lim_{\epsilon\to\infty} E\pi'_{\epsilon} = \lim_{\epsilon\to\infty} K - \int_{\mu}^{\mu+\epsilon}(\mu+\epsilon-\eta)^2 dF_{\eta} = -\infty$$

and $\lim_{\epsilon\to\infty} E\pi''_{\epsilon} = \lim_{\epsilon\to\infty} K - (\frac{\sigma^2+\epsilon^2}{2}) = -\infty$. Hence increasing ϵ without bounds cannot be optimal. This proves the claim.
(III) For the sake of clarity, we first state all derivatives needed for the evaluation of

(2.10):

$$\begin{aligned}
\frac{\partial e}{\partial \epsilon} &= \frac{A\left[\epsilon - 2\Delta F_\epsilon \left\{\epsilon - \delta\right\}\right]}{g''(e)} \\
\frac{\partial^2 e}{\partial \epsilon^2} &= \frac{A(1-2\Delta F_\epsilon)}{g''(e)} - \frac{A^2\left[\epsilon - 2\Delta F_\epsilon \left\{\epsilon - \delta\right\}\right]^2}{g''(e)} \frac{g'''(e)}{g''(e)^2} \\
\frac{\partial}{\partial \epsilon} E(\pi'_\epsilon - \pi''_\epsilon) &= \left[\epsilon - 2\Delta F_\epsilon \left\{\epsilon - \delta\right\}\right] \\
\frac{\partial^2}{\partial \epsilon^2} E(\pi'_\epsilon - \pi''_\epsilon) &= 1 - 2\Delta F_\epsilon
\end{aligned}$$

The principal should increase ϵ (locally) at any point as long as

$$-\epsilon + e\left[\epsilon - 2\Delta F_\epsilon \left\{\epsilon - \delta\right\}\right] + \frac{A\left[\epsilon - 2\Delta F_\epsilon \left\{\epsilon - \delta\right\}\right]}{g''(e)} E(\pi'_\epsilon - \pi''_\epsilon) \geq 0$$

To see whether the principal wants to set $\epsilon > 0$ at all, we first evaluate the derivative of the principal's payoff function at $\epsilon = 0$:

$$\lim_{\epsilon \to 0} -\epsilon + e\left[\epsilon - 2\Delta F_\epsilon \left\{\epsilon - \delta\right\}\right] + \frac{A\left[\epsilon - 2\Delta F_\epsilon \left\{\epsilon - \delta\right\}\right]}{g''(e)} E(\pi'_\epsilon - \pi''_\epsilon) = 0$$

because $\lim_{\epsilon \to 0} \left[\epsilon - 2\Delta F_\epsilon \left\{\epsilon - \delta\right\}\right] = 0$. So we must check whether $\epsilon = 0$ constitutes a local minimum or a local maximum. Define

$$\digamma := e\left\{\epsilon - 2\Delta F_\epsilon(\epsilon - \delta)\right\} + \frac{A(\epsilon - 2\Delta F_\epsilon(\epsilon - \delta))}{g''} E(\pi'_\epsilon - \pi''_\epsilon). \tag{2.16}$$

The second derivative of the principal's payoff function is then $-1+\frac{\partial \digamma}{\partial \epsilon}$. Hence if $\lim_{\epsilon \to 0} \frac{\partial \digamma}{\partial \epsilon} > 1$, it pays for the principal to increase ϵ. By l'Hôpital $\lim_{\epsilon \to 0} \frac{\partial \digamma}{\partial \epsilon} = \lim_{\epsilon \to 0} \frac{\digamma}{\epsilon}$. Thus the principal sets $\epsilon > 0$ if:

$$\lim_{\epsilon \to 0} \left[\begin{array}{c} \left(e + A\frac{(\sigma^2+\epsilon^2) - 2\int_\mu^{\mu+\epsilon} (\mu+\epsilon-\eta)^2 dF_\eta}{2g''(e)} \right) \times \\ \frac{\epsilon - 2\Delta F_\epsilon\{\epsilon - \delta\}}{\epsilon} \end{array} \right] > 1$$

By lemma 2.1, $\epsilon - 2(\epsilon-\delta)\Delta F_\epsilon \geq 0$. By l'Hôpital[23]

$$\lim_{\epsilon\to 0}\frac{\epsilon - 2\Delta F_\epsilon\{\epsilon-\delta\}}{\epsilon} = \lim_{\epsilon\to 0} 1 - 2\Delta F_\epsilon = 1. \tag{2.17}$$

$\lim_{\epsilon\to 0}\left((\sigma^2+\epsilon^2) - 2\int_\mu^{\mu+\epsilon}(\mu+\epsilon-\eta)^2 dF_\eta\right) = \sigma^2$. Thus the principal can benefit from ϵ-controls if $e_f^{sb} + \frac{A\sigma^2}{2g''(e_f^{sb})}$ is larger than 1, which proves the claim.

(IVa) Uniqueness: From (2.10), the optimal solution ϵ^* must satisfy the first order necessary condition

$$\{\epsilon - 2\Delta F_\epsilon(\epsilon-\delta)\}\left(e + \frac{AE(\pi'_\epsilon - \pi''_\epsilon)}{g''(e)}\right) = \epsilon \tag{2.18}$$

and the second order sufficient condition

$$\begin{aligned} 1 \;>\; & \left(e + \frac{AE(\pi'_\epsilon - \pi''_\epsilon)}{g''}\right)(1-2\Delta F_\epsilon) \\ & + \frac{A\left(\epsilon - 2\Delta F_\epsilon(\epsilon-\delta)\right)^2}{g''}\left[2 - \frac{Ag'''E(\pi'_\epsilon - \pi''_\epsilon)}{(g'')^2}\right] \end{aligned} \tag{2.19}$$

The left hand side of (2.18) was defined in (2.16) as $\mathcal{F}$. Obviously $\mathcal{F}(0) = 0$. and for $\epsilon^* > 0$, (2.18) must hold with equality too. Hence the optimum is an intersection of $\mathcal{F}$ with the 45° degree line. Moreover, from the second order condition, the slope of $\mathcal{F}$ at this intersection must be smaller than 1. $\frac{\partial \mathcal{F}}{\partial\epsilon}$ is given by the right hand side of (2.19). Assume $\frac{g'''(e)}{(g''(e))^2}\frac{A\sigma^2}{2} > 2$. We are going to show now that there cannot be equilibria with $\epsilon > \epsilon^{critical}$ for some $\epsilon^{critical}$ when this assumption holds. A sufficient condition for this is that $\frac{\partial\mathcal{F}}{\partial\epsilon} < 1 \forall\epsilon > \epsilon^{critical}$. Consider first the second line of the expression on the right hand side of (2.19): $\frac{A(\epsilon-2\Delta F_\epsilon(\epsilon-\delta))^2}{g''}\left[2 - \frac{Ag'''E(\pi'_\epsilon-\pi''_\epsilon)}{(g'')^2}\right] =: \Omega$. By assumption, Ω is nonpositive $\forall\epsilon : \lim_{\epsilon\to 0}\Omega = 0$, because $\frac{g'''(e)}{(g''(e))^2}$ attains a fixed value for $\epsilon \to 0$, and $\lim_{\epsilon\to 0}(\epsilon - 2\Delta F_\epsilon(\epsilon-\delta)) = 0$. For $\epsilon > 0 : \Omega < 0$. Consider next the expression in the first line. Distinguish two cases: case (i): $e_f^{sb} + \frac{A\sigma^2}{2g''(e_f^{sb})} < 1$. Then $\frac{\partial\mathcal{F}}{\partial\epsilon}|_{\epsilon=0} < 1$ by (III).

[23] It has been argued in the text that $\lim_{\epsilon\to 0}\frac{\partial}{\partial\epsilon}[EU'_\epsilon - EU''_\epsilon] \approx \epsilon$ for ϵ very small. This is the formal explanation for this result.

Moreover, $\frac{\partial}{\partial\epsilon}\left[(1-2\Delta F_\epsilon)\left(e+\frac{AE(\pi'_\epsilon-\pi''_\epsilon)}{g''(e)}\right)\right] =$

$$-2f_{\mu+\epsilon}\left(e+\frac{AE(\pi'_\epsilon-\pi''_\epsilon)}{g''(e)}\right)+\left\{2-Ag'''\frac{E(\pi'_\epsilon-\pi''_\epsilon)}{(g''(e))^2}\right\}\times$$
$$\frac{(1-2\Delta F_\epsilon)A(\epsilon-2\Delta F_\epsilon(\epsilon-\delta)}{g''(e)}$$

<0 by the fact that $\frac{g'''(e)}{(g''(e))^2}\frac{A\sigma^2}{2}>2$. But then $(1-2\Delta F_\epsilon)\left(e+\frac{AE(\pi'_\epsilon-\pi''_\epsilon)}{g''(e)}\right)<1;\forall\epsilon>0$, hence $\frac{\partial F}{\partial\epsilon}<1;\forall\epsilon>0$ so and $\epsilon^{critical}=0$ and there can be no other intersection of F with the 45° degree line. Case (ii): $e_f^{sb}+\frac{A\sigma^2}{2g''(e_f^{sb})}>1$. Then $\frac{\partial F}{\partial\epsilon}\mid_{\epsilon=0}>1$ and $\epsilon^*>0$. At $\epsilon^*>0$, the smallest, strictly positive, value of ϵ which solves (2.18) with equality, we can conclude from (2.18) that

$$\begin{aligned}&\left(e+A\frac{E(\pi'_\epsilon-\pi''_\epsilon)}{g''(e)}\right)(1-2\Delta F_\epsilon)-1\\ = \;& -\left(e+A\frac{E(\pi'_\epsilon-\pi''_\epsilon)}{g''(e)}\right)2\Delta F_\epsilon\frac{\delta}{\epsilon}<0.\end{aligned}$$

Hence we know that at ϵ^*, $\left(e+A\frac{E(\pi'_\epsilon-\pi''_\epsilon)}{g''(e)}\right)(1-2\Delta F_\epsilon)<1$. We already know that $\frac{\partial}{\partial\epsilon}\left(e+A\frac{E(\pi'_\epsilon-\pi''_\epsilon)}{g''(e)}\right)(1-2\Delta F_\epsilon)<0$, so again $\frac{\partial F}{\partial\epsilon}<1;\forall\epsilon\geq\epsilon^*$. So $\epsilon^*=\epsilon^{critical}$ and there is again a unique maximum.

(IVb) follows immediately from (IVa). This completes the proof. ■

Proof of Proposition 2.3. The solution satisfies $F(A,\sigma,\epsilon)-\epsilon=0$ with

$$F:=e\left\{\epsilon-2\Delta F_\epsilon(\epsilon-\delta)\right\}+\frac{A(\epsilon-2\Delta F_\epsilon(\epsilon-\delta))}{g''}E(\pi'_\epsilon-\pi''_\epsilon).$$

$F(A,\sigma,\epsilon)$ is strictly decreasing in A iff $e+\frac{AE(\pi'_\epsilon-\pi''_\epsilon)}{g''(e(\epsilon))}$ is strictly decreasing in A. Thus

$$\begin{aligned}\frac{\partial}{\partial A}\left(e+A\frac{E(\pi'_\epsilon-\pi''_\epsilon)}{g''(e(\epsilon))}\right) &= \\ \frac{E(\pi'_\epsilon-\pi''_\epsilon)}{g''(e)}\left(2-Ag'''(e)\frac{E(\pi'_\epsilon-\pi''_\epsilon)}{g''(e)^2}\right) &< 0\Leftarrow\frac{g'''(e)}{g''(e)^2}>\frac{4}{A\sigma^2};\forall e,A\end{aligned}$$

Since $F(A,\sigma,\epsilon)\mid_{\epsilon=0}-0=0$ and $\lim_{\epsilon\to\infty}F(A,\sigma,\epsilon)-0<0$ and the equilibrium is unique, the same argument as in Theorem 1 of Milgrom and Roberts (1994) applies. Hence $\epsilon^*(A)$

is monotone nonincreasing in A. The argument of the effect of σ^2 is exactly the same. Now let ϵ^* be strictly positive. Define the function

$$H(A,\epsilon(A)) := \frac{\partial E\pi''}{\partial\epsilon} + e\frac{\partial}{\partial\epsilon}\left[E(\pi'_\epsilon - \pi''_\epsilon)\right] + \frac{\partial e}{\partial\epsilon}E(\pi'_\epsilon - \pi''_\epsilon)$$

By the implicit function theorem: $\frac{\partial\epsilon^*}{\partial A} = -\frac{\frac{\partial H}{\partial A}}{\frac{\partial H}{\partial\epsilon}}$. The denominator is the second order condition, thus negative.

$$\frac{-\partial H}{\partial A} = \frac{(\epsilon - 2\Delta F_\epsilon(\epsilon-\delta))\,E(\pi'_\epsilon - \pi''_\epsilon)}{g''(e)}\left(\frac{g'''(e)AE(\pi'_\epsilon - \pi''_\epsilon)}{g''(e)^2} - 2\right) > 0$$

by $\frac{g'''(e)A\sigma^2}{(g''(e))^2 2} > 2, \forall e, A$. Hence $\frac{\partial\epsilon^*}{\partial A} < 0$.

■

$E\pi'_{\epsilon,\lambda}(\rho)$ is not directly accessible to analysis. Lemma 2.2 below derives a more workable form. Define

$$L_{\epsilon,\lambda} := \begin{bmatrix} \int_\mu^{\mu+\epsilon}\int(\mu+\epsilon-\varphi)^2 dF_{\varphi|\eta}dF_\eta + \int_{\mu+\epsilon}^{\mu+\lambda}\int(\eta-\varphi)^2 dF_{\varphi|\eta}dF_\eta \\ +\int_{\mu+\lambda}^{\infty}\int(\mu+\lambda-\varphi)^2 dF_{\varphi|\eta}dF_\eta \end{bmatrix} \tag{2.20}$$

then $E\pi'_{\epsilon,\lambda}(\rho) = K - L_{\epsilon,\lambda}$.

Lemma 2.2

$$L_{\epsilon,\lambda} = \left\{ \begin{array}{c} (1-\rho)\sigma^2 + (2\rho-1)F_{\mu-\lambda}Var_\lambda + (2\rho-1)\Delta F_\epsilon Var_\epsilon + \\ F_{\mu-\lambda}(\nu-\lambda)^2 + \Delta F_\epsilon(\epsilon-\delta)^2 - 2(1-\rho)F_{\mu-\lambda}\nu(\nu-\lambda) + \\ 2(1-\rho)\Delta F_\epsilon\delta(\epsilon-\delta). \end{array} \right\} \tag{2.21}$$

Proof of Lemma 2.2. Through a series of four steps:

Step 1: $\int_{\mu+\epsilon}^{\mu+\lambda}\int(\eta-\varphi)^2 dF_{\varphi|\eta}dF_\eta =$

$$(1-\rho)\sigma^2 - \int_\mu^{\mu+\epsilon}\int(\eta-\varphi)^2 dF_{\varphi|\eta}dF_\eta - \int_{\mu+\lambda}^{\infty}\int(\eta-\varphi)^2 dF_{\varphi|\eta}dF_\eta$$

since $2\int\int_{\mu}^{\infty}(\eta-\varphi)^2 dF_{\varphi|\eta}dF_{\eta} = \int\int(\eta-\varphi)^2 dF_{\varphi|\eta}dF_{\eta} = 2(1-\rho)\sigma^2$.

Step 2: $\int\int_{\mu+\lambda}^{\infty}(\mu+\lambda-\varphi)^2 dF_{\varphi|\eta}dF_{\eta}$

$$\begin{aligned} = \;& \int\int_{\mu+\lambda}^{\infty}(\mu+\lambda-\eta)^2 dF_{\varphi|\eta}dF_{\eta} + \int\int_{\mu+\lambda}^{\infty}(\eta-\varphi)^2 dF_{\varphi|\eta}dF_{\eta} + \\ & 2\int\int_{\mu+\lambda}^{\infty}(\mu+\lambda-\eta)(\eta-\varphi)dF_{\varphi|\eta}dF_{\eta}. \end{aligned}$$

Expand $\int\int_{\mu}^{\mu+\epsilon}(\mu+\epsilon-\varphi)^2 dF_{\varphi|\eta}dF_{\eta}$ the same way. Then, by step 1 and 2: $L_{\epsilon,\lambda} = (1-\rho)\sigma^2 +$

$$\begin{aligned} & \int\int_{\mu}^{\mu+\epsilon}(\mu+\epsilon-\eta)^2 dF_{\varphi|\eta}dF_{\eta} + 2\int\int_{\mu}^{\mu+\epsilon}(\mu+\epsilon-\eta)(\eta-\varphi)dF_{\varphi|\eta}dF_{\eta} + \\ & \int\int_{\mu+\lambda}^{\infty}(\mu+\lambda-\eta)^2 dF_{\varphi|\eta}dF_{\eta} + 2\int\int_{\mu+\lambda}^{\infty}(\mu+\lambda-\eta)(\eta-\varphi)dF_{\varphi|\eta}dF_{\eta} \end{aligned} \quad (2.22)$$

Step 3: consider the first terms in each of the last two lines of (2.22).

$$\int\int_{\mu}^{\mu+\epsilon}(\mu+\epsilon-\eta)^2 dF_{\varphi|\eta}dF_{\eta} + \int\int_{\mu+\lambda}^{\infty}(\mu+\lambda-\eta)^2 dF_{\varphi|\eta}dF_{\eta} =$$

$$\begin{aligned} & \Delta F_{\epsilon}\left\{Var(\eta\,|\eta\in[\mu,\mu+\epsilon])+(\epsilon-\delta)^2\right\} + \\ & F_{\mu-\lambda}\left\{Var(\eta\,|\eta\in[\mu+\lambda,\infty])+(\lambda-\nu)^2\right\} \end{aligned}$$

since $\int dF_{\varphi|\eta} = 1$ and $F_{\mu-\lambda}\int_{\mu+\lambda}^{\infty}(\mu+\lambda-\eta)^2 dG_{\lambda} =$
$F_{\mu-\lambda}\left\{\int_{\mu+\lambda}^{\infty}\eta^2-(\mu+\nu)^2 dG_{\lambda}+(\lambda-\nu)^2\right\}$, likewise for the term involving ϵ.

Step 4: $2\int\int_{\mu+\lambda}^{\infty}(\mu+\lambda-\eta)(\eta-\varphi)dF_{\varphi|\eta}dF_{\eta}$

$$\begin{aligned} &= 2F_{\mu-\lambda}\left\{\int_{\mu+\lambda}^{\infty}(\mu+\lambda-\eta)\eta-(\mu+\lambda-\eta)\left[(1-\rho)\mu+\rho\eta\right]dG_{\lambda}\right\} \\ &= -2F_{\mu-\lambda}\left\langle(1-\rho)\left\{Var(\eta\,|\eta\geq\mu+\lambda)+\nu^2-\lambda\nu\right\}\right\rangle \end{aligned}$$

by the fact that $\int\varphi dF_{\varphi|\eta} = (1-\rho)\mu+\rho\eta$. Proceed likewise for the term $2\int\int_{\mu}^{\mu+\epsilon}(\mu+\epsilon-\eta)(\eta-\varphi)dF_{\varphi|\eta}dF_{\eta}$ and add all up to get (2.21). ■

Lemma 2.3 $E\pi'_{\epsilon,\lambda}(\rho)$ *is a single-peaked and quasiconcave function of* λ. *For* $1 > \rho > 0$, *it has an interior maximum, say at* $\lambda = \bar{\lambda}(\rho)$ *is concave for* $\lambda \leq \tilde{\lambda}(\rho)$ and convex for $\lambda > \tilde{\lambda}(\rho)$.

Proof of Lemma 2.3. Straightforward differentiation of (2.21) shows that the derivative of $E\pi'_{\epsilon,\lambda}(\rho)$ with respect to λ equals:

$$\frac{\partial E\pi'_{\epsilon,\lambda}(\rho)}{\partial \lambda} = 2F_{\mu-\lambda}(\rho\nu - \lambda)$$

We want to show that for $\rho < 1$, there exists a $\bar{\lambda}$, such that $\frac{\partial E\pi'_{\epsilon,\lambda}(\rho)}{\partial \lambda} < 0; \forall \lambda \geq \bar{\lambda}$. This can only be the case if there exists a λ, such that $\rho\nu < \lambda$; $\forall \lambda \geq \lambda$. Consider the equation $\rho\nu = \lambda$ or

$$\rho\sigma \frac{\phi(\frac{\lambda}{\sigma})}{1 - \Phi(\frac{\lambda}{\sigma})} = \lambda$$

where ϕ and Φ are the p.d.f. and c.d.f., respectively, of the standard normal and we make use of the fact that $\mu + \nu$ is the first moment of a truncated normal distribution.[24] At $\lambda = 0$ the right hand side of the equation is zero obviously while the left hand side is strictly positive. The right hand side is increasing in λ with slope 1. The derivative of the left hand side is

$$\begin{aligned} & -\rho \frac{\frac{\lambda}{\sigma}\phi(\frac{\lambda}{\sigma})}{1 - \Phi(\frac{\lambda}{\sigma})} + \frac{\rho\phi(\frac{\lambda}{\sigma})^2}{(1 - \Phi(\frac{\lambda}{\sigma}))^2} \\ = & \frac{f_{\mu+\lambda}}{(1 - F_{\mu+\lambda})}\rho(\nu - \lambda) \end{aligned}$$

thus increasing in λ. Start with the case $\rho = 1$: for $\rho = 1$ the equation cannot have a solution because $\nu \geq \lambda, \forall \lambda$. We can also state that $\nu' \leq 1, \forall \lambda$. To see this consider the

[24]see Johnson and Kotz (1970) p. 81

change in the slope of the left hand side, i.e. ν'' :

$$\begin{aligned}\frac{\partial}{\partial\lambda}\left\{\frac{f}{1-F}\rho(\nu-\lambda)\right\} &= \rho\frac{\partial}{\partial\lambda}\left\{\frac{f}{1-F}\right\}(\nu-\lambda)+\rho\frac{f}{1-F}\times \\ &\quad \left\{\frac{f}{1-F}(\nu-\lambda)-1\right\}\end{aligned}$$

It is well known that $\frac{\partial}{\partial\lambda}\left\{\frac{f}{1-F}\right\} > 0$ for a normal distribution. Thus if $\frac{f}{1-F}(\nu-\lambda) > 1$ the slope of $\rho\nu$ will increase more and more. Therefore $\nu'(\lambda) > 1$, for any finite $\check{\lambda}$, implies that $\nu'(\lambda) > 1$ for any $\lambda \geq \check{\lambda}$. But eventually this implies $\lim_{\lambda\to\infty}(\nu-\lambda) > 0$.

Now

$$\nu-\lambda = \sigma\frac{\phi(\frac{\lambda}{\sigma})}{1-\Phi(\frac{\lambda}{\sigma})}-\lambda$$

Using l'Hôpital twice we see that

$$\lim_{\lambda\to\infty}\frac{\sigma\phi(\frac{\lambda}{\sigma})}{\lambda(1-\Phi(\frac{\lambda}{\sigma}))} = 1 \tag{2.23}$$

and therefore

$$\lim_{\lambda\to\infty}\nu-\lambda = 0$$

(2.23) says $\lim_{\lambda\to\infty}\frac{\nu}{\lambda} = 1$. But then also $\lim_{\lambda\to\infty}\nu' = 1$, again by l'Hôpital. This establishes that $\nu'(\lambda) \leq 1, \forall\lambda$.

Decrease now ρ from 1. This has two effects: it shifts the function $\rho\nu$ down and it decreases its slope to $\rho\nu'$. Since $\nu' \leq 1, \forall\lambda$; $\rho\nu' < 1$ for $\rho < 1$; $\forall\lambda$. Thus, the equation has exactly one solution, $\lambda(\rho)$, for $\rho < 1$ and $\rho\nu-\lambda$ is negative for all $\lambda \geq \bar{\lambda}(\rho)$. This finally establishes that $E\pi'_{\epsilon,\lambda}(\rho)$ is a singlepeaked, quasiconcave function. Consider now $\frac{\partial^2\pi'_{\epsilon,\lambda}(\rho)}{\partial\lambda^2} = 2\left((1-\rho)f_{\mu+\lambda}\lambda - F_{\mu-\lambda}\right)$. This is positive if $\frac{f_{\mu+\lambda}}{1-F_{\mu+\lambda}} > \frac{1}{(1-\rho)\lambda}$. Since the left hand side is increasing in λ, the right hand side is decreasing in λ, we will -if $\rho < 1$- always find a value $\tilde{\lambda}$ such that $E\pi'_{\epsilon,\lambda}(\rho)$ is convex for all $\lambda > \tilde{\lambda}$. This proves the claim. ■

Proof of Proposition 2.4. For the sake of clarity, we first state all derivatives. The results follow from straightforward differentiation of (2.21) and

$$e = h\left[A\left(\frac{(\sigma^2+\epsilon^2)}{2} - \int_{\mu}^{\mu+\epsilon}(\mu+\epsilon-\eta)^2 - \int_{\mu+\lambda}^{\infty}(\mu+\lambda-\eta)^2\right)\right]$$

Derivatives with respect to ϵ :

$$\begin{aligned}
\frac{\partial}{\partial\epsilon}E\pi'_{\epsilon,\lambda}(\rho) &= 2\Delta F_\epsilon(\rho\delta-\epsilon)\\
\frac{\partial^2}{\partial\epsilon^2}E\pi'_{\epsilon,\lambda}(\rho) &= -2\Delta F_\epsilon - 2(1-\rho)f_{\mu+\epsilon}\epsilon\\
\frac{\partial e}{\partial\epsilon} &= \frac{A\left[\epsilon-2\Delta F_\epsilon\left\{\epsilon-\delta\right\}\right]}{g''(e)}\\
\frac{\partial^2 e}{\partial\epsilon^2} &= \frac{A(1-2\Delta F_\epsilon)}{g''(e)} - \frac{A^2\left[\epsilon-2\Delta F_\epsilon\left\{\epsilon-\delta\right\}\right]^2}{g''(e)}\frac{g'''(e)}{g''(e)^2}
\end{aligned}$$

Derivatives with respect to λ :

$$\begin{aligned}
\frac{\partial E\pi'_{\epsilon,\lambda}(\rho)}{\partial\lambda} &= 2F_{\mu-\lambda}(\rho\nu-\lambda)\\
\frac{\partial^2}{\partial\lambda^2}E\pi'_{\epsilon,\lambda}(\rho) &= 2\left((1-\rho)f_{\mu+\lambda}\lambda - F_{\mu-\lambda}\right)\\
\frac{\partial e}{\partial\lambda} &= \frac{2AF_{\mu-\lambda}(\nu-\lambda)}{g''(e)}\\
\frac{\partial^2 e}{\partial\lambda^2} &= -\frac{g'''}{g''}\left[\frac{2A}{g''}F_{\mu-\lambda}(\nu-\lambda)\right]^2 - \frac{2A}{g''}F_{\mu-\lambda}
\end{aligned}$$

The first order conditions are given by FOC_ϵ :

$$e(\epsilon-2\Delta F_\epsilon(\epsilon-\rho\delta)) + \frac{A\left[\epsilon-2\Delta F_\epsilon\left\{\epsilon-\delta\right\}\right]\left(E(\pi'_{\epsilon,\lambda}(\rho)-\pi''_{\epsilon,\lambda})\right)}{g''(e)} = \epsilon$$

and FOC_λ :

$$e2F_{\mu-\lambda}(\rho\nu-\lambda) + \frac{2AF_{\mu-\lambda}(\nu-\lambda)\left[E(\pi'_{\epsilon,\lambda}(\rho)-\pi''_{\epsilon,\lambda})\right]}{g''(e)} = 0$$

(ia) Existence of a solution follows from Weierstrass's Theorem. Before we prove uniqueness a.e. (since it uses the same arguments) let us first prove (ii). The change in the principal's payoff due to a change in ρ is given by

$$\begin{aligned}
&\left(\frac{\partial}{\partial\epsilon}E\pi''_{\epsilon,\lambda}+e\frac{\partial}{\partial\epsilon}\left[E(\pi'_{\epsilon,\lambda}(\rho)-\pi''_{\epsilon,\lambda})\right]+\frac{\partial e}{\partial\epsilon}E(\pi'_{\epsilon,\lambda}(\rho)-\pi''_{\epsilon,\lambda})\right)\frac{\partial\epsilon}{\partial\rho}+\\
&\left(e\frac{\partial}{\partial\lambda}E\pi'_{\epsilon,\lambda}(\rho)+\frac{\partial e}{\partial\lambda}E(\pi'_{\epsilon,\lambda}(\rho)-\pi''_{\epsilon,\lambda})\right)\frac{\partial\lambda}{\partial\rho}+\\
&e\frac{\partial}{\partial\rho}\left[E(\pi'_{\epsilon,\lambda}(\rho)-\pi''_{\epsilon,\lambda})\right]
\end{aligned}$$

The first line is always equal to zero by the envelope theorem and the fact that FOC_ϵ always hold with equality whether $\epsilon=0$ or positive. For $0<\lambda<\infty$ the second line is zero again by the envelope theorem. By (iv) below, $\lambda^*=0$ for $\rho\leq 0$. So $\frac{\partial\lambda}{\partial\rho}=0$ $\rho<0$. At $\lambda=\infty$ $e\frac{\partial}{\partial\lambda}E\pi'_{\epsilon,\lambda}(\rho)+\frac{\partial e}{\partial\lambda}E(\pi'_{\epsilon,\lambda}(\rho)-\pi''_{\epsilon,\lambda})\geq 0$, so the second line is nonnegative if $\frac{\partial\lambda}{\partial\rho}\geq 0$. This is established in (vi) below. It remains to be shown that $\frac{\partial}{\partial\rho}\left[E(\pi'_{\epsilon,\lambda}(\rho)-\pi''_{\epsilon,\lambda})\right]>0$; $\forall\lambda,\epsilon\geq 0;\lambda\geq\epsilon$. $\frac{\partial}{\partial\rho}\left[E(\pi'_{\epsilon,\lambda}(\rho)-\pi''_{\epsilon,\lambda})\right]=$

$$\sigma^2-2F_{\mu-\lambda}Var_\lambda-2\Delta F_\epsilon Var_\epsilon-2F_{\mu-\lambda}\nu(\nu-\lambda)+2\Delta F_\epsilon\delta(\epsilon-\delta)$$

As $\frac{\partial^2}{\partial\epsilon\partial\rho}\left[E(\pi'_{\epsilon,\lambda}(\rho)-\pi''_{\epsilon,\lambda})\right]=2\Delta F_\epsilon\delta>0$ and $\frac{\partial^2}{\partial\lambda\partial\rho}\left[E(\pi'_{\epsilon,\lambda}(\rho)-\pi''_{\epsilon,\lambda})\right]=2(1-F_{\mu+\lambda})v>0$, the expression is minimized at $\epsilon=\lambda=0$: $\frac{\partial}{\partial\rho}^{\min}\left[E(\pi'_{\epsilon,\lambda}(\rho)-\pi''_{\epsilon,\lambda})\right]=$

$$\begin{aligned}
&\sigma^2-\int_\mu^\infty\eta^2-(\mu+\nu)^2\frac{dF_\eta}{\frac{1}{2}}-\nu^2\\
=\ &\sigma^2-\left(\sigma^2-\nu^2\right)-\nu^2=0.
\end{aligned}$$

This proves the claim.

(ib) Consider now uniqueness. First in the case of ϵ. For ρ sufficiently close to 1 the same arguments as above continue to hold and there must be a unique optimal ϵ. Moreover, it will be shown below that $\epsilon>0$ if and only if ρ is close to 1. Consider then uniqueness in the case of λ : A sufficient condition for uniqueness is that the problem be globally concave. However, this need not be the case. Consider the second order condition:

$$\frac{\partial^2 e}{\partial \lambda^2} E(\pi'_{\epsilon,\lambda}(\rho) - \pi''_{\epsilon,\lambda}) + 2\frac{\partial e}{\partial \lambda}\frac{\partial E\pi'_{\epsilon,\lambda}(\rho)}{\partial \lambda} + e\frac{\partial^2 E\pi'_{\epsilon,\lambda}(\rho)}{\partial \lambda^2} < 0 \tag{2.24}$$

From Lemma 2.3, we know that $\frac{\partial E\pi'_{\epsilon,\lambda}(\rho)}{\partial \lambda} < 0$ for all $\lambda > \bar{\lambda}(\rho)$. Moreover, we will show below, that λ^* must necessarily be larger than $\bar{\lambda}(\rho)$. Hence, focussing attention on the relevant range, the first and the second term in (2.24) are negative as $\frac{\partial E\pi'_{\epsilon,\lambda}(\rho)}{\partial \lambda} < 0$, $\frac{\partial e}{\partial \lambda} > 0$, and

$$\frac{\partial^2 e}{\partial \lambda^2} = -\frac{g'''}{g''}\left[\frac{2A}{g''}F_{\mu-\lambda}(\nu-\lambda)\right]^2 - \frac{2A}{g''}F_{\mu-\lambda} < 0$$

However, if $\rho < 1$, $\frac{\partial^2 E\pi'}{\partial \lambda^2} = 2\left((1-\rho)f_{\mu+\lambda}\lambda - F_{\mu-\lambda}\right)$ is larger than zero for λ sufficiently large. It is thus possible - and we have found no way to exclude this possibility in general - that there are several interior solutions. However, it is shown below that this is not an issue for often used effort cost functions. (see below). Assume nevertheless for the sake of the argument that there are several local extrema. In this case the principal's payoff must be evaluated at all these values of λ satisfying the first and second order condition.

In this case we can only show uniqueness a.e.: let $\lambda_1^*(\rho)$ and $\lambda_2^*(\rho)$ be two values of λ which both maximize the principal's payoff with $\lambda_1^*(\rho) < \lambda_2^*(\rho)$. This cannot be the case except at a finite number of disconnected points in $[0, 1]$, i.e. values of ρ. To see this, consider the change in the principal's payoff for any given value of λ or at an interior solution for λ as ρ increases. By the argument in (ii) this is equal to

$$e\frac{\partial E(\pi'_{\epsilon,\lambda}(\rho) - \pi''_{\epsilon,\lambda})}{\partial \rho} > 0$$

By the fact that $\frac{\partial^2 E(\pi'_{\epsilon,\lambda}(\rho) - \pi''_{\epsilon,\lambda})}{\partial \rho \partial \lambda} > 0$ (see above) the change in expected payoff for the principal is always higher at higher values of λ. Let now $\lambda_1'(\rho + d\rho)$ and $\lambda_2'(\rho + d\rho)$ be the local maximizers for $\rho + d\rho$. By the above argument λ_2' must give the principal a strictly higher payoff than λ_1'. This proves the claim.

(iii)The first order condition in this case is given by

$$-\epsilon + e\left\{\epsilon - 2\Delta F_\epsilon(\epsilon - \rho\delta)\right\} \qquad (2.25)$$
$$+\frac{\partial e}{\partial \epsilon} \times E(\pi'_{\epsilon,\lambda} - \pi''_{\epsilon,\lambda}) \stackrel{!}{=} 0$$

while the second order condition is

$$-1 + e\left(1 - 2\Delta F_\epsilon - 2(1-\rho) f_{\mu+\epsilon}\epsilon\right) \qquad (2.26)$$
$$+2\frac{\partial e}{\partial \epsilon}\left\{\epsilon - 2\Delta F_\epsilon(\epsilon - \rho\delta)\right\}$$
$$+\frac{\partial^2 e}{\partial \epsilon^2} E(\pi'_{\epsilon,\lambda} - \pi''_{\epsilon,\lambda})$$

< 0. Define

$$H(\epsilon, \epsilon(\rho)) := \frac{\partial \pi''}{\partial \epsilon} + e\frac{\partial}{\partial \epsilon}\left[E(\pi'_{\epsilon,\lambda}(\rho) - \pi''_{\epsilon,\lambda})\right] + \frac{\partial e}{\partial \epsilon} E(\pi'_{\epsilon,\lambda}(\rho) - \pi''_{\epsilon,\lambda}).$$

By the implicit function theorem $\frac{\partial \epsilon}{\partial \rho} = \frac{-\frac{\partial H(\epsilon)}{\partial \rho}}{\frac{\partial H(\epsilon)}{\partial \epsilon}}$. The denominator is the second order condition. It was proved that the solution is a maximum for $\rho = 1$. By continuity this is also true for ρ close to 1. Thus, the denominator is negative. Therefore $sign\left[\frac{\partial \epsilon}{\partial \rho}\right] = sign\left[\frac{\partial H(\epsilon)}{\partial \rho}\right]$.

$$\frac{\partial H(\epsilon)}{\partial \rho} = e\frac{\partial^2}{\partial \epsilon \partial \rho}\left[E(\pi'_{\epsilon,\lambda}(\rho) - \pi''_{\epsilon,\lambda})\right] + \frac{\partial e}{\partial \epsilon}\frac{\partial}{\partial \rho}\left[E(\pi'_{\epsilon,\lambda}(\rho) - \pi''_{\epsilon,\lambda})\right].$$

$$\frac{\partial^2}{\partial \epsilon \partial \rho}\left[E(\pi'_{\epsilon,\lambda}(\rho) - \pi''_{\epsilon,\lambda})\right] = e2\Delta F_\epsilon \delta > 0.$$

$\frac{\partial}{\partial \rho}\left[E(\pi'_{\epsilon,\lambda}(\rho) - \pi''_{\epsilon,\lambda})\right] \geq 0; \forall \lambda, \epsilon \geq 0; \lambda \geq \epsilon$ by (ii) above. This proves the claim.

(iv)To make delegation worthwhile at all, the principal must benefit if his agent is well informed. This requires that $E\left(\pi'_{\epsilon,\lambda}(\rho) - \pi''_{\epsilon,\lambda}\right) \geq 0$. But from lemma 2.3, $\frac{\partial E\pi'_{\epsilon,\lambda}(\rho)}{\partial \lambda} \leq 0; \forall \lambda$ $\rho \leq 0$. Hence $\lambda^* = 0$.

(v)λ^* satisfies

$$e\frac{\partial E\pi'_{\epsilon,\lambda}(\rho)}{\partial\lambda}+\frac{\partial e}{\partial\lambda}E\left(\pi'_{\epsilon,\lambda}(\rho)-\pi''_{\epsilon,\lambda}\right)\stackrel{!}{=}0.$$

The agent reacts positively on increased freedom of action ($\frac{\partial e}{\partial\lambda}>0$). If a solution exists, we must therefore necessarily have $\frac{\partial E\pi'_{\epsilon,\lambda}(\rho)}{\partial\lambda}<0$, (since $E\left(\pi'_{\epsilon,\lambda}(\rho)-\pi''_{\epsilon,\lambda}\right)\geq 0$, too). By Lemma 2.3, $E\pi'_{\epsilon,\lambda}(\rho)$ is a quasiconcave function of λ. This implies, for $0<\rho<1$, together with the arguments above that $\lambda^*>\bar{\lambda}$.

λ^* is finite for $\rho<1:e$ and $E(\pi'_{\epsilon,\lambda}(\rho)-\pi''_{\epsilon,\lambda})$ both attain fixed values for $\lambda\to\infty$. Consider $\lim_{\lambda\to\infty}(\frac{\frac{\partial E\pi'_{\epsilon,\lambda}(\rho)}{\partial\lambda}}{\frac{\partial e}{\partial\lambda}})$:

$$\lim_{\lambda\to\infty}\frac{2(1-F_{\mu+\lambda})(\rho\nu-\lambda)}{\frac{2A}{g(e)''}(1-F_{\mu+\lambda})(\nu-\lambda)}=-\infty$$

Thus, the increase in $E\pi'_{\epsilon,\lambda}(\rho)$ overcompensates the decrease in e. Essentially,for $\rho<1$, the principal can, by decreasing λ, increase $E\pi'_{\epsilon,\lambda}(\rho)$ a bit without reducing effort supply a lot. This proves the claims.

(vi) To see that $\frac{d\lambda}{d\rho}>0$, consider first local changes. Define the function

$$H(\rho,\lambda(\rho)):=e\frac{\partial E\pi'_{\epsilon,\lambda}(\rho)}{\partial\lambda}+\frac{\partial e}{\partial\lambda}E(\pi'_{\epsilon,\lambda}(\rho)-\pi''_{\epsilon,\lambda}).$$

By the familiar argument: $\frac{\partial\lambda}{\partial\rho}=-\frac{\frac{\partial H}{\partial\rho}}{\frac{\partial H}{\partial\lambda}}$. Because the second order condition holds locally at a maximum, the denominator is negative and $sign\left[\frac{\partial\lambda}{\partial\rho}\right]=sign\left[\frac{\partial H}{\partial\rho}\right]$.

$$\frac{\partial H}{\partial\rho}=\left[\frac{\partial e}{\partial\lambda}\frac{\partial E\pi'_{\epsilon,\lambda}(\rho)}{\partial\rho}+e\frac{\partial^2 E\pi'_{\epsilon,\lambda}(\rho)}{\partial\lambda\partial\rho}\right]$$

$\frac{\partial E\pi'_{\epsilon,\lambda}(\rho)}{\partial\rho}>0$ (all $\epsilon,\lambda\geq 0;\lambda\geq\epsilon$) by the argument in (ii). There it has also been established that $\frac{\partial^2 E\pi'_{\epsilon,\lambda}(\rho)}{\partial\lambda\partial\rho}=2(1-F_{\mu+\lambda})\nu>0$. Hence $\frac{\partial\lambda^*}{\partial\rho}>0$.

Assume again that the principal's problem has two interior solutions. Even in this case we must have $\frac{d\lambda}{d\rho}\geq 0$ by the same argument as above, because $\frac{\partial^2 E(\pi'_{\epsilon,\lambda}(\rho)-\pi''_{\epsilon,\lambda})}{\partial\rho\partial\lambda}>0$ any discrete jumps will be rightwards. ■

An Example to Proposition 2.4: The problem has only one local extremum in the relevant range if we restrict attention to the family of effort cost functions $g = c\frac{e^{l+1}}{l+1}; l+1 \geq 3$.[25]

In this case we have

$$e = \left[\frac{A}{c}\left\{\frac{\sigma^2}{2} - (1 - F_{\mu+\lambda})(Var_\lambda + (\nu - \lambda)^2)\right\}\right]^{\frac{1}{l}}$$

$$\frac{\partial e}{\partial \lambda} = \frac{1}{l}e^{\frac{1}{l}-1}\frac{2A}{c}(1 - F_{\mu+\lambda})(\nu - \lambda)$$

The first order condition can then be expressed as

$$z(\lambda) \stackrel{!}{=} -\frac{(\nu - \lambda)}{(\rho\nu - \lambda)} \tag{2.27}$$

where

$$z(\lambda) = (l)\frac{\frac{\sigma^2}{2} - (1 - F_{\mu+\lambda})(Var_\lambda + (\nu - \lambda)^2)}{E(\pi'_{\epsilon,\lambda}(\rho) - \pi''_{\epsilon,\lambda})}$$

The solution to the problem must satisfy $\frac{\partial E\pi'_{\epsilon,\lambda}(\rho)}{\partial \lambda} < 0$ and $E(\pi'_{\epsilon,\lambda}(\rho) - \pi''_{\epsilon,\lambda}) \geq 0$ thus both the left hand side and the right hand side of (2.27) are nonnegative. Let $\bar{\lambda}$ solve $\rho\nu - \lambda = 0$ and $\hat{\lambda} := \min[\lambda^{\max}, \grave{\lambda}]$, where $\grave{\lambda}$ solves $E(\pi'_{\epsilon,\lambda}(\rho) - \pi''_{\epsilon,\lambda}) = 0$. Now for $\rho > 0$, $\grave{\lambda} > \bar{\lambda}$. Next observe that $\lim_{\lambda\to\bar{\lambda}} -\frac{(\nu-\lambda)}{(\rho\nu-\lambda)} = \infty$, while $\infty > z(\bar{\lambda}) > 0$; $\lim_{\lambda\to\hat{\lambda}} z(\lambda) = \infty$ iff $\hat{\lambda} = \grave{\lambda}$, while $\infty > -\frac{(\nu(\grave{\lambda})-\grave{\lambda})}{(\rho\nu(\grave{\lambda})-\grave{\lambda})} > 0$. As $\frac{(\nu-\lambda)(\rho\nu'-1)-(\nu'-1)(\rho\nu-\lambda)}{(\rho\nu-\lambda)^2} < 0$ and $\frac{\partial z}{\partial \lambda} = (l)\frac{\frac{1}{A}\frac{\partial(U'-U'')}{\partial\lambda}E(\pi'_{\epsilon,\lambda}-\pi''_{\epsilon,\lambda})-\frac{1}{A}(U'-U'')\frac{\partial E(\pi'_{\epsilon,\lambda}(\rho)-\pi''_{\epsilon,\lambda})}{\partial\lambda}}{(E(\pi'_{\epsilon,\lambda}(\rho)-\pi''_{\epsilon,\lambda}))^2} > 0$ (2.27) has exactly one solution. If $\hat{\lambda} = \lambda^{\max}$ then we must let λ go to infinity: $\lim_{\lambda\to\infty} z(\lambda) = \frac{l}{2\rho-1}$, while $\lim_{\lambda\to\infty} \frac{\nu-\lambda}{\lambda-\rho\nu} = 0$. Observe that $\frac{l}{2\rho-1} > 0$ if $\rho > \frac{1}{2}$. But then there always exists exactly one solution $\lambda < \lambda^{\max}$ that solves (2.27) as long as $\rho < 1$. In particular this also shows that $\lambda = \lambda^{\max}$ if and only

[25] Instead of the Inada condition $\lim_{e\to 1} g' = \infty$ we ensure $e^* < 1$ by choice of c.

if $\rho = 1$. This example shows that everything works well, i.e. the product $eE(\pi'_{\epsilon,\lambda}(\rho) - \pi''_{\epsilon,\lambda})$ can achieve only one extremum in the relevant range under quite mild assumptions about effort costs.

2.8 Appendix B

We first show that it can't be optimal to implement first best optimal decisions if $e_f^{sb} + \frac{A\sigma^2}{g''(e_f^{sb})2}(1-A) > 1$. Thereafter we study the properties of the optimal contract as described in the text. We search for the optimal contract by solving for the optimal, incentive compatible direct revelation mechanism. Thus the agent reports his information and the principal implements a decision contingent on this information. Messages from the agent to the principal take the form $m \in m_\emptyset \cup \text{supp}_\eta$. I.e. the agent may report a specific realization of $\tilde{\eta}$ or that he did not learn anything ($m_\emptyset$). With a slight abuse of notation we will henceforth write $\emptyset$ for $m_\emptyset$. The principal then implements for each η an action $x(\eta)$ and a transfer $\tau(\eta)$ for the case where the agent is informed and for the case where the agent is ignorant an action payment pair $x(\emptyset), \tau(\emptyset)$.

Step 1: If $e_f^{sb} + \frac{A\sigma^2}{g''(e_f^{sb})2}(1-A) > 1$, implementing $x = \eta$ cannot be part of an optimal policy.

The proof to this statement uses essentially the same logic as the proof to proposition 2.2. As a first step, note that under any contract that implements first best project choices, the agent will supply a level of effort satisfying $g'(e) = \frac{A\sigma^2}{2}$. To see this, note that the agent chooses his report in order to

$$\max_{\hat{\eta}} -\frac{A}{2}(x(\hat{\eta}) - \eta)^2 + \tau(\hat{\eta}).$$

But then $x(\hat{\eta}) = \eta$ implies $\frac{\partial \tau}{\partial \eta} = 0$. It is also easy to see that the transfer function cannot have any steps, because this would violate truthtelling over some range. But then, the transfer function cannot be used at all to increase the agent's effort choice when first best project choices shall be implemented.

Consider next the following (suboptimal) contract: if the agent reports $\eta \in (\mu - \epsilon', \mu)$ the principal implements $x = \mu - \epsilon'$. If the report is $\eta \in [\mu, \mu + \epsilon')$ or $\emptyset$ the principal implements $x = \mu + \epsilon'$. For any other report the principal implements $x = \eta$. The agent receives a constant wage τ' independent of the report. By the argument above the agent reports his knowledge truthfully.

Next we compare the performance of these two contracts. The second contract is clearly suboptimal because the principal does not use all his contracting possibilities. But the logic employed is this: if we can find a suboptimal contract that dominates any contract that implements first best optimal project decisions then an optimal contract *must* involve a departure from first best project decisions. This is already sufficient to argue that it will not be optimal for the principal to implement first best project choices even if there are no conflicts irrespective of whether the principal can use monetary contracts or not.[26]

If the principal sets $\epsilon' > 0$, the agent's effort choice increases by lemma 2.1. To compensate the higher effort costs the wage must be increased. Assume that the principal compensates the entire additional effort costs. (By proposition 2.2 $eEU' + (1 - e)EU''$ will be higher if $\epsilon' > 0$ than if $\epsilon = 0$ so this is actually more than necessary.) Setting $\epsilon' > 0$ will afford an increase in the transfer according to

$$\frac{\partial \tau}{\partial \epsilon} = \frac{\partial g}{\partial e} \frac{\partial e}{\partial \epsilon}.$$

By the same logic as in the proof to proposition 2.2 (see appendix A), the principal will benefit from setting $\epsilon > 0$ if

$$e_f^{sb} + \frac{A\sigma^2}{g''(e_f^{sb})2} - g'(e_f^{sb}) \frac{A}{g''(e_f^{sb})} > 1$$

[26] This is not the same as to argue that social surplus increases when the agent supplies more effort. It might be too costly for the principal to induce a higher effort choice by the agent by means of a second best optimal contract.

or since $g'(e_f^{sb}) = \frac{A\sigma^2}{2}$

$$e_f^{sb} + \frac{A\sigma^2}{g''(e_f^{sb})2}(1-A) > 1$$

This condition now compares both direct and indirect costs: with a probability $1-e$ the principal just shoots himself in the foot by setting $\epsilon > 0$. In addition he now bears additional wage costs of $\frac{A^2\sigma^2}{g(e)''2}$. The benefit is as before, the marginally avoided risk: $\frac{A\sigma^2}{g(e)''2}$.[27]

The optimal choice of contracts: statement of the problem.

We consider the direct mechanism in which the agent is asked to reveal his type and is given incentives to do so truthfully. The principal's problem is to

$$\max_{x(\eta),\tau(\eta);x_\emptyset,\tau_\emptyset,e} K + (1-e)\left\{\int -\frac{1}{2}(x_\emptyset-\eta)^2 dF - \tau_\emptyset\right\} + e\left\{\int\left(-\frac{1}{2}(x(\eta)-\eta)^2 - \tau(\eta)\right)dF\right\}$$

st.

$$\int -\frac{A}{2}(x_\emptyset-\eta)^2 dF + \tau_\emptyset \geq \int -\frac{A}{2}(x(\hat\eta)-\eta)^2 dF + \tau(\hat\eta); \forall\hat\eta \tag{M.IC1}$$

$$-\frac{A}{2}(x(\eta)-\eta)^2 + \tau(\eta) \geq -\frac{A}{2}(x_\emptyset-\eta)^2 + \tau_\emptyset; \forall\eta \tag{M.IC2}$$

$$-\frac{A}{2}(x(\eta)-\eta)^2 + \tau(\eta) \geq -\frac{A}{2}(x(\hat\eta)-\eta)^2 + \tau(\hat\eta); \forall\eta,\hat\eta \tag{M.IC3}$$

[27] One may again wonder what this condition means in terms of the primitives of the model. The derivative of the left hand side of this inequality is

$$(1-A)\frac{\sigma^2}{g(e)''2}\left\{2 - \frac{A\sigma^2}{2}\frac{g(e)'''}{g(e)''^2}\right\} < 0.$$

Therefore the condition is easier to satisfy for small values of A.

and

$$\begin{aligned} 0 &\leq \int \left(-\frac{A}{2}(x(\eta)-\eta)^2 + \tau(\eta) \right) dF \\ &+ \int \frac{A}{2}(x_\emptyset - \eta)^2 dF - \tau_\emptyset - g'(e) \end{aligned} \tag{2.28}$$

$$\begin{aligned} 0 &\leq k + (1-e)\left(\int -\frac{A}{2}(x_\emptyset - \eta)^2 dF + \tau_\emptyset \right) \\ &+ e\left(\int \left(-\frac{A}{2}(x(\eta)-\eta)^2 + \tau(\eta) \right) dF \right) - g(e) \end{aligned} \tag{2.29}$$

$M.IC1$ states that the agent prefers to truthfully reveal his ignorance, $M.IC2$ requires that the informed agent shall prefer the contract intended for his type vis a vis the contract intended for the ignorant agent. $M.IC3$ requires that the agent truthfully reveals his type given that he knows it. In the second set of constraints (2.28) states that the agent's effort choice must be incentive compatible while the last constraint ensures that the agent is willing to participate. Note that we require individual rationality to hold ex ante only. That is some types may have negative utility ex post.

A simplification of the problem: the maximization problem is equivalent to the following problem:

$$\begin{aligned} \max_{x(\eta),t(\eta);x_\emptyset,t_\emptyset,e} \; & K + (1-e)\left\{ \int_{-\infty}^{\infty} -\frac{A+1}{2}(x_\emptyset - \eta)^2 dF - Eu_\mu \right\} + \\ e & \left\{ \int_{\mu}^{\infty} 2\left(-\frac{A+1}{2}(x(\eta)-\eta)^2 - u_\mu - \int_\mu^\eta A\left(x(\eta)-\eta\right) d\eta \right) dF \right\} \end{aligned}$$

s.t.

$$x_\emptyset = x(\mu) \tag{IC Ignorance}$$

$$\frac{\partial x}{\partial \eta} \geq 0 \tag{Monotonicity}$$

$$\frac{A\sigma^2}{2} + 2\int_{\mu}^{\infty}\int_{\mu}^{\eta} A\left(x(\eta)-\eta\right) d\eta dF_{\eta} - g'(e) \geq 0 \tag{M.ICe'}$$

$$\begin{aligned} 0 \quad \leq \quad & k + e\left(u_{\mu} + \int_{\mu}^{\infty}\int_{\mu}^{\eta} A\left(x(\eta)-\eta\right) d\eta dF_{\eta}\right) \\ & +(1-e)Eu_{\mu} - g(e) \end{aligned} \tag{M.IR'}$$

Proof of equivalence of the problems:

claim (ia): "$M.IC1$ and $M.IC2$" $\Rightarrow x_{\emptyset} = x(\mu)$:

Taking $M.IC1$ and $M.IC2$ at $\eta = \mu$ and rearranging the constraints yields

$$\int \frac{A}{2}(x(\mu)-\eta)^2 dF - \int \frac{A}{2}(x_{\emptyset}-\eta)^2 dF \geq \tau(\mu) - \tau_{\emptyset}$$

and

$$\tau(\mu) - \tau_{\emptyset} \geq \frac{A}{2}(x(\mu)-\mu)^2 - \frac{A}{2}(x_{\emptyset}-\mu)^2$$

Both inequalities can only hold simultaneously if $\tau(\mu) = \tau_{\emptyset}$. But then we must also have $x(\mu) = x_{\emptyset}$. Hence, $M.IC1$ and $M.IC2$ imply $x(\mu) = x_{\emptyset}$. claim (ib): "$M.IC3 \Rightarrow \frac{\partial x}{\partial \eta} \geq 0$."

$M.IC3$ taken locally:

$$\max_{\hat{\eta}} -\frac{A}{2}\left(x(\hat{\eta})-\eta\right)^2 + \tau(\hat{\eta})$$

gives, assuming that $x()$ and $\tau()$ are differentiable a.e. that

$$A\left(x(\eta)-\eta\right)\frac{\partial x}{\partial \eta} = \frac{\partial \tau}{\partial \eta} \text{ a.e.} \tag{2.30}$$

First totally differentiating (2.30) with respect to $\eta, \hat{\eta}$ and second deriving the agent's second order condition, one concludes that only decision rules satisfying $\frac{\partial x}{\partial \eta} \geq 0$ are implementable.

Next we claim that the reverse reasoning is also true. We first eliminate transfers in order to work with the indirect utility function, (u) and the decision function $x(\eta)$ only. (By symmetry we need only consider the upper half of the support of η.) The agent with the lowest type, i.e. the one with the worst type, is agent μ. By the envelope theorem and incentive compatibility, $A\left(x(\eta)-\eta\right)\frac{\partial x}{\partial \eta}=\frac{\partial \tau}{\partial \eta}$, the indirect utility of an agent of type η is given by[28]

$$u_\eta = u_\mu + \int_\mu^\eta A\left(x(t)-t\right)dt. \tag{2.31}$$

Now, we argue that (2.31) together with the constraints $x_\emptyset = x(\mu)$ and $\frac{\partial x}{\partial \eta} \geq 0$ imply $M.IC1-3$. Rewrite the agent's utility as

$$-\frac{A}{2}(x(\eta)-\eta)^2 + \tau(\eta) = u_\mu + \int_\mu^\eta A\left(x(t)-t\right)dt. \tag{2.32}$$

The agent's utility when lying, $-\frac{A}{2}(x(\hat{\eta})-\eta)^2+\tau(\hat{\eta})$ can be expressed as

$$-\frac{A}{2}(x(\hat{\eta})-\hat{\eta})^2 + \tau(\hat{\eta}) - \frac{A}{2}(\hat{\eta}-\eta)^2 - A(\hat{\eta}-\eta)(x(\hat{\eta})-\hat{\eta})$$

or

$$u_\mu + \int_\mu^{\hat{\eta}} A\left(x(t)-t\right)dt - \frac{A}{2}(\hat{\eta}-\eta)^2 - A(\hat{\eta}-\eta)(x(\hat{\eta})-\hat{\eta}).$$

Then, letting $\hat{\eta} < \eta$

$$\begin{aligned}
& -\frac{A}{2}(x(\eta)-\eta)^2 + \tau(\eta) - \left\{-\frac{A}{2}(x(\hat{\eta})-\eta)^2 + \tau(\hat{\eta})\right\} \\
= & \int_{\hat{\eta}}^{\eta} A\left(x(t)-t\right)dt + \frac{A}{2}(\hat{\eta}-\eta)^2 + A(\hat{\eta}-\eta)(x(\hat{\eta})-\hat{\eta}) \\
= & \int_{\hat{\eta}}^{\eta} A\left(x(t)-t\right)dt + \frac{A}{2}(\hat{\eta}-\eta)^2 - \int_{\hat{\eta}}^{\eta} A(x(\hat{\eta})-\hat{\eta})dt \\
= & \int_{\hat{\eta}}^{\eta} A\left(x(t)-x(\hat{\eta})\right)dt + \int_{\hat{\eta}}^{\eta} A(\hat{\eta}-t)dt + \frac{A}{2}(\hat{\eta}-\eta)^2 \geq 0.
\end{aligned}$$

[28]Note that types η and $-\eta$ get the same indirect utility.

Since

$$\int_{\hat{\eta}}^{\eta} A(\hat{\eta}-t)dt = -\frac{A}{2}(\eta-\hat{\eta})^2$$

and

$$\int_{\hat{\eta}}^{\eta} A\left(x(t)-x(\hat{\eta})\right)dt \geq 0$$

by $\frac{\partial x}{\partial \eta} \geq 0$. The same reasoning applies to the case $\hat{\eta} > \eta$. Last, note that once the constraint $x_{\emptyset} = x(\mu)$ is respected, $M.IC1$ and $M.IC2$ are nothing but special cases of $M.IC3$. This completes the argument. Hence the problem can be simplified in that one directly gives indirect utility to the agent and respects, instead of the constraints $M.IC1-3$, the constraints

$$x_{\emptyset} = x(\mu) \text{ and } \frac{\partial x}{\partial \eta} \geq 0. \tag{ICM}$$

Let us now consider these constraints in turn:

$$x_{\emptyset} = x(\mu) \tag{IC Ignorance}$$

$$\frac{\partial x}{\partial \eta} \geq 0 \tag{Monotonicity}$$

$$\frac{A\sigma^2}{2} + 2\int_{\mu}^{\infty}\int_{\mu}^{\eta} A\left(x(\eta)-\eta\right)d\eta dF_{\eta} - g'(e) \geq 0 \tag{M.ICe'}$$

$$\begin{aligned} 0 \leq\ & k + e\left(u_{\mu} + \int_{\mu}^{\infty}\int_{\mu}^{\eta} A\left(x(\eta)-\eta\right)d\eta dF_{\eta}\right) \\ & +(1-e)Eu_{\mu} - g(e) \end{aligned} \tag{M.IR'}$$

Concerning the monotonicity constraint, we follow the standard procedure and solve the problem in the first instance ignoring the constraint and check afterwards whether the so derived solution satisfies monotonicity. Consider then $(M.IR')$: because the transfer τ is unbounded below, u_μ can be lowered till the constraint holds with equality. Let ϑ denote the multiplier attached to this constraint.

Solution part I: a constrained subproblem:

Although the problem has been simplified a lot it is not yet straightforward to solve it. The peculiarities of the problem are reflected in the constraints $(IcIgnorance)$ and $(M.ICe')$ and the fact that e is a choice variable. But as long as e and $x_\emptyset$ are held constant the problem corresponds to a standard isoperimetric problem and can be solved by pointwise maximization techniques. So, as a first step, we simply take as given some value of $x_\emptyset$ and e and solve this constrained problem. Hence we ask in the first step: what is the cheapest way to implement a specific level of e, satisfying $g'(e) > \frac{A\sigma^2}{2}$, *if the principal is constrained to set a specific* $x_\emptyset$. In the subsequent step we solve for the optimal $x_\emptyset^*$.

Recall again that by $e_f^{sb} + \frac{A\sigma^2}{g''(e_f^{sb})2}(1-A) > 1$ the constraint $(M.ICe')$ is binding. If $(M.ICe')$ were not binding the solution would be trivial and given by $x = \eta$. We let $\theta(x_\emptyset, e)$ denote the multiplier attached to the incentive compatibility constraint. This - a bit unusual notation - reflects that in a constrained problem, θ must reflect the level of e that shall be implemented and the imposed constraint on $x_\emptyset$. The reason for this will be explained below.

To further simplify the problem note that upon integration by parts:

$$\int_\mu^\infty \int_\mu^\eta A\left(x(\eta)-\eta\right) d\eta dF_\eta = \int_\mu^\infty A\left(x(\eta)-\eta\right)\left(1-F_\eta\right) d\eta.$$

Furthermore, note that in case the agent is ignorant, the principal's utility is

$$\left\{\int_{-\infty}^\infty -\frac{A+1}{2}(x_\emptyset-\eta)^2 dF - Eu_\mu\right\} = \frac{-\sigma^2}{2} - \frac{(A+1)(x_\emptyset-\mu)^2}{2} - u_\mu$$

by the fact that $Eu_\mu = u_\mu - A\frac{\sigma^2}{2}$.

Hence the principal solves

$$
\begin{aligned}
&\max_{x(\eta)} K-u_{\mu}+(1-e)\left\{\frac{-\sigma^2}{2}-\frac{(A+1)(x_{\emptyset}-\mu)^2}{2}\right\}+\\
&e\left\{\int_{\mu}^{\infty}2\left(-\frac{A+1}{2}(x(\eta)-\eta)^2-A\left(x(\eta)-\eta\right)\frac{1-F_{\eta}}{f_{\eta}}\right)f_{\eta}d\eta\right\}+\\
&\vartheta\left\{\begin{array}{c} k+u_{\mu}+e\left(2\int_{\mu}^{\infty}A\left(x(\eta)-\eta\right)\frac{1-F_{\eta}}{f_{\eta}}f_{\eta}d\eta\right) \\ +(1-e)\left(-\frac{A\sigma^2}{2}\right)-g(e)\end{array}\right\}+\\
&\theta\left\{\frac{A\sigma^2}{2}+2\int_{\mu}^{\infty}A\left(x(\eta)-\eta\right)\frac{1-F_{\eta}}{f_{\eta}}f_{\eta}d\eta-g'(e)\right\}\\
&+\xi(\eta)\left\{x(\eta)-x_{\emptyset}\right\}
\end{aligned}
$$

Pointwise maximization delivers the first order necessary condition

$$
\begin{aligned}
0 &= -(1+A)(x^*(\eta)-\eta)-A\frac{1-F_{\eta}}{f_{\eta}}+\vartheta A\frac{1-F_{\eta}}{f_{\eta}} \\
&\quad +\frac{\theta(e)}{e}A\frac{1-F_{\eta}}{f_{\eta}}
\end{aligned}
\tag{2.33}
$$

and the complementary slackness conditions

$$
\theta\left\{\frac{A\sigma^2}{2}+2\int_{\mu}^{\infty}A\left(x(\eta)-\eta\right)\frac{1-F_{\eta}}{f_{\eta}}f_{\eta}d\eta-g'(e)\right\}=0
$$

and

$$
\xi(\eta)\left\{x(\eta)-x_{\emptyset}\right\}=0.
$$

Consider (2.33). A simple argument reveals that $\vartheta=1$: the principal can always use nondistorting transfers to shift around utility between him and the agent. Therefore our solution must yield $x\neq\eta$ if and only if $\theta>0$. Consequently one has $\vartheta=1$. By the fact that $\xi(\eta)\left\{x(\eta)-x_{\emptyset}\right\}=0$ the optimal decision schedule is given by

$$
\max\left\{x_{\emptyset},\eta+\frac{\theta(x_{\emptyset},e)}{e}\frac{A}{1+A}\frac{1-F_{\eta}}{f_{\eta}}\right\}.
$$

Figure 2.3 illustrates this contract:

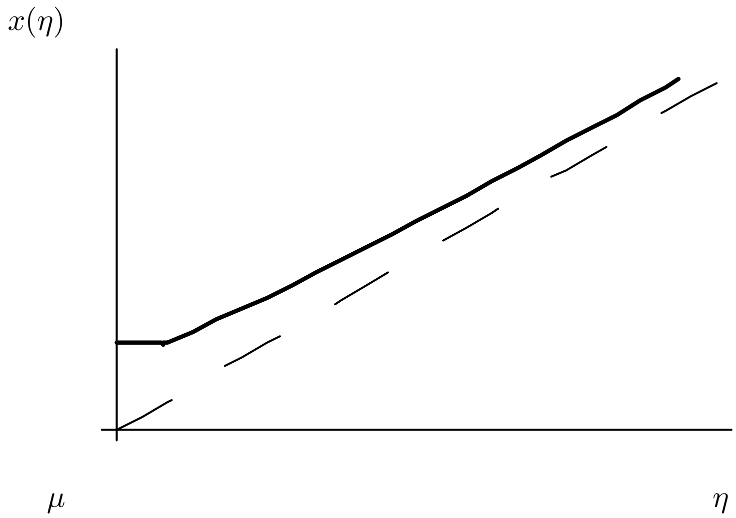

Figure 2.3

The flat part is due to the required monotonicity of the contract and the imposed constraint on $x_\emptyset$. Note that a solution can only exist if $x_\emptyset$ satisfies $x_\emptyset \geq \mu + \frac{\theta(x_\emptyset,e)}{e}\frac{A}{1+A}\frac{1-F_\mu}{f_\mu}$. If $x_\emptyset$ is set such that $x_\emptyset < \mu + \frac{\theta(x_\emptyset,e)}{e}\frac{A}{1+A}\frac{1-F_\mu}{f_\mu}$ the constraint is inconsistent with implementing the effort level! Note also that if $x_\emptyset$ is set very large, the constraint will be slack and $\theta = 0$.

The dependence of θ on $x_\emptyset$ and e :

$\frac{\partial}{\partial x_\emptyset}\theta(x_\emptyset, e) \leq 0$ with strict inequality if $\theta > 0$. To see this, take the binding incentive compatibility constraint on the agent's effort choice:

$$\frac{A\sigma^2}{2} + 2\int_\mu^\infty A\left(x(\eta) - \eta\right)\frac{1-F_\eta}{f_\eta}f_\eta d\eta - g'(e).$$

Making use of the derived optimal decision schedule this constraint can be expressed as

$$\begin{aligned}\frac{1}{2}\left[\frac{A\sigma^2}{2} - g'(e)\right] = &\int_\mu^{\breve{\eta}(x_\emptyset)} A\left(x_\emptyset - \eta\right)\frac{1-F_\eta}{f_\eta}f_\eta d\eta \\ &+ \int_{\breve{\eta}(x_\emptyset),}^\infty A\left(\frac{\theta(x_\emptyset,e)}{e}\frac{A}{1+A}\frac{1-F_\eta}{f_\eta}\right)\frac{1-F_\eta}{f_\eta}f_\eta d\eta\end{aligned}$$

where $\breve{\eta}(x_\emptyset)$ is defined such that

$$x_\emptyset = \breve{\eta}(x_\emptyset) + \frac{\theta(x_\emptyset,e)}{e}\frac{A}{1+A}\frac{1-F_{\breve{\eta}(x_\emptyset)}}{f_{\breve{\eta}(x_\emptyset)}}.$$

An increase in $x_\emptyset$ increases the integrand in the first line. Hence to keep the total rent constant (because the level of e that is to be implemented is held constant) the integrand in the second line must decrease. Hence $\frac{\partial}{\partial x_\emptyset}\theta(x_\emptyset, e) \leq 0$. (The effects on the limits of integration obviously cancel out.) A given level of e can be implemented introducing relatively large distortions around the mean and relatively small distortions at the tails of the distribution or the other way round.

That $\frac{\partial}{\partial e}\left[\frac{\theta}{e}\right] \geq 0$ with a strict inequality if $\theta > 0$ follows trivially from the binding incentive constraint and the fact that $g''(e) > 0$.

Solution part II: the optimal $x_\emptyset^*$:

Interestingly enough, as we now show, the principal always wants to smooth out the distortions as much as possible, or in other words to set

$$x_\emptyset^* = \mu + \frac{\theta(x_\emptyset, e)}{e}\frac{A}{1+A}\frac{1-F_\mu}{f_\mu}.$$

To see this, note that, by the fact that $\vartheta = 1$, one can rewrite the principal's maximization problem as

$$\begin{aligned}
&\max_{x_\emptyset} K + k - g(e) + (1-e)\left\{-\frac{(1+A)\sigma^2}{2} - \frac{(1+A)(x_\emptyset - \mu)^2}{2}\right\} \\
&+e\left\{2\int_\mu^\infty \frac{-(1+A)(x(\eta)-\eta)^2}{2}dF_\eta\right\} \\
&+\theta\left\{\frac{A\sigma^2}{2} + 2\int_\mu^\infty A\left(x(\eta)-\eta\right)\frac{1-F_\eta}{f_\eta}f_\eta d\eta - g'(e)\right\}.
\end{aligned} \tag{2.34}$$

s.t.

$$x(\eta) = \max\left\{x_\emptyset, \eta + \frac{\theta(x_\emptyset, e)}{e}\frac{A}{1+A}\frac{1-F_\eta}{f_\eta}\right\}. \tag{2.35}$$

The principal's problem is to maximize social surplus subject to the constraint that the agent's effort choice is incentive compatible. Taking the first derivative gives

$$
\begin{aligned}
&-(1-e)(1+A)(x_\emptyset-\mu)+e\left\{2\int_{\mu}^{\check{\eta}_{(x_\emptyset)}}-(1+A)(x_\emptyset-\eta)dF_\eta\right\}+ \\
&e\left\{2\int_{\check{\eta}_{(x_\emptyset)}}^{\infty}-(1+A)(x(\eta)-\eta)\frac{\partial\left[x(\eta)-\eta\right]}{\partial\theta}\frac{\partial\theta}{\partial x_\emptyset}dF_\eta\right\}+ \\
&\theta\left\{2\int_{\mu}^{\check{\eta}_{(x_\emptyset)}}A\frac{1-F_\eta}{f_\eta}f_\eta d\eta\right\}+\theta\left\{2\int_{\check{\eta}_{(x_\emptyset)}}^{\infty}A\frac{\partial\left[x(\eta)-\eta\right]}{\partial\theta}\frac{\partial\theta}{\partial x_\emptyset}\frac{1-F_\eta}{f_\eta}f_\eta d\eta\right\}.
\end{aligned} \tag{2.36}
$$

The effects on the limit of integration, $\check{\eta}_{(x_\emptyset)}$, have been omitted, since they obviously cancel out. Making use of (2.35) one has:

$$(1+A)e(x(\eta)-\eta)=\theta A\frac{1-F_\eta}{f_\eta}.$$

Therefore the third and the fifth term in (2.36) just cancel out as well! But then, since (by definition) for all $\eta\in[\mu,\check{\eta}_{(x_\emptyset)}]$ one has $x_\emptyset\geq x(\eta)$ and hence

$$\theta A\frac{1-F_\eta}{f_\eta}\leq(1+A)e(x(\eta)-\eta).$$

Consequently, the second effect in (2.36) dominates the fourth. Since the first effect is negative anyway, this establishes that the first derivative of (2.34) is negative for any value of $x_\emptyset>x(\mu)$. Therefore, the principal cannot gain anything if he "concentrates" distortions around the mean. Instead he smooths out as much as possible. Consequently, we have finally established that the optimal decision schedule satisfies

$$x(\eta)=\eta+\frac{\theta(e)}{e}\frac{A}{1+A}\frac{1-F_\eta}{f_\eta},$$

which is the expression in the text.

The final step of the optimization involves solving for the optimal level of e that shall be implemented. It is easy to prove that a solution exists. We shall not do this here since the trade-off is quite clear: implementing a higher e comes at the cost of more distortions. The trade-offs involved are the same as in the main text.

Chapter 3

Procurement with an Endogenous Type Distribution

3.1 Introduction

This paper studies a principal agent relationship à la Baron and Myerson (1982) with *endogenous characteristics* and *endogenous information asymmetries.* A buyer, the principal, contracts with a supplier, the agent, to obtain parts. In the seminal contribution of Baron and Myerson (BM henceforth) the distribution of characteristics is taken as given. In contrast, this distribution is endogenous in the present model and determined by a productive prior choice, henceforth effort, of the supplier. Two canonical examples of such a specification are presented. In the first we think of effort as a cost reducing investment. In the second we think of effort as learning. We assume that neither the effort nor the outcome of these productive prior choices are observable to the buyer. Consequently, there is moral hazard in the agent's effort decision and adverse selection at the time when contracts are signed. We ask how optimal contracts differ from the BM solution.

We show that standard properties like "no distortion at the top" and "underproduction below the top" may no longer hold. The reason is that contracts involve inference about unobserved effort as well as unobserved characteristics. An announcement of very low costs (at the top) is good news about the agent's effort. To create incentives for costly effort provision the principal may want to increase production of very efficient firms. In cases where at the top the inference about the agent's effort is very exact we show that distortions and overproduction arise at the top.

Overproduction at the top can be understood as a costly incentive device. The principal commits to a level of production over and above the level that would be justified ex post. The benefit of this commitment is to provide incentives for costly effort provision ex ante. Hence, our analysis can explain the emergence of unusually high production levels for the most efficient suppliers in terms of premia for unobserved prior investments.

The more informative messages at the top about unobserved effort the larger such premia. In the limiting case where the inference at the top is almost exact distortions at the top get arbitrarily large. At the same time, the principal can almost "get rid of the incentive problem". That is, his expected payoff is arbitrarily close but never actually equal to the payoff that he could obtain were the agent's prior effort contractable. Put differently, very simple contracts sometimes perform very well. Incentives for investments are provided through standard Baron Myerson contracts plus a reward to be paid when a minimum level of performance is reached.

In an investment setting the principal always benefits from a marginal increase in investment; the agent exerts a positive externality on the principal. The reason is that the principal prefers to trade with more efficient types ex post and a higher level of investment increases the likelihood of more efficient types.

In the learning context we explain the emergence of high powered incentive schemes. More precisely, incentives for costly information acquisition are provided through production schemes that make more use of the agent's information relative to the Baron Myerson contract. As a consequence, messages at the top are rewarded over and above the BM-quantity while messages at the bottom are punished with low levels of production. As in the investment context, rewards at the top may be so large as to generate distortions at the top and overproduction at the top.

Interestingly, the incentive problem is not always to motivate more information acquisition by the agent[1]. The externality that the agent exerts on the principal may also be negative. We show that the agent's thirst for knowledge is excessive when the principal is a quasi risk averter while the agent learns too little when the principal is a quasi risk lover. If the agent's thirst for knowledge is excessive the principal responds with an incentive scheme that makes less use of the agent's information than the Baron Myerson contract in that production is equalized across types.

This is obviously not the first paper to consider the interdependence of prior effort incentives and subsequent information and information rents. As extensions of Baron and Myerson (1982) (Laffont and Tirole (1986), respectively) most closely related are Crémer et.al. (1998a) and Lewis and Sappington (1997). These papers consider the impact of learning on optimal contracting. However, the learning technologies considered are somewhat special and do not allow for a general treatment of distortions at the top.[2] Demsky and Sappington (1986) and Szalay (2000) consider the implications of proper learning incentives for distortions in contracts when there is no conflict of interest ex post between principal and agent. Even without conflicts ex post the ex ante incentive problem of inducing the agent to acquire information may spill over into the solution to the problem ex post. That is, in the context of the delegation problem in Szalay (2000) the agent may optimally be forced to depart from ex post optimal choices. The idea is to force the agent to take a clear stance in one direction or the other and thereby enhance incentives to collect information. The present analysis differs by allowing for conflicts of interest and focussing attention to distortions at the top.

With respect to the investment model the idea that subsequent information rents provide incentives for private investments appears in Lewis and Sappington (1993a). However,

[1] Our model shares this feature with Crémer et al (1998).

[2] In these papers, as in the present one, information is productive. The literature has also emphasized other motives for information acquisition, notably strategic ones and has also treated other setups concerning the timing. Crémer et.al. (1998b) look at a situation where the seller of a good can acquire information on his costs before the buyer offers contracts. The information acquired is, however, socially useless because the seller will learn it anyway after he signed the contract. In Crémer and Khalil (1992) the seller can acquire information after the contract has been offered. Again, information acquisition is socially useless because the information will be available later on anyway. In the model of Kessler (1998) an agent has the possibility to commit to a certain information acquisition technology before the principal offers a contract. She shows that ignorance is of a strategic value to the agent.

they do not discuss distortions at the top. Finally, the influence of information rents on private investments is also discussed in Rogerson (1992) or Gul (2001). Timing and commitment structures are however quite different from the situation considered here.

The paper proceeds as follows: section 3.2 introduces the model and describes investment and the learning technology, respectively. Section 3.3 contains the main part of the analysis. Sections 3.4 presents optimal contracts, while sections 3.5 and 3.6 discuss their properties in the context of investments and learning, respectively. The final section concludes. All proofs are relegated to the appendix.

3.2 The Model

A principal contracts with an agent for the production of a good. The good is perfectly divisible so output can be produced in any quantity, q. q is perfectly observable and contractable. The agent receives a monetary transfer t from the principal. Both parties are risk neutral with respect to transfers. The principal derives gross surplus $V(q)$ from consumption, where $V(\cdot)$ is defined on $[0, \infty]$ with[3] $V_q(\cdot) > 0$, $V_{qq}(\cdot) < 0$, $\lim_{q \to 0} V_q(\cdot) = \infty$, $\lim_{q \to \infty} V_q(\cdot) = 0$. Thus the principal's net utility is $V(q) - t$.

The agent's payoff from receiving the transfer t and producing the amount q is given as $t - \beta q$. β is taken to be the realization of a random variable $\tilde{\beta}$. For the analysis of contracting we're actually less interested in $\tilde{\beta}$ and β than in a random variable $\tilde{\theta}$ which denotes the conditional expectation that the agent has about $\tilde{\beta}$ when he decides on the contract offer that the principal has proposed. The distribution of $\tilde{\theta}$ and in some constellations of $\tilde{\beta}$ as well depends on the agent's prior effort e. We consider the following time structure of the game:

[3] Throughout the paper subscripts will denote derivatives of the function with respect to the respective argument.

$---+------+-- --+-- ----+---$

P offers contract (a menu of contracts)	A takes effort e	θ realized and observed by A	production and transfers take place

First, the principal offers a contract. Then the agent chooses an effort level, e, that determines the distribution $F(\cdot, \cdot)$ of $\tilde{\theta}$. θ is realized and observed by the agent. Given this information he decides whether to sign the contract. If the agent does not sign a contract the game ends. If contracts are signed production and transfers take place.

We assume that the agent's choice of effort is not observable to the principal and that θ is the agent's private knowledge. Effort is assumed to involve a cost $g(e)$, where $g(\cdot)$ satisfies $g_e(\cdot) > 0,\ g_{ee}(\cdot) > 0,\ \lim_{e \to 0} g_e(\cdot) = 0,\ \lim_{e \to \infty} g_e(\cdot) = \infty$.

We have two canonical examples that give rise to the structure considered in this model. In the first example we think of e as a productivity enhancing investment. In the second example we think of e as learning.

In the investment model θ has the standard interpretation of marginal costs of production and we identify θ and β, i.e. $\theta \equiv \beta$. We assume that an increase in e shifts the distribution $F(\cdot, \cdot)$ of $\tilde{\theta}$ downwards in the sense of first order stochastic dominance

$$F_e(\cdot, \cdot) \geq 0 \ \forall \theta \tag{AI}$$

with strict inequality for some θ. For $e = 0$ $F(\cdot, \cdot)$ is degenerate on $\bar{\theta}$ while it has full support $\left[\underline{\theta}, \bar{\theta}\right]$ for all $e > 0$. Moreover, for all $e > 0$ we assume that $F(\cdot, e)$ has a density $f(\cdot, e)$, which is strictly positive on $\left(\underline{\theta}, \bar{\theta}\right]$.

In the learning model we suppose that the agent receives a signal about $\tilde{\beta}$. The precision of the signal depends on e.[4] θ represents the conditional expectation of $\tilde{\beta}$ given the signal,

[4]We work with deterministic signals in the sense that a given effort level delivers a signal of a given quality. Thus we exclude the case where the agent could receive signals of different qualities with certain probabilities. An example of such a technology would be that the agent receives a infinitely precise signal with probability e and a perfectly coarse signal with probability $1 - e$. Excluding this case simplifies the

the precision of the signal, and the effort level e. By the law of iterated expectations the mean of $\tilde{\theta}$ must be independent of effort. However, we assume that an increase in e makes $\tilde{\theta}$ riskier (and the conditional distribution of $\tilde{\beta}$ given the signal, the precision of the signal, and the effort level e less risky) in the sense of second order stochastic dominance:

$$\int_{\underline{\theta}}^{\theta} F_e\left(\tilde{\theta}, e\right) d\tilde{\theta} \geq 0 \, \forall \theta \tag{AL}$$

with strict inequality for some θ. The idea is that more effort entails more information about $\tilde{\beta}$. For $e = 0$ $F(\cdot, \cdot)$ is degenerate on the prior mean $E\tilde{\beta} = E\tilde{\theta}$ where E denotes the expectation operator. We assume that for all $e > 0$ $F(\cdot, e)$ has full support $\left[\underline{\theta}, \bar{\theta}\right]$ and has density $f(\cdot, e)$, which is again assumed to be strictly positive on $\left(\underline{\theta}, \bar{\theta}\right]$.

3.3 Analysis

3.3.1 The Principal's Problem

We think of contracting in terms of mechanism design. A mechanism is a tuple $\{q(\cdot), t(\cdot)\}$ which specifies quantities of production and transfers to the agent as a function of a message m the agent sends to the principal when he signs the contract. We can apply the *Revelation Principle* to restrict attention to direct, incentive compatible mechanisms, i.e. to mechanisms $\{q(\cdot), t(\cdot)\}$ that depend only on messages $\hat{\theta}$ about perceived costs. Given the revelation principle we can write the principal's problem as follows:

$$\max_{q(\cdot),\, t(\cdot),\, e} \int_{\underline{\theta}}^{\bar{\theta}} \left(V(q(\theta)) - t(\theta)\right) dF(\theta, e) \tag{P}$$

s.t.

$$t(\theta) - \theta q(\theta) \geq t(\hat{\theta}) - \theta q(\hat{\theta}) \, \forall \theta, \hat{\theta} \tag{3.1}$$

exposition a lot. For a solution for the technology with perfectly coarse/precise signals which occur with probability $e/1 - e$ see (Szalay (2001)).

$$t(\theta) - \theta q(\theta) \geq 0 \, \forall \theta \tag{3.2}$$

$$e \in \arg\max_{e} \left\{ \int_{\underline{\theta}}^{\bar{\theta}} (t(\theta) - \theta q(\theta)) \, dF(\theta, e) - g(e) \right\} \tag{3.3}$$

In the specification above, (3.1) is the truthtelling constraint to the revelation mechanism, (3.2) is the interim individual rationality constraint, (3.3) is the incentive compatibility constraint on the agent's effort choice. Note that the interim individual rationality constraints imply that the agent is willing to participate ex ante since the costless choice $e = 0$ is always admitted.[5]

The reader may wonder why (3.3) does not pay any attention to joint incentive compatibility of effort choice and subsequent reporting strategies. The reason is simply that for any potential deviation in effort choice, by the truth telling constraint (3.1), the agent finds it optimal to continue with a truth telling strategy.[6]

In the rest of this section we bring problem P into a more tractable form. All proofs of this section are in appendix A.

Lemma 3.1: Any $q(\cdot)$, $t(\cdot)$ and $e > 0$ solves problem P if and only if $t(\theta) = \theta q(\theta) + \int_{\theta}^{\bar{\theta}} q(\tilde{\theta}) d\tilde{\theta}$ and $q(\cdot)$, $e > 0$ solves Problem P':

[5] Imposing Interim Individual Rationality amounts to a limited liability constraint on the part of the agent: he is free to leave without further sanction after he is assigned a certain reward structure. This is a crucial assumption for our results. If unlimited penalties are allowed, it is well known that the first best can be implemented if the agent is risk neutral.

[6] The reader might further wonder about our notion of "directness" in the case of the learning model where the agent is asked to state his expected costs rather than the signal. We assume that there is a one-to-one relation between signals and conditional expectations so that exactly the same information is communicated whether messages are signals or conditional expectations.

For analyses with information gathering technologies where this one-to-one relation fails to hold, see Lewis and Sappington (1993b) and Szalay (2001).

$$\max_{q(.),\, e>0} \int_{\underline{\theta}}^{\bar{\theta}} \left(V(q(\theta)) - \left(\theta + \frac{F(\theta,e)}{f(\theta,e)} \right) q(\theta) \right) f(\theta,e) \tag{P'}$$

$$s.t.$$

$$e \in \arg\max_{e} \left\{ \int_{\underline{\theta}}^{\bar{\theta}} F(\theta,e)q(\theta)d\theta - g(e) \right\}$$

$$q_{\theta}(\theta) \leq 0$$

A further simplification is obtained by introducing a first order approach for e.

Proposition 3.1 *Assume*

$$\int_{\underline{\theta}}^{\theta} F_{ee}(\tilde{\theta},\cdot)d\tilde{\theta} \leq 0\,\forall\theta, e \tag{Concavity}$$

Then, $q(\cdot)$ *and* $e > 0$ *solve problem* P' *if they solve problem* P'' :

$$\max_{q(.),\, e>0} \int_{\underline{\theta}}^{\bar{\theta}} \left(V(q(\theta)) - \left(\theta + \frac{F(\theta,e)}{f(\theta,e)} \right) q(\theta) \right) f(\theta,e)d\theta \tag{P''}$$

$$s.t.$$

$$\int_{\underline{\theta}}^{\bar{\theta}} F_e(\theta,e)q(\theta)d\theta - g_e(e) = 0 \tag{IC}$$

$$q_{\theta}(\theta) \leq 0 \tag{Monotonicity}$$

Given the monotonicity constraint, it is convenient to treat P'' as an optimal control problem with control variable $\gamma(\cdot)$, where $\gamma(\cdot) = q_{\theta}(\cdot)$. The Hamiltonian for this problem is

$$\begin{aligned} H \;:&= \left\{ V\left(q(\theta)\right) - \left(\theta + \frac{F\left(\theta,e\right)}{f\left(\theta,e\right)} \right) q(\theta) \right\} f\left(\theta,e\right) + \mu\left(F_e\left(\theta,e\right) - g_e(e)\right) \\ &\quad + \lambda\left(\theta\right)\gamma\left(\theta\right) - \nu\left(\theta\right)\gamma\left(\theta\right) \end{aligned}$$

In this expression $\nu(\theta)$ is the Lagrange multiplier on the nonpositivity constraint on the control variable, $\lambda(\theta)$ is the costate variable and μ is the Lagrange multiplier on the IC constraint. From Pontryagins Maximum Principle the necessary conditions of optimality are

$$H_\gamma = 0 = \lambda(\theta) - \nu(\theta) \tag{3.4}$$

$$\lambda_\theta(\theta) = -H_q = -\left\{V_q(q(\theta)) - \left(\theta + \frac{F(\theta,e)}{f(\theta,e)} - \frac{\mu F_e(\theta,e)}{f(\theta,e)}\right)\right\} f(\theta,e) \tag{3.5}$$

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[V(q(\theta)) - \theta q(\theta) - \int_{\theta}^{\bar{\theta}} q(\tilde{\theta})d\tilde{\theta}\right] f_e(\theta,e)d\theta + \mu\left[\int_{\underline{\theta}}^{\bar{\theta}} F_{ee}(\theta,e)q(\theta)d\theta - g_{ee}(e)\right] = 0 \tag{3.6}$$

$$\int_{\underline{\theta}}^{\bar{\theta}} F_e(\theta,e)q(\theta)d\theta - g_e(e) = 0 \tag{3.7}$$

$$q_\theta(\theta) = \gamma(\theta) \tag{3.8}$$

$$\nu(\theta) \geq 0;\ \nu(\theta)\gamma(\theta) = 0 \tag{3.9}$$

$$\lambda(\bar{\theta}) = \lambda(\underline{\theta}) = 0 \tag{3.10}$$

where (3.4), (3.5), and (3.6) are the first order conditions for $\gamma(\cdot)$, $q(\cdot)$, and e, respectively, (3.7) and (3.8) spell out the constraints, (3.9) is the complementary slackness condition on the nonpositivity constraint on $\gamma(\cdot)$, and finally, (3.10) contains the transversality conditions. I proceed with a characterization of contracts that optimally implement an effort level the principal desires. Questions of existence and uniqueness are left aside.[7] Solution(s) to the system of equations (3.4)-(3.10) will be denoted $q^*(\cdot)$, $\gamma^*(\cdot)$, and e^* (with associated multiplier μ^*).

3.4 Implementation

In the analysis of Baron and Myerson (1982) effort choice plays no role (therefore $\mu \equiv 0$ in their model). If $\theta + \frac{F(\theta)}{f(\theta)}$ is non decreasing in θ the optimal production quantity in their analysis is characterized by $V_q(q(\theta)) = \theta + \frac{F(\theta)}{f(\theta)}$. For future reference call a contract satisfying this condition a *BM-contract*. The principal wants to implement efficient production (θ) but has to offer an informational rent ($\frac{F(\theta)}{f(\theta)}$) in order to have the agent reveal his information truthfully: the optimal contract reflects this rent/efficiency trade-off. Production for the inefficient firms is distorted downwards: there is underproduction below the top. However, there is no distortion at the top. I consider how the need to provide incentives for investments and learning affects the solution to the principal's problem:[8]

Proposition 3.2 *If* $\frac{\partial}{\partial\theta}\left[\theta + \frac{F(\theta,e^*)}{f(\theta,e^*)} - \mu^* \frac{F_e(\theta,e^*)}{f(\theta,e^*)}\right] \geq 0$ *the optimal contract is characterized through the first order condition*

$$V_q(q^*(\theta)) = \theta + \frac{F(\theta, e^*)}{f(\theta, e^*)} - \mu^* \frac{F_e(\theta, e^*)}{f(\theta, e^*)}$$

where

$$\mu^* = \frac{\int_{\underline{\theta}}^{\bar{\theta}} \frac{F_e(\theta,e^*)\cdot F(\theta,e^*)}{f(\theta,e^*)} q^*_\theta(\theta) d\theta}{\int_{\underline{\theta}}^{\bar{\theta}} F_{ee}(\theta, e^*) q^*(\theta) d\theta - g_{ee}(e^*) + \int_{\underline{\theta}}^{\bar{\theta}} \frac{F_e(\theta,e^*)^2}{f(\theta,e^*)} q^*_\theta(\theta) d\theta}$$

[7] To be precise: it is now well known that for a given e the following is true: if a solution to (3.5) exists, it is unique because $\left\{V\left(q(\theta)\right) - \left(\theta + \frac{F(\theta,e)}{f(\theta,e)} - \mu \frac{F_e(\theta,e)}{f(\theta,e)}\right) q(\theta)\right\}$ is concave in $q(\cdot)$ (see Guesnerie and Laffont (1984)). Sufficient conditions for existence are provided there as well. With a suitable adjustment taking care of the endogeneity of e their results can be carried over.

[8] All proofs of sections 3.4 through 3.6 are in Appendix B.

and e^* is determined by

$$\int_{\underline{\theta}}^{\bar{\theta}} F_e(\theta, e^*) q^*(\theta) d\theta - g_e(e^*) = 0$$

Incentives for investments/learning are provided through marginal information rents. The direction of the change in the optimal contract relative to the Baron Myerson situation for any type θ depends on the sign of μ^* and the sign of $F_e(\theta, e)$. The sign of the multiplier reflects the nature of the incentive problem concerning effort choice with $\mu^* > 0$ when there is underprovision of effort and $\mu^* < 0$ in case of overprovision of effort. The sign of $F_e(\theta, e)$ is determined by (AI) for the investment model, and (AL) for the learning model, respectively. Both models will be discussed separately in sections 3.5 and 3.6, respectively.

Along the lines of Holmström (1979), the distortion introduced through the moral hazard problem can be understood in terms of an inference problem when effort is viewed as an unobserved parameter, that is, in terms of likelihood ratios. The relevant likelihood measure conditions on the agent being at least as efficient as some type θ. To understand this result, recall that the ex post rent of a type θ depends on production quantities of all types with higher or equally high costs. Therefore a marginal change of production at θ changes the rent of all types who are at least as efficient as θ. The principal's cost benefit analysis of distortions takes this into account. More specifically, observe that $\frac{F_e(\theta,e)}{f(\theta,e)} = \frac{d \log[F(\theta,e)]}{de} / \frac{f(\theta,e)}{F(\theta,e)}$. $\frac{f(\theta,e)}{F(\theta,e)}$ is the density of type θ conditional on the agent being at least as efficient as θ. The higher is $\frac{f(\theta,e)}{F(\theta,e)}$ the higher is the cost of distorting the contract. $\frac{d \log[F(\theta,e)]}{de}$ is the slope of the likelihood function for the agent being at least as efficient as θ. In that sense $\left|\frac{d \log[F(\theta,e)]}{de}\right|$ measures how strongly one is inclined to infer from message θ that the agent did not take action e and the higher is the incentive effect of a given change in $q(\cdot)$ on the agent's effort.

The value μ^* is determined by the principal's optimization w.r.t. e. The corresponding first order condition is

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[V(q(\theta)) - \theta q(\theta) - \int_{\theta}^{\bar{\theta}} q(\tilde{\theta}) d\tilde{\theta} \right] f_e(\theta, e) d\theta = \mu \left[g_{ee}(e) - \int_{\underline{\theta}}^{\bar{\theta}} F_{ee}(\theta, e) q(\theta) d\theta \right] \tag{3.11}$$

The left hand side measures the direct, beneficial, effect of an increase in effort on the principal's payoff. It can be decomposed into two separate effects: first a social surplus effect, $\int_{\underline{\theta}}^{\bar{\theta}} \left[(V(q(\theta)) - \theta q(\theta)) f_e(\theta, e) \right] d\theta$, and second a rent effect, $\int_{\underline{\theta}}^{\bar{\theta}} \int_{\theta}^{\bar{\theta}} q(\tilde{\theta}) d\tilde{\theta} f_e(\theta, e) d\theta$. Due to the envelope theorem remaining effects through a change in the optimal contract are zero. The terms on the right hand side reflect indirect cost effects that arise through the incentive constraint on effort choice. The principal can be understood as a monopsonistic buyer of effort who considers the total change in costs of effort due to an increase of his demand for effort. Such a change consists of a "price effect" and a "quantity effect". Due to the envelope theorem the price effect is zero. The quantity effect has the following interpretation: a change in effort by de affords an increase in marginal information rents to $g_e(e) + g_{ee}(e)de$. The effect on the principal's payoff is the change in marginal costs times the shadow price of marginal costs, μ. Moreover, from the concavity of the agent's information rent, it follows that an increase in effort is accompanied by a deadweight loss, causing a loss of utility of height $-\int_{\underline{\theta}}^{\bar{\theta}} \mu F_{ee}(\theta, e) q(\theta) d\theta$.
We next discuss the solutions to the implementation problems in deeper detail.

3.5 Investments

We now look at the investment model where an increase in effort corresponds to a shift in the distribution of θ involving first order stochastic dominance. We proceed as follows: we first characterize qualitative features of contracts relative to the BM-contract and discuss the nature of the incentive problem. It turns out that these two issues are closely related. Next we discuss the impact of the moral hazard problem on standard features of contracts with hidden characteristics.

Proposition 3.3 *The agent invests too little from the principal's perspective, i.e.* $\mu^* > 0$. *Incentives for cost reducing investments are provided through increased production relative to the BM-contract.*

The intuition for these results is straightforward. Since the qualitative influence of the agent's effort is the same for all types, i.e. to increase the likelihood of more efficient types, the principal will either increase ($\mu^* > 0$) production for all types relative to the BM-contract or decrease it for all types ($\mu^* < 0$). Now suppose that contrary to our proposition $\mu^* < 0$. Then, from proposition 3.2 we see that there is underproduction relative to first best. Now consider the left hand side of (3.11). Taking derivatives of the principal's ex post net utility with respect to θ we get

$$\frac{\partial}{\partial\theta}\left[\int_{\underline{\theta}}^{\bar{\theta}}\left(V(q(\theta)) - \theta q(\theta) - \int_{\theta}^{\bar{\theta}} q(\tilde{\theta})d\tilde{\theta}\right)d\theta\right] = \int_{\underline{\theta}}^{\bar{\theta}}\left[V_q(q(\theta)) - \theta\right]q_\theta(\theta)d\theta.$$

Consequently, if there is underproduction throughout, the principal's ex post net utility is a monotonic function of types. More precisely, the principal's net utility ex post is monotonically decreasing in θ so that he prefers to trade with the more efficient types ex post. Observe now that a marginal increase of the agent's effort increases the likelihood of low cost types. Therefore the direct marginal effect on the principal's net utility is unambiguously positive.[9] Consider now the indirect effects on the cost side: when $\mu^* < 0$ an increase in effort slackens the incentive constraint and furthermore also reduces the deadweight loss. So if $\mu^* < 0$ there is really no trade-off involved. Therefore necessarily $\mu^* > 0$ in any optimal solution. Then, the principal equates at the optimal solution the marginal net benefit to the marginal cost. In particular this reasoning reveals that the marginal effect on the net benefit is unambiguously positive at the optimum. In other words, the agent exerts a positive externality on the principal. Therefore at an optimal

[9] The argument is akin to the one given in Finance to establish that an investor with a strictly monotonic utility function will always prefer lotteries involving a higher expected return.

solution, the principal pays the agent for exerting more effort ($\mu^* > 0$) rather than for exerting less effort ($\mu^* < 0$). Paying for more effort in turn means that marginal incentives for effort have to be increased. Since rents are determined by cumulative production levels it follows that production is higher than in the BM-contract.

Thus inference about unobserved effort has a countervailing effect on optimal contracts as opposed to inference about unobserved characteristics. The latter tends to generate c.p. underproduction while the former tends to generate overproduction. The net effect on contracts depends on a comparison of the relative strength of the two motives. This comparison in turn depends on the size of the endogenous multiplier μ^*. However, if we assume that $\lim_{\theta \to \underline{\theta}} \frac{F(\theta,e)}{f(\theta,e)} = 0 \, \forall e > 0$,[10] then there is an ideal suspect where this endogeneity problem can be avoided because the desire to limit the agent's informational rents is entirely absent: the most efficient type $\underline{\theta}$. In what follows, for lack of a better terminology, we define *the top* as the type with the lowest marginal costs, $\underline{\theta}$.

Thus, we impose for the remainder of this section:

Assumption: $\lim_{\theta \to \underline{\theta}} \frac{F(\theta,e)}{f(\theta,e)} = 0 \, \forall e > 0$.

Proposition 3.4 *Suppose that* $\infty > \lim_{\theta \to \underline{\theta}} \frac{F_e(\theta,e)}{f(\theta,e)} > 0 \, \forall e > 0$. *Then there is a distortion at the top. More precisely, there is overproduction at the top.*

Overproduction arises whenever the desire to limit the agent's information rents is dominated by the desire to increase his effort choice. Since $\lim_{\theta \to \underline{\theta}} \frac{F(\theta,e)}{f(\theta,e)} = 0 \, \forall e > 0$, at the top the desire to limit the agent's rents is absent. Consequently, if announcing to be of type $\underline{\theta}$ is informative about unobserved effort there is a distortion at the top.[11] More precisely, there is overproduction at the top. As argued above, $\frac{F_e(\underline{\theta},e)}{f(\underline{\theta},e)}$ measures how exact an inference one can make about unobserved effort from the message $\underline{\theta}$. The more exact

[10] We have to impose $\lim_{\theta \to \underline{\theta}} \frac{F(\theta,e)}{f(\theta,e)} = 0$ since we have allowed for the case where $\lim_{\theta \to \underline{\theta}} f(\theta, e) = 0$.

[11] Observe, that the benefit in terms of the influence on the probability of being more efficient than θ, $F_e(\theta, e)$, must go to zero as we approach the top since we assumed that the support of types is independent of effort. However, it may still be the case that the cost of distorting the quantity of production, measured by the density $f(\theta, e)$ approaches zero as well, and does so faster than the benefit.

the inference the larger the distortion at the top and c.p. the larger is production at the top.

Message $\underline{\theta}$ (and by continuity some messages below the top) is rewarded over and above what would be optimal ex post because it displays a favourable benefit to cost ratio. An increase in production at the top generates a pronounced effect on the agent's effort while the cost of distorting the contract is relatively small. The ex post efficiency of the contract is impaired in order to provide incentives for investments ex ante.

In a certain sense distortions at the top are a very costly incentive device.

Proposition 3.5 *Suppose that* $\infty > \lim_{\theta \to \underline{\theta}} \frac{F_e(\theta,e)}{f(\theta,e)} > 0 \, \forall e > 0$. *Furthermore, assume that* $\frac{\partial}{\partial \theta} \frac{F(\theta,e)}{f(\theta,e)} \geq 0 \; \forall e > 0$ *and* $\frac{\partial}{\partial \theta} \frac{F_e(\theta,e)}{f(\theta,e)} \leq 0 \; \forall e > 0$. *Then* $\exists \, \check{\theta}$ *such that there is overproduction for all* $\theta < \check{\theta}$ *and underproduction for all* $\theta > \check{\theta}$. *Moreover, the principal's ex post utility is maximized at* $\check{\theta}$, *increasing in* θ *for* $\theta < \check{\theta}$ *and decreasing for all* $\theta > \check{\theta}$.

If the top is informative about effort, the distribution is logconcave and our conditional likelihood ratio is monotone (and decreasing) in θ then the following result obtains: for very efficient firms the efficiency of the ex post contract is impaired so that ex post the principal's net utility is no longer maximized at the top. The principal trades too much with very low cost types, i.e. implements too high a level of production, in order to steer the agent's effort supply. The net effect of an increase in the agent's effort is still beneficial from the principal's perspective because the distribution of types becomes more favourable from his perspective. This commitment makes sense from the ex ante perspective whenever the informational content of messages at the top is sufficiently large.

Our results so far refer to the case where the top is informative but not "too informative" about effort. More precisely, the conditional likelihood ratio measure was assumed to be bounded at the top. One may wonder what happens if at the top the inference the principal can form about unobserved effort gets arbitrarily exact. In the context of a pure moral hazard model Mirrlees (1974) has shown that precise inference at the bounds of the supports generate problems of nonexistence of optimal solutions. In his model, the principal can reach an expected payoff that is arbitrarily close but never actually equal to the expected payoff he could obtain were effort observable by using simple step contracts involving large and infrequent punishments. The same reasoning applies here as well, albeit under more stringent conditions.

Proposition 3.6 *Suppose that* $\lim_{\theta \to \underline{\theta}} \frac{F_e(\theta,e)}{f(\theta,e)} = \infty \, \forall e > 0$. *Then, an optimal contract does not exist. The principal can reach a payoff that is arbitrarily close (but never actually equal to) his payoff for the given information structure* $F(\cdot, \tilde{e})$ *by implementing the following contract:*

$$\begin{aligned} q &= \hat{q} \text{ for } \theta \in \left[\underline{\theta}, \underline{\theta} + \varepsilon\right] \\ q &= \tilde{q}(\theta) \text{ otherwise} \end{aligned}$$

where $\hat{q}$ *is very large,* ε *is very small but positive,* $\tilde{q}(\theta)$ *satisfies:*

$$V_q(\tilde{q}(\theta)) = \theta + \frac{F(\theta, \tilde{e})}{f(\theta, \tilde{e})} \text{ for } \theta \in \left(\underline{\theta} + \varepsilon, \bar{\theta}\right],$$

and $\tilde{e}$ *maximizes the principal's expected payoff when effort is observable.*

Saying that the likelihood ratio goes out of bounds at the top is equivalent to saying that the benefit of distorting the contract is huge compared to the costs of doing so. In such a situation the principal can implement any level of effort at approximately no costs in expected terms. He offers an extremely high reward for the most efficient types, or more precisely, an extremely large amount of production for the most efficient firms. However, this reward will almost never have to be paid since the probability of reaching that state is approximately zero.

As a technical side remark, observe that our condition is stronger than the one Mirrlees (1974) needed to obtain such a result in the context of a pure moral hazard model.[12] The reason why we need a stronger condition is that our step contract involves two components

[12]More precisely, assuming that $\lim_{\theta \to \underline{\theta}} \frac{F(\theta,e)}{f(\theta,e)} = 0$, $\lim_{\theta \to \underline{\theta}} \frac{F_e(\theta,e)}{f(\theta,e)} = \infty \Rightarrow \lim_{\theta \to \underline{\theta}} \frac{f_e(\theta,e)}{f(\theta,e)} = \infty$, but the reverse is not true.

of costs which both have to vanish in expected terms when the principal implements ever higher amounts of production for an ever smaller set of very efficient types whereas there is only one such cost in Mirrlees' model. More precisely, as in Mirrlees' problem motivating more effort from the agent through the contract in the proposition involves an increase in rent cost. The increase in expected rent costs is negligible in our model whenever it is Mirrlees' model, that is when $\lim_{\theta \to \underline{\theta}} \frac{f_e(\theta,e)}{f(\theta,e)} = \infty$. However, there is also the increase in *physical* production costs. These vanish if the principal implements very seldomly very high levels of production only if $\lim_{\theta \to \underline{\theta}} \frac{F_e(\theta,e)}{f(\theta,e)} = \infty$.
However, even if the expected increase in physical production costs does not vanish, inference about unobserved effort may still have a considerable influence on the contract at the top. Distortions at the top may arise even if the top is *completely uninformative* about unobserved effort.

Proposition 3.7 *Suppose that* $\lim_{\theta \to \underline{\theta}} \frac{F_e(\theta,e)}{f(\theta,e)} = 0 \ \forall e > 0$ *but that* $\lim_{\theta \to \underline{\theta}} \frac{f_e(\theta,e)}{f(\theta,e)} = \infty$ $\forall e > 0$. *Then, there is bunching at the top. If in addition,* $-\left[\theta + \frac{F(\theta,e^*)}{f(\theta,e^*)} - \mu^* \frac{F_e(\theta,e^*)}{f(\theta,e^*)}\right]$ *is singlepeaked, then there is a distortion at the top. More precisely, there is overproduction at the top.*

Even if the conditional likelihood measure is zero at the top it may still be the case that its *change* is very pronounced as we move away from the top. This change at the top is proportional to $\frac{f_e(\underline{\theta},e)}{f(\underline{\theta},e)}$. If this likelihood measure is very large, inference about unobserved effort conflicts with the monotonicity constraint so that there must be bunching in the contract.

The last propositions all require that $\lim_{\theta \to \underline{\theta}} f(\theta, e) = 0$. In contrast Baron and Myerson have imposed a strictly positive density over the whole support. Consequently, the reader might judge our results incomparable to theirs. However, this is not true. We have imposed that $\lim_{\theta \to \underline{\theta}} \frac{F(\theta,e)}{f(\theta,e)} = 0$. As a consequence we obtain *exactly the same* results as they do for the case where the distribution of types is exogenously given.

3.6 Learning

We now discuss the learning model where increases in effort are associated with shifts of the distribution, $F(\cdot,\cdot)$, involving second order stochastic dominance. We proceed as follows: we first characterize the nature of the incentive problem in the learning context and then discuss the properties of optimal contracts.

For ease of exposition of the next proposition let a nondecreasing hazard rate be *strongly concave* if it satisfies the following elasticity condition:

$$\frac{\frac{\partial}{\partial\theta}\left[\frac{F(\theta,e)}{f(\theta,e)}\right]}{\frac{F(\theta,e)}{f(\theta,e)}} \leq -\frac{\frac{\partial^2}{\partial\theta^2}\left[\theta+\frac{F(\theta,e)}{f(\theta,e)}\right]}{\frac{\partial}{\partial\theta}\left[\theta+\frac{F(\theta,e)}{f(\theta,e)}\right]}$$

Proposition 3.8 (i) Equilibrium effort can be either suboptimally low ($\mu^* > 0$) or excessive ($\mu^* < 0$). More specifically, suppose that P'' is globally concave in e for given $q(\cdot)$ and $\gamma(\cdot)$. Then:
$\mu^* > 0$ if $V_{qqq}(\cdot) \geq 0$ and $\frac{F(\theta,e)}{f(\theta,e)}$ is nondecreasing strongly concave in e;
$\mu^* < 0$ if $V_{qqq}(\cdot) \leq 0$ and $\frac{F(\theta,e)}{f(\theta,e)}$ is nondecreasing convex in e.
(ii) For any (possibly non-concave) Problem P'', if $\lim_{e\to 0} f(\theta,e) = 0$ $\forall\theta \in \left[\bar{\theta}-\varepsilon,\bar{\theta}\right]$ for some $\varepsilon > 0$, then equilibrium effort is bounded away from zero.

The intuition for these results is quite different from the one in the investment model. The marginal impact of the agent's effort on the principal's net benefit,

$$\int_{\underline{\theta}}^{\bar{\theta}}\left[V(q(\theta)) - \theta q(\theta) - \int_{\theta}^{\bar{\theta}} q(\tilde{\theta})d\tilde{\theta}\right] f_e(\theta,e)d\theta,$$

is akin to a lottery adding risk to the distribution of θ. The agent's effort moves mass from the middle of the distribution towards the outer bounds so that extreme realizations of the posterior are relatively more likely. Consequently, the question whether the principal prefers to have the agent exert more effort is equivalent to asking whether he likes to incur marginal lotteries. A sufficient condition such that he likes such marginal lotteries is that he is a "quasi risk lover". That is, if the principal's net ex post utility under the optimal contract is a convex function of types then the principal benefits from a marginal increase in effort. On the other hand if his net ex post utility is concave in types, so that he is a "quasi risk averter", the principal would prefer to have less risk in the distribution of types.[13]

It is instructive to discuss why the argument cannot be given along the same lines as in the investment model. Even if one assumes that there is underproduction for all types[14] so that the principal's ex post utility is monotonic in types, the influence of a marginal increase in the agent's effort on the principal's net utility is ambiguous. In a learning model the agent shifts mass from the middle of the distribution towards the outer bounds of the distribution. More extreme posteriors are in an ex ante sense more likely if the agent exerts more effort. Consequently, more effort reduces the likelihood of trading with mediocre types and increases the likelihood of trades with extreme types. The principal's ex post utility is decreased if the agent's marginal cost turns out to be higher than ex ante expected and increased if the agent's marginal cost turn out to be lower than expected. Thus, the nature of the incentive problem can in general be of either kind: to motivate the agent to learn more or to stop him from learning more. However, even if the principal is harmed by a marginal increase of the agent's effort at an optimal solution this will -under a minor technical condition- not go so far that the principal would like the agent to abstain from learning completely. Implementing no learning is extremely costly. To see this, note that the agent uses his information in two decisions: first in deciding on whether to work for the principal and second in deciding on what quantity to produce. To stop the agent from learning the principal has to render the marginal information completely valueless

[13] We briefly comment on why attention is restricted to the special case where P'' is concave in $\{q(\cdot), e\}$. The shape of the principal's utility depends on the shape of contracts which in turn depends on the size of the multiplier μ. For the special case, where P'' is *simply assumed* to be concave, it suffices to ask whether the principal's payoff is locally increasing around $\mu = 0$. The sufficient conditions in the propositions refer to this special example. However, the conclusion that μ can be of either sign does not depend on whether P'' is concave or not. Alternatively, the sufficient conditions could have been stated as: $\mu^* > 0$ if $V_{qqq}(\cdot) \geq 0$ and $\frac{\frac{\partial}{\partial\theta}\left[\frac{F(\theta,e)}{f(\theta,e)}\right]}{\frac{F(\theta,e)}{f(\theta,e)}} \leq \frac{-\frac{\partial^2}{\partial\theta^2}\left[\frac{F(\theta,e)}{f(\theta,e)} - \mu\frac{F_e(\theta,e)}{f(\theta,e)}\right]}{\frac{\partial}{\partial\theta}\left[\theta + \frac{F(\theta,e)}{f(\theta,e)} - \mu\frac{F_e(\theta,e)}{f(\theta,e)}\right]}$ $\forall\theta$, $\forall e > 0$ with a similar modification for the case where $\mu^* < 0$. However, since μ^* is endogenous, it is not clear what such a condition really means.

[14] Observe that our proposition and its formal proof makes no use of such an assumption.

for the agent in both decisions since the marginal cost of effort is zero at zero effort. This requires that (i) the payment to the agent is so high that the contract is acceptable to the type with the highest cost and (ii) that produced quantity is independent of the agent's costs. In such a contract the principal gives away very high rents. Consider then the alternative of lowering the transfer so that the agent would not be willing to work for the principal if he believed that his costs where very high, in the interval $[\bar{\theta} - \varepsilon, \bar{\theta}]$, say, for some ε positive but small. The virtue of this alternative contract is that the principal can now buy cheaper from all types who think that their costs are lower than $\bar{\theta} - \varepsilon$. The drawback is the loss of gains from trade with the high cost types, for types within $[\bar{\theta} - \varepsilon, \bar{\theta}]$ will not participate anymore. If it is, for e close to zero, sufficiently unlikely that the agent indeed perceives his costs to be so high as to abstain from trading with the principal the virtue of this alternative contract outweighs the drawbacks.

We next turn to a discussion of the properties of optimal contracts:

Proposition 3.9 *Suppose that $\mu^* > 0$. Then, the optimal contract implements a weakly larger production than the BM-contract for the most efficient types in a range $[\underline{\theta}, \theta_1]$ (strictly larger for some of them) and a weakly smaller production than the BM-contract for types in a range $[\theta_2, \bar{\theta}]$ (strictly smaller for some of them) for some $\underline{\theta} < \theta_1 \leq \theta_2 < \bar{\theta}$. These conclusions are reversed if $\mu^* < 0$.*

It is easy to see that production can never be increased for all types in a learning model relative to the BM-solution, regardless of the nature of the incentive problem. By the law of iterated expectations the mean of the distribution of types must be independent of the agent's effort[15]. Since the agent moves mass from mediocre types to extreme types at the margin the likelihood of types at both outer bounds must be increased. Consequently, if production is increased relative to the BM-solution at one end of the distribution this implies that it is decreased for types at the other end of the distribution.

[15] Technically: we must have $\int_{\underline{\theta}}^{\bar{\theta}} F_e(\theta, e) d\theta = 0 \, \forall e$.

Intuitively, if the principal wants to foster information acquisition, he must increase the marginal value of information to the agent. Information is the more valuable to the agent the more his ex post rent depends on his information, that is the more production depends on his information. Therefore the spread between the highest production level at the top and the lowest production levels at the bottom is increased. More generally, the marginal value of information to the agent is increased by increasing production at all values of θ where the marginal impact of the agent's effort is to move the agent's posterior towards lower costs and by decreasing production at all values of θ where the agent's effort moves the posterior towards higher costs. In cases the agent's posterior is moved towards higher costs the agent's ex post rent is decreased. In contrast the agent's ex post rent is increased when the his posterior is moved towards lower costs. The positive impact of the first effect is the higher the higher production at such values of the posterior. The negative effect of the latter effect is the smaller the smaller the quantity of production at such a value of the posterior.

This reasoning is turned upside down if the principal wants to stop the agent from learning more. Then he must decrease the marginal value of information to the agent. In this case optimal contracts must make less use of the agent's information so that the spread between extreme production levels is decreased.

To sum up: our analysis explains the emergence of incentive schemes that are more extreme than the BM-solution in cases where the principal benefits from information acquisition. In cases where the principal is harmed by more information on the part of the agent, optimal contracts are more equalized across types and make less use of the agent's information than the BM-contract.

In general, second order stochastic dominance gives too little structure to determine the global shape of contracts. However, in some special cases the local characterization given above is enough to argue that the overall incentive scheme gets more extreme:

Corollary: *Suppose that $F_e(\theta, e) \geq 0\, for\, \theta \leq E\theta$ and $F_e(\theta, e) \leq 0$ otherwise $\forall e$. In particular, this is the case when $f(\cdot, \cdot)$ is singlepeaked and symmetric $\forall e$. Then, production* *is higher than the BM-quantity for all types with conditionally expected marginal costs below prior expected marginal costs and lower than the BM-quantity otherwise.*

A similar result has been found by Crémer et.al. (1998a) in a model where the agent can acquire perfect knowledge at a fixed cost. For their case, when the agent has spend the cost, he perfectly knows the cost parameter β so that we may identify the conditional expectation θ with β. We indicate the ex ante distribution of the conditional expectation of an informed agent with $F(\cdot, 1) \equiv F(\beta)$. If he does not spend the cost he gains no additional information so that his posterior is equal to the prior with probability one and hence $F(\theta, 0) = 1_{\theta \geq E\theta}$. Consequently, $\frac{F_e(\theta,e)}{f(\theta,e)}$ is equal to $\frac{F(\theta)-1_{\theta > E\theta}}{f(\theta)}$ and it follows that the contract has the same properties as the one in the corollary.

Another immediate conclusion from proposition 3.9 is that in the limited range of types at and below the top the agent's influence on the distribution of types is identical to the one in the investment model. Consequently the same principles apply when the nature of the incentive problem is the same as in the investment model ($\mu > 0$). Again, for the remainder of this section we impose:

Assumption: $\lim_{\theta \to \underline{\theta}} \frac{F(\theta,e)}{f(\theta,e)} = 0$.

Proposition 3.10 *Suppose that* $\infty > \lim_{\theta \to \underline{\theta}} \frac{F_e(\theta,e)}{f(\theta,e)} > 0 \; \forall e > 0$. *If* $\mu^* > 0$, *then there is a distortion at the top. More precisely, there is overproduction at the top. If* $\mu^* < 0$ *then there is either a distortion at the top or bunching at the top or both. If there is no bunching at the top then there is underproduction at the top.*

The first part is identical to what we observed in the discussion of investments. Consequently, we will only comment on the new features. If $\mu < 0$ the principal wants to stop the agent from acquiring more information. Since higher effort makes more extreme outcomes more likely the principal "punishes" the extreme message $\underline{\theta}$. C.p. this leads to a reduction in produced quantity at the top.

Proposition 3.11 *Suppose that* $\lim_{\theta \to \underline{\theta}} \frac{F_e(\theta,e)}{f(\theta,e)} = \infty\ \forall e > 0$. *If* $\mu^* > 0$ *then, an optimal contract does not exist. The principal can reach a payoff that is arbitrarily close (but never actually equal to) his payoff for the given information structure* $F(\cdot, \tilde{e})$ *by implementing the following contract:*

$$
\begin{aligned}
q &= \hat{q} \text{ for } \theta \in \left[\underline{\theta}, \underline{\theta} + \varepsilon\right] \\
q &= \tilde{q}(\theta) \text{ otherwise}
\end{aligned}
$$

where $\hat{q}$ *is very large,* ε *is very small but positive,* $\tilde{q}(\theta)$ *satisfies*

$$
V_q(\tilde{q}(\theta)) = \theta + \frac{F(\theta, \tilde{e})}{f(\theta, \tilde{e})} \text{ for } \theta \in \left(\underline{\theta} + \varepsilon, \bar{\theta}\right],
$$

and $\tilde{e}$ *maximizes the principal's payoff for observable effort.*

If $\mu^* < 0$, *then an optimal contract still exists. The contract involves bunching at the top.*

Part (i) of the proposition is isomorphic to its counterpart in the investment context so we need only comment on part (ii). Contrary to the possibilities of increasing the agent's effort choice the possibilities to stop the agent from acquiring more information are limited. The maximally feasible punishment is to implement zero production. Put differently, since the agent must be willing to participate ex post, the principal cannot implement arbitrarily large (but infrequent) punishments. Consequently, non existence problems of optimal solutions of this type cannot arise. Moreover, the solution will display bunching at the top. With an accurate inference at the top the principal will use the information at the top to stop the agent from learning. If there was no bunching at the top, by monotonicity of contracts, this would imply that he implements a very low level of production for all types. If the inference is almost exact at the top this would mean that the principal implements almost zero production throughout and this cannot be optimal.

Proposition 3.12 *Suppose that* $\lim_{\theta \to \underline{\theta}} \frac{F_e(\theta,e)}{f(\theta,e)} = 0\ \forall e > 0$ *but that* $\lim_{\theta \to \underline{\theta}} \frac{f_e(\theta,e)}{f(\theta,e)} = \infty$ $\forall e > 0$. *If* $\mu^* > 0$, *then there is bunching at the top. If in addition,* $-\left[\theta + \frac{F(\theta,e^*)}{f(\theta,e^*)} - \mu^* \frac{F_e(\theta,e^*)}{f(\theta,e^*)}\right]$

is singlepeaked, then there is a distortion at the top. More precisely, there is overproduction at the top. If $\mu^* < 0$ *then there is no distortion at the top.*

The intuition is straightforward. If $\lim_{\theta \to \underline{\theta}} \frac{f_e(\theta,e)}{f(\theta,e)} = \infty$ implies a nonmonotonicity in the contract for the case $\mu^* > 0$ then it must imply that the contract is monotonic for the case $\mu^* < 0$. If in turn the contract is monotonic at the top, then optimal production is characterized by $V_q(q^*(\underline{\theta})) = \underline{\theta}$ and the conclusion follows.

3.7 Conclusion

We have discussed two canonical examples leading to endogenous type distributions in procurement problems: productivity enhancing investments and learning. In a limited range of types the influence of marginal changes in the level of investments or the intensity of learning was found to be qualitatively the same. Consequently the same qualitative features of contracts arise. In particular, inference about unobserved investments and learning may lead to distortions at the top and bunching at the top. The most interesting difference between the two examples is the nature of the incentive problems. While the principal will always benefit from marginal increases in investments the agent's thirst for knowledge may also be excessive. Put differently, the agent exerts a positive externality on the principal in the investment context and the principal pays to have the agent partly internalize this externality. In contrast, the externality may be negative in the learning context and the principal might optimally pay the agent for exerting less effort.

Our results have been derived within a first order approach. It is interesting to note that taking such an approach is less restrictive in the current setting with moral hazard *and* adverse selection being present than in problems where there is only moral hazard. The reason is quite simple: in moral hazard problems it is vital to *ensure* that contracts have monotonicity properties. This is achieved by imposing the Monotone Likelihood Ratio Property (Rogerson (1985)). Monotonicity of contracts together with a condition called Convexity of the Distribution Function then ensures that the agent's first order condition is sufficient for an optimum[16]. However, MLRP and CDFC, especially when imposed together , have been found to be quite restrictive. In contrast, monotonicity is a natural property of incentive contracts under asymmetric information (Guesnerie and Laffont (1984)). Nonmonotonic contracts are not implementable anyhow.[17]

3.8 Appendix A

Proof of Lemma 3.1. Note first that $e > 0$ is imposed so that $F(\cdot)$ has a density $f(\cdot)$. The case $e = 0$ is neglected anticipating the result that this would be suboptimal from the principal's perspective. For a formal proof of this, see proposition 3.3 (investments) where the result follows trivially from the nature of the incentive problem and proposition 3.8 (learning).

The remainder of the proof is essentially due to Mirrlees (1971): Principal and Agent have quasilinear utilities, so that transfers can be substituted out. Then one can think of the problem as if the principal directly offers an indirect utility level $u(\theta)$ to the agent. From the envelope theorem (by the optimality of the agent's report) it follows that the agent's indirect utility is given by

$$u(\theta) = U(\bar{\theta}) + \int_{\theta}^{\bar{\theta}} q(\tilde{\theta}) d\tilde{\theta}. \tag{3.12}$$

[16] Jewitt (1988) has derived somewhat weaker conditions that validate taking a first order approach.

[17] To see precisely why our conditions are less stringent, note the following: For the investment model we impose a First Order Stochastic Dominance relation (FOSD) in conjunction with the *Concavity* condition, $\int^{\theta} F_{ee}(\tilde{\theta}, e) d\tilde{\theta} \leq 0 \, \forall \theta, e$. Our first condition is unambiguously weaker: FOSD is implied by MLRP but the reverse is not true (Milgrom (1981)). The comparison of CDFC and *Concavity* is more difficult since neither condition implies the other. On the contrary, they each imply that the other, respectively, does not hold. However, there is at least an intuitive sense in which *Concavity* is weaker than CDFC: CDFC requires the distribution function to be convex in the agent's effort at any point θ. In contrast *Concavity* requires that the area under the distribution function be a concave function of the agent's effort. This obviously allows for the case where the distribution function itself can be either concave or convex in effort over some ranges.

Consequently the principal's expected net utility is by $t(\theta) = \theta q(\theta) + U(\bar{\theta}) + \int_{\theta}^{\bar{\theta}} q(\tilde{\theta})d\tilde{\theta}$ equal to

$$\int_{\underline{\theta}}^{\bar{\theta}} \left(V(q(\theta)) - \theta q(\theta) - U(\bar{\theta}) - \int_{\theta}^{\bar{\theta}} q(\tilde{\theta})d\tilde{\theta} \right) dF(\theta, e) \tag{3.13}$$

Noting that $U(\bar{\theta})$ will always be set equal to zero and performing an integration by parts, one can show that (3.13) is equivalent to

$$\int_{\underline{\theta}}^{\bar{\theta}} \left(V(q(\theta)) - \theta q(\theta)\right) dF(\theta, e) - \int_{\underline{\theta}}^{\bar{\theta}} F(\theta, e) q(\theta) d\theta.$$

To establish the reverse reasoning one has to show that $t(\theta) = \theta q(\theta) + \int_{\theta}^{\bar{\theta}} q(\tilde{\theta})d\tilde{\theta}$ together with $q_{\theta}(\theta) \leq 0$ implies the truth telling constraint. This is standard and omitted. ■

Proof of Proposition 3.1. Differentiate the agent's expected rent twice to get:

$$\int_{\underline{\theta}}^{\bar{\theta}} F_{ee}(\theta) q(\theta) d\theta - g_{ee}(e).$$

After an integration by parts the first term by parts to get

$$\left(\int_{\underline{\theta}}^{\theta} F_{ee}(\tilde{\theta}, e) d\tilde{\theta} \right) q(\theta) \Bigg|_{\underline{\theta}}^{\bar{\theta}} - \int_{\underline{\theta}}^{\bar{\theta}} \left(\int_{\underline{\theta}}^{\theta} F_{ee}(\tilde{\theta}, e) d\tilde{\theta} \right) q_{\theta}(\theta) d\theta - g_{ee}(e)$$

Since $q_{\theta}(\theta) \leq 0$ by monotonicity the problem is strictly concave in e. This completes the proof. ■

3.9 Appendix B

3.9.1 Implementation

Proof of Proposition 3.2. By direct inspection of (3.5)

$$\lambda_\theta(\theta) = -\left\{V_q(q(\theta)) - \left(\theta + \frac{F(\theta,e)}{f(\theta,e)} - \frac{\mu F_e(\theta,e)}{f(\theta,e)}\right)\right\} f(\theta,e) \tag{3.14}$$

one observes: when monotonicity is nonbinding $\lambda(\theta) = 0\,\forall\theta$ so that $\lambda_\theta(\theta) = 0\,\forall\theta$ the solution is obtained by pointwise maximization and the conclusion follows. To solve for the multiplier μ^*, differentiate P'' with respect to e. The resulting first order necessary condition is:

$$\begin{aligned}
&\int_{\underline{\theta}}^{\bar{\theta}} \left\{V_q(q(\cdot)) - \left(\theta + \frac{F(\theta,e)}{f(\theta,e)} - \mu\frac{F_e(\theta,e)}{f(\theta,e)}\right)\right\} \frac{\partial q(\cdot)}{\partial e} - \\
&\qquad \int_{\underline{\theta}}^{\bar{\theta}} F_e(\theta,e)q(\theta)d\theta + \int_{\underline{\theta}}^{\bar{\theta}} \left(V(q(\theta)) - \theta q(\theta)\right) dF_e(\theta,e) \\
&\qquad\qquad + \mu\left\{\int_{\underline{\theta}}^{\bar{\theta}} F_{ee}(\theta)q(\theta)d\theta - g_{ee}(e)\right\} = 0
\end{aligned}$$

By (3.5) the first line is zero. To see this, note that $\int_{\underline{\theta}}^{\bar{\theta}} \lambda_\theta(\theta)\frac{\partial q(\cdot)}{\partial e}d\theta = 0$. Over a region of bunching two things are true: the quantity of production is constant, so that its change w.r.t. e is also the same, and the average distortion must be zero. Then, the expression reduces to

$$\mu = \frac{\int_{\underline{\theta}}^{\bar{\theta}} \left(V(q(\cdot)) - \theta q(\cdot)\right) dF_e(\theta,e) - \int_{\underline{\theta}}^{\bar{\theta}} F_e(\theta,e)q(\theta)d\theta}{-\left(\int_{\underline{\theta}}^{\bar{\theta}} F_{ee}(\theta)q(\theta)d\theta - g_{ee}(e)\right)}. \tag{3.15}$$

After an integration by parts of the nominator in (3.15)

$$\int_{\underline{\theta}}^{\bar{\theta}} \left(V(q(\cdot)) - \theta q(\cdot)\right) dF_e(\theta, e) = \left(V(q(\cdot)) - \theta q(\cdot)\right) F_e(\theta, e)|_{\underline{\theta}}^{\bar{\theta}}$$
$$- \int_{\underline{\theta}}^{\bar{\theta}} \left(\left(V_q(q(\cdot)) - \theta\right) q_\theta(\theta) - q(\theta)\right) F_e(\theta, e) d\theta$$

Because supports are nonmoving for all $e > 0$ $F_e(\bar{\theta}, e) = F_e(\underline{\theta}, e) = 0$. After some straightforward manipulation we can reexpress (3.15) as

$$\mu = \frac{\int_{\underline{\theta}}^{\bar{\theta}} \frac{F(\theta,e)}{f(\theta,e)} q_\theta(\theta) F_e(\theta, e) d\theta}{\left\{\int_{\underline{\theta}}^{\bar{\theta}} F_{ee}(\theta, e) q(\theta) d\theta - g_{ee}(e)\right\} + \int_{\underline{\theta}}^{\bar{\theta}} \frac{F_e(\theta,e)}{f(\theta,e)} q_\theta(\theta) F_e(\theta, e) d\theta} \tag{3.16}$$

■

3.9.2 Investments

Proof of Proposition 3.3. Consider first the denominator of (3.16) the term in brackets is the second order condition of the agent's effort choice, hence negative. Since $(F_e(\theta, e))^2 \geq 0$ and $q_\theta(\theta) \leq 0$ the denominator is negative. Consequently

$$sign\,[\mu] = sign \left[- \int_{\underline{\theta}}^{\bar{\theta}} \frac{F(\theta, e)}{f(\theta, e)} q_\theta(\theta) F_e(\theta, e) d\theta\right]$$

Since $F_e(\theta, e) \geq 0$ the conclusion follows. To see the effect on produced quantities, simply observe that $\mu > 0$ and $F_e(\theta, e) \geq 0$. ■

Proof of Proposition 3.4. Obvious from (3.5). ■

Proof of Proposition 3.5. From Proposition 3.4 there is overproduction at the top. By logconcavity and monotonicity of our likelihood measure $\left[\frac{F(\theta,e)}{f(\theta,e)} - \mu \frac{F_e(\theta,e)}{f(\theta,e)}\right]$ is strictly increasing in θ. The conclusion follows. ■

Proof of Proposition 3.6. Observe as a preliminary that $\lim_{\theta \to \underline{\theta}} \frac{F_e(\theta,e)}{f(\theta,e)} = \infty\ \forall e$ and $\lim_{\theta \to \underline{\theta}} \frac{F(\theta,e)}{f(\theta,e)} = 0\ \forall e$ imply that $\lim_{\theta \to \underline{\theta}} \frac{f_e(\theta,e)}{f(\theta,e)} = \infty\ \forall e$. (Henceforth we suppress $\forall e$.)

To see this, note that by $\lim_{\theta \to \underline{\theta}} \frac{F(\theta,e)}{f(\theta,e)} = 0$ it follows that $\lim_{\theta \to \underline{\theta}} f(\theta,e) > \lim_{\theta \to \underline{\theta}} F(\theta,e)$. Consequently, $\lim_{\theta \to \underline{\theta}} \frac{F_e(\theta,e)}{f(\theta,e)} \leq \lim_{\theta \to \underline{\theta}} \frac{F_e(\theta,e)}{F(\theta,e)} = \lim_{\theta \to \underline{\theta}} \frac{f_e(\theta,e)}{f(\theta,e)}$ where the last equality follows by l'Hôptial. The conclusion follows.

The cost of the incentive contract has two components. Additional rent costs and additional production costs. We consider them in turn. Consider first rent costs: Since $\lim_{\theta \to \underline{\theta}} \frac{f_e(\theta,e)}{f(\theta,e)} = \infty$ we can always find a θ close enough to the top such that $\frac{f_e(\theta,e)}{f(\theta,e)} > K$. Consider the following contract:

$$\begin{aligned} q(\theta) &= \hat{q}\, for\, \theta \in \left[\underline{\theta}, \underline{\theta} + \varepsilon\right] \\ q(\theta) &= \tilde{q}(\theta)\, for\, \theta \in (\underline{\theta} + \varepsilon, \bar{\theta}] \end{aligned}$$

where $\tilde{q}(\theta)$ satisfies

$$V_q(\tilde{q}(\theta)) = \theta + \frac{F(\theta, \tilde{e})}{f(\theta, \tilde{e})}$$

and $\tilde{e}$ denotes the effort level that would be optimal if effort was observable. By choosing the contract in the proposition the principal can implement any effort level he wants so that the incentive constraint of the agent becomes

$$\int_{\theta_l}^{\underline{\theta}+\epsilon} F_e(\theta, \tilde{e})\hat{q}d\theta + \int_{\underline{\theta}+\epsilon}^{\bar{\theta}} F_e(\theta, \tilde{e})\tilde{q}(\theta)d\theta - g_e(\tilde{e}) = 0.$$

The additional expected information rent under this contract is

$$\begin{gathered} \int_{\underline{\theta}}^{\underline{\theta}+\epsilon} [\hat{q} - \tilde{q}(\theta)]\, F(\theta, \tilde{e})d\theta = \int_{\underline{\theta}}^{\underline{\theta}+\epsilon} [\hat{q} - \tilde{q}(\theta)] \left(\int_{\underline{\theta}}^{\theta} f(\tilde{\theta}, \tilde{e})d\tilde{\theta} \right) d\theta < \\ \int_{\underline{\theta}}^{\underline{\theta}+\epsilon} [\hat{q} - \tilde{q}(\theta)] \left(\int_{\underline{\theta}}^{\theta} \frac{f_e(\tilde{\theta}, \tilde{e})}{K} d\tilde{\theta} \right) d\theta = \frac{1}{K} \int_{\underline{\theta}}^{\underline{\theta}+\epsilon} [\hat{q} - \tilde{q}(\theta)]\, F_e(\theta, \tilde{e})d\theta = \\ \frac{1}{K} \left\{ g_e(\tilde{e}) - \int_{\underline{\theta}}^{\bar{\theta}} F_e(\theta, \tilde{e})\tilde{q}(\theta)d\theta \right\} \end{gathered}$$

Where the last equality follows from the incentive compatibility condition. Obviously

$$\lim_{K \to \infty} \frac{1}{K} \left\{ g_e(e) - \int_{\underline{\theta}}^{\bar{\theta}} F_e(\theta, \tilde{e}) \tilde{q}(\theta) d\theta \right\} = 0$$

so that the additional rent cost goes to zero. Consider now the additional physical production costs:

$$\int_{\underline{\theta}}^{\underline{\theta}+\epsilon} \theta \left[\hat{q} - \tilde{q}(\theta)\right] f(\theta, \tilde{e}) d\theta.$$

We know that for θ close enough to the top we have $\frac{F_e(\theta,e)}{f(\theta,e)} > M$. Consequently

$$\begin{aligned} \int_{\underline{\theta}}^{\underline{\theta}+\epsilon} \theta \left[\hat{q} - \tilde{q}(\theta)\right] f(\theta, \tilde{e}) d\theta \quad &< \quad \left[\underline{\theta} + \epsilon\right] \int_{\underline{\theta}}^{\underline{\theta}+\epsilon} \left[\hat{q} - \tilde{q}(\theta)\right] f(\theta, \tilde{e}) d\theta < \\ & \frac{\left[\underline{\theta} + \epsilon\right]}{M} \int_{\underline{\theta}}^{\underline{\theta}+\epsilon} \left[\hat{q} - \tilde{q}(\theta)\right] F_e(\theta, \tilde{e}) d\theta \end{aligned}$$

Now the argument can be repeated along the same lines as for the additional rent costs.

■

Proof of Proposition 3.7. As a preliminary observe that we require that both $\lim_{\theta \to \underline{\theta}} \frac{F_e(\theta,e)}{f(\theta,e)} = 0$ and that $\lim_{\theta \to \underline{\theta}} \frac{F(\theta,e)}{f(\theta,e)} = 0$. However, when $F(\theta, e)$ tends to zero much faster than $F_e(\theta, e)$ one may still have $\lim_{\theta \to \underline{\theta}} \frac{F_e(\theta,e)}{F(\theta,e)} = \infty$ which is by l'Hôpital equivalent to $\lim_{\theta \to \underline{\theta}} \frac{f_e(\theta,e)}{f(\theta,e)} = \infty$, what we assume.
Straightforward differentiation of $\theta + \frac{F(\theta,e)}{f(\theta,e)} - \mu \frac{F_e(\theta,e)}{f(\theta,e)}$ reveals

$$\frac{\partial}{\partial \theta} \left(\theta + \frac{F(\theta, e)}{f(\theta, e)} - \mu \frac{F_e(\theta, e)}{f(\theta, e)} \right) \Bigg|_{\theta = \underline{\theta}} = 2 - \mu \frac{f_e(\theta, e)}{f(\theta, e)}$$

This violates monotonicity when

$$2 < \mu \frac{f_e(\theta, e)}{f(\theta, e)}$$

In particular this is true if $\lim_{\theta \to \underline{\theta}} \frac{f_e(\theta,e)}{f(\theta,e)} = \infty$.

If $\theta + \frac{F(\theta,e)}{f(\theta,e)} - \mu \frac{F_e(\theta,e)}{f(\theta,e)}$ is first decreasing then increasing there is only one region of bunching. Consequently there must be overproduction at the top. (If there are several regions of bunching, one candidate solution would set all distortions over the individual bunching regions separately equal to zero. In such a solution there must be overproduction at the top if $2 < \mu \frac{f_e(\theta,e)}{f(\theta,e)}$. However, we must also check that there is no better solution by optimizing over all regions of bunching at the same time. Consequently, if several regions of bunching are present, one cannot be sure that there is overproduction at the top even if $\lambda_\theta(\underline{\theta}) > 0$.)

■

3.9.3 Learning

Proof of Proposition 3.8. (i) By a direct extension to the proof in proposition 3 we have to determine the sign of $-\int_{\underline{\theta}}^{\bar{\theta}} \frac{F(\theta,e)}{f(\theta,e)} q_\theta(\theta) F_e(\theta,e) d\theta$. Since $F_e(\theta,e)$ is positive for some θ and negative for some other θ the sign is ambiguous. However, if P'' is globally concave in e, it suffices to check the derivative of P'' with respect to e at the level of e associated with $\mu = 0$. By an integration by parts

$$
\begin{aligned}
& -\int_{\underline{\theta}}^{\bar{\theta}} \frac{F(\theta,e)}{f(\theta,e)} q_\theta(\theta) F_e(\theta,e) d\theta = \\
& \qquad \int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{\theta}}^{\theta} F_e(\tilde{\theta},e) d\tilde{\theta} \left(\frac{\partial}{\partial \theta} \left[\frac{F(\theta,e)}{f(\theta,e)} \right] q_\theta(\theta) + \frac{F(\theta,e)}{f(\theta,e)} q_{\theta\theta}(\theta) \right) d\theta = \\
& \qquad \int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{\theta}}^{\theta} F_e(\tilde{\theta},e) d\tilde{\theta} \left(\begin{array}{c} \frac{\partial}{\partial \theta} \left[\frac{F(\theta,e)}{f(\theta,e)} \right] \frac{\frac{\partial}{\partial \theta}\left[\theta + \frac{F(\theta,e)}{f(\theta,e)}\right]}{V_{qq}(q(\cdot))} \\ + \frac{F(\theta,e)}{f(\theta,e)} \left(\frac{-V_{qqq}(q(\cdot)) \left\{ \frac{\partial}{\partial \theta} \left[\theta + \frac{F(\theta,e)}{f(\theta,e)} \right] \right\}^2}{V_{qq}^3(q(\cdot))} + \frac{\frac{\partial^2}{\partial \theta^2} \left[\theta + \frac{F(\theta,e)}{f(\theta,e)} \right]}{V_{qq}(q(\cdot))} \right) \end{array} \right) d\theta
\end{aligned}
$$

where all derivatives are evaluated around $\mu = 0$. The sufficient conditions in the proposition guarantee that all entries have the appropriate sign.

(ii) $e = 0$ is optimal for the agent if and only if $q_\theta(\theta) = 0 \, \forall \theta$ so that $q(\theta) = \bar{q}$ and moreover $\bar{t} - \bar{\theta}\bar{q} \geq 0$. However, this contract is very costly to the principal. Consider a small decrease in the transfer, such that only types $\theta \leq \theta^c$ receive positive rent and types with higher costs choose not to participate. The change in the principal's payoff is then

$$\frac{d}{dt}\left(\int_{\underline{\theta}}^{\frac{\bar{t}}{\bar{q}}}\left(V(\bar{q})-\frac{\bar{t}}{\bar{q}}\bar{q}\right)dF(\theta,e)\right)$$

$$= \frac{(V(\bar{q})-\bar{t})}{\bar{q}}dF(\frac{\bar{t}}{\bar{q}},e)dt + \int_{\underline{\theta}}^{\frac{\bar{t}}{\bar{q}}}(V(\bar{q})-\bar{t})\,dF_e(\theta,e)\frac{\partial e}{\partial t}dt$$

$$-\int_{\underline{\theta}}^{\frac{\bar{t}}{\bar{q}}} dF(\theta,e)dt$$

$$= \frac{(V(\bar{q})-\bar{t})\,dF(\frac{\bar{t}}{\bar{q}},e)}{\bar{q}}dt - F(\frac{\bar{t}}{\bar{q}},e)dt + (V(\bar{q})-\bar{t})\,F_e(\frac{\bar{t}}{\bar{q}},e)\frac{\partial e}{\partial t}dt$$

For $\frac{\bar{t}}{\bar{q}} \to \bar{\theta}$ we have $F_e(\theta,e)=0$. Moreover, for t close to $\bar{\theta}\bar{q}$, $\frac{\partial e}{\partial t}$ is finite so that $\lim_{\frac{\bar{t}}{\bar{q}}\to\bar{\theta}}(V(\bar{q})-\bar{t})\,F_e(\frac{\bar{t}}{\bar{q}},e)\frac{\partial e}{\partial t}=0$. Hence, considering $dt>0:$ $\frac{(V(\bar{q})-\bar{t})dF(\frac{\bar{t}}{\bar{q}},e)}{\bar{q}}dt - F(\frac{\bar{t}}{\bar{q}},e)dt$ is nonpositive if $dF(\frac{\bar{t}}{\bar{q}},e)=0$. Since by assumption, for e small enough there is no mass in an interval above the bottom, the conclusion follows: implementing $e=0$ cannot be optimal. ■

Proof of Proposition 3.9. The agent's action induces a mean preserving spread in the distribution of θ. This implies that $F_e(\theta,e)\geq 0$ for some firms in $[\underline{\theta},\theta_1]$ and $F_e(\theta,e)\leq 0$ for some firms in $[\theta_2,\bar{\theta}]$. The conclusion follows easily. ■

Proof of Corollary. trivial and omitted. ■

Proof of Proposition 3.10. First part is identical to proof of proposition 3.4. For the second part, note that if there is no bunching the conclusion follows directly from (3.5). If there is bunching so that $\lambda_\theta(\theta)\geq 0$, by coincidence we could have no distortion at the to. However, then by definition there is bunching. ■

Proof of Proposition 3.11. The first part is identical to proposition 3.6. The second part follows from proposition 3.10. ■

Proof of Proposition 3.12. The first part is identical to proposition 3.7. The second part is in the text. ■

Chapter 4

Financial Contracting, R&D and Growth

4.1 Introduction

The present paper discusses the impact of financial contracting under moral hazard on industry dynamics when technical change is endogenous. Much of the literature on endogenous technical change, e.g. Aghion and Howitt (1992, 1998) or Grossman and Helpman (1991), has industry dynamics involving leapfrogging of incumbents by innovating firms. Such leapfrogging is inconsistent with the observation that in many industries there are a few dominant firms whose dominance persists through decades with significant technical change.

We show that such persistence can be explained by constraints on financial contracting. These constraints provide an absolute advantage to the incumbent firm which can rely on its retained earnings to finance its innovations.

The corporate finance literature suggests that the ability to selffinance is an important determinant of firm investment (Fazzari et.al. (1988)). Some authors suggest that this is because retained earnings themselves are the most important source of finance (Mayer

(1988)).[1] It may also be the case that retained earnings increase the firms' collateral and reduce the agency cost of loan finance (see Fazzari et.al. (1988)) In either case, incumbency provides an absolute advantage in obtaining finance.[2] The paper shows that this advantage may overcome the usual leapfrogging effect and explain the persistence of incumbency.

The basic point is first developed in the context of a static model of a patent race. The model follows Reinganum (1982, 1984) except that players have different initial wealth, whereby the initial wealth position is correlated with the players' starting positions. One player holds a monopoly in the market whereas the other player starts from scratch. The incumbent has a deep pocket, the other player needs outside funds.

Incumbency then has two countervailing effects: since incremental profits are lower for the incumbent, he has less of an incentive to invest in research. This effect - originally due to Arrow - has been termed "replacement" effect. On the other hand the initial monopoly position creates a deep pocket for future research, which is important in a world of imperfect capital markets. Challengers have to contract with outside sources to finance their research expenditures. This contracting is affected by problems of moral hazard familiar from the finance literature (see Jensen and Meckling (1976)): we assume that financial resources as well as the efforts of entrepreneurs are essential inputs in the research process. The latter, however, are unobservable to third parties. In this setup a challenger financed by outside funds has insufficient incentives to take effort to make the venture go. Because outside investors foresee this behavior, they will supply less funds to the firm. Hence the less inside finance a firm has the less it can invest. This moral hazard effect may outweigh the replacement effect so that the incumbent will actually devote more resources to R&D than the challenger.

Subsequently the paper extends the argument to a dynamic setup. If moral hazard effects are strong enough all innovations will come from incumbent firms. This contrasts with the leapfrogging patterns in, e.g., Aghion and Howitt (1992). An interesting finding is that

[1] For a contrary view, see Hackethal and Schmidt (1999).
[2] See also Myers and Majluf (1984).

there may well be a negative impact of moral hazard on growth although actual innovators themselves face no financial problems. The growth retarding effect works through the market: because potential innovators are financially constrained, the competitive threat for the incumbent firms is weak and incumbents can innovate less often without fear of being replaced, i.e. they "can rest on their laurels".

Retardation of growth through agency problems has also been discussed by Aghion, Dewatripont and Rey (1996). In their model managers take insufficient effort for research and development. In the present model insufficiency of incentives for effort taken is also important but this insufficiency arises endogenously from the need of entrepreneurs to obtain outside finance. Here, agency problems not only affect the incentives of potential competitors but also the incentives of incumbents. Given that incumbency persists, the latter is what matters for equilibrium growth.

The persistence of monopolies has also been studied by Gilbert and Newberry (1982,1984). In their analysis persistence is supported by preemptive patenting. Preemption of patents plays no role in the present paper. The point of the analysis here is that even in the absence of preemptive patenting the problems associated with financial contracting can eliminate leapfrogging.

The rest of the paper is structured as follows: The next paragraph presents the model. In section 4.3, the Nash Equilibrium of the game without any financial restrictions is characterized. Section 4.4 introduces moral hazard and financial restrictions into the model. This can change the conclusions on the characteristics of the Nash equilibrium of the game (section 4.5). Section 4.6 extends this result to a dynamic growth model in the spirit of Aghion and Howitt (1992). The final section concludes.

4.2 The model

The model is a simple static equivalent of Aghion and Howitt (1992).There are two firms and two periods. One of the firms, the incumbent, is already in the market making a monopoly profit π_1. The other firm, called the challenger henceforth, initially produces nothing. During period one they both race for some cost reducing innovation which would bring down the constant marginal cost of production from the initial high level $\bar{c}$ to a lower level $\underline{c}$. The innovation is drastic, meaning that $p^m(\underline{c}) < \bar{c}$ or in words: if the successful innovator sets its monopoly price, the previous producer is no more able to compete[3]. Let π_2 denote the profit of the producer with marginal costs $\underline{c}$ setting his monopoly price $p^m(\underline{c})$ in period two. Let $\pi_2 > \pi_1 > 0$.

Research is taking place during the first period. Production with the superior technology cannot begin before the second period. If one of the firms is successful alone, then it receives a patent for the rest of all (model-) time. If nobody is successful in the research lab the incumbent can still produce with his high-cost technology. If both competitors are successful it shall be assumed that the innovation is treated as commonable and hence not patentable. Should this happen, the players engage in Bertrand competition and make zero profits.[4]

Remark 4.1 *Patent races as well as the new growth theories are usually modelled in continuous time. This has the advantage that the event of both players "winning" the race at the same time is of measure zero. Therefore it is accounted for in the payoff functions only with probability zero. The Bertrand assumption ensures, that the event "both win" receives the same "weight" in the present static formulation as it does in continuous time.*

The gross payoffs of the two competitors are thus given by the following matrix: All variables with subscript c denote the challengers choices while subscript I stands for Incumbent.

[3] In case the incumbent is the winner, there are two possible interpretations: One can think of the marginal production costs of the challenger of being either infinity or $\bar{c}$, with both assumptions yielding the same result. In the second case, there is a spillover assumption present, meaning that past innovations are immediately common knowledge.

[4] The purpose of the assumption is explained in the remark. However, one could equally well assume that the patent agency assigns the patent randomly - by flipping a fair coin - should more than one player be successful at the same time. In this case both players would assign expected value $p_c p_I \frac{\pi_2}{2}$ to this event. The key insight is that this influences their incentives symmetrically. All conclusions of the paper hinge on forces that influence the players' incentives asymmetrically. Thus, the alternative formulation would add complexity without increasing the generality of the arguments.

Event	probability	Chall's Payoff	Inc's Payoff
both fail	$(1-p_c)(1-p_I)$	0	π_1
Inc. wins	$(1-p_c)p_I$	0	π_2
Chall. wins	$p_c(1-p_I)$	π_2	0
both "win"	$p_c p_I$	0	0

Research technologies: Decisions about research are once and for all. In order to be successful in the research lab innovators have to spend money I (invest) and effort e. Success probabilities and inputs are linked by the following Cobb-Douglas technology:

$$p_j = be_j^{\alpha} I_j^{1-\alpha}; j = c, I; \alpha \in [0,1]; e, I \geq 0. \tag{4.1}$$

The interest rate (opportunity cost of funds) is zero. Effort is privately costly. Spending effort e generates nonmonetary costs $\frac{e^2}{2}$. b is a scale factor. (see below)

The given situation corresponds to a noncooperative game in which the challenger and the incumbent as strategic players choose research success probabilities (and associated inputs). As a first benchmark, the game will be analyzed under the assumption that both players have deep pockets and moral hazard is not an issue. The payoff functions in this game are then given by 4.2 for the challenger and 4.3 for the incumbent:

$$be_c^{\alpha} I_c^{1-\alpha}(1-p_I)\pi_2 - I_c - \frac{e_c^2}{2} \tag{4.2}$$

$$(1-p_c)\pi_1 + be_I^{\alpha} I_I^{1-\alpha}(1-p_c)(\pi_2 - \pi_1) - I_I - \frac{e_I^2}{2} \tag{4.3}$$

Proposition 4.1 *The strategic game between the challenger and the incumbent, summarized by the payoff functions 4.2 and 4.3 has a Nash equilibrium in pure strategies.*

Proof: A success *probability* must satisfy the restriction $p_j \leq 1; j = c, I$. Let $\lambda_j; j = c, I$

be the multiplier on this constraint and let δ_I^e and δ_I^I be the multipliers on the constraints $e_I, I_I \geq 0$. The solution to the optimization problem of the incumbent must satisfy the Kuhn Tucker necessary conditions:

$$Foc_e : (1-\lambda_I)b\alpha e_I^{\alpha-1}I_I^{1-\alpha}(1-p_c)(\pi_2-\pi_1) - e_I + \delta_I^e \stackrel{!}{=} 0 \tag{4.4}$$

$$Foc_I : (1-\lambda_I)b(1-\alpha)e_I^{\alpha}I_I^{-\alpha}(1-p_c)(\pi_2-\pi_1) - 1 + \delta_I^I \stackrel{!}{=} 0 \tag{4.5}$$

$$\delta_I^I I_I = \delta_I^e e_I = \lambda_I(be_c^{\alpha}I_c^{1-\alpha} - 1) \stackrel{!}{=} 0 \tag{4.6}$$

Solving these equations (and the analogous ones for the challenger), gives the best response function of the incumbent and the challenger, respectively

$$p_I = \min\left\{\alpha(1-\alpha)^{\frac{2(1-\alpha)}{\alpha}} b^{\frac{2}{\alpha}}(\pi_2-\pi_1)^{\frac{2-\alpha}{\alpha}}(1-p_c)^{\frac{2-\alpha}{\alpha}}, 1\right\} \tag{4.7}$$

$$p_c = \min\left\{\alpha(1-\alpha)^{\frac{2(1-\alpha)}{\alpha}} b^{\frac{2}{\alpha}}\pi_2^{\frac{2-\alpha}{\alpha}}(1-p_I)^{\frac{2-\alpha}{\alpha}}, 1\right\} \tag{4.8}$$

which define a continuous map $[0,1]^2 \to [0,1]^2$. This map must have a fixed point: strategy sets are nonempty compact convex subsets of R^2, the payoff functions are continuous in the opponents actions and concave in the own actions. Hence theorem 1.2 Fudenberg and Tirole (1995)applies.[5]■

4.3 Properties of the Nash-equilibrium

- Assumption **A4.1**: $b < \frac{1}{\pi_2}; \pi_2 - \pi_1 > 2; \pi_1 > 0$.
- *Notation*: To save on space, let henceforth $\theta := \alpha(1-\alpha)^{\frac{2(1-\alpha)}{\alpha}} b^{\frac{2}{\alpha}}$; $\epsilon := \frac{2-\alpha}{\alpha}$; $c_c :=$

[5]The theorem is due to Debreu (1952), Glicksberg (1952) and Fan (1952). See Fudenberg and Tirole (1995) for the exact references.

$\theta\pi_2^\epsilon$; $c_I := \theta(\pi_2 - \pi_1)^\epsilon$; $\sigma := \frac{\pi_2 - \pi_1}{\pi_2}$; $\mu := \sigma^\epsilon$.

Note that $\pi_1 > 0 \Rightarrow \sigma < 1 \Rightarrow \mu < 1 \Rightarrow c_I = \mu c_c < c_c$.

Remark 4.2 *A4.1 serves two purposes: First, it excludes boundary solutions in 4.7 and 4.8, hence* $0 \le p_j^*(p_i) < 1; i, j = I, c; i \neq j$. *Second, it ensures (see lemma 1 below) uniqueness of the Nash equilibrium. It is important to note, that A4.1 is a sufficient not a necessary condition to reach these ends. Therefore the following results can reasonably be expected to hold for a broader set of parameters. Heuristically, restricting b to be small, just means that however hard you work, you can never be sure to reach success.*

In view of the simplified notation, the best response functions can now be expressed as

$$p_I = c_I(1 - p_c)^\epsilon \tag{4.9}$$

$$p_c = c_c(1 - p_I)^\epsilon \tag{4.10}$$

Appendix A establishes, that **A4.1**$\Rightarrow 1 > c_c, c_I$. Note also that $c_c > c_I$! But then, the solutions to the system 4.9 and 4.10 define the fixed point(s) of a continuous map of $[0, c_c]^2$ into $[0, c_c]^2$.

Lemma 4.1 *Under Assumption A4.1 the Nash equilibrium of the simultaneous move game is unique.*

Proof: According to system 4.9 and 4.10 Nash-equilibria of the game are solutions to[6]

$$p_I = c_I(1 - c_c(1 - p_I)^\epsilon)^\epsilon \tag{4.11}$$

Uniqueness requires that the function $f = c_I(1 - c_c(1 - p_I)^\epsilon)^\epsilon$ has only one fixed point.

[6] Of course, it does not matter, whether we search for a fixed point in p_I or p_c since both formulations contain the same information: once the solution in p_I is found, the solution in p_c can be read directly from the reaction function 4.10.

Consider the slope of f :

$$\frac{\partial f}{\partial p_I} = (1 - c_c(1 - p_I)^{\epsilon})^{\epsilon - 1}(1 - p_I)^{\epsilon - 1}\epsilon c_c \epsilon c_I$$

A sufficient condition for the uniqueness conclusion is that the right hand side of this equation is smaller than one everywhere. This is what we now show: Proceeding term by term, observe first that $p_I \in [0, c_c] \Rightarrow \{1 - c_c(1 - p_I)^{\epsilon}\} \in [1 - c_c, 1 - c_c(1 - c_c)^{\epsilon}]$. Hence $\{1 - c_c(1 - p_I)^{\epsilon}\} < 1$, $\forall p_I \in [0, c_c]$. Together with $\epsilon - 1 > 0$ this implies that $(1 - c_c(1 - p_I)^{\epsilon})^{\epsilon - 1} < 1$.
Next, observe that $p_I \in [0, c_c] \Rightarrow (1 - p_I)^{\epsilon - 1} < 1$. Appendix B finally provides a proof that **A4.1**$\Rightarrow \epsilon c_c, \epsilon c_I < 1; \forall \alpha \in [0, 1]$. $\square$

An analysis of the game without a restriction on a unique equilibrium would be interesting in its own light. The restriction on uniqueness will be discussed right after Proposition 4.2 below. Throughout the analysis, we make repeated use of a result on monotone comparative statics, which shall therefore be restated here for completeness:

Lemma 4.2 *Let $f(p, \mu) : [0, c_c] \times M \to R$, where M is any partially ordered set and where $f(0, \mu) \geq 0$ and $f(c_c, \mu) \leq 0$. Suppose that for all $\mu \in M$, f is continuous in p. (Then there exists a solution to the equation $f(p, \mu) = 0$.) $p_L(\mu) \equiv \inf\{p \,|\, f(p, \mu) \leq 0\}$ is the lowest solution of $f(p, \mu) = 0$ and $p_H(\mu) \equiv \sup\{p \,|\, f(p, \mu) \geq 0\}$ is the highest solution. Suppose further that for all $p \in [0, c_c]$, f is monotone nondecreasing in μ. Then $p_L(\mu)$ and $p_H(\mu)$ are monotone nondecreasing for all $\mu \in M$. If f is strictly increasing in μ, then $p_L(\mu)$ and $p_H(\mu)$ are strictly increasing.*

Proof: Milgrom and Roberts (1994) Theorem 1.

With these two lemmas in hands, we are finally ready to state the central result on the structure of the Nash equilibrium when both players have deep pockets.[7]

[7]This result is known very well from Reinganum (1983). As the method of the proof will be used repeatedly, the result is restated at this point.

Proposition 4.2 *In the Nash Equilibrium of the simultaneous move game, the challenger invests more and exerts more effort and thus has a higher success probability than the incumbent.*

Proof: The proof will proceed in three steps: Step (i) verifies that all conditions in lemma 4.2 are satisfied. In step (ii) we will characterize the equilibrium for the case $\pi_1 = 0$ ($\mu = 1$). Finally in step (iii) we will characterize the way the equilibrium changes if $\pi_1 > 0$ ($\mu < 1$).

Step (i): Consider the function

$$g(p_I, \mu) = \mu c_c(1 - c_c(1 - p_I)^\epsilon)^\epsilon - p_I \tag{4.12}$$

We have

$$g(0, \mu) = \mu c_c(1 - c_c)^\epsilon \geq 0$$

and

$$g(c_c, \mu) = \mu c_c(1 - c_c(1 - c_c)^\epsilon)^\epsilon - c_c \leq 0$$

For all $\mu \in M$ g is continuous in p. We already know, that a solution must exist. Note that uniqueness implies that

$$\sup\{p\,|\,f(p,\mu) \geq 0\} \equiv p_H(\mu) = p_L(\mu) \equiv \inf\{p\,|\,f(p,\mu) \leq 0\}$$

This verifies, that g has the desired properties.□

Step (ii): Suppose, that $\pi_1 = 0$ and hence($\mu(\pi_1 = 0) = 1$). Then the success probability of the incumbent in the unique equilibrium of the game must satisfy

$$g(p_I, 1) = 1c_c(1 - c_c(1 - p_I)^\epsilon)^\epsilon - p_I = 0$$

Clearly, one could find the equilibrium equivalently by looking for a fixed point in p_c. This fixed point must satisfy

$$h(p_c, 1) = c_c(1 - 1c_c(1 - p_c)^\epsilon)^\epsilon - p_c = 0$$

Let $\hat{p}_c$ be a solution to $h(p_c, 1) = 0$ and let $\hat{p}_I$ be a solution to $g(p_I, 1) = 0$. Suppose then that $\hat{p}_c \neq \hat{p}_I$. Since $h(p_c, 1) \equiv g(p_I, 1)$ this contradicts lemma 4.1. Hence $\hat{p}_c = \hat{p}_I$.$\square$
Consider finally the generalized fixed point map, parameterized by $\mu(\pi_1)$:

$$g(p_I, \mu) = \mu(\pi_1)c_c(1 - c_c(1 - p_I)^\epsilon)^\epsilon - p_I$$

Since $\pi_1 > 0 \Rightarrow \mu < 1$ we know that $g(p_I, \mu(\pi_1)) < g(p_I, 1)$ for all $\pi_1 > 0$. By lemma 4.1 and 4.2, we know then that $\hat{p}_I(\mu) < \hat{p}_I(1)$ for all $\pi_1 > 0$. Since step (ii) established symmetry of the Nash equilibrium for $\pi_1 = 0$ and by the fact that success probabilities are strategic substitutes it follows that $\hat{p}_c(\mu) > \hat{p}_I(\mu)$ for all $\pi_1 > 0$.$\blacksquare$

Because the incumbent would like to rest on his laurels, he ends up investing less, spending less effort and he therefore also produces a lower success probability than his rival does. This is because the only way to get in the position to make the higher profit π_2 is to destroy his own monopoly and hence loose π_1. In view of this simple logic, it might seem odd, to impose such strong conditions as Assumption 4.1 to ensure uniqueness. However, without uniqueness one can go no further than the general statement in lemma 4.2: the lowest and the highest fixed point in p_I will both be lower in a game with $\pi_1 > 0$ compared to a game with $\pi_1 = 0$. Thus, the conclusion that the initial condition $\pi_1 > 0$ puts the incumbent at a strategic disadvantage in the R&D race for a drastic product innovation is quite robust and general. However, from this one *cannot* conclude, that *all* Nash equilibria are asymmetric in favor of the challenger. Since it is this asymmetry prediction of the growth and patent race literature that the paper targets, I have chosen to sacrifice generality in favor of clarity of the results.
The asymmetry result has a close parallel in the Schumpeter growth theory. At the

heart of what has been termed the creative destruction mechanism lies the replacement or Arrow effect: because incumbents already make monopoly profits, their incentives to conduct research are always lower than the outsiders' incentives. To make matters simple, models like Aghion and Howitt (1992) or Grossman and Helpman (1991) and almost every other paper in the field, assume a linear (or constant returns to scale) research technology. Together with a free entry condition into the research business this gives a nice Arbitrage condition, which has to be fulfilled at each instant in time. The price of the only resource of the economy, labor, is bid up in the general equilibrium such that incumbents would make losses, were they to do research. All in all this produces the well known bang bang result that only the challengers invest money in research and that it is only a question of when and not of if the incumbent loses his business.

While in Aghion and Howitt it is sure that some challenger will win the race, in the present model this holds only true "on average". This difference in results has mainly two reasons: the partial equilibrium perspective of the present model and the strictly convex effort costs. Although the research technology here is essentially the same as in Aghion and Howitt (1992) - a constant returns to scale function - its implications are strikingly different in the present partial equilibrium model: as long as the rival does not win with probability one the optimizing choices of effort and investment will always be strictly positive, however small they turn out to be. Therefore the incumbent's success probability is bounded away from zero as long as the challenger does not choose to operate at success probability one. Section 4.6 relaxes both assumptions. There is a third difference between the models: the number of entrants. The free entry assumption in the growth literature of course assumes, that there is an infinite amount of challengers lurking around while the present paper has so far been working with only one of them. However, it is immediate to generalize proposition 4.2 to a game with many entrants and one incumbent firm:

Proposition 4.3 *(i)The extended simultaneous move game with many entrants has a unique Nash equilibrium in pure strategies in which each challenger chooses the same probability. (ii)The success probability of each challenger is bigger than the success proba-*

bility of the Incumbent. (iii) As the number of challengers goes out of bounds, p_I goes to zero.

Proof: (i) Payoff functions are still as required in the proof of proposition 4.1. Thus a Nash Existence in pure strategies exists. Consider the best response functions of any two representative (the h_{th} and the i_{th}) out of n Challengers:

$$p_{ch} = c_c(1 - p_{ci})^\epsilon \prod_{\substack{j=1 \\ j \neq i,h}}^{n} (1 - p_{cj})^\epsilon (1 - p_I)^\epsilon \tag{4.13}$$

$$p_{ci} = c_c(1 - p_{ch})^\epsilon \prod_{\substack{j=1 \\ j \neq i,h}}^{n} (1 - p_{cj})^\epsilon (1 - p_I)^\epsilon \tag{4.14}$$

Let $F := \prod_{\substack{j=1 \\ j \neq i,h}}^{n} (1 - p_{cj})^\epsilon (1 - p_I)^\epsilon$ and let $p_{cj}, j = 1, ..., n; j \neq i, h$ and p_I be fixed and exogenous for the moment. Under this restriction p_{ci} must satisfy

$$p_{ci} = c_c(1 - c_c(1 - p_{ci})^\epsilon F)^\epsilon F$$

while p_{ch} must satisfy

$$p_{ch} = c_c(1 - c_c(1 - p_{ch})^\epsilon F)^\epsilon F$$

Let $\hat{p}_{ci}$ and $\hat{p}_{ch}$ be solutions of these equations. But then, since $F < 1$ $\hat{p}_{ci} = \hat{p}_{ch}$ again by lemma 4.1. Since i, h where picked arbitrarily, this must be true for any two of the challengers, hence for all of them. Furthermore since this is true for any $F < 1$ it must also be true for $F = \prod_{\substack{j=1 \\ j \neq i,h}}^{n} (1 - \hat{p}_{cj})^\epsilon (1 - \hat{p}_I)^\epsilon$, those values that are chosen in equilibrium. This proves $p_{ci} = p_{ch} = \bar{p}_c; i, h = 1, ..., n.\square$

(ii) Consider now the best response functions of any challenger i and the Incumbent, again holding fixed the choices of all other challengers at the endogenously determined equilibrium values:

$$p_I = c_I(1-\bar{p}_c)^{\epsilon(n-1)}(1-p_{ci})^{\epsilon} \tag{4.15}$$

$$p_{ci} = c_c(1-\bar{p}_c)^{\epsilon(n-1)}(1-p_I)^{\epsilon} \tag{4.16}$$

The fact that $\hat{p}_I < \hat{p}_{ci} = \bar{p}_c$ then follows from the proof of Proposition 4.2.□

(iii) Take limits in 4.15 and 4.16 as $n \to \infty$, we see, that p_I hits the "zero line" sooner than p_c, because $c_I < c_c$ and $(1-p_c)^{\epsilon} < (1-p_I)^{\epsilon}$ (because $p_I < p_c$). For large enough n the incumbent's success probability is then approximately zero while the challengers' success probabilities are still strictly positive.■

4.4 Moral Hazard: The Impact of Outside Financing Needs of the Challenger

The last section treated both the incumbent and the challenger(s) on equal footing: Both were assumed to have deep pockets or - equivalently - capital markets were assumed to be perfect. However, if a world of imperfect capital markets is considered, it turns out that the model has a natural asymmetry built in: while the incumbent can finance his investments out of retained earnings, the challenger cannot. Assuming further on that the challenger has no wealth at all he has to contract with outside financiers for financial resources. The following set of assumptions is imposed in the sequel:

- Assumption **A4.2**: (deep pocket) $\pi_1 > \alpha(1-\alpha)^{\frac{2-\alpha}{\alpha}} \{b(\pi_2 - \pi_1)\}^{\frac{2-\alpha}{\alpha}}$

A4.2 says that the endogenously chosen height of investment is in all cases, i.e. even if the challenger should abstain from doing research altogether, smaller than the amount of retained earnings the incumbent has.

- Assumption **A4.3**: (Information) Effort choice is not observable and hence not contractible. Apart from this everything else is observable and contractible.

Thus the source of moral hazard is the challenger's effort choice. Note however, that there is no other source of moral hazard. In particular, it can be verified costlessly ex post, wether there was success or not.

- Assumption **A4.4:** (degree of competition) Financial markets are perfectly competitive: Financiers make zero profits in equilibrium.

- **The timing of events:**

$+ - - - - - - + - - - - - - - + \; - - - - - - - + - - - - - - - +$

$t = 0$	$t = \frac{1}{2}$	$t = 1$	$t = 2$
Incumbent produces	Financiers offer contracts	contracts chosen, Research phase (simultaneous)	the winner(s) produce(s)

Remark 4.3 *Financiers, when they offer contracts, have to anticipate the outcome in $t = 1$. Thanks to the uniqueness of the equilibrium this poses no problems.*

4.4.1 First best

The first best can be achieved if the financier is able to observe the effort choice of the challenger. Optimal contracts can and will then be contingent on this effort choice. Of course the contract must maximize the joint surplus. Optimal effort choice and investment levels will therefore be the same as if the challenger owned his business and would therefore be the only residual claimant.

Remark 4.4 *Having a third party, i.e. the financier, brings an element of sequentiality into the game and hence requires a new solution concept for the game. Throughout the analysis we will use weak perfect Bayesian equilibrium as our solution concept. For expositional reasons the precise meaning of the concept in the given game will be discussed in the section on second best contracts. Since we are currently in a world of perfectly enforceable contracts, we can treat w.l.o.g. the financier and the challenger as effectively one agent.*[8].

[8] *Although a schizophrenic one, because he contracts with himself.*

Then, in the simultaneous move game between the Incumbent and this "super-agent", Nash equilibrium has enough bite.

We already know from the last section what the optimum will look like. It therefore suffices here to state the contracts that will sustain this equilibrium. These contracts consist of:

1. an initial amount of money I, the financier gives to the challenger.[9]
2. an amount of effort e, the challenger must exert.
3. a repayment rule, contingent on the challenger's effort choice.

Proposition 4.4 *An optimal contract under complete information between challenger and financier specifies:* (1)$I^* = \alpha(1-\alpha)^{\frac{2-\alpha}{\alpha}} \{b\pi_2\}^{\frac{2}{\alpha}} (1-p_I)^{\frac{2}{\alpha}}$, (2)$e^* = \alpha(1-\alpha)^{\frac{1-\alpha}{\alpha}} \{b\pi_2\}^{\frac{1}{\alpha}} (1-p_I)^{\frac{1}{\alpha}}$

(3)

$$R_{e^*} = \begin{cases} (1-\alpha)\pi_2 & \text{in case of success} \\ 0 & \text{in case of failure} \end{cases} \quad if \quad e = e^*$$

$$R_e = \begin{cases} \pi_2 & \text{in case of success} \\ 0 & \text{in case of failure} \end{cases} \quad if \quad e \neq e^*$$

Proof: Under complete information, an optimal contract must maximize social surplus. This determines e^* and I^*. The rule $R(.)$ is chosen such that the challenger optimally chooses to exert e^* : Since he has no wealth, the harshest punishment to impose on him is to give him nothing in case of $e \neq e^*$. Otherwise, if $e = e^*$, financier and challenger get shares of profits $\{\beta^{fb}, 1-\beta^{fb}\}$ in case of success. β^{fb} is pinned down by the zero profit condition of financiers:

$$p_c^*(p_I)\{1-p_I\}\beta^{fb}\pi_2 = I^*(p_I)$$

[9] Upon assumption, the challenger will not walk away with the money but will invest it.

where $p_c^*(p_I)$ is given by 4.10, while $I^*(p_I)$ is the investment level agreed upon in the contract. Solve this for β^{fb} to conclude that $\beta^{fb} = 1 - \alpha$.$\blacksquare$[10]

Part (1) and (2) are easily understood. The contract just says that the two parties agree to behave as if they were one profit maximizing agent, whose behavior was already derived in the last section. Due to the risk neutrality of the parties all sharing rules with the same expected payoffs for the parties are equivalent for them. The one stated in the proposition is the one with the harshest punishment in case of failure s.t. the limited wealth of the challenger. Because in equilibrium the sharing rule R_e is not directly payoff relevant the financier's individual rationality constraint is exactly binding. Finally, the result on ex post payments, $\beta^{fb} = 1 - \alpha$ is well known for example from macro theory: with a Cobb-Douglas technology, the capitalist will get a share of total output, which corresponds to the elasticity of total output with respect to his input. This in turn is exactly what the exponent $1-\alpha$ measures.[11] Note however, that the result $\beta^{fb} = 1-\alpha$ does not mean, that the parties share the profits according to a completely inflexible rule, i.e. independently of what the Incumbent does. Ex ante, the investment of the financier varies inversely with the success probability of the Incumbent. Therefore the price of money is the higher, the tougher the Incumbents behavior in the research lab[12].

4.4.2 Unobservable effort: Second best contracts

Matters get really interesting, when the effort choice of the challenger is not observable. The new informational assumptions require a new concept of equilibrium of the game:

Definition *A profile of strategies and system of beliefs (s, m) is a perfect Bayesian equi-*

[10] The repayment is expressed as a fraction of realized profits. However, it should be noted that in the simple setting of the present paper - two realizations of the return, either 0 or π_2 - debt and equity cannot be distinguished.

[11] Note that individual rationality of the agent has to be fulfilled too. This is assumed implicitly. It can always be guaranteed, it the reservation utility of the challenger is small enough.

[12] This mechanism has been used by Fudenberg and Tirole (1985) to explain "Predation without Reputation". In the CSV-setup a la Gale and Hellwig (1985) they use, aggressive behaviour of a rich rival reduces the amount of equity of his opponent. To finance a necessary investment of fixed size he therefore has to borrow a bigger amount. C.p. this increases the likelyhood of bankruptcy and consequently the interest rate charged. This can go so far that the financially constrained firm abstains from the business altogether.

librium in extensive form game Γ if it has the following properties: (i) The strategy profile s is sequentially rational given belief system m. (ii) The system of beliefs is derived from strategy profile s through Bayes' rule whenever possible. That is, if for information set h $prob(h\,|s\,) > 0$, then $m(x) = \frac{prob(x|s\,)}{prob(h|s\,)}$ for all $x \in h$.

Remark 4.5 *The order of play is given by the time line. However, the game can be analyzed as if the Incumbent was given the first move, as long as nobody observes this move. Then, the financier(s) and the challenger get the move and act sequentially.*

Perfect Bayesian equilibrium requires for the given game, that:

1. The Incumbents choice of success probability must be optimal, given the financier‘s and the challenger‘s subsequent strategy.

2. For each belief the financier has as to which success probability the Incumbent has chosen, his offer of contract must be optimal for him, given his anticipation of the challenger‘s subsequent strategy.

3. For each belief the challenger has which success probability the Incumbent has chosen, his choice of contract must be optimal given his anticipation of his own subsequent strategy (i.e. choice of effort).

The incumbents problem is exactly the same as before. Hence his best response to each choice of success probability of his rival is still given by 4.9.
Contracts between the financiers can no more be contingent on effort choice, since misbehavior on the part of the challenger cannot be distinguished from bad luck ex post. Therefore, under this new informational assumption, contracts specify:

1. An amount of money, D, the financier hands over to the challenger.

2. A repayment rule, which is *not* contingent on effort choice.

While in the last section, there was only one contract with a unique investment level, there is no reason to believe a priori, that this will also be true in the present context.

Therefore, the repayment rule has to be generalized to:

$$R(D) = \begin{cases} \beta(D)\pi_2 & \text{in case of success} \\ 0 & \text{in case of failure} \end{cases}$$

where the ex post shares are now allowed to vary with the ex ante chosen investment level. Any contract can then be summarized by a tuple $\{D, \beta(D)\}$.[13]

Remark 4.6 *Financiers effectively are in a Bertrand competition with each other. As usual this forces them to offer a contract that maximizes the challengers expected payoff s.t. an Individual Rationality constraint for themselves. So the analysis has to focus on what contract is optimal for the challenger given that weak perfect Bayesian equilibrium strategies are played subsequently. We assume that the financiers make take-it-or-leave-it offers and the challenger either accepts or rejects them.*

Lemma 4.3 *The contracts $\{D, \beta(D)\}$ financiers offer satisfy:*

$$\varphi D^{\frac{-\alpha}{2-\alpha}} \{1 - \beta(D)\}^{\frac{\alpha}{2-\alpha}} \beta(D) - 1 = 0 \tag{4.17}$$

with

$$\varphi := (1 - p_I) b \{\alpha b (1 - p_I) \pi_2\}^{\frac{\alpha}{2-\alpha}} \pi_2$$

Proof: When offering a contract, financiers must have a belief about p_I. Since players are restricted to pure strategies, this belief must put probability mass 1 on some p_I. For any belief financiers might have when they offer a contract, they must anticipate that once a contract has been chosen, the challenger's only remaining degree of freedom is the choice

[13]Since the underlying distribution of events is a two point distribution, it is easy to see, that the repayment scheme in an *optimal* contract must look like this. Moreover, it is well known that we cannot distinguish between debt and equity in this simple world. Expressing the repayment as a share of profits made is therefore without loss of generality.

of effort. This must maximize his expected payoff, given the contract, or more formally:

$$\hat{e} = \arg\max \left\{ (1-p_I) b e^{\alpha} D^{1-\alpha} (1-\beta(D))\pi_2 - \frac{e^2}{2} \right\} \tag{4.18}$$
$$\hat{e}(\beta(D), D) = \left\{ \alpha b D^{1-\alpha}(1-p_I)(1-\beta(D))\pi_2 \right\}^{\frac{1}{2-\alpha}} \tag{4.19}$$

Moreover, the contracts they offer allow them exactly to break even on expectation if and only if the challenger chooses the incentive compatible effort level, i.e. "stays on the equilibrium path". Technically this means:

$$(1-p_I) b \hat{e}^{\alpha} D^{1-\alpha} \beta(D) \pi_2 - D = 0 \tag{4.20}$$

To offer contracts, which are more favorable to them, is meaningless due to the Bertrand assumption. The combination of 4.18 and 4.20 gives the expression in the lemma.■

Thus, Incentive Compatibility of effort choice works as an effective constraint on the set of feasible contracts, financiers should offer, if they want to break even on average. Although 4.17 summarizes all relevant information on contracts, it has the disadvantage of being only an implicit relation $\beta(D)$. To calculate closed form solutions[14] it is more convenient to work with the function $D(\cdot)$ *s.t.*

$$D(\beta) = \varphi^{\frac{2-\alpha}{\alpha}} \left\{1-\beta\right\} \beta^{\frac{2-\alpha}{\alpha}} \tag{4.21}$$

This is an explicit function $D(\beta)$, with the following properties: $D(\beta) \geqslant 0; \forall \beta \in [0,1]; D(0) = D(1) = 0; \frac{\partial D}{\partial \beta}(\frac{2-\alpha}{2}) = \frac{\partial^2 D}{\partial \beta^2}(1-\alpha) = 0$; as shown in figure 4.1:

The nonlinearity present in $D(\beta)$ deserves some economic explanation. Why isn't it the case that financiers offer ever higher amounts of money in return for an ever higher ex post share in the venture? Initially, i.e. for low values of β, this is the case. In this range, a higher β makes them willing to offer a higher amount of outside finance (and investment), since this still has the unambiguous effect of raising the probability of

[14] This causes no problems. See ftn. 16.

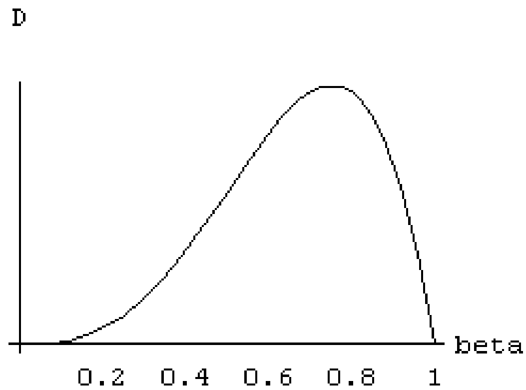

Figure 4-1:

success. However, to reach success, it takes investment of financial resources as well as the efforts of the challenger. Inspection of $\hat{e}(\beta)$ shows that $\hat{e} = 0$ for $\beta = 1$. Why should the challenger work hard if he ends up with nothing in his hands anyway? This logic tells us that $\frac{\partial e^{Ic}(\beta)}{\partial \beta}$ changes it's sign somewhere.[15] It turns out, that the outside finance capacity is reached at $\beta = \frac{2-\alpha}{2}$: $D\left(\frac{2-\alpha}{2}\right)$ is the maximum amount of finance investors are willing to provide: there is no contract with a $D > D\left(\frac{2-\alpha}{2}\right)$ which is both incentive compatible for the challenger and allows financiers to break even.

Not all contracts on the locus in the figure are relevant: consider any contract on the locus with $\beta > \frac{2-\alpha}{2}$: For each tuple $\left\{\acute{\beta}, D(\acute{\beta})\right\}$, $\acute{\beta} \in (\frac{2-\alpha}{2}, 1]$, $\exists$ a tuple $\left\{D(\tilde{\beta}), \tilde{\beta}\right\}$, such that $D(\tilde{\beta}) = D(\acute{\beta})$ but $\tilde{\beta} < \acute{\beta}$. A financier offering contracts of the form $\left\{\acute{\beta}, D(\acute{\beta})\right\}$ will thus never be able to attract any clients[16]. Having identified the locus all feasible contracts must lie on, one can finally solve for the contract which is chosen in equilibrium.

Proposition 4.5 *The second best optimal contract for the challenger is given by* $\left\{\beta^{sb}, D(\beta^{sb})\right\}$ *with* $\beta^{sb} = 1 - \alpha$.

Proof: Financiers, when offering contracts, take into account all subsequent effects on the challengers effort choice, knowing that effort will be chosen in an incentive compatible

[15] see below.

[16] The function $D(\beta)$ is one-to-one only in the range $\beta \in [0, \frac{2-\alpha}{2}]$. Working with $D(\beta)$ or $\beta(D)$ is thus equivalent iff the optimal contract has a $\beta \leq \frac{2-\alpha}{2}$. See Proposition 4.5.

way. From 4.18 and 4.21 we know, that $\hat{e}$ and D are ultimately functions of β alone. Therefore, the optimal contract for the challenger is the solution to[17]:

$$\max_{\beta}(1-p_I)\pi_2\{p_c(e(\beta),D(\beta))(1-\beta)\}-C(e(\beta))$$

From the envelope theorem - because effort choice must be optimal - it follows that

$$\left\{(1-p_I)\pi_2(1-\beta)\frac{\partial p_c}{\partial e}-\frac{\partial C}{\partial e}\right\}\frac{\partial e}{\partial \beta}=0$$

Therefore the FOC to the problem is

$$(1-p_I)\pi_2\left\{(1-\beta)\frac{\partial p_c}{\partial D}\frac{\partial D}{\partial \beta}-p_c(e(\beta),D(\beta))\right\}\stackrel{!}{=}0$$

A few algebraic transformations then show that

$$(1-\beta)(1-\alpha)D'(\beta)=D(\beta) \tag{4.22}$$

$D'(\beta)$ is straightforward to calculate from 4.21. Simplification then delivers the stated result that $\beta^{sb}=1-\alpha.$■

We can finally characterize the weak perfect Bayesian equilibrium of the game (apart from the exact value of p_I^*, which will be derived in the next section):

Proposition 4.6 *The incumbent chooses p_I^*. The financiers belief is given by $m_f(p_I^*)=1$ and he offers the challenger the contract $\{\beta=1-\alpha,\ D(\beta=1-\alpha)=(1-p_I^*)^{\frac{2}{\alpha}}(b\pi_2)^{\frac{2}{\alpha}}\alpha^2(1-\alpha)^{\frac{2-\alpha}{\alpha}}\}$. The challengers belief is given by $m_c(p_I^*)=1$. He accepts and exerts effort $e(1-\alpha)=(1-p_I)^{\frac{1}{\alpha}}(b\pi_2)^{\frac{1}{\alpha}}\alpha^2(1-\alpha)^{\frac{1-\alpha}{\alpha}}$. Off the equilibrium path, the challenger is allowed to have any belief and subsequent sequentially rational strategy.*

Proof: Since players are restricted to pure strategies and the challenger as well as the

[17] This problem is not globally concave for most parameter values. However, it can be verified, that it is quasi-concave and that at the optimum the second order condition for a maximum is fulfilled.

financier can calculate the unique outcome of the game, we must have $m_c(p_I^*) = m_f(p_I^*) = 1$. The optimality of the contract and the effort level has already been proved. The fact that the challenger is allowed to have any belief and associated sequential rational strategy is due to the Bertrand assumption: financiers cannot make more than zero profits. But then they weakly prefer to stick to their equilibrium strategy no matter what the challenger should infer from a deviation on their part.■

There are two things remarkable about the result: First, the Bertrand competition between financiers destroys the usual vulnerability of the perfect Bayesian equilibria to crazy beliefs and strategies off the equilibrium path. This makes the stated results quite robust. Second, the result $\beta^{sb} = \beta^{fb}$ is remarkable. Ex post shares in the final pie are not distorted by asymmetric information; capital's share in output still corresponds to the elasticity of output with respect to financial investment. However, it is somehow clear and it will be shown formally in the next lemma that asymmetric information will not be without its costs.

4.5 Leapfrogging reconsidered

Intuitively, we associate the ability to finance out of retained earnings with some kind of strength. On the other hand, some notion of weakness is associated with the need to take on outside finance. The next lemma formally states, what exactly is to be understood by "weakness":

Lemma 4.4 *Fix any p_I : The best response of the challenger in the second best is always smaller than in the first best.*

Proof: From 4.18 and 4.21 one can easily derive the optimal values for $D(\cdot)$ and $e(\cdot)$ at $\beta = 1-\alpha$: $D(\beta = 1-\alpha) = (1-p_I)^{\frac{2}{\alpha}}(b\pi_2)^{\frac{2}{\alpha}}\alpha^2(1-\alpha)^{\frac{2-\alpha}{\alpha}}$ and $e(1-\alpha) = (1-p_I)^{\frac{1}{\alpha}}(b\pi_2)^{\frac{1}{\alpha}}\alpha^2(1-\alpha)^{\frac{1-\alpha}{\alpha}}$. Plug these values into 4.1 and simplify to get

$$p_c = \alpha\theta\pi_2^{\epsilon}(1-p_I)^{\epsilon}$$

Since $\alpha \in [0, 1]$ this proves the claim.$\square$

Remark 4.7 *It is not obvious how to derive this reaction function: While it is clear that both the financier and the challenger can calculate the outcome of the game, it is not clear that they can coordinate their beliefs for any level of* p_I*. This is however, the implicit assumption in the above derivation.*

Financial contracting thus results (unsurprisingly) in a reduction of investment. This result emerges because the investor foresees, that an agent, who has to share the fruits of his efforts with someone else, is reluctant to exert as much effort as in the first best situation. Therefore the per Dollar price charged for each unit of investment has to be higher. But this raises a question: A priori there would be two possible routes to achieve this end. One could either change, for any given amount of investment, the share of the investor in the final profit-pie. Or, one could leave the division of profits ex post unchanged and change the amount of finance ex ante. We already know from proposition 4.5, that the latter route is taken. An economic explanation for this observation is in order: The incentive compatible effort choice, given by 4.18 reaches it's maximum at $\beta^{sb} = 1 - \alpha$.[18] Note that β^{sb} is located in the range, where 4.21 is still increasing in β. The challenger nevertheless chooses not to contract for a higher amount of finance in exchange for a higher β. Why? If he applied for more finance, the financier would recognize very well, that the challenger effectively commits himself to a *lower* level of effort thereby. Because the financier foresees any such opportunistic behavior on the part of the challenger, he has to demand a disproportionately higher share β in return for investments in excess of $D(1-\alpha)$, or in terms of 4.21, he offers an ever smaller amount of finance D in return for a higher β in the range $\beta \in [1-\alpha, 1]$. Technically this means, that 4.21 is convex in the range $[0, 1-\alpha]$ and concave thereafter. Eventually this goes so far, that the investor is reluctant to offer a higher amount of finance in exchange for a higher share: the outside finance capacity is reached at $\beta = \frac{2-\alpha}{2}$.

[18] The proof of this statement is obvious and therefore omitted.

The reduction in all endogenous values is the more pronounced the smaller α is. What is the intuition for this result? A small α means, that effort has to be raised a lot if the success probability is to be affected significantly. Still, effort is one of the essential factors in the technology, as long as $\alpha > 0$. As explained above, the need to take on outside finance weakens the agents incentives to spend effort. If for example α is very close to one, only a tiny little bit of finance has to be taken on and the reduction in effort choice is only very small. If on the other hand α is very close to zero, then the optimal choice of investment relative to the optimal choice of effort given the cost structure will be rather big. Therefore the reduction the endogenous values will be more pronounced.

In short: Due to asymmetric information all choice variables of the challenger are reduced by the factor α. While in section 4.3, the incumbent was at a strategic disadvantage due to his starting position, the strategic position of the challenger is now weakened by his "shallow" pocket. There are thus now two countervailing forces at work in the model, making it worthwhile to reconsider the statement in Proposition 4.2. Recall that $\sigma = \frac{\pi_2 - \pi_1.}{\pi_2}$ and $\mu := \sigma^{\epsilon}$.

Proposition 4.7 *If* $\alpha > \mu$ *then* $p_{*c}(p_{*I}) > p_{*I}(p_{*c})$; *if* $\alpha < \mu$ *then* $p_{*c}(p_{*I}) < p_{*I}(p_{*c})$; *finally if* $\alpha = \mu$, *then the Nash equilibrium is symmetric.*

Proof: The system of equations that determines the Nash equilibrium of the game is given by:

$$p_c = \alpha\theta\pi_2^{\epsilon}(1 - p_I)^{\epsilon}$$

$$p_I = \mu\theta\pi_2^{\epsilon}(1 - p_c)^{\epsilon}$$

The rest of the proof is obvious since the same logic as in the proof of Proposition 4.2 applies.■

In a sense both players are a bit handicapped. The incumbent, because he wants to rest on his laurels, the challenger because he has to use outside finance with its induced agency costs. The relative magnitude of α and μ determines which of the two reaction functions is shifted more heavily inwards. In words the following mechanism is at work: holding fixed a value for α and π_2, the higher the profit the incumbent already has, the lazier he gets. The other way round: holding fixed a value for π_2 and μ, the heavier the impact of the financial restrictions on the behavior of the challenger, the likelier it gets that he ends up being beaten on average. There is thus a nontrivial interaction between financial and real factors present in the model. Again, it is of interest how the results change, if a game with many entrants is considered:

Proposition 4.8 *(i)The extended simultaneous move game with many entrants has a unique Nash equilibrium in pure strategies in which each challenger chooses the same probability. (ii)If $\alpha < \mu$ the success probability of each challenger is smaller than the success probability of the Incumbent. (iii) As the number of challengers goes out of bounds, p_c goes to zero.*

Proof: analogous to Proposition 4.3.

4.6 Implications for growth

4.6.1 The model

A complete treatment of a fully dynamic model is outside the scope of the present paper. The main results, however, have straight forward extensions in the dynamic formulation. Consider the basic version of the Aghion and Howitt (1992) growth model[19]: The economy is populated by a continuous mass L of workers with linear intertemporal preferences: $u(y) = \int_0^\infty y_\tau e^{-\rho\tau} d\tau$. Each of the workers is endowed with one unit flow of labor, so that L also equals the labor supply. In addition, every worker is endowed with the ability to exert effort, e, in potentially unbounded quantities. ρ is the discount rate. Output

[19]The present model uses only a slight modification of Aghion and Howitt (1992), therefore the description of the model is very brief. For a more extensive description of the model see Aghion and Howitt (1998) chapter 2.

of the final good is produced with technology $y = Ax^\delta$, where $0 < \delta < 1$. One of the individuals, the entrepreneur, is initially the owner of the intermediate goods firm. All other individuals start without wealth. Innovations raise the efficiency parameter A by a constant factor $\gamma > 1$. The technology used in the production of the intermediate good is linear so that x also equals labor demand by that firm. The research technology is still given by 4.1. As the economy has only one resource, labor, *investing means hiring workers. The wage bill, however, has to be paid upfront.* Research costs C_t are now assumed to be linear in effort and labor input: investing an amount D_t generates costs $w_t D_t$, while spending effort e_t generates a cost $c_t e_t$.

Innovations raise the productivity of the intermediate goods sector. As in the previous part, these innovations are drastic. Once an innovation has been successful, the innovator faces the following static problem: the final goods sector will buy the intermediate good until its price equals its marginal value in production or $P_t = A_t \delta x^{\delta-1}$. The monopolist will then determine π_t and x_t such that

$$\pi_t = \max_x P_t(x)x - w_t x = \left(\frac{1-\delta}{\delta}\right) w_t x_t$$

and

$$x_t = \arg\max_x \left\{A_t \delta x^\delta - w_t x_t\right\} = \left\{\frac{\delta^2}{\frac{w_t}{A_t}}\right\}^{\frac{1}{1-\delta}}$$

The focus of the present section lies entirely on steady state behavior. In such a steady state all variables will grow at the same rate and the productivity adjusted wage rate, $\frac{w_t}{A_t}$, will be a constant, ϖ say. It is easy to see that in this case we will also have $\pi_t = A_t \pi(\varpi)$ and $x_t = x$. Assume, w.l.o.g. that $c_t = w_t$. Everything that is needed is that all variables grow with the same rate otherwise one of the inputs would shrink to zero over time. The index t denotes "model time": i.e. in model time, the length of the time interval between two subsequent innovations is 1. In real time, denoted by τ, of course this length is stochastic. Observe finally that there is only one investment opportunity in the economy.

There is thus no possibility to save apart from the research project. Research is the only way to become wealthy!

4.6.2 Optimal research policies

Assume for the moment that the entrepreneur did not have to bother about competitors doing research as well. What would his preferred research policy look like? The owner manager controls the arrival rate of a Poisson process, p_τ. As is shown in Appendix C, linearity of costs $w_t e + w_t D_t$ and constant returns to scale together imply that the arrival rate is linear in D ($p_\tau = b\left(\frac{\alpha}{1-\alpha}\right)^\alpha D_\tau := \lambda D_\tau$) and that costs of research can be expressed as a linear function of D ($C_\tau = \frac{w_\tau}{1-\alpha} D_\tau$) alone. Therefore it is convenient to optimize over D directly rather than over p.

In the steady state all costs and values grow at the same rate. There is no other influence of time other than through the arrival of innovations. Therefore the research intensity is constant during time intervals between innovations. Also because costs and values grow at the same rate, the research intensity will turn out to be independent of time across time intervals.

Following these arguments the owner of the firm faces the following stochastic dynamic programming problem:

$$V(A_t) = \max_{D_t} \left\{ \begin{array}{c} \pi(A_t)\Delta\tau - C(A_t, D_t)\Delta\tau + \\ (1+\rho\Delta\tau)^{-1}\left[\lambda D_t \Delta\tau V(A_{t+1}) + (1-\lambda D_t\Delta\tau)V(A_t)\right] \end{array} \right\}$$

Take limits as $\Delta\tau \to 0$ and recognize that in a steady state $V(A_{t+1}) = \gamma V(A_t)$. Taking into account as well that the research intensity will be constant through time and that every term in the equation can be normalized by A_t we can write (with $\frac{V(A_t)}{A_t} = V(D)$)

$$V(D) = \frac{\pi - C(D)}{\rho - \lambda D(\gamma - 1)}$$

The value of a firm is simply given by operating profits net of research costs, discounted at a rate that lies somewhat below the discount rate. This reflects the fact that innovations bring the possibility of future innovations with them.

Let D^* be the solution to the maximization program of the firm and let V^* the value of the objective function under policy D^*. Then, for any firm that is not yet in the position to make monopoly profits but assumes to stay forever in the industry, once it has entered, V^* represents the (maximal) incentive to do research. All these arguments are conditional on the assumption that the incumbent does not have to care about competitors doing research too. Since the incumbent only cares for the increase in value through the next innovation, which is given by $(\gamma - 1)V^*$, we know that without agency costs this replacement effect will generate again the well known result that industry newcomers have the highest incentives to do research. However, there is a direct extension to proposition 4.7:

Proposition 4.9 *Assume that* $\frac{\alpha^{\frac{\alpha}{1-\alpha}}}{1-\alpha+\alpha^{\frac{2-\alpha}{1-\alpha}}} < \frac{\gamma-1}{\gamma}$. *Then, in the steady state only the incumbent does research.*

Proof: See appendix C.

In this paper we will not treat the case when the condition in the Proposition does not hold. Therefore we impose for the rest of the paper:

- Assumption **A4.5:** let $\frac{\alpha^{\frac{\alpha}{1-\alpha}}}{1-\alpha+\alpha^{\frac{2-\alpha}{1-\alpha}}} < \frac{\gamma-1}{\gamma}$.

Observe that the right hand side of the inequality is increasing in γ, while the left hand side is decreasing in α. Therefore this says again that for α sufficiently small relative to γ challengers will stay out. The proof uses exactly the same arguments as the previous sections: The agency cost of finance increases the costs of producing a given probability of success[20] or decreases the probability of success for any given investment. Financial resources are needed to pay the wage bill upfront. Since labor has to be compensated according to its marginal product, the total wage bill will be a fraction $(1 - \alpha)$ of total

[20]As in Barro and Sala-i-Martin (1994), the argument is, that cost differences can make innovations profitabel for the incumbent. However, in contrast to their paper, cost differences are endogenous here.

value created in the research lab. Perfect competition between financiers then ensures that they exactly recover their costs: again $\beta = (1-\alpha)$. The partial equilibrium analysis of the last section had to use convex effort costs. This assumption can be dropped in the general equilibrium, because there is an additional free variable, the wage rate, which adjusts to a level such that the individual producer is indifferent between all levels of production. The research intensity is then determined by equilibrium in all markets. Workers are perfectly mobile and therefore paid the same wage everywhere. At the wage level that makes the incumbent indifferent between researching or not, the challenger will make losses if the assumption in the proposition holds: there is no contract that is individually rational for both financier and challenger since total costs exceed total expected gains.

There are several things to note about the condition in proposition 4.9: it was derived under the assumption that the challenger, once in the monopoly position will choose an optimal investment level over time such that V indeed reaches its maximum. This is clearly an overstatement of the matter since his effort choices are distorted downwards by the fact, that the initial contract gives the financiers the right to a share of profits of $1-\alpha$ until the indefinite future. The condition is therefore sufficient but not necessary. Note furthermore that for the argument given above it does not matter, what exact form the contract between the challenger and the incumbent takes. One can imagine many repayment schedules that all differ only in their intertemporal allocation of payments. However, ex ante they all share one feature: the expected total amount of repayment. And this is the only argument needed above.

4.6.3 The steady state

Consider now the maximization problem of the incumbent who faces no competition in research. The first order necessary condition for the optimality of his research policy is

$$-C'(D) + (\gamma - 1)\lambda V(D) \stackrel{!}{=} 0$$

(since $V' = 0$). At the optimum this implies that

$$wD = \frac{(1-\alpha)(\gamma-1)\lambda D\pi}{\rho} \tag{Arbitrage}$$

which is again the familiar condition that labor, which is paid according to its marginal product, should receive a share of total final output corresponding to the elasticity of output with respect to its input. Compensations in production and research are such that a worker is indifferent between his two opportunities. The value of the ongoing firm under this policy will then simply be $V = \frac{\pi}{\rho}$. The incumbent just values the profit stream until the indefinite future, knowing that he will do research for every new good until the costs of doing so equal the gains. The remaining share α of the expected increase in value exactly compensates the entrepreneur for his effort costs. To characterize the steady state, combine the Arbitrage condition with the resource constraint of the economy:

$$L = D + x(\varpi) \tag{Labor}$$

In the steady state the labor input into research will then satisfy:

$$1 = \frac{(1-\alpha)b(\frac{\alpha}{1-\alpha})^{\alpha}(\gamma-1)(\frac{1-\delta}{\delta})(L-D^*)}{\rho}$$

Compare this with the condition in the original Schumpeterian model: $b(\frac{\alpha}{1-\alpha})^{\alpha}$ is just the usual arrival rate λ and the factor $(1-\alpha)$ reflects the fact that labor receives only a fraction of total value added in research. Note that the condition above only indirectly reflects the impact of agency problems on growth: agency costs on the part of the newcomers generate the freedom of action for the incumbent to choose his most preferred research policy. For completeness the economy will grow (in real time) with rate

$$g^* = b(\frac{\alpha}{1-\alpha})^{\alpha} D^* \ln\gamma$$

4.6.4 Welfare

Assume that the social planner's objective was to maximize the expected value of consumption y_τ. Assume also that the planner can control the inputs of labor and effort directly but that the entrepreneur has to be compensated for his cost of exerting effort. Given the Poisson nature of the research process and taking into account the resource constraint on labor, the social planners problem is then to maximize

$$U(D) = \frac{(1-\alpha)A_0(L-D)^\delta}{\rho - (\gamma-1)b(\frac{\alpha}{1-\alpha})^\alpha D}$$

The downward adjustment of the value of innovations by the factor $(1-\alpha)$ follows from the fact that the social planner calculates with a shadow cost of effort which corresponds to the marginal product of entrepreneurial effort in the research process. The solution can be written as

$$1 = \frac{(1-\alpha)b(\frac{\alpha}{1-\alpha})^\alpha(\gamma-1)\frac{1}{\delta}(L-D)}{\rho - (\gamma-1)b(\frac{\alpha}{1-\alpha})^\alpha D}$$

There are two differences relative to private optimization: (i) The entrepreneur only cares for the part of additional consumer surplus that he can really appropriate: $(1-\delta)$ instead of 1 in case of the planner. (ii) The social planner has a lower discount rate. He values the fact, that each innovation will "stand on the shoulders of the previous innovations", the intertemporal spillover effect. The entrepreneur values this fact exactly the same way. However, he knows that all private gains from future innovations will be exactly sufficient to cover the private costs, because competition for the scarce resource drives up its price. To sum up, these two forces tend to make the intensity of research lower than the socially optimal one. The original model of Aghion and Howitt has the feature that growth can either be excessive or too slow. In the subset of parameter values that has been analyzed in this subsection growth can never be excessive. This is just because a persistent monopolist would never cannibalize his own rents: the business stealing effect is absent. No formal treatment of the complementary parameter set has been given here in the dynamic setup.

But one can conclude that the tendency for growth to be excessive is reduced there too, because the research intensity will always be smaller as the one in a world without agency costs.

If the social planner wishes to implement an optimal research intensity he faces two options. The first is to subsidize the incumbent's research. The second possibility is to finance newcomers - in *à fonds perdu* fashion - such that the economy is continuously "refreshed" by new firms entering the market. Subsidizing monopolists is certainly hard to defend from a political economic perspective - not least because it undermines the possibilities of entrants even more. But so far we cannot compare these two policy options on pure efficiency grounds. Doing this would afford an explicit characterization of the long run behavior of the economy, i.e. the distribution of wealth of the economy. This is outside the scope of the present paper and left for future work.

4.7 Conclusions

The present paper focussed on the interplay between competition in the research labs and the financial sphere of firms. The central arguments can be repeated in a nutshell: even in a situation, where all "real" incentives are in favor of a destruction of monopolies, we may end up observing a persistence of monopolies. The advantage of being able to finance all investments out of retained earnings can give the incumbent firm a strong strategic advantage vis a vis its opponents.

However, the accumulation of financial funds is not the only way how incumbents can improve upon their strategic position. The accumulation of knowledge is probably at least as important in this respect. In this sense learning and accumulating funds should be seen as complementary explanations for the observed persistence of firms. Empirically the presented model predicts - with the qualifications given in the text - that we should not observe too many vertical innovations from financially restricted industry outsiders in product lines where incumbent firms are already active. However, in industries where the elasticity of research success with respect to noncontractible inputs is high relative to the size of innovations, the real forces, i.e. the Arrow effect, will dominate the moral hazard effects and innovations should come from industry outsiders rather than from incumbents. It should be stressed that both financial restrictions and the vertical nature of innovations are crucial for the results: the model is perfectly consistent with observations like wealthy industry outsiders winning patent races for vertical innovations or financially restricted firms doing horizontal innovations.

4.8 Appendix A

Assumption A4.1 states, that $b < \frac{1}{\pi_2}; \pi_2 - \pi_1 > 2; \pi_1 > 0$. Note that this also implies that $b < \frac{1}{\pi_2 - \pi_1}$. Consider the best response functions as in the text:

$$p_I = \alpha(1-\alpha)^{\frac{2(1-\alpha)}{\alpha}} b^{\frac{2}{\alpha}} (\pi_2 - \pi_1)^{\frac{2-\alpha}{\alpha}} (1 - p_c)^{\frac{2-\alpha}{\alpha}}$$

$$p_c = \alpha(1-\alpha)^{\frac{2(1-\alpha)}{\alpha}} b^{\frac{2}{\alpha}} \pi_2^{\frac{2-\alpha}{\alpha}} (1 - p_I)^{\frac{2-\alpha}{\alpha}}$$

They can be rewritten as

$$p_I = \underbrace{\alpha(1-\alpha)^{\frac{2(1-\alpha)}{\alpha}}}_{w} \underbrace{\{b(\pi_2 - \pi_1)\}^{\frac{2}{\alpha}}}_{x} \frac{1}{\underbrace{\pi_2 - \pi_1}_{y}} \underbrace{(1 - p_c)^{\frac{2-\alpha}{\alpha}}}_{z}$$

Now $\alpha \in [0,1] \Rightarrow w < 1; y < 1$ per assumption; $b < \frac{1}{\pi_2} \Rightarrow b < \frac{1}{\pi_2 - \pi_1}; hence \quad x < 1; z < 1$ because exactly the same argument can be given for the other reaction function.

Note that $b < \frac{1}{\pi_2} \Rightarrow \lim_{\alpha \to 0} \{b(\pi_2 - \pi_1)\}^{\frac{2}{\alpha}} = 0$. Since the limit when $\alpha \to 1$ is no problem anyway, this establishes, that $0 \le p_c^*(p_I), p_I^*(p_c) < 1; \forall \alpha. \Box$

4.9 Appendix B

Consider the slope of the mapping in the text. It is given by: $\epsilon c_I (1 - c_c(1-p_I)^{\epsilon})^{\epsilon - 1} \epsilon c_c (1 - p_I)^{\epsilon - 1}$. It is to be shown, that under Assumption A4.1, this expression can never exceed

1, no matter what value α takes on. Written out explicitly, the slope looks like

$$\frac{2-\alpha}{\alpha}\alpha(1-\alpha)^{\frac{2(1-\alpha)}{\alpha}}b^{\frac{2}{\alpha}}(\pi_2-\pi_1)^{\frac{2-\alpha}{\alpha}}\cdot$$
$$\left\{1-\alpha(1-\alpha)^{\frac{2(1-\alpha)}{\alpha}}b^{\frac{2}{\alpha}}\pi_2{}^{\frac{2-\alpha}{\alpha}}(1-p_I)^{\frac{2-\alpha}{\alpha}}\right\}^{\frac{2(1-\alpha)}{\alpha}}\cdot$$
$$\frac{2-\alpha}{\alpha}\alpha(1-\alpha)^{\frac{2(1-\alpha)}{\alpha}}b^{\frac{2}{\alpha}}\pi_2{}^{\frac{2-\alpha}{\alpha}}(1-p_I)^{\frac{2(1-\alpha)}{\alpha}}$$

Written as condensed as possible:

$$(2-\alpha)^2(1-\alpha)^{\frac{4(1-\alpha)}{\alpha}}\frac{1}{\pi_2-\pi_1}\frac{1}{\pi_2}\left\{b(\pi_2-\pi_1)\right\}^{\frac{2}{\alpha}}(b\pi_2)^{\frac{2}{\alpha}}\cdot$$
$$(1-p_I)^{\frac{2-\alpha}{\alpha}}\left\{1-\alpha(1-\alpha)^{\frac{2(1-\alpha)}{\alpha}}(b\pi_2)^{\frac{2}{\alpha}}\frac{1}{\pi_2}(1-p_I)^{\frac{2-\alpha}{\alpha}}\right\}^{\frac{2(1-\alpha)}{\alpha}}$$

Under Assumption A4.1, we have $(2-\alpha)^2(1-\alpha)^{\frac{4(1-\alpha)}{\alpha}}\frac{1}{\pi_2-\pi_1}\frac{1}{\pi_2} < 1$, since even in the limit when $\alpha \to 0$ $(2-\alpha)^2(1-\alpha)^{\frac{4(1-\alpha)}{\alpha}}$ is bounded above by 4 and $\pi_2 - \pi_1 > 2$ according to Assumption A4.1. Furthermore $\{b(\pi_2-\pi_1)\}^{\frac{2}{\alpha}}$ and $(b\pi_2)^{\frac{2}{\alpha}}$ are each smaller than one due to Assumption A4.1. The second line was already dealt with in Appendix A. This establishes the claim, that the slope of the function in the text cannot exceed 1.□

Note again, that the real difficulties are present at the boundaries of the support of α. Were we not to impose the restriction on b, taking limits as $\alpha \to 0$ would involve products of the type $0\cdot\infty$, which are not well defined. Assumption A4.1 therefore serves the purpose of putting everything on the very safest side.

4.10 Appendix C

Proof of Proposition 4.9:

We have to show that - if the condition in Proposition holds - the challenger makes losses if he does research.

Consider therefore first the incumbents research costs: he combines inputs according to the cost minimizing rule $e = \frac{\alpha}{1-\alpha}D$. His total costs are therefore proportional to $D : C(D) = \frac{w}{1-\alpha}D$. Taking cost minimization into account, we can write $p_I = be^{\alpha}D^{1-\alpha} = b\left(\frac{\alpha}{1-\alpha}\right)^{\alpha}D$.

Let $\lambda := b\left(\frac{\alpha}{1-\alpha}\right)^{\alpha}$. The arrival rate, λD, is thus linear in D. (For completeness: the minimal costs of producing a probability p is given by $C_I(p) = \frac{w}{1-\alpha} b^{-1}\left(\frac{\alpha}{1-\alpha}\right)^{-\alpha} p$.)

Consider now the challenger. To characterize his costs, we must first characterize the optimal contract. The financiers problem is:

$$\max_{e,D,\beta}(1-\beta)be^{\alpha}D^{1-\alpha}V - we$$

s.t.

$$(Comp) \quad wD = \beta be^{\alpha}D^{1-\alpha}V$$

$$(IC) \quad e = \arg\max(1-\beta)be^{\alpha}D^{1-\alpha}V - we$$

where $(Comp)$ follows from perfect competition in the financial market and (IC) from the unobservability of effort. Let $\hat{e}$ be the solution to (IC). Clearly the optimal D^* is the solution to $\max_D b\hat{e}^{\alpha}D^{1-\alpha}V - wD - w\hat{e}$. Consistency with $(Comp)$ requires then that $\beta = 1-\alpha$ as claimed in the text.

One easily infers from (IC) that for any contract (D,β) the challenger's provision of effort will be distorted downwards by a factor $(1-\beta)^{\frac{1}{1-\alpha}} < 1$ compared to the first best level . Finally, for $\beta = 1-\alpha$ his total costs and success probability for a given level of D will be $C(D) = w\left(\frac{1-\alpha+\alpha^{\frac{2-\alpha}{1-\alpha}}}{1-\alpha}\right)D$ and $p_c = b\left(\frac{\alpha}{1-\alpha}\right)^{\alpha}\alpha^{\frac{\alpha}{1-\alpha}}D$, respectively. (Again for completeness, his minimal costs of producing a probability p will be $C_c(p) = w\left(\frac{1-\alpha+\alpha^{\frac{2-\alpha}{1-\alpha}}}{1-\alpha}\right)b^{-1}\left(\frac{\alpha}{1-\alpha}\right)^{-\alpha}\alpha^{\frac{-\alpha}{1-\alpha}}p > C_I(p); \forall\alpha < 1$.)

We are finally ready to compare the incentives of incumbent and challenger, respectively: the Incumbent will do research as long as the marginal gain from investing is higher than the marginal costs or

$$b\left(\frac{\alpha}{1-\alpha}\right)^{\alpha}(\gamma-1)\,VD > \frac{w}{1-\alpha}D$$

Such a constellation can never be part of a general equilibrium because it would generate infinite expected profits. To reestablish general equilibrium the wage rate has to increase until the value of the right hand side equals the value of the left hand side. Let this value be $\bar{w}$. Consider now the challengers problem. In general equilibrium all employers of labor pay the same wage. Hence, at $\bar{w}$

$$b\left(\frac{\alpha}{1-\alpha}\right)^{\alpha}\alpha^{\frac{\alpha}{1-\alpha}}\gamma V D < \bar{w}\frac{\left(1-\alpha+\alpha^{\frac{2-\alpha}{1-\alpha}}\right)}{1-\alpha}D$$

if

$$\frac{\alpha^{\frac{\alpha}{1-\alpha}}}{1-\alpha+\alpha^{\frac{2-\alpha}{1-\alpha}}} < \frac{\gamma-1}{\gamma}$$

and the total costs of investment will exceed the total gains.■

Bibliography

[1] Aghion, P. and Howitt P. (1992) "A Model of Growth through Creative Destruction" *Econometrica* **60**: 323-51

[2] Aghion, P. and Howitt P. (1998) "Endogenous Growth Theory" *MIT Press*, Cambridge MA

[3] Aghion, P., Dewatripont M. and Rey, P. (1996) "Competition, Financial Discipline and Growth" *Document de travail*, Toulouse

[4] Aghion, P. and Tirole, J. (1997) "Formal and Real Authority" *Journal of Political Economy **105**(1)*, 1-29

[5] Armstrong, Mark (1994) "Delegation and Discretion" *Discussion Papers in Economics and Econometrics* No. 9421, University of Southampton

[6] Baron, D. and Myerson, R. (1982) "Regulating a Monopolist with unknown costs" *Econometrica* **50**, 911-930

[7] Barro, R. and Sala-i-Martin, X. (1994) "Quality Improvements in Models of Growth" *National Bureau of Economic Research Working Paper 4610*

[8] Chiang, A.C. (1992) "Elements of Dynamic Optimization" *McGraw-Hill*, Inc. New York, US

[9] Crawford, V. and Sobel, J. (1982) "Strategic Information Transmission" *Econometrica* **50**(6), 1431-1451

[10] Crémer, J. and Khalil, F. (1992) "Gathering Information before Signing a Contract" *American Economic Review* **82**(3), 566-578

[11] Crémer, J., Khalil, F. and Rochet, J.-C. (1998a) "Contracts and Productive Information Gathering" *Games and Economic Behavior* **25**, 174-193

[12] Crémer, J., Khalil, F. and Rochet, J.-C. (1998b) "Strategic Information Gathering before a Contract is Offered" *Journal of Economic Theory* **81**, 163-200

[13] Demski, J. and Sappingtion, D. (1987) "Delegated Expertise" *Journal of Accounting Research* **25**(1), 68-89

[14] Fazzari, S.M., Hubbard, R.G. and Peterson, B.C. (1988) "Financing Constraints and Corporate Investment" *Brookings Papers on Economic Activity* **1**, 141-195

[15] Fudenberg, D. and Tirole, J. (1985) "Predation without Reputation" *mimeo MIT*

[16] Fudenberg, D. and Tirole, J. (1995) "Game Theory", *MIT Press*, Cambridge, MA

[17] Gale, D. and Hellwig, M. (1985) "Incentive Compatible Debt Contracts: the One Period Problem" *Review of Economic Studies* **52**, 647-64

[18] Gilbert, R. and Newbery, D. (1982), "Preemptive Patenting and the Persistence of Monopoly", *American Economic Review* **72**, 514-26

[19] Gilbert, R. and Newbery, D. (1984), "Uncertain Innovation and the Persistence of Monopoly: Comment", *American Economic Review* **74**: 238-42

[20] Grossman, G. and Helpman, E. (1991) "Quality Ladders in the Theory of Growth" *Review of Economic Studies* **58**: 43-61

[21] Grossman, S. and Hart, O. (1983) "An Analysis of thePrincipal Agent Problem" *Econometrica* **51**, 7-45

[22] Guesnerie, R. and Laffont, J.J. (1984) "A Complete Solution for a Class of Agency Problems" *Journal of Public Economics*, 329-369

[23] Gul, F. (2001) "Unobservable Investment and the Hold up Problem" *Econometrica* **69**(2), 343-376

[24] Hackethal, A. and Schmidt, R. (1999) "Financing Patterns: Measurement Concepts and Empirical Results" *Working Paper Series: Finance & Accounting* , No. 33

[25] Holmström, B. (1979) "Moral Hazard and Observability" *Bell Journal of Economics* **10**, 74-91

[26] Holmström, B. (1984), "On the Theory of Delegation" in Boyer, M. and Kihlstrom, R. (eds.), *Bayesian Models in Economic Theory* (Elsevier Science Publishers B.V.)

[27] Jensen, M. and Meckling, W. (1976) "Theory of the Firm: Managerial Behaviour, Agency Costs and Capital Structure, *Journal of Financial Economics* **3**, 305-360

[28] Jewitt, I. (1988) "Justifying the First Order Approach to Principal Agent Problems" *Econometrica* **56**(5), 1177-1190

[29] Johnson, N. and Kotz, S. (1970) "Distributions in Statistics-Continous Univariate Distributions", Vol. 2, *Wiley,* New York

[30] Kessler, A. (1998) "The Value of Ignorance", *Rand Journal of Economics*, 339-354

[31] Laffont, J.J. and Tirole, J. (1986) "Using Cost Observation to Regulate Firms" *Journal of Political Economy* **94**(3), 614-41

[32] Lewis, T. and Sappington, D. (1993a) "Choosing Workers' Qualifications: No Experience Necessary?" *International Economic Review* **34**(3), 479-502

[33] Lewis, T. and Sappington, D. (1993b) "Ignorance in Agency Problems" *Journal of Economic Theory* **61,** 169-183

[34] Lewis T. and Sappington, D. (1997) "Information Management in Incentive Problems" *Journal of Political Economy* **105**(4), 796-821

[35] Lewis, T. and Sappington, D. (1997) "Information Management in Incentive Problems" *Journal of Political Economy* **105**(4), 796-821

[36] Mayer, C. (1988) "New Issues in Corporate Finance" *European Economic Review* **32**, 1167-1188

[37] Milgrom, P. (1981) "Good News and Bad News: Representation Theorems and Applications" *Bell Journal of Economics*, 380-391

[38] Milgrom, P. and Roberts, J. (1994) "Comparing Equilibria", *American Economic Review* **84**(3):441-59

[39] Mirrlees, J. (1971) "An Exploration on the Theory of Income Taxation" *Review of Economic Studies*, 175-208

[40] Mirrlees, J. (1974) "Notes on Welfare Economics, Information and Uncertainty" in *Essays on Economic Behaviour and Uncertainty*, ed. by M. Balch, D. McFadden and S.-Y. Wu, Amsterdam North Holland

[41] Myers, S. and Majluf, N. (1984) "Corporate Financing and Investment Decisions when Firms have Information that Investors do not have" *Journal of Financial Economics* **13**, 187-221

[42] Reinganum, J. (1983) "Uncertain Innovation and the Persistence of Monopoly", *American Economic Review* **73**, 741-47

[43] Reinganum, J. (1984), "Uncertain Innovation and the Persistence of Monopoly: Reply", *American Economic Review* **74**, 243-46

[44] Rogerson, W. P. (1982) "Contractual Solutions to the Hold up Problem" *Review of Economic Studies* **59**, 777-793

[45] Rogerson, P. (1985) "The First Order Approach to Principal Agent Problems" *Econometrica*, 1357-1368

[46] Rothschild, M. and Stiglitz, J. (1970) "Increasing Risk I: a Definition", *Journal of Economic Theory* **2**, 225-243

[47] Szalay, D. (2000) "Optimal Delegation", *mimeo University of Mannheim*

[48] Szalay, D. (2001) "Endogenous Ignorance in Agency Problems", *mimeo University of Mannheim*

[49] Tirole, J. (1999) "Incomplete Contracts: Where Do We Stand?" *Econometrica* **67**(4), 741-781

Hiermit erkläre ich, dass ich die Dissertation selbständig angefertigt habe und mich anderer als der in ihr angegebenen Hilfsmittel nicht bedient habe, insbesondere, dass aus anderen Schriften Entlehnungen, soweit sie in der Dissertation nicht ausdrücklich als solche gekennzeichnet und mit Quellenangaben versehen sind, nicht stattgefunden haben.

Mannheim, 21. Mai 2001