Optimal Procurement Contracts with Pre–Project Planning

Daniel Krähmer*and Roland Strausz†

September 16, 2010

Abstract

The paper studies procurement contracts with pre–project investigations in the presence of adverse selection and moral hazard. To model the procurer's problem, we extend a standard sequential screening model to endogenous information acquisition with moral hazard. The optimal contract displays systematic distortions in information acquisition. Due to a rent effect, adverse selection induces too much information acquisition to prevent cost overruns and too little information acquisition to prevent false project cancelations. Moral hazard mitigates the distortions related to cost overruns yet exacerbates those related to false negatives. The optimal mechanism is a menu of option contracts that achieves the dual goal of providing incentives for information acquisition and truthful information revelation.

Keywords: Information acquisition, procurement, dynamic mechanism design

JEL codes: D82, H57

---

*University Bonn, Department of Economics and Hausdorff-Center for Mathematics, Adenauer Allee 24-42, D-53113 Bonn (Germany), kraehmer@hcm.uni-bonn.de.

†Humboldt-Universität zu Berlin, Institute for Microeconomic Theory, Spandauer Str. 1, D-10178 Berlin (Germany), strauszr@wiwi.hu-berlin.de. We thank the editor, two anonymous referees, Pascal Courty, Peter Esö, Regis Renault and seminar participants in Barcelona, Berlin, Bonn, Cambridge, Heidelberg, Mannheim, Michigan, Munich, Rotterdam. Both authors gratefully acknowledge financial support by the DFG (German Science Foundation) under SFB/TR-15. Roland Strausz further acknowledges support under DFG–grant STR991/2-2.

1 Introduction

Practitioners in managing procurement projects stress the importance of pre–project planning. Based on a number of case studies, Gibson and Hamilton (1994) conclude that "there does exist a positive, quantifiable relationship between effort expended during the pre-project planning phase and the ultimate success of a project."[1] The goal of pre–project investigations is to obtain more accurate cost estimates, which allow the procurer to decide more carefully about whether to implement the project. In other words, pre–project planning is a process of information acquisition before the final implementation decision is taken.[2] The project management literature not only stresses the importance of effective pre–project planning but also warns procurers to keep as much control over this process as possible. For instance, Gibson et al. (2006, p.41) write in their empirical appraisal that procurers frequently "decide to delegate the pre-project planning process entirely to contractors, often with disastrous results". Given the importance of pre–project planning and these observed "disastrous results" from delegation, we develop and analyze an economic model of pre–project planning to enhance our understanding of this process.

Our starting point is the observation that the reported "disastrous results" from delegation point to conflicting interests and incentive problems. Due to the contractor's superior expertise, we can identify three sources of incentive problems in pre–project investigations. First, the contractor is already in a better position to estimate the project's cost from the very outset. Hence, the procurer–contractor relationship typically exhibits ex ante adverse selection. Second, the contractor, as the expert, is often in the better position to evaluate the additional information which the pre–project investigation reveals. Hence, pre–project investigations lead to additional adverse selection at an interim stage. Third, the amount of information that results from pre–investigations will largely depend on the contractor's own actions such as his advice and expertise on how to perform the investigations. Hence, pre–project investigations also involve a moral hazard problem.

[1] See also, e.g., Turner (1993), Gibson et al. (2006), Kähkönen (1999) and Dahlin, Bjelm, and Svensson (1999).

[2] Gibson et. al. (1995) define pre–project planning "as the process of developing sufficient strategic information for owners to address risk and decide whether to commit resources to maximize the change for a successful capital facility project".

The presence of these three different types of incentive problems leads us to view the procurement problem as a dynamic mechanism design problem. To capture adverse selection, we adopt a sequential screening approach as in Courty and Li (2000) and Esö and Szentes (2007b). To capture moral hazard, we allow for unobservable interim information acquisition by the agent. Our results are as follows. First, we derive as a benchmark the first–best solution without any incentive problems. In this case, the principal uses pre–investigations to mitigate one of two implementation errors: If the initial information about the project is favorable and indicates a positive social value, then the principal uses pre–investigation to prevent cost overruns. For this case, we say that information acquisition prevents false positives (or type I errors). In contrast, if the initial cost information is unfavorable, then she uses pre–investigations to correct for a possibly false cancelation of the project. Information acquisition then prevents false negatives (or type II errors).

Second, we characterize the optimal contract when information acquisition is still observable but there are adverse selection problems. In this case, the principal implements the optimal contract by a menu of contracts that offers a fixed price contract and a range of option contracts. The fixed price contract obliges the contractor to carry out the project in all cost circumstances for a fixed price. The option contract gives the contractor the right to first conduct a pre–investigation and then decide whether to complete the project for an exercise price or, alternatively, quit the project. Consistent with our analysis, option contracts feature prominently in a range of real world contractual relations.

Third, we treat the case with unobservable information acquisition and characterize the conditions under which the implied moral hazard problem causes additional agency costs. In particular, we find that additional agency costs do not arise when the contractor's additional information is uncorrelated with his initial information. In general, however, moral hazard does cause additional agency costs. We demonstrate that, in this case, the principal has to adapt the contract in ways that leads to bunching and less information acquisition.

Finally, the comparison of the optimal contract under the three different contracting environments allows us to characterize the systematic distortions in information acquisition and attribute them to the different incentive problems. We first demonstrate that the adverse selection problem leads to too much information acquisition to prevent false positives (type I errors) and too little information acquisition to prevent false negatives (type II errors). Moreover, we show that the moral hazard problem unambiguously exacerbates distortions with respect to type II errors and mitigates distortions with respect to type I errors. As a result, the welfare effects of the additional moral hazard problem are ambiguous. While the principal always loses from additional agency costs, the agent's utility and, in fact, total welfare may go up in the presence of a moral hazard problem. In this case, the moral hazard problem mitigates some of the inefficiencies from adverse selection.

Our paper contributes to the growing literature on optimal dynamic mechanism design. This literature studies optimal contract design in environments in which information is privately revealed to the agents over time.[3] A recent paper by Pavan et al. (2008) provides a general framework that encompasses earlier contributions on dynamic price discrimination (e.g., Baron and Besanko (1984), Laffont and Tirole (1990, 1996), Battaglini (2005)), or on sequential screening (e.g. Courty and Li (2000), Dai et al. (2006), Esö and Szentes (2007a, b)). Our contribution to this literature is to extend the analysis to endogenous information acquisition with moral hazard.

Our basic setup is most similar to the sequential screening model by Courty and Li (2000) who study monopolistic price discrimination when consumers are uncertain about their subsequent demand but obtain additional information before actual consumption.[4] Similar to Courty and Li (2000), we impose a first order stochastic dominance ranking condition on the family of the conditional cost distributions conditional on the initial signal.[5] A well–known, fundamental problem is that, unlike in static problems, incentive compatibility in sequential screening problem cannot be characterized in terms of monotonicity of the allocation rule. Courty and Li (2000) therefore first uncover necessary conditions for incentive compatibility from "local" constraints and then use the stochastic dominance ranking to verify implementability. We give a new proof for the validity of this approach by showing that for deterministic contracts, incentive compatibility can actually be characterized in terms of monotonicity. This property then allows us to treat the case when information acquisition is endogenous. The optimal contract

[3]A number of papers deals with the design of efficient mechanisms in dynamic environments. See, e.g., Athey and Segal (2007a, 2007b), Bergemann and Välimäki (2008).

[4]Esö and Szentes (2007a) extends Courty and Li (2000) to a multi–agent setting.

[5]In contrast, Dai et al. (2006) study an environment with second order stochastic dominance ranking.

derived by Courty and Li (2000) also displays a menu of option contracts. Thus, our analysis makes clear that option contracts are a robust feature of optimal sequential screening contracts also if information acquisition is endogenous and, possibly, involves moral hazard problems. In fact, such options not only work to screen consumers, but serve the additional purpose of inducing information acquisition.

Closely related is also the model by Esö and Szentes (2007b), where the principal can acquire costly, additional information which only the agent can interpret. We differ from Esö and Szentes (2007b) in at least three respects: first, in our model it is the agent who acquires additional information. More importantly, while Esö and Szentes (2007b) restrict attention to observable information acquisition, we also analyze the natural case when it is unobservable. This leads to a moral hazard problem that changes the analysis fundamentally. A final difference is that we characterize and give an intuitive explanation for the distortions in information acquisition that are implied by adverse selection and moral hazard.[6]

Our work is also related to the literature on information acquisition in principal–agent problems (e.g., Lewis and Sappington 1997, Cremer et al. 1998, Kessler 1998, Szalay 2009). The key difference is that this literature studies the incentives of an agent to acquire private information *before* he accepts or rejects the contract. Therefore, there is only a single screening stage.

The rest of the paper is organized as follows. The next section introduces the model. Section 3 discusses the first–best benchmark. In Section 4, we present the principal's problem. In Section 5, we solve the principal's problem when information acquisition is publicly observable and derive the distortions implied by optimal contracting. In Section 6, we treat the case with unobservable information acquisition. Section 7 concludes.

2 Model

A principal seeks to procure one unit of a good from an agent. The principal's valuation for the good is $v > 0$. There are two periods. In period 1, contracting takes place, and in period 2, the agent produces the good. Before contracting, the agent privately observes a noisy signal

[6] See also Hoffmann and Inderst (2009), who study the distortions in information acquisition due to adverse selection only.

about true costs. Let the noisy signal be given by the random variable $\tilde{\gamma}$ and let true costs be given by the random variable $\tilde{c}$. The agent privately observes the realization γ of $\tilde{\gamma}$, and it is common knowledge that $\tilde{\gamma}$ is distributed with distribution function F and density $f = F'$ on the support $\Gamma = [0, \bar{\gamma}]$. We assume that f is strictly positive and differentiable on Γ and that the hazard rate $h(\gamma) \equiv F(\gamma)/f(\gamma)$ is non–decreasing in γ.

True costs are equal to the signal plus a random shock $\tilde{s}$: $\tilde{c} = \tilde{\gamma} + \tilde{s}$. The cost shock has support $\mathbb{R}$ and has zero conditional mean so that $\tilde{\gamma}$ is an unbiased estimate of costs: $E[\tilde{c}|\gamma] = \gamma + E[\tilde{s}|\gamma] = \gamma$.[7] In general, the distribution of $\tilde{c}$ conditional on γ is given by $G(\cdot|\gamma)$ with the conditional density $g(\cdot|\gamma)$.[8] We assume that the family $\{G(\cdot|\gamma)\}_\gamma$ is ranked in terms of first order stochastic dominance. That is, $G(\cdot|\gamma)$ first order stochastically dominates $G(\cdot|\gamma')$ whenever $\gamma > \gamma'$. A low γ therefore indicates a low actual cost c.

A specific version of our model is the *independent case*, where the distribution of the shock is independent of γ. Formally,

$$G(c \mid \gamma) = \hat{G}(c - \gamma), \tag{1}$$

where $\hat{G}$ is the distribution function of the shock s. For expositional clarity, we focus on the independent case whenever this does not affect results qualitatively. The independent case is natural when the cost shock is independent of the agent's average efficiency characteristics such as expertise, experience, or internal organization. A specific example is an infrastructure project and the uncertainty pertaining to soil conditions or ground water levels. The independent case automatically satisfies the first order stochastic dominance ranking of $\{G(\cdot|\gamma)\}_\gamma$.

After the contract has been signed and before production takes place, the agent can, at a cost $k > 0$, perfectly observe s and therefore learn the actual cost c. While the value of k is common knowledge, the principal cannot observe which information the agent receives. Hence, from the principal's perspective, there is adverse selection at the contracting as well as after the information acquisition stage. We will consider both observable and unobservable information

[7] The additive specification is without loss of generality because the shock can always be defined as the difference between true and expected costs. The assumption that the support of the shock, and thus of true costs, is $\mathbb{R}$ is for convenience only and standard in the literature.

[8] Throughout we assume, for technical reasons, that all partial derivatives of h, G, and g exist and are bounded.

acquisition. In the latter case, there is an additional moral hazard problem at the information acquisition stage. Finally, we assume that at the contracting stage the principal makes a take–it or leave–it offer. If the agent rejects, he gets a type–independent outside option of zero.

3 First–best

As a benchmark, we first consider a first–best world, in which information acquisition and all available information is publicly observable. In a first–best world, the two crucial questions are for which cost estimates γ the agent should acquire additional information and when to execute the public project. Hence, a first–best contract specifies, first, for each agent type γ a probability $\alpha^{FB}(\gamma)$ with which the agent acquires information. Second, it specifies a probability $q^{FB}(\gamma, c)$ with which the project is executed by an agent who has acquired information and whose true costs turn out to be c. Finally, the contract determines a probability $\bar{q}^{FB}(\gamma)$ with which an agent who has not acquired information executes the project.

The answer to the question when to execute the project is straightforward. Depending on whether true or expected costs are above or below the valuation v, the production probabilities in the first best are either zero or one. Formally,

$$q^{FB}(\gamma, c) = \begin{cases} 1 & if \quad v \geq c \\ 0 & if \quad v < c \end{cases}, \qquad \bar{q}^{FB}(\gamma) = \begin{cases} 1 & if \quad v \geq \gamma \\ 0 & if \quad v < \gamma \end{cases}. \tag{2}$$

With these efficient implementation decisions, we can determine the first–best information acquisition levels and understand how, in a first best world, information is used.

Depending on the implementation decision in the absence of additional information, the acquisition of information plays one of two roles. First, if the decision without information acquisition is to implement the project, then the procurer commits a *type I error* if c turns out to be larger than v. In this case, a loss $v - c$ is realized. Information acquisition prevents this error. Formally, the (gross) expected value of information to prevent type I errors is

$$J_I^{FB}(\gamma) \equiv \int_v^\infty (c - v)\, dG(c \mid \gamma). \tag{3}$$

Second, if the decision without information acquisition is to cancel the project, then the procurer makes a *type II error* if c turns out to be smaller than v. Information acquisition prevents this error, and thus the (gross) expected value of information to prevent type II errors is

$$J_{II}^{FB}(\gamma) \equiv \int_{-\infty}^{v} (v - c)\, dG(c \mid \gamma). \tag{4}$$

According to the first–best implementation decision (2), the procurer optimally implements the project without additional information whenever $\gamma \leq v$. Hence, the *first–best value of information* is

$$J^{FB}(\gamma) \equiv \begin{cases} J_I^{FB}(\gamma) & if \quad \gamma \in [0, v] \\ J_{II}^{FB}(\gamma) & if \quad \gamma \in (v, \bar{\gamma}]. \end{cases} \tag{5}$$

Because γ is an unbiased estimate of actual costs, J^{FB} is continuous at $\gamma = v$.[9] Moreover, the first order stochastic dominance ranking of $\{G(\cdot|\gamma)\}_\gamma$ implies that J^{FB} is increasing in the range $\gamma \in [0, v]$ and decreasing in the range $\gamma \in (v, \bar{\gamma}]$. Hence, J^{FB} is single peaked with its maximum at $\gamma = v$. Let $k^{FB} \equiv J^{FB}(v)$ denote this maximum. Figure 1 illustrates the typical shape of the curve J^{FB}.

Information acquisition is efficient for all γ with $J^{FB}(\gamma) \geq k$. Because J^{FB} attains a maximum at $\gamma = v$, information acquisition is never efficient if $k \geq k^{FB}$. Whenever $k < k^{FB}$, single–peakedness of J^{FB} implies that there exist exactly two cut–offs $\gamma_1^{FB} < v < \gamma_2^{FB}$, which satisfy

$$J_I^{FB}(\gamma_1^{FB}) = k, \qquad J_{II}^{FB}(\gamma_2^{FB}) = k. \tag{6}$$

Thus information acquisition is efficient if $\gamma \in [\gamma_1^{FB}, \gamma_2^{FB}]$. The next proposition summarizes these results, and Figure 1 illustrates.

Proposition 1 *The first–best implementation probabilities are given by (2). Moreover, it holds:*

(i) If $k < k^{FB}$, the first–best information acquisition probabilities are

$$\alpha^{FB}(\gamma) = \begin{cases} 1 & if \quad \gamma \in [\gamma_1^{FB}, \gamma_2^{FB}] \\ 0 & if \quad \gamma \notin [\gamma_1^{FB}, \gamma_2^{FB}] \end{cases}. \tag{7}$$

(ii) If $k \geq k^{FB}$, the first–best information acquisition probability is $\alpha^{FB}(\gamma) = 0$ for all γ.

[9]Note: $\int_v^\infty (c - v) dG(c \mid \gamma) = \int_{-\infty}^v (v - c) dG(c \mid \gamma) \Leftrightarrow \int_{-\infty}^\infty (c - v) dG(c \mid \gamma) = 0 \Leftrightarrow \gamma - v = 0$.

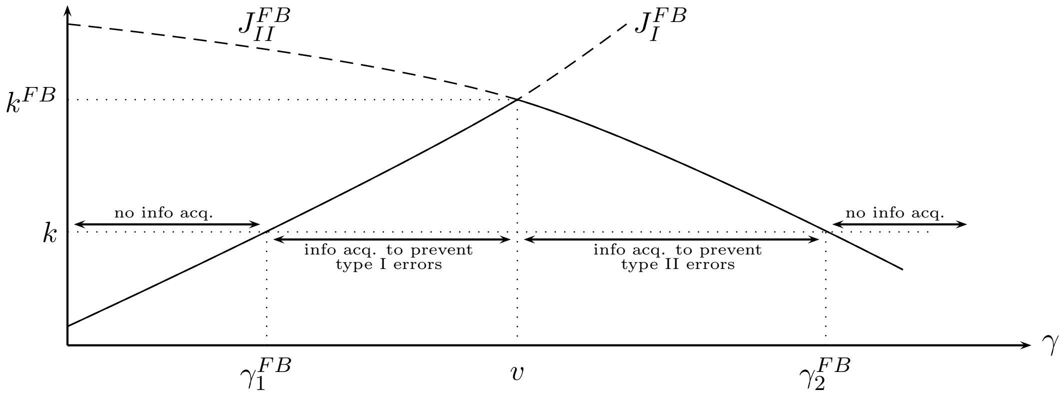

Figure 1: First–best information acquisition

4 The principal's problem

We now return to the original problem with asymmetric information and look for the contract that maximizes the principal's payoff. A contract specifies a transfer and the probability with which the agent has to produce the good in period 2, which, in principle, can be made contingent on arbitrary forms of communication between the parties. A crucial step to reduce the number of possible contracts is to apply the revelation principle for multistage games (Myerson 1986), which asserts that the optimal contract can be found in the class of direct, incentive compatible mechanisms.

A *direct* mechanism has the following structure: First, it requires the agent to submit a report $\hat{\gamma} \in \Gamma$ in period 1. Subsequently, the mechanism gives a contingent, possibly probabilistic, recommendation to the agent whether or not to acquire information. Let $\alpha(\hat{\gamma})$ be the probability with which the contract recommends information acquisition. For the case that information acquisition is not recommended, the contract specifies the transfers $\bar{t}(\hat{\gamma})$ from the principal to the agent and the probability of production $\bar{q}(\hat{\gamma})$. When information acquisition is recommended, the agent is required to submit a second report $\hat{c} \in \mathbb{R}$ in period 2, resulting in transfers $t(\hat{\gamma}, \hat{c})$ and the probability of production $q(\hat{\gamma}, \hat{c})$. Thus, a direct contract is a combination $(\alpha, \bar{t}, \bar{q}, t, q)$.

A direct contract is *incentive compatible* when it induces *truthtelling* and *obedience*. Truthtelling means that, on the equilibrium path, the agent has an incentive to report his new information truthfully. Obedience means that, on the equilibrium path, the contract must give the agent an incentive to follow the contract's recommendation whether or not to acquire information.

Formally, let $\Gamma^I = \{\gamma \in \Gamma \mid \alpha(\gamma) > 0\}$ denote the set of agent types that acquire information with a strictly positive probability. Then truthtelling in period 2 requires for all $\gamma \in \Gamma^I$ that

$$t(\gamma, c) - cq(\gamma, c) \geq t(\gamma, \hat{c}) - cq(\gamma, \hat{c}) \qquad \forall c, \hat{c} \in \mathbb{R}. \tag{AS2}$$

To state the first period truthtelling constraints, let $U(\gamma)$ denote the utility of agent type γ if he reports truthfully and obeys the contract's recommendation:

$$U(\gamma) \equiv \alpha(\gamma) \left[\int_{-\infty}^{\infty} \{t(\gamma, c) - cq(\gamma, c)\} dG(c \mid \gamma) - k \right] + (1 - \alpha(\gamma))[\bar{t}(\gamma) - \gamma \bar{q}(\gamma)]. \tag{8}$$

Incentive compatibility means that the agent cannot attain a higher utility than $U(\gamma)$ by adopting an untruthful reporting and/or disobedient information acquisition strategy. Notice that whatever the agent does in period 1, if he is required to submit a report in period 2, the second period constraints (AS2) guarantee that he reports truthfully in period 2.[10] Hence, even though the revelation principle does not require it, our setup yields truthtelling also off the equilibrium path.[11] With this in mind, we now consider all the deviations which incentive compatibility is meant to prevent and classify them in three different groups.

First, an agent type γ must not gain by reporting some type $\hat{\gamma}$ and, subsequently, obey the contract's information acquisition recommendation:

$$U(\gamma) \geq \alpha(\hat{\gamma}) \left[\int_{-\infty}^{\infty} \{t(\hat{\gamma}, c) - cq(\hat{\gamma}, c)\} dG(c \mid \gamma) - k \right] + (1 - \alpha(\hat{\gamma}))[\bar{t}(\hat{\gamma}) - \gamma \bar{q}(\hat{\gamma})], \forall \gamma, \hat{\gamma} \in \Gamma. \tag{AS1}$$

Moreover, an agent type γ must not gain by reporting $\hat{\gamma}$ and then disobeying when the contract requires him to acquire information:

$$U(\gamma) \geq \alpha(\hat{\gamma})[t(\hat{\gamma}, \gamma) - \gamma q(\hat{\gamma}, \gamma)] + (1 - \alpha(\hat{\gamma}))[\bar{t}(\hat{\gamma}) - \gamma \bar{q}(\hat{\gamma})] \qquad \forall \gamma, \hat{\gamma} \in \Gamma. \tag{MH}$$

Finally, an agent type must not gain by disobeying when the contract requires him not to acquire information. However, this cannot be optimal for any agent type, because when no

[10] E.g., expected utility of a type γ, who is ignorant of his actual costs c and reports some costs $\hat{c}$, is $t(\gamma, \hat{c}) - q(\gamma, \hat{c}) \int c dG(c|\gamma) = t(\gamma, \hat{c}) - q(\gamma, \hat{c})\gamma$. According to (AS2) his payoff is maximized for a report $\hat{c} = \gamma$.

[11] This would be different if the support of final costs c depended on the first period information γ. See Krähmer and Strausz (2008) for a discussion of the case when the supports of final costs do not overlap.

information acquisition is recommended, transfers and implementation probabilities do not condition on the additional cost information c so that the value of information for the agent is zero. Thus, the agent would only lose k.

The classification into truthtelling and obedience constraints allows us to distinguish between observable and unobservable information acquisition. If information acquisition is observable, the contract can enforce the agent's obedience directly. In this case, we are left with the truthtelling constraints (AS2) and (AS1) while the constraints (MH) are redundant. Consequently, we refer to (AS2) as *second period* and to (AS1) as *first period adverse selection constraints*. Because the obedience constraints (MH) only arise when information acquisition is unobservable, we refer to them as *moral hazard constraints*.

To guarantee the agent's participation in an incentive compatible contract, it needs to be *individually rational*:

$$U(\gamma) \geq 0 \qquad \forall \gamma \in \Gamma. \tag{IR}$$

As is standard in the sequential screening literature, we require individual rationality from an ex ante perspective only. We call an incentive compatible contract that is individually rational *feasible*.

The principal's payoff from a feasible contract is the difference between the total surplus and the agent's utility. That is, when the agent is of type γ, the principal's payoff is

$$W(\gamma) \equiv \alpha(\gamma) \left[\int_{-\infty}^{\infty} [v - c] q(\gamma, c)\, dG(c \mid \gamma) - k \right] + (1 - \alpha(\gamma))[v - \gamma]\bar{q}(\gamma) - U(\gamma), \tag{9}$$

and the principal's objective is his expected payoff

$$W \equiv \int_0^{\bar{\gamma}} W(\gamma)\, dF(\gamma). \tag{10}$$

The principal's problem with unobservable information acquisition, referred to as $\mathcal{P}$, can therefore be stated as follows:

$$\mathcal{P}: \quad \max_{(\alpha, \bar{t}, \bar{q}, t, q)} W \quad s.t. \qquad (AS2), (AS1), (MH), (IR). \tag{11}$$

With observable information acquisition, the principal's problem is a relaxed version of $\mathcal{P}$, referred to as $\mathcal{R}$:

$$\mathcal{R}: \quad \max_{(\alpha, \bar{t}, \bar{q}, t, q)} W \quad s.t. \qquad (AS2), (AS1), (IR). \tag{12}$$

5 Observable Information Acquisition

We first solve the principal's problem $\mathcal{R}$ where information acquisition is observable. Our procedure is similar to the well–known approach for solving static screening problems without additional ex post information.[12] In static problems, when the agent's cost function satisfies the single–crossing property, then incentive compatibility is equivalent to a monotone allocation rule and the fact that, up to the utility of the least efficient agent, the agent's utility is determined by the allocation alone. We begin by showing that this property carries over to the second period adverse selection constraints.

Let $u^I(\gamma, c) \equiv t(\gamma, c) - cq(\gamma, c) - k$ denote the agent's utility when he is informed. It then follows:

Lemma 1 *Let $\gamma \in \Gamma^I$. Then there are transfers $t(\gamma, c)$ such that (AS2) holds if and only if*

$$q(\gamma, c) \text{ is non–increasing in } c, \tag{MON2}$$

$$\partial u^I(\gamma, c)/\partial c = -q(\gamma, c) \quad a.e. \tag{13}$$

The proof of Lemma 1 is standard and therefore omitted. In static screening problems, the characterization of incentive compatibility in terms of monotonicity and the agent's utility implies that in the principal's problem, transfers can be eliminated both in the objective and in the constraints. In contrast, the first period adverse selection constraints (AS1) cannot be characterized in terms of monotonicity conditions of the allocation rule. The reason is that the agent's utility is given by an expectation over his cost function and so depends on the whole schedule of allocations instead of a single, type specific allocation only. This leaves the single–crossing property without bite. However, the next lemma demonstrates that (AS1) together with (13) still imply that the agent's utility is determined by the allocation alone. This will later allow us to eliminate transfers in the principal's objective but not in the constraints.

Lemma 2 *Under (AS1) and (13), the derivative $U'(\gamma)$ exists for almost all $\gamma \in \Gamma$ and whenever it exists, it equals*

$$U'(\gamma) \;=\; \alpha(\gamma) \int_{-\infty}^{\infty} \frac{\partial G(c \mid \gamma)}{\partial \gamma} q(\gamma, c)\, dc - (1 - \alpha(\gamma))\bar{q}(\gamma). \tag{14}$$

[12] This approach is also adopted in Courty and Li (2000).

Lemma 2 follows by a standard envelope argument. Since $U(\gamma) = -\int_{\gamma}^{\bar{\gamma}} U'(z)dz + U(\bar{\gamma})$, the agent's utility is thus determined up to the least efficient agent's utility $U(\bar{\gamma})$ which leaves us to determine $U(\bar{\gamma})$. In a standard static problem, the single–crossing condition implies that the utility of the least efficient type is pinned down by the individual rationality constraint. Again, since in our context the agent's utility is given by an expectation over a whole range of allocations, the single–crossing property cannot be used to determine $U(\bar{\gamma})$. Instead, we exploit that $\{G(\cdot \mid \gamma)\}_{\gamma}$ is ranked by first order stochastic dominance. It implies that $\partial G(c \mid \gamma)/\partial\gamma \leq 0$, and so by Lemma 2, $U'(\gamma) \leq 0$. We therefore obtain the following result.

Lemma 3 *Under (AS1) and (13), (IR) is equivalent to* $U(\bar{\gamma}) \geq 0$.

Clearly, at the optimal contract it must hold that $U(\bar{\gamma}) = 0$. This condition then takes care of the individual rationality constraint and determines agent type γ's utility by Lemma 2. Hence, we can substitute out $U(\gamma)$ in the objective W. By applying a common integration by parts argument, we obtain the objective as a function of the implementation and the information acquisition probabilities only.[13] This allows us to re–state the problem $\mathcal{R}$ as follows:

$$\mathcal{R}: \quad \max_{(\alpha,\bar{t},\bar{q},t,q)} \quad \int_0^{\bar{\gamma}} \left\{ \alpha(\gamma) \left(\int_{-\infty}^{\infty} \left[v - c + \frac{\partial G(c \mid \gamma)/\partial\gamma}{g(c \mid \gamma)} h(\gamma) \right] q(\gamma, c)\, dG(c \mid \gamma) - k \right) \right. \tag{15}$$

$$\left. + (1 - \alpha(\gamma))[v - \gamma - h(\gamma)]\bar{q}(\gamma) \right\} dF(\gamma)$$

$$s.t. \qquad (AS1), (MON2).$$

We stress again that even though we have inserted the utility expression (14) derived from (AS1) into the objective, this does not eliminate the constraint (AS1), because it is not equivalent to a monotonicity condition of the allocation. This is different for the constraint (AS2) which is equivalent to the constraints (MON2) and (13). Since we have inserted (13) into the objective, we are left with (MON2).

Before we solve the principal's problem, it is helpful to interpret the objective (15). We can think of the principal as maximizing total surplus where, instead of true costs, he faces higher *virtual costs* that arise because an information rent has to be conceded to the agent. If information acquisition does not take place, the virtual costs are $\gamma + h(\gamma)$. They are the same

[13] The details of the derivation are presented in the Appendix.

as the virtual costs in the static screening problem in which additional information cannot be acquired. As usual, the hazard rate $h(\gamma)$ measures the extent of asymmetric information between the agent and the principal about the expected costs γ. If information acquisition does take place, the virtual costs are $c - \psi(\gamma, c)h(\gamma)$ with

$$\psi(\gamma, c) \equiv \frac{\partial G(c \mid \gamma)/\partial \gamma}{g(c \mid \gamma)}. \tag{16}$$

They are the same as the virtual costs in a sequential screening problem in which each agent type exogenously observes the cost shock. Baron and Besanko (1984) interpret ψ as an *informativeness measure* which captures how the agent's private knowledge about the true cost distribution G changes across types. For the independent case, where the signal and the cost shock are independent, we have $\psi(\gamma, c) = -1$ so that virtual costs are simply $c + h(\gamma)$. This means that the possibility that the agent receives additional information in period 2 does not change the degree of asymmetric information in period 1. Because with observable information acquisition our qualitative results do not depend on the shape of ψ, we will, for expositional clarity, focus on the independent case. We return to the general case in Section 6, where we show that the effect of unobservable information acquisition depends crucially on ψ. For this reason, we nevertheless prove all our results for the general case.

We now turn to the solution of the principal's problem $\mathcal{R}$. In Subsection 5.1 we solve the unconstrained version of problem $\mathcal{R}$ where we ignore the constraints (AS1) and (MON2). In Subsection 5.2, we then check whether the solution actually satisfies these omitted constraints.

5.1 Solution to the unconstrained problem

The solution to the unconstrained version of problem $\mathcal{R}$ can be obtained by point–wise maximization for each γ in two steps. In the first step, the optimal implementation probabilities are determined for fixed $\alpha(\gamma)$. Clearly, this step amounts to setting the allocation $q(\cdot)$ to zero if the associated term in the squared brackets in (15) is strictly negative and setting it to one otherwise. This procedure yields:

Lemma 4 *For each γ there is a unique $c_0(\gamma) \leq v$ and there is a unique $\gamma_0 \leq v$ given by*[14]

$$v = c_0(\gamma) + h(\gamma), \qquad v = \gamma_0 + h(\gamma_0), \tag{17}$$

[14]In the general case, $c_0(\gamma)$ is determined by $v = c_0(\gamma) - \psi(\gamma, c_0(\gamma))h(\gamma)$.

such that the optimal implementation probabilities in the solution to the unconstrained version of problem $\mathcal{R}$ are given by

$$q^*(\gamma, c) = \begin{cases} 1 & if \quad c \le c_0(\gamma) \\ 0 & if \quad c > c_0(\gamma) \end{cases}, \qquad \bar{q}^*(\gamma) = \begin{cases} 1 & if \quad \gamma \le \gamma_0 \\ 0 & if \quad \gamma > \gamma_0 \end{cases}. \tag{18}$$

Lemma 4 reveals the typical distortions that are implied by the rent–efficiency trade–off that the principal faces. Observe that the second best (expected) cost thresholds γ_0 and $c_0(\gamma)$ are both smaller than the first–best cost threshold v. Hence, the principal distorts the implementation probabilities downwards relative to the first–best. The downward distortion lowers the information rent the principal needs to concede to the agent, because it reduces the extent to which a relatively efficient type could gain by mimicking a relatively inefficient type.

The second step is to determine the optimal information acquisition probabilities. Similarly to the first–best, the purpose of information acquisition is to prevent type I or type II implementation errors. Indeed, suppose that the principal implements the project in the absence of further information. In this case her payoff is $v - \gamma - h(\gamma)$. If information is available, the project is implemented if costs are smaller than $c_0(\gamma)$ in which case the principal obtains the payoff $v - c + \psi(\gamma, c)h(\gamma)$. Thus, the principal's (gross) value of information is

$$\begin{aligned} J_I^*(\gamma) &\equiv \int_{-\infty}^{c_0(\gamma)} v - c - h(\gamma)\, dG(c \mid \gamma) - [v - \gamma - h(\gamma)] & (19) \\ &= \int_{c_0(\gamma)}^{\infty} (c + h(\gamma) - v) dG(c \mid \gamma), & (20) \end{aligned}$$

where we have used that $v - \gamma - h(\gamma) = \int_{-\infty}^{\infty} v - c - h(\gamma) dG(c|\gamma)$. Expression (20) reveals that, due to information rents, the principal's value of information is distorted relative to the first–best: From the principal's perspective, a type I error occurs if *virtual* costs, $c + h(\gamma)$, are larger than v, whereas from a first–best perspective, it occurs if *true* costs, c, are larger than v. Note that the definition (17) implies that c_0 is non–increasing in γ because h is non–increasing in γ and c. Moreover, h is non–decreasing so that the first order stochastic dominance ranking of $\{G(\cdot \mid \gamma)\}_\gamma$ implies that $J_I^*(\gamma)$ is increasing.

Similarly, suppose that in the absence of information acquisition, the principal cancels the project. In this case, information acquisition prevents a type II error whenever the value of the project v exceeds the virtual costs $c + h(\gamma)$, and the principal's (gross) value of information is

$$J_{II}^*(\gamma) \equiv \int_{-\infty}^{c_0(\gamma)} (v - c - h(\gamma)) dG(c \mid \gamma). \tag{21}$$

Because h is non–decreasing and c_0 is non–increasing, the first order stochastic dominance ranking of $\{G(\cdot \mid \gamma)\}_\gamma$ implies that $J^*_{II}(\gamma)$ is decreasing.

According to Lemma 4, the principal executes the project without additional information exactly when $\gamma \leq \gamma_0$. Consequently, the principal's (gross) value of information is

$$J^*(\gamma) \equiv \begin{cases} J^*_I(\gamma) & if \quad \gamma \in [0, \gamma_0] \\ J^*_{II}(\gamma) & if \quad \gamma \in (\gamma_0, \bar{\gamma}]. \end{cases} \tag{22}$$

Notice that J^* is continuous in $\gamma = \gamma_0$.[15] Moreover, since J^*_I is increasing and J^*_{II} decreasing, $J^*(\gamma)$ is single–peaked with a maximum at γ_0 of

$$k^* \equiv J^*(\gamma_0) = \int_{-\infty}^{c_0(\gamma_0)} (v - c - h(\gamma_0))dG(c \mid \gamma_0).$$

Information acquisition by an agent type γ is optimal for the principal if and only if $J^*(\gamma) \geq k$. Since J^* is at most k^*, information acquisition is never optimal if $k \geq k^*$. If $k < k^*$, then single–peakedness of J^* implies that there are exactly two cut–offs $\gamma^*_1 < \gamma_0 < \gamma^*_2$, which satisfy

$$J^*(\gamma^*_1) = k, \qquad J^*(\gamma^*_2) = k. \tag{23}$$

Thus, it is optimal to induce information acquisition if $\gamma \in [\gamma^*_1, \gamma^*_2]$. The following lemma summarizes these results.

Lemma 5 *The optimal information acquisition probabilities in the solution to the unconstrained version of problem $\mathcal{R}$ are given as follows:*

(i) If $k < k^*$, then

$$\alpha^*(\gamma) = \begin{cases} 1 & if \quad \gamma \in [\gamma^*_1, \gamma^*_2] \\ 0 & if \quad \gamma \notin [\gamma^*_1, \gamma^*_2] \end{cases}. \tag{24}$$

(ii) If $k \geq k^*$, then $\alpha^*(\gamma) = 0$ for all γ.

5.2 Solution to the constrained problem

We now address the question whether our solution to the relaxed version of $\mathcal{R}$ actually satisfies the constraints (AS1) and (MON2). In static problems, this simply amounts to checking a

[15]Recall that by definition of γ_0: $v - \gamma_0 - h(\gamma_0) = 0$. Hence, by (19) and (21), $J^*_I(\gamma_0) = J^*_{II}(\gamma_0)$.

monotonicity condition, which in turn is guaranteed by a monotone hazard rate. Recall from our earlier discussion that in a sequential screening problem, the first period adverse selection constraints (AS1), in general, cannot be characterized in terms of monotonicity. It turns out, however, that what prevents such a characterization is that contracts are allowed to be stochastic. In fact, our next result demonstrates that for *deterministic* contracts, (AS1) can be characterized by a monotonicity condition.[16] Since the solution to the unconstrained problem is deterministic, this implies that it also solves the constrained problem.

We define a contract as *deterministic* when the decision to implement the project and the information acquisition recommendation are deterministic: $\alpha, \bar{q}, q \in \{0,1\}$. We say that a deterministic allocation $(\alpha, \bar{q}, q)$ is *non–increasing w.r.t.* $(\bar{q}, q)$ if there are γ_1, γ_2 with $\gamma_1 \leq \gamma_2$ and a non–increasing cutoff $\hat{c}_0 : [\gamma_1, \gamma_2] \to \mathbb{R}$ such that

$$\alpha(\gamma) = \begin{cases} 1 & if \quad \gamma \in [\gamma_1, \gamma_2] \\ 0 & if \quad \gamma \notin [\gamma_1, \gamma_2] \end{cases}, \quad \bar{q}(\gamma) = \begin{cases} 1 & if \quad \gamma < \gamma_1 \\ 0 & if \quad \gamma > \gamma_2 \end{cases}, \quad q(\gamma, c) = \begin{cases} 1 & if \quad c \leq \hat{c}_0(\gamma) \\ 0 & if \quad c > \hat{c}_0(\gamma). \end{cases} \tag{25}$$

We call an allocation that satisfies (25) non–increasing w.r.t. $(\bar{q}, q)$ because the implementation probabilities $(\bar{q}, q)$ are non–increasing in γ and c. The following lemma states our characterization result for deterministic contracts.

Lemma 6 *There are transfers $(\bar{t}, t)$ such that a deterministic contract $(\alpha, \bar{t}, \bar{q}, t, q)$ satisfies the adverse selection constraints (AS1) and (MON2) if and only if $(\alpha, \bar{q}, q)$ is non–increasing w.r.t. $(\bar{q}, q)$.*

Because our solution to the relaxed version of $\mathcal{R}$ displays a deterministic, non–increasing allocation w.r.t. $(\bar{q}, q)$, the lemma implies that transfers exist so that the corresponding contract satisfies (AS1) and (MON2).[17]

Proposition 2 *There are transfers $(\bar{t}^*, t^*)$ such that the contract $C^* = (\alpha^*, \bar{t}^*, \bar{q}^*, t^*, q^*)$ solves problem $\mathcal{R}$ and is, therefore, optimal.*

[16] To our knowledge, this characterization result is new to the sequential screening literature.

[17] In a model where q corresponds to a divisible quantity (and so a quantity between zero and one is still deterministic), incentive compatibility does, in general, not imply that q is non–increasing in γ and c. However, under standard regularity conditions, the optimal contract will still be deterministic. This holds also if the value of output is not linear in quantity.

5.3 Option contracts

In this subsection, we show that the implied transfers of Lemma 6 allow a reinterpretation of the contract as a menu that gives the agent the choice between a fixed price contract and various option contracts. Consequently, any deterministic, direct mechanism that satisfies the adverse selection constraints can also be implemented by an indirect, empirically more natural contract. This is true in particular of the optimal contract C^*.

To demonstrate this, we first show how the adverse selection constraints pin down transfers in Lemma 6. Consider first the range of agent types who do not acquire information. Since all types in $(\gamma_2, \bar{\gamma}]$, do not execute the project, (AS1) implies that all these types have to get the same transfer. If the individual rationality constraint is binding, this transfer is zero. On the other hand, all types in $[0, \gamma_1)$ execute the project with certainty. Therefore, (AS1) implies that transfers have to be the same for all types in $[0, \gamma_1)$, say $\bar{t}$. Hence, the first period adverse selection constraints (AS1) imply that the transfer schedule $\bar{t}(\gamma)$ exhibits

$$\bar{t}(\gamma) = \begin{cases} \bar{t} & if \quad \gamma < \gamma_1 \\ 0 & if \quad \gamma > \gamma_2. \end{cases} \tag{26}$$

Next, consider an agent type γ who acquires information. The transfer $t(\gamma, c)$ is now pinned down by the second period adverse selection constraints (AS2). Indeed, the project is executed if the cost realization c is smaller than the cutoff $\hat{c}_0(\gamma)$. Thus, (AS2) implies that all cost types $c \leq \hat{c}_0(\gamma)$ who execute the project have to get the same transfer, say $t_0(\gamma)$. Similarly, all cost types who abandon the project $(c > \hat{c}_0(\gamma))$ have to get the same transfer, say $t_1(\gamma)$. Moreover, the critical cost type $\hat{c}_0(\gamma)$ has to be indifferent between executing and abandoning the project. Otherwise, some types close to $\hat{c}_0(\gamma)$ would have incentives to lie. This pins down the difference between $t_0(\gamma)$ and $t_1(\gamma)$ so that $t_1(\gamma) = t_0(\gamma) + \hat{c}_0(\gamma)$. Hence, the second period adverse selection constraints (AS2) imply that the transfer schedule $t(\gamma, c)$ exhibits

$$t(\gamma, c) = \begin{cases} t_0(\gamma) & if \quad c > \hat{c}_0(\gamma) \\ t_0(\gamma) + \hat{c}_0(\gamma) & if \quad c \leq \hat{c}_0(\gamma). \end{cases} \tag{27}$$

Finally, the levels of t_0 and $\bar{t}$ are pinned down by the condition that the boundary type γ_1 (resp. γ_2) has to be indifferent between acquiring information and implementing (resp. not implementing) the project without acquiring information. If this was not the case, there would be types close to γ_1 (resp. γ_2) with an incentive to misreport. Observe that in pinning down transfers, we employed the *local* adverse selection constraints only, that is, that no type mimics a type close by. The proof of Lemma 6 shows that with a deterministic allocation which is non–increasing w.r.t. $(\bar{q}, q)$, these transfers actually satisfy the adverse selection constraints globally.

The shape of transfers allows the principal to implement the direct contract indirectly through a menu of contracts that consists of a fixed price contract and a range of option contracts. To see this, consider first what happens if the agent announces a type who does not acquire information. Announcing $\gamma > \gamma_2$ simply amounts to rejecting the contract, as it entails not executing the project and zero transfers. Announcing $\gamma < \gamma_1$ amounts to picking a fixed price contract which obliges the agent to complete the project at all cost circumstances for the price $\bar{t}$. Next, consider what happens if the agent announces a type who does acquire information. After his first report γ, the agent subsequently faces a choice between announcing a type $c > \hat{c}_0(\hat{\gamma})$ or a type $c \leq \hat{c}_0(\hat{\gamma})$. In the first case, the project is canceled and the agent receives $t_0(\hat{\gamma})$, while in the second case the project is implemented and the agent receives $t_0(\hat{\gamma}) + \hat{c}_0(\hat{\gamma})$. Hence, effectively the agent receives the transfer $t_0(\hat{\gamma})$ upfront and then has two options: walk away or complete the project for the price $\hat{c}_0(\hat{\gamma})$.

It consequently follows that the outcome of the optimal direct mechanism C^* is also implemented by the indirect menu contract $C' \equiv \{(\alpha = 0, \bar{t}^*), (\alpha = 1, c_0(\gamma), t_0^*(\gamma))_{\gamma \in [\gamma_1^*, \gamma_2^*]}\}$. The menu consists of the fixed price contract $(\alpha = 0, \bar{t}^*)$ and a range of option contracts $(\alpha = 1, c_0(\gamma), t_0^*(\gamma))_\gamma$. We summarize this discussion in the next proposition for the non–trivial case $k < k^*$.[18]

Proposition 3 *If $k < k^*$, then the outcome under the optimal contract C^* also obtains with the contract $C' = \{(\alpha^* = 0, \bar{t}^*), (\alpha^* = 1, c_0(\gamma), t_0^*(\gamma))_{\gamma \in [\gamma_1^*, \gamma_2^*]}\}$.*

5.4 Distortions in information acquisition

We next investigate the distortions in information acquisition. Recall that the principal, instead of maximizing overall surplus, is only interested in the share of the surplus that she can extract. Due to asymmetric information, the principal must leave a part of the surplus — the information

[18]For $k \geq k^*$, no information acquisition takes place, and the optimal contract is simply a fixed price contract.

rents — to the agent, and consequently she is also interested in how the amount of information acquisition affects the size of these rents. We now show that this *information rent effect* is intimately linked to the type of errors that information acquisition is meant to prevent. This allows us to identify the unambiguous direction of the distortions and develop a straightforward intuition for them.

First, suppose that information acquisition is used to prevent type I errors. In this case, the social value of information is $J_I^{FB}(\gamma)$ while the value to the principal is $J_I^*(\gamma)$. Because $h(\gamma) \geq 0$ and $c_0(\gamma) = v - h(\gamma) < v$, it follows

$$\begin{aligned} J_I^{FB}(\gamma) &= \int_v^{\infty} (c - v) dG(c \mid \gamma) && (28) \\ &\leq \int_v^{\infty} (c - v + h(\gamma)) dG(c \mid \gamma) \leq \int_{c_0(\gamma)}^{\infty} (c - v + h(\gamma)) dG(c \mid \gamma) && (29) \\ &= J_I^*(\gamma). && (30) \end{aligned}$$

The inequality shows that the principal overvalues information acquisition relative to the first best. That is, for preventing type I errors there is a positive information rent effect which increases the principal's value above the first best value of information.

The intuition is as follows. When additional information prevents type I errors, some projects are turned down that would otherwise have been implemented. This means that, from an ex ante perspective, information acquisition reduces the implementation probability q. A reduction in q for some cost type implies that, from a period 1 perspective, it becomes less worthwhile for a more efficient cost type to mimic this cost type. Hence, the principal has to pay lower information rents to the more efficient cost types when inducing information acquisition. As a result, the optimal contract displays excess information acquisition to prevent type I errors: $\gamma^* \leq \gamma_1^{FB}$.

Next, suppose information acquisition is used to prevent type II errors. In this case, the social value of information is $J_{II}^{FB}(\gamma)$ and the value to the principal is $J_{II}^*(\gamma)$. Because of $h(\gamma) \geq 0$ and $c_0(\gamma) = v - h(\gamma) < v$, it now follows

$$\begin{aligned} J_{II}^{FB}(\gamma) &= \int_{-\infty}^{v} (v - c) dG(c \mid \gamma) && (31) \\ &\geq \int_{-\infty}^{v} (v - c - h(\gamma)) dG(c \mid \gamma) \geq \int_{-\infty}^{c_0(\gamma)} (v - c - h(\gamma)) dG(c \mid \gamma) && (32) \\ &= J_{II}^*(\gamma). && (33) \end{aligned}$$

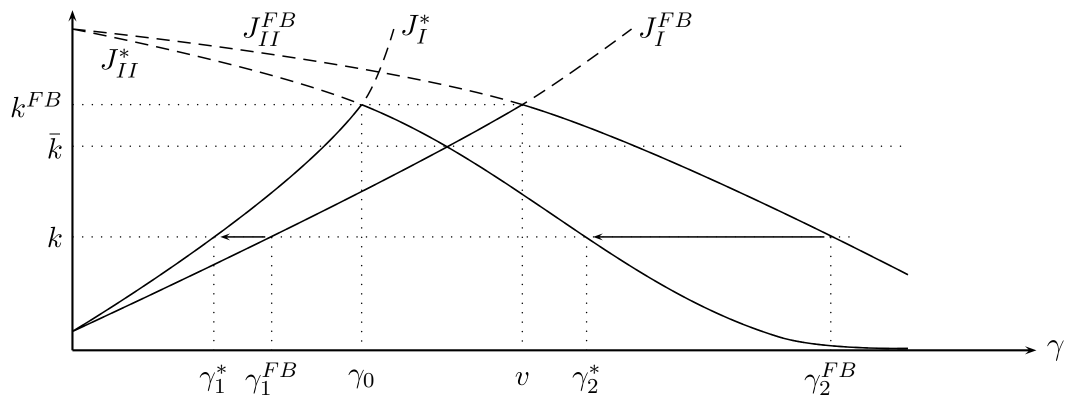

Figure 2: Distortions in information acquisition

The inequality shows that the principal undervalues information acquisition relative to the first best. That is, for preventing type II errors there is a negative information rent effect which decreases the principal's value below the first best value of information.

Although the sign of the information rent effect is now negative, the intuition behind the result follows from the same logic. When additional information prevents type II errors, some projects are implemented that would otherwise have been canceled. Hence, from an ex ante perspective, information acquisition increases the implementation probability q, and so the principal has to pay higher information rents to the more efficient cost types when inducing information acquisition. As a result, the optimal contract displays too little information acquisition to prevent type II errors: $\gamma_2^* \leq \gamma_2^{FB}$. We summarize this key result in Proposition 4.

Proposition 4 *Under the optimal contract, information acquisition is distorted so that there is excess (resp. insufficient) information acquisition to prevent types I (resp. type II) errors:* $\gamma_1^* \leq \gamma_1^{FB}$ *and* $\gamma_2^* \leq \gamma_2^{FB}$.

Figure 2 illustrates the distortions. The curve J_I^* lies higher than the curve J_I^{FB}, whereas the curve J_{II}^* lies below the curve J_{II}^{FB}. This implies that for a given k, the set of types who acquire information the second–best, $[\gamma_1^*, \gamma_2^*]$ lies to the left of the set of types who acquire information in the first–best, $[\gamma_1^{FB}, \gamma_2^{FB}]$. It further shows how the two sets shrink as k rises. As of $\bar{k}$ the sets $[\gamma_1^*, \gamma_2^*]$ and $[\gamma_1^{FB}, \gamma_2^{FB}]$ are disjoint: none of the types who acquire information in the second best acquire information in the first best. This demonstrates that distortions may lead to qualitatively different outcomes.[19]

6 Unobservable information acquisition

In this section, we examine the case when information acquisition is unobservable, and thus the optimal contract must also satisfy the moral hazard constraints (MH). The two key questions are whether the moral hazard problem causes the principal additional agency costs and how this affects distortions. We proceed in two steps. We first characterize when the optimal contract with observable information acquisition automatically satisfies (MH). Whenever this is the case, we say that *the moral hazard problem does not cause additional agency costs*. In the second step, we discuss how the optimal contract changes when the moral hazard problem causes additional agency costs and affects distortions.

Because subsequent results depend crucially on the conditional cost distributions $G(c|\gamma)$, there is a loss in restricting attention to the independent case. For this reason, we now consider more general distributions G for which the informativeness measure $\psi(\gamma, c)$ is no longer constant. In order to extend Lemma 4 to more general distributions, we impose the regularity conditions that $\partial\psi/\partial c \cdot h < 1$ and that ψ is non–increasing in γ. The proof of Lemma 4 shows that the first condition guarantees a unique cost type $c_0(\gamma)$ that solves $v - c_0(\gamma) + \psi(\gamma, c_0(\gamma))h(\gamma) = 0$. The second condition then implies that $c_0(\gamma)$ is non–increasing in γ.[20] Since $\{G(c|\gamma)\}_\gamma$ is ranked by first order stochastic dominance, $\psi(\gamma, c)$ is negative. Finally, we require that for all $\gamma \in \Gamma$:

$$\int_{-\infty}^{\infty} \psi(\gamma, c)\, dG(c \mid \gamma) = -1. \tag{34}$$

This is only a mild, technical condition. It is satisfied for any family of distribution functions with bounded support and, therefore, is a natural continuity requirement.[21] It guarantees that, similar to (22), the value of information J^* is the respective integral over the virtual surplus.

[19] Observe that in Figure 2, the first best and the optimal critical cost values are identical: $k^{FB} = k^*$. This is a special feature of the independent case which implies that $J^*(\gamma) = J^{FB}(\gamma + h(\gamma))$. The derivation of this relation is in the appendix.

[20] Differentiating the identity with respect to γ yields: $c_0'(1 - \partial\psi/\partial c \cdot h) = \partial\psi/\partial\gamma \cdot h + \psi \cdot h'$. Hence, $c_0' \leq 0$.

[21] For this reason, a sufficient condition for (34) is that $G(c|\gamma)$ converges fast enough to 1 as $c \to +\infty$ and fast enough to 0 as $c \to -\infty$. It then follows by integration by part: $\int_{-\infty}^{\infty} \psi(\gamma, c)\, dG(c|\gamma) = \frac{\partial}{\partial\gamma} \int_{-\infty}^{\infty} G(c\ \gamma)dc = \frac{\partial}{\partial\gamma}[cG(c|\gamma)]_{-\infty}^{\infty} - \frac{\partial}{\partial\gamma} \int_{-\infty}^{\infty} c\, dG(c \mid \gamma) = -1$.

6.1 Agency costs through moral hazard

We turn to the question when the optimal contract with observable information acquisition, C', automatically satisfies the moral hazard constraints (MH). In what follows we refer to C' as the *benchmark contract.* For the non–trivial case $k < k^*$, we provide two conditions for the benchmark contract to satisfy (MH), one in terms of transfers and one in terms of the informativeness measure ψ.

We first establish two straightforward necessary conditions on the cutoff $c_0(\gamma)$ and transfers $\bar{t}^*$ and $t_0^*(\gamma)$ so that the moral hazard problem does not cause additional agency costs. The first condition is that $t_0^*(\gamma) \leq 0$ for all $\gamma \in [\gamma_1^*, \gamma_2^*]$. The condition is necessary, because if $t_0^*(\gamma)$ were strictly positive for some γ, then a type $\gamma' \geq \gamma_2^*$, instead of rejecting the contract and receiving zero, would have a strict incentive to announce type γ in order to receive the upfront transfer $t_0^*(\gamma) > 0$ and, subsequently, quit the project without acquiring information. The second condition is that the net transfer when executing the project must not be higher than the transfer under the fixed price contract, that is, $\bar{t}^* \geq c_0(\gamma) + t_0^*(\gamma)$ for all $\gamma \in [\gamma_1^*, \gamma_2^*]$. The condition is necessary, because if $\bar{t}^*$ were smaller than $c_0(\gamma) + t_0^*(\gamma)$ for some γ, then a relatively efficient type $\gamma' < \gamma_1^*$ would have a strict incentive to report type γ rather than his true type γ' and then, without acquiring information, choose the option to implement the project. This strategy would allow him to complete the project for an overall transfer $c_0(\gamma) + t_0^*(\gamma)$ instead of the lower transfer $\bar{t}^*$.

We now argue that these two conditions are not only necessary but also sufficient for the moral hazard problem to not cause additional agency costs. To show sufficiency, it remains to be checked that an agent type, $\gamma \in [\gamma_1^*, \gamma_2^*]$, who is supposed to acquire information actually does so. There are two deviations to consider. First, suppose that such an agent type picks an option contract but deviates by quitting the project without acquiring information. The deviator ends up with the non–positive transfer $t_0^*(\gamma)$ and rejecting the contract would be a weakly better deviation. But since C' is individually rational, rejecting the contract cannot be profitable. Second, suppose that an agent type $\gamma \in [\gamma_1^*, \gamma_2^*]$ picks an option contract but deviates by completing the project without acquiring information. If this deviation is profitable, then, due to $\bar{t}^* \geq c_0(\gamma) + t_0^*(\gamma)$, an even more profitable deviation is to pick the fixed price contract. But because the benchmark contract satisfies the adverse selection constraints (AS1), this second deviation is also not profitable. Thus, we have established:

Lemma 7 *The moral hazard problem does not cause additional agency costs if and only if for all* $\gamma \in [\gamma_1^*, \gamma_2^*]$*:*

$$\bar{t}^* \geq c_0(\gamma) + t_0^*(\gamma) \quad \text{and} \quad t_0^*(\gamma) \leq 0. \tag{35}$$

While intuitive, the previous condition has the drawback that it depends on transfers which are endogenous. In the next lemma, we, therefore, give a sufficient condition that is only based on the primitive ψ.

Lemma 8 *If* $\partial\psi/\partial c < 0$*, then the moral hazard problem causes additional agency costs. If* $\partial\psi/\partial c \geq 0$*, then the moral hazard problem does not cause additional agency costs. In particular, the moral hazard problem does not cause additional agency costs in the independent case.*

Lemma 8 shows that the sign of the partial derivative $\partial\psi/\partial c$ determines whether or not the moral hazard problem causes additional agency costs. We shall explain the intuition behind this result in detail, as it will guide us in constructing the optimal contract when the moral hazard problem causes additional agency costs.

Intuitively, the agent has an incentive to acquire information when his value of information exceeds acquisition costs k. Note that when the agent chooses an option contract, information is valuable to the agent, because the contract allows him to choose the best option according to true cost conditions. Hence, the crucial question is to what extent the agent's value of information coincides with the principal's. To answer this, consider the value of information to some agent type γ. Given an option contract, type γ's next best alternative without information is either to execute the project for a price $p = c_0(\gamma)$ or to quit the project. Under the first alternative, an informed type γ saves $p - c$ when the additional information reveals that the cost c exceeds p. Therefore, the agent's value of information in this case is

$$J_I^A(p, \gamma) \quad = \quad \int_p^\infty c - p \, dG(c \mid \gamma), \tag{36}$$

where, under the benchmark contract, p equals $c_0(\gamma)$. When the next best alternative is to quit the project, an informed type γ gains $p - c$ when the additional information reveals that the cost c is smaller than the price p. Hence, the agent's value of information in this case is

$$J_{II}^A(p, \gamma) \quad = \quad \int_{-\infty}^p p - c \, dG(c \mid \gamma). \tag{37}$$

Now, the moral hazard problem does not cause additional agency costs when the agent's value of information is sufficiently large. More precisely, suppose that the agent type γ's value of information is larger than k whenever the principal's value of information $J^*(\gamma)$ is also larger than k. In that case, the agent's and the principal's incentives to acquire information are aligned, and the agent voluntarily acquires information. Hence, if $J_I^A(c_0(\gamma), \gamma) \geq k$ and $J_{II}^A(c_0(\gamma), \gamma) \geq k$ for all $\gamma \in [\gamma_1^*, \gamma_2^*]$, then there are no additional agency costs.

In order to see that the sign of $\partial\psi/\partial c$ determines whether this is the case, consider a situation where both the principal's and the agent's best alternative without information is to abandon the project. In that case, the principal's value of information is her virtual valuation integrated over the cost range $c < c_0(\gamma)$ while the agent's value of information is his "gross ex post rent" $c_0(\gamma) - c$ integrated over the same range. Now observe that both the virtual surplus and the gross ex post rent are exactly zero at $c_0(\gamma)$, and in the range $c < c_0(\gamma)$ the ex post gross rent increases at a rate of $+1$ as c becomes smaller, while the principal's virtual surplus increases at the rate of $1 - \partial\psi/\partial c \cdot h$. Hence, for the independent case $\partial\psi/\partial c = 0$, these rates coincide and therefore also the agent's and the principal's value of information. For $\partial\psi/\partial c > 0$, the agent's value of information is actually larger than the principal's so that the agent has a strict incentive to acquire information whenever the principal wants him to do so. If, on the other hand, $\partial\psi/\partial c < 0$, then the principal's value of information is larger than the agent's. In that case, incentives for information acquisition are misaligned for some types γ, and the benchmark contract has to be adapted to account for the moral hazard problem. We now turn to this issue.

6.2 Optimal contract when moral hazard causes agency costs

Finding the optimal contract when the moral hazard problem causes additional agency costs is complicated in general because the adverse selection constraint cannot be characterized in terms of monotonicity. However, as shown in Lemma 6, for deterministic contracts such a characterization is possible. This will allow us to solve the problem when we restrict attention to deterministic contracts. In light of Lemma 6, we can associate with any deterministic contract that satisfies the adverse selection constraints a non–increasing cutoff $\hat{c}_0 : [\gamma_1, \gamma_2] \to \mathbb{R}$. Recall from section 5.3 that any such contract can be implemented by a menu of option contracts where the option *price* to execute the project coincides with the *cutoff* $\hat{c}_0(\gamma)$. In what follows, we therefore use the terms price and cutoff interchangeably.

With deterministic contracts there are, in effect, two ways how to modify the benchmark contract when it violates the moral hazard constraints. First, the principal could limit information acquisition to those agent types whose value of information is already high enough. In this case, additional agency costs arise since information acquisition sometimes does not take place even though the principal's value of information exceeds k. The other possibility is to adapt the price $p = c_0(\gamma)$ so as to raise the agent's value of information. Due to incentive compatibility, a price increase (resp. decrease) means, however, that the agent executes (resp. abandons) projects with a negative (resp. positive) virtual surplus. Hence, additional agency costs also arise from this second possibility. The intuitive idea for finding the optimal contract is to adapt the information acquisition interval and the price so that the two different types of agency costs are optimally balanced.

We now formalize this idea. We first characterize all prices that give agents a positive incentive to acquire information and look for the optimal price in this set. Given these optimal prices, we can then determine the optimal range of information acquisition $[\gamma_1, \gamma_2]$. We begin by showing that prices that induce information acquisition by the agent lie between two bounds, p_I and p_{II}. Recall from (36) and (37) that confronted with a price p, the agent acquires information if both $J_I^A(p,\gamma)$ and $J_{II}^A(p,\gamma)$ are larger than k; otherwise he would be better off by taking the next best alternative without information. We have:

Lemma 9 *Let* $i = I, II$.
(i) For each γ*, there is a unique solution* $p_i(\gamma)$ *to* $J_i^A(p,\gamma) = k$.
(ii) p_i *is increasing in* γ.
(iii) $J_I^A(p,\gamma) \geq k$ *if and only if* $p \leq p_I(\gamma)$*, and* $J_{II}^A(p,\gamma) \geq k$ *if and only if* $p \geq p_{II}(\gamma)$.

Lemma 9 implies that the set of prices for which the agent has a positive incentive to acquire information is the band between the two p_i–curves:

$$P \equiv \{(\gamma, p) \mid p_{II}(\gamma) \leq p \leq p_I(\gamma)\}. \tag{38}$$

We can now provide a characterization analogous to Lemma 6 which includes the moral hazard constraints. The additional requirement is that the endpoints of $\hat{c}_0$ are in P.

Lemma 10 *There are transfers* $(\bar{t}, t)$ *such that a deterministic contract* $(\alpha, \bar{t}, \bar{q}, t, q)$ *satisfies the adverse selection constraints (AS2), (AS1) and the moral hazard constraints (MH) if and only if* $(\alpha, \bar{q}, q)$ *is non–increasing w.r.t.* $(\bar{q}, q)$ *and*

$$(\gamma_1, \hat{c}_0(\gamma_1)), (\gamma_2, \hat{c}_0(\gamma_2)) \in P. \tag{39}$$

As explained earlier, the moral hazard constraints are satisfied if the *whole* schedule $\hat{c}_0$ is in P. The reason why we only have to consider the endpoints is that the two p_i-curves are increasing, whereas $\hat{c}_0$ is necessarily non–increasing.[22]

Restricting attention to deterministic contracts, we can now treat problem $\mathcal{P}$ as usual. By Lemma 1 and 2, the agent's utility U is given by (14) and, therefore, decreasing in γ. Optimality then requires that the individual rationality constraint is binding for the highest type $\bar{\gamma}$. Together with (14), this pins down the agent's utility, which we can then insert in the principal's objective. Integration by parts transforms the principal's objective into the expected virtual surplus as stated in (15). Finally, by Lemma 10 and exploiting the structure of an allocation $(\alpha, \bar{q}, q)$ which is non–increasing w.r.t. $(\bar{q}, q)$ we can restate the problem $\mathcal{P}$ as follows.

$$\mathcal{S}: \max_{\gamma_1, \gamma_2, \hat{c}_0} \int_0^{\gamma_1} v - \gamma - h(\gamma) dF(\gamma) + \int_{\gamma_1}^{\gamma_2} [\int_{-\infty}^{\hat{c}_0(\gamma)} v - c + \psi(c, \gamma) h(\gamma)\; dG(c \mid \gamma) - k] dF(\gamma) \tag{40}$$

$$s.t. \qquad \text{(39) and } \hat{c}_0 : [\gamma_1, \gamma_2] \to \mathbb{R} \text{ is non–increasing.}$$

We can find a solution to $\mathcal{S}$ by the following two–step procedure: First, determine the optimal price schedule $\hat{c}_0$ for given thresholds $\gamma_1 \leq \gamma_2$ and, subsequently, determine the optimal thresholds γ_1 and γ_2 by optimizing over all optimal price schedules $\hat{c}_0$. We call $\hat{c}_0$ optimal with respect to a pair (γ_1, γ_2) if $\hat{c}_0$ is a solution of $\mathcal{S}$ for given thresholds γ_1 and γ_2. The next lemma characterizes the optimal price schedule $\hat{c}_0$ given a pair (γ_1, γ_2).

Lemma 11 *The price schedule* $\tilde{c}_0$ *is optimal with respect to a pair* (γ_1, γ_2) *where*

$$\tilde{c}_0(\gamma) \equiv \min\{p_I(\gamma_1), \max\{c_0(\gamma), p_{II}(\gamma_2)\}\}. \tag{41}$$

[22] Related to the discussion in footnote 17, it is an open question whether an analogous characterization holds in a model with divisible quantity where deterministic allocations are not either zero or one.

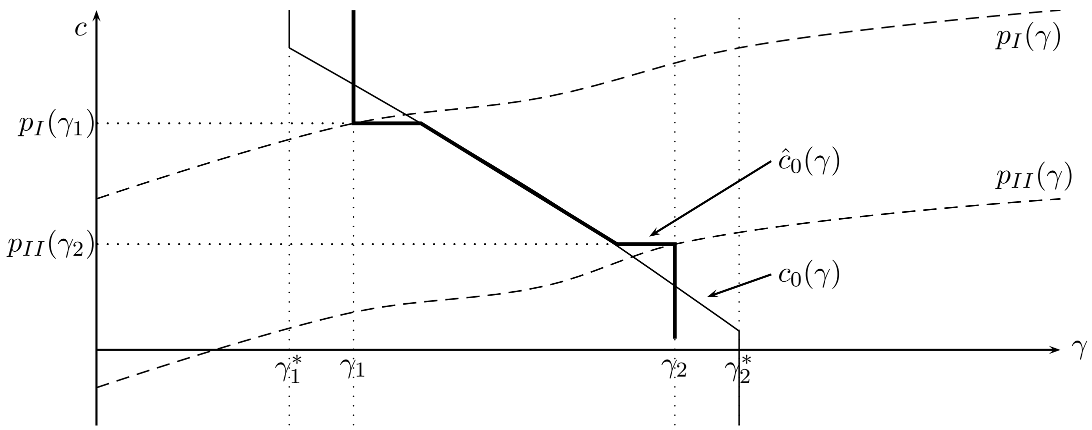

Figure 3: Bunching due to moral hazard

Intuitively, the optimal schedule $\tilde{c}_0$ coincides with the benchmark schedule c_0 whenever the price $c_0(\gamma)$ automatically induces type γ to acquire information, i.e., whenever $c_0(\gamma)$ lies in between $p_I(\gamma_1)$ and $p_{II}(\gamma_2)$. Otherwise, the price has to be adapted so as to increase the agent's value of information. Adapting the price implies a suboptimal decision by the agent in period 2 and entails losses in virtual surplus. Hence, the optimal new price is the one closest to the original price $c_0(\gamma)$ among all prices that are incentive compatible, i.e. non–increasing and in P. As illustrated in Figure 3, the optimal schedule $\tilde{c}_0$ can, therefore, be found by a bunching or ironing procedure, which flattens parts of the benchmark schedule c_0, leading to price bunching for different γ types. The remaining problem is, then, to determine over which ranges bunching takes place. The solution to this problem yields the optimal thresholds γ_1 and γ_2. While the exact optimal thresholds γ_1 and γ_2 depend on the details of the primitives, the following proposition shows how the moral hazard problem affects information acquisition:

Proposition 5 *When the moral hazard problem causes agency costs, the range of information acquisition shrinks relative to the case with observable information acquisition:*

$$\gamma_1 \geq \gamma_1^* \quad \textit{and} \quad \gamma_2 \leq \gamma_2^*. \tag{42}$$

The intuition for Proposition 5 is that outside the interval $[\gamma_1^*, \gamma_2^*]$, information acquisition costs k exceed, by definition, the principal's value of information. Hence, even if a price could be found that induces information acquisition by agent types outside of $[\gamma_1^*, \gamma_2^*]$, this would be suboptimal because information acquisition would cause a loss to the principal.

Proposition 5 allows us to identify the impact of the moral hazard problem on the distortions in information acquisition beyond those caused by adverse selection. When the moral hazard problem causes agency costs, then, relative to the case with observable information acquisition, both more type I and more type II errors are committed. Hence, information acquisition to prevent type I errors become less distorted and more efficient, while information acquisition to prevent type II errors becomes more distorted and less efficient. Consequently, the welfare effects of the additional moral hazard constraints are ambiguous and it could be that the first effect outweighs the second. In this case, the presence of the moral hazard problem would actually increase overall efficiency. Because the moral hazard problem unambiguously raises the agency costs of the principal, an increase in efficiency means that the agent gains more than the principal loses.

7 Conclusions

We study how a principal optimally deals with the incentive problems in project management when, due to large cost uncertainties, procurement projects require pre–project planning. These incentive problems typically involve both adverse selection and moral hazard. To model the procurement problem, we extend the standard sequential screening model to endogenous information acquisition with moral hazard. Our analysis offers a number of theoretical insights. In particular, we identify systematic distortions on the amount of information acquisition under the optimal contract and disentangle the distinct impact of adverse selection and moral hazard on these distortions. Moreover, we show how a bunching procedure can be employed to account for moral hazard. In doing so we present a new characterization of incentive compatibility for deterministic contracts in dynamic screening models.

From a more applied angle, our analysis extends beyond procurement to other domains. Classical examples of contracts with option clauses that adapt contracting terms to new information are labor contracts, securities such as put or call options, or ticket pricing. The existence of such contracts can be rationalized by straightforward reinterpretations of our model (e.g. viewing the agent as the buyer with uncertain demand and the principal as the seller). Option contracts are prevalent also in business-to-business procurement. Refund or buy–back contracts play a key role in vertical manufacturer retailer relations in which there is uncertainty (about demand or costs) during a sales period. The management literature reports the widespread use of option contracts in publishing, CD retailing (Kandel, 1996), for fashion goods such as apparel (Eppen and Iyer, 1997), in the semiconductor and consumer electronic industries (Milner and Rosenblatt, 2002), or in catalogue retailing (Donohue, 2000). In all these examples, private information and active, unobservable information acquisition are often intertwined.

Earlier work on dynamic mechanism design already demonstrates that option contracts represent optimal contractual structures in dynamic screening problems. Our results strengthen this view by demonstrating that option contracts do not only serve the purpose of screening agents after they obtain their additional information, but also provide the correct incentives for agents to obtain this additional information in the first place.

Appendix

Proof of Lemma 2 Let

$$\tilde{U}(\hat{\gamma};\gamma) \equiv \alpha(\hat{\gamma}) \int_{-\infty}^{\infty} u^I(\hat{\gamma}, c)\, dG(c \mid \gamma) + (1 - \alpha(\hat{\gamma}))[\bar{t}(\hat{\gamma}) - \gamma \bar{q}(\hat{\gamma})] \tag{43}$$

be the agent type γ's utility when he reports $\hat{\gamma}$ in period 1. Note that $\tilde{U}(\hat{\gamma};\gamma)$ is decreasing in γ. This follows from two observations: first $-\gamma\bar{q}(\hat{\gamma})$ is decreasing in γ. Second, the integral is decreasing in γ because of the first order stochastic dominance ranking of $\{G(c \mid \gamma)\}_\gamma$ and since $u^I(\hat{\gamma}, c)$ is non–increasing in c by (13). Moreover, we have

$$U(\gamma) = \max_{\hat{\gamma} \in \Gamma} \tilde{U}(\hat{\gamma}, \gamma). \tag{44}$$

As the maximum over decreasing functions, U is decreasing and thus differentiable almost everywhere. We now compute the derivative for all γ at which it exists. Let $\delta > 0$. Incentive compatibility (AS1) implies

$$1/\delta \cdot [U(\gamma + \delta) - U(\gamma)] \;\; \geq \;\; 1/\delta \cdot [\tilde{U}(\gamma, \gamma + \delta) - \tilde{U}(\gamma, \gamma)], \tag{45}$$

$$1/\delta \cdot [U(\gamma) - U(\gamma - \delta)] \;\; \leq \;\; 1/\delta \cdot [\tilde{U}(\gamma, \gamma) - \tilde{U}(\gamma, \gamma - \delta)]. \tag{46}$$

As $\delta \to 0$, the right hand sides of the two previous inequalities converge to

$$\alpha(\gamma) \int_{-\infty}^{\infty} u^I(\gamma, c) \frac{\partial g(c \mid \gamma)}{\partial \gamma}\, dc - (1 - \alpha(\gamma))\bar{q}(\gamma). \tag{47}$$

Therefore $U'(\gamma)$ equals (47). We conclude the proof by re–writing (47). By integration by parts, and since $\partial u^I/\partial c = -q$ by Lemma 1, we can write the integral as

$$u^I(\gamma, c) \frac{\partial G(c \mid \gamma)}{\partial \gamma}\bigg|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} q(\gamma, c) \frac{\partial G(c \mid \gamma)}{\partial \gamma}\, dc. \tag{48}$$

Now note that $\lim_{c \to -\infty} G(c \mid \gamma) = 0$ for all γ so that $\lim_{c \to -\infty} \partial G(c \mid \gamma)/\partial\gamma = 0$. Similarly, $\lim_{c \to +\infty} \partial G(c \mid \gamma)/\partial\gamma = 0$. Thus, the first term in the previous expression vanishes, and we arrive at

$$U'(\gamma) \;\; = \;\; \alpha(\gamma) \int_{-\infty}^{\infty} q(\gamma, c) \frac{\partial G(c \mid \gamma)}{\partial \gamma}\, dc - (1 - \alpha(\gamma))\bar{q}(\gamma), \tag{49}$$

which is the expression stated in Lemma 2. Q.E.D.

Derivation of the objective in (15) Integration by parts delivers

$$\int_0^{\bar\gamma} U(\gamma)\, dF(\gamma) = U(\gamma)F(\gamma)|_0^{\bar\gamma} - \int_0^{\bar\gamma} U'(\gamma)\frac{F(\gamma)}{f(\gamma)}\, dF(\gamma) \tag{50}$$

$$= -\int_0^{\bar\gamma} U'(\gamma)h(\gamma)\, dF(\gamma), \tag{51}$$

where we have used that $U(\gamma)F(\gamma)|_0^{\bar\gamma} = U(\bar\gamma) = 0$. Inserting this in the objective (10) yields that

$$W = \int_0^{\bar\gamma} \alpha(\gamma)[\int_{-\infty}^{\infty} [v-c]q(\gamma,c)\, dG(c \mid \gamma) - k] + (1-\alpha(\gamma))[v-\gamma]\bar q(\gamma) + U'(\gamma)h(\gamma)\, dF(\gamma). \tag{52}$$

If we now use the expression for U' stated in Lemma 2, we obtain the objective in (15). Q.E.D.

Proof of Lemma 4 For later reference, we prove the result for the more general case, where i) $\partial\psi/\partial c \cdot h < 1$, ii) $\partial\psi/\partial\gamma \le 0$, and iii) $\psi(\gamma,c) < 0$. From i) it follows that $c - \psi(\gamma,c)h(\gamma)$ is strictly increasing in c so that there is a unique solution $c_0(\gamma)$ to $v - c + \psi(\gamma,c)h(\gamma) = 0$, and $v - c + \psi(\gamma,c)h(\gamma) > 0$ if and only if $c < c_0(\gamma)$. From iii), we have that $c_0(\gamma) \le v$. Similarly, since $h(\gamma)$ is non–decreasing by assumption, there is a unique solution γ_0 to $v - \gamma - h(\gamma) = 0$, and $v - \gamma - h(\gamma) > 0$ if and only if $\gamma < \gamma_0$. Clearly, $\gamma_0 \le v$. For the independent case, where $\psi(\gamma,c) = -1$, condition (17) follows. In light of the objective (15), this implies that q^* and $\bar q^*$ are optimal. Q.E.D.

Proof of Lemma 6 "$\Rightarrow$": Since the contract is deterministic, the second period adverse selection constraints (AS2) imply that $q(\gamma,c)$ is given by a cut–off rule $\hat c_0(\gamma)$, where the project is implemented if and only if costs c are smaller than the cut–off. Moreover, this implies that the transfers when information is acquired are piece–wise constant: $t(\gamma,c) = t_0(\gamma)$ if $c > \hat c_0(\gamma)$ and $t(\gamma,c) = t_1(\gamma)$ if $c \le \hat c_0(\gamma)$, where the cut–off type himself has to be indifferent between implementing and quitting the project: $t_1(\gamma) = \hat c_0(\gamma) + t_0(\gamma)$.

We now show that the allocation is non–increasing w.r.t. $(\bar q, q)$. The argument is by contradiction. Suppose first that $\hat c_0$ is not non–increasing. Then there are γ, $\delta > 0$ with $\gamma, \gamma+\delta \in [\gamma_1,\gamma_2]$ and $\hat c_0(\gamma) < \hat c_0(\gamma+\delta)$. Given the transfers $t_1 = t_0 + \hat c_0$ and t_0, and the deterministic allocation rule the utility of agent type γ is

$$U(\gamma) = \int_{-\infty}^{\hat c_0(\gamma)} (t_1(\gamma) - c)dG(c \mid \gamma) + \int_{\hat c_0(\gamma)}^{\infty} t_0(\gamma)dG(c \mid \gamma) - k \tag{53}$$

$$= t_0(\gamma) + \int_{-\infty}^{\hat c_0(\gamma)} (\hat c_0(\gamma) - c)dG(c \mid \gamma) - k. \tag{54}$$

The first period adverse selection constraints (AS1) imply:

$$U(\gamma) \;\geq\; t_0(\gamma+\delta) + \int_{-\infty}^{\hat{c}_0(\gamma+\delta)} (\hat{c}_0(\gamma+\delta) - c) dG(c \mid \gamma) - k, \tag{55}$$

$$U(\gamma+\delta) \;\geq\; t_0(\gamma) + \int_{-\infty}^{\hat{c}_0(\gamma)} (\hat{c}_0(\gamma) - c) dG(c \mid \gamma+\delta) - k. \tag{56}$$

Adding the two inequalities and re–arranging terms yields:

$$\int_{-\infty}^{\infty} (\hat{c}_0(\gamma) - c) 1_{(-\infty,\hat{c}_0(\gamma))}(c) - (\hat{c}_0(\gamma+\delta) - c) 1_{(-\infty,\hat{c}_0(\gamma+\delta))}(c)\; dG(c \mid \gamma) \tag{57}$$

$$\geq \int_{-\infty}^{\infty} (\hat{c}_0(\gamma) - c) 1_{(-\infty,\hat{c}_0(\gamma))}(c) - (\hat{c}_0(\gamma+\delta) - c) 1_{(-\infty,\hat{c}_0(\gamma+\delta))}(c)\; dG(c \mid \gamma+\delta), \tag{58}$$

where $1_{(a,b)}(c)$ is an indicator function that is 1 if $c \in (a,b)$ and 0 otherwise. Notice that the integrand under the two integrals is the same. Since $\hat{c}_0(\gamma) < \hat{c}_0(\gamma+\delta)$ by assumption, the integrand is increasing and strictly so in the range of $(\hat{c}_0(\gamma), \hat{c}_0(\gamma+\delta))$. Hence, since $G(c \mid \gamma+\delta)$ first order stochastically dominates $G(c \mid \gamma)$, the second integral is strictly smaller than the first one, a contradiction. Thus, we have shown that $\hat{c}_0$ is non–increasing, as desired.

Next, we show that if there are γ, γ' with $\gamma < \gamma'$ and $\alpha(\gamma) = 1$, $\alpha(\gamma') = 0$, then we must have that $\bar{q}(\gamma') = 0$. Indeed, suppose this is not true. Then for $\delta = \gamma' - \gamma > 0$ we have $\alpha(\gamma) = 1$ and $\alpha(\gamma+\delta) = 0$, but, since $\bar{q}$ is deterministic, $\bar{q}(\gamma+\delta) = 1$. (AS1) implies that

$$t_0(\gamma) + \int_{-\infty}^{\hat{c}_0(\gamma)} (\hat{c}_0(\gamma) - c) dG(c \mid \gamma) - k \;\geq\; \bar{t}(\gamma+\delta) - \gamma, \tag{59}$$

$$\bar{t}(\gamma+\delta) - (\gamma+\delta) \;\geq\; t_0(\gamma) + \int_{-\infty}^{\hat{c}_0(\gamma)} (\hat{c}_0(\gamma) - c) dG(c \mid \gamma+\delta). \tag{60}$$

Adding the two inequalities and re–arranging terms yields:

$$\int_{-\infty}^{\infty} (\hat{c}_0(\gamma) - c) 1_{(-\infty,\hat{c}_0(\gamma))}(c) + c\; dG(c \mid \gamma) \geq \int_{-\infty}^{\infty} (\hat{c}_0(\gamma) - c) 1_{(-\infty,\hat{c}_0(\gamma))}(c) + c\; dG(c \mid \gamma+\delta). \tag{61}$$

Here, we have used $\gamma = \int_{-\infty}^{\infty} c dG(c \mid \gamma)$. Now notice that the integrand under the two integrals is the same, increasing, and strictly increasing for $c > \hat{c}_0(\gamma)$. Hence, since $G(c \mid \gamma+\delta)$ first order stochastically dominates $G(c \mid \gamma)$, the first integral is strictly smaller than the second one, a contradiction.

In a similar way, it can be shown that if there are γ, γ' with $\gamma < \gamma'$ and $\alpha(\gamma) = 0$, $\alpha(\gamma') = 1$, then we must have that $\bar{q}(\gamma) = 1$.

The previous two statements jointly imply that α can equal 1 only over an interval $[\gamma_1, \gamma_2]$, and $\bar{q}$ is 1 (resp. 0) to the left (resp. right) of the interval. But this means that the allocation is non–increasing w.r.t. $(\bar{q}, q)$, and this completes the proof of the "$\Rightarrow$"–part.

"⇐": We proceed in two steps. We first define transfers and then verify all incentive constraints.

STEP 1: (i) For $\gamma > \gamma_2$, we set for some constant $\tau \in \mathbb{R}$:

$$\bar{t}(\gamma) = \tau. \tag{62}$$

(ii) For $\gamma \in [\gamma_1, \gamma_2]$, we define $t(\gamma, c)$ piece–wise constant:

$$t(\gamma, c) = \begin{cases} t_0(\gamma) & if \quad c \geq \hat{c}_0(\gamma) \\ t_1(\gamma) & if \quad c < \hat{c}_0(\gamma) \end{cases}, \tag{63}$$

where t_0 and t_1 are defined by:

$$t_0(\gamma_2) = k - \int_{-\infty}^{\hat{c}_0(\gamma_2)} (\hat{c}_0(\gamma_2) - c) dG(c \mid \gamma_2)\, dc + \tau, \tag{64}$$

$$t_0(\gamma) = t_0(\gamma_2) + \int_{\gamma}^{\gamma_2} \hat{c}_0'(z) G(\hat{c}_0(z) \mid z)\, dz, \tag{65}$$

$$t_1(\gamma) = t_0(\gamma) + \hat{c}_0(\gamma). \tag{66}$$

(Note that $\hat{c}_0'$ is well–defined for almost all γ since $\hat{c}_0$ is decreasing by assumption.)

(iii) For $\gamma < \gamma_1$, we define $\bar{t}(\gamma)$ equal to the constant $\bar{t}$:

$$\bar{t}(\gamma) = \bar{t} = t_0(\gamma_1) - k + \gamma_1 + \int_{-\infty}^{\hat{c}_0(\gamma_1)} (\hat{c}_0(\gamma_1) - c)\, dG(c \mid \gamma_1). \tag{67}$$

STEP 2: We now verify that $(\alpha, \bar{t}, \bar{q}, t, q)$ satisfies (AS1) and (MON2). By definition of a deterministic allocation which is non–increasing w.r.t. $(\bar{q}, q)$, $q(\gamma, c)$ satisfies (MON2). To demonstrate (AS1), let $\tilde{U}(\hat{\gamma}; \gamma)$ be defined as in (43). We have to show that

$$\Delta \equiv U(\gamma) - \tilde{U}(\hat{\gamma}; \gamma) \geq 0 \qquad \forall \gamma, \hat{\gamma}. \tag{68}$$

Indeed, observe first that, by construction, transfers are continuous on Γ and differentiable with respect to γ for all γ except at γ_1 and γ_2. Together with the fact that $(\alpha, \bar{q}, q)$ is deterministic, this implies that U and $\tilde{U}$ are continuous on Γ and differentiable almost everywhere. Hence, with $U(\hat{\gamma}) = \tilde{U}(\hat{\gamma}; \hat{\gamma})$, we can write

$$\Delta = U(\gamma) - U(\hat{\gamma}) + \tilde{U}(\hat{\gamma}; \hat{\gamma}) - \tilde{U}(\hat{\gamma}; \gamma) = \int_{\hat{\gamma}}^{\gamma} U'(z)\, dz - \int_{\hat{\gamma}}^{\gamma} \frac{\partial \tilde{U}(\hat{\gamma}, z)}{\partial \gamma}\, dz. \tag{69}$$

Next, we compute U' and $\partial \tilde{U} / \partial \gamma$. As for U', note that for $\gamma \in (\gamma_2, \bar{\gamma}]$, we have $U(\gamma) = \tau$, thus $U'(\gamma) = 0$. For $\gamma \in [0, \gamma_1)$, we have $U(\gamma) = \bar{t} - \gamma$, thus $U'(\gamma) = -1$. Finally, consider

$\gamma \in (\gamma_1, \gamma_2)$. Note that by the definition of transfers, $U(\gamma) = t_0(\gamma) + \int_{-\infty}^{\hat{c}_0(\gamma)} (\hat{c}_0(\gamma) - c) dG(c \mid \gamma)$. Hence, by Leibniz' rule

$$U'(\gamma) = t_0'(\gamma) + \int_{-\infty}^{\hat{c}_0(\gamma)} \hat{c}_0'(\gamma)\, dG(c \mid \gamma) + \int_{-\infty}^{\hat{c}_0(\gamma)} (\hat{c}_0(\gamma) - c) \frac{\partial g(c \mid \gamma)}{\partial \gamma}\, dc. \tag{70}$$

If we differentiate t_0, we get $t_0'(\gamma) = -\hat{c}_0'(\gamma) G(\hat{c}_0(\gamma) \mid \gamma)$, which cancels with the second term in (70). Now apply integration by parts on the third term in (70) to get

$$U'(\gamma) = (\hat{c}_0(\gamma) - c)\, \frac{\partial G(c \mid \gamma)}{\partial \gamma}\Big|_{-\infty}^{\hat{c}_0(\gamma)} + \int_{-\infty}^{\hat{c}_0(\gamma)} \frac{\partial G(c \mid \gamma)}{\partial \gamma}\, dc \tag{71}$$

$$= \int_{-\infty}^{\hat{c}_0(\gamma)} \frac{\partial G(c \mid \gamma)}{\partial \gamma}\, dc, \tag{72}$$

where the second equality follows because $\lim_{c \to -\infty} \partial G / \partial \gamma (c \mid \gamma) = 0$.

In sum, U' exists for all $\gamma \in \Gamma$ except at γ_1 and γ_2. Because $(\alpha, \bar{q}, q)$ is deterministic, we can write U' as a function over the whole interval $[0, \bar{\gamma}]$ as

$$U'(\gamma) = \alpha(\gamma) \int_{-\infty}^{\infty} q(\gamma, c) \frac{\partial G(c \mid \gamma)}{\partial \gamma}\, dc - (1 - \alpha(\gamma)) \bar{q}(\gamma). \tag{73}$$

Moreover, similar steps that we used to derive (49) yield that for all $\gamma, \hat{\gamma}$:

$$\frac{\partial \tilde{U}(\hat{\gamma}; \gamma)}{\partial \gamma} = \alpha(\hat{\gamma}) \int_{-\infty}^{\infty} q(\hat{\gamma}, c) \frac{\partial G(c \mid \gamma)}{\partial \gamma}\, dc - (1 - \alpha(\hat{\gamma})) \bar{q}(\hat{\gamma}). \tag{74}$$

Therefore, we obtain

$$\Delta = \int_{\hat{\gamma}}^{\gamma} \int_{-\infty}^{+\infty} [\alpha(z) q(z, c) - \alpha(\hat{\gamma}) q(\hat{\gamma}, c)] \frac{\partial G(c \mid z)}{\partial \gamma}\, dc\, dz \tag{75}$$

$$- \int_{\hat{\gamma}}^{\gamma} [(1 - \alpha(z)) \bar{q}(z) - (1 - \alpha(\hat{\gamma})) \bar{q}(\hat{\gamma})]\, dz. \tag{76}$$

Now observe that in the independent case $\frac{\partial G(c|z)/\partial \gamma}{g(c|z)} = -1$. In the non–independent case, we will assume $\int \frac{\partial G(c|z)/\partial \gamma}{g(c|z)} dG(c \mid z) = -1$ (see (34)). Therefore, the second line is $\int_{\hat{\gamma}}^{\gamma} \int_{-\infty}^{+\infty} [(1 - \alpha(z)) \bar{q}(z) - (1 - \alpha(\hat{\gamma})) \bar{q}(\hat{\gamma})] \frac{\partial G(c|z)/\partial \gamma}{g(c|z)} dG(c \mid z)\, dz$, and re–arranging terms yields

$$\Delta = \int_{\hat{\gamma}}^{\gamma} \int_{-\infty}^{+\infty} \{[(1 - \alpha(z)) \bar{q}(z) + \alpha(z) q(z, c)] \tag{77}$$

$$-[(1 - \alpha(\hat{\gamma})) \bar{q}(\hat{\gamma}) + \alpha(\hat{\gamma}) q(\hat{\gamma}, c)]\} \frac{\partial G(c \mid z)/\partial \gamma}{g(c \mid z)}\, dG(c \mid z)\, dz. \tag{78}$$

We now show that Δ is non–negative if $\hat{\gamma} < \gamma$. Indeed, by assumption, $(\alpha, \bar{q}, q)$ is deterministic and non–increasing w.r.t. $(\bar{q}, q)$. This implies that $(1 - \alpha(z)) \bar{q}(z) + \alpha(z) q(z, c)$ is non–increasing in z for all c. Therefore, with $\hat{\gamma} < \gamma$, the term in the curly brackets under the integral is non–positive. Recall also that first order stochastic dominance ranking of $\{G(\cdot \mid \gamma)\}_\gamma$ implies $\partial G/\partial\gamma \leq 0$. These two observations imply that the integrand is non–negative so that $\Delta \geq 0$. For $\hat{\gamma} > \gamma$, the argument is analogous. Q.E.D.

Derivation of the relationship in footnote 19 Consider first the case $\gamma > \gamma_0$. Then in the independent case, $J^*(\gamma) = \int_{-\infty}^{c_0(\gamma)}[v - c - h(\gamma)]\hat{g}(c-\gamma)dc$. Consider the change of variable $a(c) = c + h(\gamma)$. Recall that $c_0(\gamma) = v - h(\gamma)$ so that $a(c_0(\gamma)) = v$. Moreover, $a'(c) = 1$. Thus,

$$J^*(\gamma) = \int_{-\infty}^{v}[v-a]\hat{g}(a - [\gamma + h(\gamma)])da = \int_{-\infty}^{v}[v-a]dG(a \mid \gamma + h(\gamma))da = J^{FB}(\gamma + h(\gamma)). \tag{79}$$

The argument for the case $\gamma \leq \gamma_0$ is identical. Q.E.D.

Proof of Lemma 8 We begin with the proof of an auxiliary statement: The moral hazard problem does not cause additional agency costs if and only if jointly:

$$\int_{c_0(\gamma_1^*)}^{\infty}[\psi(\gamma_1^*, c_0(\gamma_1^*)) - \psi(\gamma_1^*, c)]h(\gamma_1^*)\; dG(c \mid \gamma_1^*) \;\;\geq\;\; 0, \tag{80}$$

$$\int_{-\infty}^{c_0(\gamma_2^*)}[\psi(\gamma_2^*, c) - \psi(\gamma_2^*, c_0(\gamma_2^*))]h(\gamma_2^*)\; dG(c \mid \gamma_2^*) \;\;\leq\;\; 0. \tag{81}$$

To see this, recall that $c_0(\gamma) = v + \psi(\gamma, c_0(\gamma))h(\gamma)$. Therefore, from (64) with $\tau = 0$:

$$t_0^*(\gamma_2^*) \;\;=\;\; k - \int_{-\infty}^{c_0(\gamma_2^*)}(v - c + \psi(\gamma_2^*, c_0(\gamma_2^*))h(\gamma_2^*))\; dG(c \mid \gamma_2^*). \tag{82}$$

By definition of γ_2^*, we have that $k = \int_{-\infty}^{c_0(\gamma_2^*)}(v - c + \psi(\gamma_2^*, c)h(\gamma))\; dG(c \mid \gamma_2^*)$. Hence,

$$t_0^*(\gamma_2^*) \;\;=\;\; \int_{-\infty}^{c_0(\gamma_2^*)}[\psi(\gamma_2^*, c) - \psi(\gamma_2^*, c_0(\gamma_2^*))]h(\gamma_2^*)\; dG(c \mid \gamma_2^*). \tag{83}$$

A similar computation based on (67) and the definition of γ_1^* reveals that

$$\bar{t}^* - (c_0(\gamma_1^*) + t_0^*(\gamma_1^*)) = \bar{t}^* - t_1^*(\gamma_1^*) = \int_{c_0(\gamma_1^*)}^{\infty}[\psi(\gamma_1^*, c_0(\gamma_1^*)) - \psi(\gamma_1^*, c)]h(\gamma_1^*)\; dG(c \mid \gamma_1^*). \tag{84}$$

¿From (83) and (84) it then follows that (35) implies (80) and (81). To see the reverse, recall that (65) implies $t_0^{*\prime}(\gamma) = -c_0'(\gamma)G(c_0(\gamma) \mid \gamma) \geq 0$ (since $c_0' \leq 0$ as shown in footnote 20). Moreover, from (66) it follows $t_1^{*\prime}(\gamma) = t_0^{*\prime}(\gamma) + c_0'(\gamma) = (1 - G(c_0(\gamma) \mid \gamma))c_0'(\gamma) \leq 0$. In other words, t_0^* is non–decreasing and t_1^* non–increasing in γ. Hence, (35) is actually equivalent to $\bar{t}^* \geq t_0^*(\gamma_1^*) + c_0(\gamma_1^*)$ and $t_0^*(\gamma_2^*) \leq 0$, which is implied by (80) and (81). This completes the proof of the auxiliary statement.

By inspection of the integrands in (80) and (81), it can now easily be seen that the moral hazard problem does not (does) cause agency costs if $\partial\psi/\partial c \geq (<)0$. Q.E.D.

Proof of Lemma 9 (i) Fix γ. Since $\partial J_I^A/\partial p = -(1 - G(p \mid \gamma)) < 0$, J_I^A is strictly decreasing in p. Moreover, for $p \to -\infty$, J_I^A becomes unboundedly large, and for $p \to +\infty$, J_I^A converges to zero. Therefore, $J_I^A = k$ has a unique solution $p_I(\gamma)$. The argument for $i = II$ is similar.

(ii) First order stochastic dominance ranking of $\{G(\cdot \mid \gamma)\}_\gamma$ implies that J_I^A is increasing in γ. Together with the fact that J_I^A is decreasing in p, this implies that p_I is increasing in γ. The argument for p_{II} is analogous.

(ii) The claim follows since J_I^A is decreasing and J_{II}^A is increasing in p. Q.E.D.

Proof of Lemma 10 We first show that there are transfers $(\bar{t}, t)$ so that $(\alpha, \bar{t}, \bar{q}, t, q)$ satisfies (AS2), (AS1) if and only if $(\alpha, \bar{q}, q)$ is non–increasing w.r.t. $(\bar{q}, q)$. The "only if"-part follows from Lemma 6 and the fact that (AS2) implies (MON2) (by Lemma 1). As for the "if"-part, suppose that $(\alpha, \bar{q}, q)$ is non–increasing w.r.t. $(\bar{q}, q)$. By Lemma 6, this implies (AS1) and (MON2) for the transfers defined by (62)-(67). It is easy to verify that the resulting u^I satisfies (13). Together with (MON2) this implies (AS2) by Lemma 1.

Next, we turn to the moral hazard constraints. Given the transfers (62)-(67), the same arguments that establish Lemma 7 demonstrate that the contract satisfies the moral hazard constraints (MH) if and only if

$$\bar{t} \geq \hat{c}_0(\gamma) + t_0(\gamma) \quad \text{and} \quad t_0(\gamma) \leq \tau \quad \text{for all } \gamma \in [\gamma_1, \gamma_2]. \tag{85}$$

Because $\hat{c}_0$ is non–increasing, it follows as in the proof of Lemma 8 that t_0 is non–decreasing and t_1 non–increasing in γ. Therefore, the previous condition is actually equivalent to $\bar{t} \geq \hat{c}_0(\gamma_1) + t_0(\gamma_1)$ and $t_0(\gamma_2) \leq \tau$. Yet by definition of transfers and J_I^A, J_{II}^A, this is equivalent to (39). Q.E.D.

Proof of Lemma 11 Suppose, by contradiction, that $\tilde{c}_0$ is not optimal, then there exists some other schedule $\hat{c}_0$ that satisfies the constraints in $\mathcal{S}$ and yields a strictly higher value for its objective (40). In particular, since $\hat{c}_0$ satisfies (39), we have $\hat{c}_0(\gamma) \in [p_{II}(\gamma), p_I(\gamma)]$ for $\gamma \in \{\gamma_1, \gamma_2\}$. Because $\hat{c}_0$ is non–increasing, it also follows $p_I(\gamma_1) \geq \hat{c}_0(\gamma_1) \geq \hat{c}_0(\gamma) \geq \hat{c}_0(\gamma_2) \geq p_{II}(\gamma_2)$ for all $\gamma \in [\gamma_1, \gamma_2]$. Hence, $p_I(\gamma_1) \geq p_{II}(\gamma_2)$ and $\hat{c}_0(\gamma) \in [p_{II}(\gamma_2), p_I(\gamma_1)]$ for all $\gamma \in [\gamma_1, \gamma_2]$.

Define the sets

$$\Gamma^+ \equiv \{\gamma \in [\gamma_1, \gamma_2] \mid \hat{c}_0(\gamma) > \tilde{c}_0(\gamma)\} \text{ and } \Gamma^- \equiv \{\gamma \in [\gamma_1, \gamma_2] \mid \hat{c}_0(\gamma) < \tilde{c}_0(\gamma)\}.$$

If $\gamma \in \Gamma^+$, then, due to $\hat{c}_0(\gamma) \in [p_{II}(\gamma_2), p_I(\gamma_1)]$, we must have $\tilde{c}_0(\gamma) < p_I(\gamma_1)$, which is only the case if $c_0(\gamma) < p_I(\gamma_1)$. Hence, $\gamma \in \Gamma^+$ implies $\hat{c}_0(\gamma) > \tilde{c}_0(\gamma) \geq c_0(\gamma)$. By the definition of c_0 this then implies $v - c + \psi(\gamma, c)h(\gamma) < 0$ for all $c \in (\tilde{c}_0(\gamma), \hat{c}_0(\gamma))$. Similar arguments show that $\gamma \in \Gamma^-$ implies $v - c + \psi(\gamma, c)h(\gamma) > 0$ for all $c \in (\hat{c}_0(\gamma), \tilde{c}_0(\gamma))$.

The difference in the objective (40) between the schedule $\hat{c}_0(\gamma)$ and $\tilde{c}_0(\gamma)$ is

$$\int_{\gamma\in\Gamma^+}\int_{\tilde{c}_0(\gamma)}^{\hat{c}_0(\gamma)}(v - c + \psi(\gamma, c)h(\gamma))dG(c|\gamma)d\gamma + \int_{\gamma\in\Gamma^-}\int_{\hat{c}_0(\gamma)}^{\tilde{c}_0(\gamma)}(c - v - \psi(\gamma, c)h(\gamma))dG(c|\gamma)d\gamma \quad (86)$$

The terms under the integrands are all negative so that we obtain the contradiction that (40) is weakly smaller for $\hat{c}_0$ than for $\tilde{c}_0$. Q.E.D.

Proof of Proposition 5 We show that any solution to problem $\mathcal{S}$ exhibits $\gamma_1 \geq \gamma_1^*$ and $\gamma_2 \leq \gamma_2^*$. By contradiction, suppose the optimal contract with cut–off $\hat{c}_0$ has $\gamma_1 < \gamma_1^*$. Consider the alternative contract which differs from the optimal contract only in that $\bar{q} = 1$ on $[\gamma_1, \gamma_1^*]$, that is, its cut–off $\hat{c}_0^\dagger$ takes on the same values as $\hat{c}_0$ but its domain is $[\gamma_1^*, \gamma_2]$. We derive a contradiction by showing that the alternative contract is feasible and yields the principal more than the optimal contract.

As for feasibility, recall that p_i, $i = I, II$ is increasing in γ. Because $\hat{c}_0$ satisfies (39) and is non–increasing, this implies that the whole graph of $\hat{c}_0$ is in P. Hence, also the whole graph of $\hat{c}_0^\dagger$ is in P, in particular $\hat{c}_0^\dagger$ satisfies (39), which establishes feasibility.

To complete the argument, notice the two contracts yield the same virtual surplus except in the range $[\gamma_1, \gamma_1^*)$. The payoff difference between the alternative and the optimal contract is

$$\int_{\gamma_1}^{\gamma_1^*} v - \gamma - h(\gamma)dF(\gamma) - \int_{\gamma_1}^{\gamma_1^*}[\int_{-\infty}^{\hat{c}_0(\gamma)} v - c + \psi(c, \gamma)h(\gamma)\ dG(c \mid \gamma) - k]dF(\gamma). \quad (87)$$

By the definition of c_0 and γ_1^* we have the following ordering of virtual surpluses for $\gamma \in [\gamma_1, \gamma_1^*)$:

$$v - \gamma - h(\gamma) \;>\; \int_{-\infty}^{c_0(\gamma)}[v - c + \psi(c, \gamma)h(\gamma)dG(c \mid \gamma) - k \quad (88)$$

$$\geq\; \int_{-\infty}^{\hat{c}_0(\gamma)}[v - c + \psi(c, \gamma)h(\gamma)dG(c \mid \gamma) - k. \quad (89)$$

Hence, the alternative contract yields the principal a higher profit than the optimal contract, a contradiction. The argument for $\gamma_2 \leq \gamma_2^*$ is symmetric. Q.E.D.

References

Athey, S., and I. Segal (2007a). "Designing Efficient Mechanisms for Dynamic Bilateral Trading Games." *American Economic Review Papers and Proceedings* 97, 131-136.

Athey, S., and I. Segal (2007b). "An Efficient Dynamic Mechanism." mimeo.

Baron, D. and D. Besanko (1984). "Regulation and information in a continuing relationship." *Information Economics and Policy* 1, 267-302.

Battaglini, M. (2005). "Long-Term Contracting with Markovian Consumers." *American Economic Review* 95 (3), 637-658.

Bergemann, D. and J. Välimäki (2008). "The Dynamic Pivot Mechanism." mimeo.

Courty, P. and H. Li (2000). "Sequential Screening." *Review of Economic Studies* 67, 697-717.

Cremer, J., F. Khalil, and J.-C. Rochet (1998). "Productive Information Gathering." *Games and Economic Behavior* 25, 174-193.

Dahlin, T., L. Bjelm, and C. Svensson (1999). "Use of electrical imaging in site investigations for a railway tunnel through the Hallandsas Horst, Sweden." *Quarterly Journal of Engineering Geology and Hydrogeology* 32, 163-172.

Dai, C., T. R. Lewis, and G. Lopomo (2006). "Delegating Management to Experts." *RAND Journal of Economics* 37, 503-520.

Donohue, K. L. (2000). "Efficient Supply Contracts for Fashion Goods with Forecast Updating and Two Production Modes." *Management Science* 46, 1397-1411.

Eppen, G. D. and A. V. Iyer (1997). "Backup Agreements in Fashion Buying—The Value of Upstream Flexibility." *Management Science* 43, 1469-1484.

Esö, P. and B. Szentes (2007a). "Optimal Information Disclosure in Auctions and The Handicap Auction." *Review of Economic Studies* 74, 705-731.

Esö, P. and B. Szentes (2007b). "The Price of Advice." *Rand Journal of Economics* 38, 863-880.

Gibson, G. and M. Hamilton (1994). *Analysis of pre–project planning effort and success variables for capital facility projects.* University of Texas at Austin. Construction Industry Institute.

Gibson, G., Jr., Kaczmarowski, J., and Lore, H., Jr. (1995). "Preproject–planning process for capital facilities." J. Constr. Eng. Manage., 121(3), 312318.

Gibson, G., Y. Wang, C. Cho, and M. Pappas, (2006). "What Is Preproject Planning, Anyway?" *Journal of Management in Engineering* 22, 35-42.

Hoffmann, F. and R. Inderst (2009). "Price Discrimination and the Provision of Information." mimeo.

Kähkönen, K. (1999). "Multi-character Model of the Construction Project Definition Process." *Automation in Construction* 8 (6), p. 625-632.

Kandel, E. (1996). "The Right to Return." *Journal of Law and Economics* XXXIX, 329-356.

Kessler, A. S. (1998). "The Value of Ignorance." *Rand Journal of Economics* 29(2), 339-354.

Krähmer, D. and R. Strausz (2008). "Ex Post Private Information and Monopolistic Screening." *The B.E. Journal of Theoretical Economics, Topics* 8(1), Article 25.

Laffont, J.-J. and J. Tirole (1990). "Adverse Selection and Renegotiation in Procurement." *Review of Economic Studies* 57(4), 597-625.

Laffont, J.-J. and J. Tirole (1996). "Pollution Permits and Compliance Strategies." *Journal of Public Economics* 62(1-2), 85-125.

Lewis, T. R. and D. E. M. Sappington (1997). "Information Management in Incentive Problems." *Journal of Political Economy* 105 (4), 796-821.

Milner, J. M. and M. J. Rosenblatt (2002). "Flexible Supply Contracts for Short Life–Cycle Goods: The Buyer's Perspective." *Naval Research Logistics* 49, 25-45.

Myerson, R. (1986). "Multistage Games with Communication." *Econometrica* 54, 323-58.

Pavan, A., I. Segal, and J. Toikka (2008). "Dynamic Mechanism Design: Revenue Equivalence, Profit Maximization and Information Disclosure." mimeo.

Szalay, D. (2009). "Contracts with Endogenous Information." *Games and Economic Behavior* 65, 586-625.

Turner, J.R. (1993). *The Handbook of Project-Based Management*, McGraw-Hill Book, London, UK.