EUI Working Papers

ECO 2012/23
DEPARTMENT OF ECONOMICS

ASYMMETRIC AWARENESS AND MORAL HAZARD

Sarah Auster

EUROPEAN UNIVERSITY INSTITUTE, FLORENCE
DEPARTMENT OF ECONOMICS

Asymmetric Awareness and Moral Hazzard

SARAH AUSTER

EUI Working Paper **ECO** 2012/23

ISSN 1725-6704

Printed in Italy
European University Institute
Badia Fiesolana
I – 50014 San Domenico di Fiesole (FI)
Italy
www.eui.eu
cadmus.eui.eu

Asymmetric Awareness and Moral Hazard

Sarah Auster[*]

European University Institute

First Version: November 2011 (ECO 2011-31),
This Version: October 2012

Abstract

This paper introduces asymmetric awareness into the classical principal-agent model and discusses the optimal contract between a fully aware principal and an unaware agent. The principal enlarges the agent's awareness strategically when proposing the contract. He faces a trade-off between participation and incentives. Leaving the agent unaware allows him to exploit the agent's incomplete understanding of the world. Making the agent aware enables the principal to use the revealed contingencies as signals about the agent's action choice. The optimal contract reveals contingencies that have low probability but are highly informative about the agent's effort.
JEL Classification: D01, D83, D86.
Keywords: Unawareness, Moral Hazard, Incomplete Contracts.

1 Introduction

The standard moral hazard model analyzes the optimal contract between a principal and an agent in the presence of privately observable effort. As in standard economic models, the underlying assumption is that all decision makers are fully aware. That is, both the principal and the agent know every possible outcome realization and its distribution conditional on the agent's effort. However, in reality there are contracting situations where one party has a better understanding of the underlying uncertainties than the other. The question I want to address in this paper is whether the party with superior awareness can use his better understanding of the world strategically in the presence of moral hazard.

[*]Department of Economics, European University Institute, Via dei Roccettini 9, I-50014 San Domenico di Fiesole (FI) - Italy, E-mail: sarah.auster@eui.eu.

To illustrate this, consider the owner of a firm who wants to hire a manager. It is possible that the firm owner is aware of more opportunities and liabilities concerning his firm than the manager. Suppose, for example, that there is the possibility that one of the firm's products has adverse effects on the health of consumers. As a consequence, it is possible that the firm has to recall the product and faces severe legal liabilities. Whether the product's potential health threat becomes public or not is uncertain and depends on the manager's effort. Suppose further that the possibility of this event never crossed the manager's mind. The question is then, under which conditions it is profitable for the firm owner to disclose a possible recall and its legal consequences to the manager when offering the contract. The model shows that there is a trade off between participation and incentives. First, if the firm owner does not reveal the potential threat to the manager, he designs the contract such that the manager receives the minimal payment if the recall is realized. Since the manager does not take into account such an event, the participation constraint is less costly to satisfy. The size of the participation effect depends on the probability of the event. If the potential recall and its legal consequences are highly probable, it is easier to hire the manager without disclosing the possibility of the product's health threat. Second, since the probability of the health threat becoming public depends on the effort of the manager, its realization is a signal about the manager's effort. Disclosing the potential threat makes the incentive constraint less costly to satisfy. The size of the incentive effect depends on how much the manager's effort affects the probability of the event. If the manager can reduce the probability of the potential recall significantly, it is optimal to make his payment contingent on its realization.

This paper proposes a theoretical model which introduces asymmetric awareness in the canonical moral hazard model. The model analyzes the optimal contract between a fully aware principal and an unaware agent. A decision maker is called unaware when there exist contingencies that he does not know, and he does not know that he does not know, and so on *ad infinitum* (Modica and Rustichini, 1994). In the proposed model, the agent is assumed to be unaware of some relevant events, which means there are contingencies that affect the agent's payoff in some states but that have never crossed his mind. Further, the agent is assumed to be unaware of his unawareness, so he believes that his description of the world is correct and complete. This implies that the agent is oblivious to the possibility that the principal is aware of contingencies that he is unaware of. The principal, on the other hand, is assumed to be fully aware. Moreover, the principal knows that the agent is unaware and he knows what the agent is unaware of. When writing the contract the principal can make the agent aware of some or all relevant contingencies.

In line with the standard principal-agent model, the risk-neutral principal proposes a contract to the risk-averse agent. The principal is the owner of a risky project, whose outcome is a function of the realization of a finite number of elementary contingencies of which the agent only knows a subset. These contingencies can be thought of as elementary propositions that can be either true or false. The probability of a contingency to be realized depends on the agent's privately taken action. The agent's effort can be high or low and it is assumed that implementing high effort is always optimal. Since the principal cannot observe the agent's effort, the terms of the contract have to be such that it is in the agent's best interest to exert the level of effort the principal wishes to implement. The compensation scheme is made contingent on the observable and verifiable outcome, rewarding the agent for outcome realizations that are relatively likely under high effort. The agent is assumed to have limited liability, thus transfers have to be non-negative in each state of the world. If the principal leaves the agent unaware of some contingencies, there is a non-empty set of possible outcomes that the agent does not take into account. It is optimal for the principal to construct the contract such that the agent receives the minimal payment whenever an unforeseen outcome is realized.

The main question this paper addresses is whether and under which conditions the principal enlarges the agent's awareness. The rationale for leaving the agent unaware is what I refer to as the participation effect. If the agent is unaware, his beliefs are systematically biased, which is exploited by the principal. The principal pays in expectation less than the agent's reservation utility, because there is positive probability that he pays zero and because the agent does not take this into account. The rationale for making the agent aware is what I refer to as the incentive effect. Since the probability of a contingency to be realized depends on the effort of the agent, including it in the contract allows the principal to use its realization as a signal about the agent's action choice. This implies that the information structure is richer and providing incentives is less costly.

When contemplating the announcement of a contingency the principal faces the following trade-off. The cost of announcement is the payment to the agent in the states where the respective contingency is realized. The cost increases with the probability of the unforeseen contingency given the optimal level of effort. The gain of enhancing the agent's awareness is a richer information structure. The gain increases with the informativeness of the signal. The principal includes contingencies in the contract for which the gains outweigh the losses. Roughly speaking, these are contingencies that are very unlikely but highly informative. The characterization of the trade-off between participation and incentive effect is the key contribution of this paper.

If the agent is unaware after reading the contract, his perception of the world differs from the principal's perception of the world. The question arises as to whether the agent can rationalize the proposed contract given his beliefs or whether he should get suspicious. To answer this question I analyze the principal's optimization problem from the viewpoint of the agent. The solution to this problem coincides with the proposed contract whenever the principal's expected profit evaluated at the agent's beliefs is non-negative and the optimal effort choice is the same for both beliefs. Given these relatively weak conditions, the proposed contract is fully rationalizable for the agent, i.e. the agent has no reason to become suspicious upon reading the contract. The reason for this is that the principal uses the signals within the agent's awareness optimally. Since the agent is unaware of the existence of other relevant contingencies, the proposed contract maximizes the principal's expected payoff given the agent's beliefs.

Next, I allow for competition among principals. In the benchmark model without unawareness, principals engage in a Bertrand competition over the compensation scheme. In equilibrium, principals make zero profits and the whole surplus goes to the agent. I call this equilibrium full awareness equilibrium. Without requiring the contract to be rationalizable for the agent, the full awareness equilibrium is the unique symmetric Nash equilibrium in my environment. The rationale behind this result is that principals engage in a Bertrand competition over the compensation scheme for every level of awareness. Once profits are small enough it is optimal to deviate and reveal another contingency to the agent in order to capture the whole market. If the contract is required to be rationalizable for the agent, this equilibrium still exists. It is unique when the number of competing principals is large enough.

Finally, some generalizations are discussed. Throughout the main part of the analysis I assume that contingencies are independent of each other. Giving up this assumption the optimal compensation scheme varies with the ratio of perceived to objective probability. Further, I assume that output is a discrete one-to-one mapping from contingencies to real numbers. I discuss how results change when more general forms of output functions are considered. Next, the analysis abstracts from the optimal action choice. I show that whenever the principal wishes to implement low effort, it is optimal to not reveal any contingencies, because there is no incentive effect. Further, I discuss the optimal contract when the agent is the residual claimant. There is an additional effect on the participation constraint because the agent's evaluation of the project generally depends on his level of awareness. It is favorable to the principal to disclose negative outcome shocks, because their revelation lowers the agent's outside option. Lastly, a frequently raised concern is whether unawareness is observationally equivalent to full awareness with zero probability beliefs. I discuss in what sense my model can be interpreted as a standard principal-agent model with heterogeneous priors.

Section 2 gives an overview of the related literature. Section 3 introduces the theoretical model. In Section 4 the optimal contract with observable effort is characterized as a benchmark. The main part of the paper, Section 5, is devoted to the analysis of the optimal contract with unobservable effort. I proceed in two steps: First, I characterize the optimal compensation scheme for a given revelation strategy. Second, I analyze the optimal revelation strategy and examine the basic trade-off between participation and incentives. Section 6 introduces competition among principals and Section 7 discusses generalizations of the basic model. Section 8 offers some concluding remarks. All proofs can be found in the appendix.

2 Related Literature

It is not possible to incorporate non-trivial unawareness in the standard state space model. This has been shown in the seminal paper by Dekel et al. (1998). In response, Heifetz et al. (2006), Li (2009), Board and Chung (2011) and Galanis (2011) have proposed generalized state space models that do allow for non-trivial unawareness. My model adopts the generalized state space model introduced by Heifetz et al. (2006). Their unawareness structure consists of a lattice of state spaces, ordered according to their expressive power, where each state space captures a particular horizon of propositions. In a companion work Heifetz et al. (2011) introduce probabilistic beliefs to the model. For ease of exposition, my model foregoes the formal introduction of state spaces, projections among them and events. Appendix A.1 explains how the basic model is built on the unawareness structure proposed by Heifetz et al. (2006/2011).

Filiz-Ozbay (2012) was one of the first to incorporate unawareness into contracting problems. She considers a contracting situation between a fully aware insurer and an unaware insuree. The key difference between my work and her paper is the presence of moral hazard and the assumption on beliefs. In Filiz-Ozbay (2012) there is no hidden action and consequently no incentive effect, which implies that her set up restricts my framework to the case where the agent is the residual claimant and the revelation of new states involves a participation effect only. However, Filiz-Ozbay (2012) allows for a wider range of equilibrium beliefs. She assumes that the agent assigns arbitrary probability beliefs to newly revealed states with the restriction that the principal's payoff evaluated at the agent's beliefs is non-negative and that relative probability beliefs previously held are unchanged. Given this assumption, it is possible that the agent's beliefs deviate stronger from objective probabilities when becoming aware than before. Due to the insurance motive and the wider range of equilibrium beliefs, the effect of revelation on the participation constraint in her environment is ambiguous. Depending on the effect on the participation constraint disclosure can be profitable or not. Also Ozbay (2008) analyzes a setting where the decision maker is unaware of some events and a fully aware announcer strategically mentions contingencies before the decision maker takes an action. Both Filiz-Ozbay (2012) and Ozbay (2008) explore the possibility that the unaware agent is able to reason why the other agent proposed the observed contract.

A second strand of literature analyzes contracting problems with unawareness of actions. In these models agents are aware of all of nature's moves but are unaware of their own action space.

Von Thadden and Zhao (2011 and 2012) propose a moral hazard model with a fully aware principal and an unaware agent. At first glance this set up may seem similar to mine, however the underlying intuition and the results are very distinct. In contrast to my model, the agent in their model understands all relevant contingencies but is unaware of his own action space. Von Thadden and Zhao (2012) assume that if the principal leaves the agent unaware, the agent chooses some default action unconsciously, but assesses his expected utility with respect to such default action correctly. The principal decides whether to make the agent aware of his own action space or whether to leave him unaware. In a standard moral hazard framework this can be interpreted as the decision whether to restrict the agent's action choice to some sub-optimal level ex-ante or whether to leave the action choice to the agent's discretion. Making the agent aware enlarges the agent's action space and consequently relaxes the participation constraint. However, enlarging the agent's action space adds further incentive constraints to the principal's optimization problem. Consequently, the principal faces a trade off between participation and incentives, but the effects are reversed compared to my model. Their main result is, that it is optimal to leave the agent unaware whenever the default action is close enough to the first best effort level. Also Zhao (2008) considers a moral hazard problem with unawareness of actions and default actions. In his set-up both the principal and the agent can be unaware of their action space.

Finally, this work is related to the literature on moral hazard and heterogeneous priors. Santos-Pinto (2008) analyzes a principal-agent model with an agent that holds wrong beliefs about the impact of his effort and calls such biased beliefs self-image. He shows that if positive self-image and effort are complements, the impact of positive self-image is favorable to the principal.[1] If unawareness in my model is interpreted as assigning probability zero to certain outcomes, the resulting distribution does not satisfy the imposed restrictions in Santos-Pinto (2008). Consequently his results do not apply in my framework.[2] Also De la Rosa (2011) analyzes a moral hazard problem with overconfidence.

[1]The agent is said to have a positive self-image, whenever there is first-order stochastic dominance of the agent's perceived distribution over the actual distribution for any action. He defines effort and self-image as complements if first-order stochastic dominance is stronger for high effort than for low effort.

[2]In my set up the agent implicitly has a positive self-image if he is unaware of outcome decreasing contingencies. If I assume that the true distribution conditional on high effort first-order stochastically dominates the true distribution conditional on low effort and that the agent is unaware of negative outcome shocks only, unawareness implies that effort and positive self-image are substitutes instead of complements. This is because first-order stochastic dominance implies that the probability of low outcomes is more likely under low than high effort. Unawareness implicitly implies that the agent assign probability zero to some of these outcomes given high and low effort. Thus, the agent's beliefs deviate stronger for low than high effort.

3 The Model

There are two individuals involved in the contracting problem: A principal and an agent. The principal is risk neutral and the agent is risk averse. The agent receives utility from monetary transfers C and disutility from effort e. I assume that the utility function is separable in money and effort: $U(C,e) = \upsilon(C) - e$, where υ satisfies the Inada conditions. Effort can take two possible values $e \in \{e^L, e^H\}$, where $e^L < e^H$.

The uncertainty of the environment is captured by a finite set of elementary contingencies, denoted by Θ. A contingency $\theta \in \Theta$ is a random variable with realizations 0 and 1. It can be thought of as an elementary proposition that can be either true or false. The probability of $\theta = 1$ depends on the effort of the agent, denoted by e. Throughout the main part of the analysis it will be assumed that, given e, the contingencies in Θ are conditionally independent of each other.

Assumption 1 *The random variables θ and θ' are conditionally independent given e, for any $\theta, \theta' \in \Theta$.*

Awareness Structure: Unlike in the standard moral hazard problem, the agent is unaware of some contingencies. The subset the agent is aware of is denoted by $\Theta_A \subset \Theta$. The principal is aware of the entire set Θ. Further, he knows that the agent is unaware and he knows which contingencies the agent is unaware of. The agent is unaware of his unawareness and is unaware of the principal's superior awareness. When the principal writes the contract he can enlarge the agent's awareness by mentioning contingencies in the contract, denoted by $X \subseteq \Theta \backslash \Theta_A$. The agent updates his awareness and considers henceforth the contingencies in the set $\widehat{\Theta} = \Theta_A \cup X$.

State Spaces: A state of the world in this environment can be thought of as a sequence of 0's and 1's of length $|\Theta|$ that specifies the realization of each $\theta \in \Theta$. Let S denote the collection of these sequences. Since the agent is unaware of some contingencies, he does not perceive the actual state space but a less expressive one. A state in the agent's subjective state space can be thought of as a sequence of 0's and 1's of length $|\Theta_A| < |\Theta|$ that specifies the realization of each $\theta \in \widehat{\Theta}$. Let $\widehat{S}$ denote the collection of these sequences. For example, let $\Theta = \{\theta_1, \theta_2\}$ and $\widehat{\Theta} = \{\theta_1\}$. Objectively there are four states of the world $S = \{(0,0), (0,1), (1,0), (1,1)\}$, but the agent only perceives two $\widehat{S} = \{(0), (1)\}$. The set of contingencies the agent is aware of determines the dimension of his subjective state space. Disclosing a contingency in the contract implies adding another dimension to his subjective state space.

Outcomes: There is a project with stochastic outcome Y, which is observable and verifiable. Outcome is a function of the contingencies in Θ. Since state $s \in \mathcal{S}$ specifies the realization of each contingency $\theta \in \Theta$, we can define outcome realization y directly as a function of the state $s \in \mathcal{S}$:

$$y = f(s), \quad s \in S.$$

Let $\mathcal{Y}$ denote the range of function f. Since the agent does not know the whole set Θ and consequently does not know the objective state space, he cannot know $\mathcal{Y}$. Instead he perceives outcome as a function of the contingencies he is aware of, i.e. a mapping from his subjective state space to $\mathbb{R}$. I assume that the agent's perceived outcome is equivalent to the objective outcome when the contingencies the agent is unaware of are not realized, denoted by $\theta = 0$. Note that $\theta = 0$ can refer to an elementary proposition being true or false.[3] A way to think about this assumption is that there are events that the agent has never observed and that have never crossed his mind. Instead, the agent has some implicit assumptions about the underlying state of the world, but is unaware of these implicit assumptions (Li, 2008). Consequently, he cannot imagine a world in which a proposition implicitly assumed to be true (false) turns out to be false (true). This assumption is prevalent in the literature of unawareness.[4]

The agent's outcome function is

$$y = \widehat{f}(s), \quad s \in \widehat{S}.$$

where $\widehat{f}(s) = f(s, 0, 0, 0, ...), s \in \widehat{S}$.[5] Let $\widehat{\mathcal{Y}}$ denote the range of function $\widehat{y}$.

Assumption 2 $|\mathcal{Y}| = 2^{|\Theta|}$.

A2 imposes that outcome differs across every state of the world, which implies that the agent knows a subset of possible outcomes whenever he is not fully aware, $\widehat{\mathcal{Y}} \subset \mathcal{Y}$. The assumption that f is a one-to-one function is important for tractability of the characterization of the trade-off between participation and incentives.[6] More general outcome functions are discussed in Section 7.2.

Probability Measures: Let $\pi(y|e)$ denote the probability of $y \in \mathcal{Y}$ given effort e and assume $\pi(y|e) > 0, \forall y \in \mathcal{Y}$. The distribution over $\mathcal{Y}$ is known to the principal. The agent is assumed to have correct beliefs over the distribution of contingencies within his awareness. So if the agent is aware of $\theta \in \Theta$, but unaware of $\theta' \in \Theta$, he assesses the probability of contingency θ to be realized

[3] For example, if a decision maker is unaware of the concept of gravity and θ is the elementary statement that there is gravity, then $\theta = 0$ means that the elementary statement is true. Similarly, if a decision maker is unaware of global warming and θ is the elementary statement that there is global warming, then $\theta = 0$ means that the elementary statement is false.

[4] See for example Modica et al. (1998) and Heifetz et al. (2011).

[5] Assume that any sequence $s \in S$ is ordered such that $s = \{s', s''\}$, with $s' \in \widehat{S}$.

[6] A2 implies that disclosing contingencies allows the principal to use a more informative outcome distribution, which implies that the effect of disclosure on incentives is always positive. If this assumption is given up, the incentive effect is ambiguous.

correctly. This implies that he assigns the correct conditional probability to every $y \in \widehat{\mathcal{Y}}$ given that none of the unforeseen contingencies are realized. Under the assumption of independence the probability that $\theta = 0$ for all $\theta \notin \widehat{\Theta}$ is constant across all $y \in \widehat{\mathcal{Y}}$. Let $\Pi(\widehat{\Theta}|e) := \prod_{\theta \notin \widehat{\Theta}} \Pr[\theta = 0|e]$ denote the probability that $y \in \widehat{\mathcal{Y}}$. Then the agent assigns probability

$$\widehat{\pi}(y|e) := \frac{\pi(y|e)}{\Pi(\widehat{\Theta}|e)}, \quad y \in \widehat{\mathcal{Y}},$$

to outcome $y \in \widehat{\mathcal{Y}}$ conditional on effort e.[7]

The Contract: As in the standard principal agent problem, effort is assumed to be non-observable; thus, the principal offers a contract based on the observable and verifiable Y. The distribution of Y depends on the effort of the agent. It is assumed that $E[Y|e^H] > E[Y|e^L]$ and that $E[Y|e^H] - E[Y|e^L]$ is large enough, so that it is always optimal to induce high effort. This allows us to abstract the analysis from the choice of effort. In Section 7.3 the optimal action choice will be discussed. The agent is assumed to have limited liability; thus, the outcome contingent compensation C is non-negative for all $y \in \mathcal{Y}$.

Definition 1 *A contract is a pair* $(\widehat{\Theta}, C)$ *with* $\Theta_A \subseteq \widehat{\Theta} \subseteq \Theta$ *and* $C : \mathcal{Y} \longrightarrow \mathbb{R}_0^+$*.*

Let $(\widehat{\Theta}^*, \widehat{C}^*)$ denote the contract that maximizes the principal's expected payoff. Following Filiz-Ozbay (2012), we can introduce a notion of incompleteness.

Definition 2 *A contract* $(\widehat{\Theta}, C)$ *is incomplete if* $\widehat{\Theta} \neq \Theta$*. Otherwise it is complete.*

Now suppose the contract is incomplete, so that $\mathcal{Y} \backslash \widehat{\mathcal{Y}}$ is non-empty. The principal can construct the contract such that he pays zero to the agent when an unforeseen level of outcome is realized, e.g. by finding a functional form of the compensation scheme on $\mathcal{Y}$ satisfying zero payments for all $y \in \mathcal{Y} \backslash \widehat{\mathcal{Y}}$ or by including a "zero payment otherwise" clause in the contract. I will abstract from the question of how the principal can implement zero payments in the unforeseen states, but analyze a reduced form of this model. Zero payments at $y \in \mathcal{Y} \backslash \widehat{\mathcal{Y}}$ are optimal because any positive payment in these states will leave the agent's expected utility unaffected, but make the principal strictly worse off. It is important to note that zero payments facilitate notation considerably, but that the results hold for any other minimal payment as long as it is low enough.[8] Thus, the optimal contract in this environment can be interpreted as a contract that promises a fixed payment and that rewards the agent with boni for certain outcomes. Whenever a contingency is realized that is not anticipated by the agent, the bonus is not paid.

[7] Generally one can depart from the assumption that the agent assigns correct probabilities to events within his awareness and allow for a wider range of beliefs. This adds effects of heterogeneous beliefs or ambiguity to the problem.

[8] Low enough means that the minimal payment constraint is not binding for outcomes within the agent's awareness. Otherwise the trade-off for the respective outcomes changes.

Expected Payoffs: The principal's outside option in the case of rejection is assumed to be zero. His expected payoff is given by

$$EU_P = \begin{cases} \sum_{y \in \mathcal{Y}} \pi(y|e) \left[y - C(y) \right] & \text{if the agent accepts,} \\ 0 & \text{if the agent rejects.} \end{cases}$$

The agent assesses his expected utility with respect to his restricted state space. The outside option of rejecting the contract is $\bar{U}$:

$$EU_A = \begin{cases} \sum_{y \in \widehat{\mathcal{Y}}} \widehat{\pi}(y|e) \upsilon \left(C \left(\widehat{y} \right) \right) - e & \text{if the agent accepts,} \\ \overline{U} & \text{if the agent rejects.} \end{cases}$$

4 The Optimal Contract with Observable Effort

In order to have a benchmark it is useful to first characterize the contract when effort is observable. If effort is observable and verifiable the contract can be made directly contingent on the action of the agent. The principal solves the problem:

$$\max_{\widehat{\Theta}, C(\cdot)} \quad \sum_{y \in \mathcal{Y}} \pi(y|e^H) \left[y - C(y) \right]$$

subject to

$$\sum_{y \in \widehat{\mathcal{Y}}} \widehat{\pi}(y|e^H) \upsilon \left(C(y) \right) - e^H \geq \overline{U}$$

$$C(y) \geq 0, \quad \forall y \in \mathcal{Y}.$$

We know that when the contract is complete it is optimal to give the agent full insurance. This can be seen from the first order conditions

$$\frac{1}{\upsilon'(C(y))} = \lambda, \qquad \forall y \in \mathcal{Y}.$$

The agent receives $C^{FB} = \upsilon^{-1} \left(\bar{U} + e^H \right)$ independent of the realization of Y. The first best is achieved. Now suppose the principal leaves the agent unaware of some contingencies. The first order conditions for $C(y), y \in \widehat{\mathcal{Y}}$ are

$$\frac{1}{\upsilon'(C(y))} = \lambda \frac{1}{\Pi(\widehat{\Theta}|e^H)}.$$

The first-order conditions imply that the transfer across $y \in \widehat{\mathcal{Y}}$ is constant. The optimal compensation scheme is simply $\widehat{C}^*(y) = C^{FB}, \forall y \in \widehat{\mathcal{Y}}$ and $\widehat{C}^*(y) = 0, \forall y \in \mathcal{Y} \backslash \widehat{\mathcal{Y}}$. The expected payment to the agent is $\Pi(\widehat{\Theta}|e^H)C^{FB}$. It is minimized when the probability of paying, $\Pi(\widehat{\Theta}|e^H)$, is minimized, which is achieved when the agent's awareness level is lowest. Consequently, if effort is observable, it is optimal to reveal nothing to the agent.

Proposition 4.1 *Under A1, A2 and observable effort,* $\widehat{\Theta}^* = \Theta_A$.

The reason for result 4.1 is that disclosing contingencies to the agent makes the participation constraint more costly to satisfy. Since effort is observable there is no incentive effect and only the participation effect matters.

5 The Optimal Contract with Unobservable Effort

Under unobservable effort the principal maximizes his expected profit subject to the participation constraint, the incentive constraint and the limited liability constraints. The participation constraint assures that the agent accepts the contract. The incentive constraint leads the agent to exert high effort e^H. The principal solves:

$$\max_{\widehat{\Theta}, C(\cdot)} \sum_{y \in \mathcal{Y}} \pi(y|e^H) \left[y - C(y)\right] \tag{1}$$

subject to

$$\sum_{y \in \widehat{\mathcal{Y}}} \widehat{\pi}(y|e^H) \upsilon\left(C\left(y\right)\right) - e^H \geq \overline{U} \tag{2}$$

$$e^H \in \arg\max_e \left\{ \sum_{y \in \widehat{\mathcal{Y}}} \widehat{\pi}(y|e) \upsilon\left(C\left(y\right)\right) - e \right\} \tag{3}$$

$$C(y) \geq 0, \qquad \forall y \in \mathcal{Y}, \tag{4}$$

where (2) is the participation constraint, (3) is the incentive constraint and (4) is the limited liability constraint. The analysis of the optimal contract can be divided into two steps. In step one the principal chooses the optimal compensation scheme $\widehat{C}$ given announcement $\widehat{\Theta}$. In step two he chooses the optimal level of awareness $\widehat{\Theta}$.

5.1 Step 1: Optimal Compensation Scheme given Awareness $\widehat{\Theta}$

The optimal compensation scheme given awareness $\widehat{\Theta}$ is characterized by (2), (3) and the necessary condition

$$\frac{1}{\upsilon'(C(y))} = \frac{1}{\Pi(\widehat{\Theta}|e^H)} \left(\lambda + \gamma \left[1 - \frac{\widehat{\pi}\left(y|e^L\right)}{\widehat{\pi}\left(y|e^H\right)} \right] \right), \quad \forall y \in \widehat{\mathcal{Y}}, \tag{5}$$

as well as $C(y) = 0, \forall y \in \mathcal{Y} \backslash \widehat{\mathcal{Y}}$.[9] Note that the optimal compensation scheme varies with the likelihood ratio $\frac{\widehat{\pi}(y|e^L)}{\widehat{\pi}(y|e^H)}$ of the restricted information structure $\widehat{\Theta}$ instead of Θ. Since the agent is unaware of the contingencies in $\Theta \backslash \widehat{\Theta}$, these signals cannot be used to induce e^H. As in the standard moral-hazard problem both participation and incentive constraint hold with equality. Let $\widehat{C}$ denote the solution to this system of equations.

Lemma 5.1 *Assume A1 and A2. Under $\widehat{C}$, both $\lambda > 0$ and $\gamma > 0$.*

Proof See Appendix A.2.2.

The optimal compensation scheme $\widehat{C}(\cdot)$ across $y \in \widehat{\mathcal{Y}}$ coincides with the the optimal compensation scheme of the standard principal-agent model with symmetric awareness and restricted information structure $\widehat{\Theta}$. To see this suppose for a moment that $\widehat{S}$ is the objective state space and that both, the principal and the agent, are symmetrically aware of $\widehat{S}$. Then the principal solves

$$\min_{C(\cdot)} \sum_{y \in \widehat{\mathcal{Y}}} \widehat{\pi}(y|e^H) C(y) \tag{6}$$

subject to (2) and (3). Let $C^C_{\widehat{\Theta}}$ denote the solution to this problem. $C^C_{\widehat{\Theta}}$ is the optimal complete contract under symmetric awareness and information structure $\widehat{\Theta}$. Under asymmetric awareness, the expected payment to the agent is

$$\sum_{y \in \mathcal{Y}} \pi(y|e^H) C(y) = \Pi(\widehat{\Theta}|e^H) \sum_{y \in \widehat{\mathcal{Y}}} \widehat{\pi}(y|e^H) C(y) - (1 - \Pi(\widehat{\Theta})) \cdot 0,$$

which is equivalent to (6) except for the scaling factor $\Pi(\widehat{\Theta}|e^H)$. Since $\Pi(\widehat{\Theta}|e^H)$ is nothing but a constant for a given $\widehat{\Theta}$, the two optimization problems are equivalent and $\widehat{C}(y) = C^C_{\widehat{\Theta}}(y)$ for all $y \in \widehat{\mathcal{Y}}$. The expected profit of the optimal incomplete contract is simply the expected payment of $C^C_{\widehat{\Theta}}$ weighted by the probability that none of the unforeseen contingencies are realized:

$$E[\widehat{C}(Y)|e^H] = \Pi(\widehat{\Theta}|e^H) E[C^C_{\widehat{\Theta}}(Y)|e^H].$$

5.2 Step 2: Optimal Awareness $\widehat{\Theta}^*$

The optimal level of disclosure is characterized by identifying the basic trade off between participation and incentives of disclosing contingencies to the agent. To separate the effect on incentives from the effect on participation, it is useful to compare the expected payment of complete contracts under different information structures.

[9] Inada conditions assure that (4) is not binding for $C(y), y \in \widehat{\mathcal{Y}}$.

Lemma 5.2 *Let Z be a non-empty subset of $\Theta \backslash \widehat{\Theta}$. Then*

$$\Delta C_{\widehat{\Theta}}^{Z} := E\left[C_{\widehat{\Theta}}^{C}(Y)|e^{H}\right] - E\left[C_{\widehat{\Theta} \cup Z}^{C}(Y)|e^{H}\right] \geq 0,$$

with strict inequality if and only if $\exists \theta \in Z$ *such that* $\Pr[\theta = 1|e^{H}] \neq \Pr[\theta = 1|e^{L}]$.

Proof See Appendix A.2.3.

This result is in line with Holmström's *Sufficient Statistic Theorem* (1979), which states that a signal θ is valuable if and only if it is informative.[10] Valuable means that both principal and agent can be made better off by including θ or simply that agency costs are reduced. Under independence θ is informative if and only if $\Pr[\theta = 1|e^{H}] \neq \Pr[\theta = 1|e^{L}]$.

5.2.1 The Basic Trade-Off

To understand the effect on participation and incentives of disclosing a subset of $\Theta \backslash \Theta_A$, compare the expected payoff of the principal under Θ_A and $\widehat{\Theta} \supset \Theta_A$:

$$\Pi(\Theta_A|e^{H}) E\left[C_{\Theta_A}^{C}(Y)|e^{H}\right] \gtrless \Pi(\widehat{\Theta}|e^{H}) E\left[C_{\widehat{\Theta}}^{C}(Y)|e^{H}\right], \tag{7}$$

or simply[11]

$$\prod_{\theta \in X} \Pr\left[\theta = 0 \mid e^{H}\right] E\left[C_{\Theta_A}^{C}(Y)|e^{H}\right] \gtrless E\left[C_{\widehat{\Theta}}^{C}(Y)|e^{H}\right].$$

Using $\Delta C_{\Theta_A}^{X} = E\left[C_{\Theta_A}^{C}(Y)|e^{H}\right] - E\left[C_{\widehat{\Theta}}^{C}(Y)|e^{H}\right]$ we can restate (7) in terms of gains and losses of revealing set X:

$$\Delta C_{\Theta_A}^{X} \gtrless \left(1 - \prod_{\theta \in X} \Pr\left[\theta = 0|e^{H}\right]\right) E\left[C_{\Theta_A}^{C}(Y)|e^{H}\right]. \tag{8}$$

$\Delta C_{\Theta_A}^{X}$ is the incentive gain of disclosing set X. The announcement of an informative contingency allows the principal to use the contingency as a signal about the agent's effort. Disclosing an informative contingency to the agent consequently relaxes the incentive constraint and the information gain is captured by $\Delta C_{\Theta_A}^{X}$. $(1 - \prod_{\theta \in X} \Pr[\theta = 0|e^{H}]) E[C_{\Theta_A}^{C}(Y)|e^{H}]$ is the participation loss of disclosing X. With probability $1 - \prod_{\theta \in X} \Pr\left[\theta = 0|e^{H}\right]$ one of the contingencies in X is realized. Under announcement $\Theta_A \cup X$ the principal has to pay the agent a positive wage, whereas under Θ_A he pays zero. Disclosing a contingency to the agent consequently tightens the participation

[10] Holmström (1979) shows this for continuous outcome and continuous effort.

[11] Remember that $\widehat{\Theta} = \Theta_A \cup X$.

constraint. Under observable effort there is no incentive effect, which is why it is optimal to keep the agent unaware. Under unobservable effort the principal chooses $\widehat{\Theta}$ such that the net gain of revelation is maximized. Since Θ is finite, he compares a finite number of announcement strategies and their respective expected payoffs. Whenever there exists a $\widehat{\Theta}$ such that the incentive effect outweighs the participation effect, the principal enlarges the agent's awareness. Whether this is the case or not generally depends on the exogenous parameters of the model.

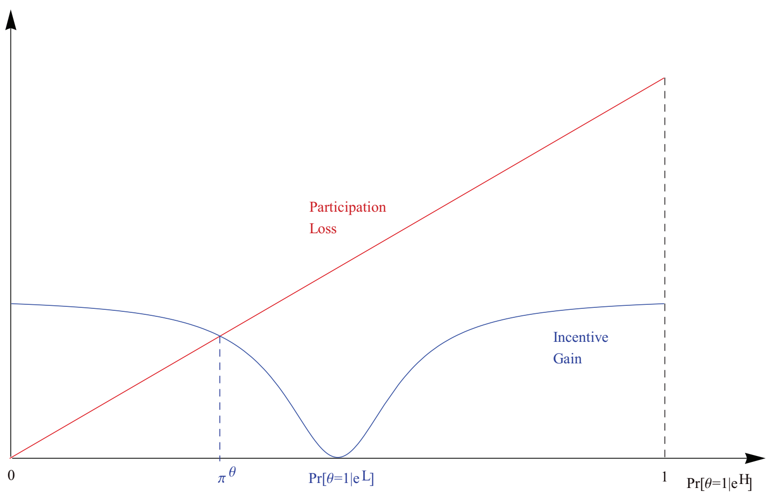
5.2.2 Single Announcements

To find some conditions on the trade-off in (8) it useful to first analyze one step deviations from set Θ_A. Consider the announcement of $\theta \notin \Theta_A$ and define $\Delta^\theta := \left|\Pr[\theta = 1|e^H] - \Pr[\theta = 1|e^L]\right|$. Δ^θ is a measure of informativeness of signal θ. The larger Δ^θ the more informative is the realization of θ about the effort of the agent. Let $V_{\widehat{\Theta}} := E[C^C_{\widehat{\Theta}}(Y)|e^H]$ denote the principal's expected payment to the agent given full awareness, restricted information structure $\widehat{\Theta}$ and optimal compensation scheme $C^C_{\widehat{\Theta}}$. $V_{\widehat{\Theta}}$ is decreasing in the informativeness of signals $\theta \in \widehat{\Theta}$ due to decreasing agency costs.

Lemma 5.3 *Assume A1 and A2. $V_{\widehat{\Theta}}$ is monotonically decreasing in $\Delta^\theta, \forall \theta \in \widehat{\Theta}$.*

Proof See Appendix A.2.4

Figure 1: Trade-Off



Lemma 5.3 allows us to find conditions on the trade-off between participation and incentives for single announcements. Consider the trade-off of disclosing θ, illustrated in Figure 1:[12]

$$\Delta C^{\theta}_{\Theta_A} \quad \text{vs.} \quad \Pr\left[\theta = 1|e^H\right] E\left[C^C_{\Theta_A}(Y)|e^H\right].$$

The incentive gain $\Delta C^{\theta}_{\Theta_A} = E[C^C_{\Theta_A}] - E[C^C_{\Theta_A \cup \theta}]$ as a function of $\Pr[\theta = 1|e^H]$ is illustrated by the blue curve. It is decreasing on $[0, \Pr[\theta = 1|e^L])$ and increasing on $(\Pr[\theta = 1, e^L], 1]$. To see this, note that $E[C^C_{\Theta_A}]$ is independent of the distributional properties of θ and $E[C^C_{\Theta_A \cup \theta}]$ is decreasing in Δ^{θ} (Lemma 5.3). At $\Pr[\theta = 1|e^L]$ signal θ is uninformative and $\Delta C^{\theta}_{\Theta_A} = 0$. The participation loss $\Pr\left[\theta = 1|e^H\right] E\left[C^C_{\Theta_A}(Y)|e^H\right]$ is illustrated by the red line. It is linearly increasing in $\Pr[\theta = 1|e^H]$.

Figure 1 shows that if probability $\Pr[\theta = 1|e^H]$ is small enough, the incentive gain outweighs the participation loss. To see that this is generally the case, note that

$$\lim_{\Pr[\theta=1|e^H]\to 0} \quad \Delta C^{\theta}_{\Theta_A} = E[C^C_{\Theta_A}(Y)|e^H] - C^{FB} > 0,$$
$$\lim_{\Pr[\theta=1|e^H]\to 0} \quad \Pr\left[\theta = 1|e^H\right] E\left[C^C_{\Theta_A}(Y)|e^H\right] = 0.$$

Hence, for $\Pr[\theta = 1|e^H]$ close to zero the incentive gain is strictly greater than the participation loss, whereas for $\Pr[\theta = 1|e^H] = \Pr[\theta = 1|e^L]$ the incentive gain is strictly smaller than the participation loss. Since the incentive gain is monotonically decreasing on $(0, \Pr[\theta = 1|e^L])$ and the participation loss is linearly increasing on $(0, \Pr[\theta = 1|e^L])$, we know that there is a threshold on that interval, such that the incentive effect outweighs the participation effect whenever $\Pr[\theta = 1|e^H]$ is smaller than this threshold.

Proposition 5.4 *Assume A1 and A2. For every $\theta \in \Theta \backslash \Theta_A$ there exists a threshold $\pi^{\theta} \in (0, \Pr[\theta = 1|e^L])$ such that if $\Pr[\theta = 1|e^H] < \pi^{\theta}$, the incentive effect outweighs the participation effect.*

Proof Omitted.

Note that if there exists a contingency θ outside the agent's awareness such that $\Pr[\theta = 1|e^H] < \pi^{\theta}$ the principal strictly prefers to reveal θ to the agent than to reveal nothing.

5.2.3 Combined Announcements

In the previous section I characterized the gains and losses of revealing a contingency θ for a given set Θ_A. When the agent is unaware of more than one contingency the trade-off associated with the

[12]This figure shows $\Delta C^{\theta}_{\Theta_A}$ and $\Pr[\theta = 1|e^H] E[C^C_{\Theta_A}(Y)|e^H]$ for the following specification: $\upsilon(C) = \frac{C^{1-\sigma}}{1-\sigma}, \sigma = 0.5, e^H = 1, e^L = 0, \bar{U} = 5$. There are two contingencies. Contingency $\theta \in \Theta_A$ with $\Pr[\theta = 1|e^H] = 0.65$ and $\Pr[\theta = 1|e^L] = 0.5$ and contingency $\theta' \notin \Theta_A$ with $\Pr[\theta' = 1|e^L] = 0.5$.

revelation of one of these contingencies generally depends on the revelation of other contingencies. To see how the trade-off changes as other contingencies are revealed I consider the special case of symmetric contingencies. I specify conditions under which the net gain of making the agent aware decreases in the number of contingencies revealed. A decreasing net gain implies that whenever the revelation of any single contingency yields a negative net gain, a combined revelation cannot be profitable. With this, it is straight forward to find the optimal set of revelation $\widehat{\Theta}^*$.

Suppose the agent is unaware of N symmetric contingencies, $\Theta \backslash \Theta_A = \{\theta_1, ..., \theta_N\}$, where symmetric means that every contingency has the same distribution. Let $\alpha := \Pr[\theta_i = 0|e^H], i = 1, ...N$ and $\widehat{\Theta}_n := \Theta_A \cup \theta_1 \cup ... \cup \theta_n$.[13] The net gain of revealing the nth contingency is:

$$NG(\theta_n) := \alpha^{N-n} \left(\Delta C^{\theta_n}_{\widehat{\Theta}_{n-1}} - (1-\alpha) E[C^C_{\widehat{\Theta}_{n-1}}(Y)|e^H] \right).$$

The value of $NG(\theta_n)$ is determined by two factors. First, there is the difference between incentive gain and participation loss as we saw in the previous section. This difference is weighted by the probability that the gains and losses are realized, α^{N-n}. Note that both $\Delta C^{\theta_n}_{\widehat{\Theta}_{n-1}}$ and $(1-\alpha)E[C^C_{\widehat{\Theta}_{n-1}}(Y)|e^H]$ are decreasing in n, whereas α^{N-n} is exponentially increasing in n. Figure 2(a) shows the weighted incentive gain $\alpha^{N-n}\Delta C^{\theta_n}_{\widehat{\Theta}_{n-1}}$ and the weighted participation loss $\alpha^{N-n}(1-\alpha)E[C^C_{\widehat{\Theta}_{n-1}}(Y)|e^H]$ as a function of n. Generally, $NG(\theta_n)$ can decrease or increase in n.

Lemma 5.5 *Assume A1, A2 and symmetric contingencies. $NG(\theta_n)$ is decreasing in n if*

$$E[C^C_{\widehat{\Theta}_n}(Y)|e^H]^2 < E[C^C_{\widehat{\Theta}_{n-1}}(Y)|e^H] E[C^C_{\widehat{\Theta}_{n+1}}(Y)|e^H] \quad \textit{for all} \quad n \in \{1, ..., N-1\}. \tag{9}$$

Proof See Appendix A.2.5

The intuitive interpretation of condition (9) is that the expected payment of the complete contract $C^C_{\widehat{\Theta}_n}$ as a function of n is sufficiently convex, i.e. the inclusion of signal θ_n in the information structure reduces agency costs strongly when n is small but only marginally when n is large. Condition (9) implies that disclosing θ_{n+1} yields a lower net gain than disclosing θ_n, depicted in Figure 2(b).[14]

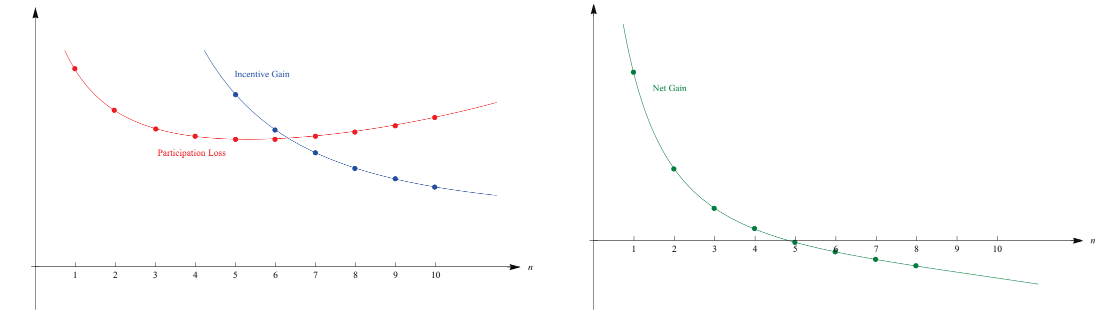
If condition (9) holds, it is possible to determine the optimal level of revelation $\widehat{\Theta}^*$. First, note that the net gain of any set $Z = \{\theta_{\bar{n}}, \theta_{\bar{n}+1}, ..., \theta_{\bar{n}+z}\}$ is equal to the sum over individual net gains $NG(\theta_n), n = \bar{n}, ..., \bar{n}+z$.

$$NG(Z) = \alpha^{N-(\bar{n}-1)} E[C^C_{\widehat{\Theta}_{\bar{n}-1}}(Y)|e^H] - \alpha^{N-(\bar{n}+z)} E[C^C_{\widehat{\Theta}_{\bar{n}+z}}(Y)|e^H] = \sum_{n=\bar{n}}^{\bar{n}+z} NG(\theta_n).$$

[13] For notational convenience let $\widehat{\Theta}_0 = \Theta_A$.

[14] Condition (9) is satisfied, for example when $v(C) = \frac{C^{1-\sigma}}{1-\sigma}$, $\sigma = 0.5$, $e^H = 1$, $e^L = 0$, $\bar{U} = 9$, $N = 5$, $\alpha = 0.6$ and $\Pr[\theta = 0|e^L] = \Pr[\theta = 1|e^H] = 0.5$.

Figure 2: Trade-Off θ_n



(a) Weighted Incentive Gain and Participation Loss of Revealing θ_n

(b) Net Gain of Revealing θ_n

Profits are maximized when the net gain of revelation $\sum_{n=1}^{n^*} NG(\theta_n)$ is maximized. This is achieved by including all contingencies θ_n such that $NG(\theta_n) \geq 0$.

Corollary 5.6 *Assume A1, A2 and symmetric contingencies. If condition* (9) *is satisfied and if* $NG(\theta_N) < 0$*, the optimal announcement* $\widehat{\Theta}^*$ *is* $\widehat{\Theta}_{n^*}$*, where* n^* *is such that*

$$NG(\theta_{n^*}) \geq 0 \quad \textit{and} \quad NG(\theta_{n^*+1}) < 0.$$

Otherwise $\widehat{\Theta}^* = \Theta$.

Proof Omitted.

5.3 Justifiability of the Contract

If the contract is incomplete, the agent's perception of the world differs from the principal's. An important question is whether the proposed contract can elicit suspicion on the side of the agent. When receiving a contract the agent may ask herself whether announcement $\widehat{\Theta}$ and compensation scheme C maximizes the principal's payoff. The agent can only contemplate contracts within his awareness. These are contracts that specify any $\tilde{\Theta}, \Theta_A \subseteq \tilde{\Theta} \subseteq \widehat{\Theta}$ and any $C \in \mathbb{R}_0^+$ for all $y \in \widehat{\mathcal{Y}}$. Filiz-Ozbay (2012) introduces an equilibrium refinement, which requires that the equilibrium contract maximizes the principal's expected payoff from the viewpoint of the agent. In line with her equilibrium refinement, I define justifiability of a contract as follows.

Definition 3 *A contract* $(\widehat{\Theta}, C)$ *is called justifiable if it is a solution to the optimization problem*

$$\max_{\tilde{\Theta} \subseteq \widehat{\Theta}, C(\cdot), e} EU_P^A := \sum_{y \in \widehat{\mathcal{Y}}} \widehat{\pi}(y|e) \left[y - C(y)\right]$$

subject to (2), (3) and $C(y) \geq 0, \forall y \in \widehat{\mathcal{Y}}$.

EU_P^A is the expected payoff of the principal evaluated at the agent's beliefs. Definition 3 requires contract $(\widehat{\Theta}, C)$ to be optimal for the principal from the viewpoint of the agent. It turns out that under relatively mild conditions the optimal contract $(\widehat{\Theta}^*, \widehat{C}^*)$ indeed maximizes the principal's payoff evaluated at the agent's beliefs.

Proposition 5.7 *Assume A1 and A2.* $(\widehat{\Theta}^*, \widehat{C}^*)$ *is justifiable according to Definition 3 if and only if* $EU_P^A(\widehat{\Theta}^*, \widehat{C}^*, e^H) \geq 0$ *and* $EU_P^A(\widehat{\Theta}^*, \widehat{C}^*, e^H) \geq EU_P^A(\widehat{\Theta}^*, \widehat{C}^*, e^L)$.

Proof See Appendix A.2.6.

A necessary and sufficient condition for the optimal contract to be justifiable is that the principal's expected utility is non-negative and that e^H is the optimal action choice from the viewpoint of the agent. Up to now the optimization problem was disconnected from the actual outcome levels in $\mathcal{Y}$ because only the likelihood ratio associated with each outcome level is relevant for the induction of e^H. If we require the contract to be justifiable, this is no longer necessarily the case. Whenever the optimal contract is incomplete, the agent perceives only a subset of possible outcomes. It is possible that the optimal contract leaves the agent unaware of high outcomes such that $EU_P^A(\widehat{\Theta}^*, \widehat{C}^*, e^H) < 0$ whereas $EU_P(\widehat{\Theta}^*, \widehat{C}^*, e^H) \geq 0$. Similarly, it is possible that the optimal contract leaves the agent unaware of outcomes that are strongly correlated with effort such that $EU_P^A(\widehat{\Theta}^*, \widehat{C}^*, e^H) \geq EU_P^A(\widehat{\Theta}^*, \widehat{C}^*, e^L)$ whereas $EU_P(\widehat{\Theta}^*, \widehat{C}^*, e^H) \geq EU_P(\widehat{\Theta}^*, \widehat{C}^*, e^L)$. If any of the two conditions in Proposition 5.7 is violated, the refinement introduces another dimension in the trade-off. The principal is no longer only concerned with the distributional properties of the contingencies in $\Theta \backslash \Theta_A$, but also with the outcomes an announcement reveals.

Given that the expected profit evaluated at the agent's beliefs is non-negative and higher at e^H than at e^L, the optimal contract $(\widehat{\Theta}^*, \widehat{C}^*)$ is justifiable. Justifiability of $\widehat{C}^*$ is straight forward. If the optimal contract is complete, principal and agent share the same belief. Hence, $\widehat{C}^*$ maximizes the principal's expected payoff from the agent's perspective. If the optimal contract is incomplete, we know that the transfer rule for outcomes within the agent's awareness coincides with the optimal compensation scheme of the complete contract given information structure $\widehat{\Theta}$. Since the agent thinks that the contract is complete, $\widehat{C}^*$ solves the principal's optimization problem given the agent's beliefs.

To see justifiability of $\widehat{\Theta}^*$, remember that the agent can only consider the announcement of contingencies within his awareness. Suppose $\Theta_A \subset \widehat{\Theta}^*$. The agent evaluates the principal's expected payoff for every announcement $\widehat{\Theta}^* \backslash Z$, where $Z \in \widehat{\Theta}^* \backslash \Theta_A$. The agent knows that given announcement $\widehat{\Theta}^* \backslash Z$ the optimal compensation scheme is $C(y) = C^C_{\widehat{\Theta}^* \backslash Z}(y)$ for all y within the agent's hypothetical awareness $\widehat{\Theta}^*\backslash Z$ and zero otherwise. Thus, after reading the contract, the agent understands the principal's optimal contract for any level of awareness lower or equal than his actual awareness, but he does not understand that there may remain contingencies that he is unaware of. For $\widehat{\Theta}^*$ to be justifiable it has to be true that

$$E\left[C^C_{\widehat{\Theta}^*}(Y)|e^H\right] \leq \prod_{\theta\in\widehat{\Theta}^*\backslash Z} \Pr[\theta=0|e^H] E\left[C^C_{\widehat{\Theta}^*\backslash Z}(Y)|e^H\right], \tag{10}$$

for any $Z \subseteq \widehat{\Theta}^*\backslash\Theta_A$. This coincides with the optimality condition of the principal. Consequently, (10) is fulfilled and $\widehat{\Theta}^*$ can be rationalized by the agent.

6 Competing Principals

In the basic set-up I analyze the optimization problem of a monopolistic principal. This section addresses the question of how these results change when principals compete against each other. Suppose there are N principals that are all aware of Θ. They make simultaneous offers, denoted by $(\widehat{\Theta}_i, C_i), i = 1, ..., N$. The agent updates his awareness after hearing all of the offers and accepts at most one. He considers henceforth every contingency in $\widehat{\Theta}_1 \cup ... \cup \widehat{\Theta}_N$. If the agent is indifferent between two or more contracts he accepts each contract with equal probability. I focus on symmetric equilibria in pure strategies, $(\widehat{\Theta}_i, C_i) = (\widehat{\Theta}, C), i = 1, ..., N$. Principal $i's$ payoff function is

$$EU_i = \frac{1}{N}\left(E[Y|e^H] - E[C(Y)|e^H]\right), \qquad i = 1, ..., N.$$

In the absence of asymmetric awareness, principals engage in a Bertrand competition over transfer rule C. The second-best allocation is achieved, where principals make zero profits and the surplus goes to the agent. Let C^* denote the equilibrium compensation scheme under full awareness.

Proposition 6.1 *Assume A1 and A2 and assume that the second-best surplus is strictly positive. There is a unique symmetric Nash equilibrium in which the agent is fully aware and each principal offers the complete zero profit contract* (Θ, C^*)*.*

Proof See Appendix A.2.7.

The intuition for Proposition 6.1 is that whenever the announcements of the other principals promote full awareness the own announcement is not payoff relevant. Further, any change in the compensation scheme yields either negative or zero expected profits for the standard Bertrand competition argument. Thus, the equilibrium exist. The reason for the equilibrium to be unique is simply that for any level of awareness, principals engage in a Bertrand competition over the compensation scheme. As profits become small enough it is profitable to deviate and enlarge the agent's awareness. This deviation allows the deviator to capture the whole market and make positive profits, because the agent recognizes the zero payments of the other contracts in the newly revealed states.

When we require the contract to be justifiable according to Definition 3, the symmetric zero profits Nash equilibrium is no longer necessarily unique. Suppose each principal announces $\widehat{\Theta}_i = \widehat{\Theta} \subset \Theta, i = 1, ...N$ and offers compensation $C_i = \widehat{C}, i = 1, ..., N$ such that the agent's utility is maximized across the states the agent is aware of, the incentive constraint is satisfied and the expected payoff of the principals from the viewpoint of the agent is equal to zero:

$$\sum_{y \in \widehat{\mathcal{Y}}} \widehat{\pi}(y|e^H)[y - \widehat{C}(y)] = 0.$$

The agent believes that he receives an expected transfer of $\sum_{y \in \widehat{\mathcal{Y}}} \widehat{\pi}(y|e^H)\widehat{C}(y)$, whereas the actual expected transfer is $\Pi(\widehat{\Theta}|e^H)E[\widehat{C}(Y)|e^H]$. Note that $\Pi(\widehat{\Theta}|e^H)E[\widehat{C}(Y)|e^H]$ is strictly smaller than $\sum_{y \in \widehat{\mathcal{Y}}} \widehat{\pi}(y|e^H)\widehat{C}(y)$, because the agent does not take into account the zero payments in $y \in \mathcal{Y} \backslash \widehat{\mathcal{Y}}$. Consequently, it is possible that $(\widehat{\Theta}, \widehat{C})$ yields a positive expected payoff for all principals:

$$EU_i = \frac{1}{N} \sum_{y \in \mathcal{Y}} \pi(y|e)[y - \widehat{C}(y)] > 0.$$

To see that $(\widehat{\Theta}, \widehat{C})$ can be indeed an equilibrium contract, note that any deviation in the compensation scheme C_i either makes the agent worse off or makes the contract non-justifiable according to Definition 3. A joint deviation in $\widehat{\Theta}_i$ and C_i, allows principal i to capture the whole market. Whether this deviation yields a higher expected payoff than offering $(\widehat{\Theta}, \widehat{C})$ depends on the number of principals among other parameters of the model. The larger the number of principals, the smaller is the expected profit when no one is deviating. This implies that we can find an N large enough such that deviating is always profitable until the agent is fully aware.

Proposition 6.2 *Assume A1 and A2, assume that the second-best surplus is strictly positive and assume that the agent rejects any non-justifiable contract according to Definition 3. There exists a critical $\bar{N} \in \mathbb{N}$, such that for every $N > \bar{N}$ in equilibrium the agent is fully aware.*

Proof See Appendix A.2.8.

The result that competition promotes awareness if the number of competitors is large enough was first shown by Filiz-Ozbay (2012). The line of argument is essentially the same. The profit of each principal in a symmetric equilibrium with unawareness decreases with the number of principals. Enlarging the agent's awareness in Filiz-Ozbay's (2012) environment implies changing the beliefs of the agent, which allows the deviator to capture the whole market. Hence, whenever the number of principals is large enough, revealing new states to the agent is profitable and in equilibrium the agent is fully aware.

In contrast to my result, Gabaix and Laibson (2006) propose a model where unawareness prevails even under intense competition. In their model firms compete for consumers of whom some are fully aware and the others are unaware. Firms can educate a fraction of the unaware consumers, but they have no incentive to do so. The reason is that aware consumers profit from the existence of unaware consumers, so whenever an unaware agent is made aware by a competing firm he has no incentives to switch, because he profits from the remaining unaware consumers.

7 Discussion

7.1 Dependence between Contingencies

Throughout the main part of the analysis it is assumed that contingencies are conditionally independent of each other. Giving up this assumption, unawareness has an additional effect on the optimal compensation scheme. To see this, consider first the case of observable effort. When $\widehat{\Theta} \neq \Theta$, the necessary condition is

$$\frac{1}{\upsilon'(C(y))} = \lambda \frac{\widehat{\pi}(y|e^H)}{\pi(y|e^H)}, \qquad \forall y \in \widehat{\mathcal{Y}}.$$

$\frac{\widehat{\pi}(y|e^H)}{\pi(y|e^H)}$ is the ratio of the agent's perceived probability to the objective probability. Under conditional dependence this ratio depends on y, because the probability that none of the unforeseen contingencies are realized varies across $y \in \widehat{\mathcal{Y}}$. It is optimal to promise a high payment at y if the probability that one of the unforeseen contingencies is realized is relatively high, because the probability that the principal has to keep his promise is relatively low. This makes a deviation from the optimal risk sharing rule profitable. Since the agent cannot rationalize a non-constant compensation scheme, such a contract is not justifiable according to Definition 3.

When effort is not observable, the first order condition is

$$\frac{1}{\upsilon'(C(y))} = \frac{\widehat{\pi}(y|e^H)}{\pi(y|e^H)} \left\{ \lambda + \gamma \left[1 - \frac{\widehat{\pi}(y|e^L)}{\widehat{\pi}(y|e^H)} \right] \right\}, \quad \forall y \in \widehat{\mathcal{Y}}.$$

It shows that there are two sources of variation in the optimal compensation scheme $\widehat{C}$ across $y \in \widehat{\mathcal{Y}}$. First, $\widehat{C}$ varies with the ratio of perceived to objective probability just as in the case of observable effort. The second source of variation comes from the agent's perceived likelihood ratio as in the basic model. Again, the agent cannot rationalize the first source of variation and the optimal contract is not necessarily justifiable according to Definition 3. If we require the contract to be justifiable, the optimal compensation scheme ignores the variation in the probability of $y \in \widehat{\mathcal{Y}}$ and is equivalent to $(\widehat{\Theta}^*, \widehat{C}^*)$ under conditional independence.

7.2 The Output Function

In the basic model, output is discrete and differs across every state of the world. If y is not one-to-one and contingencies are not observable, both participation loss and incentive gain of revelation are affected. The participation loss is generally diminished, because the revelation of an unforeseen contingency does not necessarily imply the revelation of an unforeseen outcome. The incentive gain is no longer unambiguous, because the agent's perceived distribution of outcome can be more informative about the agent's effort than the actual distribution of outcome.

To see this, suppose there are only two possible realizations of outcome. The project can be either a success or a failure, $\mathcal{Y} = \{s, f\}$ with $s > f$. Whether the project is a success or a failure depends on the realization of Θ. If $\widehat{\Theta} \subset \Theta$ the agent is aware of outcomes $\{s, f\}$ but believes probability distribution $\widehat{\pi}(.|e)$.[15] Now consider the trade-off the principal faces when disclosing $\theta \notin \Theta_A$. Since the revelation of θ does not reveal any new outcomes the participation loss of announcement is zero. The effect on incentives depends on how the revealed contingency affects the perceived distribution of the agent. To see that the incentive effect can be negative, suppose $\Theta = \{\theta_1, \theta_2\}$, $\Theta_A = \{\theta_1\}$ and assume θ_2 is not informative, i.e. $\Pr[\theta_2 = 1|e^H] = \Pr[\theta_2 = 1|e^L]$. In the basic model $\Pr[\theta = 1|e^H] = \Pr[\theta = 1|e^L]$ implies $\Delta C^{\theta}_{\Theta_A} = 0$, because the principal has the choice to ignore the realization of θ_2. Under $\mathcal{Y} = \{s, f\}$ this is no longer the case. Suppose that $y = s$ whenever $\theta_1 = \theta_2 = 0$ and $y = f$ otherwise. Solving for the optimal compensation scheme, the expected payment under $\widehat{\Theta}^* = \Theta_A$ is smaller than the expected payment under $\widehat{\Theta}^* = \Theta$.[16] There is an incentive loss of revealing θ_2 because the perceived distribution of the unaware agent is more informative about the action choice than the true distribution. A further discussion on

[15] Generally it is possible that the agent is only aware of one outcome. In that case the principal has to make the agent aware of the other outcome, otherwise it is impossible to induce e^H.

[16] If the principal leaves the agent unaware, the optimal compensation scheme is

$$C^u(s) = \upsilon^{-1}\left(\bar{U} + \frac{\Pr[\theta_1 = 1|e^L]e^H - \Pr[\theta_1 = 1|e^H]e^L}{\Pr[\theta_1 = 1|e^L] - \Pr[\theta_1 = 1|e^H]}\right),\ C^u(f) = \upsilon^{-1}\left(\bar{U} - \frac{\Pr[\theta_1 = 0|e^L]e^H - \Pr[\theta_1 = 0|e^H]e^L}{\Pr[\theta_1 = 1|e^L] - \Pr[\theta_1 = 1|e^H]}\right).$$

If the principal reveals contingency θ_2, the optimal compensation scheme is

$$\begin{aligned} C^a(s) &= \upsilon^{-1}\left(\bar{U} + \frac{\left(1 - \Pr[\theta_1 = 0|e^L]\Pr[\theta_2 = 0|e^L]\right)e^H - \left(1 - \Pr[\theta_1 = 0|e^H]\Pr[\theta_2 = 0|e^H]\right)e^L}{\Pr[\theta_1 = 0|e^H]\Pr[\theta_2 = 0|e^H] - \Pr[\theta_1 = 0|e^L]\Pr[\theta_2 = 0|e^L]}\right) \\ C^a(f) &= \upsilon^{-1}\left(\bar{U} - \frac{\Pr[\theta_1 = 0|e^L]e^H - \Pr[\theta_1 = 0|e^H]e^L}{\Pr[\theta_1 = 1|e^L] - \Pr[\theta_1 = 1|e^H]}\right). \end{aligned}$$

So $C^u(f) = C^a(f)$ and $\upsilon(C^u(s)) - \upsilon(C^a(s)) = -\frac{\Pr[\theta_2=1|e^H]}{\Pr[\theta_2=0|e^H]}\frac{e^H - e^L}{\Pr[\theta_1=0|e^H]-\Pr[\theta_1=0|e^L]} < 0$, which implies that $E[C^u(Y)|e^H] < E[C^a(Y)|e^H]$.

heterogeneous priors and moral hazard can be found in Santos-Pinto (2008) and De la Rosa (2011).

7.3 Optimal Action Choice

Throughout the analysis I assumed that $E[Y|e^H] - E[Y|e^L]$ is large enough such that the principal always finds it optimal to induce e^H. Unawareness makes incentives more costly, hence it is generally possible that for different levels of awareness different levels of effort are optimal. Since effort can only be high or low, the analysis of the optimal contract under low effort is straightforward.

Proposition 7.1 *Assume A1 and A2. If e^L is the action choice, the optimal contract is (Θ_A, C^L) with $C^L(y) = \upsilon^{-1}(\bar{U} - e^L), \forall y \in \widehat{\mathcal{Y}}$ and $C^L(y) = 0, \forall y \in \mathcal{Y} \backslash \widehat{\mathcal{Y}}$.*

Proof See Appendix A.2.9.

The optimal contract inducing low effort leaves the agent unaware because there is no incentive effect. The principal induces low effort in equilibrium if

$$E[Y|e^L] - \prod_{\theta \in \Theta \backslash \Theta_A} \Pr[\theta = 0|e^L]\upsilon^{-1}(\bar{U} - e^L) > E[Y|e^H] - \prod_{\theta \in \Theta \backslash \widehat{\Theta}^*} \Pr[\theta = 0|e^L]E[C^C_{\widehat{\Theta}^*}(Y)|e^H].$$

Whether this is the case or not depends on the distributional properties of the random variable Y, but it is easy to find examples where e^H is the optimal action choice under full awareness and e^L is the optimal action choice under asymmetric awareness.

7.4 The Agent as the Residual Claimant

The basic model assumes that the principal is the residual claimant. In the presence of asymmetric awareness ownership of the project matters, because the agent's valuation of the project depends on his level of awareness. If the agent is the residual claimant, unawareness generally affects his perceived outside option $\sum_{y \in \widehat{\mathcal{Y}}} \widehat{\pi}(y|e^H)\upsilon(y)$. The principal solves:

$$\max_{\widehat{\Theta}, C(\cdot)} \sum_{y \in \mathcal{Y}} \pi(y|e^H)[P - C(y)]$$

subject to

$$\sum_{y \in \widehat{\mathcal{Y}}} \widehat{\pi}(y|e^H)\upsilon\left(y + C(y) - P\right) \geq \sum_{y \in \widehat{\mathcal{Y}}} \widehat{\pi}(y|e^H)\upsilon(y)$$

$$e^H \in \arg\max_e \left\{ \sum_{y \in \widehat{\mathcal{Y}}} \widehat{\pi}(y|e)\upsilon(y + C(y) - P) - e \right\}$$

$$C(y) \geq 0, \qquad \forall y \in \mathcal{Y},$$

where P is the premium paid by the agent and $C(y)$ is the outcome contingent transfer.[17] As in in the basic model, revealing a contingency $\theta \notin \Theta_A$ involves a trade-off between participation and incentives. In addition, enlarging the agent's awareness affects his perceived outside option. This effect is favorable to the principal if

$$\sum_{y \in \tilde{\mathcal{Y}}} \tilde{\pi}(y|e^H)v(y) < \sum_{y \in \widehat{\mathcal{Y}}} \widehat{\pi}(y|e^H)v(y),$$

where $\tilde{\Theta} = \widehat{\Theta} \cup \theta, \theta \in \Theta \backslash \widehat{\Theta}$. This is the case if $E[Y|\theta = 1] < E[Y|\theta = 0]$, i.e. if $\theta = 1$ is a negative outcome shock. Consequently, if the agent is the residual claimant, not only do the distributional properties of $\theta \notin \Theta_A$ matter but also the outcome its announcement reveals. Roughly speaking, when the agent is the residual claimant the principal includes contingencies in the contract that are very unlikely, highly informative and that reveal "bad" outcomes. We can think of this setting as a contract between an insurer and an insuree, where the insuree is partly unaware and effort affects the probability of incurring a loss. The additional participation effect gives the insurer incentives to reveal severe calamities to the insuree, such that the insuree is willing to buy insurance at a higher price.

7.5 Unawareness and Zero Probability Beliefs

A frequently raised concern is whether unawareness is observationally equivalent to full awareness with zero probability beliefs (Li, 2008). Epistemically, unawareness has very different properties from zero probability beliefs. An agent is unaware if and only if he assigns probability zero to an event and to its negation (Heifetz et al., 2011). Schipper (2012) shows how this feature also implies behavioral differences between unawareness and zero probability beliefs. However, the results derived in my model can be generated in a framework with full awareness and zero probability beliefs. Under the interpretation of heterogeneous priors, there are some caveats to be taken into account.

In order to derive my results in a framework with full awareness and zero probability beliefs, the agent needs to update a zero probability prior to a non-zero posterior. Note that such updating cannot be interpreted as a consequence of the arrival of new information, since information in the standard state space model expands the set of null states instead of narrowing it. Generating my results in the standard state space framework requires a model that allows for manipulation of beliefs rather than revelation of information. Allowing the principal to manipulate the agent's beliefs without the presence of hard information is rather difficult to motivate. Further, it is important to note that reporting the true distribution is not incentive compatible for the principal if lying is possible. Consequently, there are strong assumptions on the message set available to the

[17] Due to the limited liability constraint the principal would like to scale up both P and C. In order to have a solution, one has to assume that such a contract elicits suspicion on the side of the agent.

principal necessary.[18] Given these caveats, asymmetric awareness seems to be a more natural way to think about this environment and the arising trade-off.

8 Conclusion

This paper incorporates asymmetric awareness in the classical principal-agent model. It shows that the principal makes the agent strategically aware and that the optimal contract can be incomplete. Enlarging the agent's awareness involves a trade-off between participation and incentives. The cost of disclosing contingencies to the agent is the payment in the states that the agent is initially unaware of. The gain of disclosing contingencies to the agent is the richer information structure that is used to induce incentives. Hence, it is profitable to announce contingencies that have a low probability but are highly correlated with the effort of the agent. Under relatively mild assumptions the optimal contract is justifiable for the agent, which means that the optimal contract maximizes the principal's expected profit evaluated at the beliefs of the agent.

If we allow for competition among principals, in the unique symmetric Nash equilibrium the agent is fully aware and principals make zero profits. If the contract is required to be justifiable for the agent, this equilibrium is no longer necessarily unique but when the number of principals is large enough uniqueness is restored.

In the proposed model, the principal is able to implement zero payments whenever there is an event the agent is initially unaware of. This may not be feasible in real-life contracting situations. Restricting the set of feasible contracts allows us to generate results that are closer to observed contracts, but the basic trade-off prevails. If, for example, the compensation scheme is restricted to be monotone in the outcome, there are three effects driving the optimal revelation strategy: In addition to the participation and incentive effect we saw in the basic model also the revealed outcomes matter, where it is most costly to disclose low outcomes to the contracting partner. Restricting the set of feasible contracts adds interesting features to the optimal compensation scheme and revelation strategy, but as long as the contracting partner with superior awareness is able to profit from the other's limited understanding of the underlying uncertainties the basic trade-off prevails.

[18] One may argue that also in the framework with unawareness an implicit assumption is that the principal can only reveal "true" contingencies. One could imagine a case, where the principal can include virtual events (e.g. the possibility of a dragon), but since I assume that the agent, once aware, completely understands all consequences and the probability distribution of a contingency, the restriction to "true" contingencies seems to be much more natural.

Acknowledgements

I am very grateful to my supervisor Piero Gottardi. I would also like to thank Arpad Abraham, Wouter Dessein, Erik Eyster, Leonardo Felli, Paul Heidhues, Aviad Heifetz, Martin Meier, Salvatore Modica, Stephen Morris, In-Uck Park, Nicola Pavoni, Michele Piccione, Andrew Postlewaite, Andrea Prat, Luis Santos-Pinto, Balazs Szentes and Jean-Marc Tallon for very helpful comments.

A Appendix

A.1 Theoretical Foundations of the Unawareness Structure

My model adopts the unawareness structure introduced by Heifetz et al. (2006). They propose a generalized state space model that allows for non-trivial unawareness in multi-agent settings and strong properties of knowledge. In Heifetz et al. (2008), the authors provide complete and sound axiomatization for their class of unawareness structures. For ease of exposition, the basic set up in this paper is a strongly simplified version of the original model, foregoing the formal introduction of state spaces, projections among them, events, etc. This section does not present the generalized state space model in detail, but provides some insights on how my basic model is built on the foundational literature.

The unawareness structure proposed by Heifetz et al. (2006/2011) consists of a lattice of disjoint space spaces $\mathcal{S} = \{S_{\widehat{\Theta}}\}_{\widehat{\Theta} \subseteq \Theta}$, with a partial order $\succeq$ on $\mathcal{S}$. $S_{\Theta'} \succeq S_{\Theta''}$ means that $S_{\Theta'}$ is more expressive than $S_{\Theta''}$, so the spaces are ordered according to their richness in terms of facts that they can describe. The upmost state space S_Θ is interpreted as the objective state space. Let $\Omega := \bigcup_{\widehat{\Theta} \subseteq \Theta} S_{\widehat{\Theta}}$ denote the union of all state spaces with typical element ω. Any $\omega \in S_{\widehat{\Theta}}$ can be interpreted as a vector of realizations of the random variables in $\widehat{\Theta}$, the starting point of the basic model.

For any $S_{\Theta'}, S_{\Theta''} \in \mathcal{S}$ such that $S_{\Theta'} \succeq S_{\Theta''}$, there is a surjective projection $r^{S_{\Theta'}}_{S_{\Theta''}} : S_{\Theta'} \to S_{\Theta''}$, where $r^{S_{\Theta'}}_{S_{\Theta''}}(\omega)$ is the restricted description of $\omega \in S_{\Theta'}$ in the limited vocabulary of $S_{\Theta''}$. Let $g(S_{\Theta''}) = \{S_{\Theta'} \in \mathcal{S} : S_{\Theta'} \succeq S_{\Theta''}\}$ denote the set of state spaces that are at least as expressive as $S_{\Theta''}$. Further, given a set of states $D \subseteq S_{\Theta''}$, let $D^\uparrow = \bigcup_{S_{\Theta'} \in g(S_{\Theta''})} (r^{S_{\Theta'}}_{S_{\Theta''}})^{-1}(D)$ denote all $\omega \in \Omega$ that describe D in at least as expressive vocabulary as $S_{\Theta''}$. Then an event is a pair $(D^\uparrow, S_{\Theta''})$ with $D \subseteq S_{\Theta''}$ and $S_{\Theta''} \in \mathcal{S}$. Back to the basic model an elementary event that some proposition $\theta \in \Theta$ is true is denoted by $(\omega^\uparrow, S_\theta)$, where ω is the state in S_θ in which θ is true (note that S_θ has two elements). Consequently, $\omega^\uparrow$ is the set of states in Ω where the proposition θ is expressible and true, $(S_\theta \backslash \omega)^\uparrow$ is the set of states in Ω where proposition θ is expressible and false and $\Omega \backslash \{\omega^\uparrow \cup (S \backslash \omega)^\uparrow\}$ is the set of states where neither the event of θ being true nor false is expressible. If the agent is aware of $\widehat{\Theta}$ and $\theta \notin \widehat{\Theta}$ than $S_{\widehat{\Theta}} \subset \Omega \backslash \{\omega^\uparrow \cup (S \backslash \omega)^\uparrow\}$.

Similar to the basic model, define $y = f(\omega), \omega \in S_\Theta$, where f is a one-to-one function. The agent cannot express S_Θ and consequently cannot know function f. It is assumed that the agent's perceived outcome is equal to the actual outcome when no unforeseen contingency is realized. Let $(\omega_{\theta=0}{}^{\uparrow}, S_\theta), \omega_{\theta=0} \in S_\theta$ denote the event that $\theta = 0$. Then the agent's perceived outcome function is defined by

$$\widehat{f}(\omega) := f\left(\left\{\left(r^{S_\Theta}_{S_{\widehat{\Theta}}}\right)^{-1}(\omega)\right\} \cap \left\{\bigcap_{\theta \in \Theta \setminus \widehat{\Theta}} (\omega_{\theta=0})^{\uparrow}\right\}\right), \quad \omega \in S_{\widehat{\Theta}}.$$

It is easy to check that this coincides with the original specification of the outcome function.

In Heifetz et al. (2011) the generalized state space model is augmented by probabilistic beliefs. Let μ denote a probability measure on S_Θ. Then the marginal $\widehat{\mu}$ of μ on $S_{\widehat{\Theta}}$ is defined by

$$\widehat{\mu}(\omega) := \mu\left(\left(r^{S_\Theta}_{S_{\widehat{\Theta}}}\right)^{-1}(\omega)\right), \quad \omega \in S_{\widehat{\Theta}}.$$

In the basic set up we are interested in conditional probabilities on effort, so define $\mu(\omega|e), \omega \in S_\Theta$ and $\widehat{\mu}(\omega|e), \omega \in S_{\widehat{\Theta}}$ analogously. Then let $\pi(y|e) := \mu\left(f^{-1}(y)|e\right)$ and $\widehat{\pi}(y) = \widehat{\mu}\left(\widehat{f}^{-1}(y)|e\right)$. Under the assumption of independence, this yields $\pi(y|e) = \Pi(\widehat{\Theta}|e)\widehat{\pi}(y|e)$, the original definition in the model.

A.2 Proofs

A.2.1 Proof of Proposition 4.1

Suppose $\Theta_A \subset \widehat{\Theta}^*$. The optimal compensation scheme is $\widehat{C}^*(y) = v^{-1}(\bar{U} + e^H)$ for all $y \in \widehat{\mathcal{Y}}$ and $\widehat{C}^*(y) = 0$ for all $y \in \mathcal{Y} \setminus \widehat{\mathcal{Y}}$. Then the expected payment is

$$\prod_{\theta \in \Theta \setminus \widehat{\Theta}^*} \Pr[\theta = 0|e^H] v^{-1}(\bar{U} + e^H),$$

which is clearly greater than $\prod_{\theta \in \Theta \setminus \Theta_A} \Pr[\theta = 0|e^H] v^{-1}(\bar{U} + e^H)$, the expected payment under $\widehat{\Theta}^* = \Theta_A$, due to the assumption $\pi(y|e) > 0, \forall y \in \mathcal{Y}$. Hence $\widehat{\Theta}^*$ cannot be optimal. ∎

A.2.2 Proof of Lemma 5.1

Suppose $\lambda = 0$. Since $\sum_{y \in \widehat{\mathcal{Y}}} \widehat{\pi}(y|e^H) = \sum_{y \in \widehat{\mathcal{Y}}} \widehat{\pi}(y|e^L) = 1$ and $\widehat{\pi}(.|e^H) \neq \widehat{\pi}(.|e^L)$ there must exist some $y \in \widehat{\mathcal{Y}}$ such that $\widehat{\pi}(y|e^H) - \widehat{\pi}(y|e^L) < 0$. But since $\gamma \geq 0$, $\lambda = 0$ would imply that $\frac{1}{v'(C(y))} \leq 0$

for some $y \in \widehat{\mathcal{Y}}$, which violates the assumption $\upsilon'(\cdot) > 0$. Hence $\lambda > 0$.

Now suppose $\gamma = 0$. Then, the first-order conditions of the optimization problem imply that compensation is fixed across outcomes within the agent's awareness. But this implies that the incentive constraint is no longer satisfied. Hence, $\gamma > 0$. ∎

A.2.3 Proof of Lemma 5.2

Let $\tilde{S}$ denote the state space, let $\tilde{y}$ denote the output function with range $\tilde{\mathcal{Y}}$ and let $\tilde{\pi}(.|e)$ denote the probability belief given awareness $\widehat{\Theta} \cup Z$. Further, let $\rho : \tilde{\mathcal{Y}} \longrightarrow \widehat{\mathcal{Y}}$ be the mapping from set $\tilde{\mathcal{Y}}$ to set $\widehat{\mathcal{Y}}$, where $\rho(\tilde{y}(\tilde{s})) = \widehat{y}(\widehat{s})$ and $\widehat{s} \in \widehat{S}$ is a subsequence of $\tilde{s} \in \tilde{S}$. Now consider the compensation scheme $\tilde{C}$ with $\tilde{C}(y) = C^C_{\widehat{\Theta}}(\rho(y))$ for all $y \in \tilde{\mathcal{Y}}$. Note that $\tilde{C}$ satisfies both participation and incentive constraint with equality.

Suppose $\Pr[\theta = 1|e^H] = \Pr[\theta = 1|e^L], \forall \theta \in Z$. Then, for any $y \in \tilde{\mathcal{Y}}$ we have

$$\frac{\tilde{\pi}(y|e^L)}{\tilde{\pi}(y|e^H)} = \frac{\widehat{\pi}(y|e^L)}{\widehat{\pi}(y|e^H)}.$$

Hence, $C^C_{\widehat{\Theta}}$ satisfies the first order conditions and consequently solves the optimization problem. $\Delta C^Z_{\widehat{\Theta}} = 0$.

Now, suppose $\Pr[\theta = 1|e^H] \neq \Pr[\theta = 1|e^L]$ for some $\theta \in Z$. Then, there must exist some $y, y' \in \rho^{-1}(y), y \in \widehat{\mathcal{Y}}$ such that

$$\frac{\tilde{\pi}(y|e^L)}{\tilde{\pi}(y|e^H)} \neq \frac{\tilde{\pi}(y'|e^L)}{\tilde{\pi}(y'|e^H)}.$$

Consequently, $C^C_{\widehat{\Theta}}$ does not satisfy the first-order conditions. Hence, $C^C_{\widehat{\Theta}}$ is feasible but not optimal, which implies that $E[C^C_{\widehat{\Theta} \cup Z}(Y|e^H)] < E[C^C_{\widehat{\Theta}}(Y|e^H)]$ and $\Delta C^Z_{\widehat{\Theta}} > 0$. ∎

A.2.4 Proof of Lemma 5.3

W.l.o.g. we can assume $\Pr[\theta = 1|e^H] > \Pr[\theta = 1|e^L], \theta \in \widehat{\Theta}$. $C^C_{\widehat{\Theta}}$ is the compensation scheme that solves the principal's optimization problem given full awareness and information structure $\widehat{\Theta}$. Let $E_{\widehat{\Theta}}$ denote the expectation operator with respect to awareness $\widehat{\Theta}$. Then we have:

$$E_{\widehat{\Theta}}\left[\upsilon(C^C_{\widehat{\Theta}}(Y))|e, \theta = 1\right] > E_{\widehat{\Theta}}\left[\upsilon(C^C_{\widehat{\Theta}}(Y))|e, \theta = 0\right], \quad e = e^L, e^H,$$

which follows directly from the first-order conditions. The incentive constraint can be rewritten as

$$\begin{aligned} & \Pr[\theta=1|e^H]E_{\widehat{\Theta}}\left[\upsilon(C^C_{\widehat{\Theta}}(Y))|e^H,\theta=1\right]+\Pr[\theta=0|e^H]E_{\widehat{\Theta}}\left[\upsilon(C^C_{\widehat{\Theta}}(Y))|e^H,\theta=0\right]-e^H \\ = \;& \Pr[\theta=1|e^L]E_{\widehat{\Theta}}\left[\upsilon(C^C_{\widehat{\Theta}}(Y))|e^L,\theta=1\right]+\Pr[\theta=0|e^L]E_{\widehat{\Theta}}\left[\upsilon(C^C_{\widehat{\Theta}}(Y))|e^L,\theta=0\right]-e^L. \end{aligned}$$

Now, consider probability $\Pr[\theta = 1|e^L] - \varepsilon$ with $\varepsilon > 0$. Under $C^C_{\widehat{\Theta}}$ and $\Pr[\theta = 1|e^L] - \varepsilon$ the participation constraint is clearly satisfied with equality. Looking at the incentive constraint it is easy to see that

$$\begin{aligned} & \Pr[\theta=1|e^H]E_{\widehat{\Theta}}\left[\upsilon(C^C_{\widehat{\Theta}}(Y))|e^H,\theta=1\right]+\Pr[\theta=0|e^H]E_{\widehat{\Theta}}\left[\upsilon(C^C_{\widehat{\Theta}}(Y))|e^H,\theta=0\right]-e^H \\ > \;& \Pr[\theta=1|e^L]E_{\widehat{\Theta}}\left[\upsilon(C^C_{\widehat{\Theta}}(Y))|e^L,\theta=1\right]+\Pr[\theta=0|e^L]E_{\widehat{\Theta}}\left[\upsilon(C^C_{\widehat{\Theta}}(Y))|e^L,\theta=0\right]-e^L \\ & -\varepsilon\left(E_{\widehat{\Theta}}\left[\upsilon(C^C_{\widehat{\Theta}}(Y))|e^L,\theta=1\right]-E_{\widehat{\Theta}}\left[\upsilon(C^C_{\widehat{\Theta}}(Y))|e^L,\theta=0\right]\right). \end{aligned}$$

We know that under the optimal transfer rule both constraints are satisfied with equality. Consequently, given $\Pr[\theta = 1|e^L] - \varepsilon$, $C^C_{\widehat{\Theta}}$ is feasible but not optimal, so we have $V_{\widehat{\Theta}}(\Pr[\theta = 1|e^L]) > V_{\widehat{\Theta}}(\Pr[\theta = 1|e^L] - \varepsilon)$. The same line of reasoning applies to $\Pr[\theta = 1|e^H]$. In that case both constraints are slack. Hence, the $V_{\widehat{\Theta}}$ is decreasing in Δ^θ for all $\theta \in \widehat{\Theta}$. ∎

A.2.5 Proof of Lemma 5.5

$$NG(\theta_n)-NG(\theta_{n+1})=\alpha^{N-n-1}\left(\alpha^2E[C^C_{\widehat{\Theta}_{n-1}}(Y)|e^H]-2\alpha E[C^C_{\widehat{\Theta}_n}(Y)|e^H]+E[C^C_{\widehat{\Theta}_{n+1}}(Y)|e^H]\right)$$

This function is continuous in α. Moreover, as α goes to zero $NG(\theta_n)-NG(\theta_{n+1})$ is strictly positive. Now, if $NG(\theta_n)-NG(\theta_{n+1}) < 0$ for some α, there has to be an α such that $NG(\theta_n)-NG(\theta_{n+1}) = 0$ by the *Intermediate Value Theorem*. Solving for α we have

$$\alpha_{1/2}=\frac{2E[C^C_{\widehat{\Theta}_n}(Y)|e^H]\pm\sqrt{4E[C^C_{\widehat{\Theta}_n}(Y)|e^H]^2-4E[C^C_{\widehat{\Theta}_{n-1}}(Y)|e^H]E[C^C_{\widehat{\Theta}_{n+1}}(Y)|e^H]}}{2E[C^C_{\widehat{\Theta}_{n-1}}(Y)|e^H]}.$$

If the discriminant $4(E[C^C_{\widehat{\Theta}_n}(Y)|e^H]^2-E[C^C_{\widehat{\Theta}_{n-1}}(Y)|e^H]E[C^C_{\widehat{\Theta}_{n+1}}(Y)|e^H]) < 0$ there are no real roots. Hence, whenever $E[C^C_{\widehat{\Theta}_n}(Y)|e^H]^2 < E[C^C_{\widehat{\Theta}_{n-1}}(Y)|e^H]E[C^C_{\widehat{\Theta}_{n+1}}(Y)|e^H]$, $NG(\theta_n) - NG(\theta_{n+1}) > 0$ for all $\alpha \in (0, 1)$. ∎

A.2.6 Proof of Proposition 5.7

It is clear that whenever $EU^A_P(\widehat{\Theta}^*, \widehat{C}^*, e^H) < 0$, the agent thinks that the principal would be strictly better off by not offering the contract. Similarly if $EU^A_P(\widehat{\Theta}^*, \widehat{C}^*, e^H) \geq EU^A_P(\widehat{\Theta}^*, \widehat{C}^*, e^L)$ the agent cannot rationalize why the principal proposes an incentive compatible contract. Hence, $(\widehat{\Theta}^*, \widehat{C}^*)$

cannot be justifiable.

Now suppose $EU_P^A(\widehat{\Theta}^*, \widehat{C}^*, e^H) \geq 0$ and $EU_P^A(\widehat{\Theta}^*, \widehat{C}^*, e^H) \geq EU_P^A(\widehat{\Theta}^*, \widehat{C}^*, e^L)$.

Justifiability of $\widehat{C}^*$*:* If $\widehat{\Theta}^* = \Theta$, principal and agent share the same beliefs. Hence, $\widehat{C}^*$ maximizes the principal's expected payoff from the agent's perspective. If $\widehat{\Theta}^* \neq \Theta$, $\widehat{C}^*(y) = C^C_{\widehat{\Theta}^*}(y)$ for all $y \in \widehat{\mathcal{Y}}$. Since the agent thinks that the contract is complete, $\widehat{C}^*$ maximizes the principal's payoff according to the beliefs of the agent.

Justifiability of $\widehat{\Theta}^*$*:* $\widehat{\Theta}^*$ is optimal for the principal given the agent's beliefs if

$$E\left[C^C_{\widehat{\Theta}^*}(Y|e^H)\right] \leq \prod_{\theta \in \widehat{\Theta}^* \setminus Z} \Pr[\theta = 0|e^H] E\left[C^C_{\widehat{\Theta}^* \setminus Z}(Y)|e^H\right],$$

for any $Z \subseteq \widehat{\Theta}^* \setminus \Theta_A$. This coincides with the optimality condition of the principal. Hence, whenever $EU_P^A(\widehat{\Theta}^*, \widehat{C}^*, e^H) \geq 0$ and $EU_P^A(\widehat{\Theta}^*, \widehat{C}^*, e^H) \geq EU_P^A(\widehat{\Theta}^*, \widehat{C}^*, e^L)$, $(\widehat{\Theta}^*, \widehat{C}^*)$ is justifiable. ∎

A.2.7 Proof of Proposition 6.1

Existence: Consider a deviation of principal i. The strategies are $\widehat{\Theta}_j = \Theta$ and $C_j = C^*$ for all $j = 1, ..., i-1, i+1, ..., N$. A deviation in $\widehat{\Theta}_i$ clearly leaves the expected payoff unaffected because the agent is aware of $\widehat{\Theta}_1 \cup ... \cup \widehat{\Theta}_N = \Theta$ for all $\widehat{\Theta}_i \subseteq \Theta$. A deviation in C_i is not profitable for the standard Bertrand argument. C^* maximizes the agent's expected utility given the zero outside option of the principal. A deviation $C_i \neq C^*$ must make either the agent or the principal worse off. If the agent is worse off, he rejects the contract and the expected payoff is zero. If the principal is worse off, he has a negative expected payoff. Hence, deviating in $(\widehat{\Theta}_i, C_i)$ is not profitable.

Uniqueness: It is easy to see that this is the only symmetric Nash equilibrium. If $\widehat{\Theta}_i = \Theta, i = 1, ..., N$ principals engage in a standard Bertrand competition. The unique symmetric equilibrium is $C_i = C^*$ for all $i = 1, ..., N$. Now suppose $\widehat{\Theta}_i = \tilde{\Theta} \subset \Theta, i = 1, ..., N$. Given $\tilde{\Theta}$ principals engage in a Bertrand competition over the compensation rule. But as profits become small enough it is profitable to announce a contingency in $\Theta \setminus \tilde{\Theta}$. This deviation allows principal i to capture the whole market and make positive profits, because the agent realizes that he receives zero in some states if he accepts one of the other contracts. Hence, $\tilde{\Theta}$ cannot be an equilibrium announcement and the unique symmetric Nash equilibrium is $(\widehat{\Theta}_i, C_i) = (\Theta, C^*), i = 1, ..., N$. ∎

A.2.8 Proof of Proposition 6.2

Suppose $\widehat{\Theta}_i = \widehat{\Theta} \subset \Theta, i = 1, ..., N$. Let $\bar{C}$ denote the solution to the problem

$$\max_{C} \sum_{y \in \widehat{\mathcal{Y}}} \widehat{\pi}(y|e^H) \upsilon\left(C\left(y\right)\right) - e^H$$

subject to

$$\sum_{y \in \widehat{\mathcal{Y}}} \widehat{\pi}(y|e)\left[y - C(y)\right] \geq 0$$

$$e^H \in \arg\max_{e} \left\{ \sum_{y \in \widehat{\mathcal{Y}}} \widehat{\pi}(y|e) \upsilon\left(C\left(y\right)\right) - e \right\},$$

and let $\widehat{C}(y) = \bar{C}(y), \forall y \in \widehat{\mathcal{Y}}$ and $\widehat{C}(y) = 0, \forall y \in \mathcal{Y} \backslash \widehat{\mathcal{Y}}$. Then, principal i's expected profit is $EU_i(\widehat{\Theta}, \widehat{C}) = \frac{1}{N} \sum_{y \in \mathcal{Y}} \pi(y|e) \left[y - \widehat{C}(y)\right]$. Since $\widehat{C}$ is independent of N, the expected profit is decreasing in N. If i announces $\theta \in \Theta \backslash \widehat{\Theta}$, he offers a compensation scheme $\tilde{C}$, which solves the following optimization problem

$$\max_{C} \sum_{y \in \mathcal{Y}} \pi(y|e)\left[y - C(y)\right]$$

subject to

$$\sum_{y \in \tilde{\mathcal{Y}}} \tilde{\pi}(y|e^H) \upsilon\left(C(y)\right) \geq \sum_{y \in \tilde{\mathcal{Y}}} \tilde{\pi}(y|e^H) \upsilon\left(\widehat{C}(y)\right)$$

$$e^H \in \arg\max_{e} \left\{ \sum_{y \in \tilde{\mathcal{Y}}} \tilde{\pi}(y|e) \upsilon\left(C\left(y\right)\right) - e \right\},$$

where $\tilde{y}$ is the agent's perceived outcome function with range $\tilde{\mathcal{Y}}$ and probability belief $\tilde{\pi}(.|e)$ associated to set $\tilde{\Theta} := \widehat{\Theta} \cup \theta$. Principal i's expected profit is $EU_i(\tilde{\Theta}, \tilde{C}) = \sum_{y \in \mathcal{Y}} \pi(y|e) \left[y - \tilde{C}(y)\right]$, which is independent of N. Hence, there always exists an $\bar{N}$ such that

$$\frac{1}{N} \sum_{y \in \mathcal{Y}} \pi(y|e) \left[y - \widehat{C}(y)\right] < \sum_{y \in \mathcal{Y}} \pi(y|e) \left[y - \tilde{C}(y)\right]$$

for all $N \geq \bar{N}$. ∎

A.2.9 Proof of Proposition 7.1

When e^L is optimal, the principal solves

$$\max_{\widehat{\Theta}, C(\cdot)} \quad \sum_{y \in \mathcal{Y}} \pi(y|e^L) \left[y - C(y)\right]$$

subject to

$$\sum_{y \in \widehat{\mathcal{Y}}} \widehat{\pi}(y|e^L) \upsilon \left(C\left(y\right)\right) - e^L \geq \overline{U}$$

$$C(y) \geq 0, \quad \forall y \in \mathcal{Y}.$$

The optimal compensation scheme for a given $\widehat{\Theta}$ is $C^L(y) = \upsilon^{-1}(\bar{U} - e^L), \forall y \in \widehat{\mathcal{Y}}$ and $C^L(y) = 0, \forall y \in \mathcal{Y} \backslash \widehat{\mathcal{Y}}$. The expected payment is

$$E[C^L(Y)] = \prod_{\theta \in \Theta \backslash \widehat{\Theta}} \Pr[\theta = 0|e^L] \upsilon^{-1}(\bar{U} - e^L),$$

which is clearly minimized for $\widehat{\Theta} = \Theta_A$. ∎

References

[1] Board, O.J., Chung, K.S., 2011. Object-Based Unawareness: Theory and Applications. Mimeo.

[2] Dekel, E., Lipman, B., Rustichini, A., 1998. Standard State-Space Models Preclude Unawareness. Econometrica 66(1), 159-173.

[3] Filiz-Ozbay, E., 2012. Incorporating Unawareness into Contract Theory. Games and Economic Behavior, *forthcoming*.

[4] Gabaix, X., Laibson, D., 2006. Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets. Quarterly Journal of Economics 121(2), 505-540.

[5] Galanis, S., 2011. Unawareness of Theorems. Economic Theory, *forthcoming*.

[6] Grossman, S., Hart, O., 1983. An Analysis of the Principal-Agent Problem. Econometrica 51(1), 7-45.

[7] Harris, R., 1978. Ex-Post Efficiency and Resource Allocation under Uncertainty. The Review of Economic Studies 45(3), 427-436.

[8] Heifetz, A., Meier, M., Schipper, B.C., 2006. Interactive Unawareness. Journal of Economic Theory 130(1), 78-94.

[9] Heifetz, A., Meier, M., Schipper, B.C., 2008. A Canonical Model for Interactive unawareness. Games and Economic Behavior, 62, 304-324

[10] Heifetz, A., Meier, M., Schipper, B.C., 2011. Unawareness, Beliefs and Speculative Trade. Mimeo.

[11] Holmström, B., 1979. Moral Hazard and Observability. Bell Journal of Economics 10(1), 74-91.

[12] Li, J., 2008. A Note on Unawareness and Zero Probability. Mimeo.

[13] Li, J., 2009. Information Structures with Unawareness. Journal of Economic Theory 144(3), 977-993.

[14] Modica, S., Rustichini, A. 1994. Awareness and Partitional Information Structures. Theory and Decision 37, 107-124.

[15] Modica, S., Rustichini, A., Tallon, J.M., 1998. Unawareness and Bankruptcy: A General Equilibrium Model. Economic Theory 12(2), 259-292.

[16] Ozbay, E., 2008. Unawareness and Strategic Announcements in Games with Uncertainty. Mimeo.

[17] De la Rosa, L.E., 2011. Overconfidence and Moral Hazard. Games and Economic Behavior 73(2), 429-451.

[18] Santos-Pinto, L., 2008. Positive Self-Image and Incentives in Organizations. Economic Journal 118(531), 1315-1332.

[19] Schipper, B.C., 2012. Awareness-Dependent Subjective Expected Utility. International Journal of Game Theory, *forthcoming*.

[20] Von Thadden, E.L., Zhao, X.J., 2011. Multi-Task Agency with Unawareness. Mimeo.

[21] Von Thadden, E.L., Zhao, X.J., 2012. Incentives for Unaware Agents. Review of Economic Studies, *forthcoming*.

[22] Zhao, X.J., 2008. Moral Hazard with Unawareness. Rationality and Society 20(4), 471-496.

European
University
Institute
European University Press