

# Incentives for Unaware Agents

ERNST-LUDWIG VON THADDEN

*University of Mannheim and CEPR*

and

XIAOJIAN ZHAO

*Hong Kong University of Science and Technology*

*First version received February 2009; final version accepted October 2011 (Eds.)*

The paper introduces the problem of unawareness into principal–agent theory and discusses optimal incentive contracts when the agent may be unaware of her action space. Depending on the agent’s default behaviour, it can be optimal for the principal to propose an incomplete contract (that keeps the agent unaware) or a complete contract. The key trade-off is that of enlarging the agent’s choice set versus adding costly incentive constraints. If agents differ in their unawareness, optimal contracts show a self-reinforcing pattern: if there are few unaware agents in the economy optimal contracts promote awareness, if unawareness is wide spread optimal contracts shroud the contracting environment, thus keeping the agent unaware.

*Key words:* Moral hazard, Screening, Incomplete contracts, Unawareness

*JEL Codes:* D01, D86, D82, D83

## 1. INTRODUCTION

The classical model of moral hazard between a principal and an agent as developed by Mirrlees (1975), Holmström (1979), and Grossman and Hart (1983) assumes that the agent takes an unobservable action for which a properly designed contract provides the right incentives. However, as Milgrom and Roberts (1992) have put it: “Explicit incentive contracts are not nearly as widely used in employment relations as simple theory might lead one to expect” (p. 392). In response to this observation, a growing literature has argued that high-powered incentives are potentially counter-productive because they can crowd out intrinsic motivation, can lead to communication problems, or hamper cooperation.<sup>1</sup>

We address the problem of incentive contracting from a new angle. Our goal is twofold. First, we want to introduce a concept into contract theory that has received wide attention in sociology and psychology, namely that human activity often follows heuristics and is rule guided (Vanberg, 2002; Fiske and Taylor, 2007). Second, we study how such rule-guided behaviour interacts with classical optimizing behaviour and what implications this has for incentive contracting. We therefore assume that agents may simply be unaware of the full effort problem that they are facing and provide a certain default level of effort if not instructed and motivated otherwise. At the individual level, some level of cleanliness, friendliness, or diligence may be simply part of an employee’s character or be part of a social norm and therefore does not need to be stimulated by incentives. People have work routines that they follow because of habit or because

1. See, e.g. Falk and Fehr (1999) or Bénabou and Tirole (2003).

they have been trained this way, such as an engineer who favours technical solutions over marketing approaches even if this is inefficient or an employee who has been brought up to keep her tools in good condition even if not rewarded for it. Making agents aware of their routines, biases, and habits and then make them change them typically requires training and explicit incentives. In the human resources literature, for instance, [Baron and Kreps \(1999\)](#) argue that “if a task is not formally recognized in a worker’s incentive pay, he or she has less incentive to pay attention to that task” (p. 269).

From the point of view of the employer, what is the trade-off between rule-guided behaviour that is not subject to contractual rewards and compensation-driven behaviour that relies on explicit financial incentives? We address this question in the framework of the standard principal–agent model at a fairly general level. When the agent is unaware of the effort problem and operates by default, it may not be optimal for the principal to specify the agent’s actions contractually and regulate them by means of incentive–compatibility constraints. “Incomplete contracts” might be a better alternative for the principal. For instance, if the employer knows that the employee is unaware of some shirking behaviour, then it may not be optimal for the employer to regulate this type of activity in the contract since this makes the employee aware of the activity and necessitates the provision of costly incentives.

We first discuss the contracting problem under the assumption that the principal knows that the agent is unaware of the effort problem and therefore takes an unconscious default action. It is easy to see that if this default action is sufficiently close to the first best level, the principal will write an incomplete contract where the description of the agent’s action is missing. If the agent’s default behaviour is sufficiently far from the first best level, the principal will optimally make the agent aware. The incompleteness economizes on the costs of incentive provision that arises if the principal wants to regulate the agent’s behaviour explicitly. We thus identify a new trade-off in the principal–agent problem: the benefit of enlarging the agent’s choice set versus the cost of adding an incentive constraint.

We then extend the analysis to an environment with heterogeneous awareness of agents, where the principal cannot distinguish whether the agent is aware or not. In such an environment, the contract that is optimal for an unaware agent is typically not viable because aware agents will exploit its low pay-performance sensitivity. Thus, the principal has to screen agents. The problem is similar to traditional screening problems (see, *e.g.* [Bolton and Dewatripont, 2005](#), and [Laffont and Martimort, 2002](#)) but has less structure in the constraints. The biggest difference is that in our problem, the extent of unawareness (the population mix) is not exogenous because the principal has the option to make the agent aware of the effort problem by proposing a complete contract with explicit performance targets. Yet, despite this difficulty, the comparative statics of the optimal contract are surprisingly simple: in populations with a relatively large extent of unawareness optimal contracts will be incomplete, thus preserving unawareness. On the other hand, in populations with a relatively small extent of unawareness, contracts will be complete and eliminate the remaining unawareness. This result is similar to that of [Gabaix and Laibson \(2006\)](#) that firms strategically shroud expensive attributes of their products if and only if there are enough unaware consumers in the market. For organizations, this result can explain the observation in the human resources literature (*e.g.* [Lawler, 1990](#)) that there are corporate cultures with very little emphasis on performance pay and those that make heavy use of it (such as the famous Lincoln electric case<sup>2</sup>).

In fact, a central insight of the literature on organizational behaviour is that organizational culture can “let employees internalize organizational values and norms and let these values

2. See [Milgrom and Roberts \(1992\)](#), p. 393.

and norms guide their decisions and actions” (Meyer *et al.*, 2007, p. 370). In such an environment, explicit supervision and financial incentives play a lesser role than in organizations where agents understand the mechanisms by which individual effort leads to measurable results and where management fosters such effort taking through monetary performance contracts. Our comparative statics result implies that there can be a tipping point in organizations such that if more and more employees realize the costs and benefits of their actions, the organization will switch from largely rule-guided behaviour and low-powered incentives to explicit performance contracting.<sup>3</sup> It can be argued that such a tipping point occurred in the financial industry in the late 1980/1990s, leading to a wide-spread shift to strong performance-based pay thereafter.

A similar observation may apply at the societal level. In developing economies such as China, firms have traditionally shrouded work benefits or career opportunities, which has led to a culture of systemic overprovision of effort. While possibly optimal in the past, there is evidence of rising awareness of work-related malpractice in the population and a change in firm policies.<sup>4</sup>

In order to probe more deeply into the relation between effort incentives, rents, and rule-guided behaviour, we then specialize our model in Section 5 to the case of two outcomes. In this specification, one can characterize the solution of the full screening problem fairly completely and identify several features of the optimal contract that are of interest. The analysis shows that there is a trade-off between the rent conceded to aware agents and the inefficiency in terms of risk exposure of unaware agents and that the optimal resolution of this trade-off depends on the population mix. In particular, when unawareness is wide spread, contracts will be largely performance insensitive and aware agents will obtain large rents. If the share of aware agents increases, this rent decreases, and unaware agents will bear more inefficient performance risk.

The rest of the paper is organized as follows. The next section reviews some of the related literature. Section 3 introduces the general principal–agent model with unawareness. Section 4 derives the basic insights of contract design in the presence of unawareness. Section 5 focusses on the case of two outcomes and also illustrates this problem numerically. Section 6 discusses several extensions and concludes with a number of robustness checks.

## 2. LITERATURE

### 2.1. *Decisions under unawareness*

Following Modica and Rustichini (1994), a growing literature has studied unawareness of events in the standard state space model. Since our work is on agents’ unawareness of their action set, these theories do not bear directly on our work, although we rely conceptually on theories of the unawareness of theorems (Galani, 2009) or interactive unawareness (Heifetz, Meier and Schipper, 2006; Li, 2009). For theories of unawareness of actions, the work by Rego and Halpern (2007), who provide a general setting for studying games with unawareness of actions, is fundamental. The principal–agent model that we discuss in our paper uses a simple dynamic game and fits naturally in their approach.

3. Hendrikse (2003) describes this interaction between individual and organizational behaviour as follows: “Organizations have an effect on human behavior, and human behavior shapes organizations. ... Institutions influence the way in which information is presented, which information is communicated and how the information is interpreted. Customs and routines of organizational members are based on this information. Organizations can therefore also channel human behaviour to a certain extent” (pp. 361–362).

4. A recently famous example is Foxconn, see <http://www.guardian.co.uk/technology/2011/apr/30/apple-chinese-factory-workers-suicides-humiliation>.

Unawareness requires a theory of restricted decision making by agents. As discussed above, we follow Hayek (1967), Mayr (1992), and Vanberg (2002) and others by assuming that unaware agents choose a default action instead of optimizing over (all of) their action space. The default action is an instance of rule-guided or automatic behaviour that is not determined by rational choice. There is ample evidence of such behaviour in the sociological and psychological literature that documents various forms of automatic versus controlled behaviour (see, e.g. Fiske and Taylor, 2007). This creates an exogenous bias in behaviour that we model, greatly simplified, as a constant parameter in our analysis.

### 2.2. *Contracts with unaware agents*

A central question of our work is the interaction between aware and unaware contracting parties. There are difficulties at the conceptual level and in applying the concept. Conceptually, Filiz-Ozbay (2008) and Ozbay (2008) have argued that a satisfactory theory of unawareness should account for the possibility that an agent wonders why observed contracts are as they are. We address this issue of justifiability in Section 6.

In terms of applications, Gabaix and Laibson (2006) analyse the interaction between firms and unaware consumers. Consumers who are unaware of later add-on prices are exploited by the firms. In our paper, the agent is only unaware of her own actions, so there is no issue of exploitation. However, like Gabaix and Laibson (2006), we find that keeping agents unaware is optimal if, and only if, there are sufficiently many unaware agents in the population. Filiz-Ozbay (2008) models the interaction between a rational insurer and an insuree who is unaware of some contingencies. Although the interaction is strategic, unawareness concerns the classical issue of missing future contingencies, which is not the focus of our paper. Closest to our theory in this respect is the work by Eliaz and Spiegel (2006) who study a contract-theoretic model of screening consumers' awareness of their future changed tastes. Eliaz and Spiegel (2011) also study how firms can use marketing devices to manipulate the consumer's perceived choice set when the consumer is unaware of some products and analyse the behavioural implications in the context of a competitive market model. In our paper, the principal is also confronted with agents of different awareness and designs contracts to exploit these differences. But differently from their work, we focus on the provision of incentives and on how the presence of unaware agents affects the incentives of aware agents.

Squintani (2003) considers unawareness of the game form instead of the action space. In his model, principals and agents may not be aware of the possibility of renegotiating the contract after the agent's effort choice. His model is explicitly dynamic and argues that if there are too many unaware agents in the population, principals will offer agents high-powered incentive contracts and possibly renegotiate them later. On the other hand, if there are too few unaware agents, principals will not use incentive contracts at all and rather monitor agents, which, through forgetting, increases the number of unaware agents. These two effects taken together show that the number of unaware agents in the population can be stable over time. This result, albeit obtained in a different framework, is similar to our result about the stability of unawareness under incentive contracting. But since Squintani (2003) does not consider menus of contracts, there is no contractual externality of unaware on aware agents in his model, and thus, no reason for the population to tip towards full awareness if the number of aware agents is too large.

### 2.3. *Incomplete contracts*

Our work also contributes to the literature on the foundations of contract incompleteness. Incomplete contracts traditionally refer to missing contingencies, and the literature has proposed several reasons why contracting parties may not specify everything that is relevant for the interaction

in the contract.<sup>5</sup> Our theory extends the notion of contractual incompleteness from contingencies to actions and endogenizes the incompleteness of contracts with respect to actions. In terms of law and economics, our theory is concerned with the “front end of contracting” rather than the back end (Scott and Triantis, 2006). But in our theory, contract incompleteness arises from a concern with incentive costs rather than drafting costs. This justification of contractual vagueness does not seem to have been studied in the literature on law and economics, and it would be interesting to extend our approach to issues at the “back end of contracting”, such as costly litigation over incomplete contracts as in Scott and Triantis (2006).

Recent research has endogenized contractual incompleteness by limited cognition and strategic investment in cognition by the contracting parties (Bolton and Faure-Grimaud, 2010; Tirole, 2009). These papers take a less radical approach towards unawareness than the literature following Modica and Rustichini (1994) and our paper, as they assume that agents are aware of the fact that they may be unaware of some relevant elements of the contracting environment. Investment in cognition is an interesting extension of our analysis that we discuss in Section 6.

#### 2.4. *Low-powered incentives*

We show that unaware agents should face low-powered incentives because they would not understand the risk associated with outcome-dependent pay. This is consistent with the recent literature in economic psychology and experimental economics that highlights the importance of low-powered incentives and intrinsic motivation. The classical reference in management science here is Dunnette and Hough (1990).

In their classical paper on experimental wage competition, Falk and Fehr (1999) argue that employers shy away from explicit incentive contracting because this erodes reciprocity, which they identify as a key element in employment contracts. Building on Twain (1876), Bénabou and Tirole (2003) model intrinsic motivation as arising from information about one’s aptitude for the job that is transmitted in a principal–agent model with an informed principal. If in such a model, the principal offers high-powered incentives, the agent will infer that the principal believes that the task is difficult for the agent and may consequently refrain from putting too much effort into it. In this sense, extrinsic motivation (incentive pay) can crowd out intrinsic motivation. Bénabou and Tirole (2006) further point out that agents may feel shame if they work harder for high-powered incentives. Similarly, Falk and Kosfeld (2006) argue that restrictions on the agent’s action space or explicit incentives signal the principal’s distrust of the agent, which leads the agent to perform less well. They produce experimental evidence consistent with this interpretation of intrinsic motivation.

Our work provides another psychological foundation of low-powered incentives that is consistent with the experimental results. In our interpretation, agents who would provide a decent default level of effort in a context with no explicit effort incentives become aware of the material side of the effort problem when confronted with monetary incentives and explicit directions. In equilibrium, such behaviour can coexist with incentive-driven effort provision as long as there are sufficiently many unaware agents in the population.

### 3. THE MODEL

In this section, we consider the standard principal–agent problem of Mirrlees (1975), Holmström (1979), and the subsequent literature and modify it in one important respect: unawareness.

5. Most notable are probably arguments invoking verifiability (Grossman and Hart, 1986), signaling (Aghion and Bolton, 1987), and explicit writing costs (Dye, 1985; Anderlini and Felli, 1999).

There are two parties, a risk-neutral principal and a risk-averse agent. The principal proposes a contract to the agent to work for him. The agent's work involves effort, denoted by  $a \in \mathcal{A}$ , where  $\mathcal{A} \subset \mathbb{R}$  is an interval of the real line,<sup>6</sup> and generates a stochastic output  $x \in \mathbb{R}$ . The choice of  $a$  is not observable by the principal, but  $x$  is verifiable. Output is distributed according to a distribution with c.d.f.  $F(\cdot; a)$ . We make the standard assumption that output is produced by decreasing returns to scale:<sup>7</sup>

$$E[x|a] \text{ is increasing and concave in } a. \quad (1)$$

The principal remunerates the agent by a compensation rule  $w(\cdot)$ , where payments satisfy  $w \in (\underline{w}, \bar{w})$ , with  $-\infty \leq \underline{w} < \bar{w} \leq \infty$ .<sup>8</sup> In standard contract theory, a contract is a pair  $(w, a)$ . Although  $a$  is not observable by the principal, it is usually included in the contract; its choice must be supported by an appropriate incentive constraint. Alternatively, the parties can write an incomplete contract  $w$  that induces the agent to choose a certain level of effort by virtue of her incentive constraint. If both parties understand the contracting problem, complete and incomplete contracts are equivalent.

The timing is as follows:

1. The principal proposes a contract.
2. The agent decides whether to accept it. If the agent rejects the contract, the game is over.
3. If the agent accepts the contract, she exerts effort  $a$ .
4. The outcome  $x$  is realized and the contractual compensation is paid.

The principal is risk-neutral with objective function

$$\int (x - w(x)) dF(x, a)$$

if the agent is hired. His outside utility is assumed to be sufficiently low to hire the agent under an optimal contract. The agent's utility is

$$V(w(x)) - C(a),$$

where  $V$  is increasing and strictly concave and  $C$  increasing and strictly convex. Outside the principal-agent relationship, the agent can obtain utility  $\underline{u}$ .<sup>9</sup>

The first best solution to the contracting problem provides full insurance to the agent:  $w(x) = W^{\text{FB}}$  for all  $x$ . The first best effort  $a^{\text{FB}}$  and  $W^{\text{FB}}$  are given by

$$V(W^{\text{FB}}) - C(a^{\text{FB}}) = \underline{u}$$

6. The precise nature of this interval (open or closed, bounded or unbounded) matters for existence problems but is of little interest for our question.

7. This is condition 2.10b in Jewitt (1988). In Jewitt (1988), this is the most innocuous one in a set of assumptions that validate the first-order Approach (FOA). Since we do not need the FOA, we only impose assumption (1). Concavity here is meant in the weak sense, hence the assumption covers the case of additive effort.

8. Consistent with unlimited liability, we assume that transfers lie in an open interval, as is standard in the literature. If  $\underline{w} > -\infty$ , this can lead to a trivial non-existence problem. However, if in this case, the agent's utility  $V(\underline{w})$  is well defined by continuity and finite, we can use transfers  $w \in [\underline{w}, \bar{w})$ .

9. Note that we do not rule out rents by the agent, unlike standard analyses of the insurance-incentive trade-off (such as Grossman and Hart, 1983). However, some of our more detailed analysis becomes significantly simpler if the relevant participation constraints bind. We make this assumption in Propositions 3 and 4, namely that  $V(\underline{w})$  is sufficiently low compared to  $\underline{u}$ , which is implied, e.g. by Grossman and Hart's (1983) unlimited-liability assumption that  $V(\underline{w}) = -\infty$ .

and

$$a^{FB} = \arg \max_a E[x|a] - V^{-1}(C(a) + \underline{u}).$$

We assume that these two equations indeed have a unique solution. To make the problem interesting, we assume that  $a^{FB} \neq \min \mathcal{A}$ .

The innovation in our paper is the assumption that the agent may be unaware of the possibility of choosing  $a$  before contracting. If she is still unaware of  $a$  after contracting, the agent will unconsciously choose the *default* effort or status quo choice  $a = \tau \in \mathcal{A}$ . The agent's choice of the default action is not based on rational calculation,  $\tau$  is rule-guided behaviour (see, e.g. Hayek, 1967, and Vanberg, 2002, in economics and sociology) or program-based behaviour (see, e.g. Mayr, 1992, in biology).

There are two ways of interpreting the notion of unawareness and the implied rule-guided behaviour in our context. The first assumes that the agent is simply unaware of all the activities summarized by  $a$ . This may be the utilization of a certain type of equipment that improves output (in which case the default level  $\tau$  of not using this equipment is inefficiently low) or some form of amenity or perquisite that makes work more pleasant (in which case the default level may be too high or too low).

The second (broader) interpretation of unawareness assumes that the agent is aware of the activities summarized by  $a$  but unaware of their causes and consequences, which Galanis (2009) calls unawareness of theorems. In this case, the agent chooses  $\tau$  according to some habits or routines that do not respond to incentives. Examples for this type of activity are unobservable investments into maintaining equipment or the work environment, the agent's effort in personal customer relations and other forms of personal conduct (where more of these activities correspond to higher  $\tau$ ) or work routines that are cumbersome, slow, or dysfunctional (where more of these activities correspond to lower  $\tau$ ).<sup>10</sup>

We allow for the explicit communication of the agent's choice set prior to contracting. We summarize this communication, which can be informal or take the form of systematic training, by the notion of a complete contract  $(w, a)$  in Stage 1. If the principal communicates the full effort problem, the agent updates her awareness and understands the impact of her effort. On the other hand, if the principal proposes an incomplete contract where  $a$  is missing, the agent remains unaware of it. Thus, if the agent is unaware of the effort problem, complete contracts and incomplete contracts are different instruments.

#### 4. ANALYSIS

##### 4.1. Homogenous awareness

If the agent is aware of the effort problem, the principal faces the standard moral hazard problem:

$$\max_{w(\cdot), a} \int (x - w(x)) dF(x, a) \tag{2}$$

$$\text{s.t. } a \in \arg \max_{a'} \int V(w(x)) dF(x, a') - C(a') \tag{3}$$

$$\int V(w(x)) dF(x, a) - C(a) \geq \underline{u}, \tag{4}$$

10. Note that the agent here is only unaware of the relationship between the costs and benefits of her actions but knows her utility (through experience), although she cannot optimize over the variable of the utility function.

where condition (3) is the incentive-compatibility constraint for the aware agent, and (4) is her participation constraint.

We assume that a solution to this problem exists; it can be obtained by standard methods (see, e.g. Bolton and Dewatripont, 2005, or Laffont and Martimort, 2002). Denote by  $\bar{\pi}^A$  the value of the problem, which is independent of the default effort level  $\tau$ . Furthermore, let  $(\bar{w}^A, \bar{a}^A)$  denote any optimal compensation rule and action (these may not be unique).

If the agent is unaware, we assume that the principal knows  $\tau$ , which should be interpreted as a typical status quo choice taken by unaware agents. Thus, the principal's problem is

$$\max_{w(\cdot)} \int (x - w(x)) dF(x, \tau) \quad (5)$$

$$\text{s.t. } \int V(w(x)) dF(x, \tau) - C(\tau) \geq \underline{u}, \quad (6)$$

where condition (6) is the participation constraint for the unaware agent.<sup>11</sup>

Denote by  $\bar{\pi}^U = \bar{\pi}^U(\tau)$  the value of this problem. Since  $\tau$  is exogenous, there is no effort problem, and the optimal contract provides full insurance:  $\bar{w}^U(x) = W^U$  for all  $x$ . Hence, problem (5)–(6) has a solution if  $\tau$  is sufficiently close to  $a^{FB}$ .

**Lemma 1.** *When the agent is unaware, the principal's profit  $\bar{\pi}^U$  is concave in  $\tau$  with a unique maximum at  $\tau = a^{FB}$ .*

*Proof.* At the optimum, the principal's profit is

$$\begin{aligned} \bar{\pi}^U &= \int (x - W^U) dF(x, \tau) \\ &= E[x|\tau] - W^U. \end{aligned}$$

If  $\underline{u} + C(\tau) > V(\underline{w})$ , the participation constraint (6) binds and we have  $W^U = V^{-1}(\underline{u} + C(\tau))$ , which by assumption is strictly convex in  $\tau$ . If  $\underline{u} + C(\tau) \leq V(\underline{w})$ , then  $W^U = \underline{w}$ . Hence,  $\bar{\pi}^U(\tau) = E[x|\tau] - \max\{\underline{w}, V^{-1}(\underline{u} + C(\tau))\}$ , which is continuous and concave. The second part of the lemma follows because  $\underline{u} + C(a^{FB}) \geq V(\underline{w})$  and the first best is unique.  $\parallel$

Since the principal's profit  $\bar{\pi}^A$  does not depend on  $\tau$  and is strictly smaller than the first best profit, Lemma 1 immediately implies the following proposition.

**Proposition 1.** *There exist  $\tau_{\min} < a^{FB}$  and  $\tau_{\max} > a^{FB}$  such that if the principal knows that the agent is unaware, he optimally proposes an incomplete contract for values  $\tau \in (\tau_{\min}, \tau_{\max})$  and proposes a complete contract otherwise.*

Put simply, Proposition 1 says that if and only if the agent unconsciously is far too lazy or far too diligent in her work, the principal will optimally make her aware of the effort problem and regulate her activity through incentive pay. It is quite plausible that if the agent intrinsically provides very little effort, it is better for the principal to make the agent aware of it and subject her to explicit effort incentives. For instance, if the agent is unaware of using a certain type of technology or behaviour to improve the output, the principal will point this out to the agent and

11. We assume that the aware agent and the unaware agent derive the same utility level from their outside option. In particular, we rule out the possibility that the agent can improve the value of her outside option when being aware of  $a$ .



give her incentives to operate differently. Interestingly, however, even if the agent provides too much effort, say because she focusses too much on certain less productive activities or because she is unaware of certain on-the-job amenities, the principal also has an incentive to make the agent aware of this. The reason is that the agent bears the cost even of the actions she undertakes unconsciously. However, making the agent aware of the effort problem also has a cost: it adds an incentive constraint to the agent's choice problem, with a corresponding reduction of surplus.

Proposition 1 has a simple and general interpretation: *the principal's decision between a complete and an incomplete contract balances the benefit of enlarging the agent's choice set against the cost of adding incentive constraints.*

There are many examples in the management or psychology literature of agents' providing too little or too much effort since they misunderstand or are unaware of some aspects of the agency relationship. Hendrikse (2003) provides an interesting survey of some of these biases and discusses possible corporate responses. In particular, he argues that people often allocate their cognitive resources to only a few areas of interest, while other areas receive little or no attention. For example, a manager with an engineering background may focus excessively on making technologically superior products, whose level of sophistication goes beyond what is demanded by the market. As most individuals are generally slow to react to new information because of excessive confidence in their established assumptions or because of their unawareness of other explanations (Fischhoff, Slovic and Lichtenstein, 1977; Russo and Schoemaker, 1989), the employer has to make such managers actively aware of the right effort choices, if the manager's behaviour is too far off the desired level.

Personal attitudes, such as friendliness, cleanliness, or punctuality, are further examples of effort dimensions that can greatly influence commercial success, in particular in service-related industries. While there are people or cultures for whom such effort appears more natural, others may have to be made aware of these dimensions and given explicit incentives. Firms' policies with respect to mandatory work benefits such as sick leave are another example. Often the firm's policy in this respect is not communicated explicitly to the employees, in the expectation that workers will use the "right amount" of such benefits. This is the case if this "right amount" is sufficiently close to the firm's desired level and if detailed behavioural rules, monitoring strategies, and incentive programs would be too costly.<sup>12</sup>

#### 4.2. Heterogeneous awareness

We now generalize the principal-agent analysis by assuming that the principal does not know whether the agent is aware of the effort problem or not. Formally, we assume that there is a fraction  $\lambda \in [0, 1]$  of agents who are fully aware (type A) and  $1 - \lambda$  of the agents who are unaware (type U), but that the principal cannot distinguish them. To make the problem interesting, we assume that the default effort level of the unaware agent lies in the interval  $(\tau_{\min}, \tau_{\max})$  in which it is in principle better for the principal to keep the unaware agent unaware. Without this assumption, the principal would only offer full-awareness contracts.

We have identified the contracts  $(\bar{w}^A, \bar{a}^A)$  for the aware agent and  $\bar{w}^U$  for the unaware agent that the principal would optimally offer to each of these two types if he knew their type. However, as the following observation shows, if the principal offers both these contracts agents of different types may not self-select:

12. For some interesting background information, see <http://www.enotes.com/small-business-encyclopedia/sick-leave-personal-days>.

**Observation 1.** Assume that  $V$  is unbounded below and the unaware agent provides a default effort larger than the minimum. If the principal proposes the contracts  $(\bar{w}^A, \bar{a}^A)$  and  $\bar{w}^U$ , the aware agent will choose  $\bar{w}^U$ .

Hence, aware agents pretend to be unaware in order to exploit the insurance provided by the  $U$ -contract. To show Observation 1 note that if  $V$  is unbounded below the participation constraint (4) in the aware problem (2)–(4) binds. Hence, under  $(\bar{w}^A, \bar{a}^A)$ , the aware agent receives a payoff of  $\underline{u}$ . But if she chooses  $\bar{w}^U$ , she receives<sup>13</sup>

$$\begin{aligned} & \max_a \int V(W^U) dF(x, a) - C(a) \\ & = \max(V(\underline{w}), \underline{u} + C(\tau)) - \min C(a) \\ & > \underline{u} + C(\tau) - C(\tau) = \underline{u}. \end{aligned}$$

Hence, we need to determine the menu of contracts into which agents select themselves according to their type.<sup>14</sup> Yet, there is a second problem. If the principal proposes a menu of contracts of which one specifies the effort level  $a^A$ , then unaware agents will become aware of  $a$  because  $a$  is explicitly announced in the menu of contracts. But the principal can easily circumvent this problem by proposing two incomplete contracts  $w^A$  and  $w^U$ . As discussed before, there is no conceptual difference between an incomplete and a complete contract in our setting if the agent is aware of the effort problem. The corresponding effort is automatically implied by the agent's optimization given the contracts. From now on we shall therefore only consider menus  $w^A$  and  $w^U$  of incomplete contracts.<sup>15</sup>

The principal's (screening) problem is as follows.

$$\max_{w^A(\cdot), a^A, w^U(\cdot)} \lambda \int (x - w^A(x)) dF(x, a^A) + (1 - \lambda) \int (x - w^U(x)) dF(x, \tau) \quad (\text{OBJ})$$

$$\text{s.t. } a^A \in \arg \max_a \int V(w^A(x)) dF(x, a) - C(a) \quad (\text{ICA})$$

$$\int V(w^A(x)) dF(x, a^A) - C(a^A) \geq \underline{u} \quad (\text{PCA})$$

$$\int V(w^U(x)) dF(x, \tau) - C(\tau) \geq \underline{u} \quad (\text{PCU})$$

$$\int V(w^A(x)) dF(x, a^A) - C(a^A) \geq \max_a \left\{ \int V(w^U(x)) dF(x, a) - C(a) \right\} \quad (\text{ICA-U})$$

$$\int V(w^U(x)) dF(x, \tau) - C(\tau) \geq \int V(w^A(x)) dF(x, \tau) - C(\tau). \quad (\text{ICU-A})$$

Here, (ICA) is the aware agent's incentive constraint for the choice of effort  $a^A$ , given the contract  $w^A$ , (PCA) and (PCU) are the agent's participation constraints for the aware and the

13. If  $C$  does not attain its minimum because of an open-set problem, the argument needs to be modified in the obvious way.

14. One may consider the possibility that the principal selects a subset of the agents, takes them on the side, communicates only to them through a complete contract, and leaves the other agents unaware. Although a procedure like this would extend the set of feasible strategies of the principal, it can never be optimal because the composition of each part of the population is the same.

15. In order to simplify the exposition, we do not insist on technicalities such as measurability problems. Throughout we only consider  $w^U$  that are measurable with respect to  $dF(x, \tau)$  and  $w^A$  that are measurable with respect to  $dF(x, a^A)$ .

unaware type, respectively, and (ICA-U) and (ICU-A) the incentive-compatibility constraints that make sure that the aware and the unaware agent select the appropriate contracts. Note that the unaware agent does not know why the principal proposes the menu in question. Thus, there is no higher order mutual knowledge of the interaction in our model. We will address this question of the justifiability of contracts in Section 6.<sup>16</sup>

Let  $\pi_S(\lambda, \tau)$  be the value of the above screening problem. Note that  $\pi_S(0, \tau) = \bar{\pi}^U$  because the optimal unaware contract  $\bar{w}^U$  can be offered as a pooling contract, which satisfies all constraints of the full problem. However, if  $\lambda = 1$ , it is not necessarily possible to complement the optimal aware contract  $\bar{w}^A$  by an acceptable unaware contract. Hence, a priori  $\pi_S(1, \tau) \leq \bar{\pi}^A$ .

However, the optimal screening solution is not all the principal can achieve because in contrast to classical screening problems, the principal can manipulate  $\lambda$ . Eventually, the principal compares the screening result to the full-awareness outcome since the principal can make all agents aware. Let  $\pi(\lambda, \tau) = \max\{\pi_S(\lambda, \tau), \bar{\pi}^A\}$  be the value of the overall problem. Note that  $\bar{\pi}^A$  is independent of  $\lambda$  and  $\tau$ .

Despite the generality of the problem, the following proposition shows that the principal's incentives to make agents aware have a surprisingly general structure.

**Proposition 2.** *For each  $\tau \in (\tau_{\min}, \tau_{\max})$ , there exists a  $\lambda^*(\tau) \in (0, 1]$  such that the principal leaves unaware agents unaware if  $\lambda < \lambda^*(\tau)$  and makes all agents aware if  $\lambda > \lambda^*(\tau)$ .*

*Proof.* Fix  $\tau \in (\tau_{\min}, \tau_{\max})$ . Denote the objective function of the screening problem (OBJ)-(ICU-A) by

$$h(w^A, a^A, w^U; \lambda) = \lambda \int (x - w^A(x))dF(x, a^A) + (1 - \lambda) \int (x - w^U(x))dF(x, \tau).$$

For each  $(w^A, a^A, w^U)$ ,  $h$  is absolutely continuous and differentiable everywhere in  $\lambda$ . Furthermore, the profit from the aware agent clearly satisfies

$$\int (x - w^A(x))dF(x, a^A) \leq \bar{\pi}^A \tag{7}$$

because the constraint set (ICA)-(ICU-A) includes the constraint set of problem (3)–(4). Similarly,

$$\int (x - w^U(x))dF(x, \tau) \leq \bar{\pi}^U.$$

Hence,

$$\begin{aligned} |h_\lambda(w^A, a^A, w^U; \lambda)| &= \left| \int (x - w^A(x))dF(x, a^A) - \int (x - w^U(x))dF(x, \tau) \right| \\ &\leq \bar{\pi}^A + \bar{\pi}^U \end{aligned}$$

for all  $(w^A, a^A, w^U)$  and  $\lambda$ .

The integral envelope theorem of Milgrom and Segal (2002, Theorem 2) therefore implies that  $\pi_S$  is absolutely continuous in  $\lambda$  and that

$$\frac{d\pi_S(\lambda, \tau)}{d\lambda} = \int (x - w^A(x))dF(x, a^A) - \int (x - w^U(x))dF(x, \tau) \tag{8}$$

16. The revelation principle does not apply in our context since the unaware agent does not know her type. Because the principal wants to keep the unaware agent unaware, he will not use direct revelation mechanisms.

whenever the derivative exists. Equation (8) is the difference between the principal's profit from the aware agent and that from the unaware agent.

Consider any  $\lambda$  such that  $\frac{d\pi_S(\lambda, \tau)}{d\lambda} > 0$ . By (8), the latter inequality implies that the principal makes strictly larger profits on aware agents than on unaware ones. Hence, the optimal values of  $(w^A, a^A, w^U)$  satisfy

$$h(w^A, a^A, w^U; \lambda) < \int (x - w^A(x)) dF(x, a^A).$$

Combining this with equation (7) yields  $\pi_S(\lambda, \tau) < \bar{\pi}^A$ .

Since  $\pi_S$  is continuous in  $\lambda$ , so is  $\pi$ . Hence, the preceding argument implies that if there is a  $\hat{\lambda}$  such that the principal optimally chooses the full-awareness contract, then he will do so for all  $\lambda > \hat{\lambda}$ . This proves the existence of a threshold  $\lambda^*(\tau) \in [0, 1]$ . Since  $\tau \in (\tau_{\min}, \tau_{\max})$ , we have  $\bar{\pi}^U > \bar{\pi}^A$ . Since  $\pi_S(0, \tau) = \bar{\pi}^U$ , the continuity of  $\pi_S$  implies  $\lambda^*(\tau) > 0$ .  $\parallel$

Proposition 2 states that in populations with a relatively large extent of unawareness, contracts will be incomplete, thus preserving unawareness. On the other hand, in populations with a relatively small extent of unawareness, it can be optimal to eliminate the remaining unawareness and offer complete contracts to all agents. What is “large” and what “small” depends on the effect of unawareness as summarized by  $\tau$ .

The proposition thus identifies a self-reinforcing pattern. This is similar to the finding in Gabaix and Laibson (2006) where the shrouding of product attributes is shown to be optimal if there are many unaware consumers in the market. In their model of competitive markets, this happens because educating unaware consumers allows them to profit from the firms' otherwise competitive pricing, thus exposing firms to losses if there are many unaware consumers. In our model, the reason is very different: shrouding is optimal because it economizes on the costs of providing effort incentives. If the principal offers incomplete contracts (*i.e.* shrouds the effort problem), he can utilize the relatively efficient default effort level of the unaware agents at two sorts of costs. First, the unaware agents do not do exactly what the principal would like them to do, and second, the principal cannot incentivize the aware agents as much as he likes because then these agents would pretend to be unaware. If the proportion of aware agents is sufficiently small, this latter cost is small, and the principal prefers incomplete contracts. If this group is large, it can become optimal to make all agents aware and use the corresponding incentives.

Proposition 2 leaves open the question whether this latter option is actually used. The next proposition shows that in a large class of examples, this is indeed the case.

**Proposition 3.** *Suppose that the participation constraint in the full-awareness problem (2)-(4) binds. If  $\tau \in (\tau_{\min}, \tau_{\max})$  and  $\tau \neq \bar{a}^A$ , then  $\lambda^*(\tau) < 1$ .*

*Proof.* First, suppose that  $\pi_S(1, \tau) < \bar{\pi}^A$ . Since  $\tau \in (\tau_{\min}, \tau_{\max})$  implies that  $\pi_S(0, \tau) = \bar{\pi}^U > \bar{\pi}^A$ , the continuity of  $\pi_S$  implies  $\lambda^*(\tau) < 1$ .

Now suppose that  $\pi_S(1, \tau) = \bar{\pi}^A$ . For  $\lambda = 1$ , the objective function (OBJ) is the same as (2). Hence, if  $(\bar{a}, \bar{w}^A, \bar{w}^U)$  is a solution to the screening problem (OBJ)-(ICU-A) at  $\lambda = 1$ ,  $(\bar{a}, \bar{w}^A)$  solves the full-awareness problem (2)-(4). By assumption, in this problem, (PCA) binds. Together with (ICA-U), this implies

$$\max_a \left\{ \int V(\bar{w}^U(x)) dF(x, a) - C(a) \right\} \leq \underline{u}. \quad (9)$$

By (PCU), the maximum in equation (9) is attained at  $a = \tau$  and (PCU) must bind

$$\max_a \left\{ \int V(\bar{w}^U(x))dF(x, a) - C(a) \right\} = \int V(\bar{w}^U(x))dF(x, \tau) - C(\tau) = \underline{u}. \tag{10}$$

The first equality of equation (10) means that  $\bar{w}^U$  implements  $\tau$  in the full-awareness problem. Since by assumption,  $\tau$  is not part of a solution to the full awareness problem, the profit the principal gets from implementing  $\tau$  is strictly lower than  $\bar{\pi}^A$ .

On the other hand, this profit is exactly the profit the principal makes on the unaware agent in the screening problem at  $\lambda = 1$ . Hence, at the optimum of the screening problem, the profit the principal gets from unaware agents is strictly less than that from aware agents unless  $\tau = \bar{a}^A$ .

Now the differential envelope theorem of Milgrom and Segal (2002, Theorem 3) implies that  $\pi_S(\cdot, \tau)$  is left-hand differentiable at  $\lambda = 1$  and that

$$\frac{d\pi_S(1-, \tau)}{d\lambda} = \int (x - \bar{w}^A(x))dF(x, \bar{a}^A) - \int (x - \bar{w}^U(x))dF(x, \tau).$$

By the previous argument, this is strictly positive if  $\tau \neq \bar{a}^A$ , which implies  $\lambda^*(\tau) < 1$  because  $\pi_S(\lambda, \tau)$  is continuous. ||

The proof of Proposition 3 only uses the fact that the participation constraint in the full awareness problem (2)–(4) binds. This is the case, e.g. if  $V(\underline{w}) = -\infty$  as in Grossman and Hart (1983), but holds more generally, as the example in the next section shows. If  $\tau = \bar{a}^A$ , unaware agents provide second best effort without being incentivized. In this special case, it is certainly optimal to never make unaware agents aware. The proposition shows that in all other cases, it is optimal for the principal to make everybody aware if the extent of unawareness in the population is sufficiently small (i.e. if  $\lambda$  is sufficiently close to 1).

### 5. SPECIAL CASE: TWO OUTCOMES

In this section, we specialize our general model to the case of two output realizations, which is particularly suited for the discussion of performance incentives, and illustrate some possible contracting outcomes numerically. The example highlights four qualitative features of the agency problem that hold more generally but can be described more clearly in the specific example: (1) The problem of finding the optimal contracts for the aware and the unaware agent can be split and the two subproblems be solved separately, (2) there is a trade-off between the rent left to the aware agent and inefficient risk-bearing by the unaware agent, (3) the optimal resolution of this trade-off depends monotonically on the population parameter  $\lambda$ , (4) even if there are only few aware agents in the population, unaware agents will bear some inefficient performance risk, but these incentives are low powered.

As in the general model, the agent’s effort is an arbitrary number, which we take to be non-negative,  $a \geq 0$ . The agent’s effort generates a stochastic output  $x \in \{0, X\}$ ,  $X > 0$ , with the probability measure  $P(x = 0|a) = 1/(1 + a)$  and  $P(x = X|a) = a/(1 + a)$ . We assume that  $X$  is sufficiently large for a solution to exist that is acceptable to the principal and for  $a = 0$  not to be optimal in the different contracting problems discussed below.

Clearly,  $E[x|a]$  is increasing and concave in  $a$ . Moreover,  $x$  satisfies the monotone likelihood ratio property (MLRP) and the convex distribution function condition (CDFC), which are sufficient conditions for the first-order approach in the pure moral hazard problem (Rogerson, 1985).<sup>17</sup>

17. (MLRP): For any  $a' < a$ ,  $P(x|a)/P(x|a')$  is increasing in  $x$ .

The principal remunerates the agent by a compensation  $w_0$  for  $x = 0$  and  $w_1$  for  $x = X$ . His objective function is

$$\frac{a}{1+a}(X - w_1) - \frac{1}{1+a}w_0.$$

The simple outcome structure allows us to interpret  $w_0$  as the agent's base salary and  $w_1 - w_0$  as her bonus. The agent's monetary utility function  $V$  is increasing, strictly concave, and defined over  $w \in (0, \infty)$ , with  $V(0) = \lim_{w \rightarrow 0} V(w)$  sufficiently small for the relevant participation constraints to bind. A standard example is CRRA utility

$$V(w) = \begin{cases} \frac{1}{\gamma}(w^\gamma - 1) & \gamma \neq 0, \gamma < 1 \\ \ln w & \gamma = 0 \end{cases} \quad (11)$$

with  $\gamma \leq 0$ , where utility is unbounded below. However, below we also document simulations for the case of  $\gamma > 0$  to show that our results also apply to cases of bounded utility (which is more complicated to characterize analytically because of the possibility of a trade-off between incentives and rents).

The agent's utility cost of effort is  $C(a) = ka^\beta$  with  $k > 0$  and  $\beta > 1$ . Hence, the agent's objective is

$$V(w) - ka^\beta.$$

Following Grossman and Hart (1983), it is convenient to work in utility rather than money terms and thus to set

$$v_i = V^{-1}(w_i)$$

for  $i = 0, 1$ .

The full-awareness problem (2)–(4) and the unaware problem (5)–(6) can be re-formulated easily and yield  $\bar{\pi}^A$  and  $\bar{\pi}^U(\tau)$  as discussed in Proposition 1 from which one calculates  $(\tau_{\min}, \tau_{\max})$ .

For the general screening problem, as discussed in Section 4.2, we assume  $\tau \in (\tau_{\min}, \tau_{\max})$ . In utility terms, the principal's problem is as follows:<sup>18</sup>

$$\begin{aligned} \max_{a \geq 0, v_0^A, v_1^A, v_0^U, v_1^U} & \lambda \left( \frac{a}{1+a}(X - V^{-1}(v_1^A)) - \frac{1}{1+a}V^{-1}(v_0^A) \right) \\ & + (1 - \lambda) \left( \frac{\tau}{1+\tau}(X - V^{-1}(v_1^U)) - \frac{1}{1+\tau}V^{-1}(v_0^U) \right) \end{aligned} \quad (\text{OBJ})$$

---

In our example, this reduces to the condition  $P(x = X|a)/P(x = X|a') > P(x = 0|a)/P(x = 0|a')$ , which holds for any  $a' < a$ .

(CDFC):  $F(x, a)$  is convex in  $a$  for all  $x$ .

In our example,

$$F(x, a) = \begin{cases} 0 & \text{if } x < 0 \\ 1/(1+a) & \text{if } 0 \leq x < X \\ 1 & \text{if } x \geq X \end{cases}$$

18. In what follows, we do not add the constraint that  $v_i \in \text{range}(V)$  explicitly for better readability. It is, of course, implicit in the formulation of the objective function.

$$\text{s.t. } a \in \arg \max_{a' \geq 0} \frac{a'}{1+a'} v_1^A + \frac{1}{1+a'} v_0^A - ka'^\beta \tag{ICA}$$

$$\frac{1}{1+a} v_0^A + \frac{a}{1+a} v_1^A - ka^\beta \geq \underline{u} \tag{PCA}$$

$$\frac{1}{1+\tau} v_0^U + \frac{\tau}{1+\tau} v_1^U - k\tau^\beta \geq \underline{u} \tag{PCU}$$

$$\frac{1}{1+a} v_0^A + \frac{a}{1+a} v_1^A - ka^\beta \geq \max_{b \geq 0} \left\{ \frac{1}{1+b} v_0^U + \frac{b}{1+b} v_1^U - kb^\beta \right\} \tag{ICA-U}$$

$$\frac{1}{1+\tau} v_0^U + \frac{\tau}{1+\tau} v_1^U \geq \frac{1}{1+\tau} v_0^A + \frac{\tau}{1+\tau} v_1^A. \tag{ICU-A}$$

It is straightforward to see that (PCA) is implied by (ICA-U) and (PCU). This implies by our assumption about  $V(0)$  that (PCU) and (ICA-U) must bind. A simple application of Jensen’s inequality shows that  $v_1^U \geq v_0^U$ . By concavity, one can then decompose (ICA-U) into (ICA1) and (ICA2) by introducing a choice variable  $b \geq 0$  to obtain the following reformulation of the incentive-compatibility constraint:

$$\frac{1}{1+a} v_0^A + \frac{a}{1+a} v_1^A - ka^\beta = \frac{1}{1+b} v_0^U + \frac{b}{1+b} v_1^U - kb^\beta \tag{ICA1}$$

$$v_1^U - v_0^U = k\beta b^{\beta-1} (1+b)^2. \tag{ICA2}$$

$b$  measures the efficiency loss due to risk-bearing by the unaware agent. If  $b = 0$ , the unaware agent is fully insured, which is efficient.

**Lemma 2.** *If there is an optimum with  $v_1^A < v_0^A$ , then there is one with  $v_1^A = v_0^A$ .*

*Proof.* If  $v_1^A \leq v_0^A$ ,  $a = 0$ . Hence,  $v_1^A$  only matters for (ICU-A), by relaxing this constraint. However, if  $a = 0$ , it is straightforward to show that (ICU-A) does not bind at the optimum.  $\square$

By Lemma 2, we can restrict attention to the case  $v_1^A \geq v_0^A$  and therefore write the incentive constraint (ICA) as a first-order constraint. The screening problem thus reduces to the following problem:

$$\max_{a, b \geq 0, v_0^A, v_1^A, v_0^U, v_1^U} \lambda \left( \frac{a}{1+a} (X - V^{-1}(v_1^A)) - \frac{1}{1+a} V^{-1}(v_0^A) \right) \tag{OBJ}$$

$$+ (1-\lambda) \left( \frac{\tau}{1+\tau} (X - V^{-1}(v_1^U)) - \frac{1}{1+\tau} V^{-1}(v_0^U) \right)$$

$$\text{s.t. } v_1^A - v_0^A = k\beta a^{\beta-1} (1+a)^2 \tag{ICA'}$$

$$v_0^U + \tau v_1^U = (1+\tau)(\underline{u} + k\tau^\beta) \tag{PCU'}$$

$$\frac{1}{1+a} v_0^A + \frac{a}{1+a} v_1^A - ka^\beta = \frac{1}{1+b} v_0^U + \frac{b}{1+b} v_1^U - kb^\beta \tag{ICA1}$$

$$v_1^U - v_0^U = k\beta b^{\beta-1} (1+b)^2 \tag{ICA2}$$

$$v_0^U + \tau v_1^U \geq v_0^A + \tau v_1^A. \tag{ICU'}$$

In order to analyze this problem, it is convenient to let, for  $(z, \tau) \in \mathbb{R}_0^+ \times (\tau_{\min}, \tau_{\max})$ ,

$$h(z, \tau) := \frac{\tau}{1+\tau} k\beta z^{\beta-1} (1+z)^2 - k\tau^\beta.$$

Straightforward computation shows that  $h$  has the following properties.

**Lemma 3.** For all  $(z, \tau) \in \mathbb{R}^+ \times (\tau_{\min}, \tau_{\max})$ ,

1.  $h_z(z, \tau) > 0$ ,
2.  $\frac{d}{dz}(h(z, z) - h(z, \tau)) < 0 \Leftrightarrow z < \tau$  and  $\frac{d}{dz}(h(z, z) - h(z, \tau)) > 0 \Leftrightarrow z > \tau$ .

By inserting (ICA') and (ICA2) into (ICA1) and (ICA2) into (PCU'), problem (OBJ)–(ICU') can be rewritten as follows:

$$\begin{aligned} \max_{a, b \geq 0, v_0^A, v_1^A, v_0^U, v_1^U} \lambda & \left( \frac{a}{1+a} (X - V^{-1}(v_1^A)) - \frac{1}{1+a} V^{-1}(v_0^A) \right) \\ & + (1-\lambda) \left( \frac{\tau}{1+\tau} (X - V^{-1}(v_1^U)) - \frac{1}{1+\tau} V^{-1}(v_0^U) \right) \\ \text{s.t. } & v_0^U + h(b, \tau) = \underline{u} \end{aligned} \quad (12)$$

$$v_0^A + h(a, a) = v_0^U + h(b, b) \quad (13)$$

$$v_0^A + h(a, \tau) \leq \underline{u} \quad (14)$$

$$v_1^A = v_0^A + k\beta a^{\beta-1} (1+a)^2 \quad (15)$$

$$v_1^U = \underline{u} + (1+\tau)k\tau^{\beta-1} + \frac{1}{\tau} h(b, \tau). \quad (16)$$

Here, constraint (12) combines (PCU') and (ICA2) and is essentially the unaware participation constraint, constraint (13) is the aware incentive constraint (ICA1), and constraint (14) combines (ICU'), (ICA'), and (PCU') and is essentially the unaware incentive constraint. Constraint (15), which is (ICA'), and equation (16), which combines (PCU') and (ICA2), define the residual variables  $v_1^A$  and  $v_1^U$ .

Due to the additive nature of the objective function and the particular structure of the constraints, this problem can be decomposed into two subproblems and thus solved in two stages. For any  $b \geq 0$ , define

$$r := h(b, b) - h(b, \tau). \quad (17)$$

The first stage then consists of finding the incentive contract for aware agents,  $a, v_0^A, v_1^A$ , to

$$\max \frac{a}{1+a} (X - V^{-1}(v_1^A)) - \frac{1}{1+a} V^{-1}(v_0^A) \quad (\text{P-A1})$$

$$\text{s.t. } v_0^A + h(a, a) = r + \underline{u} \quad (\text{P-A2})$$

$$(14), (15). \quad (\text{P-A3})$$

Note that this problem only depends on the variables of the unaware subproblem through  $r$ . By equation (12),  $r$  can be rewritten as  $r = v_0^U + h(b, b) - \underline{u}$ . Hence, constraint (P-A2) is the aware incentive constraint (ICA1), rewritten in terms of  $r$ . Since  $v_0^A + h(a, a)$  is the aware agent's equilibrium utility, (P-A2) shows that  $r$  measures the aware agent's information rent. Hence,  $r$  must be non-negative, and if  $r = 0$ , the aware agent's participation constraint (PCA) binds (which only happens if  $\lambda = 1$ ). Let  $H(r)$  denote the value of problem (P-A1)–(P-A3).



**Lemma 4.** *If  $r \leq k\tau^\beta$ , problem (P-A1)–(P-A3) has a solution, and  $H(r)$  is continuous and strictly decreasing in  $r$ .*

*Proof.* By substituting for  $v_0^A$  from (P-A2) and combining (P-A2) and equation (14), the first-stage problem is equivalent to the problem of finding  $a \geq 0$  to maximize

$$\frac{a}{1+a}(X - V^{-1}(v_1^A)) - \frac{1}{1+a}V^{-1}(v_0^A)$$

s.t.  $h(a, a) - h(a, \tau) \geq r$  (18)

where

$$v_0^A = \underline{u} + r - ka^\beta(\beta(1+a) - 1). \tag{19}$$

$$v_1^A = \underline{u} + r + ka^{\beta-1}(\beta(1+a) + a). \tag{20}$$

By Lemma 3.2, the left-hand side of equation (18) is strictly quasi-convex with a unique minimum at  $a = \tau$ . Therefore, the constraint sets defined by equation (18) are strictly shrinking for increasing  $r$ :

$$\{a \geq 0; h(a, a) - h(a, \tau) \geq r\} \subset \{a \geq 0; h(a, a) - h(a, \tau) \geq r'\}$$

for  $r' < r$ . Furthermore, they are non-empty if  $r \leq h(0, 0) - h(0, \tau) = k\tau^\beta$ . Since the objective function decreases strictly for all large enough  $a$ , we can restrict attention to  $a$  in a compact interval  $[0, A]$ . By continuity, a solution therefore exists if  $r \leq k\tau^\beta$ .  $H$  is continuous by the maximum theorem.

The derivative of the objective function (P-A1) with respect to  $r$  is

$$-\frac{a}{1+a}V^{-1'}(v_1^A) - \frac{1}{1+a}V^{-1'}(v_0^A) < 0.$$

Hence, decreasing  $r$  strictly increases the objective function and strictly increases the set of feasible maximizers. This proves the lemma.  $\parallel$

Note that standard envelope theorems, including the constrained envelope theorem of (Milgrom and Segal, 2002, p. 597), do not apply to the above problem because of lack of convexity. Our argument relies on the monotonicity of the constraint sets (in the sense of set inclusion) with respect to the parameter.

The second-stage problem is to find the compensation contract for unaware agents,  $b, v_0^U, v_1^U$ , which determines the information rent  $r$ , to

$$\max(1 - \lambda) \left( \frac{\tau}{1 + \tau} (X - V^{-1}(v_1^U)) - \frac{1}{1 + \tau} V^{-1}(v_0^U) \right) + \lambda H(r) \tag{P-U1}$$

s.t. (12), (16), and (17). (P-U2)

A fundamental insight in classical screening models is the trade-off between allocational efficiency and informational rents (see, e.g. Bolton and Dewatripont, 2005 and Laffont and Martimort, 2002). Interestingly, the same trade-off can be established in our theory of heterogeneous awareness. Furthermore, the optimal resolution of the trade-off depends monotonically on  $\lambda$ .

**Proposition 4.** *At the optimum, the aware agents' rent  $r$  decreases and the unaware agents' risk exposure  $b$  increases strictly with the aware population share  $\lambda$ .<sup>19</sup>*

*Proof.* Substituting for  $v_0^U$ ,  $v_1^U$ , and  $r$ , the second-stage problem is equivalent to the problem of finding  $b \geq 0$  to maximize

$$M(b; \lambda) := (1 - \lambda)G(b) + \lambda H(r), \quad (21)$$

where

$$G(b) := \frac{\tau}{1 + \tau} \left( X - V^{-1} \left( v_1^U \right) \right) - \frac{1}{1 + \tau} V^{-1} \left( v_0^U \right)$$

$$v_1^U = \underline{u} + k\tau^{\beta-1}(1 + \tau) + \frac{1}{\tau}h(b, \tau)$$

$$v_0^U = \underline{u} - h(b, \tau)$$

$$r = h(b, b) - h(b, \tau).$$

$G$  is the profit the principal makes on the unaware agent.  $G$  is differentiable, and we have

$$G'(b) = -\frac{1}{1 + \tau} h_z(b, \tau) (V^{-1'}(v_1^U) - V^{-1'}(v_0^U)).$$

Since  $h_z(b, \tau) > 0$  for  $b > 0$  by Lemma 3.1, since  $V^{-1'}$  is strictly increasing, and  $v_0^U < v_1^U$  for all  $b > 0$ , we have  $G'(b) < 0$  for all  $b > 0$ .

If  $b > \tau$ , then  $\frac{d}{db}(h(b, b) - h(b, \tau)) > 0$  by Lemma 3.2. Lemma 4 therefore implies that  $H(r(b))$  is strictly decreasing in  $b$  for all  $b \geq \tau$ . Hence,  $M$  is strictly decreasing in  $b$  for all  $b \geq \tau$ . Therefore, if there is an optimum,  $b < \tau$ . Conversely, when restricting attention to  $b \in [0, \tau]$ , then  $r \leq k\tau^\beta$  by Lemma 3.2,  $H$  is continuous by the maximum theorem, and thus a solution exists.

Next, let  $\lambda' < \lambda$  and  $b' < b \leq \tau$ . Then

$$M(b', \lambda) - M(b', \lambda') < M(b, \lambda) - M(b, \lambda') \quad (22)$$

since  $H(r(b))$  is strictly increasing in  $b$  for  $b < \tau$  by equation (17), Lemma 3.2, and Lemma 4, and since  $G'(b) < 0$  for all  $b > 0$ . Equation (22) means that  $M$  has strictly increasing differences (see, e.g. Vives, 1999, Chapter 2). Standard monotone comparative statics results (e.g. Vives, 1999, Theorem 2.3) therefore imply that  $b$  is strictly increasing in  $\lambda$ .

Because  $b < \tau$ , this together with equation (17) and Lemma 3.2 implies that  $r$  is strictly decreasing in  $\lambda$ .  $\parallel$

Proposition 4 implies that whenever there are some aware agents in the population ( $\lambda > 0$ ), unaware agents are inefficiently incentivized ( $b > 0$ ). But if there are many unaware agents ( $\lambda$  small), this distortion is small and incentives are low powered ( $b$  is small).<sup>20</sup> Conversely, when there are hardly any unaware agents in the population ( $\lambda$  close to 1), the principal wants to extract the maximum rent from the aware agents ( $r$  close to 0), which means by equation (17) that unaware agents bear a maximum of risk ( $b$  close to  $\tau$ ), so that aware agents can be

19. Since the maximum may not be unique, this statement must be understood in terms of correspondences: if solutions to the unaware problem are denoted by  $b^*(\lambda)$ , the correspondence  $\lambda \mapsto b^*(\lambda)$  is strictly increasing, which means that all selections of  $b^*$  are strictly increasing. Analogously for  $r$ .

20. Remember that  $\pi_S(0, \tau) = \bar{\pi}^U$  and  $b = 0$  in the fully unaware problem.

optimally incentivized. However, as we know from Propositions 2 and 3, this is so expensive that the principal prefers to make the remaining unaware agents aware and induce the full-awareness effort  $\bar{a}^A$ . This is optimal, although the principal concedes practically no rent to the aware agent in the screening solution.

Interestingly, despite the unaware agent’s bounded rationality, much of the preceding analysis is as in standard screening models. For example, the “bad” type (the aware agent) gets a positive rent, while the “good” type is kept to her reservation utility. Also, here as in standard models, the incentive constraint of the “bad” type binds. Furthermore, the aware agents’ rents are decreasing in their population share  $\lambda$ . On the other hand, it can be shown that in the present model, the single-crossing property fails to hold, efficiency losses can arise for both types, and the incentive-compatibility constraints of both types can bind in a separating solution.

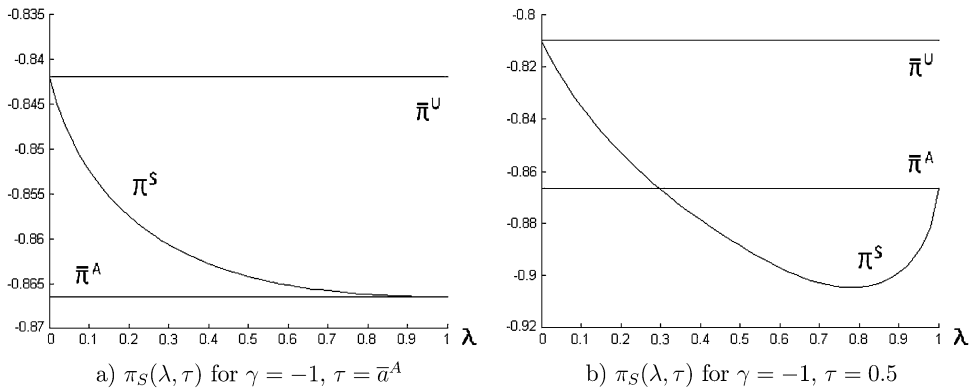


FIGURE 1  
 (a)  $\pi_S(\lambda, \tau)$  for  $\gamma = -1, \tau = \bar{a}^A$  and (b)  $\pi_S(\lambda, \tau)$  for  $\gamma = -1, \tau = 0.5$

We conclude this section by illustrating Proposition 2 and 3 numerically in our example.

For the production and cost parameters, we choose  $\beta = 2, k = 0.5$ , and  $X = 1$ , the agent’s reservation utility is  $\underline{u} = 0$ . Utility is CRRA, given by equation (11).

We first consider the case of utility that is unbounded below and let  $\gamma = -1$ . In this case,  $(\tau_{\min}, \tau_{\max}) = (0.18, 0.64)$ , and the second-best effort level in the incentive problem under awareness, equations (2)–(4) is  $\bar{a}^A = 0.23$ .<sup>21</sup>

According to Proposition 3, there are two cases of interest.

- (1)  $\tau = \bar{a}^A$ : Figure 1(a) shows indeed a corner  $\lambda^*(\tau) = 1$ .
- (2)  $\tau \neq \bar{a}^A$ : Figure 1(b) uses  $\tau = 0.5$  and shows indeed an interior  $\lambda^*(\tau) < 1$ .

We now consider the case of utility that is bounded below and let  $\gamma = 1/2$ . To simplify the numbers, we also let  $X = 10$ . One can check that the participation constraint in the aware problem binds, and we find  $\bar{a}^A = 0.81$ . Furthermore, because of the new feasibility constraint  $v_i \geq -2$ ,<sup>22</sup> if  $\lambda = 1$ , we cannot necessarily lower the unaware agent’s utility sufficiently to find  $v_0^U, v_1^U$  that make  $(v_0^U, v_1^U, \bar{v}_0^A, \bar{v}_1^A)$  incentive-compatible in the full screening problem. The simulation shows that this is possible if and only if  $\tau \leq \bar{\tau} = 1.11$ . Hence, if  $\tau > 1.11$ , we have  $\pi_S(1, \tau) < \bar{\pi}^A$ .

21. All numbers are rounded.  
 22. Equation (11) and  $w \geq 0$  imply that  $v \geq -1/\gamma$ .

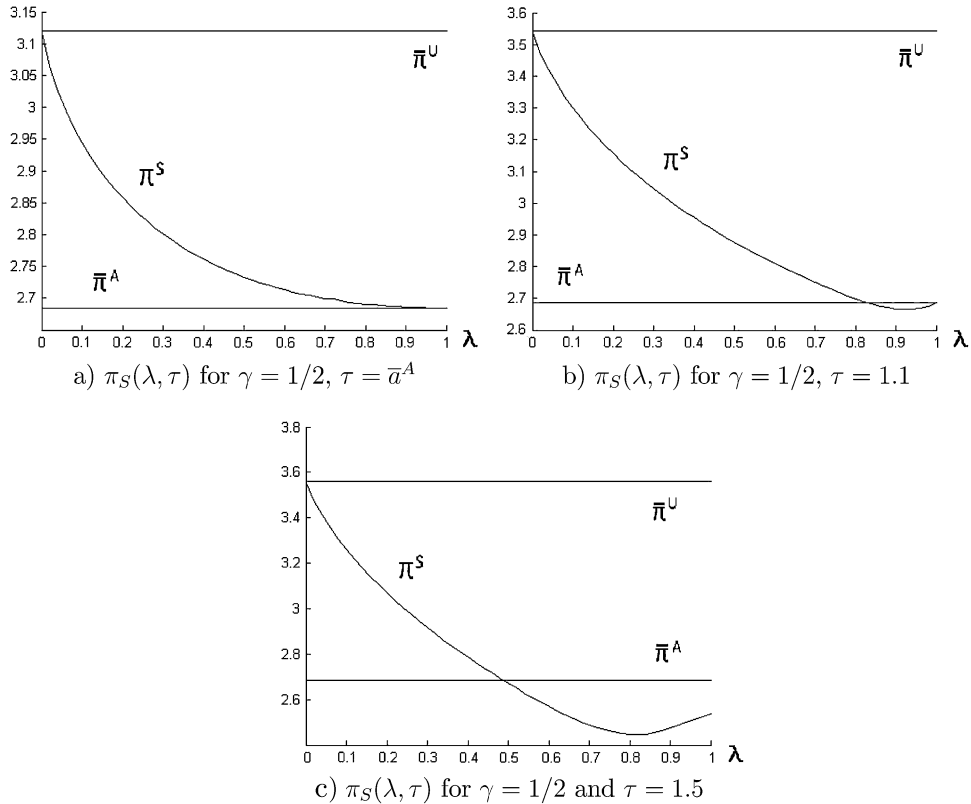


FIGURE 2

(a)  $\pi_S(\lambda, \tau)$  for  $\gamma = 1/2, \tau = \bar{a}^A$ , (b)  $\pi_S(\lambda, \tau)$  for  $\gamma = 1/2, \tau = 1.1$ , and (c)  $\pi_S(\lambda, \tau)$  for  $\gamma = 1/2, \tau = 1.5$

We therefore have several qualitatively different results depending on the value of  $\tau$ .

- (1)  $\tau = \bar{a}^A (\leq \bar{\tau})$ : In this case, the condition of Proposition 3 is violated. Indeed, Figure 2(a) shows a corner cutoff value  $\lambda^*(\tau) = 1$ . Since  $\tau < \bar{\tau}$ , we can achieve  $\pi_S(1, \tau) = \bar{\pi}^A$  in the solution to the screening problem.
- (2)  $\tau \neq \bar{a}^A$  and  $\tau < \bar{\tau}$ : Figure 2(b) uses  $\tau = 1.1$  and yields an interior  $\lambda^*(\tau) < 1$ . Since  $\tau < \bar{\tau}$ , we still have  $\pi_S(1, \tau) = \bar{\pi}^A$ .
- (3)  $\tau \neq \bar{a}^A$  and  $\tau > \bar{\tau}$ : Let  $\tau = 1.5$ . Again, we have an interior  $\lambda^*(\tau) < 1$  by Proposition 3. However, as shown in Figure 2(c),  $\pi_S(1, \tau) < \bar{\pi}^A$ .

## 6. CONCLUDING DISCUSSION

We have proposed a theory of incentive provision and communication in a principal–agent relationship if agents may not understand the full effects of their actions. While the model has several special features, we will now argue that it is robust to a number of possible extensions.

**Justifiability of Contracts.** Filiz-Ozbay (2008), Ozbay (2008), and Heifetz, Meier and Schipper (2010) argue that in the context of unawareness, a reasonable solution concept should include the requirement that the agent finds the contract justifiable, in the sense that the contract is optimal for the principal also from the agent’s point of view. Now we explore the ramifications of this added requirement for our analysis.

First, it is simple to see that the solution of the basic contracting problem in Section 4.1 is justifiable in this sense. If the optimal contract is complete the agent is aware, and the problem reduces to the standard principal–agent problem, the solution of which is even robust to common knowledge of rationality and of the contractual setting. If the optimal contract is incomplete, the agent remains unaware and unconsciously chooses  $\tau$ . Then, the agent’s objective function (6) includes a fixed-cost element, and again the contract is optimal in the agent’s mind.

However, the solution of the heterogeneous awareness problem in Section 4.2 is not necessarily justifiable. When the principal prefers incomplete contracts, the unaware agent does not understand why there are two different contracts or a single contract designed the way it is (in case of pooling). If we require the contract to be justifiable, the solution is either the complete contract outcome (full awareness outcome) or the incomplete pooling contract  $\bar{w}^U$  that makes sense for the unaware agent.<sup>23</sup> The additional justifiability constraint therefore reduces the principal’s profit from proposing incomplete contracts and makes him more likely to make all agents aware. This is because we add an additional justifiability constraint to the principal’s problem. Put differently, skepticism of unaware agents can promote their awareness.

*Precontractual cognition:* Yet, it can be argued that justifiability is too strong a requirement. If one acknowledges that the model necessarily only describes a simplified snapshot of a full (highly complex) contracting problem, it may well be reasonable to assume that the agent does not want or need to understand the reason for what she sees, as long as what she chooses is optimal for her.

In what follows, we therefore propose a weaker justifiability restriction than that of Filiz-Ozbay (2008) and Ozbay (2008). In the spirit of Bolton and Faure-Grimaud (2009) and Tirole (2009), we assume that if the observed contracts are not justifiable, the unaware agent may become aware that something is wrong with her view and starts thinking about it. This cognitive effort leads her to full awareness with probability  $\delta$ . With probability  $1 - \delta$ , the agent remains unaware and chooses (one of) the proposed contracts without further ado. In fact, after seeing a non-justifiable contract, *i.e.* a menu of incomplete contracts that is different from the single contract  $\bar{w}^U$ , the fraction of aware agents increases to  $\lambda' = \lambda + (1 - \lambda)\delta$ . Hence, non-justifiable contracts promote awareness. Thus, making all agents aware is more likely to be better than non-justifiable contracts according to Proposition 2.

In this extended model, there are three alternatives for the principal: (i) making all agents aware, (ii) proposing  $\bar{w}^U$  alone, (iii) proposing the solution of the screening problem (OBJ)–(ICU-A), with the fraction of aware agents replaced by  $\lambda' (> \lambda)$ .

From Proposition 2, we conclude that now the incomplete contracting solution (iii) becomes less attractive than the complete contracting solution (i), as  $\lambda' > \lambda$ , *i.e.* the fraction of aware agents becomes larger. Yet, alternatives (ii) and (iii) may be optimal in some circumstances. For example, when  $\lambda$  is small,  $\tau \neq \bar{a}^A$  and  $\delta$  is large, proposing  $\bar{w}^U$  alone (alternative (ii)) is optimal. In this case, non-justifiable contracts (alternative (iii)) tend to be worse than the complete contracts (i) since  $\lambda'$  is large. Furthermore, proposing  $\bar{w}^U$  alone is better than alternative (i): since  $\lambda$  is small, the loss from the rent of the aware agent is small for the principal. On the other hand, when  $\delta$  is small enough, alternative (iii) can be optimal, as  $\lambda'$  is not significantly greater than  $\lambda$ .

*Communication profiness:* We have shown that under certain conditions, the principal prefers to leave the agent unaware of the full contracting problem. Interestingly, even the aware agent has no strict incentive to make her unaware colleague aware through communication. First, the aware agent will not do so before contracting because under the optimal contract, the unaware

23. Here, we assume that the aware agent understands the full contracting problem as well as the principal. If the aware agent does not know that other agents may be unaware, this contract is not justifiable either.

agent exerts a positive externality on the aware agent by conferring a positive rent on her (as our discussion in Section 5 shows, (PCA) is usually redundant). Second, after contracting, the aware agent has no incentive to do so either because making the unaware agent aware cannot create any extra rent for herself but only hurts the principal, as the principal always earns more from the unaware agents. Hence, our result is robust to the possibility of internal communication among agents.

*Dynamic stability:* Our analysis has been static, but it lends itself to an interesting dynamic interpretation. In Proposition 2, we have shown that the optimal contracts are more likely to leave unaware agents unaware for smaller  $\lambda$  (where the measure of agents is taken with respect to  $\tau$ ). Hence, the more agents are likely to be unaware, the more will remain unaware after contracting and vice versa. This suggests a certain stability of unawareness. This observation may be particularly important as it suggests that deviations from full rationality are not necessarily doomed to die out in the long run.

*Welfare implication of public announcements:* Is there room for a benevolent policy maker to improve the outcome through promoting the agent's awareness?<sup>24</sup> We have argued that when there are only unaware agents in the population, a public policy of making agents aware is not welfare enhancing. The reason is that such an announcement would force the principal to provide explicit incentives to the agent, which is costly. Since the principal maximizes total surplus, the fact that he does not choose to make the agent aware shows that the costs outweigh the benefits. In Section 4.2, when there are heterogeneous agents, this conclusion still holds. If the principal prefers to leave the agent unaware, the outcome Pareto-dominates the outcome of making all agents aware. To see this point, note that when the principal prefers to leave the agent unaware, he must earn a higher expected profit than making her aware. Furthermore, the aware agent earns a positive rent while she would earn zero rents if all agents were made aware. Finally, the unaware agent earns a zero rent in any case. Thus, there is no need for the policy maker to intervene in the heterogeneous environment as well. Of course, this no-intervention recommendation crucially depends on the assumptions that the principal knows all the choice possibilities of the agent.

*Slip-of-mind or clueless unawareness:* Board and Chung (2009) distinguish "slip-of-mind unawareness" and "clueless unawareness". Under the former, the agent becomes aware of the full problem as soon as she sees its description in the contract under the latter she is "hard wired" in her choice of the default action. Our model can encompass both forms of unawareness if we introduce costs of communication through a complete contract ("training"). Then, the two cases of Board and Chung (2009) represent two extremes, one with zero costs, and the other with infinitely high cost. Compared to our model, the existence of the cost only makes the alternative of making all agents aware less attractive for the principal.

*Acknowledgment.* We would like to thank Bruno Biais and four anonymous referees for very helpful comments and suggestions, as well as Suddipto Bhattacharya, Patrick Bolton, Jacques Crémer, Roberta Dessí, Mathias Dewatripont, Kfir Eliaz, Antoine Faure-Grimaud, Leonardo Felli, Emel Filiz-Ozbay, Spyros Galanis, Aviad Heifetz, Philippe Jehiel, Daniel Kräbmer, Pei Kuang, Jing Li, Patrick Rey, Klaus Schmidt, Ilya Segal, Dagmar Stahlberg, Jidong Zhou, and participants in several conferences and seminars for useful discussions and comments. We also thank the German Science Foundation (DFG) for financial support.

## REFERENCES

- AGHION, P. and BOLTON, P. (1987), "Contracts as a Barrier to Entry", *American Economic Review*, **77**, 388–401.  
 ANDERLINI, L. and FELLI, L. (1999), "Incomplete Contracts and Complexity Costs", *Theory and Decision*, **46**, 23–50.  
 BARON, J. and KREPS, D. (1999), *Strategic Human Resources: Frameworks for General Managers* (New York: Wiley).

24. This has been suggested, e.g. by Korobkin (2003) in the context of unawareness about other agents' actions.

- BÉNABOU, R. and TIROLE, J. (2003), "Intrinsic and Extrinsic Motivation", *Review of Economic Studies*, **70**, 489–520.
- BÉNABOU, R. and TIROLE, J. (2006), "Incentives and Prosocial Behavior", *American Economic Review*, **96**, 1652–1678.
- BOARD, O. and CHUNG, K. S. (2009), "Object-Based Unawareness: Theory and Applications", mimeo, University of Pittsburgh and University of Minnesota.
- BOLTON, P. and DEWATRIPONT, M. (2005), *Contract Theory* (Cambridge: MIT Press).
- BOLTON, P. and FAURE-GRIMAUD, A. (2009), "Thinking Ahead: The Decision Problem", *Review of Economic Studies*, **76**, 1205–1238.
- BOLTON, P. and FAURE-GRIMAUD, A. (2010), "Satisficing Contracts", *Review of Economic Studies*, **77**, 937–971.
- DUNNETTE, M. and HOUGH, L. (1990), *Handbook of Industrial and Organizational Psychology* (Palo Alto: Consulting Psychologists Press).
- DYE, R. (1985), "Costly Contract Contingencies", *International Economic Review*, **26**, 233–250.
- ELIAZ, K. and SPIEGLER, R. (2006), "Contracting with Diversely Naive agents", *Review of Economic Studies*, **73**, 689–714.
- ELIAZ, K. and SPIEGLER, R. (2011), "Consideration Sets and Competitive Marketing", *Review of Economic Studies*, **78**, 235–262.
- FALK, A. and FEHR, E. (1999), "Wage Rigidity in a Competitive Incomplete Contract Market", *Journal of Political Economy*, **107**, 106–134.
- FALK, A. and KOSFELD, M. (2006), "The Hidden Costs of Control", *American Economic Review*, **96**, 1611–1630.
- FILIZ-OZBAY, E. (2008), "Incorporating Unawareness into Contract Theory", mimeo, University of Maryland.
- FISCHHOFF, B., SLOVIC, P. and LICHTENSTEIN, S. (1977), "Knowing with Certainty: The Appropriateness of Extreme Confidence", *Journal of Experimental Psychology: Human Perception and Performance*, **3**, 552–564.
- FISKE, S. and TAYLOR, S. (2007), *Social Cognition—from Brain to Culture* (Maidenhead: McGraw-Hill).
- GABAIX, X. and LAIBSON, D. (2006), "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets", *Quarterly Journal of Economics*, **121**, 505–540.
- GALANIS, S. (2009), "Unawareness of Theorems", mimeo, University of Southampton.
- GROSSMAN, S. and HART, O. (1983), "An Analysis of the principal-agent Problem", *Econometrica*, **51**, 7–45.
- GROSSMAN, S. and HART, O. (1986), "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration", *Journal of Political Economy*, **94**, 691–719.
- HALPERN, J. and REGO, L. C. (2006), "Extensive Games with Possibly Unaware Players", in *Proceedings of Fifth International Joint Conference on Autonomous agents and Multiagent Systems*, 744–751.
- HAYEK, F. A. (1967), "Rules, Perception and Intelligibility", in Hayek, F. A. (ed.) *Studies in Philosophy, Politics and Economics* (Chicago: University of Chicago Press) 43–65.
- HEIFETZ, A., MEIER, M. and SCHIPPER, B. C. (2006), "Interactive Unawareness", *Journal of Economic Theory*, **130**, 78–94.
- HEIFETZ, A., MEIER, M. and SCHIPPER, B. C. (2010), "Dynamic Unawareness and Rationalizable Behavior", mimeo, University of California, Davis.
- HENDRIKSE, G. (2003), *Economics and Management of Organizations: Coordination, Motivation and Strategy* (Maidenhead: McGraw-Hill).
- HOLMSTRÖM, B. (1979), "Moral Hazard and Observability", *Bell Journal of Economics*, **10**, 74–91.
- JEWITT, I. (1988), "Justifying the First-Order Approach to principal-agent Problems", *Econometrica*, **56**, 1177–1190.
- KOROBKIN, R. (2003), "Bounded Rationality, Standard Form Contracts, and Unconscionability", *University of Chicago Law Review*, **70**, 1203–1295.
- LAFFONT, J. and MARTIMORT, D. (2002), *The Theory of Incentives: The principal-agent Model* (Princeton University Press).
- LAWLER, E. E. (1990), *Strategic Pay: Aligning Organizational Strategies and Pay Systems* (San Francisco: Jossey-Bass Publishers).
- LI, J. (2009), "Information Structures with Unawareness", *Journal of Economic Theory*, **144**, 977–993.
- MAYR, E. (1992), "The Idea of Teleology", *Journal of the History of Ideas*, **53**, 117–35.
- MEYER, E., ASHLEIGH, M., GEORGE, J. and JONES, G. (2007), *Contemporary Management*, European Edition (Maidenhead: McGraw-Hill).
- MILGROM, P. and ROBERTS, J. (1992), *Economics, Organization, and Management* (Englewood Cliffs: Prentice Hall).
- MILGROM, P. and SEGAL, I. (2002), "Envelope Theorems for Arbitrary Choice Sets", *Econometrica*, **70**, 583–601.
- MIRRLEES, J. (1975), "The Theory of Moral Hazard and Unobservable Behaviour: Part 1", *Review of Economic Studies*, **66**, 3–21.
- MODICA, S. and RUSTICHINI, A. (1994), "Awareness and Partitional Information Structure", *Theory and Decision*, **37**, 107–124.
- OZBAY, E. (2008), "Unawareness and Strategic Announcements in Games with Uncertainty", mimeo, University of Maryland.
- REGO, L. C. and HALPERN, J. (2007), "Generalized Solution Concepts in Games with Possibly Unaware Players", in Dov Samet (ed) *Theoretical Aspects of Rationality and Knowledge*, Proceedings of the 11th Conference (TARK 2007) (Louvain: Presses Universitaires De Louvain) 253–262.
- ROGERSON, W. (1985), "The First-Order Approach to principal-agent Problems", *Econometrica*, **53**, 1357–1368.
- RUSSO, J. E. and SCHOEMAKER, P. J. H. (1989), *Decision Traps* (New York: Doubleday).

- SCOTT, R. and TRIANTIS, G. (2006), "Anticipating Litigation in Contract Design", *Yale Law Journal*, **115**, 814–879.
- SQUINTANI, F. (2003), "Moral Hazard, Renegotiation, and Forgetfulness", *Games and Economic Behavior*, **44**, 98–113.
- TIROLE, J. (2009), "Cognition and Incomplete Contracts", *American Economic Review*, **99**, 265–294.
- TWAIN, M. (1876), *The Adventures of Tom Sawyer* (New York: Harper and Brothers).
- VANBERG, V. (2002), "Rational Choice vs. Program-Based Behavior: Alternative Theoretical Approaches and their Relevance for the Study of Institution", *Rationality and Society*, **14**, 7–53.
- VIVES, X. (1999), *Oligopoly Pricing* (Cambridge: MIT Press).