ESSAYS IN MICROECONOMICS

Inaugural-Dissertation
zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften
durch die
Rechts- und Staatswissenschaftliche Fakultät
der Rheinischen Friedrich-Wilhelms-Universität
Bonn

vorgelegt von
Philipp Weinschenk
aus Fellbach

Bonn 2009

Dekan:	Porf. Dr. Christian Hillgruber
Erstreferent:	Prof. Martin Hellwig, Ph.D.
Zweitreferent:	Prof. Paul Heidhues, Ph.D.

Tag der mündlichen Prüfung: 04.11.2009

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn
(http://hss.ulb.uni-bonn.de/diss_online) elektronisch publiziert.

ACKNOWLEDGEMENTS

I thank Martin Hellwig for his valuable and precise advice. The interaction with Paul Heidhues taught me a lot about industrial organization and behavioral economics. Jörg Budde's topics course brought me and my co-authors solid knowledge about principal agent models and the basis for our common paper.

I thank my co-authors and friends Fabian Herweg and Daniel Müller for the close and excellent collaboration. I benefited a lot from our discussions and their help. Despite his aversion to industrial organization, Felix Bierbrauer helped me a lot during my studies.

I am grateful for the comments and suggestions from: Patrick Bolton, Stefanie Brilon, Christoph Engel, Jos Jansen, Botond Kőszegi, Michael Kurschilgen, Matthias Lang, Tymofiy Mylovanov, Klaas Schulze, Nora Szech, and numerous seminar participants. I thank Brian Cooper for correcting my texts on several occasions. I enjoyed and still enjoy the research atmosphere at the Bonn Graduate School of Economics and the Max Planck Institute for Research on Collective Goods.

I am grateful to my parents that they not only begot me, but also constantly supported me. I thank my girlfriend ♡Andrea Wittke♡ for her lovability and encouragements.

Contents

I INTRODUCTION **1**

II PERSISTENCE OF MONOPOLY AND RESEARCH SPECIALIZATION **7**

II.1 INTRODUCTION 7
II.2 MODEL 10
II.3 ANALYSIS 11
II.3.1 BERTRAND PROFITS 11
II.3.2 RESEARCH DECISIONS 12
II.3.3 RESULTS 14
II.4 DISCUSSION 16
II.4.1 EMPIRICAL EVIDENCE 16
II.4.2 THE ROLE OF ASSUMPTION A1 17
II.4.3 COMPARISON TO THE LITERATURE 18
II.4.4 TIMING AND ROBUSTNESS 19
II.5 WELFARE ANALYSIS 20
II.6 EXTENSION: AUCTION SETTING 22
II.7 CONCLUSIONS 23

III ENTRY AND INCUMBENT INNOVATION **25**

III.1 INTRODUCTION 25
III.2 MODEL 28
III.3 ANALYSIS 30
III.3.1 BERTRAND COMPETITION AND ENTRY . . . 30
III.3.2 INVESTMENT 30
III.4 ALTERNATIVE TIMING 34

III.5 HOW IMPORTANT ARE THE INCUMBENT'S INITIAL COSTS? 37
III.6 CONCLUSIONS 38

IV AMBIGUITY IN A PRINCIPAL-AGENT MODEL 41
IV.1 INTRODUCTION 41
IV.2 MODELS 44
IV.2.1 THE STANDARD LEN MODEL 44
IV.2.2 THE MODEL WITH AMBIGUITY AND AMBIGUITY AVERSION 45
IV.3 ADDITIONAL INFORMATION: VIOLATION OF THE INFORMATIVENESS PRINCIPLE 51
IV.4 DISCUSSION 52
IV.4.1 WHEN PARTIES "AGREE TO DISAGREE" . . . 52
IV.4.2 MULTIPLE PRINCIPALS AND COMPETITION . 53
IV.4.3 AGENT'S WELFARE 53
IV.4.4 HISTORY DEPENDENCE 54
IV.5 CONCLUSIONS 54

V THE OPTIMALITY OF SIMPLE CONTRACTS: MORAL HAZARD AND LOSS AVERSION 57
V.1 INTRODUCTION 57
V.2 MODEL 66
V.3 ANALYSIS 71
V.3.1 TWO POLAR CASES: PURE RISK AVERSION VS. PURE LOSS AVERSION 77
V.3.2 THE GENERAL CASE: LOSS AVERSION AND RISK AVERSION 84
V.4 IMPLEMENTATION PROBLEMS, TURNING A BLIND EYE, AND STOCHSTIC CONTRACTS 90
V.4.1 THE CASE OF A BINARY MEASURE OF PERFORMANCE 90
V.4.2 INVALIDITY OF THE FIRST-ORDER APPROACH 91
V.4.3 TURNING A BLIND EYE 93
V.4.4 BLACKWELL REVISITED 95

V.5 ROBUSTNESS, EXTENSIONS, AND CONCLUSION . 97

VI TECHNOLOGY OF SKILL FORMATION AND HIDDEN INFORMATION 101

VI.1 INTRODUCTION 101

VI.2 MODELS 102

VI.2.1 CUNHA AND HECKMAN'S MODEL 102

VI.2.2 THE MODEL WITH HIDDEN INFORMATION . . 103

VI.3 CONCLUSIONS & DICUSSION 106

VII APPENDICES 109

VII.1 APPENDIX TO CHAPTER II 109

VII.2 APPENDIX TO CHAPTER III 117

VII.3 APPENDIX TO CHAPTER IV 125

VII.4 APPENDIX TO CHAPTER V 127

VII.5 APPENDIX TO CHAPTER VI 146

I. INTRODUCTION

This thesis contains five papers. Two belong to the field of industrial organization. Both are on innovation and the investment in it. The paper "Persistence of Monopoly and Research Specialization" takes a new look at monopoly persistence and explores which effects determine it. In "Entry and Incumbent Innovation", we investigate how the threat of entry influences the innovation activity of an incumbent. Two papers belong to the field of contract theory. They are on principal-agent relationships with moral hazard. In "Ambiguity in a Principal-Agent Model", we assume that the agent's knowledge about the statistical properties of the performance measure is ambiguous and that the agent is averse to ambiguity. We investigate how ambiguity and ambiguity aversion change the use of information and the power of the incentives which are optimally set. In "The Optimality of Simple Contracts: Moral Hazard and Loss Aversion", Fabian Herweg, Daniel Müller, and I explore how an agent's loss aversion changes the complexity of the optimal contract. The last paper is on the economics of education and human capital formation. In "The Technology of Skill Formation and Hidden Information", we consider a model of child development, where the formation of human capital occurs in multiple stages via investments. We explore how hidden information about how to treat a young child best changes the optimal investment plan. In the remainder of the introduction, we will explain the papers and their results in more detail.

In Chapter II, we investigate the persistence of monopolies in markets with innovations. The extensive literature on the subject has discussed this issue in terms of the efficiency effect and the replacement effect. Since competition destroys profits, the efficiency effect predicts that the incumbent's incentive to remain a monopolist through innovating is at least as great as the entrant's incentive to become a duopolist. The replacement effect predicts the opposite: the entrant's incentive to innovate is higher than the incumbent's because only the incumbent takes into account that innovating replaces its existing technology. According to the classical work of Gilbert and Newbery (1982), the efficiency effect determines the outcome, whereas in a seminal paper by Reinganum (1983) it is the replacement effect.

We build a unifying model in which both effects are present. In our model, the first-moving incumbent may be able to discourage the potential entrant from investing in research by investing itself. The outcomes of research activities are uncertain. Therefore, preemption is less than perfect. For high success probabilities, we obtain a result in the spirit of Gilbert and Newbery (1982): preemption is almost perfect and so the efficiency effect is the driving force. For low success probabilities the incumbent can hardly preempt and so the replacement effect predominates. This result is in the spirit of Reinganum (1983). A rough intuition is that research is a powerful preemption device if and only if it is likely to succeed.

The former results imply that research with a high success probability is more likely done by the incumbent than by the potential entrant, and it is vice versa for research with a low success probability. In this sense, incumbents specialize in "safe" research, and potential entrants in "risky" research. We also show that research undertaken by potential entrants is, on average, "riskier" than that of incumbent firms. Moreover, the probability of entry has an inverted U-shape in the success probability of research. Since even at the peak the probability of entry is only a quarter, the persistence of monopoly is high. We also explore the normative aspects of our model and show that, apart from one exception, firms never overinvest and may underinvest. When the incumbent preempts the potential entrant and the innovation is non-drastic, overinvestment may occur.

In Chapter III, we explore how the threat of entry influences an incumbent's investments in R&D. This question is important since innovations are central to growth. The literature has found two counteracting effects. First, the Schumpeterian effect. A larger entry threat reduces the incumbent's expected profit and therefore also its investment. Second, the entry deterrence effect. To deter entry, or to make entry unlikely, a greater threat requires a larger investment. Combining both effects yields that the incumbent's investment is hump-shaped in the entry threat. When the entry threat is small and increases, the incumbent invests more to deter entry or to make entry unlikely. This is due to the entry deterrence effect. However, when the threat becomes huge, entry can no longer profitably be deterred or made unlikely and the investment becomes small. Then the Schumpeterian effect dominates. We show that these results are robust to different timings. In one time structure, the incumbent does not know the rivals' production costs when deciding about investment. In the alternative timing, the costs are known.

Chapter IV is motivated by the following empirical findings: wage schemes sometimes do not depend on performance or the dependence is rather weak. Additionally, the Informativeness Principle, according to which the principal uses all information in a compensation contract which is correlated with performance, is often violated. These findings are in contrast to the theoretical literature. We show that considering ambiguity and ambiguity aversion can bring theory in line with empirics.

There is also a conceptual justification to consider ambiguity. In standard models of principal-agent relationships, it is implicitly assumed that the agent knows precisely the statistical properties of the performance measure. However, we think that this is a rather strong—and in real life often unrealistic—assumption. We relax it by assuming that the agent faces ambiguity with respect to the performance measure. Due to ambiguity, the agent's beliefs about the distribution of the shock on the performance measure are not represented by a single probability function, but instead by a set of probability functions. Since Ellsberg (1961), at the very latest, we know that subjects are averse to ambiguity. Therefore we assume that the agent is not only risk-averse, but also ambiguity-averse. We show that with ambiguity the agent is pessimistic about the distribution of the shock on the performance measure. As a consequence, the compensation demanded by the agent is relatively high, compared to the standard model (which neither considers ambiguity nor ambiguity aversion). Because the principal has to ensure participation of the agent, this implies that the principal's cost of providing incentives is relatively high. As a result, the principal sets relatively weak incentives which yield a relatively low expected payoff. It may even be the case that the optimal contract is a fixed wage.

When there are two performance measures, it can be optimal for the principal to ignore one of them (and potentially both), even though both measures are informative. The reason is that with ambiguity, the inclusion of a measure into a wage scheme causes costs which are not negligible even when the wage depends on the measure only to a small extent. Hence, the Informativeness Principle does not hold in our model.

In Chapter V, Fabian Herweg, Daniel Müller, and I consider a principal-agent relationship with moral hazard. Empirically, wage schemes sometimes consist of remarkably few different levels. Sometimes there are even only two levels: a base wage and a lump-sum bonus. The observed simplicity of contracts, however, is at odds with predictions made by economic theory. We show that considering loss aversion can solve this puzzle.

Our model is standard but for one twist: the agent is assumed to be loss-averse in the sense of Kőszegi and Rabin (2006, 2007). With the tradeoff between incentive provision and risk-sharing being at the heart of moral hazard, allowing for a richer description of the agent's risk preferences that goes beyond standard risk aversion seems a natural starting point to gain deeper insights into contract design. Our main finding is that a simple lump-sum bonus scheme is optimal when loss aversion is the driving force of the agent's risk preferences. This is in stark contrast to the findings for a standard risk-averse agent. The intuition is as follows: an agent who is risk- but not loss-averse exhibits local risk-neutrality. This implies that paying slightly different wages for different signals improves incentives at negligible cost. Therefore simple contracts cannot be optimal. With a loss-averse agent, this is no longer true. With the reference point being multidimensional under the concept of Kőszegi and Rabin, the agent is first-order risk-averse at all possible wage levels. In consequence, paying even slightly different wages reduces the agent's expected utility, for which in turn he demands to be compensated. Thus, by offering a simple contract that specifies only few different wage levels, the principal can lower the expected payment necessary to compensate the agent.

In Chapter VI, we contribute to the literature on the formation of human capital. Cunha and Heckman (2007) consider an economic model of child development, where the formation of human capital occurs in multiple stages via investments. They solve for the optimal intertemporal investment plan, which is important to derive policy implications. We extend their framework by assuming that children are differentiated in the sense that a child's type determines what type of investment is most productive for him/her, and that this information is not available when a child is young. When a child is old, the type is revealed.

How does the optimal investment plan change as a result of hidden information? There are two intuitive guesses. (i) It is optimal to invest less in the first and a more in the second phase of childhood, because in the second one can tailor the investments to a child's type and therefore yield a high return of investment. (ii) It is optimal to invest more in the first and less in the second phase to make sure that, despite low returns in the first phase, the effective investment in the first phase is not "too bad". We show that the answer crucially depends on the substitutability of investment between phases: when investments are easily substitutable (easier than Cobb-Douglas), intuition (i) is right; when substitution is difficult (more difficult than Cobb-Douglas), (ii) is right. More specifically, hidden information weakens the importance of early investments in children when inter-phase investments are easily substitutable, but strengthen them when substitution is difficult.

II. PERSISTENCE OF MONOPOLY AND RESEARCH SPECIALIZATION

We examine the persistence of monopolies in markets with innovations when the outcome of research is uncertain. We show that for low success probabilities of research, the incumbent can seldom preempt the potential entrant. Then the efficiency effect outweighs the replacement effect. It is vice versa for high probabilities. Moreover, the incumbent specializes in "safe" research and the potential entrant in "risky" research. We also show that the probability of entry has an inverted U-shape in the success probability. Since even at the peak entry is rather unlikely, the persistence of the monopoly is high.

II.1. INTRODUCTION

This paper takes a new look at monopoly persistence in markets with innovations. The extensive literature on the subject has discussed this issue in terms of the efficiency effect and the replacement effect. Since competition destroys profits, the efficiency effect predicts that the incumbent's incentive to remain a monopolist through innovating is at least as great as the entrant's incentive to become a duopolist. The replacement effect (Arrow, 1962) predicts the opposite: The entrant's incentive to innovate is higher than the incumbent's because only the incumbent takes into account that innovating replaces its existing technology. According to the classical work of Gilbert and Newbery (1982; henceforth GN) the efficiency effect determines the outcome, whereas in a seminal paper by

Reinganum (1983; henceforth RE) it is the replacement effect.[1] We build a unifying model in which both effects are present. We allow for uncertainty with respect to the outcomes of innovative activities and show that the success probability of research determines the relative strength of both effects.

We consider a monopolized market, where the first-moving incumbent may be able to discourage the potential entrant from investing in research by investing itself.[2] The outcomes of research activities are uncertain. Therefore, preemption is less than perfect. For high success probabilities we obtain a result in the spirit of GN: preemption is almost perfect so the efficiency effect is the driving force. Intuitively, since the success probability is high the potential entrant's expected profit from research greatly depends on the incumbent's research decision. Hence, it is very likely that the incumbent can and does preempt the potential entrant. For low success probabilities the same argument applies in reverse, i.e., the incumbent can hardly preempt and so the replacement effect predominates. This result is in the spirit of RE.

These results imply that research with a high success probability is more likely done by the incumbent than by the potential entrant, and it is vice versa for research with a low success probability. In this sense, incumbents specialize in "safe" research, and potential entrants in "risky" research. We also show that research undertaken by potential entrants is, on average, "riskier" than that of incumbent firms. Moreover, the probability of entry has—at least roughly—an inverted U-shape in the success probability of research. Since even at the peak the probability of entry is only a quarter, the persistence of monopoly is high.

We also explore the normative aspects of our model. We consider the second best world in which pricing cannot be regulated and show that, apart from one exception, firms never overinvest and may underinvest.

[1]See also the debate in Reinganum (1984) and Gilbert and Newbery (1984).

[2]The idea that a dominant firm might use its investment decision as a strategic device to persuade a potential entrant not to enter stems from Spence (1977) and Dixit (1980). They consider capacity investments.

When the incumbent preempts the potential entrant and the innovation is non-drastic, overinvestment may occur.[3] Intuitively, this holds when in case that incumbent's research is successful (i) the monopoly price is almost the same as when the incumbent would have the old technology (so that the consumer surplus is hardly increased) and (ii) the incumbent's profit only slightly improves relative to the investment costs.

The research process considered by GN is commonly interpreted as an auction. As an extension, we integrate such an auction process into our model. This changes our results: regardless of the success probability, the incumbent will always outbid the entrant if the innovation is non-drastic. So entry will never occur. This replicates GN's result in a more general framework which allows for uncertainty of the research process.

Our paper is related to the literature on the persistence of monopoly in markets with innovations, which is surveyed by Gilbert (2006). The relation of our model to GN and RE is discussed later. Denicolo (2001) and Etro (2004) consider a research process of the RE type where the replacement effect disappears, since the aggregate R&D effort is independent of the incumbent's decision. Fudenberg and Tirole (1986) show that in RE's model, when the innovation is non-drastic, the efficiency effect may outweigh the replacement effect. With different outcomes being possible, there is, however, no clear-cut result.[4] Our model delivers clear and intuitive results without any assumption on whether the innovation is drastic or not.

We offer a novel explanation to the question why entrants do riskier research than incumbents. Existing literature on this question emphasizes other explanations. While Kihlstrom and Laffont (1979) look at differences in the risk-attitudes of firms, De Meza and Southey (1996)

[3]An innovation is called drastic if it is so large that the innovative entrant is effectively unconstrained by incumbent's competition. It can charge monopoly prices and yield monopoly profits.

[4]See Fudenberg and Tirole (1986, Ch. 3) and Tirole (1988, pp. 397-398). Also Beath, Katsoulacos, and Ulph (1989a, 1989b) show that both effects can play a role in a model similar to RE's. But again, no clear and simple results can be yielded (1989a, p. 167).

consider excessive optimism of entrepreneurs. Scherer and Ross (1990, Ch. 17) blame the bureaucracy in large companies. Baumol (2004) highlights educational differences between researchers in incumbent firms and entrepreneurs that engage in research. In Rosen (1991) the ex ante high-cost firm must spend more than the ex ante low-cost firm to yield the same cost level. Through this asymmetry, the former chooses a riskier research project than the latter.

We present, analyze, and discuss the model in Sections II.2, II.3, and II.4, respectively. A welfare analysis is in Section II.5. After considering an auction setting in Section II.6, we conclude in Section II.7. Proofs are in the Appendix.

II.2. MODEL

There are two firms, an incumbent I, and a potential entrant E. At stage 1 the incumbent decides whether or not to invest in a firm specific research project. Investing causes expected costs of $k > 0$ and yields an innovation with probability $p \in (0, 1]$. At stage 2 firm E faces the same decision. In order to focus on the replacement and the efficiency effect we set both firms on equal footing and assume that both firms' projects have the same costs and success probabilities. At stage 3 nature independently determines whether each firm's project is successful or not. A successful firm gets a process innovation that enables production at per-unit costs of $\underline{c}$. If I does not invest or its project fails, it can produce at per-unit costs of $\bar{c}$ by using its old technology, where $\bar{c} > \underline{c} > 0$. In contrast, if E does not invest or its project fails, it cannot produce at all. Finally, at stage 4, firms compete à la Bertrand. For reasons that will become clear later, we assume that there is also a stage 0 where first the success probability p is drawn from density g, and then the cost k is drawn from conditional density h.[5]

Firms are risk neutral and cannot collude. There is perfect information. The solution concept is subgame perfect Nash equilibrium. Con-

[5]When k is drawn before or simultaneously to p, we can ignore the k value until p is drawn, and so preserve the vision that p is drawn first.

sumer demand is given by the function $D(\phi)$, where ϕ is the consumer price. We assume that $D(\phi)$ is falling in ϕ, positive for $\phi = \bar{c}+\epsilon$ (where ϵ is small and positive), and $D(\underline{c})$ is finite. These assumptions allow us to borrow Tirole's (1988) analysis of Bertrand profits; see below. To yield clear-cut normative results we have to assume that monopolist's optimal price is unique.[6]

II.3. ANALYSIS

We solve the model by backward induction. In this section we first describe the Bertrand profits of firms. Then we determine their research decisions. Finally we present the results.

II.3.1. BERTRAND PROFITS

Bertrand profits are uniquely determined by the firms' production costs and therefore can be expressed as $\pi^J(c^I, c^E)$, where $J \in \{I, E\}$, $c^I \in \{\bar{c}, \underline{c}\}$, and $c^E \in \{\underline{c}, -\}$. The symbol "–" indicates that E cannot produce at all. We normalize the maximal profit $\pi^I(\underline{c}, -)$ to 1. The following lemma on the structure of firms' profits is due to Tirole (1988).[7]

LEMMA 1:

(i) $\pi^E(\bar{c}, -) = \pi^E(\underline{c}, -) = \pi^E(\underline{c}, \underline{c}) = \pi^I(\underline{c}, \underline{c}) = \pi^I(\bar{c}, \underline{c}) = 0$,

(ii) $1 > \pi^I(\bar{c}, -) > 0$,

(iii) $1 \geq \pi^E(\bar{c}, \underline{c}) > 1 - \pi^I(\bar{c}, -)$.

Part (i) describes the well-known result that when a firm cannot produce or has the same or even higher per-unit costs than its competitor its Bertrand profit is zero. Part (ii) states that a monopolist is strictly better off with low than with high per-unit costs; nonetheless, a monopolist with high per-unit costs makes a positive profit. The first inequality of

[6]Hermalin (2009) offers some weak assumptions on $D(\cdot)$ that guarantee inter alia uniqueness; see his Proposition 3. The key assumption is that $D(\cdot)$ is log-concave.

[7]Tirole partially summarizes existing literature. He does not consider the Nash equilibria found by Blume (2003) where one firm plays a weakly dominated strategy; see Tirole (1988, p. 234, footnote 37).

part (iii) contains the *efficiency effect*: since competition destroys industry profits, I's incentive to remain a monopolist through innovating [and yield a Bertrand profit of 1] is at least as great as E's incentive to become a duopolist [which yields a Bertrand profit of $\pi^E(\bar{c}, \underline{c})$]. However, when I takes into account that its old technology is replaced when it innovates [the net-effect of innovating on its Bertrand profit is just $1 - \pi^I(\bar{c}, -)$] E's incentive to innovate is higher than I's. This is the *replacement effect* which is captured by the last inequality. When the first weak inequality of (iii) is strict we say that the innovation is *non-drastic*; in case of equality the innovation is called *drastic*.

II.3.2. RESEARCH DECISIONS

Firm J's research decision, $J \in \{I, E\}$, is denoted by $a^J \in \{0, 1\}$, where 0 denotes no investment and 1 investment. We assume that in case of indifference a firm does not invest.[8]

Potential Entrant's Research Decision.— Since a^I is either 0 or 1, there are two subgames. Letting $b^E(a^I)$ denote E's best responses to a^I, we have

$$b^E[0] = 1 \iff k < \pi^E(\bar{c}, \underline{c})p =: \bar{k}(p); \tag{II.1}$$

$$b^E[1] = 1 \iff k < \pi^E(\bar{c}, \underline{c})(1-p)p =: \underline{k}(p). \tag{II.2}$$

Thus, E invests if and only if the costs k are sufficiently low. More specifically, if $k < \underline{k}(p)$, then $a^E = 1$ is E's dominant strategy and when $k \geq \bar{k}(p)$ the dominant strategy is $a^E = 0$. For $k \in [\underline{k}(p), \bar{k}(p))$ we have $b^E(0) = 1$ and $b^E(1) = 0$.

Incumbent's Research Decision.— I chooses the optimal action foreseeing E's later responses. If $k \geq \bar{k}(p)$, then

$$a^I = 1 \iff k < (1 - \pi^I(\bar{c}, -))p =: \hat{k}(p); \tag{II.3}$$

if $k < \underline{k}(p)$, then

$$a^I = 1 \iff k < (1 - \pi^I(\bar{c}, -))(1-p)p =: \tilde{k}(p); \tag{II.4}$$

[8]To rule out that cases of indifference drive our results we will later introduce an assumption that guarantees that cases of indifference have a measure of zero.

and if $k \in [\underline{k}(p), \bar{k}(p))$, then

$$a^I = 1 \iff k < p. \tag{II.5}$$

Equilibrium.— Using the previous formulas the construction of the equilibria is straightforward. Figure II.1 shows the equilibrium research decisions, which we denote by $a^* = (a^{I*}, a^{E*})$. The following lemma summarizes formally.

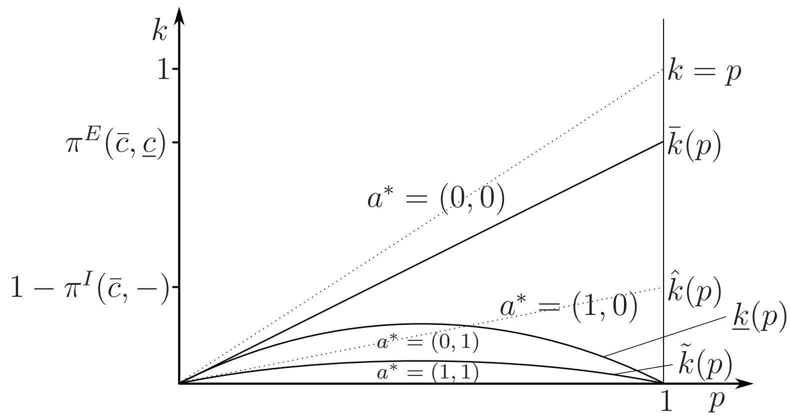

Figure II.1: Research decisions in equilibrium.

LEMMA 2: The firms' research decision in equilibrium are

- $a^* = (1, 1)$ for $k \in [0, \tilde{k}(p))$;
- $a^* = (0, 1)$ for $k \in [\tilde{k}(p), \underline{k}(p))$;
- $a^* = (1, 0)$ for $k \in [\underline{k}(p), \bar{k}(p))$;
- $a^* = (0, 0)$ for $k \geq \bar{k}(p)$.

Intuition.— Investments are strategic substitutes. In the parameter area where $k \in [\underline{k}(p), \bar{k}(p))$ E invests if and only if I has not invested. Hence, I can preempt E. But is it profitable for I to preempt? Yes it is, due to the efficiency effect: in this parameter area $k < p$, i.e., (II.5) holds. Roughly speaking, I prefers the risk to replace its old technology itself to the risk of being replaced by E. Hence in this parameter set the equilibrium is $a^* = (1, 0)$. Note, in this parameter set regardless whether I invests or not, exactly one firm will invest and so eventually make I's old technology obsolete. Consequently, I does not take into account that its existing technology may be replaced by its *own* investment and the replacement effect plays no role.

For low p values preemption is possible only for a relatively small range of costs, compared to the range of costs where at least one firm invests. The line of arguments is as follows. Observe that the probability that only E receives the innovation when it invests—and thus receives a positive Bertrand profit—hardly changes through I's decision. Therefore, E's expected profit from research hardly depends on I's decision. Hence, preemption is seldom possible. In contrast, for high p values, preemption is mostly possible; the arguments stated before apply in reverse. So research is a powerful preemption device if and only if it is likely to succeed.

Due to the replacement effect, there is a set of parameters where E is willing to invest, irrespective of what I has done, but where I is no longer motivated to invest, given that E will invest. Hence in this set $a^* = (0, 1)$. As explained before, when the success probability p is high, E's expected profit from research—and so its willingness to invest—is very sensitive upon I's investment decision. Hence the replacement effect loses its power when p becomes large. In the extreme case of $p = 1$, it has no power at all: E never invests when I has invested.

When the costs are very low both firms are always willing to invest and hence $a^* = (1, 1)$. In the remaining parameter set, costs are so high that $a^* = (0, 0)$.

II.3.3. RESULTS

From Figure II.1 or the arguments made before it is intuitive that when the success probability p is low the replacement effect is "more important" than the efficiency effect, whereas for high p values it is vice versa. To formalize this intuition, we assume that prior to the firms' investment decisions the costs k and the success probability p are drawn. This is kind of a comparative statics analysis which allows us to determine how "important" the different equilibria and effects are.

In order to obtain concrete results we make the following assumption which says that k is uniformly distributed in the "relevant" set.

A S S U M P T I O N A 1 : The conditional density of k, $h(k|p)$, is uniform in k for $(p, k) \in \mathcal{S} := \{(p, k)|0 < p \leq 1, 0 < k \leq \bar{k}(p)\}$.

With this assumption we can establish our main result which is about the probability of investments, conditional on p. Note, we take the perspective that k is not yet drawn.

P R O P O S I T I O N 1 : Suppose A1 holds. If p is sufficiently high, I is more likely to invest in research than E. It is vice versa if p is sufficiently low.

Intuitively, given that a high p value was drawn, it is much less likely that a k will be drawn which lies in the small interval where $a^* = (0, 1)$ rather than in the large interval where $a^* = (1, 0)$; see Figure II.1. For a low p value it is vice versa. Put differently, when the success probability is high it is likely that the incumbent preempts the potential entrant and so that the efficiency effect determines the outcome. This is not true when the success probability is low. Then the replacement effect predominates. So both effects are important in our model, and the success probability determines their relative power.

Taking another view by looking at a large number of I-E-pairs, Proposition 1 predicts that most of the "risky" research is done by potential entrants but not by incumbents. Incumbents on the other hand specialize in "safe" research, which is undertaken rarely by potential entrants.[9]

Perhaps not surprising—albeit nontrivial to prove—is that potential entrants' research is "riskier" than that of the incumbent. We measure the likelihood of failure when neither p nor k is yet drawn. Additionally to A1 we assume that p is distributed with positive and non-atomic density.

P R O P O S I T I O N 2 : Suppose A1 holds and that $g(p)$ has full support and is finite $\forall p \in (0, 1]$. Then E's research is more likely to fail than I's.

The sketch of the proof is as follows: First we show that the likelihood that I invests relative to the likelihood that E invests is increasing in p.

[9] We say that research is *"risky"* (*"safe"*) when p is low (high). This denomination is cheeky because when p is low the outcome of the research process can be less uncertain than when p is high.

Then we show that this implies that the distribution of p, conditional that I invests, first-order stochastically dominates the distribution of p, conditional that E invests. This enables us to finally prove that E's research is more likely to fail than I's.

What remains to determine is the persistence of monopoly. When E has invested and is successful, it competes with I on the market. Then the monopoly does not persist. Otherwise the monopoly persists. We can deliver the following result about the probability of entry (i.e., non-persistence of the monopoly), conditional on p.

PROPOSITION 3: Suppose A1 holds. Then the probability of entry is at most $p(1-p)$.

The intuition is simple: When p is low, research is seldom successful and entry rarely occurs. When p is high, on the other hand, I preempts E except for a small interval of cost realizations; see Proposition 1. Therefore entry is unlikely, too. For intermediate values of p, however, it is likely that E invests and is successful. Hence, the probability of entry has roughly an inverted U-shape in p.[10] By Proposition 3, the probability of entry is at most $1/4$, which implies that the probability for the monopoly to persist is at least $3/4$.[11] Hence we conclude that the persistence of the monopoly is high.

II.4. DISCUSSION

II.4.1. EMPIRICAL EVIDENCE

Baumol (2004) finds that "risky" research is most often done by entrepreneurs and not by incumbents, and that it is vice versa for "safe" research.[12] This fits Proposition 1. Also Vinod Koshla notes that "[r]isk

[10]Ignoring the probability mass outside $\mathcal{S}$ (this is possible when one redefines the densities g and h accordingly) yields an exact inverted U-shape.

[11]The reason why the persistence is not lower in our model when the replacement effect is powerful is that the replacement effect is only strong when research often fails—and failure of potential entrant's research is another reason for persistence.

[12]See also Scherer and Ross (1990, p. 653).

and acceptance of failure are central to innovation, [...] but the dinosaurs typically avoid both" (Statement in *The Economist*, 2007, p. 3). For industries in which failures are common, our model predicts that most of the research is done by potential entrants and not by incumbents. This seems to be the case, for example, in the IT sector.

Proposition 2, which states that E's research is more likely to fail than I's, is supported by Astebo's (2003, p. 227) finding that

> "the average probability that an independent inventor succeeds in commercialising his/her invention is estimated to about 0.07. In comparison, the probability of commercial success of conducting R&D in established firms is approximately 0.27,"

where the later value is from Mansfield et al. (1977). Further evidence is provided by Baumol (2004) and Bianchi and Henrekson (2005, p. 367).

Empirically, the persistence of monopolies seems to be high (Geroski, 1995), as Proposition 3 predicts.

II.4.2. THE ROLE OF ASSUMPTION A1

Scherer and Ross (1990) elucidate how the costs of R&D are distributed. However, since we use expected costs and normalize them we cannot use their insights. To us, A1 seems a natural starting point. This assumption is not as restrictive as it may seem for several reasons. First, A1 assumes that $h(\cdot)$ is uniform in k for parameters of the set $\mathcal{S}$, but $h(\cdot)$ may still depend upon p. Second, when the average density of k conditional on p in the different equilibrium sets of $\mathcal{S}$ is the same, the proofs, and so also our results, stay unchanged.[13] Third, A1 is sufficient, but not necessary, for our results.

To see the last point, suppose that A1 does not hold. Proposition 1 also holds under the alternative, weak assumption: the average density

[13]More technically this means that for all $p \in (0,1]$, $\bar{h}(k \in a^* = (1,0)|p) = \bar{h}(k \in a^* = (0,1)|p) = \bar{h}(k \in a^* = (1,1)|p)$, where $\bar{h}(k \in a^* = i|p)$ is the average conditional density of k when k is in equilibrium set i.

of k, conditional on p and that either only I or E invests, is bounded between two positive constants.[14] Additionally observe that then the probability of entry approaches 0, as p approaches 0 or 1. So at least roughly the probability of entry has an inverted U-shape in the success probability of research; cf. Proposition 3.

A1 is important for Proposition 2. When A1 does not hold, the likelihood that I invests relative to the likelihood that E invests can be locally decreasing in p. Hence, the result stated in Proposition 2 can reverse. From Figure II.1 it is, however, intuitive that for "many distributions" of k and p the result holds.[15]

II.4.3. COMPARISON TO THE LITERATURE

In our model each firm possess one idea, which can be interpreted as a box. Each box contains either nothing or the innovation.[16] To open its box a firm has to invest.[17] This conception implies that the success probability of research is exogenous, and if a box turns out empty there

[14]Formally, $\bar{h}(k \in a^* = i|p)$ has an infimum and a supremum which are in $\mathbb{R}^{++}$ for all $i \in \{(1,0),(0,1)\}$ and for all $p \in (0,1]$. To understand why this condition is sufficient consider the following example. The supremum is twice as large as the infimum. Look at Figure II.1 and determine the p value for which the interval of k values with equilibrium $a^* = (1,0)$ is twice as large as the interval with $a^* = (0,1)$. When a p is drawn which lies above this critical value we can be sure that it is more likely that the costs k will lie in the interval with $a^* = (1,0)$ than in the interval with $a^* = (0,1)$. So it is more likely that I invests in research than that E does. One can also easily construct a lower critical value of p and show that it is vice versa when p is low enough.

[15]Think of a joint probability distribution lying over Figure II.1. Ignore the density when I does not invest. Then calculate I's center of mass. Make the same steps for E. For "many distributions", I's center is further to the right than E's.

[16]With this interpretation nature determines success or failure already at Stage 1. Since firms do not know the realization until Stage 3 this modification of the timing does not change the model in any way.

[17]This description is in line with Scotchmer's (2004) statements that "[a]n innovation requires both an idea and an investment in it" (p. 39) and that "some research efforts do not pay off with certainty ... [and] failures obviously cannot be identified in advance" (p. 40).

is no way for that firm to get the innovation. Since each box can be opened at most once, the game is not repeated.[18]

The models differ greatly with respect to uncertainty. GN consider no uncertainty in the research process. In RE the uncertainty effectively concerns only the arrival date of the innovation, because the game is repeated unless one firm is successful. In our model, on the other hand, it is uncertain whether a firm's idea is realizable; see above. This type of uncertainty is extremely important in reality, see Freeman and Soete (1997, Ch. 8), Scotchmer (2004, pp. 40, 55), or DiMasi (2001).

In our model research is a powerful preemption device if and only if it is likely to succeed. In contrast, in GN preemption is always possible, in RE never. When preemption is possible, it is worthwhile due to the efficiency effect. Consequently, the efficiency effect is the driving force in GN, in our model when the success probability is high, and does not play a role in RE. When preemption is not possible, the efficiency effect is not important, and the replacement effect steps in. Hence it predominates in RE, in our model when the success probability is low, and not at all in GN. Moreover, the probability that a monopoly persists is below one-half in RE, equal to one in GN, and between three-quarters and one in our model. Hence, regarding the importance of the relevant effects and the persistence of monopoly, we take a position between RE and GN.

II.4.4. TIMING AND ROBUSTNESS

We assumed that firms decide sequentially about investing in research. When instead they decide simultaneously, preemption is not possible, and our results are no longer valid. While arbitrary from a theoretical point of view, the assumption of a sequential investment game is not unusual in the literature and also has an intuitive appeal: First, in contrast to

[18] An interesting extension would be that firms have several boxes. However, even when one assumes that each firm can open at most one box the analysis gets cumbersome because the number of cases multiplies. Another interesting extension would be that firms can manipulate the type of their box, or that they can influence upfront what type they likely will receive.

the incumbent, the potential entrant might need some time to gather information about the market or to obtain funding. Second, Freeman and Soete (1997, p. 202) argue that "a firm which is closely in touch with the requirements of its customers may recognize potential markets". So the incumbent but not the potential entrant, may be quicker in developing new ideas. Third, with the incumbent being already prominent among market participants, its activity may be visible for everyone, while the entrepreneur's may not. Hence, only the incumbent may be able to credibly preannounce its investment decision.

In an earlier version of the paper we considered a different timing, where the potential entrant observes incumbent's success or failure before it decides about its investment. This does not change our results substantially. The same is true for the following extensions: (i) heterogenous research costs or success probabilities, (ii) patents,[19] (iii) Cournot competition, (iv) product innovations, and (v) correlated success probabilities.

II.5. WELFARE ANALYSIS

In Section II.3 we analyzed the positive aspects of our model. Now we explore the normative implications. In the second best world where prices cannot be regulated we seek to answer the question whether there is too much or too little investment from a welfare point of view. The literature (see Tirole 1988, p. 399) has found two counteracting effects. First, there is the nonappropriability of social surplus effect: firms may underinvest because the innovator typically does not receive the whole

[19]Our non-extended model can be interpreted in two ways: (i) There are no patents and each firm keeps details of its innovation secret so that an outsider cannot imitate. This interpretation is empirically justified because "patents are regarded as a necessary incentive for innovation in only a few industries" (Cohen 1995, p. 227). See also Scotchmer (2004, Ch. 9). (ii) There are patents but both firms innovations are different in the sense that each firm can get a patent on its technology. Additionally note, that there are patents in RE and GN is not crucial for their results: Without patents and with Bertrand competition a firm no longer wants to engage in research when its competitor was already successful.

social surplus created by its innovation. Second, firms may overinvest due to the business stealing effect: the innovator may not take into account that it steals the rival's business.

PROPOSITION 4: When the innovation is non-drastic and the incumbent preempts the potential entrant then firms may overinvest. In all other cases firms do not overinvest and may underinvest.

We first give the intuition for the case where the innovation is non-drastic and the incumbent preempts the potential entrant. When incumbent's research is successful it may set a price which is almost the same as when it would have the old technology. So the consumer surplus is hardly affected through the innovation. That is, the nonappropriability of social surplus effect is weak. Observe that through the threat of entry, the incumbent is "forced" to steal its own business. Hence the business stealing effect is powerful and may dominate. Put differently, no investment of both firms may be socially desirable.

In contrast, when the innovation is drastic, the successful incumbent sets a much lower price than it would set without the innovation. So the expected consumer surplus increases greatly and the nonappropriability effect dominates. Similar arguments apply for the case where only the potential invests.

One may presume that when both firms invest this may not be socially desirable: research effort is duplicated and so both firms may yield the innovation. This suspicion is false. When both firms are successful the consumer price is only $\underline{c}$, which results in a dominant nonappropriability effect.

Subsidies.— Suppose that the only policy instrument of a government is a research subsidy. Through subsidies the government can change firms' investment decisions, since firms determine their investments on the basis of the net costs. Proposition 4 shows that subsidies are especially relevant to support drastic innovations because for these innovations firms sometimes underinvest but never overinvest.

Targeting subsidies to potential entrants and not to incumbents has two potential advantages. First, the equilibrium $a^* = (0, 1)$ is socially weakly preferred to $a^* = (1, 0)$.[20] Second, promising a subsidy to E in case that it invests can push I to preempt E. Hence, the subsidy is not paid. Nonetheless a previously unexplored research project may now be investigated.

II.6. EXTENSION: AUCTION SETTING

GN's research process is commonly interpreted as a first-price auction with complete and perfect information (Reinganum, 1984). Next, we integrate such an auction setting and show that our results change substantially.

Suppose I and E bid for the service of a firm which implements a research project for the winner.[21] The auction is held before it is clear whether the research project will be successful.[22] A firm's valuation is its willingness to pay for victory, i.e., the difference in its expected profit between winning and losing.

PROPOSITION 5: The incumbent always wins if the innovation is non-drastic.

So with an auction setting and a non-drastic innovation there never is entry. This result coincides completely with GN.[23] Intuitively, since I can always outbid E, preemption is always possible. Given that the

[20] When the innovation is non-drastic the preference is strict: total welfare is higher when there is a duopoly in which E has the innovation than in a monopoly where I has it.

[21] This need not be taken literally. GN's interpretation is that the firms are in a race, and the firm which invests most wins. Another is that firms compete for scarce and essential resources, and so only the firm which invests the most gets them.

[22] The alternative timing is that the auction is held afterwards. Then either a project with a success probability of one or zero is auctioned. In the latter case holding an auction is superfluous. The former case is a just special case in the setting of the original timing. Hence, the alternative timing needs no separate investigation.

[23] We allow for an uncertain research process. GN consider uncertainty only verbally, but it is not clear to us what type of uncertainty they mean. Yi (1995) couples an auction with RE's model, and his result is that the entrant will never do research.

innovation is non-drastic, then, by virtue of the efficiency effect, I's valuation is strictly higher than E's, and preemption is indeed worthwhile. In contrast, when the innovation is drastic firms' valuations are the same and one has to specify a tie-breaking rule. However, if there is only a bit of uncertainty whether an innovation is indeed drastic, I's valuation is higher than E's, and so I will win the auction. Hence generically, E never does research, and entry never occurs. This insight is new.

II.7. CONCLUSIONS

We presented a simple model in which both, the replacement and the efficiency effect are present. We showed that research is a powerful preemption device if and only if it is likely to succeed. This results in the predominance of the efficiency effect when the success probability of research is high and the predominance of the replacement effect when it is low.

III. ENTRY AND INCUMBENT INNOVATION

We explore how the threat of entry influences the innovation activity of an incumbent. We find that the incumbent's investment is hump-shaped in the entry threat. When the entry threat is small and increases, the incumbent invests more to deter entry, or to make it unlikely. This is due to the entry deterrence effect. However, when the threat becomes huge, entry can no longer profitably be deterred or made unlikely and the investment becomes small. Then the Schumpeterian effect dominates.

III.1. INTRODUCTION

Even though innovations are central to growth, the question whether more competition leads to greater R&D investments is not settled. While we do not try to answer this general question, we seek to explore the more specific question how the threat of entry influences an incumbent's investments in R&D. We build a simple model that captures two important but counteracting effects. First, a Schumpeterian effect. A larger entry threat reduces the incumbent's expected profit and therefore also its investment. Second, an entry deterrence effect.[1] To deter entry, or to make entry unlikely, a greater threat requires a larger investment.

Combining the effects, we find that the incumbent's investment is

[1]For the importance of entry in the United States, see Aghion and Howitt (2006, p. 279). Entry deterrence is empirically relevant: "Most R&D investments made by private firms are aimed at securing market advantage" (Scotchmer 2004, p. 1). See also the empirical study of Goolsbee and Syverson (2008).

hump-shaped in the entry threat. When the entry threat is small and increases, the incumbent invests more to deter entry or to make entry unlikely. Then the entry deterrence effect dominates the Schumpeterian effect. However, when the threat becomes huge, entry can no longer profitably be deterred or made unlikely and the investment becomes small. Then the Schumpeterian effect dominates.

We show that the hump-shaped relationship between incumbent's R&D investment and the entry threat is robust to different timings. In one time structure, the incumbent does not know the rivals' production costs when deciding about investment. In the alternative timing the costs are known.

Aghion, Blundell, Griffith, Howitt, and Prantl (2009) also explore how the R&D investment of an incumbent depends on the strength of the entry threat.[2] A difference is that they measure the entry threat by entry costs, whereas we measure it by the number and quality of potential entrants. Additionally, in their model, the leading incumbent is not only threatened by a potential entrant, but also by another incumbent. They show that a higher entry threat increases the leading incumbent's investment when the firm is initially close to the technological frontier; this is due to the escape-entry effect. It is the other way round if the leading incumbent is further behind the frontier; this is due to the discouragement effect. So in contrast to our model, for a certain type of firm, only one effect is present and the influence of a higher threat on the investment is monotone. The model has some problems. First, in a dynamic model it is not appealing that firms only consider the profit of the next period, but not at all profits of later periods. Second, why should there be a technological frontier which moves exogenously? Is it not more plausible that firms themselves determine how the technological frontier moves? Indeed, regarding the former two points, Aghion, Bloom, Blundell, Griffith, and Howitt (2005) take the completely opposite route. Third, it is assumed that innovations occur step-by-step and that entry can only take place at the new technological frontier. This has the un-

[2] For a similar model, see Aghion, Burgess, Redding, and Zilibotti (2005).

plausible consequences that an incumbent which is close to the frontier and innovates must not fear entry at all. In contrast, an incumbent which is further below the frontier cannot prevent entry, no matter how much it invests.

In Aghion, Bloom, Blundell, Griffith, and Howitt (2005) the interplay between an escape-competition effect and the Schumpeterian effect generates an inverted-U relationship between R&D investment and product market competition. In their model, the leader (of a duopoly) does not want to innovate because this does not change its profit. This is in contrast to our model where we only consider the incentives of the leader (which is in our case a monopolist). Moreover, they do not consider entry.

There is a discussion in competition policy about the optimal patent breadth, how costly imitation should be, and when competition law should require a firm with market power to share its property.[3] In our model, stronger patent protection, higher costs of imitation, or stricter property rights can be interpreted as a weakening of the entry threat. Empirically, there is no clear evidence that patents provide strong positive incentives to invest in innovation. The picture is rather mixed.[4] This is a puzzling result (Lerner 2009, p. 347). Our model delivers a simple and intuitive solution. We predict that the incumbent's investment is hump-shaped in the entry threat.[5]

Our model can also be interpreted as one in which the incumbent is a home firm that is threated by foreign firms. The empirical results are mixed but point to a positive relationship between foreign competition

[3]See, for example, Gallini (1992), Scotchmer (2004), Segal and Whinston (2007), and Vickers (2009).

[4]See the survey of Bessen and Meurer (2008), the study on the role of the patent system in the British Industrial Revolution of Mokyr (2009), or Lerner's (2009) study on the impacts of shifts in patent policy across 60 countries.

[5]Also Segal and Whinston (2007) show that in some cases "policies that protect entrants necessarily raise the rate of innovation" (p. 1703); they concentrate on innovations made by potential entrants. In Boldrin and Levine's (2009) model investments in R&D are higher in a competitive equilibrium than in a monopolistic equilibrium.

and innovation in the home market.[6]

The relationship between the intensity of competition and R&D investment is generally regarded as ambiguous in theoretical models. This is due to the large variety of relevant effects and of the definition of competitiveness.[7]

The paper proceeds as follows. In the next section, we present the model. In Section III.3, we analyze it. The alternative timing is considered in Section III.4. Section III.5 explores the question how important the incumbent's initial production costs are for the relationship between incumbent's R&D investments and the entry threat. Section III.6 concludes.

III.2. MODEL

There is an incumbent, firm 0, and N rivals, firms $1, ..., N$. Rivals can enter at cost $t > 0$. They threat the monopoly position of the incumbent. By investing in R&D the incumbent can lower its production costs which makes entry less likely. We will explore how the incumbent's optimal investment varies with the quality and the number of rivals.

At Stage 1, the incumbent chooses its R&D investment $k \geq 0$. The incumbent's per-unit production costs are

$$c_0(k) = C - h(k),$$

where $C > 0$. We assume that the function h is twice differentiable and satisfies the following mild assumptions.

[6]See Gilbert and Sunshine (1995), Lelarge and Nefussi (2008), MacDonald (1994), Pavcnik (2002), Javorcik (2004), and Aitken and Harrison (1999). Aghion, Blundell, Griffith, Howitt, and Prantl (2009) find mixed results, in accordance with their theoretical predictions: whether there is a positive or negative effect depends on the distance of the incumbent to the technological frontier.

[7]See, for example, Lee and Wild (1980) vs. Delbono and Denicolo (1991), Gilbert and Sunshine (1995), Belleflamme and Vergari (2006), Sacco and Schmutzler (2007), Schmutzler (2007), Denicolo and Zanchettin (2008), and Vives (2008). For a survey, see Aghion and Griffith (2005) or Gilbert (2006).

A S S U M P T I O N A 2: (i) $h'(k) > 0$, (ii) $h''(k) < 0$, (iii) $h(0) = 0$, (iv) $\lim_{k\to\infty} C - h(k) > t$, (v) $h'(0) > 1$.

In words, (i) says that a higher investment lowers production costs; (ii) that there are decreasing returns to scale; (iii) that no investment yields no cost reduction; (iv) that it is not possible to yield production costs that make entry impossible; (v) that when there is no entry threat, investing at least a tiny amount is optimal, see below. Parts (i) and (iv) imply $C > t$.

At Stage 2 the rivals' per-unit production costs $(c_1, ..., c_N)$ are drawn.[8] We will later make concrete assumptions on the distributions.

At Stage 3, the rivals decide upon entry in an arbitrary order, potentially simultaneous. We assume that in case of indifference a rival does not enter.

At Stage 4 all rivals that entered and the incumbent compete à la Bertrand. All firms produce a homogenous good. Figure III.1 summarizes the timing.

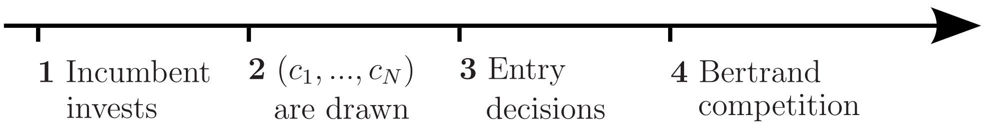

Figure III.1: Timing

Consumers have unit demand and a willingness to pay of one.[9] If there are two or more cheapest firms, they buy from the firm with the lower production costs. This assumption is solely made to avoid open set problems. So that the market is always served, we assume that $C < 1$.

There is perfect information, and each firm maximizes its expected profit. Our solution concept is subgame perfect Nash equilibrium. We consider only pure strategy Nash equilibria.

[8]Also Hopenhayn (1992) considers productivity shocks. But he does not consider innovations.

[9]That is, the incumbent's revenue is at most 1. So any investment $k > 1$ is dominated by $k = 0$. So one could relax Assumption A2(iv) to $C - h(1) > t$.

III.3. ANALYSIS

III.3.1. BERTRAND COMPETITION AND ENTRY

Standard analysis of the Bertrand game yields that rival j's total profit when it entered the market is

$$\pi_j^{\text{entry}}\big|_{\mathcal{M}} = \max\{0, \min\{c_i\}_{i\in\mathcal{M}\setminus j} - c_j\} - t,$$

where $j \in \{1, ..., N\}$ and $\mathcal{M} \subseteq \{0, ..., N\}$ is the set of all firms which are in the market, that is, the incumbent plus the rivals that entered. The first term of the formula is the Bertrand profit. If firm j does not have the lowest production costs among all firms in a market, it makes a Bertrand profit of zero. Otherwise its Bertrand profit is the minimum per-unit cost of all other firms in the market minus its own production costs. The second term is cost of entry.

When a rival does not enter it makes zero profits. Hence, no rival wants to enter in equilibrium, if and only if

$$\pi_j^{\text{entry}}\big|_{\mathcal{M}=\{0\}} \leq 0 \quad \forall j \in \{1, ..., N\}. \tag{III.1}$$

This can be rewritten as

$$c_0(k) \leq c_j + t \quad \forall j \in \{1, ..., N\}.$$

When this condition does not hold, the equilibrium is such that some rival j enters. Then $c_j < c_0(k)$ and the incumbent will make a Bertrand profit of zero. Hence,

$$\pi_0(k) = \begin{cases} \pi_0^{\text{no entry}}(k) = 1 - c_0(k) - k & \text{for } c_0(k) \leq \underline{c} + t, \\ \pi_0^{\text{entry}}(k) = -k & \text{otherwise,} \end{cases} \tag{III.2}$$

where we defined $\underline{c} := \min_{j\in\{1,...,N\}}\{c_j\}$ as the minimum production costs of rivals.

III.3.2. INVESTMENT

The incumbent's expected profit when it invests k is

$$\mathbb{E}[\pi_0(k)] = \pi_0^{\text{no entry}}(k)\text{Prob}^{\text{no entry}}(k) + \pi_0^{\text{entry}}(k)\text{Prob}^{\text{entry}}(k).$$

Using (III.2) we can rewrite this as

$$\mathbb{E}[\pi_0(k,F)] = (1 - C + h(k))(1 - F(C - h(k) - t)) - k,$$

where F is the distribution function from which $\underline{c}$ is drawn. The incumbent's incentives to invest are determined by the marginal effect of investment on its expected profit.

Consider first the benchmark case where entry never occurs:

$$\frac{d\mathbb{E}[\pi_0(k,F)]}{dk} = h'(k) - 1. \tag{III.3}$$

The first term on the right hand sight captures the marginal effect of investment on production costs, the second describes the marginal cost of investment.

Next, consider the more interesting case where entry may occur:

$$\begin{aligned} \frac{d\mathbb{E}[\pi_0(k,F)]}{dk} &= h'(k)(1 - F(C - h(k) - t)) \\ &+ (1 - C + h(k))f(C - h(k) - t))h'(k) - 1, \end{aligned} \tag{III.4}$$

where f is the density function which belongs to F. What has changed through the entry threat? On the one hand, the return of investment is lower: it becomes less likely that the investment is actually used in production; see the first term. Put differently, it is less likely that the investment "pays off". This is called the *Schumpeterian effect*.[10] On the other hand, the return of investment is higher: investing more makes entry less likely; see the second term. We call this the *entry deterrence effect* of investment.[11] Since both effects run in different directions it may well be the case that the incumbent's incentive to invest is higher with an entry threat than without one.

[10] See also Aghion, Harris, Howitt, and Vickers (2001) and Aghion, Bloom, Blundell, Griffith, and Howitt (2005). The Schumpeterian effect is closely related to the discouragement effect in Aghion, Blundell, Griffith, Howitt, and Prantl (2009).

[11] Aghion and Griffith (2005) call this the Darwinian effect of competition. Our entry deterrence effect is similar to the escape-entry effect considered by Aghion, Blundell, Griffith, Howitt, and Prantl (2009) and the escape-competition effect developed in Aghion, Harris, Howitt, and Vickers (2001) and Aghion, Bloom, Blundell, Griffith, and Howitt (2005).

For concreteness, we assume that the production costs $(c_1, ..., c_N)$ of the rivals are independently drawn from exponential density functions. Rival j's costs are drawn from density

$$f_j(c_j) = \lambda_j e^{-\lambda_j c_j} \tag{III.5}$$

with $\lambda_j > 0$ and corresponding distribution function

$$F_j(c_j) = 1 - e^{-\lambda_j c_j}. \tag{III.6}$$

The nice feature when all c_j's are independently and exponentially distributed is that $\underline{c}$ is exponentially distributed, too:[12]

$$F(\underline{c}) = 1 - e^{-\lambda \underline{c}}, \text{ with } \lambda := \sum_{j=1}^{N} \lambda_j.$$

Hence we can allow for heterogeneity of the rivals through different λ_js without complicating the analysis. The parameter λ captures the strength of the entry threat. It increases with the number N and quality λ_j of rivals. When $\lambda = 0$ there is no entry threat.

Under the exponential distribution of the rivals' per-unit costs, we get

$$\mathbb{E}\left[\pi_0(k, \lambda)\right] = (1 - c_0(k))\, e^{-\lambda(c_0(k)-t)} - k. \tag{III.7}$$

Let the optimal investment be given by the function $k^*(\lambda)$.

P R O P O S I T I O N 6: When there is no entry threat the incumbent invests a positive amount: $k^*(0) = h'^{-1}(1) > 0$. When the entry threat is huge ($\lambda \to \infty$) the incumbent does not invest. An investment of at least $\hat{k}$, where $\hat{k}$ is an arbitrary positive investment level, cannot be optimal when λ is sufficiently high.

PROOF: See Appendix.

[12]Technically we need the distribution of the first-order statistics. This distribution can also be derived when each rival's costs are not drawn from an exponential distribution. Then the distribution of the first-order statistics is in general more complicated.

That is, the incumbent invests some positive amount when there is no entry threat. In contrast, when the threat is overwhelming, entry occurs for sure. Then the incumbent does not invest at all. The Schumpeterian effect dominates the entry deterrence effect. The intuition for the last point is as follows. When the entry threat is large, entry is very likely, even when the incumbent invests $\hat{k}$ or more. So the incumbent invests an amount less than $\hat{k}$ to save investment costs.

Next we explore whether it is possible that a higher entry threat increases the optimal investment. The next proposition offers a sufficient condition such that this is true.

PROPOSITION 7: Suppose that $C < \frac{1+t}{2}$. The optimal investment $k^*(\lambda)$ is increasing in λ for $\lambda \to 0$.

PROOF: See Appendix.

Since we assumed that $C \in (0, 1)$ and $t > 0$, this sufficient condition can easily be met. When the initial production costs C are low, the incumbent's monopoly profit is high. Then the incumbent invests more when there is a small entry threat than when there is no threat, because it wants to defend its monopoly. That is, for low entry threats the entry deterrence effect dominates the Schumpeterian effect. In contrast, when C is high, the monopoly profit is low and so the incumbent may have few incentives to defend its monopoly position. Then the Schumpeterian effect dominates the entry deterrence effect even for small entry threats.

The next Proposition follows directly from Propositions 6 and 7.

PROPOSITION 8: Suppose that $C < \frac{1+t}{2}$. The optimal investment $k^*(\lambda)$ is hump-shaped in λ.

To sum up, when the entry threat is small and increases, the incumbent invests more to make entry unlikely. This is due to the entry deterrence effect. However, when the threat becomes huge, entry can no longer profitably be made unlikely and the investment becomes small. Then the Schumpeterian effect dominates.

In the Appendix, we consider an alternative specification where $\underline{c}$ is drawn from the uniform distribution and show that our results are robust. The robustness of our results is also shown in the next section.

III.4. ALTERNATIVE TIMING

We now consider an alternative timing where Stage 1 and 2 are interchanged. That is, the incumbent already knows $(c_1, ..., c_N)$ when deciding about investment. The optimal investment is denoted by $k^{**}(\underline{c})$.

When the rivals' production costs are infinite we know from the previous analysis that entry never occurs. From (III.2) we get that the incumbent's investment is then

$$k^{**}(\infty) = h'^{-1}(1) > 0. \tag{III.8}$$

Note, also for $\underline{c} \geq 1 - t$ there is also no entry threat, because even when the incumbent does not invest, no rival would enter. So without an entry threat the incumbent's investment is the same in both timings.

The profit function, given that entry is deterred, is concave in k:

$$d^2\pi_0^{\text{no entry}}(k)/dk^2 = h''(k) < 0.$$

So when the incumbent deters entry, it either invests $k^{**}(\infty)$ or, if that is not enough, just enough to deter entry:

$$k^{\text{deter entry}}(\underline{c}) = \begin{cases} k^{**}(\infty) = h'^{-1}(1) & \text{for } \underline{c} \geq c_0(k^{**}(\infty)) - t, \\ h^{-1}(C - t - \underline{c}) & \text{otherwise.} \end{cases} \tag{III.9}$$

When the incumbent does not want to deter entry, $\pi_0(k) = -k$, see (III.2). So the optimal investment is

$$k^{\text{do not deter entry}}(\underline{c}) = 0. \tag{III.10}$$

This yields zero profits.[13]

[13]Note, even a zero investment may deter entry. So the previous equation is only sensible when entry occurs, given a zero investment.

Does the incumbent want to deter entry or not? Denote the investment, above which entry deterrence yields a loss, by $\bar{k}$. It is implicitly given by

$$1 - C + h(\bar{k}) - \bar{k} = 0. \qquad \text{(III.11)}$$

Through the assumptions made before, existence and uniqueness are guaranteed.[14]

So when investing according to (III.9) yields an investment which is at most $\bar{k}$, it is optimal to deter entry and to follow this investment rule. Otherwise, not deterring entry and zero investments are optimal, see (III.10). The following lemma summarizes our findings. They are illustrated in Figure III.2.[15]

L E M M A 3: When $\underline{c}$ is below $c_0(\bar{k}) - t$ the incumbent does not invest and entry occurs. Otherwise the incumbent invests according to (III.9) and entry is deterred.

Figure III.2 shows a hump-shaped relationship between the incumbent's investment and $\underline{c}$. But to make the results comparable to the one yielded under the original timing we seek to answer the following question: How large is the average investment of the incumbent, given λ? Again, we assume that rivals cost are drawn from an exponential density function.

[14] $\bar{k}$ exists because h is continuous in k, and $\pi_0^{\text{no entry}}(k = 1) < 0$, $\pi_0^{\text{no entry}}(k = 0) > 0$. This value is unique because $\pi_0^{\text{no entry}}(k = 0) > 0$ and $\pi_0^{\text{no entry}}(k)$ is a concave function of k through A1(ii).

[15] The following properties of $k^{\text{deter entry}}(\underline{c})$ are useful to construct the Figure.

(i) $k^{\text{deter entry}}(\underline{c})$ is continuous at $\underline{c} = c_0(k^{**}(\infty)) - t$,

(ii) it has a kink at $\underline{c} = c_0(k^{**}(\infty)) - t$:
$\lim_{\underline{c} \searrow c_0(k^{**}(\infty))-t} dk^{\text{deter entry}}(\underline{c})/d\underline{c} = 0$ and
$\lim_{\underline{c} \nearrow c_0(k^{**}(\infty))-t} dk^{\text{deter entry}}(\underline{c})/d\underline{c} = -1$,

(iii) $k^{\text{deter entry}}(\underline{c})$ is constant in $\underline{c}$ for $\underline{c} > c_0(k^{**}(\infty)) - t$, and

(iv) $k^{\text{deter entry}}(\underline{c})$ is decreasing and convex in $\underline{c}$ for $\underline{c} < c_0(k^{**}(\infty)) - t$.

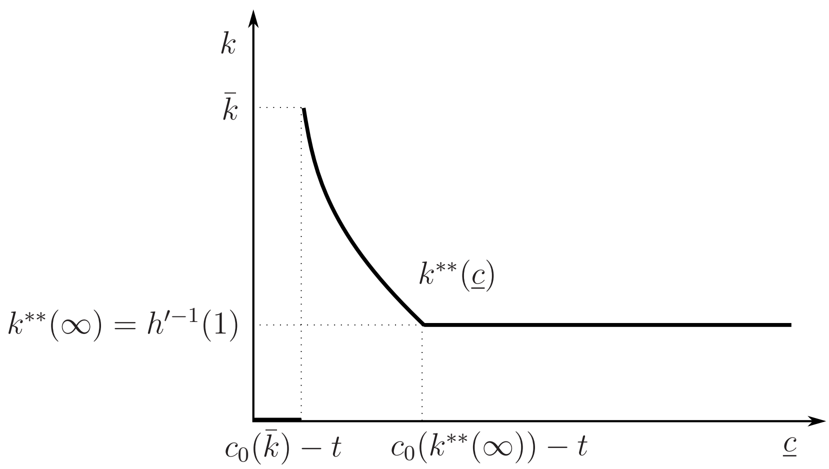

Figure III.2: Incumbent's investment decision

P R O P O S I T I O N 9: $\lim_{\lambda \to 0} \mathbb{E}[k|\lambda] = k^{**}(\infty)$, $\lim_{\lambda \to \infty} \mathbb{E}[k|\lambda] = 0$. If

$$\frac{\big(c_0(k^{**}(\infty)) - c_0(\bar{k})\big)^2}{2} - (c_0(\bar{k}) - t)k^{**}(\infty) > 0,$$

then the maximum of $\mathbb{E}[k|\lambda]$ is greater than $k^{**}(\infty)$.

PROOF: See Appendix.

Intuitively, when $\lambda \to 0$ the probability that $\underline{c}$ will lie in the right region of Figure III.2 approaches one. Therefore, the incumbent's expected investment is $k^{**}(\infty)$. When $\lambda \to \infty$ the probability that $\underline{c}$ will lie in the left region of the Figure approaches one. Therefore, the incumbent's expected investment is zero.

The intuition for the sufficient condition is as follows: If λ is low, it is very likely that $\underline{c}$ will lie in the right region of Figure III.2. Then $\mathbb{E}[k|\lambda] \approx k^{**}(\infty)$. When λ increases, it gets more likely that $\underline{c}$ is in the left or in the middle region of Figure III.2. When $k^{**}(\infty)$ is small enough, $\mathbb{E}[k|\lambda]$ increases with λ for small λs. The reason is that the reduction of the probability to get a medium investment $k^{**}(\infty)$ is overcompensated through an increased probability to get a high investment.[16]

[16]Similarly, given some $k^{**}(\infty)$, the middle region must be large enough, so that $\big(c_0(k^{**}(\infty)) - c_0(\bar{k})\big)$ must be high enough. The size of the left region is $c_0(\bar{k}) - t$.

The results with this alternative timing are qualitatively the same as with the original timing: For a low entry threat the incumbent's average investment is moderate. For a medium threat the incumbent's average investment is, under some conditions, relatively high. For a huge entry threat the incumbent's expected investment approaches zero. Therefore, the relationship between the incumbent's investment and the entry threat is again hump-shaped.

Additionally note, that the LHS of the sufficient condition in Proposition 9 is decreasing in C:

$$\frac{d\text{LHS}}{dC} = \big(-c_0(k^{**}(\infty)) + c_0(\bar{k}) - k^{**}(\infty)\big)\,\frac{dc_0(\bar{k})}{d\bar{k}}\,\frac{d\bar{k}}{dC} - k^{**}(\infty) < 0.$$

Hence, with lower initial production costs C it is more likely that we get a hump-shaped relationship between the incumbent's R&D investment and the entry threat. This finding concerning the alternative timing is parallel to the one with the original timing; see Proposition 8.

III.5. HOW IMPORTANT ARE THE INCUMBENT'S INITIAL COSTS?

Aghion, Blundell, Griffith, Howitt, and Prantl (2009) show that a higher entry threat increases incumbent's investment when the firm is initially close to the technological frontier, and it is the other way round when the incumbent is further behind the frontier. Does this result also hold in our model?

We have no frontier in our model. But closeness to the frontier corresponds, in our model, to low initial production costs C. As can be seen from (III.4),

$$\frac{d^2\mathbb{E}[\pi_0(k,F)]}{dkdC} = \\ -2h'(k)f(C-h(k)-t)) + (1-C+h(k))h'(k)f'(C-h(k)-t)).$$

The first term is negative, whereas the sign of the second term is ambiguous. Under the exponential distribution the second term is negative,

When this region is small enough, only little of the probability mass goes in the left region when λ increases, and hence $\mathbb{E}[k|\lambda]$ increases with λ for small λs.

too. Therefore we have $d^2\mathbb{E}[\pi_0(k, F)]/dkdC < 0$.[17]

Proposition 10: When there is an entry threat, the incumbent's optimal investment is decreasing in the initial costs C.

Proof: See Appendix.

Observe that without an entry threat the incumbent's optimal investment is independent of C. Therefore, Proposition 10 implies that with high initial costs it is more likely that an entry threat decreases the incumbent's optimal investment. This is similar to the finding of Aghion, Blundell, Griffith, Howitt, and Prantl (2009). The intuition for our result is as follows. The Schumpeterian effect is more likely to dominate the entry deterrence effect when the incumbent has initially high costs, because high costs make entry deterrence (a) less profitable, since the production costs are relatively high and (b) more difficult, since for a given investment entry becomes more likely.

With the alternative timing similar arguments hold. Without an entry threat the incumbent's optimal investment is independent of the initial costs. When there is an entry threat, higher initial costs lead to lower expected investments. This is true because of two effects: (i) with a higher C, the range where the incumbent does not invest at all increases, because entry can very often not be profitably deterred; (ii) the maximal investment $\bar{k}$ decreases.

III.6. CONCLUSIONS

The model formalizes the idea that an incumbent rests on its laurels when there is no threat, fights when there is some moderate threat, and gives up when the threat is huge. We measure the threat by the number and quality of rival firms which may enter the market. A higher threat may motivate an incumbent to invest more in R&D to deter entry or to make it unlikely; then the entry deterrence effect dominates. However, when the

[17]This is also true with other distributions for which $f'(\cdot)$ is negative, zero, or "not too" positive.

threat is overwhelming, the incumbent has little chances to deter entry and invests little or not at all; then the Schumpeterian effect dominates. Therefore, the relationship between the incumbent's investment and the entry threat is hump-shaped.

IV. AMBIGUITY IN A PRINCIPAL-AGENT MODEL

We consider a principal-agent model with moral hazard where the agent's knowledge about the applied performance measure is ambiguous. We show that agent's ambiguity aversion leads to weaker incentives and a lower payoff for the principal, compared to the standard model without ambiguity and without ambiguity aversion. Moreover, when there is enough ambiguity the principal sets no incentives at all. Additionally, the Informativeness Principle does not hold.

IV.1. INTRODUCTION

In principal-agent models with moral hazard, the principal motivates the agent to spend effort via a performance-dependent wage scheme. Standard theory predicts that the wage scheme highly depends on performance. However, in reality, wage schemes sometimes do not depend on performance or the dependence is rather weak.[1] Moreover, the Informativeness Principle,[2] which is a key finding of the literature, "seems to be violated in many occupations" (Prendergast 1999, p. 21). We show that considering ambiguity and ambiguity aversion in an otherwise standard model can bring theory in line with empirics.

In standard models of principal-agent relationships with moral hazard it is implicitly assumed that the agent knows precisely the statistical

[1]See Holmström and Milgrom (1991) or Prendergast (1999) and the references therein.

[2]Roughly speaking, the Informativeness Principle says that the principal wants to use all information in a compensation contract which is correlated with performance.

properties of the performance measure. However, we think that this is a rather strong—and in real life often unrealistic—assumption.[3] We relax it by assuming that the agent faces ambiguity with respect to the performance measure.

The famous Ellsberg (1961) paradox shows that individuals are averse to ambiguity. Ellsberg suggested the following experiment: there are two urns, each containing 100 balls, each of which is either red or black. Urn A contains 50 black balls and 50 red ones. There is no information on urn B. One ball will be drawn from each urn. A subject has to choose a bet; when she wins the bet she earns 100$. Empirically, subjects are indifferent between the bets "the ball drawn from urn A is black" and "... red". This also holds for urn B. However, subjects prefer bets in which urn A is involved over bets in which B is involved. This cannot hold under the rational expectations hypothesis. Gilboa and Schmeidler (1989) propose the following solution: "In case of urn B, the subject has too little information to form a prior. Hence, (s)he considers a *set* of priors as possible. Being uncertainty averse, (s)he takes into account the *minimal* expected utility (over all priors in the set) while evaluating a bet" (p. 142; italics provided). That is, subjects dislike bets with ambiguity (unknown probabilities). They are ambiguity-averse.[4]

Maybe the simplest principal-agent model with moral hazard is the LEN model: the wage scheme is *l*inear, the agent's utility is *e*xponential, and the shock on the performance measure is *n*ormally distributed. We use the standard LEN model except that we assume that the agent's

[3] As Gollier (2008) notes, "[i]n many circumstances, it is difficult to assess the precise probability distribution to describe the uncertainty faced by a decision maker." This view is also supported by Ghirardato (1994, see p. 3). Post, van den Assem, Baltussen, and Thaler (2008, p. 39) emphasize that "real-life choices rarely come with precise probabilities."

[4] In contrast, according to the rational expectations hypothesis, in such circumstances, a subject nonetheless has a single probability measure in conformity with the Bayesian model; see Savage's (1954) axiomatization and the nontechnical discussion of Gilboa, Postlewaite, and Schmeidler (2008). For a neural empirical study on ambiguity, see Hsu, Bhatt, Adolphs, Tranel, and Camerer (2005). Note, ambiguity is sometimes also called Knightian uncertainty following Knight (1921).

knowledge about the performance measure is ambiguous and that he is ambiguity-averse. Due to ambiguity the agent's beliefs about the distribution of the shock on the performance measure are not represented by a single probability function, but instead by a set of probability functions. We use Gilboa and Schmeidler's (1989) ambiguity aversion concept in which an act is evaluated by the probability distribution that yields the lowest expected utility. In our moral hazard framework this means that the agent is pessimistic about the distribution of the shock on the performance measure whenever rewards are subject to stochastics, which is necessary in order to create incentives. As a consequence, the compensation demanded by the agent is relatively high, compared to the standard LEN model (which neither considers ambiguity nor ambiguity aversion). Since the principal has to ensure participation of the agent, this implies that the principal's cost of providing incentives is relatively high. As a result, the principal sets relatively weak incentives which yield a relatively low expected payoff. It may even be the case that the optimal incentive scheme is a fixed wage. In the standard LEN model this can never happen.[5]

When there are two performance measures it can be optimal for the principal to ignore one of them (and potentially both), even though both measures are informative. The reason is that with ambiguity, the inclusion of a measure into a wage scheme causes costs which are not negligible even when the wage depends on the measure only to a small extent. Hence, the Informativeness Principle does not hold in our model.

In contrast to the finance literature,[6] the ambiguity concept is rarely used in principal-agent theory. There are a few exceptions. Mukerji (2003) inquires into the impact of ambiguity in procurement contracts under cost uncertainty. He shows that the optimal linear contract sets no financial incentives at all to induce exertion of cost-reducing effort

[5] Also, in Holmström and Milgrom's (1991) multiple-tasks model, it can be optimal to set weak or no monetary incentives. We show that this can also arise in a one-task model when there is ambiguity.

[6] See, for example, Dow and Werlang (1992) or Epstein and Wang (1994).

when ambiguity is sufficiently high. In contrast to our model, only a binary shock realization is considered and the agent is assumed to be risk-neutral. In Lang (2007) the agent is—as in Mukerji (2003)—risk-neutral. He considers a framework with two tasks where only the performance of one task can be rewarded via a linear contract. He shows that ambiguity may lead to the provision of weak incentives. Ghirardato (1994) and Karni (2006) also belong to this literature, but have different foci then we have.

In the next section we first present and analyze the standard LEN model. Then we extend it to incorporate ambiguity and ambiguity aversion. In Section IV.3 we show that the Informativeness Principle does not hold in our model. In Section IV.4 we discuss our model and its results. Section IV.5 concludes.

IV.2. MODELS

IV.2.1. THE STANDARD LEN MODEL

Consider the LEN model specified by Bolton and Dewatripont (2005, Ch. 4.2). There is a risk-neutral principal and a risk-averse agent with an exponential (i.e., CARA) utility function. The former makes a take-it-or-leave-it offer to the latter. It is assumed that the wage contract can only be linear in the realization of the performance measure q: $w = t+sq$, where t is the fix component and s is the variable component.[7] When the agent rejects the contract he gets a monetary payoff of $\bar{w}$ and the principal of $\bar{\pi}$, with $\bar{w}+\bar{\pi} < 0$.[8] When the agent accepts, he has to choose an action (also called effort) $a \in \mathbb{R}_{\geq 0}$. The effort costs are $\psi(a) = \frac{1}{2}ca^2$, where $c > 0$ is a cost parameter. Performance depends on the agent's effort as well as on a shock: $q = a+\epsilon$, where $\epsilon \sim N(0, \sigma^2)$. After the shock is realized, the

[7]Linear contracts are not optimal in this setting. They are, however, optimal in the dynamic setting of Holmström and Milgrom (1987). We plan to check whether or not this remains true when there is ambiguity.

[8]The inequality guarantees that the principal prefers hiring an agent with the fixed wage contract $w = \bar{w}$ over not hiring an agent. We add this assumption to make sure that the principal always hires an agent.

wage payment is made in accordance with the contract. The principal's payoff is $q - w$. The agent's payoff is $u(\cdot) = -e^{-\eta[w-\psi(a)]}$, where $\eta > 0$ is the agent's coefficient of absolute risk aversion.

Straightforward calculations yield that the agent chooses the effort

$$a = \begin{cases} s/c & \text{for } s \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

The principal sets

$$s^*_{\text{standard}} = 1/(1 + \eta c \sigma^2) \tag{IV.1}$$

which generates her an expected payoff of

$$\mathbb{E}[q - w]^*_{\text{standard}} = \frac{1}{2c(1 + \eta c \sigma^2)} - \bar{w}. \tag{IV.2}$$

IV.2.2. THE MODEL WITH AMBIGUITY AND AMBIGUITY AVERSION

How the Agent Evaluates a Wage Scheme.— Let there be a finite number of probability distributions of the shock on the performance measure that are plausible, given the agent's knowledge. We assume that the plausible probability distributions are stochastically independent. For the sake of consistency we assume that aggregating all these plausible distributions yields the objective distribution $\epsilon \sim N(0, \sigma^2)$. Given stochastic independence, and since the objective distribution is normal, Cramér's (1936) Theorem implies that the plausible probability distributions are normal, too.

Each plausible distribution $[i, j]$ is characterized by its mean μ_i and variance σ_j^2. Suppose that there are $N \times n$ plausible distributions and that the parameters describing the agent's plausible distributions can be ordered in a grid. See Figure IV.1 for a 4×3 example. Formally, the set of parameters characterizing the plausible distributions is

$$\mathcal{S} := \{(\mu_i, \sigma_j^2) | \mu_i \in \boldsymbol{\mu}, \sigma_j^2 \in \boldsymbol{\sigma^2}\},$$

with $\boldsymbol{\mu} := (\mu_1, ..., \mu_i, ..., \mu_N)$ and $\boldsymbol{\sigma^2} := (\sigma_1^2, ..., \sigma_j^2, ..., \sigma_n^2)$, where $\mu_i < \mu_{i+1}$, $\sigma_j^2 < \sigma_{j+1}^2$.[9] According to the agent's knowledge, the probability

[9]We specified the set of the plausible distributions such that the analysis is kept

that the shock is drawn from the plausible distribution $[i, j]$ is $p_{i,j}$, where $\sum_{\mathcal{S}} p_{i,j} = 1$. We assume that $p_{i,j} > 0$ for all i, j.

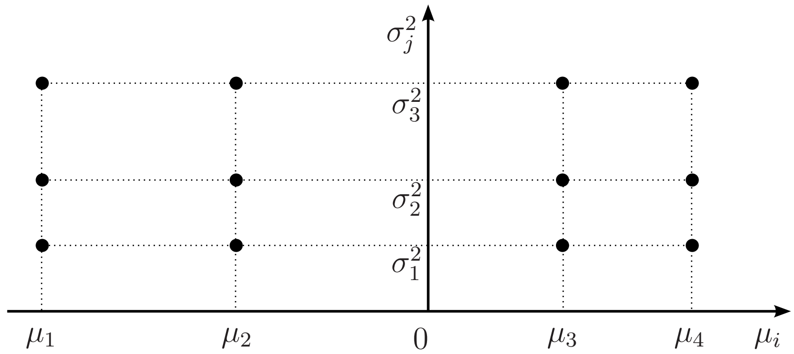

Figure IV.1: A 4×3 example of plausible distributions.

When there is only one plausible distribution we say that there is no ambiguity.

DEFINITION 1: When $N \times n = 1$ there is *no ambiguity* and when $N \times n > 1$ there is *ambiguity*.

DEFINITION 2: There is *more ambiguity* when μ_1 decreases and σ_n^2 does not decrease, or μ_1 does not increase and σ_n^2 increases.

The motivation for Definition 2 is as follows: in our model, less precise knowledge (i.e., more ambiguity) translates into a wider dispersion of $\boldsymbol{\mu}$ and $\boldsymbol{\sigma^2}$, i.e., into a decrease of μ_1 and an increase of σ_n^2.

The following lemma will be useful.

LEMMA 4: Suppose $N \times n > 1$. Then $\mu_1 \leq 0$ and $\sigma_n^2 > \sigma^2$. The first inequality is strict for $N \geq 2$.

PROOF: See Appendix.

simple. Our model, however, can easily be generalized to other specifications. Whenever there exists a plausible distribution which has weakly the lowest mean and weakly the highest variance, nothing changes at all. When there is no such plausible distribution, the analysis gets more cumbersome: the distribution according to which the agent evaluates the wage scheme depends on the incentives set by the principal.

In words, the plausible distribution with the lowest mean has a mean which is lower or equal to the mean of the objective distribution and the plausible distribution with the highest variance has a variance which is greater than the variance of the objective distribution.[10]

Being ambiguity-averse in the sense of Gilboa and Schmeidler (1989), the agent assigns to each possible effort and contract the plausible distribution which leads to the lowest expected utility.[11] That is, the agent evaluates a wage scheme according to the most pessimistic plausible distribution.[12]

Suppose that the agent evaluates a contract according to the plausible distribution $[i, j]$ and takes effort a. Calculations similar to the ones in Bolton and Dewatripont (2005) yield that the certainty equivalent of the agent is then

$$\hat{w}(\cdot) = t + sa + s\mu_i - \frac{1}{2}\eta s^2\sigma_j^2 - \frac{1}{2}ca^2. \qquad \text{(IV.3)}$$

But according to which plausible distribution does the agent evaluate the contract? Consider the case where $s > 0$. One directly sees that $d\hat{w}(\cdot)/d\mu_i > 0$ and $d\hat{w}(\cdot)/d\sigma_j^2 < 0$. Hence, the lowest expected utility is yielded when the variance is maximal and the mean minimal. So the agent evaluates the wage scheme according to the plausible distribution $[1, n]$. For $s = 0$, all plausible distributions yield the same expected payoff. Hence, the plausible distribution used to evaluate the wage scheme is arbitrary. Without loss of generality we can assume that also in this case the agent uses the plausible distribution $[1, n]$. Moreover, it is straight-

[10]It is easy to show that Lemma 4 is also true for the case in which the realizations of the plausible distributions are correlated (and the plausible distributions are still normal). There is one exception: when the correlation is 1 and $n = 1$ we have $\sigma^2 = \sigma_n^2$.

[11]There are other, more complicated concepts, for example Ghirardato, Maccheroni, and Marinacci (2004) or Klibanoff, Marinacci, and Mukerji (2005). See Eichberger and Kelsey (2009) for an overview.

[12]The feature of Gilboa and Schmeidler's (1989) model that the agent assigns to each effort the lowest expected value over his set of priors is not important for our results. Suppose instead that the agent takes multiple priors into account and overweights the "bad" priors. Then a condition similar to Lemma 4 can be derived. Therefore, all of our results stay qualitatively valid.

forward that for $s < 0$ the agent uses the plausible distribution $[N, n]$.

LEMMA 5: The agent evaluates a wage scheme with $s \geq 0$ according to the plausible distribution $[1, n]$ and a scheme with $s < 0$ according to the plausible distribution $[N, n]$.

Principal's Problem.— The principal knows that $\epsilon \sim N(0, \sigma^2)$.[13] Moreover, in accordance with the standard LEN model, we assume that the principal knows the agent's preferences and his plausible distributions.

The principal solves the following program:

$$\max_{s,t} \mathbb{E}[q - w] \quad \text{subject to}$$

(i) the incentive constraint that the desired effort level $a^*_{\text{ambiguity}}$ is chosen by the agent: $a^*_{\text{ambiguity}} \in \text{argmax}_a \hat{w}(a)$,

(ii) the participation constraint which guarantees that the agent signs the wage scheme: $\hat{w}(a^*_{\text{ambiguity}}) \geq \bar{w}$, and

(iii) the constraint that the agent evaluates the wage scheme with distribution $[1, n]$ for $s \geq 0$ and with $[N, n]$ for $s < 0$ (see Lemma 5).

We now seek to simplify the maximization problem. First, let us look at the incentive constraint. From (IV.3) we directly get that the agent chooses—as in the standard LEN model—

$$a = \begin{cases} s/c & \text{for } s \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Moreover, in the optimum the participation constraint holds with equality; otherwise the principal can decrease t so that her payoff increases. For both, $s = 0$ and $s < 0$, the agent chooses $a = 0$. It is straightforward to show that setting $s < 0$ is strictly dominated by $s = 0$ for the principal. So we can neglect the case $s < 0$. Using the previous insights, the

[13]Thus, we consider a non-common prior setting (see also Eliaz and Spiegler 2006, 2008). Alternatively, we can also assume that the principal has—as the agent—ambiguous information, but is not ambiguity averse.

principal's maximization problem simplifies to

$$\max_s \frac{s}{c} + s\mu_1 - \eta s^2\sigma_n^2/2 - \frac{s^2}{2c} - \bar{w} \quad \text{s.t. } s \geq 0. \tag{IV.4}$$

The interpretation is simple: the first part is the effort induced by the contract; the terms $-(s\mu_1 - \eta s^2\sigma_n^2/2)$ are the combined risk and ambiguity premium which can be decomposed into a risk premium $\eta s^2\sigma^2/2$ and an ambiguity premium $-s\mu_1 + \eta s^2(\sigma_n^2 - \sigma^2)/2$; the term $\frac{s^2}{2c}$ represents the equilibrium effort costs; and $\bar{w}$ is the monetary equivalent of the agent's outside option.

Simple optimization of (IV.4) yields that the principals sets the following incentives:

$$s^*_{\text{ambiguity}} = \begin{cases} \frac{1+c\mu_1}{1+\eta c\sigma_n^2} & \text{for } 1 + c\mu_1 > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{IV.5}$$

From Lemma 4, we know that $\mu_1 \leq 0$ and $\sigma_n^2 > \sigma^2$. Hence, we have the following results.

P R O P O S I T I O N 11: With ambiguity the incentives set by the principal are weaker than without ambiguity. With ambiguity the principal may set no incentives at all.

To understand both results, we come back to the interpretation of (IV.4). Due to ambiguity there is another premium beside the risk premium which makes providing incentives more costly for the principal. This explains why the principal sets weaker incentives.

Moreover, due to ambiguity the marginal costs of providing incentives are not negligible even for weak incentives whenever $\mu_1 < 0$. Formally, $\lim_{s\to 0} d(\text{ambiguity premium})/ds = -\mu_1$. At the same time the marginal benefit of setting weak incentives (via inducing a higher effort) is $1/c$. Hence, whenever it is sufficiently costly to provide even weak incentives the principal sets no incentives at all.[14]

[14]When no incentives are set, zero effort is implemented. It may be more plausible to assume that the agent wants to spend some effort even when no incentives are set. This is indeed the case if, for example, $\psi'(a) \leq 0$ for $a \in [0, \bar{a}]$, see Holmström and Milgrom (1991, pp. 33-34).

Without ambiguity there is no ambiguity premium. Then the marginal costs of providing weak incentives are negligible.[15] Therefore the principal always sets incentives, see (IV.1).[16]

Plugging (IV.5) into (IV.4) yields the principal's expected payoff in equilibrium:

$$\mathbb{E}[q-w]^*_{\text{ambiguity}} = \begin{cases} \frac{(1+c\mu_1)^2}{2c(1+\eta c\sigma_n^2)} - \bar{w} & \text{for } 1+c\mu_1 > 0, \\ -\bar{w} & \text{otherwise.} \end{cases} \tag{IV.6}$$

Since $\mu_1 \leq 0$ and $\sigma_n^2 > \sigma^2$ (see Lemma 4) we get the following result.

PROPOSITION 12: Ambiguity leads to a lower expected payoff for the principal.

Intuitively, with ambiguity the principal has to compensate the agent not only for his risk premium but also for his ambiguity premium. This increases the principal's cost of implementing any level of effort, and hence also the principal's cost when the optimal effort is implemented.

Comparative Statics.— More ambiguity leads to the provision of weaker incentives: $ds^*_{\text{ambiguity}}/d\mu_1 > 0$ and $ds^*_{\text{ambiguity}}/d\sigma_n^2 < 0$. Intuitively, more ambiguity increases the marginal costs of providing incentives. Hence, the principal sets weaker incentives.

More ambiguity leads to a lower expected payoff of the principal: $d\mathbb{E}[q-w]^*_{\text{ambiguity}}/d\mu_1 \geq 0$ and $d\mathbb{E}[q-w]^*_{\text{ambiguity}}/d\sigma_n^2 \leq 0$, where for $s^*_{\text{ambiguity}} > 0$ the inequalities are strict. The logic is simple: The principal can set the same incentives with little ambiguity than with a lot of ambiguity. Since more ambiguity leads to higher costs for the principal in order to provide a certain level of incentives (see arguments before), the principal must be better off with little ambiguity.

Finally note that in the standard LEN model as well as in the enriched LEN model with ambiguity and ambiguity aversion a higher degree of

[15]For these insights it is not important that the effort costs are quadratic: it is sufficient that the marginal effort costs are negligible for low efforts. In models with limited liability of the agent, setting weak incentives can be costly for the principal. See, for example, Besley and Ghatak (2008, Section 1).

[16]Note, risk aversion plays no role for weak incentives since the agent is first-order risk neutral. Technically, $\lim_{s\to 0} d(\text{risk premium})/ds = 0$.

risk aversion η (at least weakly) decreases the optimal incentives and the principal's payoff.

IV.3. ADDITIONAL INFORMATION: VIOLATION OF THE INFORMATIVENESS PRINCIPLE

Suppose there are two performance measures Y and Z with realizations y and z, respectively. Similar to the case with one performance measure we assume that

$$y = a + \epsilon_Y, \quad \text{where } \epsilon_Y \sim N(0, \sigma_Y^2),$$
$$z = a + \epsilon_Z, \quad \text{where } \epsilon_Z \sim N(0, \sigma_Z^2),$$

and that the principal's expected payoff is $a - \mathbb{E}[w]$. We assume that ϵ_Y and ϵ_Z are uncorrelated. Moreover, we assume that the agent only feels ambiguity with respect to the performance measure Z and he is again ambiguity-averse. The contract takes the form $w = t + s_Y y + s_Z z$. Simple calculations yield that it is not always optimal for the principal to include both performance measures into the contract.[17]

PROPOSITION 13: The performance measure Z is included in the wage scheme if and only if

$$\mu_{Z,1} > -\frac{\eta\sigma_Y^2}{1 + \eta c \sigma_Y^2}.$$

The measure Y is always used.

PROOF: See Appendix.

In the standard model (without ambiguity and without ambiguity aversion) it is always optimal to use both measures. The intuition is simple (see Holmström 1979, p. 87). Suppose the optimal contract depends

[17]Note that this result does qualitatively neither depend on the assumption that the agent only feels ambiguity with respect to the performance measure Z, nor on the assumption that ϵ_Y and ϵ_Z are uncorrelated. When he feels ambiguity with respect to both measures the principal may optimally neglect both measures; that is, set a fixed wage. This finding and the intuition is as in the case with only one performance measure. When ϵ_Y and ϵ_Z are correlated it gets less worthwhile to include both measures into a contract.

only on the measure Y. By making the contract slightly dependent on the measure Z the agent's incentives can be improved. Because the agent is locally risk-neutral, this does marginally not increase the agent's risk premium. Hence, the optimal contract uses Y and Z.

This argument does not hold with ambiguity. The costs of making a contract even slightly dependent on the measure Z are non-negligible whenever we have an ambiguity-averse agent who is ambiguous about Z. Therefore, when there is a lot of ambiguity (i.e., when $\mu_{Z,1}$ is low) the principal does not use the measure Z, even though Z reveals information; see Proposition 13.[18] This finding is valuable since "the Informativeness Principle, i.e., that all factors correlated with performance should be included in a compensation contract, seems to be violated in many occupations" (Prendergast 1999, p. 21).

IV.4. DISCUSSION

IV.4.1. WHEN PARTIES "AGREE TO DISAGREE"

The model stays mathematically equivalent when there is neither ambiguity nor ambiguity aversion, but it is instead the case that the agent and the principal have different opinions about the distribution of the shock on the performance measure. That is, both parties "agree to disagree". Suppose that the principal thinks that the shock is normally distributed with variance σ^2 and (normalized) mean 0, whereas the agent thinks that the shock is normally distributed with variance σ_n^2 and mean μ_1. When the agent is more pessimistic about the distribution than the principal we have $\mu_1 < 0$ and $\sigma_n^2 > \sigma^2$, as in the case with ambiguity and ambiguity

[18]The previous finding that with one performance measure a fixed wage may be optimal (see Proposition 11) can also be seen as a violation of the Informativeness Principle: there is a performance measure which reveals information about the agent's effort, but the measure may not be used. Similarly, without ambiguity, one performance measure, and linear effort costs of $\tilde{c}a$, a fixed wage contract is optimal whenever $\tilde{c} \geq 1$. Therefore we prefer to interpret the Informativeness Principle as follows: *Whenever some incentives are set*, any measure of performance that reveals information on the effort is included in the contract.

aversion.[19]

IV.4.2. MULTIPLE PRINCIPALS AND COMPETITION

Suppose there are multiple principals who compete for the agent. When a principal offers a wage scheme, she essentially offers the agent a certain expected utility level. Hence, we can depict competition simply as a variation of the reservation monetary payoff $\bar{w}$. Since our results do not depend on $\bar{w}$—as long as the principal wants to hire the agent, which is also true for competition—our results are robust to the introduction of competition.

IV.4.3. AGENT'S WELFARE

Given some agent who is either ambiguity-averse (and there is ambiguity) or not. When is he better off? Since the principal makes a take-it-or-leave-it offer, the agent's expected ex-ante utility is in both cases equal to the utility from the reservation payoff $\bar{w}$. However, when the agent is ambiguity-averse, there is a difference between expected ex-ante utility and expected ex-post utility: the agent is in expectation better off ex-post than expected ex-ante. Hence, from an ex-post perspective the agent yields a higher surplus with ambiguity aversion than without it. The agent's higher surplus is at the principal's expense, see Proposition 12.[20]

Do these results also hold with multiple principals? Suppose multiple principals compete perfectly for the agent. Then, in equilibrium all principals must make zero expected profits. This is true with or without ambiguity aversion of the agent. Hence, the principals' surpluses do not change due to agent's ambiguity aversion. Next we show that the agent

[19]The agent may also be more optimistic than the principal so that we have $\mu_1 > 0$ and $\sigma_n^2 < \sigma^2$. Then the principal sets very high-powered incentives and receives a very high expected payoff, compared to the case with a common prior.

[20]That a party profits from its distorted prior is non-standard. For example, in DellaVigna and Malmendier (2004), the party with the undistorted prior exploits the party with the distorted prior.

is ex-ante and ex-post worse off with ambiguity aversion than without. The expected ex-post total surplus is $\mathbb{E}[q-w]+\tilde{w}$, where $\tilde{w}$ is the ex-post certainty equivalent which equals $t+sa-\frac{1}{2}\eta s^2\sigma^2-\frac{1}{2}ca^2$. Maximization of the expected ex-post total surplus yields that $s^*_{\text{total}}=1/(1+\eta c\sigma^2)$. Observe that $s^*_{\text{total}}=s^*_{\text{standard}}$. With ambiguity aversion the principal sets different incentives: $s^*_{\text{total}}\neq s^*_{\text{ambiguity}}$. Hence, with an ambiguity-averse agent, the expected ex-post total surplus is lower than without ambiguity aversion. From before we know that with perfect competition the principal's surplus does not change due to ambiguity aversion of the agent. Hence, from an ex-post perspective the agent is worse off with ambiguity aversion than without it. This result is the reverse yielded with only one principal.

IV.4.4. HISTORY DEPENDENCE

Our insights can be applied further. Suppose a principal has to select one out of two performance measures to design a linear wage contract. The first is well known by the agent (there is no ambiguity), but is very inexact (i.e., it has a high variance); the properties of the second are the other way round. One can use the formulas of Proposition 12 to compare them. It may well turn out that the principal prefers the first measure, although it is less exact.

When substituting an old with a new measure in a wage scheme, the agent's knowledge about the new is scarcer than about the old.[21] Hence, old measures have a comparative advantage over new measures.

IV.5. CONCLUSIONS

We considered a principal-agent relationship with moral hazard where the agent's knowledge of the performance measure is ambiguous and he is averse to ambiguity. This has the effect that the agent is pessimistic

[21]This argument is supported by the findings described by Rustichini (2005, p. 1625) "that the ambiguity premium declines as subjects repeat their choices: People slowly adjust to ambiguity".

about the distribution of the shock on the performance measure. Hence, the compensation demanded by the agent is relatively high, compared to the standard model. This implies that the principal's cost of providing incentives are relatively high. Therefore, the principal sets relatively weak incentives which yield a relatively low expected payoff. The principal may even set a fixed wage. This can never be optimal in the standard model. Moreover, in the enriched model with two performance measures the famous Informativeness Principle does not hold: It can be optimal for the principal to ignore one measure (or potentially both), even though both measures are informative.

V. THE OPTIMALITY OF SIMPLE CONTRACTS: MORAL HAZARD AND LOSS AVERSION

Joint work with
Fabian Herweg and Daniel Müller

This paper extends the standard principal-agent model with moral hazard to allow for agents having reference-dependent preferences according to Kőszegi and Rabin (2006, 2007). When loss aversion is the predominant determinant of the agent's risk preferences, the principal optimally offers a simple bonus contract, i.e., when the agent's performance exceeds a certain threshold, he receives a fixed bonus payment. Also when risk aversion becomes more important, the optimal contract displays less complexity than predicted by orthodox theory. Thus, loss aversion introduces an endogenous complexity cost into contracting.

V.1. INTRODUCTION

The recent literature provides very strong evidence that contractual forms have large effects on behavior. As the notion that "incentive matters" is one of the central tenets of economists of every persuasion, this should be comforting to the community. On the other hand, it raises an old puzzle: if contractual form matters so much, why do we observe such a prevalence of fairly simple contracts?

— *Bernard Salanié (2003)*

A lump-sum bonus contract, with the bonus being a payment for achieving a certain level of performance, is probably the most simple incentive scheme for employees one can think of. According to Steenburgh (2008), salesforce compensation plans provide incentives mainly via a lump-sum bonus for meeting or exceeding the annual sales quota.[1] Simple contracts are commonly found not only in labor contexts but also in insurance markets. A prevalent form of insurance contracts is a straight-deductible contract, widely used, for example, in automobile insurance. The observed plainness of contractual arrangements, however, is at odds with predictions made by economic theory, as nicely stated in the above quote by Salanié. While Prendergast (1999) already referred to the discrepancy between theoretically predicted and actually observed contractual form, over time this question was raised again and again, recently by Lazear and Oyer (2007), and the answer still is not fully understood.[2]

Beside this gap between theoretical prediction and observed practice, both theoretical as well as empirical studies demonstrate that these simple contractual arrangements create incentives for misbehavior of the agent that is outside the scope of most standard models. As Oyer (1998) points out, facing an annual sales quota provides incentives for sales-

[1]Incentives for salespeople in the food manufacturing industry often are solely created by a lump-sum bonus, see Oyer (2000). Moreover, in his book about designing effective sales compensation plans, Moynahan (1980) argues that for a wide range of industries lump-sum bonus contracts are optimal. For a survey on salesforce compensation plans see Joseph and Kalwani (1998).

[2]For evidence on deductibles in the automobile insurance, see Puelz and Snow (1994) or Chiappori et al. (2006). As was shown by Rothschild and Stiglitz (1976), the use of deductibles can theoretically be explained if the insurance market is subject to adverse selection. Besides adverse selection, however, moral hazard plays an important role in automobile insurance. Deductibles were found to be optimal under moral hazard by Holmström (1979) if the insured person's action influences only the probability of an accident but not its severity. As pointed out by Winter (2000), however, "[d]riving a car more slowly and carefully reduces both the probability of an accident and the likely costs of an accident should it occur." Thus, existing theories cannot explain the prevalence of deductibles in these markets.

people to manipulate prices and timing of business to maximize their own income rather than firms' profits. For insurance markets, Dionne and Gagné (2001) show that "deductible contracts can introduce perverse effects when falsification behavior is potentially present".[3] These observations raise "the interesting question of why these [...] contracts are so prevalent. [...] It appears that there must be some benefit of these contracts that outweighs these apparent costs" (Lazear and Oyer, 2007).

To give one possible explanation for the widespread use of these contractual arrangements, we consider a principal-agent model with moral hazard, framed as an employer-employee relationship, which is completely standard but for one twist: the agent is assumed to be loss-averse in the sense of Kőszegi and Rabin (2006, 2007). With the tradeoff between incentive provision and risk sharing being at the heart of moral hazard, allowing for a richer description of the agent's risk preferences that goes beyond standard risk aversion seems a natural starting point to gain deeper insights into contract design. Our main finding is that a simple (lump-sum) bonus scheme is optimal when loss aversion is the driving force of the agent's risk preferences.[4] This is in stark contrast to the findings for a standard risk-averse agent. An agent who is risk but not loss-averse exhibits local risk neutrality, which implies that paying slightly different wages for different signals improves incentives at negligible cost. A loss-averse agent, on the other hand, is first-order risk-averse. Since losses loom larger than equally-sized gains, in expectations the agent suffers from deviations from his reference point. With the reference point being multidimensional under the concept of Kőszegi and Rabin, the agent is first-order risk-averse at all possible wage levels. In

[3]For evidence on fraudulent claims being a major problem in the car insurance market see Caron and Dionne (1997), who estimated the cost of fraud in the Québec automobile insurance market in 1994 at $100 million, just under 10% of total claims. For an estimation of the costs of fraudulent claims in the United States, see Foppert (1994).

[4]In the following, we will use the terms bonus contract and bonus scheme interchangeably to refer to a contract that specifies exactly two distinct wage payments, a base wage and a lump-sum bonus.

consequence, paying even slightly different wages reduces the agent's expected utility, for which in turn he demands to be compensated. Thus, by offering a simple contract, that specifies only few different wage levels, the principal can lower the expected payment necessary to compensate the agent for the induced losses.

We present our model of a principal-agency that is subject to moral hazard in Section V.2. The principal, who is both risk and loss neutral, does not observe the agent's effort directly. Instead, she observes a measure of performance that is correlated with the agent's effort decision. Following Kőszegi and Rabin, we posit that a decision maker – next to intrinsic consumption utility from an outcome – also derives gain-loss utility from comparing the actual outcome with his rational expectations about outcomes. More precisely, the sensation of gains and losses is derived by comparing a given outcome to all possible outcomes. To illustrate this point, consider an employee who receives a wage of \$5000 for good performance, a wage of \$4400 for mediocre performance, and a wage of \$4000 for bad performance. If the employee's performance is mediocre, this generates mixed feelings, a loss of \$600 and a gain of \$400.[5] The key feature of the Kőszegi-Rabin model is that expectations matter in determining the reference point.[6] While mainly based on findings in the psychological literature,[7] evidence for this assumption is provided also by two recent contributions to the economic literature. In a real-effort experiment, Abeler et al. (2009) find strong evidence for individuals taking their expectations as a reference point, rather than the status quo.[8] Similarly, analyzing decision making in a large-stake game show, Post

[5]For at least suggestive evidence on mixed feelings, see Larsen et al. (2004).

[6]The feature that the reference point is determined by the decision maker's forward-looking expectations is shared with the disappointment aversion models of Bell (1985), Loomes and Sugden (1986), and Gul (1991).

[7]For instance, Mellers et al. (1999) and Breiter et al. (2001) document that both the actual outcome and unattained possible outcomes affect subjects' satisfaction with their payoff.

[8]The status quo was most often assumed as reference point in the wake of Kahneman and Tversky's (1979) original formulation of prospect theory.

et al. (2008) come to the conclusion that observed behavior "is consistent with the idea that the reference point is based on expectations." Regarding applications, the Kőszegi-Rabin concept is used by Heidhues and Kőszegi (2005, 2008) to introduce consumer loss aversion into otherwise standard models of industrial organization. While the former paper explains why monopoly prices react less sensitive to cost shocks than predicted by orthodox theory, the latter provides an answer to the question why non-identical competitors charge identical prices for differentiated products.

As a benchmark, in Section V.3 we reconsider the case of a purely risk-averse agent: Under the optimal contract, signals that are more indicative of higher effort are rewarded strictly higher, thereby giving rise to a strictly increasing wage profile. We then turn to the analysis of a purely loss-averse agent who does not exhibit risk aversion in the usual sense. After providing sufficient conditions for the first-order approach to be valid, we establish our main result: when the agent is loss-averse, it is optimal to offer a bonus contract. No matter how rich the set of possible realizations of the performance measure, the optimal contract comprises of only two different wage payments. We already touched on the intuition underlying this finding: With the agent's action being unobservable, the necessity to create incentives makes it impossible for the principal to bear the complete risk. With losses looming larger than equally sized gains, this ex ante imposes an expected net loss on the agent, which equals the sum over the ex ante expected wage differences weighted by the product of the corresponding probabilities. To illustrate, let us return to the example introduced above. Suppose the agent expects to perform well, moderately, or poorly with probability p_G, p_M and p_B, respectively. Then, ex ante, the agent expects a wage difference – or net loss – of \$600 with probability $p_M p_G$, a net loss of \$400 with probability $p_B p_M$, and a net loss of \$1000 with probability $p_B p_G$. The agent demands to be compensated for his overall expected net loss, which the principal therefore seeks to minimize. Consider, for the sake of argument, a principal who wants to strengthen incentives to provide effort, starting out from a not fully differentiated wage scheme. There are two ways to do so.

First, the principal can introduce a new wage spread, i.e., pay slightly different wages for two signals that were rewarded equally in the original wage scheme, while keeping the differences between all other neighboring wages constant. Secondly, the principal can increase an existing wage spread, holding constant all other spreads between neighboring wages. Both procedures increase the overall expected net loss by increasing the size of some of the expected losses without reducing others. Introducing a new wage spread, however, additionally increases the overall expected net loss by increasing the ex ante expected probability of experiencing a loss. Therefore, in order to improve incentives, it is advantageous to increase an existing wage spread without adding to the contractual complexity in the sense of increasing the number of different wages. In this sense, reference-dependent preferences according to Kőszegi and Rabin introduce an endogenous complexity cost into contracting based on psychological foundations.

Thereafter, we establish several properties displayed by the optimal contract. Let a signal that is the more likely to be observed the higher the agent's effort be referred to as a *good* signal. We find that the subset of signals that are rewarded with the high wage contains either only good signals, or all good signals and possibly a few bad signals as well.[9] When abstracting from integer-programming problems, it is optimal for the principal to order the signals according to their relative informativeness (likelihood ratio), i.e., the agent receives the high wage for all signals that are more indicative of high effort than a cutoff signal. Last, we show that an increase in the agent's degree of loss aversion may allow the principal to use a lower-powered incentive scheme in order to implement a desired level of effort. The reason is that a higher degree of loss aversion may be associated with a stronger incentive for the agent to choose a high effort

[9]The theoretical prediction that inferior performance may also well be rewarded with a bonus is in line with both Joseph and Kalwani (1998)'s suggestion that organizations tend to view the payment of a bonus as a reward for good or even acceptable performance rather than an award for exceptional performance, and Churchill et al. (1993)'s prescription that bonuses should be based on objectives that can be achieved with reasonable rather than Herculean efforts.

in order to reduce the probability of incuring a loss. The overall cost of implementation, however, increases in the agent's degree of loss aversion.

In the last part of Section V.3, we analyze the general case in which the agent is both risk and loss-averse. It is shown that our results are robust towards a small degree of risk aversion. Moreover, we give a heuristic reasoning why a reduction in the complexity of the contract is also to be expected to be optimal for a non-negligible degree of risk aversion, and confirm our conjecture by means of a numerical example.[10]

Returning to the case of a purely loss-averse agent, in Section V.4 we relax the assumptions that guaranteed validity of the first-order approach. Here, to keep the analysis tractable, we focus on binary measures of performances. If the agent's degree of loss aversion is sufficiently high and if the performance measure is sufficiently informative, then only extreme actions – work as hard as possible or do not work at all – are incentive compatible. Put differently, the principal may face severe problems in fine-tuning the agent's incentives. These implementation problems, however, can be remedied if the principal can commit herself to stochastically ignoring the low realization of the performance measure, i.e., by turning a blind eye from time to time. Besides alleviating implementation problems, turning a blind eye may also lower the cost of implementing a certain action. Thus, the sufficiency part of Blackwell's theorem does not hold when the agent has reference-dependent preferences.

After briefly summarizing our main findings, Section V.5 concludes by discussing robustness of our results with respect to imposed assumptions. All proofs are given in the appendix.

RELATED LITERATURE

Before presenting our model, we relate our paper to the small but steadily growing literature that analyzes the implications of loss aversion on in-

[10]This finding also relates to the observation that, within a firm, pay for individuals often seems to be less variable than productivity, as recently surveyed by Lazear and Shaw (2007). Our model suggests an alternative explanation for this pay compression outside the realms of inequity aversion, tournament theory, and influence activities.

centive design.[11] With reference-dependent preferences being at the heart of loss aversion on the one hand, but with no unifying approach provided how to determine a decision maker's reference point on the other hand, it is little surprising that all contributions differ in this particular aspect. While Dittmann et al. (2007) posit that the reference income is exogenously given by the previous year's fixed wage, Iantchev (2005), who considers a market environment with multiple principals competing for the services of multiple agents, applies the concept of Rayo and Becker (2007). Here, an agent's reference point is endogenously determined by the equilibrium conditions in the market. When focusing on a particular principal-agent pair, however, both the principal and the agent take the reference point as exogenously given. An exogenous reference point does not always seem plausible. Starting out from the premise that the reference point is forward looking and depends on the distributions of outcomes, as suggested by ample evidence, De Meza and Webb (2007) consider both exogenous as well as endogenous formulations of the reference point. Concluding that the disappointment concept of Gul (1991), which equates the reference point with the certainty equivalent of the income distribution, does yield some questionable implications,[12] De Meza and Webb propose that the reference income is the median income, which

[11]Beside loss aversion there are other behavioral biases that are incorporated into contracting problems with moral hazard. Non-standard risk preferences in a moral hazard framework are analyzed by Schmidt (1999), who applies Yaari's (1987) concept of dual expected utility theory. Englmaier and Wambach (2006) characterize the optimal contract for the case of an inequity-averse agent in the sense of Fehr and Schmidt (1999). A multi-agent contracting problem in which agents care about their own status is investigated by Besley and Ghatak (2008) in a static context, and by Auriol and Renault (2008) in a dynamic setting. By introducing worker overconfidence into a multi-agent moral-hazard problem, Fang and Moscarini (2005) show that it can be optimal not to screen workers according to their skills. For a review of behavioral economics of organizations see Camerer and Malmendier (2007).

[12]De Meza and Webb consider two otherwise identical agents who differ only in their degree of loss aversion. They point out that with the certainty equivalent as reference point, there are situations where the less loss-averse agent experiences a loss, but the more loss-averse agent does not.

captures the idea that the agent incurs a loss at all incomes for which it is odds-on that a higher income would be drawn. Taking median income as reference income, however, suffers from the drawback that it is discontinuous in the underlying probability distribution.[13]

All of the aforementioned contributions explore questions of both empirical importance as well as theoretical interest: Dittmann et al. (2007) find that a loss aversion model dominates an equivalent risk aversion model in explaining observed CEO compensation contracts if the reference point is equal to the previous year's fixed wage. Iantchev (2005) finds evidence for his theoretically predicted results in panel data from Safelite Glass Corporation. Last, by explaining why bonuses are paid for good performance rather than penalties for poor performance, De Meza and Webb (2007) provide a theoretical underpinning for the frequent usage of option-like incentive schemes in CEO compensation. The contractual form predicted by these papers, however, is rather complex: while the optimal contract typically displays a range where pay is independent of performance, for performance above this range payment varies with performance in a fairly complex way, depending crucially on the underlying distribution of signals. Theoretical predictions differ in whether or not the optimal contract includes punishment for very poor performance or where in the wage schedule the optimal contract features discontinuities. Thus, none of these papers provides a rationale for the prevalence of fairly simple contracts, bonus contracts in particular.[14]

To the best of our knowledge, Daido and Itoh (2007) is the only paper that also applies the concept of reference dependence à la Kőszegi and

[13]For example, suppose that with a probability of .51 a manager earns \$1m and with a probability of .49 he earns \$2m. With median income as reference point the manager will never suffer a loss because his reference income is \$1m. A small shift in probabilities, however, makes the median income equal to \$2m. Now, the agent suffers a loss in almost 50% of all cases.

[14]De Meza and Webb (2007) find conditions under which a simple bonus contract is optimal. For this to be the case, however, they assume that the reference point is exogenously given and that all wage payments are in the loss region, where the agent is assumed to be risk loving.

Rabin to a principal-agent setting. The focus of Daido and Itoh, however, greatly differs from ours. Assuming that the performance measure comprises of only two signals, two types of self-fulfilling prophecy are explained, the Galatea and the Pygmalion effects.[15] While sufficient to capture these two effects, the assumption of a binary measure of performance does not allow one to inquire into the form that contracts take under moral hazard.

V.2. MODEL

There are two parties, a principal and an agent.[16] The principal offers a one-period employment contract to the agent, who has an outside employment opportunity (or reservation utility) yielding expected utility $\bar{u}$. If the agent accepts the contract, then he chooses an effort level $a \in \mathcal{A} \equiv [0, 1]$. The agent's action a equals the probability that the principal receives a benefit $B > 0$. The principal's expected net benefit is

$$\pi = aB - E[W],$$

where W is the compensation payment the principal pays to the agent.[17] The principal is assumed to be risk and loss neutral, thus she maximizes π. We wish to inquire into the form that contracts take under moral hazard and loss aversion. Therefore, we focus on the cost minimization problem to implement a certain action $\hat{a} \in (0, 1)$.[18]

[15]Roughly speaking, the former effect refers to empirical findings that an agent's self-expectation about his performance is an important determinant of his actual performance, whereas the latter effect refers to the phenomenon that a principal's expectation about the agent's performance has an impact on the agent's actual performance.

[16]The framework is based on MacLeod (2003), who analyzes subjective performance measures without considering loss-averse agents.

[17]The particular functional form of the principal's profit function is not crucial for our analysis. We assume this specific structure since it allows for a straight-forward interpretation of the performance measure.

[18]The second-best action maximizes the principal's expected benefit, aB, minus the minimum cost of implementing action a. The overall optimal contract exhibits

The action choice $a \in \mathcal{A}$ is private information of the agent and unobservable for the principal. Furthermore, the realization of B is not directly observable. A possible interpretation is that B corresponds to a complex good whose quality cannot be determined by a court, thus a contract cannot depend on the realization of B. Instead the principal observes a contractible measure of performance, $\hat{\boldsymbol{\gamma}}$, with $s \in \mathcal{S} \equiv \{1, \ldots, S\}$ being the realization of the performance measure, also referred to as signal. Let $S \geq 2$. The probability of observing signal s conditional on B being realized is denoted by γ_s^H. Accordingly, γ_s^L is the probability of observing signal s conditional on B not being realized. Hence, the unconditional probability of observing signal s for a given action a is $\gamma_s(a) \equiv a\gamma_s^H + (1-a)\gamma_s^L$. For technical convenience, we make the following assumption.

ASSUMPTION A3: For all $s, \tau \in \mathcal{S}$ with $s \neq \tau$,

(i) $\gamma_s^H / \gamma_s^L \neq 1$ (informative signals),

(ii) $\gamma_s^H, \gamma_s^L \in (0,1)$ (full support),

(iii) $\gamma_s^H / \gamma_s^L \neq \gamma_\tau^H / \gamma_\tau^L$ (different signals).

Assumption (i) guarantees that any signal s is either a good or a bad signal, in the sense that the overall probability of observing that signal unambiguously increases or decreases in a. Part (ii) ensures that for all $a \in \mathcal{A}$, all signals occur with positive probability. Last, with assumption (iii) signals can unambiguously be ranked according to the relative impact of an increase in effort on the probability of observing a particular signal.[19]

The contract which the principal offers to the agent consists of a payment for each realization of the performance measure, $\{w_s\}_{s=1}^S \in$

the same characteristics as the contract that minimizes the cost of implementing an arbitrary action $\hat{a}$.

[19]Formally, for all $a \in [0,1]$, $(\gamma_s^H - \gamma_s^L)/\gamma_s(a) > (\gamma_\tau^H - \gamma_\tau^L)/\gamma_\tau(a) \iff \gamma_s^H/\gamma_s^L > \gamma_\tau^H/\gamma_\tau^L$.

$\mathbb{R}^S$.[20]

The agent is assumed to have reference-dependent preferences in the sense of Kőszegi and Rabin (2006): Overall utility from consuming $\boldsymbol{x} = (x_1, \ldots, x_K) \in \mathbb{R}^K$ – when having reference level $\boldsymbol{r} = (r_1, \ldots, r_K) \in \mathbb{R}^K$ for each dimension of consumption – is given by

$$v(\boldsymbol{x}|\boldsymbol{r}) \equiv \sum_{k=1}^{K} m_k(x_k) + \sum_{k=1}^{K} \mu(m_k(x_k) - m_k(r_k)).$$

Put verbally, overall utility is assumed to have two components: consumption utility and gain-loss utility. Consumption utility, also called intrinsic utility, from consuming in dimension k is denoted by $m_k(x_k)$. How a person feels about gaining or losing in a dimension is assumed to depend in a universal way on the changes in consumption utility associated with such gains and losses. The universal gain-loss function $\mu(\cdot)$ satisfies the assumptions imposed by Tversky and Kahneman (1991) on their "value function".[21] In our model, the agent's consumption space comprises of two dimensions, money income ($x_1 = W$) and effort ($x_2 = a$).[22] The agent's intrinsic utility for money is assumed to be a strictly increasing, (weakly) concave, and unbounded function. Formally, $m_1(W) = u(W)$ with $u'(\cdot) > \varepsilon > 0$, $u''(\cdot) \leq 0$. The intrinsic disutility from exerting effort $a \in [0, 1]$ is a strictly increasing, strictly convex function of effort, $m_2(a) = -c(a)$ with $c'(0) = 0$, $c'(a) > 0$ for $a > 0$, $c''(\cdot) > 0$, and $\lim_{a \to 1} c(a) = \infty$. We assume that the gain-loss function is piece-wise

[20]Restricting the principal to offer non-stochastic wage payments is standard in the principal-agent literature and also in accordance with observed practice. In a later section we comment on this assumption.

[21]Roughly speaking, $\mu(z)$ is strictly increasing, continuous for all z, twice differentiable for all $z \neq 0$ with $\mu(0) = 0$, convex over the range of losses, and concave over the range of gains. For a more formal statement of these properties, see Bowman et al. (1999).

[22]We implicitly assume that the agent is a "narrow bracketer" in the sense that he ignores that the risk from the current employment relationship is incorporated with substantial other risk.

linear,

$$\mu(m) = \begin{cases} m \ , & \text{for } m \geq 0 \\ \lambda m, & \text{for } m < 0 \end{cases} .$$

The parameter λ characterizes the weight put on losses relative to gains.[23] The weight on gains is normalized to one. When $\lambda > 1$, the agent is loss-averse in the sense that losses loom larger than equally-sized gains.[24]

Following Kőszegi and Rabin (2006, 2007), the agent's reference point is determined by his rational expectations about outcomes. A given outcome is then evaluated by comparing it to all possible outcomes, where each comparison is weighted with the ex-ante probability with which the alternative outcome occurs. With the actual outcome being itself uncertain, the agent's ex ante expected utility is obtained by averaging over all these comparisons.[25] We apply the concept of choice-acclimating personal equilibrium (CPE) as defined in Kőszegi and Rabin (2007), which assumes that a person correctly predicts his choice set, the environment

[23] Alternatively, one could assume that $\mu(m) = \eta m$ for gains and $\mu(m) = \eta\lambda m$ for losses, where $\eta \geq 0$ can be interpreted as the weight attached to gain-loss utility relative to intrinsic utility. Our implicit normalization $\eta = 1$ is without loss of generality due to the applied concept of choice-acclimating personal equilibrium (CPE). Carrying η through the whole analysis would only replace $(\lambda - 1)$ by $\eta(\lambda - 1)$ in all formulas.

[24] The assumption of a piece-wise linear gain-loss function is not uncommon in the literature on incentive design with loss-averse agents, see De Meza and Webb (2007), Daido and Itoh (2007). In their work on asset pricing, Barberis et al. (2001) also apply this particular functional form, reasoning that "curvature is most relevant when choosing between prospects that involve only gains or between prospects that involve only losses. For gambles that can lead to both gains and losses, [...] loss aversion at the kink is far more important than the degree of curvature away from the kink."

[25] Suppose the actual outcome $\boldsymbol{x}$ and the vector of reference levels $\boldsymbol{r}$ are distributed according to distribution functions F and G, respectively. As introduced above, overall utility from two arbitrary vectors $\boldsymbol{x}$ and $\boldsymbol{r}$ is given by $v(\boldsymbol{x}|\boldsymbol{r})$. With the reference point being distributed according to probability measure G, the utility from a certain outcome is the average of how this outcome feels compared to all other possible outcomes, $U(\boldsymbol{x}|G) = \int v(\boldsymbol{x}|\boldsymbol{r}) \ dG(\boldsymbol{r})$. Last, with $\boldsymbol{x}$ being drawn according to probability measure F, utility is given by $E[U(F|G)] = \iint v(\boldsymbol{x}|\boldsymbol{r}) \ dG(\boldsymbol{r})dF(\boldsymbol{x})$. Since we use choice acclimating personal equilibrium, $F = G$.

he faces, in particular the set of possible outcomes and how the distribution of these outcomes depends on his decisions, and his own reaction to this environment. The eponymous feature of CPE is that the agent's reference point is affected by his choice of action. As pointed out by Kőszegi and Rabin, CPE refers to the analysis of risk preferences regarding outcomes that are resolved long after all decisions are made. This environment seems well-suited for many principal-agent relationships: Often the outcome of a project becomes observable, and thus performance-based wage compensation feasible, long after the agent finished working on that project. Under CPE, the expectations relative to which a decision's outcome is evaluated are formed at the moment the decision is made and, therefore, incorporate the implications of the decision. More precisely, suppose the agent chooses action a and that signal s is observed. The agent receives wage w_s and incurs effort cost $c(a)$. While the agent expected signal s to come up with probability $\gamma_s(a)$, with probability $\gamma_\tau(a)$ he expected signal $\tau \neq s$ to be observed. If $w_\tau > w_s$, the agent experiences a loss of $\lambda(u(w_s) - u(w_\tau))$, whereas if $w_\tau < w_s$, the agent experiences a gain of $u(w_s) - u(w_\tau)$. If $w_s = w_\tau$, there is no sensation of gaining or losing involved. The agent's utility from this particular outcome is given by

$$u(w_s)+\sum_{\{\tau|w_\tau<w_s\}}\gamma_\tau(a)(u(w_s)-u(w_\tau))+\sum_{\{\tau|w_\tau\geq w_s\}}\gamma_\tau(a)\lambda(u(w_s)-u(w_\tau))-c(a).$$

Averaging over all possible outcomes yields the agent's expected utility from choosing action a:

$$E[U(a)] = \sum_{s=1}^{S}\gamma_s(a)\Bigg\{u(w_s) + \sum_{\{\tau|w_\tau<w_s\}}\gamma_\tau(a)(u(w_s) - u(w_\tau)) + \sum_{\{\tau|w_\tau\geq w_s\}}\gamma_\tau(a)\lambda(u(w_s) - u(w_\tau))\Bigg\} - c(a).$$

Note that since the agent's expected and actual effort choice coincide, there is neither a gain nor a loss in the effort dimension.

We conclude this section by briefly summarizing the underlying timing.

1) The principal makes a take-it-or-leave-it offer to the agent.

2) The agent either accepts or rejects the contract. If the agent rejects the game ends and each party receives her/his reservation payoff. If the agent accepts the game moves to the next stage.

3) The agent chooses his action and forms rational expectations about the monetary outcomes. The agent's rational expectations about the realization of the performance measure determine his reference point.

4) Both parties observe the realization of the performance measure and payments are made according to the contract.

V.3. ANALYSIS

Let the inverse function of the agent's intrinsic utility of money be $h(\cdot)$, i.e., $h(\cdot) := u^{-1}(\cdot)$. Thus, the monetary cost for the principal to offer the agent utility u_s is $h(u_s) = w_s$. Due to the assumptions imposed on $u(\cdot)$, $h(\cdot)$ is a strictly increasing and weakly convex function. Following Grossman and Hart (1983), we regard $\boldsymbol{u} = \{u_1, \ldots, u_S\}$ as the principal's control variables in her cost minimization problem to implement action $\hat{a} \in (0,1)$. The principal offers the agent a contract that specifies for each signal a monetary payment or, equivalently, an intrinsic utility level. With this notation, the agent's expected utility from exerting effort a is given by

$$E[U(a)] = \sum_{s \in \mathcal{S}} \gamma_s(a) u_s - (\lambda - 1) \sum_{s \in \mathcal{S}} \sum_{\{\tau | u_\tau > u_s\}} \gamma_\tau(a) \gamma_s(a)(u_\tau - u_s) - c(a). \quad \text{(V.1)}$$

For $\lambda = 1$ the agent's expected utility equals expected net intrinsic utility. Thus, for $\lambda = 1$ we are in the standard case without loss aversion. Moreover, from the above formulation of the agent's utility it becomes clear that λ captures not only the weight put on losses relative to gains, but $(\lambda - 1)$ also characterizes the weight put on gain-loss utility relative to intrinsic utility. Thus, for $\lambda \leq 2$, the weight attached to gain-loss utility is below the weight attached to intrinsic utility. For a given contract $\boldsymbol{u}$, the agent's marginal utility of effort

$$E[U'(a)] = \sum_{s\in\mathcal{S}}(\gamma_s^H - \gamma_s^L)u_s$$
$$-(\lambda-1)\sum_{s\in\mathcal{S}}\sum_{\{\tau|u_\tau>u_s\}}[\gamma_\tau(a)(\gamma_s^H-\gamma_s^L)+\gamma_s(a)(\gamma_\tau^H-\gamma_\tau^L)](u_\tau-u_s)-c'(a). \tag{V.2}$$

Suppose the principal wants to implement action $\hat{a} \in (0, 1)$. The optimal contract minimizes the expected wage payment to the agent subject to the usual incentive compatibility and individual rationality constraints:

$$\min_{u_1,\dots,u_S} \sum_{s\in\mathcal{S}} \gamma_s(\hat{a})h(u_s)$$
$$\text{subject to} \quad E[U(\hat{a})] \geq \bar{u}\ , \tag{IR}$$
$$\hat{a} \in \arg\max_{a\in\mathcal{A}} E[U(a)]. \tag{IC}$$

As a first benchmark, consider the case where the agent's action choice is observable and contractible, i.e., the incentive constraint (IC) is absent. In order to implement action $\hat{a}$ in this first-best situation, the principal pays the agent $u^{FB} = \bar{u} + c(\hat{a})$ irrespective of the realization of the performance measure if the agent chooses the desired action, thereby compensating him for his outside option and his effort cost. In the presence of moral hazard, on the other hand, the principal faces the classic tradeoff between risk sharing and providing incentives: When the agent is anything but risk and loss neutral, it is neither optimal to have the agent bear the complete risk, nor to fully insure the agent.

At this point we simplify the analysis by imposing two assumptions. These assumptions are sufficient to guarantee that the principal's cost minimization problem exhibits the following two properties: First, there are incentive-compatible wage contracts, i.e., contracts under which it is optimal for the agent to choose the desired action $\hat{a}$. Existence of such contracts is not generally satisfied with the agent being loss-averse. Second, the first-order approach is valid, i.e., the incentive constraint to implement action $\hat{a}$ can equivalently be represented as $E[U'(\hat{a})] = 0$. The first assumption that we introduce requires that the weight attached to gain-loss utility does not exceed the weight put on intrinsic utility.

Assumption A4: No dominance of gain-loss utility, $\lambda \leq 2$.

As carefully laid out in Kőszegi and Rabin (2007), CPE implies a strong notion of risk aversion, in the sense that a decision maker may choose stochastically dominated options when $\lambda > 2$.[26] The reason is that, with losses looming larger than gains of equal size, the person ex ante expects to experience a net loss. In consequence, if reducing the scope of possibly incuring a loss is the decision maker's primary concern, the person would rather give up the slim hope of experiencing a gain at all in order to avoid the disappointment in case of not experiencing this gain. In our model, if the agent is sufficiently loss-averse, the principal may be unable to implement any action $\hat{a} \in (0,1)$. The reason is that the agent minimizes the ex ante expected net loss by choosing one of the two extreme actions. The values of λ for which this behavior is optimal for the agent depend on the precise structure of the performance measure. Assumption A4 is sufficient, but not necessary, to ensure that there is a contract such that $\hat{a} \in (0,1)$ satisfies the necessary condition for incentive compatibility. Moreover, the tendency to choose stochastically dominated options seems counterintuitive.[27] Next to ensuring existence of an incentive compatible contract, A4 rules out that our findings are driven by such counterintuitive behavior of the agent. It is worthwhile to emphasize, that our main findings (Propositions 15 and 19) still hold

[26]Suppose a loss-averse person has to choose between two lotteries: lottery 1 pays x for sure; lottery 2 pays $x+y$ with probability p, where $y > 0$, and x otherwise. Then, for each $\lambda > 2$, the decision maker prefers the dominated lottery 1 if $p < (\lambda-2)/(\lambda-1)$. For further details on this point, see Kőszegi and Rabin (2007).

[27]The "uncertainty effect" identified by Gneezy et al. (2006) refers to people valuing a risky prospect less than its worst possible outcome. While this may be interpreted as experimental evidence for people having preferences for stochastically dominated options, this finding crucially relies on the lottery currency not being stated in purely monetary terms. Therefore, we believe that in the context of wage contracts most people do not choose dominated options.

for $\lambda > 2$ as long as existence and validity of the first-order approach are guaranteed. In Section V.4 we relax Assumption A4 and discuss in detail the implications of higher degrees of loss aversion.

To keep the analysis tractable we impose the following assumption.

ASSUMPTION A5: Convex marginal cost function, $\forall\ a \in [0,1]$: $c'''(a) \geq 0$.

Given A4, Assumption A5 is a sufficient but not a necessary condition for the first-order approach to be applicable.[28] Alternatively, it would also suffice to have λ sufficiently small, or the slope of the marginal cost function sufficiently steep. In fact, our results only require the validity of the first-order approach, not that Assumption A5 holds. In Section V.4 we consider the case in which the first-order approach is invalid.

LEMMA 6: Suppose A3-A5 hold, then the constraint set of the principal's cost minimization problem is nonempty for all $\hat{a} \in (0,1)$.

The above lemma states that there are wage contracts such that the agent is willing to accept the contract and then chooses the desired action. Moreover, we will show that a second-best optimal contract exists. This, however, is shown separately for the three cases analyzed in this section: pure risk aversion, pure loss aversion, and the intermediate case.

Sometimes it will be convenient to state the constraints in terms of increases in intrinsic utilities instead of absolute utilities. Note that whatever contract $\{\hat{u}_s\}_{s\in\mathcal{S}}$ the principal offers, we can relabel the signals such that this contract is equivalent to a contract $\{u_s\}_{s=1}^S$ with $u_{s-1} \leq u_s$ for all $s \in \{2,\ldots,S\}$. This, in turn, allows us to write the contract as $u_s = u_1 + \sum_{\tau=2}^s b_\tau$, where $b_\tau = u_\tau - u_{\tau-1} \geq 0$ is the increase in intrinsic utility for money when signal τ instead of signal $\tau - 1$ is observed. Let $\boldsymbol{b} = (b_2,\ldots,b_S)$. Using this notation allows us to rewrite the individual

[28]The validity of the first-order approach under assumptions A3-A5 is rigorously proven in the appendix. The reader should be aware, however, that the proof requires some notation introduced later on. We therefore recommend to defer reading the proof until having read the preliminary considerations up to Section V.3.1.

rationality constraint as follows:

$$u_1 + \sum_{s=2}^{S} b_s \left[\sum_{\tau=s}^{S} \gamma_\tau(\hat{a}) - \rho_s(\hat{\boldsymbol{\gamma}}, \lambda, \hat{a}) \right] \geq \bar{u} + c(\hat{a}) \ , \qquad \text{(IR}'\text{)}$$

where

$$\rho_s(\hat{\boldsymbol{\gamma}}, \lambda, \hat{a}) := (\lambda - 1) \left[\sum_{\tau=s}^{S} \gamma_\tau(\hat{a}) \right] \left[\sum_{t=1}^{s-1} \gamma_t(\hat{a}) \right].$$

Let $\boldsymbol{\rho}(\hat{\boldsymbol{\gamma}}, \lambda, \hat{a}) = (\rho_2(\hat{\boldsymbol{\gamma}}, \lambda, \hat{a}), \ldots, \rho_S(\hat{\boldsymbol{\gamma}}, \lambda, \hat{a}))$. The first part of the agent's utility, $u_1 + \sum_{s=2}^{S} b_s(\sum_{\tau=s}^{S} \gamma_\tau(\hat{a}))$, is the expected intrinsic utility for money. Due to loss aversion, however, the agent's utility has a second negative component, the term $\boldsymbol{b}'\boldsymbol{\rho}(\hat{\boldsymbol{\gamma}}, \lambda, \hat{a})$. Where does this term come from? With bonus b_s being paid to the agent whenever a signal higher or equal to s is observed, the agent expects to receive b_s with probability $\sum_{\tau=s}^{S} \gamma_\tau(\hat{a})$. With probability $\sum_{t=1}^{s-1} \gamma_t(\hat{a})$, however, a signal below s will be observed, and the agent will not be paid bonus b_s. Thus, with "probability" $[\sum_{\tau=s}^{S} \gamma_\tau(\hat{a})][\sum_{t=1}^{s-1} \gamma_t(\hat{a})]$ the agent experiences a loss of λb_s. Analogous reasoning implies that the agent will experience a gain of b_s with the same probability. With losses looming larger than gains of equal size, in expectation the agent suffers from deviations from his reference point. This ex ante expected net loss is captured by the term, $\boldsymbol{b}'\boldsymbol{\rho}(\hat{\boldsymbol{\gamma}}, \lambda, \hat{a})$, which we will refer to as the agent's "loss premium".[29] A crucial point is that the loss premium increases in the complexity of the contract. When there is no wage differentiation at all, i.e., $\boldsymbol{b} = \boldsymbol{0}$, then the loss premium vanishes. If, in contrast, the contract specifies many different wage payments, then the agent ex ante considers a deviation from his reference point very likely. Put differently, for each additional wage payment an extra negative term enters the agent's loss premium and therefore reduces his expected utility.[30]

[29]Our notion of the agent's loss premium is highly related to the average self-distance of a lottery defined by Kőszegi and Rabin (2007). Let $D(\boldsymbol{u})$ be the average self-distance of incentive scheme $\boldsymbol{u}$, then $[(\lambda - 1)/2]D(\boldsymbol{u}) = \boldsymbol{b}'\boldsymbol{\rho}(\hat{\boldsymbol{\gamma}}, \lambda, \hat{a})$.

[30]While the exact change of the loss premium from adding more and more wage payments is hard to grasp, this point can heuristically be illustrated by considering the upper bound of the loss premium. Suppose the principal sets $n \leq S$ different

Given the first-order approach is valid, the incentive constraint can be rewritten as

$$\sum_{s=2}^{S} b_s \beta_s(\hat{\boldsymbol{\gamma}}, \lambda, \hat{a}) = c'(\hat{a}), \tag{IC$'$}$$

where

$$\beta_s(\hat{\boldsymbol{\gamma}}, \lambda, \hat{a}) := \left(\sum_{\tau=s}^{S} (\gamma_\tau^H - \gamma_\tau^L) \right) - (\lambda - 1)$$
$$\cdot \left[\left(\sum_{t=1}^{s-1} \gamma_t(\hat{a}) \right) \left(\sum_{\tau=s}^{S} (\gamma_\tau^H - \gamma_\tau^L) \right) + \left(\sum_{\tau=s}^{S} \gamma_\tau(\hat{a}) \right) \left(\sum_{t=1}^{s-1} (\gamma_t^H - \gamma_t^L) \right) \right].$$

Here, $\beta_s(\cdot)$ is the marginal effect on incentives of an increase in the wage payments for signals above $s-1$. Without loss aversion, i.e., $\lambda = 1$, this expression equals the marginal probability of observing at least signal s. If the agent is loss-averse, on the other hand, an increase in the action also affects the agent's loss premium. The probability of bearing a loss of size b_s is a quadratic function of the probability of observing at least signal s. The agent's action balances the tradeoff between maximizing intrinsic utility and minimizing the expected net loss. Depending on the precise signal structure and the action to be implemented, loss aversion may facilitate as well as hamper the creation of incentives. Let $\boldsymbol{\beta}(\hat{\boldsymbol{\gamma}}, \lambda, \hat{a}) = (\beta_2(\hat{\boldsymbol{\gamma}}, \lambda, \hat{a}), \ldots, \beta_S(\hat{\boldsymbol{\gamma}}, \lambda, \hat{a}))$.

As in the standard case, incentives are created solely by increases in intrinsic utilities, $\boldsymbol{b}$. In consequence, (IR$'$) is binding in the optimum. If this was not the case, i.e., if $\boldsymbol{b}$ satisfies (IC$'$) but (IR$'$) holds with strict inequality, then the principal can lower payment u_1 up to the point where the (IR$'$) is satisfied with equality. Thus, reducing u_1 while holding $\boldsymbol{b}$ constant lowers the principal's expected wage payment while preserving incentives.

It is obvious that (IC$'$) can only be satisfied if there exists at least one $\beta_s > 0$. If, for example, signals are ordered according to their likelihood

wages. It is readily verified that the loss premium is bounded from above by $(\lambda - 1)[(u_S - u_1)/2] \times [(n-1)/n]$, and that this upper bound increases as n increases. Note, however, that even for $n \to \infty$ the upper bound of the loss premium is finite.

ratios, then $\beta_s(\cdot) > 0$ for all $s = 2, \ldots, S$. More precisely, for a given ordering of signals, under A4 the following equivalence follows immediately from the fact that $\sum_{t=1}^{s-1}(\gamma_t^H - \gamma_t^L) = -\sum_{\tau=s}^{S}(\gamma_\tau^H - \gamma_\tau^L)$:

$$\beta_s(\hat{\boldsymbol{\gamma}}, \lambda, \hat{a}) > 0 \iff \sum_{\tau=s}^{S}(\gamma_\tau^H - \gamma_\tau^L) > 0. \tag{V.3}$$

V.3.1. TWO POLAR CASES: PURE RISK AVERSION VS. PURE LOSS AVERSION

In this part of the paper we analyze the two polar cases: The standard case where the agent is only risk-averse but not loss-averse, on the one hand, and the case of a loss-averse agent with a risk-neutral intrinsic utility function, on the other hand.

Pure Risk Aversion

First consider an agent who is risk-averse in the usual sense, i.e., $h''(\cdot) > 0$, but does not exhibit loss aversion. As discussed earlier, the latter requirement corresponds to the case where $\lambda = 1$. With the agent not being loss-averse, the first-order approach is valid even without Assumption A5.

PROPOSITION 14 (HOLMSTRÖM, 1979): Suppose A3 holds, $h''(\cdot) > 0$, and $\lambda = 1$. Then there exists a second-best optimal contract to implement $\hat{a} \in (0, 1)$. The second-best contract has the property that $u_s \neq u_\tau$ $\forall s, \tau \in \mathcal{S}$ and $s \neq \tau$. Moreover, $u_s > u_\tau$ if and only if $\gamma_s^H/\gamma_s^L > \gamma_\tau^H/\gamma_\tau^L$.

Proposition 14, restates the well-known finding by Holmström (1979) for discrete signals: Signals that are more indicative of higher effort, i.e., signals with a higher likelihood ratio γ_s^H/γ_s^L, are rewarded strictly higher. Thus, the optimal wage scheme is complex in the sense that it is fully differentiated, with each signal being rewarded differently.

Pure Loss Aversion

Having considered the polar case of pure risk aversion, we now turn to the other extreme, a purely loss-averse agent. Formally, intrinsic utility of money is a linear function, $h''(\cdot) = 0$, and the agent is loss-averse, $\lambda > 1$. As we have already reasoned, whatever contract the principal offers, relabeling the signals always allows us to represent this contract as an (at least weakly) increasing intrinsic utility profile. Therefore we can decompose the principal's problem into two steps: first, for a given ordering of signals, choose a nondecreasing profile of intrinsic utility levels that implements the desired action $\hat{a}$ at minimum cost; second, choose the signal ordering with the lowest cost of implementation. As we know from the discussion at the end of the previous section, a necessary condition for an upward-sloping incentive scheme to achieve incentive compatibility is that for the underlying signal ordering at least one $\beta_s(\cdot) > 0$. In what follows we restrict attention to the set of signal orderings that are incentive feasible in the afore-mentioned sense. Nonemptiness of this set follows immediately from Lemma 6.

Consider the first step of the principal's problem, i.e., taking the ordering of signals as given, find the nondecreasing payment scheme with the lowest cost of implementation. In what follows, we write the agent's intrinsic utility in terms of additional payments, $u_s = u_1 + \sum_{\tau=2}^{S} b_\tau$. With $h(\cdot)$ being linear, the principal's objective function is $C(u_1, \boldsymbol{b}) = u_1 + \sum_{s=2}^{S} b_s(\sum_{\tau=2}^{S} \gamma_\tau(\hat{a}))$. Remember that at the optimum, (IR$'$) holds with equality. Inserting (IR$'$) into the principal's objective allows us to write the cost minimization problem for a given order of signals in the following simple way:

Program ML:

$$\min_{\boldsymbol{b} \in \mathbb{R}_+^{S-1}} \ \boldsymbol{b}'\boldsymbol{\rho}(\hat{\boldsymbol{\gamma}}, \lambda, \hat{a})$$

$$\text{subject to} \quad \boldsymbol{b}'\boldsymbol{\beta}(\hat{\boldsymbol{\gamma}}, \lambda, \hat{a}) = c'(\hat{a}) \qquad \text{(IC}'\text{)}$$

The minimization problem (ML) has a simple intuition. The principal seeks to minimize the agent's expected net loss subject to the incentive compatibility constraint. Similar to the case of pure risk aversion, where the principal would like to cut back the agent's risk premium, here she is interested in minimizing the agent's loss premium. Due to the incentive constraint, however, this loss premium has to be strictly positive.

We want to emphasize that solving Program ML also yields insights for the more general case with a concave intrinsic utility function. Even though the principal's objective will not reduce to minimizing the agent's loss premium alone when intrinsic utility is non-linear, this nevertheless remains an important aspect of the principal's problem. Since the solution to Program ML tells us how to minimize the agent's loss premium irrespective of the functional form of intrinsic utility, one should expect its properties to carry over to some extent to the solution of the more general problem.

The principal's cost minimization problem for a given order of signals is a simple linear programming problem: minimize a linear objective function subject to one linear equality constraint. Since we restricted attention to orderings of signals with $\beta_s(\cdot) > 0$ for at least one signal s, a solution to (ML) exists. Due to the linear nature of problem (ML), (generically) this solution sets exactly one $b_s > 0$ and all other $b_s = 0$. Put differently, the problem is to find that b_s which creates incentives at the lowest cost.

So far we have seen that, for a given ordering of signals, the principal considers it optimal to offer the agent a bonus contract: pay a low wage for signals below some threshold, and a high wage for signals above this threshold. What remains to do for the principal, in a second step, is to find the signal ordering that leads to the lowest cost of implementation. With the number of different orders of signals being finite, this problem clearly has a solution.

Before summarizing the above discussion more concisely, we want to relate our finding to the benchmark case of pure risk neutrality. As is well-known, with both contracting parties being risk (and loss) neutral a broad range of contracts – including simple bonus schemes – is optimal. With the agent being loss-averse even to a negligible degree, however, the unique optimal contractual form is a bonus scheme.[31]

PROPOSITION 15: Suppose A3-A5 hold, $h''(\cdot) = 0$ and $\lambda > 1$. Then there exists a second-best optimal contract to implement action $\hat{a} \in (0, 1)$. Generically, the second-best optimal incentive scheme $\{u_s^*\}_{s=1}^S$ is a bonus contract, i.e., $u_s^* = u_H^*$ for $s \in \mathcal{B}^* \subset \mathcal{S}$ and $u_s^* = u_L^*$ for $s \in \mathcal{S} \setminus \mathcal{B}^*$, where $u_H^* > u_L^*$.

According to Proposition 15, the principal considers it optimal to offer the agent a bonus contract which entails only a minimum degree of wage differentiation in the sense that, no matter how rich the signal space, the contract specifies only two different wage payments. This endeavor to reduce the complexity of the contract is plausible, since a high degree of wage differentiation increases the agent's loss premium: With the employment contract she offers to the agent, the principal determines the dimensionality of the agent's reference point. The higher the dimensionality of the reference point is, the more likely it is that the agent incurs a loss in a particular dimension. Therefore, with the concept of reference-dependent preferences developed by Kőszegi and Rabin, it truly pains a person to be exposed to numerous potential outcomes. This disutility of the agent from facing several possible (monetary) outcomes, which he demands to be compensated for, makes it costly for the principal to offer complex contracts. In consequence, the optimal contract entails only a minimum of wage differentiation. To provide a more intuitive explanation for this finding, consider a principal who – starting out from a given wage scheme – has to improve incentives. There are basically two ways to do so. On the one hand, the principal can introduce a new wage spread, i.e., pay slightly different wages for two signals that were rewarded equally

[31]If, in addition to both the principal and the agent being risk-neutral, the agent is protected by limited liability, Park (1995), Kim (1997), Oyer (2000), and Demougin and Fluet (1998) show that the optimal contract is a bonus scheme. These findings, however, immediately collapse when the agent is somewhat risk-averse as is demonstrated by Jewitt et al. (2008). Our findings, on the other hand, are robust towards introducing a slightly concave intrinsic utility function, as we will illustrate in Section V.3.2.

in the original wage scheme, while keeping the differences between all other neighboring wages constant. On the other hand, the principal can increase an existing wage spread, holding constant all other spreads between neighboring wages. Both procedures increase the loss premium by increasing the size of some of the the expected losses without reducing others. Introducing a new wage spread, however, additionally increases the loss premium by increasing the ex ante expected probability of experiencing a loss. Therefore, in order to improve incentives for a loss-averse agent, it is advantageous to increase a particular existing wage spread without adding to the contractual complexity in the sense of increasing the number of different wages. Under the standard notion of a risk-averse agent, however, one should not expect to encounter this tendency to reduce the complexity of contracts. The reason is that increasing incentives by introducing a small new wage spread is basically costless for the principal because locally the agent is risk-neutral. Therefore, under risk aversion different signals are rewarded differently.

Up to now, however, we have not specified which signals are generally included in the set $\mathcal{B}^*$. In light of the above observation, the principal's problem boils down to choosing a binary partition of the set of signals, $\mathcal{B} \subset \mathcal{S}$, which characterizes for which signals the agent receives the high wage and for which signals he receives the low wage. The wages u_L and u_H are then uniquely determined by the corresponding individual rationality and incentive compatibility constraints. The problem of choosing the optimal partition of signals, $\mathcal{B}^*$, which minimizes the principal's expected cost of implementing action $\hat{a}$ is an integer programming problem. As is typical for this class of problems, and as is nicely illustrated by the well-known "0-1 Knapsack Problem", it is impossible to provide a general characterization of the solution.[32]

[32]The "0-1 Knapsack Problem" refers to a hiker who has to select from a group of items, all of which may be suitable for her trip, a subset that has greatest value while not exceeding the capacity of her knapsack. Suppose there are n items, each item j has a value $v_j > 0$ and a weight $w_j > 0$. Let the capacity of the knapsack be $c > 0$. The 0-1 Knapsack Problem may be formulated as the following maximization problem: max $\sum_{j=1}^{n} v_j x_j$ subject to $\sum_{j=1}^{n} w_j x_j \leq c$ and $x_j \in \{0, 1\}$ for $j = 1, \ldots, n$.

Next to these standard intricacies of integer programming, there is an additional difficulty in our model: the principal's objective behaves non-monotonically when including an additional signal into the "bonus set" $\mathcal{B}$. This is due to different – possibly conflicting – targets that the principal pursues when deciding how to partition the set $\mathcal{S}$. From Program (ML) it follows that, for a given "bonus set" $\mathcal{B}$, the minimum cost of implementing action $\hat{a}$ is given by

$$C_{\mathcal{B}} = \bar{u} + c(\hat{a}) + \frac{c'(\hat{a})(\lambda - 1)P_{\mathcal{B}}(1 - P_{\mathcal{B}})}{[\sum_{s\in\mathcal{B}} \gamma_s^H - \gamma_s^L][1 - (\lambda - 1)(1 - 2P_{\mathcal{B}})]}\ , \tag{V.4}$$

where $P_{\mathcal{B}} := \sum_{s\in\mathcal{B}} \gamma_s(\hat{a})$. The above costs can be rewritten such that the principal's problem amounts to

$$\max_{\mathcal{B}\subset\mathcal{S}} \left[\sum_{s\in\mathcal{B}} (\gamma_s^H - \gamma_s^L)\right] \left\{\frac{1}{(\lambda - 1)P_{\mathcal{B}}(1 - P_{\mathcal{B}})} - \frac{1}{P_{\mathcal{B}}} + \frac{1}{1 - P_{\mathcal{B}}}\right\}. \tag{V.5}$$

This objective function illustrates the tradeoff that the principal faces when deciding how to partition the signal space. The first term, $\sum_{s\in\mathcal{B}}(\gamma_s^H - \gamma_s^L)$, is the aggregate marginal impact of effort on the probability of the bonus $b := u_H - u_L$ being paid out. In order to create incentives for the agent, the principal would like to make this term as large as possible, which in turn allows her to lower the bonus payment. This can be achieved by including only good signals in $\mathcal{B}$. The second term, on the other hand, is maximized by making the probability of paying the agent the high wage either as large as possible or as small as possible, depending on the exact signal structure and the action to be implemented. With the loss premium being given by $(\lambda - 1)P_{\mathcal{B}}(1 - P_{\mathcal{B}})b$, this is intuitive: By making the event of paying the high wage very likely or unlikely, the principal minimizes the scope for the agent to experience a loss that he demands to be compensated for. Depending on the signal structure, these two goals may conflict with each other, which makes a complete characterization of the optimal contract very intricate. Nevertheless, it can be shown that the optimal contract displays the following plausible property.

PROPOSITION 16: Let $\mathcal{S}^+ \equiv \{s \in \mathcal{S} | \gamma_s^H - \gamma_s^L > 0\}$. The optimal partition of the signals for which the high wage is paid, $\mathcal{B}^*$, has the following property: either $\mathcal{B}^* \subseteq \mathcal{S}^+$ or $\mathcal{S}^+ \subseteq \mathcal{B}^*$.

Put verbally, the optimal partition of the signal set takes one of the two possible forms: the high wage is paid out to the agent (i) either only for good signals though possibly not for all good signals, or (ii) for all good signals and possibly a few bad signals as well. Loosely speaking, if the principal considers it optimal to pay the high wage very rarely, she will reward only good signals with the extra payment b. If, on the other hand, she wants the agent to receive the high wage with high probability, then she will reward at least all good signals.

Without further assumptions, due to the discrete nature of the problem it is hard to characterize the signals that are included in $\mathcal{B}^*$. Back to the "0-1 Knapsack Problem", here it is well-established for the continuous version of the problem that the solution can easily be found by ordering the items according to their value-to-weight ratio.[33] Defining $\kappa := \max_{\{s,t\}\subseteq\mathcal{S}} |\gamma_s(\hat{a}) - \gamma_t(\hat{a})|$, we can obtain a similar result. Assuming that κ is sufficiently small, which is likely to hold if the performance measure is, for instance, sales revenues measured in cents, makes the principal's problem of choosing $\mathcal{B}^*$ similar to a continuous problem.[34] With this assumption, we can show that it is optimal to order the signals according to their likelihood ratios.

P R O P O S I T I O N 17: Suppose κ is sufficiently small, then there exists a constant K such that $\mathcal{B}^* = \{s \in \mathcal{S} \mid \gamma_s^H/\gamma_s^L \geq K\}$.

Though wage payments are only weakly increasing in the likelihood ratio, this finding resembles the standard result for a risk-averse agent, where the incentive scheme is strictly increasing in the likelihood ratio.

Before moving on to the discussion of the more general case in which the agent is both risk and loss-averse, we want to pause to point out the following comparative static results.

P R O P O S I T I O N 18: An increase in the agent's degree of loss aversion (i) strictly increases the minimum cost of implementing action $\hat{a}$;

[33] In the continuous "0-1 Knapsack Problem" the constraints on the variables $x_j \in \{0,1\}$ are relaxed to $x_j \in [0,1]$, e.g. Dantzig (1957).

[34] Here, the probability of observing a specific signal, say, sales revenues of exactly \$13,825.32 is rather small.

(ii) decreases the necessary wage spread to implement action $\hat{a}$ if and only if $P_{\mathcal{B}^*} > 1/2$, given that the change in λ does not lead to a change of $\mathcal{B}^*$.

Part (ii) of Proposition 18 relates to the reasoning by Kőszegi and Rabin (2006) that if the agent is loss-averse and expectations are the driving force in the determination of the reference point, then "in principal-agent models, performance-contingent pay may not only directly motivate the agent to work harder in pursuit of higher income, but also indirectly motivate [him] by changing [his] expected income and effort." As can be seen from (V.1), the agent's expected utility under the second-best contract comprises of two components, the first of which is expected net intrinsic utility from choosing effort level $\hat{a}$, $u_L + b^* \sum_{s \in \mathcal{B}^*} \gamma_s(\hat{a}) - c(\hat{a})$. Due to loss aversion, however, there is a second component: With losses looming larger than equally sized gains, in expectation the agent suffers from deviations from his reference point. While the strength of this effect is determined by the degree of the agent's loss aversion, λ, his action choice – together with the signal parameters – determines the probability that such a deviation from the reference point actually occurs. We refer to this probability, which is given by $P_{\mathcal{B}^*}(1 - P_{\mathcal{B}^*})$, as loss probability. Therefore, when choosing his action, the agent has to balance off two possibly conflicting targets, maximizing expected net intrinsic utility and minimizing the loss probability. The loss probability, which is a strictly concave function of the agent's effort, is locally decreasing at $\hat{a}$ if and only if $P_{\mathcal{B}^*} > 1/2$. In this case, an increase in λ, which makes reducing the loss probability more important, may lead to the agent choosing a higher effort level, which in turn allows the principal to use lower-powered incentives. The principal, however, cannot capitalize on this since, according to part (i) of Proposition 18, the overall cost of implementation strictly increases in the agent's degree of loss aversion.

V.3.2. THE GENERAL CASE: LOSS AVERSION AND RISK AVERSION

We now turn to the intermediate case where the agent is both risk and loss-averse. The agent's intrinsic utility for money is a strictly increasing and strictly concave function, which implies that $h(\cdot)$ is strictly increasing and strictly convex. Moreover, the agent is loss-averse, i.e., $\lambda > 1$. From Lemma 6, we know that the constraint set of the principal's problem is nonempty. By relabeling signals, each contract can be interpreted as a contract that offers the agent a (weakly) increasing intrinsic utility profile. This allows us to assess whether the agent perceives receiving u_s instead of u_t as a gain or a loss. As in the case of pure loss aversion, we analyze the optimal contract for a given feasible ordering of signals.

The principal's problem for a given arrangement of the signals is given by

Program MG:

$$\min_{u_1,\ldots,u_S} \sum_{s=1}^{S} \gamma_s(\hat{a}) h(u_s) \qquad \text{subject to}$$

$$\sum_{s=1}^{S} \gamma_s(\hat{a}) u_s - (\lambda - 1) \sum_{s=1}^{S-1} \sum_{t=s+1}^{S} \gamma_s(\hat{a})\gamma_t(\hat{a})[u_t - u_s] - c(\hat{a}) = \bar{u} \ , \quad (\text{IR}_G)$$

$$\sum_{s=1}^{S} (\gamma_s^H - \gamma_s^L) u_s - (\lambda - 1) \cdot \sum_{s=1}^{S-1} \sum_{t=s+1}^{S} \left[\gamma_s(\hat{a})(\gamma_t^H - \gamma_t^L) + \gamma_t(\hat{a})(\gamma_s^H - \gamma_s^L)\right] [u_t - u_s] = c'(\hat{a}) \ , \quad (\text{IC}_G)$$

$$u_S \geq u_{S-1} \geq \ldots \geq u_1. \quad (\text{OC}_G)$$

Since the objective function is strictly convex and the constraints are all linear in $\boldsymbol{u} = \{u_1, \ldots, u_S\}$, the Kuhn-Tucker theorem yields necessary and sufficient conditions for optimality. Put differently, if there exists a solution to the problem (MG) the solution is characterized by the partial derivatives of the Lagrangian associated with (MG) set equal to zero.

Lemma 7: Suppose A3-A5 hold and $h''(\cdot) > 0$, then there exists a second-best optimal incentive scheme for implementing action $\hat{a} \in (0, 1)$, denoted $\boldsymbol{u}^* = \{u_1^*, \ldots, u_S^*\}$.

In order to interpret the first-order conditions of the Lagrangian to problem (MG) it is necessary to know whether the Lagrangian multipliers are positive or negative.

LEMMA 8: The Lagrangian multipliers of program (MG) associated with the incentive compatibility constraint and the individual rationality constraint are both strictly positive, i.e., $\mu_{IC} > 0$ and $\mu_{IR} > 0$.

We now give a heuristic reasoning why pooling of information may well be optimal in this more general case. For the sake of argument, suppose there is no pooling of information in the sense that it is optimal to set distinct wages for distinct signals. In this case all order constraints are slack; formally, if $u_s \neq u_{s'}$ for all $s, s' \in \mathcal{S}$ and $s \neq s'$, then $\mu_{O,s} = 0$ for all $s \in \{2, \dots, S\}$. In this case, i.e., when none of the ordering constraints is binding, then the first-order condition of optimality with respect to u_s, $\partial \mathcal{L}(\boldsymbol{u})/\partial u_s = 0$, can be written as follows:

$$h'(u_s) = \underbrace{\left(\mu_{IR} + \mu_{IC}\frac{\gamma_s^H - \gamma_s^L}{\gamma_s(\hat{a})}\right)}_{=:H_s} \underbrace{\left[1 - (\lambda - 1)\left(2\sum_{t=1}^{s-1}\gamma_t(\hat{a}) + \gamma_s(\hat{a}) - 1\right)\right]}_{=:\Gamma_s} - \underbrace{\mu_{IC}(\lambda - 1)\left[2\sum_{t=1}^{s-1}(\gamma_t^H - \gamma_t^L) + (\gamma_s^H - \gamma_s^L)\right]}_{=:\Lambda_s}. \quad \text{(V.6)}$$

For $\lambda = 1$ we have $h'(u_s) = H_s$, the standard "Holmström-formula".[35] Note that $\Gamma_s > 0$ for $\lambda \leq 2$. More importantly, irrespective of the signal ordering, we have $\Gamma_s > \Gamma_{s+1}$. The third term, Λ_s, can be either positive or negative. If the compound signal of all signals below s and the signal s itself are bad signals, then $\Lambda_s < 0$.

Since the incentive scheme is nondecreasing, when the order constraints are not binding it has to hold that $h'(u_s) \geq h'(u_{s-1})$. Thus, if $\mu_{OC,s-1} = \mu_{OC,s} = \mu_{OC,s+1} = 0$ the following inequality is satisfied:

$$H_s \times \Gamma_s - \Lambda_s \geq H_{s-1} \times \Gamma_{s-1} - \Lambda_{s-1}. \quad \text{(V.7)}$$

[35]See Holmström (1979). This formula is also referred to as the modified Borch sharing rule (Borch, 1962).

Note that for the given ordering of signals, if there exists any pair of signals $s, s-1$ such that (V.7) is violated, then the optimal contract for this ordering involves pooling of wages. Even when $H_s > H_{s-1}$, as it is the case when signals are ordered according to their likelihood ratio, it is not clear that inequality (V.7) is satisfied. In particular, when s and $s-1$ are similarly informative it seems to be optimal to pay the same wage for these two signals as can easily be illustrated for the case of two good signals: If s and $s-1$ are similarly informative good signals then $H_s \approx H_{s-1} > 0$ but $\Gamma_s < \Gamma_{s-1}$ and $\Lambda_s > \Lambda_{s-1}$, thus condition (V.7) is violated. In summary, it may well be that for a given incentive-feasible ordering of signals, and thus overall as well, the order constraints are binding, i.e., it may be optimal to offer a contract which is less complex than the signal space allows for. We illustrate this conjecture in the following with an example.

Application with Constant Relative Risk Aversion

In the general case of a risk and loss-averse agent the principal seeks to minimize the loss and the risk premium. Roughly speaking, the risk premium is increasing in the curvature of the agent's intrinsic utility function. Put differently, when the agent's intrinsic utility function becomes close to linearity the risk premium goes to zero. Thus, for a slightly concave intrinsic utility function one should expect that the principal's main objective is to minimize the loss premium, which is achieved by a bonus scheme as is shown in Section V.3.1. In the following we show that these reasoning is correct for the case of an intrinsic utility function that features constant relative risk aversion (CRRA).

Suppose $h(u) = u^r$, with $r \geq 0$ being a measure for the agent's risk aversion. More precisely, the Arrow-Pratt measure for relative risk aversion of the agent's intrinsic utility function is $R = 1 - \frac{1}{r}$ and therefore constant. The following result states that the optimal contract is still a bonus contract when the agent is not only loss-averse, but also slightly risk-averse.

P r o p o s i t i o n 19: Suppose A3-A5 hold, $h(u) = u^r$ with $r > 1$, and $\lambda > 1$. Generically, for r sufficiently small the optimal incentive scheme $\{u_s^*\}_{s=1}^S$ is a bonus scheme, i.e., $u_s^* = u_H^*$ for $s \in \mathcal{B}^* \subset \mathcal{S}$ and $u_s^* = u_L^*$ for $s \in \mathcal{S} \backslash \mathcal{B}^*$ where $u_L^* < u_H^*$.

Without loss aversion, in contrast, according to Proposition 14 the optimal contract is fully differentiated even for intrinsic utility being arbitrarily close to linearity.

Next, we demonstrate by means of an example that pooling of signals may well be optimal even for a non-negligible degree of risk aversion. Suppose the agent's effort cost is $c(a) = (1/2)a^2$ and the effort level to be implemented is $\hat{a} = \frac{1}{2}$. Moreover, we assume that the reservation utility $\bar{u} = 10$, which guarantees that all utility levels are positive.[36] To keep the example as simple as possible, it is assumed that the agent's performance can take only three values, i.e., the agent's performance is either excellent (E), satisfactory (S) or inadequate (I). We consider two specifications of the performance measure. In the first specification the satisfactory signal is a good signal, whereas in the second specification it is a bad signal. In all parameter constellations we consider, it turns out that it is always (weakly) optimal to order signals according to their likelihood ratio, i.e., $u_1 = u_I$, $u_2 = u_S$ and $u_3 = u_E$. In the first specification the conditional probabilities take the following values:

$$\gamma_E^H = 5/10 \qquad \gamma_E^L = 1/10$$
$$\gamma_S^H = 4/10 \qquad \gamma_S^L = 3/10$$
$$\gamma_I^H = 1/10 \qquad \gamma_I^L = 6/10.$$

The structure of the optimal contract for this specification and various values of r and λ is presented in Table V.1.

Table V.1 suggests that the optimal contract typically involves pooling of the two good signals, in particular when the agent's intrinsic utility is not too concave, i.e., if the agent is not too risk-averse. Table V.1 nicely illustrates the trade-off the principal faces when the agent is both, risk and loss-averse: If the agent becomes more risk-averse pooling is less

[36]Increasing $\bar{u}$ makes the agent less (absolutely) risk-averse and thus is similar to a reduction in r.

r \ λ	1.0	1.1	1.3	1.5
1.5	$u_1 < u_2 < u_3$	$u_1 < u_2 = u_3$	$u_1 < u_2 = u_3$	$u_1 < u_2 = u_3$
2	$u_1 < u_2 < u_3$	$u_1 < u_2 < u_3$	$u_1 < u_2 = u_3$	$u_1 < u_2 = u_3$
3	$u_1 < u_2 < u_3$	$u_1 < u_2 < u_3$	$u_1 < u_2 = u_3$	$u_1 < u_2 = u_3$

Table V.1: Structure of the optimal contract with two "good" signals.

likely to be optimal. If, on the other hand, he becomes more loss-averse, pooling is more likely to be optimal.[37]

In the second specification we assume that there are two bad signals. The conditional probabilities are as follows:

$$\gamma_E^H = 6/10 \qquad \gamma_E^L = 1/10$$
$$\gamma_S^H = 2/10 \qquad \gamma_S^L = 4/10$$
$$\gamma_I^H = 2/10 \qquad \gamma_I^L = 5/10.$$

The results for this case are presented in Table V.2.

r \ λ	1.0	1.1	1.3	1.5
1.5	$u_1 < u_2 < u_3$	$u_1 = u_2 < u_3$	$u_1 = u_2 < u_3$	$u_1 = u_2 < u_3$
2	$u_1 < u_2 < u_3$	$u_1 = u_2 < u_3$	$u_1 = u_2 < u_3$	$u_1 = u_2 < u_3$
3	$u_1 < u_2 < u_3$	$u_1 = u_2 < u_3$	$u_1 = u_2 < u_3$	$u_1 = u_2 < u_3$

Table V.2: Structure of the optimal contract with two "bad" signals.

In this specification, a binary statistic that pools the two bad signals seems to be optimal almost always. The reason behind this observation is that the two bad signals are very similar. In consequence, paying the same wage for satisfactory as well as inadequate performance increases

[37]For a given r, the degree of pooling does not monotonically increase in λ. As discussed at the end of Section V.3.1, a higher degree of loss aversion of the agent may help the principal to create incentives. If this is the case, a contract that contains less pooling is preferred from an incentive point of view. If this positive effect of less pooling on incentives outweighs the negative effect on the agent's loss premium, then the optimal contract consists of more distinct wage payments when λ increases. This can happen, however, only locally: at some point the degree of pooling increases in λ.

the risk premium only slightly. On the other hand, by pooling satisfactory and inadequate performance it becomes less likely for the agent ex ante to experience a loss, i.e., the loss premium is reduced. Therefore, it is optimal for the principal to use a bonus scheme even when the agent's degree of loss aversion is small.

V.4. IMPLEMENTATION PROBLEMS, TURNING A BLIND EYE, AND STOCHSTIC CONTRACTS

In this section, we relax the assumptions that guaranteed the validity of the first-order approach. In particular, in order to explore the implications of a higher degree of loss aversion, we relax A4. We restrict attention to two simplifications of the former model. First, we return to the assumption of a purely loss-averse agent. Second, only binary measures of performance are considered.

V.4.1. THE CASE OF A BINARY MEASURE OF PERFORMANCE

As before, the principal cannot observe the agent's action a or whether the benefit B was realized or not. Instead she observes a contractible binary measure of performance, i.e., $\mathcal{S} = \{1, 2\}$. For notational convenience , let $(1-\gamma^H)$ and γ^H denote the probabilities of observing signal $s = 1$ and $s = 2$, respectively, conditional on B being realized. Accordingly, $(1-\gamma^L)$ and γ^L are the probabilities of observing signal $s = 1$ and $s = 2$, respectively, conditional on B not being realized.[38] Thus, the unconditional probability of observing signal $s = 2$ for a given action a is $\gamma(a) \equiv a\gamma^H + (1-a)\gamma^L$. Let $\hat{\boldsymbol{\gamma}} = (\gamma^H, \gamma^L)$. We reformulate A3 for the binary case as follows.

ASSUMPTION A6: $\quad 1 > \gamma^H > \gamma^L > 0$.

With only two possible signals to be observed, the contract takes the form of a bonus contract: the agent is paid a base wage which yields

[38] In the notation introduced above, we have $\gamma_1^H = 1-\gamma^H$, $\gamma_2^H = \gamma^H$, $\gamma_1^L = 1-\gamma^L$ and $\gamma_2^L = \gamma^L$.

intrinsic utility u if the bad signal is observed, and he is paid the base wage plus a bonus b resulting in intrinsic utility $u + b$ if the good signal is observed. For now assume that $b \geq 0$.[39] We assume that the agent's intrinsic disutility of effort is a quadratic function, $c(a) = (k/2)a^2$.[40] The agent's expected utility from choosing effort level a then is

$$E\left[U(a)\right] = u + \gamma(a)b - \frac{k}{2}a^2 - (\lambda - 1)\gamma(a)(1 - \gamma(a))b. \tag{V.8}$$

As before, the first component is expected net intrinsic utility from choosing effort level a, that is, expected wage payment minus effort cost. The second component is the loss premium, with $\gamma(a)(1-\gamma(a))$ denoting the loss probability.

V.4.2. INVALIDITY OF THE FIRST-ORDER APPROACH

The first derivative of expected utility with respect to effort is given by

$$E\left[U'(a)\right] = \underbrace{(\gamma^H - \gamma^L)b\left[2 - \lambda + 2\gamma(a)(\lambda - 1)\right]}_{MB(a)} - \underbrace{ka}_{MC(a)}. \tag{V.9}$$

While the marginal cost, $MC(a)$, obviously is a straight line through the origin with slope k, the marginal benefit, $MB(a)$, also is a positively sloped, linear function of effort a. An increase in b unambiguously makes $MB(a)$ steeper. Letting a_0 denote the intercept of $MB(a)$ with the horizontal axis, we have

$$a_0 = \frac{\lambda - 2 - 2\gamma^L(\lambda - 1)}{2(\gamma^H - \gamma^L)(\lambda - 1)}.$$

The cases for $a_0 < 0$ and $a_0 > 0$ are depicted in Figures V.1 and V.2, respectively. Implementation problems in our sense refer to a situation where there are actions $a \in (0, 1)$ that are not incentive compatible for any bonus payment.

[39]The assumption $b \geq 0$ is made only for expositional purposes, the results hold true for $b \in \mathbb{R}$.

[40]This functional form does not fit exactly the assumptions on $c(\cdot)$ that we imposed above, but is made for expositional convenience. Allowing for more general effort cost functions does not qualitatively change the insights that are to be obtained.

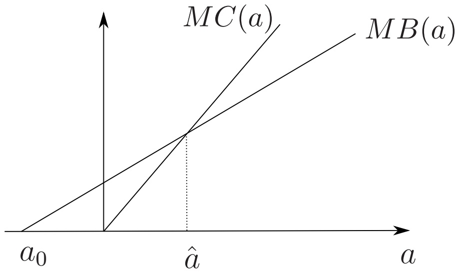

Figure V.1: $MB(a)$ and $MC(a)$ for $a_0 < 0$.

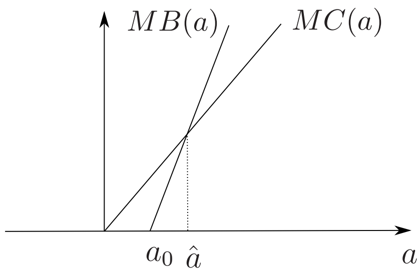

Figure V.2: $MB(a)$ and $MC(a)$ for $a_0 > 0$.

P R O P O S I T I O N 20: Suppose A6 holds, then effort level $\hat{a} \in (0, 1)$ is implementable if and only if $a_0 \leq 0$.

Implementation problems arise when $a_0 > 0$, or equivalently, when $\gamma^L < 1/2$ and $\lambda > 2(1 - \gamma^L)/(1 - 2\gamma^L) > 2$. Somewhat surprisingly, this includes performance measures with $\gamma^L < 1/2 < \gamma^H$, which (possibly) are highly informative. Informative in this context means that it is more likely to observe the bad signal if benefit B was not realized, whereas it is more likely to observe the good signal if B was realized. So, why do these implementation problems arise in the first place? Remember that the agent has two targets: First, as in classic models, he seeks to maximize net intrinsic utility, $u + b\gamma(a) - (k/2)a^2$. When the agent cares only about this net intrinsic utility (e.g., he is loss neutral) then each action can be implemented by choosing a sufficiently high bonus. Due to loss aversion, however, the agent has a second target which is minimizing the expected loss. How can the agent pursue this goal? He can do so by choosing an action such that the loss probability, $\gamma(a)(1 - \gamma(a))$, becomes small. The crucial point is that these two targets may conflict with each other in the sense that an increase in effort may increase net intrinsic utility but at the same time also increases the loss probability. First of all, note that implementation problems never arise when $\gamma^L \geq 1/2$ or $\lambda \leq 2$. For $\gamma^L \geq 1/2$, the loss probability is strictly decreasing in the agent's action. Consequently, with both targets of the agent being aligned, an increase in the bonus unambiguously leads to an increase in the agent's action. For $\lambda \leq 2$, the weight put on gain-loss utility, $\lambda - 1$, is lower than the weight put on intrinsic utility, so the agent is more interested in maximizing net intrinsic utility than in minimizing the loss probability. For $\gamma^L < 1/2$, on the other hand, implementation problems do arise when λ is sufficiently large. Roughly speaking, being sufficiently loss-averse, the agent primarily cares about reducing the loss probability. With the loss probability being inverted U-shaped, the agent achieves this by choosing one of the two extreme actions $a \in \{0, 1\}$. Therefore, the principal cannot motivate the agent to choose an action $\hat{a} \in (0, 1)$ when $\gamma^L < 1/2$ and the agent's loss aversion is sufficiently severe.

V.4.3. TURNING A BLIND EYE

As we have seen in the preceding analysis, the principal faces implementation problems whenever $a_0 > 0$. One might wonder if there is a remedy for these implementation problems. The answer is "yes", there is a remedy. The principal can manipulate the signal in her favor by not paying attention to the signal from time to time but nevertheless paying the bonus in these cases. Formally, suppose the principal commits herself to stochastically ignoring the signal with probability $p \in [0, 1)$.[41] Thus, the overall probability of receiving the bonus is given by $\gamma(a; p) \equiv p + (1 - p)\gamma(a)$. This strategic ignorance of information gives rise to a transformed performance measure $\hat{\boldsymbol{\gamma}}(p) = (\gamma^H(p), \gamma^L(p))$. As before, $\gamma^H(p)$ denotes the probability that the bonus is paid to the agent conditional on benefit B being realized. Given that B is realized, this happens either when the performance measure is ignored, or – if the principal pays attention to the performance measure – when the good signal is realized. Hence, $\gamma^H(p) = p + (1 - p)\gamma^H$. Analogously, the probability of the bonus being paid out conditional on B not being realized is given by $\gamma^L(p) = p + (1 - p)\gamma^L$. As it turns out, ignoring the whole performance measure with probability p is formally equivalent to ignoring only the

[41] Always ignoring the signal, i.e., setting $p = 1$, would be detrimental for incentives because then the agent's monetary payoff is independent of his action. Hence, he would choose the least cost action $a = 0$. Therefore, we a priori restrict the principal to choose p from the interval $[0, 1)$.

bad signal with probability p.[42] For this reason, we refer to the principal not paying attention to the performance measure as turning a blind eye on bad performance of the agent. It is readily verified that under the transformed performance measure $\hat{\boldsymbol{\gamma}}(p)$ the intercept of the $MB(a)$ function with the horizontal axis,

$$a_0(p) \equiv \frac{\lambda - 2 - 2\left[p + (1-p)\gamma^L\right](\lambda - 1)}{2(1-p)(\gamma^H - \gamma^L)(\lambda - 1)},$$

not only is decreasing in p but also can be made arbitrarily small, in particular, arbitrarily negative. Formally, $da_0(p)/dp < 0$ and $\lim_{p \to 1} a_0(p) = -\infty$. In the light of Proposition 20 this immediately implies that the principal can eliminate any implementation problems by choosing p sufficiently high, that is, by turning a blind eye sufficiently often.

Besides alleviating possible implementation problems, turning a blind eye can also benefit the principal from a cost perspective. Using the definition of $\gamma(a;p)$ it can be shown that the minimum cost of implementing action $\hat{a}$ under the transformed performance measure, $C(\hat{a};p)$, takes the following form:

$$C(\hat{a};p) = \overline{u} + \frac{k}{2}\hat{a}^2 + \frac{k\hat{a}(\lambda - 1)(1 - \gamma(\hat{a}))}{(\gamma^H - \gamma^L)} \cdot \frac{\gamma(\hat{a}) + p(1 - \gamma(\hat{a}))}{1 - (\lambda - 1)\left[1 - 2\gamma(\hat{a}) - 2p(1 - \gamma(\hat{a}))\right]} \quad \text{(V.10)}$$

Differentiating the principal's cost with respect to p reveals that sign $\{dC(\hat{a};p)/dp\}$ = sign$\{2 - \lambda\}$. Hence, an increase in the probability of ignoring the bad signal decreases the cost of implementing a certain action if and only if $\lambda > 2$. Hence, whenever the principal turns a blind eye in order to remedy implementation problems, she will do so to the

[42] In this latter case, the agent receives the bonus either when the good signal is observed, which happens with probability $\gamma(a)$, or when the bad signal is observed but is ignored, which happens with probability $(1 - \gamma(a))p$. Hence, the overall probability of the bonus being paid out is given by $\gamma(a) + (1 - \gamma(a))p$.

largest possible extent.[43,44] We summarize the preceding analysis in the following proposition.

PROPOSITION 21: Suppose the principal can commit herself to stochastic ignorance of the signal. Then each action $\hat{a} \in [0,1]$ can be implemented. Moreover, the implementation costs are strictly decreasing in p if and only if $\lambda > 2$.

We restricted the principal to offer non-stochastic payments conditional on which signal is observed. If the principal was able to do just that, then she could remedy implementation problems by paying the base wage plus a lottery in the case of the bad signal. For instance, when the lottery yields b with probability p and zero otherwise, this is just the same as turning a blind eye. This observation suggests that the principal may benefit from offering a contract that includes randomization, which is in contrast to the finding under conventional risk aversion, see Holmström (1979).[45]

V.4.4. BLACKWELL REVISITED

We conclude this section by briefly pointing out an interesting implication of the above analysis. Suppose the principal has no access to a randomization device, i.e., turning a blind eye is not possible. Then the above considerations allow a straight-forward comparison of performance

[43]Formally, for $\lambda > 2$, the solution to the principal's problem of choosing the optimal probability to turn a blind eye, p^*, is not well defined because $p^* \to 1$. If the agent is subject to limited liability or there is a cost of ignorance, however, the optimal probability of turning a blind eye is well defined.

[44]This is in the spirit of Becker and Stigler (1974), who show that despite a small detection probability of malfeasance, incentives can be maintained if the punishment is sufficiently severe.

[45]The finding that stochastic contracts may be optimal is not novel to the principal-agent literature. Haller (1985) shows that in the case of a satisficing agent, who wants to achieve certain aspiration levels of income with certain probabilities, randomization may pay for the principal. Moreover, Strausz (2006) finds that deterministic contracts may be suboptimal in a screening context.

measures $\hat{\boldsymbol{\zeta}} = (\zeta^H, \zeta^L)$ and $\hat{\boldsymbol{\gamma}} = (\gamma^H, \gamma^L)$ if $\hat{\boldsymbol{\zeta}}$ is a convex combination of $\hat{\boldsymbol{\gamma}}$ and $\mathbf{1} \equiv (1, 1)$.

COROLLARY 1: Let $\hat{\boldsymbol{\zeta}} = p\mathbf{1} + (1-p)\hat{\boldsymbol{\gamma}}$ with $p \in (0, 1)$. Then the principal at least weakly prefers performance measure $\hat{\boldsymbol{\zeta}}$ to $\hat{\boldsymbol{\gamma}}$ if and only if $\lambda \geq 2$.

The finding that the principal prefers the "garbled" performance measure $\hat{\boldsymbol{\zeta}}$ over performance measure $\hat{\boldsymbol{\gamma}}$ is at odds with Blackwell's theorem. To see this, let performance measures $\hat{\boldsymbol{\gamma}}$ and $\hat{\boldsymbol{\zeta}}$ be characterized, respectively, by the stochastic matrices

$$\boldsymbol{P}_\gamma = \begin{pmatrix} 1-\gamma^H & \gamma^H \\ 1-\gamma^L & \gamma^L \end{pmatrix} \quad \text{and} \quad \boldsymbol{P}_\zeta = \begin{pmatrix} 1-\zeta^H & \zeta^H \\ 1-\zeta^L & \zeta^L \end{pmatrix}.$$

According to Blackwell's theorem, any decision maker prefers information system $\hat{\boldsymbol{\gamma}}$ to $\hat{\boldsymbol{\zeta}}$ if and only if there exists a non-negative stochastic matrix $\boldsymbol{M}$ with $\sum_j m_{ij} = 1$ such that $\boldsymbol{P}_\zeta = \boldsymbol{P}_\gamma \boldsymbol{M}$.[46] It is readily verified that this matrix $\boldsymbol{M}$ exists and takes the form

$$\boldsymbol{M} = \begin{pmatrix} 1-p & p \\ 0 & 1 \end{pmatrix}.$$

Thus, even though comparison of the two performance measures according to Blackwell's theorem implies that the principal should prefer $\hat{\boldsymbol{\gamma}}$ over $\hat{\boldsymbol{\zeta}}$, the principal actually prefers the "garbled" information system $\hat{\boldsymbol{\zeta}}$ over information system $\hat{\boldsymbol{\gamma}}$. While Kim (1995) has already shown that the necessary part of Blackwell's theorem does not hold in the agency model, the sufficiency part was proven to be applicable to the agency framework by Gjesdal (1982).[47] Our findings, however, show that this is not the case anymore when the agent is loss-averse.

[46] See Blackwell (1951, 1953).

[47] In order to avoid confusion: The necessary part of Blackwell's theorem states that the principal being better off implies that she uses a more informative performance measure. The sufficiency part conversely states that making use of more informative performance measure implies that the principal is better off.

V.5. ROBUSTNESS, EXTENSIONS, AND CONCLUSION

In this paper, we explore the implications of reference-dependent preferences on contract design in an otherwise standard model of principal-agency. We find that introducing a loss-averse agent leads to a reduction in the complexity of the optimal contractual arrangement. When loss aversion is the predominant feature of the agent's risk preferences, the optimal contract takes the form of a simple bonus contract even if the agent is also risk-averse: some realizations of the performance measure are rewarded with a bonus payment, while others are not. Thus, loss aversion provides a theoretical rationale for bonus contracts, the wide application of which is hard to reconcile with obvious drawbacks – as seasonality effects or insurance fraud – that come along with this particular contractual form.

In the remainder of this section, we consider the robustness of our results. After a brief and semi-formal analysis of an alternative equilibrium concept, we explore the consequences of non quadratic effort costs for implementation problems. Finally, we conclude by discussing diminishing sensitivity of the gain-loss function. Throughout the whole analysis, we adopted the concept of choice-acclimating personal equilibrium (CPE). Kőszegi and Rabin (2006, 2007) provide another concept, called unacclimating personal equilibrium (UPE). The major difference between UPE and CPE is the timing of expectation formation and actual decision making. Under UPE a decision maker first forms his expectations, which determine his reference point, and thereafter, given these expectations, chooses the optimal action. To rule out that people can systematically cheat themselves, for action $\hat{a}$ to be an UPE, it must be optimal for the decision maker to choose $\hat{a}$ given that he expected to do so. In the following, we will argue that applying UPE instead of CPE does not change our main findings. Intuitively, the optimality of simple contracts is rooted in the agent's dislike of being exposed to numerous outcomes, which is reflected in the functional form of his ex ante expected utility. With expectations being met on the equilibrium path under UPE, the expected utility takes the same form under both equilibrium concepts.

Thus, one should expect simple contracts to be optimal also under UPE. For the sake of a more formal argument, consider the case of a purely loss-averse agent, i.e., suppose intrinsic utility is linear. The agent's ex ante expected utility from choosing action a when expecting action $\hat{a}$ is

$$E[U(a|\hat{a})] = \sum_{s=1}^{S} \gamma_s(a) \left[u_s + \sum_{j=1}^{s-1} \gamma_j(\hat{a})(u_s - u_j) - \lambda \sum_{t=s+1}^{S} \gamma_t(\hat{a})(u_t - u_s) \right] - c(a) + \mu(c(\hat{a}) - c(a)).$$

The agent's ex ante expected utility, and in consequence the individual rationality constraint, takes the same form under both equilibrium concepts, CPE and UPE. The incentive compatibility constraint, on the other hand, depends on the applied equilibrium concept. Given the agent expected to choose $\hat{a}$, his marginal utility from choosing a is

$$E[U'(a|\hat{a})] = \sum_{s=1}^{S}(\gamma_s^H - \gamma_s^L)u_s + \sum_{s=1}^{S}\sum_{j=1}^{s-1} \gamma_j(\hat{a})(\gamma_s^H - \gamma_s^L)(u_s - u_j) - \lambda \sum_{s=1}^{S}\sum_{j=s+1}^{S} \gamma_j(\hat{a})(\gamma_s^H - \gamma_s^L)(u_j - u_s) - c'(a) + \mu'(c(\hat{a}) - c(a)).$$

Note that either $\mu'(\cdot) = 1$ or $\mu'(\cdot) = \lambda$, depending on whether $\hat{a}$ is greater or lower than a. Even though $E[U(a|\hat{a})]$ is a strictly concave function in the agent's actual action choice a for all values of $\lambda \geq 1$, under UPE there arises the problem of multiplicity of equilibria. More precisely, for a given incentive scheme $\boldsymbol{u}$, there exists a range of actions $a \in [\underline{a}(\boldsymbol{u}), \bar{a}(\boldsymbol{u})]$ all of which constitute a UPE. This problem can be circumvented by assuming that the agent chooses the highest action which constitutes a UPE. In this case there is no need to impose additional assumptions on the cost function or to assume that λ is sufficiently small.[48] By imposing

[48]For given expectations $\hat{a}$, let EU_g and EU_l denote the agent's expected utility given that $\mu(x) = x$ and $\mu(x) = \lambda x$, respectively. Both EU_g and EU_l are strictly concave functions, with EU_g achieving its maximum at a strictly higher action than EU_l. EU_g and EU_l intersect at $\hat{a}$. Action $\hat{a}$ is an UPE if it lies between the maximizing actions of EU_g and EU_l. Therefore, expecting to choose the action which maximizes EU_g not only constitutes an UPE, but also is the highest possible UPE.

this alternative assumption, the incentive compatibility constraint can be rewritten as

$$\sum_{s=2}^{S} b_s \left\{ \left(\sum_{t=s}^{S} (\gamma_t^H - \gamma_t^L) \right) \left(1 + \sum_{j=1}^{s-1} \gamma_j(\hat{a}) \right) - \lambda \left(\sum_{t=s}^{S} \gamma_t(\hat{a}) \right) \left(\sum_{j=1}^{s-1} (\gamma_j^H - \gamma_j^L) \right) \right\} = 2c'(\hat{a}).$$

Clearly, the incentive compatibility constraint is linear in the additional payments $\boldsymbol{b} = (b_2, \ldots, b_S)$. Thus, our bonus contract result is robust with respect to this change of assumptions.

There is another way to resolve the multiplicity problem under UPE. Kőszegi and Rabin (2006, 2007) define a preferred personal equilibrium (PPE) as a decision maker's ex ante favorite plan among those plans he actually will follow through. Put differently, given incentive scheme $\boldsymbol{u}$, the agent chooses the action $a^{PPE} \in [\underline{a}(\boldsymbol{u}), \bar{a}(\boldsymbol{u})]$ that maximizes expected utility among those actions that constitute a UPE. If for all incentive-compatible incentive schemes we have $a^{PPE} \in (\underline{a}(\boldsymbol{u}), \bar{a}(\boldsymbol{u}))$ then PPE and CPE coincide, i.e., a^{PPE} is determined by the first-order condition that characterizes the agent's action under CPE. Thus, by imposing the PPE-analogue of A4 and A5 we can derive results identical to those under CPE. If $a^{PPE} \in \{\underline{a}(\boldsymbol{u}), \bar{a}(\boldsymbol{u})\}$ for all incentive-compatible incentive schemes, the optimal contract also is a bonus contract since both boundary actions are determined by functions linear in $\boldsymbol{b} = (b_2, \ldots, b_S)$.[49] In the intermediate case, however, where $a^{PPE} \in (\underline{a}(\boldsymbol{u}), \bar{a}(\boldsymbol{u}))$ for some incentive-compatible incentive schemes but $a^{PPE} \in \{\underline{a}(\boldsymbol{u}), \bar{a}(\boldsymbol{u})\}$ for others, the optimal contract is not necessarily a bonus scheme.

If the agent's action is characterized by PPE, for all actions $\hat{a} \in (0, 1)$ to be implementable we still need the assumption that λ is not too high. Put differently, similar implementation problems as discussed in Section V.4 also arise under PPE. Compared to CPE, however, these

[49]The case of $a^{PPE} = \bar{a}(\boldsymbol{u})$ corresponds to the alternative assumption to A4 discussed above. If $a^{PPE} = \underline{a}(\boldsymbol{u})$, on the other hand, then a^{PPE} maximizes EU_l, as defined in the previous footnote.

implementation problems are less severe. For instance, actions close to zero are always implementable under PPE.

For the discussion of implementation problems in Section V.4, we restricted attention to quadratic effort costs. The finding that implementation problems are a potential issue, however, holds true for a wide variety of cost functions. Depending on the particular functional form of the corresponding marginal costs, these implementation problems may be more or less severe. For instance, the result that there are implementation problems if $a_0 > 0$ holds true for all strictly increasing and strictly convex cost functions with $c'(0) = 0$. As for strictly concave marginal costs with $c'(0) = 0$, no action $\hat{a} \in (0, 1)$ is implementable if $a_0 \geq 0$; and even for $a_0 < 0$ there may be actions, in particular actions close to 1, that are not implementable.

Moreover, we kept the whole analysis simple by ignoring diminishing sensitivity, that is, by considering a piece-wise linear gain-loss function. A more general gain-loss function makes the analysis far more complicated: Both the incentive compatibility constraint and the individual rationality constraint are no longer linear functions in the intrinsic utility levels, and thus the Kuhn-Tucker conditions are not necessarily sufficient. Nevertheless, we expect that a reduction in the complexity of the contract will benefit the principal in this case as well. Diminishing sensitivity of the agent's utility implies that the sum of two net losses of two monetary outcomes exceeds the net loss of the sum of these two monetary outcomes. Therefore, in addition to the effects discussed in the paper, under diminishing sensitivity there is another channel through which melting two bonus payments into one "big" bonus affects, and in tendency reduces, the agent's expected net loss. There is, however, an argument running counter to this intuition. As we have shown, loss aversion may help the principal to create incentives. Therefore, setting many different wage payments, and thereby – in a sense – creating many kinks, proximity to which the agent strongly dislikes under diminishing sensitivity, may have favorable incentive effects. Exploring the effects of diminishing sensitivity in a principal-agent relationship with moral hazard is therefore an open question for future research.

VI. TECHNOLOGY OF SKILL FORMATION AND HIDDEN INFORMATION

VI.1. INTRODUCTION

The formation of human capital is a central issue in economics. Cunha and Heckman (2007; henceforth CH) consider an economic model of child development, where the formation of human capital occurs in multiple stages via investments. They solve for the optimal intertemporal investment plan, which is important to derive policy implications.

We extend their framework by assuming that children are differentiated in the sense that a child's type determines what type of investment is most productive for him/her, and that this information is not available when a child is young. That is, there is hidden information for the parents when the child is young (in the first phase). However, we assume that when a child is older (in the second phase) its type is revealed. How does the optimal investment plan change as a result of hidden information? Put differently, how should parents react to uncertainty about how to treat their young child best?

There are two intuitive guesses: (i) it is optimal to invest less in the first and a more in the second phase, because in the second one can tailor the investments to a child's type and therefore yield a high return of investment; (ii) invest more in the first and less in the second phase to make sure that, despite low returns in the first phase, the effective investment in the first phase is not "too bad".

We show that the answer crucially depends on the substitutability of investment between phases: when investments are easily substitutable (easier than Cobb-Douglas), intuition (i) is right; when substitution is difficult (more difficult than Cobb-Douglas), (ii) is right. More specifically, hidden information weakens the importance of early investments in children when inter-phase investments are easily substitutable, but strengthen them when substitution is difficult.

In the next section we first present and analyze CH's model. Then we extend it by introducing differentiated investments and hidden information. In Section VI.3 we conclude and discuss.

VI.2. MODELS

VI.2.1. CUNHA AND HECKMAN'S MODEL

There is a child with two periods of childhood, $t = 1, 2$. Child's adult stock of skill h, also called human capital, is given by

$$h = m\left(h^p, \theta_1, I_1, I_2\right),$$

where h^p is the skill of the parents, θ_1 the child's initial ability, and I_1 and I_2 are investments in the first and second period respectively.[1] For concreteness CH consider the following form, where I is given by a CES function:

$$h = m\left(h^p, \theta_1, I\right), \; I = \left[\gamma I_1^{\phi} + (1-\gamma) I_2^{\phi}\right]^{1/\phi}. \tag{VI.1}$$

The parameter γ, $0 < \gamma < 1$, is interpreted as a skill multiplier. It reveals the productivity of early investment not only in directly boosting h, but also in raising the productivity of I_2 by increasing the second period ability through high first-period investments; see CH (p. 38). The parameter ϕ, $\phi \leq 1$, describes how easy investments in different periods can be substituted for each other. For $\phi = 1$, we have a linear relationship: $I = [\gamma I_1 + (1-\gamma) I_2]$. That is, investments are perfect substitutes. For $\phi \to -\infty$ investments are not substitutable; the function is of the Leontief type. For $\phi = 0$ one gets the Cobb-Douglas function. The elasticity of substitution is $1/(1-\phi)$.

We assume that parents at the beginning of $t = 1$ maximize the

[1] We use a slightly different notation than CH.

present value of net wealth of their children[2]

$$\mathbb{E}[\pi] = w\mathbb{E}[h] - I_1 - \frac{1}{1+r}I_2$$

over $\{I_1, I_2\}$. The costs of second period investments are discounted by the factor $1/(1+r)$, where r is the interest rate. The life-time discounted wage per unit of skill is denoted by w. So that an optimum exists we assume that $d^2m(\cdot)/dI^2 < 0$. To guarantee that it is optimal to invest some positive amount we assume that $\lim_{I\to 0} dm(\cdot)/dI$ is "sufficiently large". For $\phi < 1$, optimization yields that the ratio of the monetary investments in period 1 relative to that in period 2 is[3]

$$\frac{I_1}{I_2} = \left(\frac{\gamma}{(1-\gamma)(1+r)}\right)^{\frac{1}{1-\phi}}. \qquad \text{(VI.2)}$$

CH interpret their formula as follows: "High productivity of initial investment (the skill multiplier γ) drives the parent toward making early investments. The interest rate drives the parent to invest late" (p. 39).

VI.2.2. THE MODEL WITH HIDDEN INFORMATION

We now extend the model by considering hidden information. We assume that there are two types of investment in every period: $\hat{I}_t$ and $\check{I}_t$.[4] For example, $\hat{I}_t$ may be the investment in child's sporting abilities whereas $\check{I}_t$ may denote investments in creative abilities. How $\hat{I}_t$ and $\check{I}_t$ combine to determine the effective investment depends on the child's type α:

$$I_t^{\text{effective}} = (1+\alpha)\hat{I}_t + (1-\alpha)\check{I}_t. \qquad \text{(VI.3)}$$

[2]This is the maximization problem considered by Cunha, Flavio, James Heckman, Lance Lochner, and Dimitriy Masterov (2005). An alternative approach is to consider a dynamic overlapping generations model (see the same paper or CH). But, as the authors note, "the main conclusions of the simple, static model ... are valid in a more fully specified economic environment." To focus on the main points we use the simple static model.

[3]For $\phi = 1$ one gets corner solutions and the ratio need not be defined.

[4]Also CH consider differentiated investments. But they do not consider hidden information.

With equal probability the type is β or $-\beta$, with $0 < \beta < 1$. Hence, when the child is of type β it is most productive to invest in $\hat{I}_t$ and not in $\check{I}_t$; it is the other way round when the type is $-\beta$. The height of the parameter β captures how strong the productivity of the different intra-period investments differs.

In $t = 1$, the child's type α is not known to the parents. There is hidden information. Hence they cannot be sure how to best tailor the investment to the child.[5] That is, the parents do not know whether they should invest in the child's sporting or creative abilities. In $t = 2$, when the child is older, the parents learn the child's type. Obviously, then it is optimal not to invest in the less productive type of investment: $\hat{I}_2 = 0$ when $\alpha = -\beta$ and $\check{I}_2 = 0$ when $\alpha = \beta$. We denote the highly productive investments by I_2^H. That is, $I_2^H = \hat{I}_2$ when $\alpha = \beta$ and $I_2^H = \check{I}_2$ when $\alpha = -\beta$.

As in CH's model we assume that the relationship between effective per period investments and effective total investment I is given by a CES function:

$$I = \left[\gamma \left(I_1^{\text{effective}}\right)^{\phi} + (1-\gamma)\left(I_2^{\text{effective}}\right)^{\phi}\right]^{1/\phi}. \tag{VI.4}$$

The parents' investment policy $\left\{\hat{I}_1, \check{I}_1, I_2^H\big|_{\alpha=\beta}, I_2^H\big|_{\alpha=-\beta}\right\}$ maximizes

$$\mathbb{E}_\alpha[\pi] = w\mathbb{E}_\alpha\left[h\right] - \hat{I}_1 - \check{I}_1 - \frac{1}{1+r} I_2^H. \tag{VI.5}$$

Note, since at $t = 2$ the child's type α is known the investment in $t = 2$ can be made contingent on α.

LEMMA 9: It is optimal to choose $\hat{I}_1 = \check{I}_1$ and $I_2^H\big|_{\alpha=-\beta} = I_2^H\big|_{\alpha=\beta}$. For $\phi < 1$ it is optimal to invest in both periods.

PROOF: See Appendix.

The first part of Lemma 9 says that it is optimal to diversify investments completely by choosing $\hat{I}_1 = \check{I}_1$. The second part states that

[5] We do not consider mechanism that reveal the child's type. This is justified because young children are simply unable to reveal their types (or maybe they cannot be convinced to take part in any kind of mechanism).

although the *type* of the second period investment depends on α, the *height* of the second period investments, i.e. I_2^H, is independent of α.

As in CH's model we would like to determine the ratio of the monetary investments in period 1 ($I_1^T := \hat{I}_1 + \check{I}_1$) relative to that in period 2 ($I_2^T := \hat{I}_2 + \check{I}_2$). From the first-order conditions of the problem (see the Appendix) and Lemma 9 one directly yields that for $\phi < 1$ the ratio is

$$\frac{I_1^T}{I_2^T} = \left(\frac{\gamma}{(1-\gamma)(1+r)(1+\beta)^{\phi}} \right)^{\frac{1}{1-\phi}}. \tag{VI.6}$$

P R O P O S I T I O N 22: Suppose that $\phi < 1$. When $\phi > (<) 0$ the ratio of the monetary investments in period 1 relative to that in period 2 with hidden information is smaller (larger) than without hidden information.

The intuition is as follows. The return of second period investments is high because they can be tailored to the child's type. When investments can be substituted easily (easier than Cobb-Douglas), a low first period investment can easily be compensated by a high second period investment. Hence, it is optimal to invest little in the first and a lot in the second period. However, when investments are difficult to substitute (more difficult than Cobb-Douglas), this is not the case. A low first period investment can only be compensated by a very high second period investment. This would be very costly. Hence, it is optimal to invest a lot in the first period to make sure that the effective first period investment is substantial.[6]

Finally, looking at (VI.2) and (VI.6) yields the following insights.[7]

[6] The results we found are mathematically closely related to Acemoglu (2002). He considers how the augmentation of one factor changes the relative marginal products of both factors of production. He shows that when the elasticity of substitution is above 1, then the relative marginal product of the factor which is augmented improves. When the elasticity of substitution is below 1, then it is the other way round. In our model, hidden information augments investments in period 2 relative to investments in period 1.

[7] As in CH's model the investment ratio need not be defined for $\phi = 1$. Some simple calculations yield that the results stated in Proposition 23 is also valid for $\phi = 1$.

PROPOSITION 23: The first period investment exceeds the second period investment in the model without hidden information if $\gamma > (1-\gamma)(1+r)$. With hidden information this is true for $\gamma > (1-\gamma)(1+r)(1+\beta)^{\phi}$. It is the other way round when the formulas hold with $<$.

So when $\phi > 0$, i.e., when substitution is easier than with a Cobb-Douglas function, the skill multiplier γ must be larger in the model with hidden information than in the model without so that the first period investment exceeds the second period investment. For $\phi < 0$, i.e., when substitution is more difficult than with a Cobb-Douglas function, the multiplier γ can be lower.

To sum up both propositions, hidden information weakens the importance of early investments in children when inter-period investments are easily substitutable. When substitution is difficult, early investments get more important.

For completeness, consider the case when there is no hidden information. Then it is optimal to invest only in the productive type of investment in both periods. The parameter β appears in both first-order conditions in the same way. Hence, the term β cancels out in the investment ratio. So (VI.6) applies when one sets $\beta = 0$. That is, the investment ratio is as in CH's model, see (VI.2).

VI.3. CONCLUSIONS & DICUSSION

We have extended the model of CH by introducing hidden information about a child's type when it is young. We have shown that hidden information weakens the importance of early investment in children when inter-period investments are easily substitutable, but strengthens their importance when substitution is difficult.

We have assumed that the differentiated investments of a period combine linear to the effective investment of a period, see (VI.3). This simplification can be defended as follows: With a more complicated functional form it stays in general true that knowing a child's type allows tailored investments which yield a higher return than not tailored ones. That is, investing later yields a return on effective second period investment which is greater, say by a factor $(1+\beta)$, than the expected return early investments have on the effective first period investment. Since the factor $(1+\beta)$ comes back we can use the linear specification as a reduced form of the more general specification.[8]

Extending the model to three periods, in which there is uncertainty only in the first period, can lead to richer patterns. Roughly speaking, the optimal relative investment of the first and the second periods is governed by the extended model with hidden information. That between the second and third periods by the model without hidden information. So the optimal investments may be non-monotone in time. Inverted-U or U patterns are possible.

Eliciting the hidden information through scientific tests allows a tailored investment policy also for young children. Then the same adult skill levels are attainable with lower investments. Alternatively, with the same investments higher skill levels can be achieved. Hence, those tests are important to improve adult skill and the effectiveness of child investments.

The model we consider can be interpreted more broadly: There is a multi-period investment problem with only initial uncertainty about the most productive way to invest.

[8]This arguments show that the *inter*-period investment problem does not change due to a more general specification. That is, the investment ratio stays the same. However, the *intra*-period investment problem changes. With a more general specification it may be optimal to always invest in both types of investments. Then it may be optimal to invest more in an old child's ability in which he/she is less talented.

VII. APPENDICES

VII.1. APPENDIX TO CHAPTER II

It is useful to define $\mu^J(p)$ as the probability that firm $J \in \{I, E\}$ will invest in research. Also define $\tau(p)$ $\big(= p\mu^E(p)\big)$ as the probability that entry will occur. Both variables are measured after p and before k is drawn.

PROOF OF PROPOSITION 1

From Lemma 2 we get that

$$\mu^I(p) = \int_0^{(1-\pi^I(\bar{c},-))(1-p)p} h(k|p)dk + \int_{\pi^E(\bar{c},\underline{c})(1-p)p}^{\pi^E(\bar{c},\underline{c})p} h(k|p)dk$$

and

$$\mu^E(p) = \int_0^{(1-\pi^I(\bar{c},-))(1-p)p} h(k|p)dk + \int_{(1-\pi^I(\bar{c},-))(1-p)p}^{\pi^E(\bar{c},\underline{c})(1-p)p} h(k|p)dk.$$

Hence, $\mu^I(p) > \mu^E(p)$ if and only if

$$\int_{\pi^E(\bar{c},\underline{c})(1-p)p}^{\pi^E(\bar{c},\underline{c})p} h(k|p)dk > \int_{(1-\pi^I(\bar{c},-))(1-p)p}^{\pi^E(\bar{c},\underline{c})(1-p)p} h(k|p)dk. \tag{VII.1}$$

So $\mu^I(p) > \mu^E(p)$ if and only if it is more likely that the replacement effect will determine the outcome than that the efficiency effect will.

Using A1, (VII.1) is

$$p^2\pi^E(\bar{c},\underline{c}) > (1-p)p\left[\pi^E(\bar{c},\underline{c}) - (1-\pi^I(\bar{c},-))\right].$$

So $\mu^I(p) > \mu^E(p)$ if (and only if)

$$p > \dot{p} := \frac{\pi^E(\bar{c},\underline{c}) - (1-\pi^I(\bar{c},-))}{2\pi^E(\bar{c},\underline{c}) - (1-\pi^I(\bar{c},-))}.$$

From Lemma 1 follows that $\dot{p} \in (0,1)$. Similarly, $\mu^I(p) < \mu^E(p)$ if (and only if) $p < \dot{p}$. Note, the "only if" part is not included in Proposition 1 because it is an artefact of the uniform assumption upon h. $\square$

PROOF OF PROPOSITION 2

Start with some notation. Let the expected p, conditional that firm J invests, be

$$P^J := \int_0^1 p g^J(p) dp,$$

where

$$g^J(p) := \frac{g(p)\mu^J(p)}{\int_0^1 g(p)\mu^J(p)dp}$$

is the density of p, conditional that $J \in \{I, E\}$ invests. The associated distribution function is denoted by $G^J(p)$. We seek to show that $P^I > P^E$.

Step 1. Using Lemma 2 and A1,

$$\frac{d\left(\frac{\mu^I(p)}{\mu^E(p)}\right)}{dp} = \frac{1}{(1-p)^2}.$$

Since $\frac{g^I(p)}{g^E(p)} = \frac{\mu^I(p)}{\mu^E(p)} \frac{\int_0^1 g(r)\mu^E(r)dr}{\int_0^1 g(r)\mu^I(r)dr}$, it follows that $d\left(\frac{g^I(p)}{g^E(p)}\right)/dp$ is positive and finite $\forall p \in (0,1)$.

Step 2. CLAIM: $g^I(p) < g^E(p)$ for $p \to 0$ and $g^I(p) > g^E(p)$ for $p \to 1$.

PROOF: Since $g(p)$ and $\mu^J(p)$ are positive and finite $\forall p \in (0,1]$, it follows that $g^J(p)$ is positive and finite $\forall p \in (0,1]$ as well. From step 1, $d\left(\frac{g^I(p)}{g^E(p)}\right)/dp > 0$ $\forall p \in (0,1)$, and by definition $\int_0^1 g^I(p)dp = 1$ and $\int_0^1 g^E(p)dp = 1$. Hence it must hold that $g^I(p) < g^E(p)$ for $p \to 0$, and $g^I(p) > g^E(p)$ for $p \to 1$.

Step 3. CLAIM: there exists a $\tilde{p} \in (0,1)$ such that $g^I(p) = (<,>) g^E(p)$ for $p = (<,>) \tilde{p}$.

PROOF: Step 1 says that $\frac{d\frac{g^I(p)}{g^E(p)}}{dp}$ is positive and finite $\forall p \in (0,1)$. So $\frac{g^I(p)}{g^E(p)}$ is continuous and increasing in p, $\forall p \in (0,1)$. This implies, together with the intermediate value theorem and step 2, that there exists a $\tilde{p} \in (0,1)$ such that $g^I(\tilde{p}) = g^E(\tilde{p})$. Since $\frac{g^I(p)}{g^E(p)}$ is increasing in p $\forall p \in (0,1)$, and $g^I(p) < (>) g^E(p)$ for $p \to 0(1)$ by step 2, it holds that $g^I(p) < (>) g^E(p)$ for $p < (>) \tilde{p}$.

Step 4. By Proposition 6.D.1 in Mas-Colell, Whinston, and Greene (1995, p. 195) $g^I(p)$ first-order stochastically dominates $g^E(p)$ if and only if

$$\int_0^x g^I(p)dp \leq \int_0^x g^E(p)dp \quad \forall x \in (0,1]. \tag{VII.2}$$

We seek to show a slightly different property.

Claim: The inequality in (VII.2) is strict $\forall x \in (0,1)$.

Proof: Suppose there exists a $\hat{x} \in (0,1)$ such that $\int_0^{\hat{x}} g^I(p)dp \geq \int_0^{\hat{x}} g^E(p)dp$.

Case 1: $\hat{x} \leq \tilde{p}$. From step 3, $g^I(p) < g^E(p)$ $\forall p \in (0,\tilde{p})$, and $g^I(\tilde{p}) = g^E(\tilde{p})$. Hence, $\int_0^{\hat{x}} g^I(p)dp \geq \int_0^{\hat{x}} g^E(p)dp$ is false.

Case 2: $\hat{x} > \tilde{p}$. If $\int_0^{\hat{x}} g^I(p)dp \geq \int_0^{\hat{x}} g^E(p)dp$, then

$$\int_{\hat{x}}^1 g^I(p)dp \leq \int_{\hat{x}}^1 g^E(p)dp, \tag{VII.3}$$

since $\int_0^1 g^I(p)dp = 1$ and $\int_0^1 g^E(p)dp = 1$. From step 3, $\forall p \in (\tilde{p},1)$ it is true that $g^I(p) > g^E(p)$. From Lemma 2 follows that $g^E(1) = 0$. By definition $g^I(1) \geq 0$. Hence, (VII.3) is false.

Step 5. Using the definition of P^E and P^I, we get by integrating by parts that

$$P^I = 1 - \int_0^1 G^I(p)dp, \; P^E = 1 - \int_0^1 G^E(p)dp.$$

From step 4 we know that $G^I(p) < G^E(p)$ $\forall p \in (0,1)$. By definition $G^I(1) = G^E(1)$. Hence, $\int_0^1 G^I(p)dp < \int_0^1 G^E(p)dp$ and so $P^I > P^E$. □

PROOF OF PROPOSITION 3

A1 implies that $h(k|p) = h(p)$ $\forall k \in \mathcal{S}$. By the definition of a density,

$$\int_0^{p\pi^E(\bar{c},\underline{c})} h(p)dk \leq 1$$

or rewritten

$$h(p) \leq \frac{1}{p\pi^E(\bar{c}, \underline{c})}.$$

Together with Lemma 2 this implies that

$$\mu^E(p) = h(p)\pi^E(\bar{c}, \underline{c})(1-p)p \leq \frac{1}{p\pi^E(\bar{c}, \underline{c})}\pi^E(\bar{c}, \underline{c})(1-p)p = 1-p.$$

Since, by definition $\tau(p) = p\mu^E(p)$, it follows that $\tau(p) \leq p(1-p)$. □

PROOF OF PROPOSITION 4

The total welfare, excluding potential investments in research, when the incumbent has production costs c^I and the potential entrant costs c^E is denoted by $t(c^I, c^E)$. It consists of the firms' Bertrand profits and consumer welfare (for sake of clearness we do not use the normalization $\pi^I(\underline{c}, -) = 1$):

$$\begin{aligned}
t(\bar{c}, -) &= \int_{\phi^I(\bar{c},-)}^{\infty} D(\phi)d\phi + \pi^I(\bar{c}, -), \\
t(\underline{c}, -) &= \int_{\phi^I(\underline{c},-)}^{\infty} D(\phi)d\phi + \pi^I(\underline{c}, -), \\
t(\underline{c}, \underline{c}) &= \int_{\underline{c}}^{\infty} D(\phi)d\phi + 0, \\
t(\bar{c}, \underline{c}) &= \int_{\phi^E(\bar{c},\underline{c})}^{\infty} D(\phi)d\phi + \pi^E(\bar{c}, \underline{c}),
\end{aligned}$$

where $\phi^J(c^I, c^E)$ is the price set by firm $J \in \{I, E\}$. Let $T(a^I, a^E)$ be the expected total welfare taking into account firms' investments. Straightforward calculations yield:

$$\begin{aligned}
T(0,0) &= t(\bar{c}, -), \\
T(0,1) &= pt(\bar{c}, \underline{c}) + (1-p)t(\bar{c}, -) - k, \\
T(1,0) &= pt(\underline{c}, -) + (1-p)t(\bar{c}, -) - k, \\
T(1,1) &= p^2 t(\underline{c}, \underline{c}) + p(1-p)t(\underline{c}, -) \\
&\quad + p(1-p)t(\bar{c}, \underline{c}) + (1-p)^2 t(\bar{c}, -) - 2k.
\end{aligned}$$

Observe that for a drastic innovation $\phi^E(\bar{c}, \underline{c}) = \phi^I(\underline{c}, -)$ and so $T(0,1) = T(1,0)$. However, when the innovation is non-drastic $\phi^E(\bar{c}, \underline{c}) < \phi^I(\underline{c}, -)$

and so $T(0,1) > T(1,0)$. Next we explore the question whether there is socially too much or too little investment. We look at the second best world in which pricing cannot be regulated.

Overinvestment when $a^* = (1,0)$*?—* Consider first the case that in equilibrium only the incumbent invests. Is it socially desirable that no firm invests? Using the formulas derived before yields

$$T(1,0) - T(0,0) = p\left[\int_{\phi^I(\underline{c},-)}^{\phi^I(\bar{c},-)} D(\phi)d\phi + \pi^I(\underline{c},-) - \pi^I(\bar{c},-)\right] - k.$$

From Lemma 2 we know that for $a^* = (1,0)$,

$$k = p\pi^E(\bar{c},\underline{c}) - (+\text{term}),$$

where the positive term can be arbitrarily small. Then

$$T(1,0) - T(0,0) = \\ p\left[\int_{\phi^I(\underline{c},-)}^{\phi^I(\bar{c},-)} D(\phi)d\phi + \pi^I(\underline{c},-) - \pi^I(\bar{c},-) - \pi^E(\bar{c},\underline{c})\right] + (+\text{term}). \tag{VII.4}$$

For a drastic innovation, $\pi^I(\underline{c},-) = \pi^E(\bar{c},\underline{c})$. The profit term $\pi^I(\bar{c},-)$ can be rewritten as

$$\int_{\bar{c}}^{\phi^I(\bar{c},-)} D(\phi^I(\bar{c},-))d\phi.$$

Since we assumed that the monopolist' optimal price is unique we have that for a drastic innovation $\phi^I(\underline{c},-) \leq \bar{c}$. Hence, $T(1,0) - T(0,0) > 0$.

However, when the innovation is non-drastic, then it can be the case that $T(1,0)-T(0,0) < 0$. This is true in the following example: (+term) is small; $\phi^I(\bar{c},-) \approx \phi^I(\underline{c},-)$ so that the integral term in (VII.4) is small; let $\pi^I(\underline{c},-) - \pi^I(\bar{c},-)$ be well below $\pi^E(\bar{c},\underline{c})$ (this is possible due to the replacement effect).

Overinvestment when $a^* = (0,1)$*?—* When in equilibrium only E invests, is it socially desirable that no firm invests?

$$T(0,1) - T(0,0) = p\left[\int_{\phi^E(\bar{c},\underline{c})}^{\phi^I(\bar{c},-)} D(\phi)d\phi + \pi^E(\bar{c},\underline{c}) - \pi^I(\bar{c},-)\right] - k.$$

From Lemma 2, when $a^* = (0,1)$ then

$$k = \pi^E(\bar{c}, \underline{c})(1-p)p - (+\text{term}),$$

and so

$$T(0,1) - T(0,0) = p\left[\int_{\phi^E(\bar{c},\underline{c})}^{\phi^I(\bar{c},-)} D(\phi)d\phi + p\pi^E(\bar{c}, \underline{c}) - \pi^I(\bar{c}, -)\right] + (+\text{term}).$$

Are the terms in the square brackets also positive if $p \to 0$? Yes, since $\phi^I(\bar{c}, -) > \bar{c} \geq \phi^E(\bar{c}, \underline{c})$ and the demand is decreasing in the price:

$$\int_{\phi^E(\bar{c},\underline{c})}^{\phi^I(\bar{c},-)} D(\phi)d\phi \geq \int_{\bar{c}}^{\phi^I(\bar{c},-)} D(\phi)d\phi > \int_{\bar{c}}^{\phi^I(\bar{c},-)} D(\phi^I(\bar{c}, -))d\phi = \pi^I(\bar{c}, -).$$

Obviously, for all p the terms in the square brackets are positive, and so it holds that $T(0,1) - T(0,0) > 0$.

Overinvestment when $a^* = (1,1)$?— Next consider the case when both firms invest in research. Is it socially desirable that instead only the potential entrant invests?

$$\begin{aligned} T(1,1) - T(0,1) & \\ &= p\left[p\left(\int_{\underline{c}}^{\phi^E(\bar{c},\underline{c})} D(\phi)d\phi - \pi^E(\bar{c}, \underline{c})\right) + \right. \\ &\left. (1-p)\left(\int_{\phi^I(\underline{c},-)}^{\phi^I(\bar{c},-)} D(\phi)d\phi + \pi^I(\underline{c}, -) - \pi^I(\bar{c}, -)\right)\right] - k. \end{aligned}$$

From Lemma 2 we know that when $a^* = (1,1)$, then

$$k = p(1-p)(\pi^I(\underline{c}, -) - \pi^I(\bar{c}, -)) - (+\text{term}).$$

Using this we get that

$$\begin{aligned} T(1,1) - T(0,1) = p\left[p\left(\int_{\underline{c}}^{\phi^E(\bar{c},\underline{c})} D(\phi)d\phi - \pi^E(\bar{c}, \underline{c})\right)\right. \\ \left. + (1-p)\left(\int_{\phi^I(\underline{c},-)}^{\phi^I(\bar{c},-)} D(\phi)d\phi\right)\right] + (+\text{term}). \end{aligned}$$

The second integral is nonnegative because $\phi^I(\underline{c},-) \leq \phi^I(\bar{c},-)$, see Tirole (1988, p. 66). The term $\int_{\underline{c}}^{\phi^E(\bar{c},\underline{c})} D(\phi)d\phi - \pi^E(\bar{c},\underline{c})$ can be rewritten as

$$\int_{\underline{c}}^{\phi^E(\bar{c},\underline{c})} D(\phi)d\phi - \int_{\underline{c}}^{\phi^E(\bar{c},\underline{c})} D(\phi^E(\bar{c},\underline{c}))d\phi,$$

which is positive because demand is decreasing in the price and $\phi^E(\bar{c},\underline{c}) > \underline{c}$. Hence, $T(1,1) - T(0,1) > 0$.

We know that $T(0,1) \geq T(1,0)$. Hence also $T(1,1) - T(1,0) > 0$.

Finally, is $T(1,1)$ socially preferred to $T(0,0)$?

$$\begin{aligned} T(1,1) - T(0,0) = p^2 \left[t(\underline{c},\underline{c}) - t(\underline{c},-) - t(\bar{c},\underline{c}) + t(\bar{c},-)\right] \\ + p\left[t(\underline{c},-) + t(\bar{c},\underline{c}) - 2t(\bar{c},-)\right] - 2k. \end{aligned}$$

Again, for $a^* = (1,1)$ it holds that

$$k = p(1-p)(\pi^I(\underline{c},-) - \pi^I(\bar{c},-)) + (+\text{term}).$$

Hence, after some calculations,

$$\begin{aligned} &T(1,1) - T(0,0) \\ &= p^2\left[\int_{\underline{c}}^{\phi^E(\bar{c},\underline{c})} D(\phi)d\phi - \int_{\phi^I(\underline{c},-)}^{\phi^I(\bar{c},-)} D(\phi)d\phi - \pi^E(\bar{c},\underline{c}) + \pi^I(\underline{c},-) - \pi^I(\bar{c},-)\right] \\ &+p\left[\int_{\phi^I(\underline{c},-)}^{\phi^I(\bar{c},-)} D(\phi)d\phi + \int_{\phi^E(\bar{c},\underline{c})}^{\phi^I(\bar{c},-)} D(\phi)d\phi + \pi^E(\bar{c},\underline{c}) - \pi^I(\underline{c},-)\right] + (+\text{term}). \end{aligned} \tag{VII.5}$$

If $p = 1$, one gets that

$$T(1,1) - T(0,0) = \int_{\underline{c}}^{\phi^I(\bar{c},-)} D(\phi)d\phi - \pi^I(\bar{c},-) + (+\text{term}),$$

which is obviously positive.

If $p \to 0$, only the last line of (VII.5) is important. The replacement effect holds and so $\pi^E(\bar{c},\underline{c}) - \pi^I(\underline{c},-) = -\pi^I(\bar{c},-) + (+\text{term})$, see Lemma 1. So the last line of (VII.5) can be rewritten as

$$p\left[\int_{\phi^I(\underline{c},-)}^{\phi^I(\bar{c},-)} D(\phi)d\phi + \int_{\phi^E(\bar{c},\underline{c})}^{\phi^I(\bar{c},-)} D(\phi)d\phi - \pi^I(\bar{c},-)\right] + (+\text{term}).$$

From before (see the case $a^* = (0,1)$) we know that $\int_{\phi^E(\bar{c},\underline{c})}^{\phi^I(\bar{c},-)} D(\phi)d\phi > \pi^I(\bar{c},-)$, and so the last line is positive. Hence, also for $p \to 0$, $T(1,1) - T(0,0) > 0$. Since for $p = 1$ and for $p \to 0$ we have $T(1,1) - T(0,0) > 0$, and through the functional form of (VII.5), we get that $T(1,1) - T(0,0) > 0 \; \forall p \in (0,1]$.

Underinvestment.— Note, when $a^* = (0,0)$ the previous analysis of the case $a^* = (1,0)$ applies expect that there is no positive term, but instead a negative term. When k is close to the boundary where $a^* = (1,0)$, the negative term is close to zero. Hence, our analysis shows that it may be socially desirable that one firm invests when in fact no firm invests. The same arguments hold when in fact one firm invests and one asks the question whether investments of both firms is socially preferable. Hence, firm may invest too little. □

PROOF OF PROPOSITION 5

Denote firm J's valuation as v^J, $J \in \{I, E\}$. I's valuation v^I is given through the k which equates (II.5), and so $v^I = p$. Similarly, from (II.1) we get $v^E = \pi^E(\bar{c},\underline{c})p$. If the innovation is non-drastic, $\pi^E(\bar{c},\underline{c}) < 1$, and hence $v^I > v^E \; \forall p \in (0,1]$. Since we consider a first-price auction with complete and perfect information, I will always win. □

VII.2. APPENDIX TO CHAPTER III

PROOF OF PROPOSITION 6

The derivative of (III.7) is

$$\frac{d\mathbb{E}\left[\pi_0(k,\lambda)\right]}{dk} = h'(k)e^{-\lambda(C-h(k)-t)}+\left(1-C+h(k)\right)e^{-\lambda(C-h(k)-t)}\lambda h'(k)-1. \tag{VII.6}$$

First consider the case of no entry threat. This is captured by $\lambda = 0$. Then $d\mathbb{E}\left[\pi_0(k,0)\right]/dk = h'(k) - 1$. The optimal investment $k^*(0)$ solves $h'(k^*(0)) = 1$. So $k^*(0) = h'^{-1}(1)$, which is positive through Assumption A2(v).

Next, consider the other extreme of $\lambda \to \infty$. Then Assumption A2(iv) implies that entry occurs for sure for all investment levels k. So $\mathbb{E}\left[\pi_0(k,\lambda)\right] = -k$ for all k. Therefore the incumbent chooses not to invest. That is, $k^*(\infty) = 0$.

Finally, we prove the last part of Proposition 6. By Assumption 2 (iv) and $t > 0$ we have $c_0(k) > 0$ $\forall k$. Moreover, the expected revenue is at most 1. Hence, $k \geq 1$ leads to a loss for the incumbent. This is dominated by $k = 0$, which yields a nonnegative profit. Therefore, an investment of $k \geq 1$ can never be optimal.

From (III.7) we get that $\forall k \in [\hat{k}, 1]$ we have

$$\begin{aligned}
\mathbb{E}\left[\pi_0(k,\lambda)\right] &= \left(1-c_0(k)\right)e^{-\lambda(c_0(k)-t)} - k \\
&\leq \left(1-c_0(1)\right)e^{-\lambda(c_0(1)-t)} - k \\
&\leq \left(1-c_0(1)\right)e^{-\lambda(c_0(1)-t)} - \hat{k}.
\end{aligned}$$

When λ is sufficiently high

$$\left(1-c_0(1)\right)e^{-\lambda(c_0(1)-t)} - \hat{k} < 0$$

and therefore also

$$\mathbb{E}\left[\pi_0(k,\lambda)\right] < 0 \quad \forall k \in [\hat{k}, 1].$$

This is dominated by $k = 0$. Hence, an investment of $k \in [\hat{k}, \infty)$ cannot be optimal. □

PROOF OF PROPOSITION 7

We look at the effect of λ on the marginal return of investment:

$$\frac{d^2\mathbb{E}\left[\pi_0(k,\lambda)\right]}{dkd\lambda} = \begin{array}{r} -h'(k)e^{-\lambda(C-h(k)-t)}(C-h(k)-t) \\ +\left(1-C+h(k)\right)e^{-\lambda(C-h(k)-t)}h'(k) \\ -\left(1-C+h(k)\right)e^{-\lambda(C-h(k)-t)}\lambda h'(k)(C-h(k)-t). \end{array}$$

The first term on the RHS is negative: a higher λ increases the probability of entry. That is, the return to investment decreases due to the Schumpeterian effect. The remaining terms capture the effect on the entry deterrence effect. The sign of the sum of the remaining terms is ambiguous. That is, entry deterrence may or may not become more attractive when λ increases. Given some k, when λ is small (large), the remaining terms are positive (negative). Since we seek to explore whether it is possible that a higher entry threat increases the optimal investment, we consider the case $\lambda \to 0$:

$$\lim_{\lambda\to 0}\frac{d^2\mathbb{E}\left[\pi_0(k,\lambda)\right]}{dkd\lambda} = -h'(k)(C-h(k)-t)+\left(1-C+h(k)\right)h'(k),$$

which has the same sign as

$$1-2C+2h(k)+t.$$

So when

$$C < h(k)+\frac{1+t}{2},$$

we have that $\lim_{\lambda\to 0}\frac{d^2\mathbb{E}[\pi_0(k,\lambda)]}{dkd\lambda} > 0$. When

$$C < \frac{1+t}{2},$$

this is true for all k.

$$\mathbb{E}\left[\pi_0(k^*(0),0)\right] \geq \mathbb{E}\left[\pi_0(\dot{k},0)\right]$$

$\forall\dot{k} \in \mathbb{R}^{++}$ by the optimality of $k^*(0)$. The inequality also holds for the subset $\dot{k} \in [0, k^*(0))$, which can equivalently be written as

$$\int_{\dot{k}}^{k^*(0)} \frac{d\mathbb{E}\left[\pi_0(k,0)\right]}{dk}ds \geq 0 \quad \forall\dot{k} \in [0,k^*(0)).$$

Let λ be small. Since $C < \frac{1+t}{2}$ we know from before that $\frac{d^2\mathbb{E}[\pi_0(k,\lambda)]}{dkd\lambda} > 0$. Hence,

$$\int_{\dot{k}}^{k^*(0)} \frac{d\mathbb{E}\left[\pi_0(k,0)\right]}{dk} ds > 0 \quad \forall \dot{k} \in [0, k^*(0)).$$

This implies that

$$\mathbb{E}\left[\pi_0(k^*(0),\lambda)\right] > \mathbb{E}\left[\pi_0(\dot{k},\lambda)\right] \quad \forall \dot{k} \in [0, k^*(0)).$$

This proofs that $k^*(\lambda) \geq k^*(0)$.

Moreover, since $\frac{d^2\mathbb{E}[\pi_0(k,\lambda)]}{dkd\lambda} > 0$ we have that

$$d\,\mathbb{E}\left[\pi_0(k,\lambda)\right]|_{k^*(0)}/dk > 0.$$

Hence, $k^*(\lambda) \neq k^*(0)$. Together with the previous finding and the fact that through the assumptions made a $k^*(\lambda)$ exists, this implies that $k^*(\lambda) > k^*(0)$. □

PROOF OF PROPOSITION 9

With help of Lemma 3 we yield that the incumbent's expected investment is

$$\begin{aligned}\mathbb{E}[k|\lambda] = \quad & \int_0^{c_0(\bar{k})-t} 0 dF(\underline{c}) + \int_{c_0(\bar{k})-t}^{c_0(k^{**}(\infty))-t} k^{**}(\cdot)dF(\underline{c}) \quad \text{(VII.7)} \\ & + \int_{c_0(k^{**}(\infty))-t}^{\infty} k^{**}(\infty)dF(\underline{c}),\end{aligned}$$

where $k^{**}(\infty) = h'^{-1}(1)$ and $k^{**}(\cdot) = h^{-1}(C - t - \underline{c})$ for the second integral, see (III.9). All three integrals always have a positive probability mass because A1(iv) implies that $c_0(\bar{k}) - t > 0$ for all k.

The analysis when $\lambda \to 0$ or $\lambda \to \infty$ is especially easy. When $\lambda \to 0$, the probability that $\underline{c}$ is so high that the incumbent will invest $k^{**}(\infty)$ approaches one. Moreover, for other levels of $\underline{c}$ the investments are $\in [0, \bar{k}]$. Hence, $\lim_{\lambda\to 0} \mathbb{E}[k|\lambda] = k^{**}(\infty)$. On the contrary, if $\lambda \to \infty$, the probability that $\underline{c}$ is in the region where the incumbent does not invest approaches one. This implies that $\lim_{\lambda\to\infty} \mathbb{E}[k|\lambda] = 0$.

But how large is $\mathbb{E}[k|\lambda]$ if we have a medium λ value? Since all areas have a positive weight, $k^{**}(\infty)$ as well as $k^{**}(\cdot)$ are nonnegative and at most $\bar{k}$, it follows that $\bar{k} > \mathbb{E}[k|\lambda] > 0$. But under what conditions can $\mathbb{E}[k|\lambda]$ exceed $k^{**}(\infty)$?

For medium values of $\underline{c}$, i.e., when $c_0(\bar{k}) - t < \underline{c} < c_0(k^{**}(\infty)) - t$, the optimal investment is given by (see Lemma 3)

$$h^{-1}\left(C - t - \underline{c}\right).$$

Denote this part of the investment function by $k^{**}_{\text{medium}}(\cdot)$. The first order Taylor approximation of $k^{**}_{\text{medium}}(\cdot)$ is

$$k^{**}_{\text{medium}}(\cdot) \approx k^{**}(\infty) + c_0(k^{**}(\infty)) - \underline{c} - t.$$

Since $d^2 k^{**}_{\text{medium}}(\cdot)/d\underline{c}^2 > 0$ (see previous analysis) we do not overestimate $k^{**}_{\text{medium}}(\cdot)$ by the approximation. Next, we put the approximation of $k^{**}_{\text{medium}}(\cdot)$ into (VII.7), so

$$\begin{aligned}\mathbb{E}[k|\lambda] \geq \quad & \int_0^{c_0(\bar{k})-t} 0 dF(\underline{c}) \\ & + \int_{c_0(\bar{k})-t}^{c_0(k^{**}(\infty))-t} \left(k^{**}(\infty) + c_0(k^{**}(\infty)) - \underline{c} - t\right) dF(\underline{c}) \\ & + \int_{c_0(k^{**}(\infty))-t}^{\infty} k^{**}(\infty) dF(\underline{c}).\end{aligned}$$

For the exponential density function this is

$$\begin{aligned}\mathbb{E}[k|\lambda] \geq \quad & \int_{c_0(\bar{k})-t}^{c_0(k^{**}(\infty))-t} \left(c_0(k^{**}(\infty)) - \underline{c} - t\right) \lambda e^{-\lambda \underline{c}} d\underline{c} \\ & + \int_{c_0(\bar{k})-t}^{\infty} k^{**}(\infty) \lambda e^{-\lambda \underline{c}} d\underline{c}.\end{aligned} \quad \text{(VII.8)}$$

Since the exponential density is decreasing in $\underline{c}$ and the term in brackets $(c_0(k^{**}(\infty)) - \underline{c} - t)$ is linear in $\underline{c}$, an approximation and a lower bound of the first integral of (VII.8) is

$$\begin{aligned}&\int_{c_0(\bar{k})-t}^{c_0(k^{**}(\infty))-t} \frac{(c_0(k^{**}(\infty)) - c_0(k^{**}(\infty)) + t - t) + \left(c_0(k^{**}(\infty)) - c_0(\bar{k}) + t - t\right)}{2} \lambda e^{-\lambda \underline{c}} d\underline{c} \\ &\qquad = \int_{c_0(\bar{k})-t}^{c_0(k^{**}(\infty))-t} \frac{c_0(k^{**}(\infty)) - c_0(\bar{k})}{2} \lambda e^{-\lambda \underline{c}} d\underline{c}.\end{aligned}$$

With (VII.8) we get

$$\begin{aligned}\mathbb{E}[k|\lambda] \geq \quad & \frac{c_0(k^{**}(\infty)) - c_0(\bar{k})}{2} \left[e^{-\lambda(c_0(\bar{k})-t)} - e^{-\lambda(c_0(k^{**}(\infty))-t)}\right] \\ & + k^{**}(\infty) e^{-\lambda(c_0(\bar{k})-t)} =: \Omega(\lambda)\end{aligned}$$

Obviously, $\lim_{\lambda\to 0}\Omega(\lambda) = k^{**}(\infty)$, and

$$\frac{d\Omega(\lambda)}{d\lambda}e^{\lambda(c_0(\bar{k})-t)} = \frac{c_0(k^{**}(\infty)) - c_0(\bar{k})}{2} \left[-(c_0(\bar{k}) - t) + (c_0(k^{**}(\infty)) - t)e^{-\lambda(c_0(k^{**}(\infty))-c_0(\bar{k}))}\right] - (c_0(\bar{k}) - t)k^{**}(\infty).$$

By A1(v), $c_0(\bar{k}) > t$. From the analysis before, $c_0(k^{**}(\infty)) > c_0(\bar{k})$. Hence, $c_0(k^{**}(\infty)) > t$. Therefore the RHS is decreasing in λ. Moreover, since $e^{\lambda(c_0(\bar{k})-t)} > 0$ the LHS has the same sign as $d\Omega(\lambda)/d\lambda$.

We seek to explore whether $\Omega(\lambda)$ can exceed $k^{**}(\infty)$. Observe that

$$\lim_{\lambda\to 0}\text{RHS} > 0 \iff \frac{\left(c_0(k^{**}(\infty)) - c_0(\bar{k})\right)^2}{2} - (c_0(\bar{k}) - t)k^{**}(\infty) > 0.$$

When this condition holds, $\Omega(\lambda)$ is increasing in λ when λ is sufficiently small. Together with $\lim_{\lambda\to 0}\Omega(\lambda) = k^{**}(\infty)$, this implies that the maximum of $\Omega(\lambda)$ is greater than $k^{**}(\infty)$. Since $\mathbb{E}[k|\lambda] \geq \Omega(\lambda)$, then also the maximum of $\mathbb{E}[k|\lambda]$ is greater than $k^{**}(\infty)$. This establishes the result that under this condition the maximum of $\mathbb{E}[k|\lambda]$ is greater than $k^{**}(\infty)$.
□

PROOF OF PROPOSITION 10

Let the function $k^{***}(C|\lambda)$ denote the optimal investment depending on incumbent's initial costs C, given some λ. We next proof by contradiction. Let $C^I < C^{II}$ and suppose that $k^{***}(C^{II}|\lambda) \geq k^{***}(C^I|\lambda)$. One can rewrite the incumbent's expected profit as

$$\mathbb{E}[\pi_0(k^{***}(C^{II}|\lambda)|C^{II})] = \mathbb{E}[\pi_0(k^{***}(C^{I}|\lambda)|C^{II})] + \int_{k^{***}(C^I|\lambda)}^{k^{***}(C^{II}|\lambda)} \frac{\partial\mathbb{E}[\pi_0(k|C^{II})]}{\partial k}dk.$$

The optimality of $k^{***}(C^{II}|\lambda)$ requires that the term with the integral is nonnegative. The optimality of $k^{***}(C^I|\lambda)$ requires that

$$\mathbb{E}[\pi_0(k^{***}(C^I|\lambda)|C^I)] \geq \mathbb{E}[\pi_0(k^{***}(C^{II}|\lambda)|C^I)],$$

or rewritten that

$$\int_{k^{***}(C^I|\lambda)}^{k^{***}(C^{II}|\lambda)} \frac{\partial \mathbb{E}[\pi_0(k, C^I)]}{\partial k} dk \leq 0.$$

But since

$$k^{***}(C^{II}|\lambda) \geq k^{***}(C^I|\lambda), C^I < C^{II}, \text{ and } d^2\mathbb{E}[\pi_0(k, F)]/dkdC < 0,$$

it follows that

$$\int_{k^{***}(C^I|\lambda)}^{k^{***}(C^{II}|\lambda)} \frac{\partial \mathbb{E}[\pi_0(k, C^I)]}{\partial k} dk > \int_{k^{***}(C^I|\lambda)}^{k^{***}(C^{II}|\lambda)} \frac{\partial \mathbb{E}[\pi_0(k, C^{II})]}{\partial k} dk,$$

which is a contradiction.

Observe that this result does not require that $\underline{c}$ is exponentially distributed. We only used $d^2\mathbb{E}[\pi_0(k, F)]/dkdC < 0$, which also holds with other distributions. □

UNIFORM DISTRIBUTION

We assumed that consumers have a unit demand and that rivals' costs are exponentially distributed. Both assumptions are made to simplify calculations. Since the intuition for our results seems to be quite general, we are optimistic that our results also hold with alternative assumptions. To illustrate, we next consider the case where $\underline{c}$ is uniformly distributed.

$$F(\underline{c}) = \begin{cases} 0 & \text{for } \underline{c} < 0, \\ \lambda \underline{c} & \text{for } \underline{c} \in [0, 1/\lambda], \text{with } \lambda \geq 0, \\ 1 & \text{otherwise.} \end{cases}$$

The strength of the entry threat is again captured by the parameter λ. Since we want to yield explicit solutions we consider a concrete functional form of $h(k)$. Let

$$h(k) = \alpha k^{1/2}.$$

This functional form violates Assumption A2(iv). But this causes no problems when α is small enough. It is useful to define

$$\check{k} \equiv k^{1/2}.$$

We now can write the incumbent's objective function as

$$\mathbb{E}[\pi_0(\check{k})] = (1 - C + \alpha\check{k})(1 - F(C - \alpha\check{k} - t)) - \check{k}^2.$$

Suppose that for the optimal investment entry is possible with a probability strictly between 0 and 1. Then $F(C - \alpha\check{k} - t) = \lambda(C - \alpha\check{k} - t)$. The first order condition $d\mathbb{E}[\pi_0(\check{k})]/d\check{k} = 0$ yields

$$\check{k}^{\mathrm{FOC}} = \frac{\alpha}{2}\,\frac{1 - 2\lambda C + \lambda t + \lambda}{1 - \alpha^2\lambda}. \tag{VII.9}$$

Observe that

$$\frac{d\check{k}^{\mathrm{FOC}}}{d\lambda} = \frac{\alpha}{2}\,\frac{1 - 2C + t + \alpha^2}{(1 - \alpha^2\lambda)^2}.$$

It is easy to see that, as in the original specification with an exponential distribution of rivals' production costs, $C \leq 1/2$ is a sufficient condition that the optimal investment—if it is given by the first-order condition—is increasing in the degree of the entry threat.

It is not obvious whether it is optimal to invest according to the first-order condition. Corner solutions may be preferable. Additionally, it must be checked whether with (VII.9) the probability of entry is indeed strictly between zero and one. To simplify the exercise we consider the very special case with

$$C = \frac{1}{2},\; t = \frac{1}{4},\; \alpha = \frac{1}{4}.$$

Then

$$\check{k}^{\mathrm{FOC}} = \frac{4 + \lambda}{32 - 2\lambda}. \tag{VII.10}$$

Since $k \in \mathbb{R}^{++}$, investing according to formula (VII.10) is possible only for $\lambda \in [0, 16)$. The check whether with (VII.10) the probability of entry is indeed strictly between zero and one yields that we have to restrict the application of formula (VII.10) further to $\lambda \in [0, 28/3]$.

Setting $\mathbb{E}[\pi_0(\check{k}^{\mathrm{FOC}})] = 0$ yields three solutions: $\lambda_1 = 44/9$, $\lambda_2 = 12$, $\lambda_3 = 16$. Observe that only $\lambda_1 \in [0, 28/3]$. As Figure VII.1 shows, only for the subinterval $\lambda \in [0, 44/9]$ yields investing $\check{k}^{\mathrm{FOC}}$ nonnegative expected profits.

Is not investing at all preferable to investing according to (VII.10)? One can show that no investment is strictly dominated by using (VII.10)

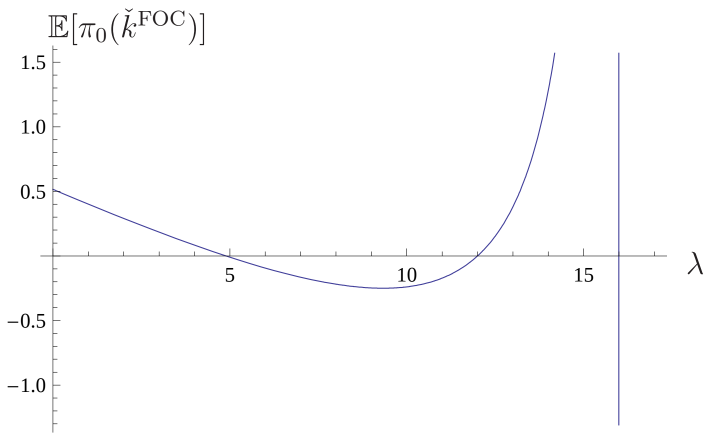

Figure VII.1: Incumbent's expected profit when following (VII.10).

for all $\lambda \in (-4, 16)$. The other corner solution is that entry is deterred for sure. Then $\check{k}$ has to be at least 1 for all $\lambda > 0$, because then $c_0(k) \leq 1/4 = t$. But then the incumbent makes at most a profit of zero. This is worse than not investing because this yields at least a small positive expected profit. So for $\lambda > 44/9$ the incumbent is best off by not investing; for $0 < \lambda < 44/9$ the optimal investment is given by (VII.10); for $\lambda = 44/9$ the incumbent is indifferent between not investing and investing according to (VII.10). We assume that in case of indifference the incumbent chooses to invest. It can, moreover, easily be shown that for $\lambda = 0$, investing according to (VII.10) is optimal. So we conclude that

$$k^*(\lambda) = \begin{cases} \left(\frac{4+\lambda}{32-2\lambda}\right)^2 & \text{for } \lambda \in [0, 44/9], \\ 0 & \text{for } \lambda > 44/9. \end{cases}$$

That is, the relationship between incumbent's R&D investment and the entry threat is again hump-shaped, and for large entry threats the incumbent does not invest at all.

VII.3. APPENDIX TO CHAPTER IV

PROOF OF LEMMA 4

Stochastic independence of the plausible distributions implies that

$$\sum_{\mathcal{S}} p_{i,j}\epsilon_{i,j} \sim N\left(\sum_{\mathcal{S}} p_{i,j}\mu_i, \sum_{\mathcal{S}} p_{i,j}^2\sigma_j^2\right).$$

We assumed that aggregating the plausible distributions yields the objective distribution. So

$$\sum_{\mathcal{S}} p_{i,j}\mu_i \stackrel{!}{=} 0 \quad \text{and} \quad \sum_{\mathcal{S}} p_{i,j}^2\sigma_j^2 \stackrel{!}{=} \sigma^2.$$

The following chain shows that $\sigma^2 < \sigma_n^2$:

$$\sigma^2 = \sum_{\mathcal{S}} p_{i,j}^2\sigma_j^2 < \sum_{\mathcal{S}} p_{i,j}\sigma_j^2 \leq \sum_{\mathcal{S}} p_{i,j}\sigma_n^2 = \sigma_n^2.$$

For $N \geq 2$ we can show that $\mu_1 < 0$:

$$0 = \sum_{\mathcal{S}} p_{i,j}\mu_i > \sum_{\mathcal{S}} p_{i,j}\mu_1 = \mu_1.$$

For $N = 1$, the chain holds with equality. Therefore, $\mu_1 = 0$. □

PROOF OF PROPOSITION 13

Consider the case with an ambiguity-averse agent. The agent's belief about the distribution of ϵ_Z is characterized by $(\mu_{Z,i}, \sigma_{Z,j}^2)$. Simple calculations yield that the agent's certainty equivalent is

$$\hat{w}(\cdot) = t + (s_Y + s_Z)a + s_Z\mu_{Z,i} - \frac{1}{2}\eta\left(s_Y^2\sigma_Y^2 + s_Z^2\sigma_{Z,j}^2\right) - \frac{1}{2}ca^2. \quad \text{(VII.11)}$$

Therefore, the agent chooses

$$a = \begin{cases} \frac{s_Y+s_Z}{c} & \text{for } s_Y + s_Z \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

It is straightforward that setting $s_Z < 0$ makes no sense for the principal. For the same reasons as in case with only one performance measure we then have $\mu_{i,Z} = \mu_{Z,1}$ and $\sigma^2_{Z,j} = \sigma^2_{Z,n}$. The principal's maximization problem can be simplified to

$$\max_{s_Y,s_Z} \frac{s_Y + s_Z}{c} + s_Z\mu_{Z,1} - \frac{1}{2}\eta(s_Y^2\sigma_Y^2 + s_Z^2\sigma^2_{Z,n}) - \frac{1}{2}\frac{(s_Y+s_Z)^2}{c} - \bar{w} \text{ s.t. } s_Z \geq 0. \tag{VII.12}$$

The derivatives are

$$\begin{aligned}\frac{d...}{ds_Y} &= \frac{1}{c} - \eta s_Y\sigma_Y^2 - \frac{s_Y + s_Z}{c},\\ \frac{d...}{ds_Z} &= \frac{1}{c} - \eta s_Z\sigma^2_{Z,n} - \frac{s_Y + s_Z}{c} + \mu_{Z,1}.\end{aligned}$$

It is straightforward that $s_Y = 0$ and $s_Z > 0$ cannot be optimal: then $d.../ds_Y > d.../ds_Z$, which is a contradiction when $s_Y = 0$ and $s_Z > 0$. Moreover, $s_Y = 0$ and $s_Z = 0$ cannot be optimal: then $d.../ds_Y$ is positive for $s_Y = 0$ and $s_Z = 0$. Hence, $s_Y > 0$ and $d.../ds_Y = 0$. So

$$s_Y = \frac{1 - s_Z}{1 + c\eta\sigma_Y^2}.$$

Using this yields

$$\frac{d...}{ds_Z} = c\left(1 + c\mu_{Z,1} - \frac{1}{1+\eta c\sigma_Y^2} + s_Z\left(\frac{1}{1+\eta c\sigma_Y^2} - 1 - \eta c\sigma^2_{Z,n}\right)\right).$$

One directly sees that the RHS is decreasing in s_Z.

Hence, when $\left.\frac{d...}{ds_Z}\right|_{s_Z=0} \leq 0$ the principal sets $s_Z = 0$. Otherwise he sets $s_Z > 0$. So $s_Z > 0$ if and only if

$$\mu_{Z,1} > -\frac{\eta\sigma_Y^2}{1+\eta c\sigma_Y^2}.$$

□

VII.4. APPENDIX TO CHAPTER V

PROOF OF LEMMA 6

Suppose that signals are ordered according to their likelihood ratio, that is, $s > s'$ if and only if $\gamma_s^H/\gamma_s^L > \gamma_{s'}^H/\gamma_{s'}^L$. Consider a contract of the form

$$u_s = \begin{cases} \underline{u} & \text{if} \quad s < \hat{s} \\ \underline{u} + b & \text{if} \quad s \geq \hat{s} \end{cases},$$

where $b > 0$ and $1 < \hat{s} \leq S$. Under this contractual form and given that the first-order approach is valid, (IC) can be rewritten as

$$b \Bigg\{ \left[\sum_{s=\hat{s}}^{S} (\gamma_s^H - \gamma_s^L) \right] \left(1 - (\lambda - 1) \sum_{s=1}^{\hat{s}-1} \gamma_s(\hat{a}) \right) - (\lambda - 1) \left(\sum_{s=1}^{\hat{s}-1} (\gamma_s^H - \gamma^L) \right) \left(\sum_{s=\hat{s}}^{S} \gamma_s(\hat{a}) \right) \Bigg\} = c'(\hat{a}).$$

Since signals are ordered according to their likelihood ratio, we have $\sum_{s=\hat{s}}^{S} (\gamma_s^H - \gamma_s^L) > 0$ and $\sum_{s=1}^{\hat{s}-1} (\gamma_s^H - \gamma^L) < 0$ for all $1 < \hat{s} \leq S$. This implies that the term in curly brackets is strictly positive for $\lambda \leq 2$. Hence, with $c'(\hat{a}) > 0$, b can alway be chosen such that (IC) is met. Rearranging the participation constraint,

$$\underline{u} \geq \bar{u} + c(\hat{a}) - b \left(\sum_{s=\hat{s}}^{S} \gamma_s(\hat{a}) \right) \left[1 - (\lambda - 1) \left(\sum_{s=1}^{\hat{s}-1} \gamma_s(\hat{a}) \right) \right],$$

reveals that (IR) can be satisfied for any b by choosing $\underline{u}$ appropriately. This concludes the proof. □

PROOF OF PROPOSITION 14

It is readily verified that Assumptions 1-3 from Grossman and Hart (1983) are satisfied. Thus, the cost-minimization problem is well defined, in the sense that for each action $a \in (0, 1)$ there exists a second-best incentive scheme. Suppose the principal wants to implement action $\hat{a} \in (0, 1)$ at minimum cost. Since the agent's action is not observable, the principal's problem is given by

$$\min_{\{u_s\}_{s=1}^{S}} \sum_{s=1}^{S} \gamma_s(\hat{a}) h(u_s) \qquad \text{(MR)}$$

subject to

$$\sum_{s=1}^{S} \gamma_s(\hat{a})u_s - c(\hat{a}) \geq \bar{u} \,, \tag{IR$_R$}$$

$$\sum_{s=1}^{S} (\gamma_s^H - \gamma_s^L)u_s - c'(\hat{a}) = 0 \,. \tag{IC$_R$}$$

where the first constraint is the individual rationality constraint and the second is the incentive compatibility constraint. Note that the first-order approach is valid, since the agent's expected utility is a strictly concave function of his effort. The Lagrangian to the resulting problem is

$$\mathcal{L} = \sum_{s=1}^{S} \gamma_s(a)h(u_s) - \mu_0 \left\{ \sum_{s=1}^{S} \gamma_s(a)u_s - c(a) - \bar{u} \right\} - \mu_1 \left\{ \sum_{s=1}^{S} (\gamma_s^H - \gamma_s^L)u_s - c'(a) \right\},$$

where μ_0 and μ_1 denote the Lagrange multipliers of the individual rationality constraint and the incentive compatibility constraint, respectively. Setting the partial derivative of $\mathcal{L}$ with respect to u_s equal to zero yields

$$\frac{\partial \mathcal{L}}{\partial u_s} = 0 \iff h'(u_s) = \mu_0 + \mu_1 \frac{\gamma_s^H - \gamma_s^L}{\gamma_s(\hat{a})}, \quad \forall s \in \mathcal{S}. \tag{VII.13}$$

Irrespective of the value of μ_0, if $\mu_1 > 0$, convexity of $h(\cdot)$ implies that $u_s > u_{s'}$ if and only if $(\gamma_s^H - \gamma_s^L)/\gamma_s(\hat{a}) > (\gamma_{s'}^H - \gamma_{s'}^L)/\gamma_{s'}(\hat{a})$, which in turn is equivalent to $\gamma_s^H/\gamma_s^L > \gamma_{s'}^H/\gamma_{s'}^L$. Thus it remains to show that μ_1 is strictly positive. Suppose, in contradiction, that $\mu_1 \leq 0$. Consider the case $\mu_1 = 0$ first. From (A.1) it follows that $u_s = u^f$ for all $s \in \mathcal{S}$, where u^f satisfies $h'(u^f) = \mu_0$. This, however, violates (IC$_R$), a contradiction. Next, consider $\mu_1 < 0$. From (A.1) it follows that $u_s < u_{s'}$ if and only if $(\gamma_s^H - \gamma_s^L)/\gamma_s(\hat{a}) > (\gamma_{s'}^H - \gamma_{s'}^L)/\gamma_t(\hat{a})$. Let $\mathcal{S}^+ \equiv \{s | \gamma_s^H - \gamma_s^L > 0\}$, $\mathcal{S}^- \equiv \{s | \gamma_s^H - \gamma_s^L < 0\}$, and $\hat{u} \equiv \min\{u_s | s \in \mathcal{S}^-\}$. Since $\hat{u} > u_s$ for all

$s \in \mathcal{S}^+$, we have

$$\begin{aligned}\sum_{s=1}^{S}(\gamma_s^H - \gamma_s^L)u_s &= \sum_{\mathcal{S}^-}(\gamma_s^H - \gamma_s^L)u_s + \sum_{\mathcal{S}^+}(\gamma_s^H - \gamma_s^L)u_s \\ &< \sum_{\mathcal{S}^-}(\gamma_s^H - \gamma_s^L)\hat{u} + \sum_{\mathcal{S}^+}(\gamma_s^H - \gamma_s^L)\hat{u} \\ &= \hat{u}\sum_{s=1}^{S}(\gamma_s^H - \gamma_s^L) \\ &= 0,\end{aligned}$$

again a contradiction to (IC$_R$). Hence, $\mu_1 > 0$ and the desired result follows. □

PROOF OF PROPOSITION 15

The problem of finding the optimal contract $\boldsymbol{u}^*$ to implement action $\hat{a} \in (0, 1)$ is decomposed into two subproblems. First, for a given incentive feasible ordering of signals, we derive the optimal nondecreasing incentive scheme that implements action $\hat{a} \in (0, 1)$. Then, in a second step, we choose the ordering of signals for which the ordering specific cost of implementation is lowest.

Step 1: Remember that the ordering of signals is incentive feasible if $\beta_s(\cdot) > 0$ for at least one signal s. For a given incentive feasible ordering of signals, in this first step we solve Program ML. First, note that it is optimal to set $b_s = 0$ if $\beta_s(\cdot) < 0$. To see this, suppose, in contradiction, that in the optimum (IC$'$) holds and $b_s > 0$ for some signal s with $\beta_s(\cdot) \leq 0$. If $\beta_s(\cdot) = 0$, then setting $b_s = 0$ leaves (IC$'$) unchanged, but leads to a lower value of the objective function of Program ML, contradicting that the original contract is optimal. If $\beta_s(\cdot) < 0$, then setting $b_s = 0$ not only reduces the value of the objective function, but also relaxes (IC$'$), which in turn allows to lower other bonus payments, thereby lowering the value of the objective function even further. Again, a contradiction to the original contract being optimal. Let $\mathcal{S}_\beta \equiv \{s \in \mathcal{S} | \beta_s(\cdot) > 0\}$ denote the set of signals for which $\beta_s(\cdot)$ is strictly positive under the considered ordering of signals, and let S_β denote the number of elements in this set. Thus, Program (ML) can be rewritten as

Program ML$^+$:

$$\begin{array}{llll} & & \displaystyle\min_{\{b_s\}_{s\in\mathcal{S}_\beta}} \sum_{s\in\mathcal{S}_\beta} b_s\rho_s(\hat{\boldsymbol{\gamma}},\lambda,\hat{a}) & \\ \text{subject to} & (i) & \displaystyle\sum_{s\in\mathcal{S}_\beta} b_s\beta_s(\hat{\boldsymbol{\gamma}},\lambda,\hat{a}) = c'(\hat{a}) & (\text{IC}^+) \\ & (ii) & b_s \geq 0, \quad \forall s\in\mathcal{S}_\beta. & \end{array}$$

Program ML$^+$ is a linear programming problem. It is well-known that if a linear programming problem has a solution, it must have a solution at an extreme point of the constraint set. Generically, there is a unique solution and this solution is an extreme point. Since the constraint set of Program ML$^+$, $\mathcal{M} \equiv \{\{b_s\}_{s\in\mathcal{S}_\beta} \in \mathbb{R}_+^{S_\beta} \,|\, \sum_{s\in\mathcal{S}_\beta} b_s\beta_s(\hat{\boldsymbol{\gamma}},\lambda,\hat{a}) = c'(\hat{a})\}$, is closed and bounded, Program ML$^+$ has a solution. Hence, generically $\sum_{s\in\mathcal{S}_\beta} b_s\rho_s(\hat{\boldsymbol{\gamma}},\lambda,\hat{a})$ achieves its greatest lower bound at one of the extreme points of $\mathcal{M}$. (We comment on genericity below.) With $\mathcal{M}$ describing a hyperplane in $\mathbb{R}_+^{S_\beta}$, all extreme points of $\mathcal{M}$ are characterized by the following property: $b_s > 0$ for exactly one signal $s \in S_\beta$ and $b_t = 0$ for all $t \in S_\beta$, $t \neq s$. It remains to determine for which signal the bonus is set strictly positive. The size of the bonus payment, which is set strictly positive, is uniquely determined by (IC$^+$):

$$b_s\beta_s(\hat{\boldsymbol{\gamma}},\lambda,\hat{a}) = c'(\hat{a}) \iff b_s = \frac{c'(\hat{a})}{\beta_s(\hat{\boldsymbol{\gamma}},\lambda,\hat{a})}. \tag{VII.14}$$

Therefore, from the objective function of Program ML$^+$ it follows that, for the signal ordering under consideration, the optimal signal for which the bonus is set strictly positive, $\hat{s}$, is characterized by

$$\hat{s} = \arg\min_{s\in\mathcal{S}_\beta} \frac{c'(\hat{a})}{\beta_s(\hat{\boldsymbol{\gamma}},\lambda,\hat{a})}\rho_s(\hat{\boldsymbol{\gamma}},\lambda,\hat{a}).$$

Step 2: From all incentive feasible signal orders, the principal chooses the one which minimizes her cost of implementation. With the number of incentive feasible signal orders being finite, this problem clearly has a solution. Let s^* denote the resulting cutoff, i.e.,

$$u_s^* = \begin{cases} u^* & \text{if} \quad s < s^* \\ u^* + b^* & \text{if} \quad s \geq s^* \end{cases},$$

where

$$b^* = c'(\hat{a})/\beta_{s^*}(\hat{\boldsymbol{\gamma}}, \lambda, \hat{a}) \text{ and}$$

$$u^* = \bar{u} + c(\hat{a}) - b^* \left[\sum_{\tau=s^*}^{S} \gamma_\tau(\hat{a}) - \rho_{s^*}(\hat{\boldsymbol{\gamma}}, \lambda, \hat{a}) \right].$$

Letting $u_L^* = u^*$, $u_H^* = u^* + b^*$, and $\mathcal{B}^* = \{s \in \mathcal{S} | s \geq s^*\}$ establishes the desired result.

On genericity: We claimed that, for any given feasible ordering of signals, generically Program ML$^+$ has a unique solution at one of the extreme points of the constraint set. To see this, note that a necessary condition for the existence of multiple solutions is $\beta_s/\beta_{s'} = \rho_s/\rho_{s'}$ for some $s, s' \in \mathcal{S}_\beta$, $s \neq s'$. This condition is characterized by the action to be implemented, $\hat{a}$, the structure of the performance measure, $\{(\gamma_s^H, \gamma_s^L)\}_{s=1}^S$, and the agent's degree of loss aversion, λ. Now, fix $\hat{a}$ and $\{(\gamma_s^H, \gamma_s^L)\}_{s=1}^S$. With both $\beta_s > 0$ and $\rho_s > 0$ for all $s \in \mathcal{S}_\beta$, it is readily verified, that exactly one value of λ equates $\beta_s/\beta_{s'}$ with $\rho_s/\rho_{s'}$. Since λ is drawn from the interval $(1, 2]$, and with the number of signals being finite, this necessary condition for Program ML$^+$ having multiple solutions for a given feasible ordering of signals generically will not hold. With the number of feasible orderings being finite, generic optimality of a corner solution carries over to the overall problem. □

PROOF OF PROPOSITION 16

$\mathcal{B}^*$ maximizes $X(\mathcal{B}) := \left[\sum_{s\in\mathcal{B}}(\gamma_s^H - \gamma_s^L)\right] \times Y(P_\mathcal{B})$, where

$$Y(P_\mathcal{B}) := \frac{1}{(\lambda - 1)P_\mathcal{B}(1 - P_\mathcal{B})} - \frac{1}{P_\mathcal{B}} + \frac{1}{1 - P_\mathcal{B}}.$$

Suppose for the moment that $P_\mathcal{B}$ is a continuous decision variable. Accordingly,

$$\frac{dY(P_\mathcal{B})}{dP_\mathcal{B}} = \frac{1}{P_\mathcal{B}^2(1 - P_\mathcal{B})^2} \left[2P_\mathcal{B}^2 + \frac{2 - \lambda}{\lambda - 1}(2P_\mathcal{B} - 1) \right]. \qquad \text{(VII.15)}$$

It is readily verified that $dY(P_\mathcal{B})/dP_\mathcal{B} < 0$ for $0 < P_\mathcal{B} < \bar{P}(\lambda)$ and $dY(P_\mathcal{B})/dP_\mathcal{B} > 0$ for $\bar{P}(\lambda) < P_\mathcal{B} < 1$, where

$$\bar{P}(\lambda) \equiv \frac{\lambda - 2 + \sqrt{\lambda(2 - \lambda)}}{2(\lambda - 1)}.$$

Note that for $\lambda \leq 2$ the critical value $\bar{P}(\lambda) \in [0, 1/2)$. Hence, excluding a signal of $\mathcal{B}$ increases $Y(P_{\mathcal{B}})$ if $P_{\mathcal{B}} < \bar{P}(\lambda)$, whereas including a signal to $\mathcal{B}$ increases $Y(P_{\mathcal{B}})$ if $P_{\mathcal{B}} \geq \bar{P}(\lambda)$. With these insights the next two implications follow immediately.

(i) $P_{\mathcal{B}^*} < \bar{P}(\lambda) \implies \mathcal{B}^* \subseteq \mathcal{S}^+$

(ii) $P_{\mathcal{B}^*} \geq \bar{P}(\lambda) \implies \mathcal{S}^+ \subseteq \mathcal{B}^*$

We prove both statements in turn by contradiction. (i) Suppose $P_{\mathcal{B}^*} < \bar{P}(\lambda)$ and that there exists a signal $\hat{s} \in \mathcal{S}^-$ which is also contained in $\mathcal{B}^*$, i.e., $\hat{s} \in \mathcal{B}^*$. Clearly, $\sum_{s \in \mathcal{B}^*}(\gamma_s^H - \gamma_s^L) < \sum_{s \in \mathcal{B}^* \setminus \{\hat{s}\}}(\gamma_s^H - \gamma_s^L)$ because $\hat{s}$ is a bad signal. Moreover, $Y(\mathcal{B}^*) < Y(\mathcal{B}^* \setminus \{\hat{s}\})$ because $Y(\cdot)$ increases when signals are excluded of $\mathcal{B}^*$. Thus $X(\mathcal{B}^*) < X(\mathcal{B}^* \setminus \{\hat{s}\})$, a contradiction to the assumption that $\mathcal{B}^*$ is the optimal partition. (ii) Suppose $P_{\mathcal{B}^*} \geq \bar{P}(\lambda)$ and that there exists a signal $\tilde{s} \in \mathcal{S}^+$ that is not contained in $\mathcal{B}^*$, i.e., $\mathcal{B}^* \cap \{\tilde{s}\} = \emptyset$. Since $\hat{s}$ is a good signal $\sum_{s \in \mathcal{B}^*}(\gamma_s^H - \gamma_s^L) < \sum_{s \in \mathcal{B}^* \cup \{\hat{s}\}}(\gamma_s^H - \gamma_s^L)$. $P_{\mathcal{B}^*} \geq \bar{P}(\lambda)$ implies that $Y(\mathcal{B}^* \cup \{\tilde{s}\}) > Y(\mathcal{B}^*)$. Thus, $X(\mathcal{B}^*) < X(\mathcal{B}^* \cup \{\tilde{s}\})$ a contradiction to the assumption that $\mathcal{B}^*$ maximizes $X(\mathcal{B}^*)$. Finally, since for any $\mathcal{B}^*$ we are either in case (i) or in case (ii), the desired result follows. □

PROOF OF PROPOSITION 17

Suppose, in contradiction, that in the optimum there are signals $s, t \in \mathcal{S}$ such that $s \in \mathcal{B}^*$, $t \notin \mathcal{B}^*$ and $\frac{\gamma_s^H - \gamma_s^L}{\gamma_s(\hat{a})} < \frac{\gamma_t^H - \gamma_t^L}{\gamma_t(\hat{a})}$. We derive a contradiction by showing that exchanging signal s for signal t reduces the principal's cost, which implies that the original contract cannot be optimal. Let $\bar{\mathcal{B}} \equiv (\mathcal{B}^* \setminus \{s\}) \cup \{t\}$.It suffices to show that $X(\bar{\mathcal{B}}) > X(\mathcal{B}^*)$, where $X(\mathcal{B})$ is defined as in the proof of Proposition 16. $X(\bar{\mathcal{B}}) > X(\mathcal{B}^*)$ is equivalent to

$$\left(\sum_{j \in \mathcal{B}^*}(\gamma_j^H - \gamma_j^L) + (\gamma_t^H - \gamma_t^L) - (\gamma_s^H - \gamma_s^L)\right)\left[\frac{1-(\lambda-1)(1-2P_{\bar{\mathcal{B}}})}{(\lambda-1)P_{\bar{\mathcal{B}}}(1-P_{\bar{\mathcal{B}}})}\right] >$$
$$\left(\sum_{j \in \mathcal{B}^*}(\gamma_j^H - \gamma_j^L)\right)\left[\frac{1-(\lambda-1)(1-2P_{\mathcal{B}^*})}{(\lambda-1)P_{\mathcal{B}^*}(1-P_{\mathcal{B}^*})}\right].$$

Rearranging yields

$$\left[(\gamma_t^H - \gamma_t^L) - (\gamma_s^H - \gamma_s^L)\right]\left[\frac{1-(\lambda-1)(1-2P_{\bar{\mathcal{B}}})}{(\lambda-1)P_{\bar{\mathcal{B}}}(1-P_{\bar{\mathcal{B}}})}\right] > \left(\sum_{j\in\mathcal{B}^*}(\gamma_j^H - \gamma_j^L)\right)\left[\frac{1-(\lambda-1)(1-2P_{\mathcal{B}^*})}{(\lambda-1)P_{\mathcal{B}^*}(1-P_{\mathcal{B}^*})} - \frac{1-(\lambda-1)(1-2P_{\bar{\mathcal{B}}})}{(\lambda-1)P_{\bar{\mathcal{B}}}(1-P_{\bar{\mathcal{B}}})}\right]. \tag{VII.16}$$

With $Y(P_{\mathcal{B}})$ being defined as in the proof of Proposition 16, we have to consider two cases, (i) $dY(P_{\mathcal{B}^*})/P_{\mathcal{B}} \geq 0$, and (ii) $dY(P_{\mathcal{B}^*})/P_{\mathcal{B}} < 0$.

Case (i): Since $\gamma_s(\hat{a}) - \gamma_t(\hat{a}) \leq \kappa$, we have $P_{\mathcal{B}^*} \leq P_{\bar{\mathcal{B}}} + \kappa$. With $Y(P_{\mathcal{B}})$ being (weakly) increasing at $P_{\mathcal{B}^*}$, inequality (VII.16) is least likely to hold for $P_{\mathcal{B}^*} = P_{\bar{\mathcal{B}}} + \kappa$. Inserting $P_{\mathcal{B}^*} = P_{\bar{\mathcal{B}}} + \kappa$ into (VII.16) yields

$$\left[(\gamma_t^H - \gamma_t^L) - (\gamma_s^H - \gamma_s^L)\right]\left[\frac{1-(\lambda-1)(1-2P_{\bar{\mathcal{B}}})}{(\lambda-1)P_{\bar{\mathcal{B}}}(1-P_{\bar{\mathcal{B}}})}\right] > \left(\sum_{j\in\mathcal{B}^*}(\gamma_j^H - \gamma_j^L)\right)\left[\frac{1-(\lambda-1)(1-2P_{\bar{\mathcal{B}}}-2\kappa)}{(\lambda-1)[P_{\bar{\mathcal{B}}}(1-P_{\bar{\mathcal{B}}}) + \kappa(1-2P_{\bar{\mathcal{B}}})] - \kappa^2} - \frac{1-(\lambda-1)(1-2P_{\bar{\mathcal{B}}})}{(\lambda-1)P_{\bar{\mathcal{B}}}(1-P_{\bar{\mathcal{B}}})}\right]. \tag{VII.17}$$

The right-hand side of (VII.17) becomes arbitrarily close to zero for $\kappa \to 0$, thus it remains to show that

$$\left[(\gamma_t^H - \gamma_t^L) - (\gamma_s^H - \gamma_s^L)\right]\left[\frac{1-(\lambda-1)(1-2P_{\bar{\mathcal{B}}})}{(\lambda-1)P_{\bar{\mathcal{B}}}(1-P_{\bar{\mathcal{B}}})}\right] > 0. \tag{VII.18}$$

For (VII.18) to hold, we must have $(\gamma_t^H - \gamma_t^L) - (\gamma_s^H - \gamma_s^L) > 0$. From the proof of Proposition 16 we know that $\mathcal{S}^+ \subseteq \mathcal{B}^*$ if $Y(P_{\mathcal{B}})$ is increasing at $\mathcal{B}^*$. Since the principal will end up including all good signals in the set $\mathcal{B}^*$ anyway, the question of interest is whether she can benefit from swapping two bad signals. Therefore, we consider case $s, t \in \mathcal{S}^-$, where $\mathcal{S}^- \equiv \{s \in \mathcal{S} | \gamma_s^H - \gamma_s^L < 0\}$. With $s, t \in \mathcal{S}^-$, we have

$$\left[(\gamma_t^H - \gamma_t^L) - (\gamma_s^H - \gamma_s^L)\right] \geq \gamma_t(\hat{a})\gamma_s(\hat{a}) \cdot \left[\frac{1}{\gamma_s(\hat{a})}\frac{\gamma_t^H - \gamma_t^L}{\gamma_t(\hat{a})} - \frac{1}{\gamma_s(\hat{a}) + \kappa}\frac{\gamma_s^H - \gamma_s^L}{\gamma_s(\hat{a})}\right], \tag{VII.19}$$

where the inequality holds because $\gamma_t(\hat{a}) - \gamma_s(\hat{a}) \leq \kappa$. Note that for $\kappa \to 0$ the right-hand side of (VII.19) becomes strictly positive, thus

$(\gamma_t^H - \gamma_t^L) - (\gamma_s^H - \gamma_s^L) > 0$ for $\kappa \to 0$. Hence, for κ sufficiently small, $X(\mathcal{B}^*) < X(\bar{\mathcal{B}})$, a contradiction to $\mathcal{B}^*$ being optimal.

Case (ii): Since $\gamma_t(\hat{a}) - \gamma_s(\hat{a}) \leq \kappa$, we have $P_{\mathcal{B}^*} \geq P_{\bar{\mathcal{B}}} - \kappa$. With $Y(P_{\mathcal{B}})$ being decreasing at $P_{\mathcal{B}^*}$, inequality (VII.16) is least likely to hold for $P_{\mathcal{B}^*} = P_{\bar{\mathcal{B}}} - \kappa$. Inserting $P_{\mathcal{B}^*} = P_{\bar{\mathcal{B}}} - \kappa$ into (VII.16), and running along the lines of case (i) allows us to establish that, for κ sufficiently small, $X(\mathcal{B}^*) < X(\bar{\mathcal{B}})$, a contradiction to $\mathcal{B}^*$ being optimal.

To sum up, for κ sufficiently small we have

$$\max_{s \in \mathcal{S} \setminus \mathcal{B}^*} \{(\gamma_s^H - \gamma_s^L)/\gamma_s(\hat{a})\} < \min_{s \in \mathcal{B}^*} \{(\gamma_s^H - \gamma_s^L)/\gamma_s(\hat{a})\} ,$$

or equivalently,

$$\max_{s \in \mathcal{S} \setminus \mathcal{B}^*} \{\gamma_s^H/\gamma_s^L\} < \min_{s \in \mathcal{B}^*} \{\gamma_s^H/\gamma_s^L\}.$$

Letting $K \equiv \min_{s \in \mathcal{B}^*}\{\gamma_s^H/\gamma_s^L\}$ establishes the desired result. □

PROOF OF PROPOSITION 18

We first prove part (ii). Suppose that a small change in λ leaves the optimal partition $\mathcal{B}^*$ of the set of all signals unchanged. Rearranging (IC′) yields

$$b^* = \frac{c'(\hat{a})}{\sum_{s \in \mathcal{B}^*}(\gamma_s^H - \gamma_s^L) - (\lambda - 1)\left[\sum_{s \in \mathcal{B}^*}(\gamma_s^H - \gamma_s^L)\right][1 - 2P_{\mathcal{B}^*}]}.$$

Straight-forward differentiation reveals that

$$\frac{db^*}{d\lambda} = \frac{c'(\hat{a})\left[\sum_{s \in \mathcal{B}^*}(\gamma_s^H - \gamma_s^L)\right][1 - 2P_{\mathcal{B}^*}]}{\left\{\sum_{s \in \mathcal{B}^*}(\gamma_s^H - \gamma_s^L) - (\lambda - 1)\left[\sum_{s \in \mathcal{B}^*}(\gamma_s^H - \gamma_s^L)\right][1 - 2P_{\mathcal{B}^*}]\right\}^2}.$$

Since under the second-best contract $\sum_{s \in \mathcal{B}^*}(\gamma_s^H - \gamma_s^L) > 0$, the desired result follows.
To prove part (i), let $\mathcal{B}^+ \equiv \left\{\mathcal{B} \subset \mathcal{S} \,|\, \sum_{s \in \mathcal{B}}(\gamma_s^H - \gamma_s^L) > 0\right\}$. For any $\tilde{\mathcal{B}} \in \mathcal{B}^+$, let

$$b_{\tilde{\mathcal{B}}} = \frac{c'(\hat{a})}{\sum_{s \in \tilde{\mathcal{B}}}(\gamma_s^H - \gamma_s^L) - (\lambda - 1)\left[\sum_{s \in \tilde{\mathcal{B}}}(\gamma_s^H - \gamma_s^L)\right][1 - 2P_{\tilde{\mathcal{B}}}]}$$

and

$$\underline{u}_{\tilde{\mathcal{B}}} = \bar{u} + c(\hat{a}) - b_{\tilde{\mathcal{B}}} P_{\tilde{\mathcal{B}}} + (\lambda - 1) P_{\tilde{\mathcal{B}}}(1 - P_{\tilde{\mathcal{B}}}) b_{\tilde{\mathcal{B}}}.$$

The cost of implementing action $\hat{a}$ when paying $\underline{u}_{\tilde{\mathcal{B}}}$ for signals in $\mathcal{S} \setminus \tilde{\mathcal{B}}$ and $\underline{u}_{\tilde{\mathcal{B}}} + b_{\tilde{\mathcal{B}}}$ for signals in $\tilde{\mathcal{B}}$ is given by

$$C_{\tilde{\mathcal{B}}} = \underline{u}_{\tilde{\mathcal{B}}} + b_{\tilde{\mathcal{B}}} P_{\tilde{\mathcal{B}}} = \bar{u} + c(\hat{a}) + \frac{c'(\hat{a})(\lambda - 1)P_{\tilde{\mathcal{B}}}(1 - P_{\tilde{\mathcal{B}}})}{\left[\sum_{s \in \tilde{\mathcal{B}}}(\gamma_s^H - \gamma_s^L)\right]\left[1 - (\lambda - 1)(1 - 2P_{\tilde{\mathcal{B}}})\right]}.$$

Differentiation of $C_{\tilde{\mathcal{B}}}$ with respect to λ yields

$$\frac{dC_{\tilde{\mathcal{B}}}}{d\lambda} = \frac{c'(\hat{a})P_{\tilde{\mathcal{B}}}(1 - P_{\tilde{\mathcal{B}}})}{\left[\sum_{s \in \tilde{\mathcal{B}}}(\gamma_s^H - \gamma_s^L)\right]\left[1 - (\lambda - 1)(1 - 2P_{\tilde{\mathcal{B}}})\right]^2}.$$

Obviously, $dC_{\tilde{\mathcal{B}}}/d\lambda > 0$ for all $\mathcal{B} \in \mathcal{B}^+$. Since the optimal partition of $\mathcal{S}$ may change as λ changes, the minimum cost of implementing action $\hat{a}$ is given by

$$C(\hat{a}) = \min_{\mathcal{B} \in \mathcal{B}^+} C_{\mathcal{B}}.$$

Put differently, $C(\hat{a})$ is the lower envelope of all $C_{\mathcal{B}}$ for $\mathcal{B} \in \mathcal{B}^+$. With $C_{\mathcal{B}}$ being continuous and strictly increasing in λ for all $\mathcal{B} \in \mathcal{B}^+$, it follows that also $C(\hat{a})$ is continuous and strictly increasing in λ. This completes the proof. □

PROOF OF LEMMA 7

We show that program (MG) has a solution, that is, $\sum_{s=1}^{S} \gamma_s(\hat{a})h(u_s)$ achieves its greatest lower bound. First, from Lemma 6 we know that the constraint set of program (MG) is not empty for action $\hat{a} \in (0, 1)$. Next, note that from (IR_G) it follows that $\sum_{s=1}^{S} \gamma_s(\hat{a})u_s$ is bounded below. Following the reasoning in the proof of Proposition 1 of Grossman and Hart (1983), we can artificially bound the constraint set – roughly spoken because unbounded sequences in the constraint set make $\sum_{s=1}^{S} \gamma_s(\hat{a})h(u_s)$ tend to infinity by a result from Bertsekas (1974). Since the constraint set is closed, the existence of a minimum follows from Weierstrass' theorem. □

PROOF OF LEMMA 8

Since (IR_G) will always be satisfied with equality due to an appropriate adjustment of the lowest intrinsic utility level offered, relaxing (IR_G) will always lead to strictly lower costs for the principal. Therefore, the shadow value of relaxing (IR_G) is strictly positive, so $\mu_{IR} > 0$.

Next, we show that relaxing (IC_G) has a positive shadow value, $\mu_{IC} > 0$. We do this by showing that a decrease in $c'(\hat{a})$ leads to a reduction in the principal's minimum cost of implementation. Let $\{u_s^*\}_{s\in\mathcal{S}}$ be the optimal contract under (the original) Program MG, and suppose that $c'(\hat{a})$ decreases. Now the principal can offer a new contract $\{u_s^N\}_{s\in\mathcal{S}}$ of the form

$$u_s^N = \alpha u_s^* + (1-\alpha)\sum_{t=1}^{S}\gamma_t(\hat{a})u_t^* \ , \tag{VII.20}$$

where $\alpha \in (0,1)$, which also satisfies (IR_G), the relaxed (IC_G), and (OC_G), but yields strictly lower costs of implementation than the original contract $\{u_s^*\}_{s\in\mathcal{S}}$.

Clearly, for $\hat{\alpha} \in (0,1)$, $u_s^N < u_{s'}^N$ if and only if $u_s^* < u_{s'}^*$, so (OC_G) is also satisfied under contract $\{u_s^N\}_{s\in\mathcal{S}}$.

Next, we check that the relaxed (IC_G) holds under $\{u_s^N\}_{s\in\mathcal{S}}$. To see this, note that for $\alpha = 1$ we have $\{u_s^N\}_{s\in\mathcal{S}} \equiv \{u_s^*\}_{s\in\mathcal{S}}$. Thus, for $\alpha = 1$, the relaxed (IC_G) is oversatisfied under $\{u_s^N\}_{s\in\mathcal{S}}$. For $\alpha = 0$, on the other hand, the left-hand side of (IC_G) is equal to zero, and the relaxed (IC_G) in consequence is not satisfied. Since the left-hand side of (IC_G) is continuous in α under contract $\{u_s^N\}_{s\in\mathcal{S}}$, by the intermediate-value theorem there exists $\hat{\alpha} \in (0,1)$ such that the relaxed (IC_G) is satisfied with equality.

Last, consider (IR_G). The left-hand side of (IR_G) under contract $\{u_s^N\}_{s\in\mathcal{S}}$ with $\alpha = \hat{\alpha}$ amounts to

$$\begin{aligned}
&\sum_{s=1}^{S}\gamma_s(\hat{a})u_s^N - (\lambda-1)\sum_{s=1}^{S-1}\sum_{t=s+1}^{S}\gamma_s(\hat{a})\gamma_t(\hat{a})\left[u_t^N - u_s^N\right] \\
&= \sum_{s=1}^{S}\gamma_s(\hat{a})u_s^* - \tilde{\alpha}(\lambda-1)\sum_{s=1}^{S-1}\sum_{t=s+1}^{S}\gamma_s(\hat{a})\gamma_t(\hat{a})\left[u_t^* - u_s^*\right] \\
&> \sum_{s=1}^{S}\gamma_s(\hat{a})u_s^* - (\lambda-1)\sum_{s=1}^{S-1}\sum_{t=s+1}^{S}\gamma_s(\hat{a})\gamma_t(\hat{a})\left[u_t^* - u_s^*\right] \\
&= \bar{u} + c(\hat{a}) \ ,
\end{aligned} \tag{VII.21}$$

where the last equality follows from the fact that $\{u_s^*\}_{s\in\mathcal{S}}$ fulfills the (IR_g) with equality. Thus, contract $\{u_s^N\}_{s\in\mathcal{S}}$ is feasible in the sense that all constraints of program (MG) are met. It remains to show that the principal's costs are reduced. Since $h(\cdot)$ is strictly convex, the principal's objective function is strictly convex in α, with a minimum at $\alpha = 0$. Hence, the principal's objective function is strictly increasing in α for $\alpha \in (0, 1]$. Since $\{u_s^N\}_{s\in\mathcal{S}} \equiv \{u_s^*\}_{s\in\mathcal{S}}$ for $\alpha = 1$, for $\alpha = \hat{\alpha}$ we have

$$\sum_{s=1}^{S} \gamma_s(\hat{a}) h(u_s^*) > \sum_{s=1}^{S} \gamma_s(\hat{a}) h(u_s^N),$$

which establishes the desired result. □

PROOF OF PROPOSITION 19

For the agent's intrinsic utility function being sufficiently linear, the principal's costs are approximately given by a second-order Taylor polynomial about $r = 1$, thus

$$C(\boldsymbol{u}|r) \approx \sum_{s\in\mathcal{S}} \gamma_s(\hat{a}) u_s + \Omega(\boldsymbol{u}|r) \ ,$$

where

$$\Omega(\boldsymbol{u}|r) \equiv \sum_{s\in\mathcal{S}} \gamma_s(\hat{a}) \Big[(u_s \ln u_s)(r-1) + (1/2) u_s (\ln u_s)^2 (r-1)^2 \Big].$$

Relabeling signals such that the wage profile is increasing allows us to express the incentive scheme in terms of increases in intrinsic utility. The agent's binding participation constraint implies that

$$u_1 = \bar{u} + c(\hat{a}) - \sum_{s=2}^{S} b_s \left\{ \sum_{\tau=s}^{S} \gamma_\tau(\hat{a}) - (\lambda - 1) \left[\sum_{\tau=s}^{S} \gamma_\tau(\hat{a}) \right] \left[\sum_{t=1}^{s-1} \gamma_t(\hat{a}) \right] \right\} \equiv u_1(\boldsymbol{b})$$

and $u_s = u_1(\boldsymbol{b}) + \sum_{t=2}^{s} b_t \equiv u_s(\boldsymbol{b})$ for all $s = 2, \ldots, S$. Inserting the binding participation constraint into the above cost function and replacing $\Omega(\boldsymbol{u}|r)$ equivalently by $\tilde{\Omega}(\boldsymbol{b}|r) \equiv \Omega(u_1(\boldsymbol{b}), \ldots, u_S(\boldsymbol{b})|r)$ yields

$$C(\boldsymbol{b}|r) \approx \bar{u} + c(\hat{a}) + (\lambda - 1) \sum_{s=2}^{S} b_s \left[\sum_{\tau=s}^{S} \gamma_\tau(\hat{a}) \right] \left[\sum_{t=1}^{s-1} \gamma_t(\hat{a}) \right] + \tilde{\Omega}(\boldsymbol{b}|r).$$

Hence, for a given increasing wage profile the principal's cost minimization problem is:

Program ME:

$$\min_{\boldsymbol{b}\in\mathbb{R}_+^{S-1}} \boldsymbol{b}'\boldsymbol{\rho}(\hat{\boldsymbol{\gamma}},\lambda,\hat{a}) + \tilde{\Omega}(\boldsymbol{b}|r)$$

$$\text{subject to} \quad \boldsymbol{b}'\boldsymbol{\beta}(\hat{\boldsymbol{\gamma}},\lambda,\hat{a}) = c'(\hat{a}) \qquad \text{(IC}'\text{)}$$

If r is sufficiently close to 1, then the incentive scheme that solves program ML also solves program ME. Note that generically program ME is solved only by bonus schemes. Put differently, even if there are multiple optimal contracts for program ML, all these contracts are generically simple bonus contracts. Thus, from Proposition 15 it follows that generically for r close to 1 the optimal incentive scheme entails a minimum of wage differentiation. Note that for $\lambda = 1$ the principal's problem is to minimize $\tilde{\Omega}(\boldsymbol{b}|r)$ even for r sufficiently close to 1. □

PROOF OF PROPOSITION 20

First consider $b \geq 0$. We divide the analysis for $b \geq 0$ into three subcases.

Case 1 ($a_0 < 0$): For the effort level $\hat{a}$ to be chosen by the agent, this effort level has to satisfy the following incentive compatibility constraint:

$$\hat{a} \in \arg\max_{a\in[0,1]} \; u + \gamma(a)b - \gamma(a)(1-\gamma(a))b(\lambda-1) - \frac{k}{2}a^2 \qquad \text{(IC)}$$

For $\hat{a}$ to be a zero of $dE\left[U(a)\right]/da$, the bonus has to be chosen according to

$$b^*(\hat{a}) = \frac{k\hat{a}}{(\gamma^H - \gamma^L)\left[2-\lambda+2\gamma(\hat{a})(\lambda-1)\right]}.$$

For $a > a_0$, $b^*(a)$ is a strictly increasing and strictly concave function with $b^*(0) = 0$. Hence, each $\hat{a} \in [0,1]$ can be made a zero of $dE\left[U(a)\right]/da$ with a non-negative bonus. By choosing the bonus according to $b^*(\hat{a})$, $\hat{a}$ satisfies, by construction, the first-order condition. Inserting $b^*(\hat{a})$ into the $d^2E\left[U(a)\right]/da^2$ shows that expected utility is strictly concave function if $a_0 < 0$. Hence, with the bonus set equal to $b^*(\hat{a})$, effort level $\hat{a}$ satisfies the second-order condition for optimality and therefore is incentive compatible.

Case 2 ($a_0 = 0$): Just like in the case where $a_0 < 0$, each effort level $a \in [0, 1]$ turns out to be implementable with a non-negative bonus. To see this, consider bonus

$$b_0 = \frac{k}{2(\gamma^H - \gamma^L)^2(\lambda - 1)}.$$

For $b \leq b_0$, $dE\left[U(a)\right]/da < 0$ for each $a > 0$, that is, lowering effort increases expected utility. Hence, the agent wants to choose an effort level as low as possible and therefore exerts no effort at all. If, on the other hand, $b > b_0$, then $dE\left[U(a)\right]/da > 0$. Now, increasing effort increases expected utility, and the agent wants to choose effort as high as possible. For $b = b_0$, expected utility is constant over all $a \in [0, 1]$, that is, as long as his participation constraint is satisfied, the agent is indifferent which effort level to choose. As a tie-breaking rule we assume that, if indifferent between several effort levels, the agent chooses the effort level that the principal prefers.

Case 3 ($a_0 > 0$): If $a_0 > 0$, the agent either chooses $a = 0$ or $a = 1$. To see this, again consider bonus b_0. For $b \leq b_0$, $dE\left[U(a)\right]/da < 0$ for each $a > 0$. Hence, the agent wants to exert as little effort as possible and chooses $a = 0$. If, on the other hand, $b > b_0$, then $d^2E\left[U(a)\right]/da^2 > 0$, that is, expected utility is a strictly convex function of effort. In order to maximize expected utility, the agent will choose either $a = 0$ or $a = 1$ depending on whether $E\left[U(0)\right]$ exceeds $E\left[U(1)\right]$ or not.

Negative Bonus: $b < 0$

Let $b^- < 0$ denote the monetary punishment that the agent receives if the good signal is observed. With a negative bonus, the agent's expected utility is

$$\begin{aligned} E\left[U(a)\right] = u + \gamma(a)b^- + \gamma(a)(1 - \gamma(a))\lambda b^- \\ + (1 - \gamma(a))\gamma(a)(-b^-) - \frac{k}{2}a^2. \quad \text{(VII.22)} \end{aligned}$$

The first derivative with respect to effort,

$$\frac{dE\left[U(a)\right]}{da} = \underbrace{(\gamma^H - \gamma^L)b^- \left[\lambda - 2\gamma(a)(\lambda - 1)\right]}_{MB^-(a)} - \underbrace{ka}_{MC(a)},$$

reveals that $MB^-(a)$ is a positively sloped function, which is steeper the harsher the punishment is, that is, the more negative b^- is. It is worthwhile to point out that if bonus and punishment are equal in absolute value, $|b^-| = b$, then also the slopes of $MB^-(a)$ and $MB(a)$ are identical. The intercept of $MB^-(a)$ with the horizontal axis, a_0^- again is completely determined by the model parameters:

$$a_0^- = \frac{\lambda - 2\gamma^L(\lambda - 1)}{2(\gamma^H - \gamma^L)(\lambda - 1)}.$$

Note that $a_0^- > 0$ for $\gamma^L \leq 1/2$. For $\gamma^L > 1/2$ we have $a_0^- < 0$ if and only if $\lambda > 2\gamma^L/(2\gamma^L - 1)$. Proceeding in exactly the same way as in the case of a non-negative bonus yields a familiar results: effort level $\hat{a} \in [0, 1]$ is implementable with a strictly negative bonus if and only if $a_0^- \leq 0$. Finally, note that $a_0 < a_0^-$. Hence a negative bonus does not improve the scope for implementation. □

PROOF OF PROPOSITION 21

Throughout the analysis we restricted attention to non-negative bonus payment. It remains to be shown that the principal cannot benefit from offering a negative bonus payment: implementing action $\hat{a}$ with a negative bonus is at least as costly as implementing action $\hat{a}$ with a positive bonus. In what follows, we make use of notation introduced in the paper as well as in the proof of Proposition 20. Let $a_0(p)$, $a_0^-(p)$, $b^*(\hat{a}; p)$, and $u^*(\hat{a}; p)$ denote the expressions obtained from a_0, a_0^-, $b^*(\hat{a})$, and $u^*(\hat{a})$, respectively, by replacing $\gamma(\hat{a})$, γ^L, and γ^H with $\gamma(\hat{a}; p)$, $\gamma^L(p)$, and $\gamma^H(p)$. From the proof of Proposition 19 we know that (i) action $\hat{a}$ is implementable with a non-negative bonus (negative bonus) if and only if $a_0(p) \leq 0$ ($a_0^-(p) \leq 0$), (ii) $a_0^-(p) \leq 0$ implies $a_0(p) < 0$. We will show that, for a given value of p, if $\hat{a}$ is implementable with a negative bonus then it is less costly to implement $\hat{a}$ with a non-negative bonus.

Consider first the case where $a_0^-(p) < 0$. The negative bonus payment satisfying incentive compatibility is given by

$$b^-(\hat{a};p) = \frac{k\hat{a}}{(\gamma^H(p) - \gamma^L(p))\left[\lambda - 2\gamma(\hat{a};p)(\lambda-1)\right]}.$$

It is easy to verify that the required punishment to implement $\hat{a}$ is larger in absolute value than than the respective non-negative bonus which is needed to implement $\hat{a}$, that is, $b^*(\hat{a};p) < |b^-(\hat{a};p)|$ for all $\hat{a} \in (0,1)$ and all $p \in [0,1)$. When punishing the agent with a negative bonus $b^-(\hat{a};p)$, $u^-(\hat{a};p)$ will be chosen to satisfy the corresponding participation constraint with equality, that is,

$$u^-(\hat{a};p) = \bar{u} + \frac{k}{2}\hat{a}^2 - \gamma(\hat{a};p)b^-(\hat{a};p)\left[\lambda - \gamma(\hat{a},p)(\lambda-1)\right].$$

Remember that, if $\hat{a}$ is implemented with a non-negative bonus, we have

$$u^*(\hat{a};p) = \bar{u} + \frac{k}{2}\hat{a}^2 - \gamma(\hat{a};p)b^*(\hat{a};p)\left[2 - \lambda + \gamma(\hat{a};p)(\lambda-1)\right].$$

It follows immediately that the minimum cost of implementing $\hat{a}$ with a non-negative bonus is lower than the minimum implementation cost with a strictly negative bonus:

$$\begin{aligned}
C^-(\hat{a};p) &= u^-(\hat{a};p) + \gamma(\hat{a};p)b^-(\hat{a};p) \\
&= \bar{u} + \frac{k}{2}\hat{a}^2 - \gamma(\hat{a};p)b^-(\hat{a};p)\left[\lambda - \gamma(\hat{a};p)(\lambda-1) - 1\right] \\
&> \bar{u} + \frac{k}{2}\hat{a}^2 + \gamma(\hat{a};p)b^*(\hat{a};p)\left[\lambda - \gamma(\hat{a};p)(\lambda-1) - 1\right] \\
&= \bar{u} + \frac{k}{2}\hat{a}^2 - \gamma(\hat{a};p)b^*(\hat{a};p)\left[1 - \lambda + \gamma(\hat{a};p)(\lambda-1)\right] \\
&= \bar{u} + \frac{k}{2}\hat{a}^2 - \gamma(\hat{a};p)b^*(\hat{a};p)\left[2 - \lambda + \gamma(\hat{a};p)(\lambda-1)\right] \\
&\quad + \gamma(\hat{a};p)b^*(\hat{a};p) \\
&= u^*(\hat{a};p) + \gamma(\hat{a};p)b^*(\hat{a};p) \\
&= C(\hat{a};p).
\end{aligned}$$

The same line of argument holds when $a_0^- = 0$: the bonus which satisfies the (IC) is

$$b_0^-(\hat{a};p) = -\frac{k}{2(\gamma^H(p) - \gamma^L(p))^2(\lambda-1)},$$

and so $b^*(\hat{a};p) < |b_0^-(\hat{a};p)|$ for all $\hat{a} \in (0,1)$ and all $p \in [0,1)$. □

PROOF OF COROLLARY 1

Let $p \in (0,1)$. With $\hat{\boldsymbol{\zeta}}$ being a convex combination of $\hat{\boldsymbol{\gamma}}$ and $\mathbf{1}$ we have $(\zeta^H, \zeta^L) = p(1,1)+(1-p)(\gamma^H, \gamma^L) = (\gamma^H + p(1-\gamma^H), \gamma^L + p(1-\gamma^L))$. The desired result follows immediately from Proposition 16: Consider $\lambda > 2$. Implementation problems are less likely to be encountered under $\hat{\boldsymbol{\zeta}}$ than under $\hat{\boldsymbol{\gamma}}$. Moreover, if implementation problems are not an issue under both performance measures, then implementation of a certain action is less costly under $\hat{\boldsymbol{\zeta}}$ than under $\hat{\boldsymbol{\gamma}}$. For $\lambda = 2$ implementation problems do not arise and implementation costs are identical under both performance measures. Last, if $\lambda < 2$, implementation problems are not an issue under either performance measure, but the cost of implementation is strictly lower under $\hat{\boldsymbol{\gamma}}$ than under $\hat{\boldsymbol{\zeta}}$. □

VALIDITY OF THE FIRST-ORDER APPROACH

L E M M A 10: Suppose A3-A5 hold, then the incentive constraint in the principal's cost minimization problem can be represented as $E[U'(\hat{a})] = 0$.

PROOF: Consider a contract $(u_1, \{b_s\}_{s=2}^S)$ with $b_s \geq 0$ for $s = 2, \ldots, S$. In what follows, we write β_s instead of $\beta_s(\hat{\boldsymbol{\gamma}}, \lambda, \hat{a})$ to cut back on notation. The proof proceeds in two steps. First, we show that for a given contract with the property $b_s > 0$ only if $\beta_s > 0$, all actions that satisfy the first-order condition of the agent's utility maximization problem characterize a local maximum of his utility function. Since the utility function is continuous and all extreme points are local maxima, if there exists some action that fulfills the first-order condition, this action corresponds to the unique maximum. In the second step we show that under the optimal contract we cannot have $b_s > 0$ if $\beta_s \leq 0$.

Step 1: The second derivative of the agent's utility with respect to a is

$$E[U''(a)] = -2(\lambda - 1) \sum_{s=2}^{S} b_s \sigma_s - c''(a) \; , \tag{VII.23}$$

where $\sigma_s := (\sum_{i=1}^{s-1} \gamma_i^H - \gamma_i^L)(\sum_{i=s}^{S} \gamma_i^H - \gamma_i^L) < 0$. Suppose action $\hat{a}$ satisfies the first-order condition. Formally

$$\sum_{s=2}^{S} b_s \beta_s = c'(\hat{a}) \iff \sum_{s=2}^{S} b_s \frac{\beta_s}{\hat{a}} = \frac{c'(\hat{a})}{\hat{a}}. \qquad \text{(VII.24)}$$

Action $\hat{a}$ locally maximizes the agent's utility if

$$-2(\lambda - 1) \sum_{s=2}^{S} b_s \sigma_s < c''(\hat{a}). \qquad \text{(VII.25)}$$

Under Assumption A5, we have $c''(\hat{a}) > c(\hat{a})/\hat{a}$. Therefore, if

$$\sum_{s=2}^{S} b_s \big[-2(\lambda - 1)\sigma_s - \beta_s/\hat{a}\big] < 0 \,, \qquad \text{(VII.26)}$$

then (VII.24) implies (VII.25), and each action $\hat{a}$ satisfying the first-order condition of the agent's maximization problem is a local maximum of his expected utility. Inequality (VII.26) obviously is satisfied if each element of the sum is negative. Summand s is negative if and only if

$$\begin{aligned} &- 2(\lambda - 1) \left(\sum_{i=1}^{s-1} (\gamma_i^H - \gamma_i^L) \right) \left(\sum_{i=s}^{S} (\gamma_i^H - \gamma_i^L) \right) \hat{a} \\ &\qquad - \left(\sum_{\tau=s}^{S} (\gamma_\tau^H - \gamma_\tau^L) \right) \left[1 - (\lambda - 1) \left(\sum_{t=1}^{s-1} \gamma_t(\hat{a}) \right) \right] \\ &\qquad\qquad - (\lambda - 1) \left[\sum_{\tau=s}^{S} \gamma_\tau(\hat{a}) \right] \left(\sum_{t=1}^{s-1} (\gamma_t^H - \gamma_t^L) \right) < 0. \end{aligned}$$

Rearranging the above inequality yields

$$\begin{aligned} &\left(\sum_{i=s}^{S} (\gamma_i^H - \gamma_i^L) \right) \left\{ \lambda + 2(\lambda - 1) \left[\hat{a} \sum_{i=1}^{s-1} (\gamma_i^H - \gamma_i^L) - \sum_{i=1}^{s-1} \gamma_i(\hat{a}) \right] \right\} > 0 \\ &\iff \left(\sum_{i=s}^{S} (\gamma_i^H - \gamma_i^L) \right) \left\{ \lambda \left(1 - \sum_{i=1}^{s-1} \gamma_i^L \right) + (2 - \lambda) \sum_{i=1}^{s-1} \gamma_i^L \right\} > 0 \end{aligned} \qquad \text{(VII.27)}$$

The term in curly brackets is positive, since $\lambda \leq 2$ and $\sum_{i=1}^{s-1} \gamma_i^L < 1$. Furthermore, note that $\sum_{i=s}^{S} (\gamma_i^H - \gamma_i^L) > 0$ since $\beta_s > 0$ for all $b_s > 0$. This completes the first step of the proof.

Step 2: Consider a contract with $b_s > 0$ and $\beta_s \leq 0$ for at least one signal $s \in \{2, \ldots, S\}$ that implements $\hat{a} \in (0, 1)$. Then, under this contract, (IC′) is satisfied and there exists at least one signal t with $\beta_t > 0$ and $b_t > 0$. Obviously, the principal can reduce both b_s and b_t without violating (IC′). This reasoning goes through up to the point where (IC′) is satisfied and $b_s = 0$ for all signals s with $\beta_s \leq 0$. From the first step of the proof we know that the resulting contract implements $\hat{a}$ incentive compatibly. Next, we show that reducing any spread, say b_k, always reduces the principal's cost of implementation.

$$C(\boldsymbol{b}) = \sum_{s=1}^{S} \gamma_s(\hat{a}) h \left(u_1(\boldsymbol{b}) + \sum_{t=2}^{s} b_s \right) \text{, where}$$

$$u_1(\boldsymbol{b}) = \bar{u} + c(\hat{a}) - \sum_{t=2}^{S} b_s \left[\sum_{\tau=s}^{S} \gamma_\tau(\hat{a}) - (\lambda - 1) \left(\sum_{\tau=s}^{S} \gamma_\tau(\hat{a}) \right) \left(\sum_{t=1}^{s-1} \gamma_t(\hat{a}) \right) \right].$$

The partial derivative of the cost function with respect to an arbitrary b_k is

$$\frac{\partial C(\boldsymbol{b})}{\partial b_k} = \sum_{s=1}^{k-1} \gamma_s(\hat{a}) h' \left(u_1(\boldsymbol{b}) + \sum_{t=2}^{s} b_s \right) \left[\frac{\partial u_1}{\partial b_k} \right] + \sum_{s=k}^{S} \gamma_s(\hat{a}) h' \left(u_1(\boldsymbol{b}) + \sum_{t=2}^{s} b_s \right) \left[\frac{\partial u_1}{\partial b_k} + 1 \right].$$

Rearranging yields

$$\frac{\partial C(\boldsymbol{b})}{\partial b_k} = \sum_{s=1}^{k-1} \gamma_s(\hat{a}) h'(u_s) \cdot \underbrace{\left[(\lambda - 1) \left(\sum_{\tau=k}^{S} \gamma_\tau(\hat{a}) \right) \left(\sum_{t=1}^{k-1} \gamma_t(\hat{a}) \right) - \sum_{\tau=k}^{S} \gamma_\tau(\hat{a}) \right]}_{<0} + \sum_{s=k}^{S} \gamma_s(\hat{a}) h'(u_s) \underbrace{\left[(\lambda - 1) \left(\sum_{\tau=k}^{S} \gamma_\tau(\hat{a}) \right) \left(\sum_{t=1}^{k-1} \gamma_t(\hat{a}) \right) - \sum_{\tau=k}^{S} \gamma_\tau(\hat{a}) + 1 \right]}_{>0}.$$

Note $u_s \leq u_{s+1}$ which implies that $h'(u_s) \leq h'(u_{s+1})$. Thus, the following inequality holds

$$\frac{\partial C(\boldsymbol{b})}{\partial b_k} \geq \sum_{s=1}^{k-1} \gamma_s(\hat{a}) h'(u_k) \cdot \left[(\lambda - 1) \left(\sum_{\tau=k}^{S} \gamma_\tau(\hat{a}) \right) \left(\sum_{t=1}^{k-1} \gamma_t(\hat{a}) \right) - \sum_{\tau=k}^{S} \gamma_\tau(\hat{a}) \right]$$
$$+ \sum_{s=k}^{S} \gamma_s(\hat{a}) h'(u_k) \left[(\lambda - 1) \left(\sum_{\tau=k}^{S} \gamma_\tau(\hat{a}) \right) \left(\sum_{t=1}^{k-1} \gamma_t(\hat{a}) \right) - \sum_{\tau=k}^{S} \gamma_\tau(\hat{a}) + 1 \right].$$

The above inequality can be rewritten as follows

$$\frac{\partial C(\boldsymbol{b})}{\partial b_k} \geq h'(u_k) \left[(\lambda - 1) \left(\sum_{\tau=k}^{S} \gamma_\tau(\hat{a}) \right) \left(\sum_{t=1}^{k-1} \gamma_t(\hat{a}) \right) \right] > 0.$$

Since reducing any bonus lowers the principal's cost of implementation, it cannot be optimal to set $b_s > 0$ for $\beta_s \leq 0$. This completes the second step of the proof. In combination with step 1, this establishes the desired result. $\square$

VII.5. APPENDIX TO CHAPTER VI

It is useful to know the following properties of a CES function, where we take the specification of (VI.1):

(i) $\frac{dI}{dI_t}$ is positive and homogenous of degree 0.
(ii) $\frac{d^2I}{dI_1 dI_2}$ is positive for $\phi < 1$ and zero for $\phi = 1$.
(iii) $\frac{d^2I}{dI_t^2}$ is negative for $\phi < 1$ and zero for $\phi = 1$.
(iv) For $\phi < 1$ and $I_{-t} > 0$, $\lim_{I_t \to 0} \frac{dI}{dI_t} = \infty$.

PROOF OF LEMMA 9

The first-order conditions of (VI.5) are:

$$\frac{d\pi}{dI_2^H} = w \left. \frac{dm(\cdot)}{dI} \right|_{\alpha} [\cdot]_{\alpha}^{1/\phi-1} (1-\gamma)((1+\beta)\, I_2^H\big|_{\alpha})^{\phi-1}(1+\beta) - \frac{1}{1+r} = 0;$$

$$\begin{aligned} \frac{d\mathbb{E}_\alpha[\pi]}{d\hat{I}_1} &= \frac{1}{2} w \left. \frac{dm(\cdot)}{dI} \right|_{\alpha=\beta} [\cdot]_{\alpha=\beta}^{1/\phi-1} \gamma \left((1+\beta)\hat{I}_1 + (1-\beta)\check{I}_1\right)^{\phi-1} (1+\beta) \\ &+ \frac{1}{2} w \left. \frac{dm(\cdot)}{dI} \right|_{\alpha=-\beta} [\cdot]_{\alpha=-\beta}^{1/\phi-1} \gamma \left((1-\beta)\hat{I}_1 + (1+\beta)\check{I}_1\right)^{\phi-1} (1-\beta) \\ &- 1 = 0; \end{aligned}$$

$$\begin{aligned} \frac{d\mathbb{E}_\alpha[\pi]}{d\check{I}_1} &= \frac{1}{2} w \left. \frac{dm(\cdot)}{dI} \right|_{\alpha=\beta} [\cdot]_{\alpha=\beta}^{1/\phi-1} \gamma \left((1+\beta)\hat{I}_1 + (1-\beta)\check{I}_1\right)^{\phi-1} (1-\beta) \\ &+ \frac{1}{2} w \left. \frac{dm(\cdot)}{dI} \right|_{\alpha=-\beta} [\cdot]_{\alpha=-\beta}^{1/\phi-1} \gamma \left((1-\beta)\hat{I}_1 + (1+\beta)\check{I}_1\right)^{\phi-1} (1+\beta) \\ &- 1 = 0. \end{aligned}$$

$[\cdot]_\alpha$ is the square bracket of (VI.4) evaluated at α. For the first-order conditions of $\hat{I}_1$ and $\check{I}_1$ we have used the Envelope theorem. Note, we cannot be sure that in the optimum the first-order conditions must be satisfied.

Due to the assumptions on $m(\cdot)$ it cannot be optimal not to invest at all. Additionally, due to the concavity of $m(\cdot)$ an optimum exists.

Part 1: It holds that $\hat{I}_1 = \check{I}_1$.

The case $\phi < 1$. Property (iv) of the CES function implies that it is optimal to invest a positive amount in both periods. Which proves the last part of Lemma 9 and implies that the first-order condition of I_2^H must be fulfilled in the optimum. Since $\frac{d^2I}{dI_t^2} < 0$ for $\phi < 1$ and $d^2m(\cdot)/dI^2 < 0$

the optimal I_2^H is unique. Moreover, since $\frac{dI}{dI_t}$ is homogenous of degree 0 and $d^2m(\cdot)/dI^2 < 0$ in the optimum $\frac{I_2^H}{I_1^{\text{effective}}}$ is decreasing in $I_1^{\text{effective}}$.

Suppose that $\hat{I}_1 < \check{I}_1$. Then $\frac{d\mathbb{E}_\alpha[\pi]}{d\hat{I}_1} \leq \frac{d\mathbb{E}_\alpha[\pi]}{d\check{I}_1}$. Hence,

$$\left.\frac{dm(\cdot)}{dI}\right|_{\alpha=\beta} [\cdot]_{\alpha=\beta}^{1/\phi-1} \left((1+\beta)\hat{I}_1 + (1-\beta)\check{I}_1\right)^{\phi-1}$$
$$\leq \left.\frac{dm(\cdot)}{dI}\right|_{\alpha=-\beta} [\cdot]_{\alpha=-\beta}^{1/\phi-1} \left((1-\beta)\hat{I}_1 + (1+\beta)\check{I}_1\right)^{\phi-1}.$$

From before we know that the first-order condition of I_2^H must be fulfilled in the optimum. Inserting it into the previous inequality yields

$$\left(\frac{I_2^H\big|_{\alpha=\beta}}{I_1^{\text{effective}}\big|_{\alpha=\beta}}\right)^{1-\phi} \leq \left(\frac{I_2^H\big|_{\alpha=-\beta}}{I_1^{\text{effective}}\big|_{\alpha=-\beta}}\right)^{1-\phi}. \tag{VII.28}$$

Due to $\hat{I}_1 < \check{I}_1$ we have $I_1^{\text{effective}}\big|_{\alpha=\beta} < I_1^{\text{effective}}\big|_{\alpha=-\beta}$. Since in the optimum $\frac{I_2^H}{I_1^{\text{effective}}}$ is decreasing in $I_1^{\text{effective}}$, see before, (VII.28) cannot be fulfilled. Also $\hat{I}_1 > \check{I}_1$ yields a contradiction. Hence, $\hat{I}_1 = \check{I}_1$.

The case $\phi = 1$. With $\phi = 1$ there are either corner solutions in which it is optimal to invest in only one period, or there is an indifference. In the latter case it is weakly optimal to choose $\hat{I}_1 = \check{I}_1$. In the former case, it is either optimal (i) not to invest in period 1, or (ii) it is optimal not to invest in period 2. In case (i) $\hat{I}_1 = \check{I}_1 = 0$. In case (ii) we must have $\hat{I}_1, \check{I}_1 > 0$ which implies $\frac{d\mathbb{E}_\alpha[\pi]}{d\hat{I}_1} = \frac{d\mathbb{E}_\alpha[\pi]}{d\check{I}_1}$. So

$$\frac{1}{2}\left.\frac{dm(\cdot)}{dI}\right|_{\alpha=\beta}(1+\beta) + \frac{1}{2}\left.\frac{dm(\cdot)}{dI}\right|_{\alpha=-\beta}(1-\beta)$$
$$= \frac{1}{2}\left.\frac{dm(\cdot)}{dI}\right|_{\alpha=\beta}(1-\beta) + \frac{1}{2}\left.\frac{dm(\cdot)}{dI}\right|_{\alpha=-\beta}(1+\beta)$$

which simplifies to

$$\left.\frac{dm(\cdot)}{dI}\right|_{\alpha=\beta} = \left.\frac{dm(\cdot)}{dI}\right|_{\alpha=-\beta}.$$

Since $I_2^H = 0$ this requires $\hat{I}_1 = \check{I}_1$.

Part 2: It holds that $I_2^H\big|_{\alpha=-\beta} = I_2^H\big|_{\alpha=\beta}$.

The case $\phi < 1$. From Part 1 we know that $\hat{I}_1 = \check{I}_1 > 0$. This directly implies, see the first-order condition of I_2^H, that in the optimum, although the type of the second period investment depends on α, the height of the second period investments I_2^H is independent of α.

The case $\phi = 1$. When it is optimal to invest only in the second period the same arguments as with $\phi < 1$ apply. When it is optimal only to invest in the first period we have $I_2^H = 0$ for both, $\alpha = \beta$ and $\alpha = -\beta$. When there is a case of indifference it is weakly optimal to choose $\hat{I}_1 = \check{I}_1$, see Part 1. Then the the first-order condition of I_2^H implies that I_2^H is independent of α. □

LITERATURE

Abeler, J., A. Falk, L. Götte, and D. Huffman (2009). Reference Points and Effort Provision. *IZA Discussion Paper* No. 3939, http://ftp.iza.org/dp3939.pdf.

Acemoglu, D. (2002). Directed Technical Change. *Review of Economic Studies* 69, 781-809.

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt (2005). Competition and Innovation: An Inverted-U Relationship. *Quarterly Journal of Economics* 120, 701-728.

Aghion, P., R. Blundell, R. Griffith, P. Howitt, and S. Prantl (2009). The Effects of Entry on Incumbent Innovation and Productivity. *Review of Economics and Statistics* 91, 20-32.

Aghion, P., R. Burgess, S. Redding, and F. Zilibotti (2005). Entry Liberalization and Inequality in Industrial Performance. *Journal of the European Economic Association* 3, 291-302.

Aghion, P. and R. Griffith (2005). *Competition and Growth. Reconciling Theory and Evidence*. Cambridge: MIT Press.

Aghion, P., C. Harris, P. Howitt, and J. Vickers (2001). Competition, Imitation and Growth with Step-by-Step Innovation. *Review of Economic Studies* 68, 467-492.

Aghion, P. and P. Howitt (2006). Appropriate Growth Policy: A Unifying Framework. *Journal of the European Economic Association* 4, 269-314.

Aitken, B.J. and A.E. Harrison (1999). Do Domestic Firms Benefit from Direct Foreign Investment? Evidence from Venezuela. *American Economic Review* 89, 605-618.

Arrow, K. (1962). Economic Welfare and the Allocation of Resources for Inventions. In *The Rate and Direction of Inventive Activity,* ed. R. Nelson. NJ: Princeton University Press, 609-624.

Astebo, T. (2003). The Return to Independent Invention: Evidence of Unrealistic Optimism, Risk Seeking or Skewness Loving? *Economic Journal* 113, 226-239.

Auriol, E. and R. Renault (2008). Status and incentives. *RAND Journal of Economics* 39, 305-326.

Barberis, N., M. Huang and T. Santos (2001). Prospect Theory and Asset Prices. *Quarterly Journal of Economics* 116, 1-53.

Baumol, W. (2004). Education for Innovation: Entrepreneurial Breakthroughs vs. Corporate Incremental Improvements. *NBER Working Paper* 10578.

Beath, J., Y. Katsoulacos, and D. Ulph (1989a). The Game-Theoretic Analysis of Innovation: A Survey. *Bulletin of Economic Research* 41, 163-184.

Beath, J., Y. Katsoulacos, and D. Ulph (1989b). Strategic R&D Policy. *Economic Journal* 99, 74-83.

Becker, G.S. and G.J. Stigler (1974). Law Enforcement, Malfeasance, and Compensation of Enforcers. *Journal of Legal Studies* 3, 1-18.

Bell, D.E. (1985). Disappointment in Decision Making under Uncertainty. *Operations Research* 33, 1-27.

Belleflamme, P. and C. Vergari (2006). Incentives to Innovate in Oligopolies. *CORE Discussion Paper* 2006/14.

Bertsekas, D. (1974). Necessary and Sufficient Conditions for Existence of an Optimal Portfolio. *Journal of Economic Theory* 8, 235-247.

Besley, T. and M. Ghatak (2008). Status Incentives. *American Economic Review* 98, 206-211.

Bessen, J.E. and M.J. Meurer (2008). Do Patents Perform Like Property? *Boston University School of Law Working Paper* 08-08.

Bianchi, M. and M. Henrekson (2005). Is Neoclassical Economics still Entrepreneurless? *Kyklos* 58, 353-377.

Blackwell, D. (1951). Comparison of Experiments. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, J. Neyman, editor, University of California Press, Berkeley.

Blackwell, D. (1953). Equivalent Comparison of Experiments. *Annals of Mathematics and Statistics* 24, 265-272.

Blume, A. (2003). Bertrand with fudge. *Economics Letters* 78, 167-168.

Boldrin, M. and D.K. Levine (2009). A Model of Discovery. *American Economic Review: Papers & Proceedings* 99, 337-342.

Bolton, P. and M. Dewatripont (2005). *Contract Theory*. Boston: MIT Press.

Borch, K. (1962). Equilibrium in a Reinsurance Market. *Econometrica* 30, 424-444.

Bowman, D., D. Minehart and M. Rabin (1999). Loss-Aversion in a Consumptions-Savings Model. *Journal of Economic Behavior and Organization* 38, 155-178.

Breiter, H.C., I. Aharon, D. Kahneman, A. Dale and P. Shizgal (2001). Functional Imaging of Neural Responses to Expectancy and Experience of Monetary Gains and Losses. *Neuron* 30, 619-639.

Camerer, C.F. and U. Malmendier (2007). Behavioral Economics of Organizations. In *Behavioral Economics and its Applications*, P. Diamond and H. Vartiainen, editors, Princeton University Press, Princeton, 235-281.

Caron, L. and G. Dionne (1997). Insurance Fraud Estimation: More Evidence from the Québec Automobile Insurance Industry. *Assurances* 64, 567-578.

Chiappori, P.-A., B. Jullien, B. Salanié, and F. Salanié (2006). Asymmetric information in insurance: general testable implications. *RAND Journal of Economics* 37, 783-798.

Churchill, G.A., N.M. Ford, and O.C. Walker (1993). *Sales Force Management*, Irwin, Homewood.

Cohen, W. (1995). Empirical Studies of Innovative Activity. In *Handbook of the Economics of Innovation and Technological Change,* ed. Paul Stoneman. Oxford: Blackwell Publishers, 182-264.

Cramér, H, (1936). Über eine Eigenschaft der normalen Verteilungsfunktion. *Mathematische Zeitschrift* 41, 405-414.

Cunha, F. and J. Heckman (2007). The Technology of Skill Formation. *American Economic Review: Papers and Proceedings* 97, 31-47.

Cunha, F., J. Heckman, L. Lochner, and D. Masterov (2005). Interpreting the Evidence on Life Cycle Skill Formation. In Erik Hanushek and F. Welch (Ed.), *Handbook of the Economics of Education* 1, North Holland.

Daido, K. and H. Itoh (2007). The Pygmalion and Galatea Effects: An Agency Model with Reference-Dependent Preferences and Applications to Self-Fulfilling Prophecy. *Working Paper* Hitotsubashi University.

Dantzig, G.B. (1957). Discrete Variable Extremum Problems. *Operations Research* 5, 266-277.

Delbono, F. and V. Denicolo (1991). Incentives to Innovate in a Cournot Oligopoly. *Quarterly Journal of Economics* 106, 951-961.

DellaVigna, S. and U. Malmendier (2004). Contract Design and Self-Control: Theory and Evidence. *Quarterly Journal of Economics* 119, 353-402.

De Meza, D. and C. Southey (1996). The Borrower's Curse: Optimism, Finance and Entrepreneurship. *Economic Journal* 106(435), 375-386.

De Meza, D. and D.C. Webb (2007). Incentive Design Under Loss Aversion. *Journal of the European Economic Association* 5, 66-92.

Demougin, D. and C. Fluet (1998). Mechanism Sufficient Statistic in the Risk-Neutral Agency Problem. *Journal of Institutional and Theoretical Economics* 154, 622-639.

Denicolo, V. (2001). Growth with non-drastic innovations and the persistence of leadership. *European Economic Review* 45, 1399-1413.

Denicolo, V. and P. Zanchettin (2008). Competition, Market Selection and Growth. *Working Paper*.

DiMasi, J.A. (2001). Risks in new drug development: Approval success rates for investigational drugs. *Clinical Pharmacology and Therapeutics* 69, 297-307.

Dionne, G. and R. Gagné (2001). Deductible Contracts Against Fraudulent Claims: Evidence from Automobile Insurance. *Review of Economics and Statistics* 83, 290-301.

Dittmann, I., E. Maug and O. Spalt (2007). Executive Stock Options When Managers Are Loss Averse. *Working Paper* University of Mannheim.

Dixit, A. (1980). The Role of Investment in Entry-Deterrence. *Economic Journal* 90(357), 95-106.

Dow, J. and S.R. Da Costa Werlang (1992). Uncertainty Aversion, Risk Aversion, and the Optimal Choice of Portfolio. *Econometrica* 60, 197-204.

Economist, The (2007). Something new under the sun. A special report on innovation, 13th October 2007, 3-4.

Eichberger, J. and D. Kelsey (2009). Ambiguity. In *The Oxford Handbook of Rational and Social Choice,* ed. P. Anand, P. Pattanaik, and C. Puppe. Oxford: Oxford University Press.

Eliaz, K. and R. Spiegler (2006). Contracting with Diversely Naive Agents. *Review of Economic Studies* 73, 689-714.

Eliaz, K. and R. Spiegler (2008). Consumer Optimism and Price Discrimination. *Theoretical Economics* 3, 459-497.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics* 75, 643-669.

Englmaier, W. and A. Wambach (2006). Optimal Incentive Contracts under Inequity Aversion. *Working Paper* University of Cologne.

Epstein, L.G. and T. Wang (1994). Intertemporal Asset Pricing Under Knightian Uncertainty. *Econometrica* 62, 283-322.

Etro, F. (2004). Innovation by Leaders. *Economic Journal* 114, 281-303.

Fang, H. and G. Moscarini (2005). Morale hazard. *Journal of Monetary Economics* 52, 749-777.

Fehr, E. and K.M. Schmidt (1999). A Theory of Fairness, Competition, and Cooperation. *Quarterly Journal of Economics* 114, 817-868.

Foppert, D. (1994). Waging War Against Fraud. *Best's Review: Property-Causality*, Ed. 94.

Freeman, C.and L. Soete (1997). *The Economics of Industrial Innovation.* Third edition. Cambridge: MIT Press.

Fudenberg, D. and J. Tirole (1986). Dynamic Models of Oligopoly. In Volume 3 of *Fundamentals of Pure and Applied Economics*, ed. Lesourne and Sonnenschein. London: Harwood.

Gallini, N.T. (1992). Patent policy and costly imitation. *RAND Journal of Economics* 23 52-63.

Geroski, P. (1995). What do we know about entry? *International Journal of Industrial Organization* 13, 421-440.

Ghirardato, P. (1994). Agency Theory with Non-Additive Uncertainty. *Mimeo.*

Ghirardato, P., F. Maccheroni, and M. Marinacci (2004). Differentiating Ambiguity and Ambiguity Attitude. *Journal of Economic Theory* 118, 133-173.

Gilbert, R.J. (2006). Looking for Mr. Schumpeter: Where Are We in the Competition-Innovation Debate? In *Innovation Policy and the Economy,* ed. Adam B. Jaffe, Josh Lerner, and Scott Stern. Cambridge: MIT Press, 159-215.

Gilbert, R.J. (2006). Competition and Innovation. *Journal of Industrial Organization Education* 1, 1-23.

Gilbert, R.J. and D.M.G. Newbery (1982). Preemptive Patenting and the Persistence of Monopoly. *American Economic Review* 72, 514-526.

Gilbert, R.J. and D.M.G. Newbery (1984). Uncertain Innovation and the Persistence of Monopoly: Comment. *American Economic Review* 74, 238-242.

Gilbert, R.J. and S.C. Sunshine (1995). Incorporating Dynamic Efficiency Concerns in Merger Analysis: The use of Innovation Markets. *Antitrust Law Journal* 62, 569-601.

Gilboa, I., A.W. Postlewaite, and David Schmeidler (2008). Probability and Uncertainty in Economic Modeling. *Journal of Economic Perspectives* 22, 173-188.

Gilboa, I. and D. Schmeidler (1989). Maxmin Expected Utility with Non-Unique Prior. *Journal of Mathematical Economics* 18, 141-153.

Gjesdal, F. (1982). Information and Incentives: The Agency Information Problem. *Review of Economic Studies* 49, 373-390.

Gollier, C. (2008). Does ambiguity aversion reinforce risk aversion? Applications to Portfolio Choices and Asset Prices. *Mimeo.*

Goolsbee, A. and C. Syverson (2008). How do Incumbents Respond to the Threat of Entry? Evidence from the Major Airlines. *Quarterly Journal of Economics* 123, 1611-1633.

Grossman, S. and O. Hart (1983). An Analysis of the Principal-Agent Problem. *Econometrica* 51, 7-45.

Gul, F. (1991). A Theory of Disappointment Aversion. *Econometrica* 59, 667-686.

Haller, H. (1985). The Principal-Agent Problem with a Satisficing Agent. *Journal of Economic Behavior and Organization* 6, 359-379.

Heidhues, P. and B. Kőszegi (2005). The Impact of Consumer Loss Aversion on Pricing. *CEPR Discussion Paper* No. 4849, http://ssrn.com/abstract=717706.

Heidhues, P. and B. Kőszegi (2008). Competition and Price Variation when Consumers are Loss Averse. *American Economic Review* 98, 1245-1268.

Hermalin, B. (2009). Lecture Notes for Economics. http://faculty.haas.berkeley.edu/hermalin/LectureNotes201b_v5.pdf.

Holmström, B. (1979). Moral Hazard and Observability. *Bell Journal Of Economics* 10, 74-91.

Holmström, B. and P. Milgrom (1987). Aggregation and Linearity in the Provision of Intertemporal Incentives. *Econometrica* 55, 303-328.

Holmström, B. and P. Milgrom (1991). Multitask Principal Agent Analyses: Incentive Contracts, Asset Ownership and Job Design. *Journal of Law, Economics, and Organization* 7 (special issue), 24-52.

Hopenhayn, H.A. (1992). Entry, Exit, and firm Dynamics in Long Run Equilibrium. *Econometrica* 60, 1127-1150.

Hsu, M., M. Bhatt, R. Adolphs, D. Tranel, and C.F. Camerer (2005). Neural Systems Responding to Degrees of Uncertainty in Human Decision-Making. *Science* 310, 1680-1683.

Iantchev, E.P. (2005). Optimal Contracts with Prospect Theory Preferences in the Presence of Moral Hazard. *Working Paper* University of Chicago.

Javorcik, B.S. (2004). Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers Through Backward Linkages. *American Economic Review* 94, 605-627.

Jewitt, I., O. Kadan, and J.M. Swinkels (2008). Moral Hazard with Bounded Payments. *Journal of Economic Theory* 143, 59-82.

Joseph, K. and M.U. Kalwani (1998). The Role of Bonus Pay in Salesforce Compensation Plans. *Industrial Marketing Management* 27, 147-159.

Kahneman, D. and A. Tversky (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47, 263-291.

Karni, E. (2006). Agency Theory with Maxmin Expected Utility Players. *Mimeo.*

Kihlstrom, R.E. and J.J. Laffont (1979). A General Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion. *Journal of Political Economy* 87, 719-748.

Kim, S.K. (1995). Efficiency of an Information System in an Agency Model. *Econometrica* 63, 89-102.

Kim, S.K. (1997). Limited Liability and Bonus Contracts. *Journal of Economics and Management Strategy* 6, 899-913.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005). A Smooth Model of Decision Making Under Ambiguity. *Econometrica* 73, 1849-1892.

Knight, F. H. (1921). *Risk, Uncertainty and Profit.* Boston: Houghton Mifflin.

Kőszegi, B. and M. Rabin (2006). A Model of Reference-Dependent Preferences. *Quarterly Journal of Economics* 121, 1133-1165.

Kőszegi, B. and M. Rabin (2007). Reference-Dependent Risk Preferences. *American Economic Review* 97, 1047-1073.

Lang, M. (2007). Ambiguity Aversion in a Multitask Moral Hazard Framework. *Mimeo.*

Larsen, J.T., A.P. Mc Graw, B.A. Mellers, and J.T. Cacioppo (2004). The Agony of Victory and Thrill of Defeat: Mixed Emotional Reactions to Disappointing Wins and Relieving Losses. *Psychological Science* 15, 325-330.

Lazear, E.P. and K.L. Shaw (2007). Personnel Economics: The Economist's View of Human Resources. *Journal of Economic Perspectives* 21, 91-114.

Lazear, E.P. and P. Oyer (2007). Personnel Economics. *NBER Working Paper 13480*, http:// www.nber.org/papers/w13480.

Lee, T. and L.L. Wilde (1980). Market Structure and Innovation: A Reformulation. *Quarterly Journal of Economics* 94, 429-436.

Lelarge, C. and B. Nefussi (2008). Exposure to Low-Wage Competition, Activity Changes and Quality Upgrading: An Empirical Assessment. *Mimeo.*

Lerner, J. (2009). The Empirical Impact of Intellectual Property Rights on Innovation: Puzzles and Clues. *American Economic Review Papers & Proceedings* 99, 343-348.

Loomes, G. and R. Sugden (1986). Disappointment and Dynamic Consistency in Choice under Uncertainty. *Review of Economic Studies* 53, 271-282.

MacDonald, J.M. (1994). Does Import Competition Force Efficient Production? *Review of Economics and Statistics* 76, 721-727.

MacLeod, B. (2003). Optimal Contracting with Subjective Evaluation. *American Economic Review* 93, 216-240.

Mansfield, E., J. Rapaport, A. Romeo, E. Villani, S. Wagner, and F. Husic (1977). *The production and application of new industrial technology.* New York: Norton.

Mas-Colell, A., M.D. Whinston, and J.R. Greene (1995). *Microeconomic Theory.* Oxford: Oxford University Press.

Mellers, B., A. Schwartz and H. Ritov (1999). Emotion-Based Choice. *Journal of Experimental Psychology: General* 128, 332-345.

Mokyr, J. (2009). Intellectual Property Rights, the Industrial Revolution, and the Beginnings of Modern Economic Growth. *American Economic Review Papers & Proceedings* 99, 349-355.

Moynahan, J.K. (1980). *Designing an Effective Sales Compensation Program.* AMACOM, New York.

Mukerji, S.(2003). Ambiguity Aversion and Cost-Plus Procurement Contracts. *University of Oxford, Department of Economics Discussion Paper* 171.

Oyer, P. (1998). Fiscal Year Ends and Non-Linear Incentive Contracts: The Effect on Business Seasonality. *Quarterly Journal of Economics* 113, 149-188.

Oyer, P. (2000). A Theory of Sales Quotas with Limited Liability and Rent Sharing. *Journal of Labor Economics* 18, 405-426.

Park, E.-S. (1995). Incentive Contracting under Limited Liability. *Journal of Eco-*

nomics and Management Strategy 4, 477-490.

Pavcnik, N. (2002). Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants. *Review of Economic Studies* 69, 245-276.

Prendergast, C. (1999). The Provision of Incentives in Firms. *Journal of Economic Literature* 37, 7-63.

Post, T., M.J. Van den Assem, G. Baltussen and R.H. Thaler (2008). Deal Or No Deal? Decision Making Under Risk in a Large-Payoff Game Show. *American Economic Review* 98, 38-71.

Puelz, R. and A. Snow (1994). Evidence on Adverse Selection: Equilibrium Signalling and Cross-Subsidization in the Insurance Market. *Journal of Political Economy* 102, 236-257.

Rayo, L. and G.S. Becker (2007). Evolutionary Efficiency and Happiness. *Journal of Political Economy* 115, 302-337.

Reinganum, J. (1983). Uncertain Innovation and the Persistence of Monopoly. *American Economic Review* 73, 741-748.

Reinganum, J. (1984). Uncertain Innovation and the Persistence of Monopoly: Reply. *American Economic Review* 74, 243-246.

Rosen, R.J. (1991). Research and Development with Asymmetric Firm Sizes. *RAND Journal of Economics* 22, 411-429.

Rothschild, M. and J.E. Stiglitz (1976). Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. *Quarterly Journal of Economics* 90, 639-649.

Rustichini, A. (2005). Emotion and Reason in Making Decisions. *Science* 310, 1624-1625.

Sacco, D. and A. Schmutzler (2007). The Effects of Competition in Investment Games. *University of Zurich Working Paper* No. 0608.

Salanié, B. (2003). Testing Contract Theory. *CESifo Economic Studies* 49, 461-477.

Savage, L. (1954). *The Foundations of Statistics*. New York: John Wiley.

Scherer, F.M. and D. Ross (1990). *Industrial Market Structure and Economic Performance*. Third edition. Boston: Houghton Mifflin Company.

Schmidt, U. (1999). Moral Hazard and First-Order Risk Aversion. *Journal of Economics* Supplement 8, 167-179.

Schmutzler, A. (2007). The Relationship Between Competition and Innovation - Why is it Such a Mess? *University of Zurich Working Paper* No. 0716.

Scotchmer, S. (2004). *Innovation and Incentives*. Cambridge: MIT Press.

Segal, I. and M.D. Whinston (2007). Antitrust in Innovative Industries. *American Economic Review* 97, 1703-1730.

Spence, A.M. (1977). Entry, Capacity, Investment and Oligopolistic Pricing. *Bell Journal of Economics* 8, 534-544.

Steenburgh, T.J. (2008). Effort or Timing: The Effect of Lump-Sum Bonuses. *Quantitative Marketing and Economics* 6, 235-256.

Strausz, R. (2006). Deterministic Versus Stochastic Mechanisms in Principal-Agent Models. *Journal of Economic Theory* 128, 306-314.

Tirole, J. (1988). *The Theory of Industrial Organization.* Cambridge: MIT Press.

Tversky, A. and D. Kahneman (1991). Loss Aversion in Riskless Choice: A Reference-Dependent Model. *Quarterly Journal of Economics* 106, 1039-1061.

Vickers, J. (2009). Competition Policy and Property Rights. *Discussion Paper* University of Oxford, Department of Economics.

Vives, X. (2008). Innovation and Competitive Pressure. *Journal of Industrial Economics* 61, 419-469. Winter, R.A. (2000). Optimal Insurance under Moral Hazard.

In *Handbook of Insurance.* G. Dionne, editor, Kluwer Academic Publishers, Boston, 155-183.

Yaari, M.E. (1987). The Dual Theory of Choice under Risk. *Econometrica* 55, 95-115.

Yi, S.S. (1995). Uncertain innovation and persistence of monopoly revisited. *Economics Letters* 49, 319-322.