Gathering Information before Signing a Contract: The Case of Imperfect Information*

Stefan Terstiege[†]

April 30, 2015

Abstract

I study information gathering for rent-seeking purposes in contracting. In my model, an agent learns his payoff type only after accepting a contract, but can at costs acquire imperfect information while deliberating whether to accept. I show that the principal deters the acquisition if *and only if* the costs are high. The result stands in contrast to a finding by Crémer and Khalil (1992), who demonstrate that the acquisition of perfect information will always be deterred. A key insight is that the case of imperfect information is an instance of a sequential-screening problem.

Keywords: Principal agent; information acquisition; rent seeking; sequential screening

JEL Codes: D82; D83; D86

1   Introduction

This paper offers a new perspective on information gathering for rent-seeking purposes in contracting. Consider the following procurement relationship. A principal seeks to buy parts, which an agent can produce. The agent's production costs are at first unknown to both parties. *After* accepting a contract, the agent privately learns the exact state when he makes the necessary preparations for production. In particular, he learns his costs early enough so that the output level could still be adjusted—if the contract allows this. *While deliberating* whether to accept, he can acquire preliminary information, but at an extra expense.

The acquisition of such precontractual information would be a rent-seeking activity: From a social perspective, the information is redundant, and its acquisition thus wasteful, given that uncertainty resolves in any case before production. To the agent, on the other hand, the information may be valuable, because it would allow him to forecast more precisely whether or not the offered contract would be profitable for himself. A similar situation prevails with the sale of experience goods. Here, the consumers may be able to gather information about their valuation before making the purchase decision, which has little social value if the sellers can take back and resell the good.

How does the possibility for such rent seeking affect contract design? Specifically, to what distortions does it lead? That is the question that I address in this paper. In a seminal paper, Crémer and Khalil (1992) (henceforth CK), demonstrate that if precontractual information would already remove all uncertainty, the principal will design the contract such that the agent accepts without acquiring information. I consider the case of imperfect information. I show that, there, the principal deters the acquisition if *and only if* the agent's investigation costs exceed some cutoff. A key insight is that the case of imperfect information is an instance of a sequential-screening problem.

The result—and the importance of the imperfection of the information—can be explained as follows. Ex post, the agent will earn a rent, given that he will learn his payoff type (i.e., in the above procurement setting, his production costs) privately. Since he will learn the type only after the signing of the contract, the principal can try to extract the *expected* rent with a participation fee. But now, by acquiring information, the agent has the costly option to forecast more precisely whether his *actual* rent

is likely to be larger than the fee, and to sign the contract only then. Unless such rent seeking is prohibitively expensive, the principal must make a trade-off between efficiency and surplus extraction.

The crucial step is to recognize that precontractual information is relevant for the agent's expectation of his rent, and thus his willingness to pay participation fees. If the acquisition entails low costs, so that the contract must anyway be designed almost as if the agent *had* the information, it is therefore better to make the terms of trade contingent on the information—and thus to induce the acquisition. In particular, if the agent does acquire information, the principal can implement a more efficient contract menu with a larger participation fee conditional on the agent receiving good news about his payoff type, and a less efficient menu with a smaller fee in case of bad news. The corresponding contract screens the agent sequentially, so as to elicit not just the payoff type but also the posterior belief thereof upon information acquisition.

Of course, precontractual information would be relevant for the agent's expectation of his rent also if the information was perfect, as in CK's analysis. But in that case, the information would be *identical* to the one that the agent obtains after signing, when he finally learns his payoff type. Hence, a contract that deters the acquisition could nevertheless condition on the information. This explains why my result differs from the one by CK.

The paper thus identifies a possible form of inefficiency in contracting: before the signing of the contract, parties possibly waste resources to acquire information about parameters that, after signing, they learn anyway. Even though this act is inefficient, the social surplus may be larger if the acquisition does take place. Specifically, in my model, the principal's best contracts that induce information acquisition implement different, possibly more efficient terms of trade than the best contracts that deter information acquisition. The exact welfare properties of these contracts highly depend on the details of the model, and are not studied here.

The analysis suggests implications for the use of trial subscriptions as a marketing device, for example by newspapers. A trial subscription allows consumers to learn about their valuation before deciding whether to buy a regular subscription. Effectively, such a test has little social value if subscription plans can be canceled at short notice.

But in case that the consumers can easily obtain precontractual information anyway, then, according to this paper, a supplier may find it optimal to induce information gathering—for which free trial subscriptions are an effective means.

A key insight is that the contracting problem is closely related to the ones considered in the sequential-screening literature (e.g., Courty and Li 2000; Esö and Szentes 2007; Krähmer and Strausz 2015). There, agents gradually receive private information over time before the allocation takes place.[1] In the seminal paper by Courty and Li (2000), in particular, the agent exogenously *has* imperfect information from the outset and learns his exact payoff type after the signing of the contract. That setting is equivalent to the special case of my model in which information acquisition entails zero costs.[2] For this case, the question of whether optimal contracts induce information acquisition can be rephrased as whether they condition on the information—and thus indeed screen the agent sequentially. Courty and Li provide a *complete* characterization of optimal contracts, using regularity assumptions concerning the probability distributions of their model. Under these assumptions, the optimal contracts do screen sequentially. I focus on the optimality of sequential screening and verify this property assuming just a first-order stochastic dominance ordering of the posterior distributions. The optimality of sequential screening is also studied by Krähmer and Strausz (2015). They show that if the agent has the right to withdraw from the contract when he learns his payoff type, optimal contracts are static, and only condition on the payoff type.

With nonzero investigation costs, the contracting problem differs from the one in Courty and Li (2000) by a moral hazard issue. Specifically, to induce information acquisition the principal may have to provide extra incentives; contracts that do not condition on the information, on the other hand, need not be designed as if the agent had it. The polar case—endogenous *post*contractual information—is studied by Kräh-

---

[1]Another closely related literature considers settings with dynamic adverse selection and *multiple* allocations (see, e.g., Battaglini 2005 and Boleslavsky and Said 2013). Occasionally, I refer to both strands as the *dynamic contracting literature*. See Pavan et al. (2014) for a unified framework.

[2]In Courty and Li's model, the agent receives precontractual information already before the contract is offered, rather than only before the signing decision. This difference is irrelevant: in either case, the agent begins to act with the signing decision. A technical difference is that Courty and Li assume a continuum of types whereas I, to facilitate the comparison with CK, assume a finite number.

mer and Strausz (2011). There, the agent's incentives to acquire information differ, since he cannot quit the contract afterwards.

Various papers analyze profit-maximizing contracts for related settings in which an agent can acquire information before signing (see Bergemann and Välimäki 2002 for *surplus*-maximizing mechanisms in a general mechanism-design framework with endogenous information). In particular, Crémer et al. (1998a), Lewis and Sappington (1997), and Szalay (2009) assume that the agent never learns his payoff type for free (see Shi 2012 for an auction setting). Crémer and Khalil (1994) as well as Crémer et al. (1998b), on the other hand, consider the case that information gathering must take place already before the contract is offered, and thus cannot be induced or deterred by contract design. Finally, Compte and Jehiel (2008) demonstrate that if, in CK's setting, several agents compete for a single, bilateral contract, the principal possibly induces information gathering to find a suitable candidate.

The paper is organized as follows. The next section presents the model. Sections 3 and 4 describe the efficient outcome and briefly review CK's result, respectively. In Section 5, I first establish the main result, according to which optimal contracts may induce information gathering. Afterwards, I study optimal contracts in more detail and analyze the comparative static. Section 6 concludes. All proofs are in the appendix.

2   Model

I use a variant of the procurement model by CK. Specifically, a principal seeks to purchase some quantity of a good, which an agent can produce. Given output $q \geq 0$, marginal production costs β, and transfer $t \in \mathbb{R}$, the agent's payoff is $t - \beta q$. The principal's payoff is $V(q) - t$, where V is strictly concave, continuously differentiable, and satisfies $\lim_{q \to 0} V'(q) = \infty$ and $\lim_{q \to \infty} V'(q) = 0$.

Initially, both parties do not know the agent's marginal costs. The common prior is that β equals β_i with probability $\bar{\gamma}(\beta_i) > 0$, where $i \in I := \{1, \ldots, n\}$. Suppose $1 < n < \infty$, and let $0 < \beta_1 < \cdots < \beta_n < \infty$. The agent learns the true value of β before production takes place, but only after the date at which he must decide whether to accept the contract.

At investigation costs $e \geq 0$, the agent can acquire a signal s of β while deliberating

whether to accept. The signal equals s_j with (unconditional) probability $\pi_j > 0$, where $j \in J := \{1, \ldots, m\}$. Signal realization s_j gives rise to the posterior probability $\gamma_j(\beta_i)$ of β_i. Suppose $1 < m < \infty$, and let the possible posteriors be ordered in terms of strict first-order stochastic dominance: for any $i < n$, the cumulative posterior probability that β_i obtains,

$$\Gamma_j(\beta_i) := \sum_{k=1}^{i} \gamma_j(\beta_k),$$

strictly decreases in j.[3] A low signal s thus indicates low costs β.

A contract stipulates the terms of trade (t, q), possibly contingent on communication. The principal cannot observe whether the agent acquires the signal. Moreover, she cannot verify any reports that the agent might submit about his private information. To distinguish the two possible pieces of private information, I refer to $j \in J$ as the agent's *posterior type* and to $i \in I$ as his *cost type*. Finally, I refer to the signal s as (*precontractual*) *information*.

In detail, the timing of the interaction is as follows:

1. Principal offers contract

2. Agent can acquire information

3. Agent must accept or reject contract

4. If contract accepted: agent learns cost type before production takes place

   If contract rejected: interaction ends with zero payoffs

Importantly, precontractual information is imperfect. Specifically, let $\gamma_j(\beta_i) > 0$ for all i, j, so that the agent always still deems each cost type possible upon acquiring information. This is the only difference to the model by CK: there, s equals β, so that precontractual information reveals the cost type perfectly.

3 Efficiency

From an efficiency perspective, the contract should maximize the expected surplus of the interaction, that is, the expected difference between $V(q) - \beta q$ and the incurred

---

[3]First-order stochastic dominance orderings are common in the dynamic-contracting literature; see Pavan et al. (2014) for a discussion. The seminal paper by Courty and Li (2000) considers both first- and second-order stochastic dominance.

investigation costs, if any.[4] The efficient output level depends on the agent's production costs and equals $\hat{q}(\beta_i) := V'^{-1}(\beta_i)$. Clearly, precontractual information is redundant, and its acquisition thus socially inefficient, given that the agent anyway learns his costs—perfectly, for free, and before production takes place. I denote the maximum expected surplus by

$$\hat{W} := \sum_j \pi_j \sum_i \gamma_j(\beta_i) \left[V(\hat{q}(\beta_i)) - \beta_i \hat{q}(\beta_i) \right].$$

Note that if information acquisition was impossible, the principal could implement the efficient output level and fully extract all gains from trade, as the agent learns his production costs only after the signing of the contract.[5]

4   Benchmark: the case of perfect information

Henceforth, I study the contracting problem from the perspective of the principal, who wants to maximize her expected payoff. I focus on the basic question of whether optimal contracts induce or deter information acquisition. As a benchmark, I first review the result by CK. They show that if the information is perfect, it cannot be advantageous to induce the acquisition (see their Lemma 1).

The key insight is that, with perfect information, to each contract that induces information acquisition there corresponds a contract that does not and that implements the same terms of trade. To see this, consider a contract that induces information acquisition. Modify it by allowing the agent to exit once his learns his production costs exogenously, and by postponing any further choice until after that date. If precontractual information is perfect, the agent has then the same information as under the original contract when he decides (ultimately) about his participation and makes any further choice. By revealed preferences, the modified contract therefore implements the same terms of trade as the original one. But the agent has clearly no incentive to acquire information.

---

[4]Unless stated differently, I always mean the expectation *based on the prior.*

[5]E.g., this could be achieved with a "sell-the-firm" contract, which lets the agent choose output and specifies the transfer $\hat{t}(q) = V(q) - \hat{W}$.

If precontractual information is perfect, it is *identical* to the information that the agent obtains after signing. Therefore, a contract that deters the acquisition can nevertheless condition on the information. With imperfect information, this is clearly not the case: there, the agent must indeed conduct the acquisition if the terms of trade are to depend on precontractual information.

5   The case of imperfect information

I now return to the original setting, where precontractual information is imperfect, and show that the principal may there prefer to induce information acquisition.

5.1   Contracts that deter information acquisition

Invoking the revelation principle for dynamic games (see Myerson 1986), I restrict attention to direct, incentive-compatible contracts. The contracts that deter information acquisition have the form

$$(\bar{\mathbf{t}}, \bar{\mathbf{q}}) := (\bar{t}(\beta_k), \bar{q}(\beta_k))_{k \in I} \,.$$

Once the agent learns his cost type, he must announce it with a report $k \in I$. Given the report, the terms of trade are $(\bar{t}(\beta_k), \bar{q}(\beta_k))$. The contract is incentive-compatible if and only if the agents finds it best to dispense with information acquisition and to report the cost type truthfully.

In detail, the contract must satisfy the following conditions. First (moving backwards), the agent must report the cost type truthfully. Using for the agent's payoff the notation

$$\bar{U}(\beta_i) := \bar{t}(\beta_i) - \beta_i \bar{q}(\beta_i),$$

this condition reads

$$\bar{U}(\beta_i) \geq \bar{U}(\beta_k) + (\beta_k - \beta_i)\bar{q}(\beta_k) \quad \forall i, k \in I. \tag{1}$$

Second, the agent must accept the contract. Since he does not yet know the cost type when the participation decision is due, this condition only requires that the contract guarantees him a nonnegative payoff in expectation, rather than for each particular type. Importantly, the expectation derives from the prior, as the agent is supposed

8

to dispense with information gathering:

$$\sum_j \pi_j \sum_i \gamma_j(\beta_i)\bar{U}(\beta_i) \geq 0. \tag{2}$$

Finally, the agent must indeed not gather information. Precontractual information is valuable to him if and only if, with some *posterior types*, the contract yields a negative expected payoff: if he could update his expectation, he would be able to avoid a likely loss by rejecting the contract offer. The value of information must be smaller than the investigation costs:

$$\sum_j \pi_j \sum_i \gamma_j(\beta_i)\bar{U}(\beta_i) \geq \sum_j \pi_j \max\left\{\sum_i \gamma_j(\beta_i)\bar{U}(\beta_i), 0\right\} - e. \tag{3}$$

Consider now the principal's objective. She seeks to maximize her expected payoff. Thus, the best contracts that deter information acquisition are the solutions to

$$\bar{\mathcal{P}}: \quad \max_{(\bar{\mathbf{t}},\bar{\mathbf{q}})} \sum_j \pi_j \sum_i \gamma_j(\beta_i)[V(\bar{q}(\beta_i)) - \bar{t}(\beta_i)] \quad s.t. \quad (1)\text{--}(3).$$

5.2   Contracts that induce information acquisition

The contracts that induce information acquisition have the form

$$(\mathbf{t}, \mathbf{q}) := (t_l(\beta_k), q_l(\beta_k))_{k \in I, l \in J}.$$

First, the agent must submit a report $l \in J$ about the posterior type. Later on, when he learns his production costs, he must submit a second report, $k \in I$, this time about the cost type. The two reports lead to the terms of trade $(t_l(\beta_k), q_l(\beta_k))$. A contract $(\mathbf{t}, \mathbf{q})$ is incentive-compatible if and only if the agent prefers to gather information and to submit two truthful reports.

Since the agent obtains two pieces of private information—first the posterior type, later on the cost type—he must also report twice, and is thus screened sequentially. Note that if precontractual information was perfect rather than imperfect, so that posterior and cost type were equivalent, there would be no reason to let the agent report twice. Indeed, the revelation principle would then demand just one report (to be submitted before the agent learns his costs exogenously).

9

In detail, the contract must satisfy the following conditions. First, the agent must again report the cost type truthfully. Using for the agent's payoff the notation

$$U_j(\beta_i) := t_j(\beta_i) - \beta_i q_j(\beta_i),$$

this condition reads:

$$U_j(\beta_i) \geq U_j(\beta_k) + (\beta_k - \beta_i)q_j(\beta_i) \quad \forall i, k \in I; j \in J. \tag{4}$$

Second, the agent must also report the posterior type truthfully:

$$\sum_i \gamma_j(\beta_i)U_j(\beta_i) \geq \sum_i \gamma_j(\beta_i)U_l(\beta_i) \quad \forall j, l \in J. \tag{5}$$

Third, the agent must participate. In contrast to the contracts $(\bar{\mathbf{t}}, \bar{\mathbf{q}})$, this condition requires that the contract is acceptable conditional on the posterior type, rather than the prior, because the agent is supposed to gather information:

$$\sum_i \gamma_j(\beta_i)U_j(\beta_i) \geq 0 \quad \forall j \in J. \tag{6}$$

Finally, the agent must indeed gather information. Precontractual information can be valuable to him because it allows to report the posterior type truthfully.[6] The following condition ensures that the value of information is larger than the investigation costs:

$$\sum_j \pi_j \sum_i \gamma_j(\beta_i)U_j(\beta_i) - e \geq \max\left\{\sum_j \pi_j \sum_i \gamma_j(\beta_i)U_l(\beta_i), 0\right\} \quad \forall l \in J. \tag{7}$$

The best contracts that induce information acquisition are thus the solutions to

$$\mathcal{P}: \quad \max_{(\mathbf{t},\mathbf{q})} \sum_j \pi_j \sum_i \gamma_j(\beta_i)[V(q_j(\beta_i)) - t_j(\beta_i)] \quad s.t. \quad (4)\text{--}(7).$$

---

[6]Without (6), the individual-rationality condition, the information could be valuable also with respect to the participation decision. (Precisely, it would be valuable if and only if the contract was acceptable with some, but not all posterior types.) Requiring individual rationality is as usual without loss of generality: given any incentive-compatible contract $(\mathbf{t}', \mathbf{q}')$ that is acceptable only with posterior types $j \in J' \subset J$, the contract $(\mathbf{t}'', \mathbf{q}'')$ with $(t_l''(\beta_k), q_l''(\beta_k))_k = (t_l'(\beta_k), q_l'(\beta_k))_k$ for all $l \in J'$ and $(t_l''(\beta_k), q_l''(\beta_k))_k = (0,0)$ for all $l \notin J'$ is incentive-compatible, individually rational, and implements the same terms of trade.

The principal's *optimal* contracts, finally, can be found as follows: First, derive the best contracts that deter and induce information acquisition, respectively. Second, maximize over this binary choice.[7]

5.3   Information gathering for rent-seeking purposes

I now show that unless information gathering is prohibitively costly, the principal must make a trade-off between efficiency and surplus extraction to design an optimal contract. So suppose she chooses to deter the acquisition of the socially useless information. The following lemma simplifies the relevant conditions.

Lemma 1. *Among the contracts* $(\bar{\mathbf{t}}, \bar{\mathbf{q}})$*, for each contract that satisfies (1)–(3) there is a contract with identical expected payoffs for both parties that satisfies*

$$\bar{q}(\beta_i) - \bar{q}(\beta_{i+1}) \geq 0 \quad \forall i < n \tag{8}$$

$$\sum_j \pi_j \left[\bar{U}(\beta_n) + \sum_{i=1}^{n-1} \Gamma_j(\beta_i)(\beta_{i+1} - \beta_i)\bar{q}(\beta_{i+1}) \right] \geq 0 \tag{9}$$

$$\sum_{j=l}^m \pi_j \left[\bar{U}(\beta_n) + \sum_{i=1}^{n-1} \Gamma_j(\beta_i)(\beta_{i+1} - \beta_i)\bar{q}(\beta_{i+1}) \right] + e \geq 0 \quad \forall l \in \{2, \ldots, m\} \tag{10}$$

together with

$$\bar{U}(\beta_i) - \bar{U}(\beta_{i+1}) = (\beta_{i+1} - \beta_i)\,\bar{q}(\beta_{i+1}) \quad \forall i < n. \tag{11}$$

Moreover, (8)–(11) imply (1)–(3).

According to (11), the agent earns an extra payoff for not exaggerating his production costs. By (8), on the other hand, reporting lower costs obliges to produce more output. These two, standard conditions make sure that the agent reports his cost type truthfully.

(9) is the participation condition. Note that by (11), the agent's payoff with cost type i equals

$$\bar{U}(\beta_n) + \sum_{k=i}^{n-1}(\beta_{i+1} - \beta_i)\bar{q}(\beta_{i+1}).$$

---

[7]The existence of these contracts follows from standard existence theorems for convex optimization problems and is therefore not proved here.

Henceforth, I refer to the summand on the right-hand side as the agent's *rent* with cost type i, and to $-\bar{U}(\beta_n)$ as a *participation fee*. The agent finds the contract acceptable if and only if, conditional on the prior, the participation fee does not exceed the expected rent.

(10), finally, guarantees that the agent does not gather information. Precontractual information is valuable to the agent if and only if the participation fee exceeds the expected rent conditional on some *posterior types*. To understand the condition, note that the rent is larger the lower the cost type. By the first-order stochastic dominance ranking of the posteriors, the *expected* rend is thus larger the lower the posterior type. This implies that if the participation fee exceeds the expected rent with some posterior type l, then so it does for all types $j > l$.

Condition (11) can be inserted directly into the principal's objective function, which I now write as the difference between the expected surplus and the expected payoff to the agent:

$$\bar{\Delta} := \sum_j \pi_j \sum_i \gamma_j(\beta_i)[V(\bar{q}(\beta_i)) - \beta_i \bar{q}(\beta_i)]$$
$$- \sum_j \pi_j \left[\bar{U}(\beta_n) + \sum_{i=1}^{n-1} \Gamma_j(\beta_i)(\beta_{i+1} - \beta_i)\bar{q}(\beta_{i+1}) \right].$$

Note that by (11), the transfers are pinned down by the output schedule up to the participation fee. I may therefore regard contracts that deter information acquisition alternatively as combinations $(-\bar{U}(\beta_n), \bar{\mathbf{q}})$ of participation fee and output schedule. The best contract, denoted by $(-\bar{U}^*(\beta_n), \bar{\mathbf{q}}^*)$, is the (unique) solution to

$$\bar{\mathscr{P}}: \quad \max_{(-\bar{U}(\beta_n), \bar{\mathbf{q}})} \bar{\Delta} \quad s.t. \quad (8)-(10).$$

I can now show that the agent's possibility to gather information for rent-seeking purposes may force the principal to make a trade-off between efficiency and surplus extraction. Suppose information acquisition was impossible. In that case, condition (10) could be ignored. Clearly, the principal would implement the efficient output schedule, $\bar{\mathbf{q}} = \hat{\mathbf{q}}$, and fully extract all gains from trade with a participation fee equal to the agent's expected rent,

$$-\bar{U}(\beta_n) = -\hat{U}(\beta_n) := \sum_j \pi_j \sum_{i=1}^{n-1} \Gamma_j(\beta_i)(\beta_{i+1} - \beta_i)\hat{q}(\beta_{i+1}).$$

The main result of this section is that this contract may violate condition (10):

Proposition 1. *There exists a cutoff level of investigation costs $\hat{e} > 0$ such that $(-\hat{U}(\beta_n), \hat{\mathbf{q}})$ satisfies (10) if and only if $e \geq \hat{e}$.*

Since the participation fee $-\hat{U}(\beta_n)$ equals the expected rent conditional on the prior, it must exceed it conditional on some posterior types. Hence, precontractual information is valuable to the agent. The described contract satisfies (10), the no-information-acquisition condition, thus only for high investigation costs.

By acquiring information, the agent has the costly option to check his payoff from the contract before deciding whether to accept it. Proposition 1 says that unless this rent-seeking activity is prohibitively expensive, the principal must make a trade-off between efficiency and surplus extraction to find an optimal contract. I now demonstrate that it can then be advantageous to make the terms of trade contingent on the information—and thus to induce the acquisition.

5.4 Use of contracts that induce information acquisition

This section presents the main result of the paper, according to which the principal may prefer to induce information acquisition. In the remaining analysis, I denote by $\bar{W}(e)$ and $W(e)$ the principal's expected payoff from the best contracts that deter and induce information acquisition, respectively, depending on the investigation costs.

My derivation will not involve a full characterization of these best contracts. In particular, I will not provide an analog of Lemma 1 and simplify the conditions that the contracts that induce information acquisition must satisfy. This is because, as is well known in the literature on dynamic contracting with exogenous information, truthtelling conditions for sequential-screening contracts generally lack useful characterizations (see, e.g., Pavan et al. 2014). The usual approach is to make stringent regularity assumptions that allow to focus on a relaxed problem (see Battaglini and Lamba 2014 for details and a discussion). Specifically, the assumptions would concern the probability distributions of the model, and thus the quality of precontractual information. I omit them because my result holds generally with imperfect information. In fact, in the present setting (and within the general model), a characterization of the

best sequential-screening contracts $(\mathbf{t}, \mathbf{q})$ would entail additional complications, given that precontractual information is endogenous.[8]

I start with a partial comparison.

Lemma 2. *The functions* $\bar{W}$ *and* W *have the following properties:*

1. $\bar{W}$ *is nondecreasing and* W *nonincreasing;*

2. *they are continuous;*

3. *they have a unique intersection.*

The first and the last statement are intuitive: Contracts that induce information acquisition have two disadvantages. First, they must satisfy a more restrictive participation condition. Specifically, they must be profitable for the agent conditional on the posterior type, whereas a contract that deters the acquisition only needs to be profitable conditional on the prior, as long as the value of information does not exceed the investigation costs. The second disadvantage is that information acquisition must be incentive-compatible, to compensate the agent for the investigation costs.[9] Now, these disadvantages are less substantial the lower the investigation costs. In particular, if $e = 0$, the agent effectively *has* precontractual information anyway. In fact, in that case, the only difference between the contracts $(\bar{\mathbf{t}}, \bar{\mathbf{q}})$ and $(\mathbf{t}, \mathbf{q})$ is that the latter ones allow to make the terms of trade contingent on the information.

The key step is now to recognize that if information acquisition entails zero costs, so that the contract must anyway be designed as if the agent had the information, it is strictly optimal to make the terms of trade contingent:

Lemma 3. $W(0) > \bar{W}(0)$.

Remark 1. If information acquisition entails zero costs, the contracting problem is in fact equivalent to the one considered in the seminal paper on sequential screening by Courty and Li (2000), where the agent exogenously has imperfect precontractual information from the outset. Courty and Li provide a complete characterization of optimal contracts, using regularity assumptions. Under these assumptions, the optimal contracts in their model do condition on precontractual information as well.

---

[8]Section 5.5 provides a characterization for the special case of binary posterior and cost types.

[9]Hence, $W(e) \leq \hat{W} - e$ for all e, whereas by Proposition 1 $\bar{W}(e) = \hat{W}$ for $e \geq \hat{e}$.

In the following, I give a detailed illustration of the lemma for the case of binary posterior and cost types ($m = n = 2$). I refer to $j = 1, 2$ as the *optimistic* and the *pessimistic* posterior type, respectively, and write $j = O, P$. Consider first the contracts that deter information acquisition. As in the previous section, I regard them as combinations $(-\bar{U}(\beta_2), \bar{\mathbf{q}})$ of participation fee and output schedule. The agent earns a rent of $(\beta_2 - \beta_1)\bar{q}(\beta_2)$ if he has low production costs, and no rent in case of high costs. Conditional on posterior type j, his expected payoff thus equals

$$\bar{U}(\beta_2) + \gamma_j(\beta_1)(\beta_2 - \beta_1)\bar{q}(\beta_2).$$

Now, since information gathering entails zero costs, the contract must be acceptable for each posterior type. With the optimistic type, however, the agent is more confident to earn the rent: $\gamma_O(\beta_1) > \gamma_P(\beta_1)$. Hence, the principal cannot extract the expected rent fully—an illustration of Proposition 1. As a consequence, the best contract that deters information acquisition stipulates an inefficiently low output level $\bar{q}(\beta_2)$.

The key insight is that the terms of trade should never be distorted for *both* posterior types in this fashion. With the optimistic type, the agent has a larger valuation for rent, given that he is more confident to earn it. Hence, if—with the optimistic type—he was to produce efficiently and pay an extra participation fee that fully extracts the extra expected rent, whereas—with the pessimistic type—he was to stick with the original contract, he would comply.

In detail, suppose the agent may choose among the best contract that deters information acquisition, $(-\bar{U}^*(\beta_2), \bar{\mathbf{q}}^*)$, and a contract $(-\bar{U}'(\beta_2), \bar{\mathbf{q}}')$ which only differs in that $\bar{q}'(\beta_2) = \hat{q}(\beta_2)$ and

$$-\bar{U}'(\beta_2) = -\bar{U}^*(\beta_2) + \gamma_O(\beta_1)(\beta_2 - \beta_1)\left[\hat{q}(\beta_2) - \bar{q}^*(\beta_2)\right].$$

Under the alternative contract, the agent thus produces efficiently, and hence earns an extra rent. However, he must also pay an extra participation fee, which equals the extra *expected* rent conditional on the optimistic posterior type. With that type, he would consequently be indifferent, and thus willing to choose the alternative contract:

$$
\begin{aligned}
& \bar{U}'(\beta_2) + \gamma_O(\beta_1)(\beta_2 - \beta_1)\bar{q}'(\beta_2) \\
= {}& \bar{U}^*(\beta_2) - \gamma_O(\beta_1)(\beta_2 - \beta_1)\left[\hat{q}(\beta_2) - \bar{q}^*(\beta_2)\right] + \gamma_O(\beta_1)(\beta_2 - \beta_1)\hat{q}(\beta_2) \\
= {}& \bar{U}^*(\beta_2) + \gamma_O(\beta_1)(\beta_2 - \beta_1)\bar{q}^*(\beta_2).
\end{aligned}
$$

With the pessimistic type, in contrast, the agent would stick with the original contract, given that he is less confident to earn the extra rent:

$$\bar{U}'(\beta_2) + \gamma_P(\beta_1)(\beta_2 - \beta_1)\bar{q}'(\beta_2)$$

$$= \bar{U}^*(\beta_2) - \gamma_O(\beta_1)(\beta_2 - \beta_1)\left[\hat{q}(\beta_2) - \bar{q}^*(\beta_2)\right] + \gamma_P(\beta_1)(\beta_2 - \beta_1)\hat{q}(\beta_2)$$

$$< \bar{U}^*(\beta_2) - \gamma_P(\beta_1)(\beta_2 - \beta_1)\left[\hat{q}(\beta_2) - \bar{q}^*(\beta_2)\right] + \gamma_P(\beta_1)(\beta_2 - \beta_1)\hat{q}(\beta_2)$$

$$= \bar{U}^*(\beta_2) + \gamma_P(\beta_1)(\beta_2 - \beta_1)\bar{q}^*(\beta_2).$$

Clearly, this scheme amounts to a contract $(\mathbf{t}, \mathbf{q})$ that generates strictly more expected surplus than the best contract that deters information acquisition but provides the agent with the same expected payoff. The proof extends the reasoning to the original model, where the numbers of posterior and cost types are arbitrary, and shows that the principal can strictly improve over the best contract that deters information acquisition with a contract $(\mathbf{t}, \mathbf{q})$ that exhibits "no distortion at the top", that is, which implements the efficient output schedule if the agent has the posterior type $j = 1$.[10,11]

Thus, if information gathering entails zero costs, so that the contract must anyway be designed as if the agent *had* the information, it is strictly optimal to make the terms of trade contingent because the information determines the agent's expectation of the rent that he will earn, and hence his willingness to pay participation fees.

Lemmas 2 and 3 directly imply the main result of the paper. Of course, it would be best if the agent could not obtain precontractual information at all. But given that he has this possibility, it can be advantageous not to deter the acquisition:

Theorem 1. *There exists a cutoff level of investigation costs $\bar{e} > 0$ such that*
- *if $e < \bar{e}$, optimal contracts induce information acquisition ($W(e) > \bar{W}(e)$);*
- *if $e > \bar{e}$, optimal contracts deter information acquisition ($W(e) < \bar{W}(e)$).*

---

[10]The reasoning is reminiscent of Battaglini (2005). In a setting with multiple allocations, dynamic adverse selection, and binary states, he shows that once the agent reports for the first time to be in the "good" state, all subsequent allocations are efficient (his "generalized no-distortion-at-the-top principle").

[11]The reasoning might suggest that given $e = 0$, optimal contracts generally implement more efficient output levels the lower the posterior type. Results from the dynamic-contracting literature show that this is not the case: distortions in optimal contracts highly depend on the details of the model (see in particular Battaglini and Lamba 2014).

5.5 Design of contracts that induce information acquisition

Henceforth, I restrict attention to the case of binary posterior and cost types. The present section provides a full characterization of the best contracts that induce information acquisition—the optimal contracts for low investigation costs. The best contracts that *deter* the acquisition are characterized in CK's analysis, whose results easily extend to imperfect information.

Analogously to Lemma 1, I first simplify the relevant conditions.

Lemma 4. *Among the contracts* $(\mathbf{t}, \mathbf{q})$*, for each contract that satisfies (4)–(7) there is a contract with identical expected payoffs for both parties that satisfies*

$$q_j(\beta_1) - q_j(\beta_2) \geq 0 \quad \forall j = O, P \tag{12}$$

$$U_P(\beta_2) + \gamma_P(\beta_1)(\beta_2 - \beta_1)q_P(\beta_2) \geq 0 \tag{13}$$

$$\pi_O \pi_P \left[\gamma_O(\beta_1) - \gamma_P(\beta_1)\right](\beta_2 - \beta_1)\left[q_O(\beta_1) - q_P(\beta_2)\right] - e \geq 0 \tag{14}$$

together with

$$U_O(\beta_1) - U_O(\beta_2) = (\beta_2 - \beta_1)q_O(\beta_1) \tag{15}$$

$$U_P(\beta_1) - U_P(\beta_2) = (\beta_2 - \beta_1)q_P(\beta_2) \tag{16}$$

$$U_O(\beta_2) - U_P(\beta_2) = -\gamma_O(\beta_1)(\beta_2 - \beta_1)\left[q_O(\beta_1) - q_P(\beta_2)\right] + \frac{e}{\pi_O}. \tag{17}$$

Moreover, (12)–(17) imply (4)–(7).

(12), (15), and (16) together ensure that the agent reports his production costs truthfully. As with the contracts that deter information acquisition, this requires that the agent earns a rent in case that the costs are low. The agent's expected payoff thus again equals the difference between expected rent and participation fee. However, both rent and fee now depend on posterior type.

(13) is the participation condition. By incentive-compatibility, it suffices that the agent finds the contract profitable if he has the pessimistic posterior type, with which he is less confident to earn rent.

(14) and (17), finally, together ensure that the agent acquires information and reports the posterior type truthfully.[12] Specifically, (17) requires an extra payoff for

---

[12]It is a special feature of the case of binary posterior types that the information-acquisition condition, (7), implies the truthtelling condition (5).

the optimistic type that, in particular, compensates for the investigation costs. Any extra expected rent that this type earns with his terms of trade, however, can be extracted with an extra participation fee of

$$\gamma_O(\beta_1)(\beta_2 - \beta_1)\left[q_O(\beta_1) - q_P(\beta_2)\right],$$

as illustrated in the previous section. (14) requires extra output, to rule out that the agent *pretends* to have the optimistic type. Note that extra payoff and output increase in the investigation costs. This is a common feature of contracts with endogenous information (see in particular Lewis and Sappington 1997, who coined the term "super high-powered incentive scheme", and the general analysis by Szalay 2009).

Conditions (15)–(17) can be inserted directly into the objective function, which I now write as the difference between the expected surplus and the expected payoff to the agent:

$$\Delta := \sum_j \pi_j \sum_i \gamma_j(\beta_i)\left[V(q_j(\beta_i)) - \beta_i q_j(\beta_i)\right] - e$$
$$- \sum_j \pi_j\left[U_P(\beta_2) + \gamma_j(\beta_1)(\beta_2 - \beta_1)q_P(\beta_2)\right].$$

Note that by (15)–(17), the transfers are pinned down by the output schedules up to the participation fee for the pessimistic posterior type. I may therefore regard contracts that induce information acquisition alternatively as combinations $(-U_P(\beta_2), \mathbf{q})$ of this participation fee and the menu of output schedules. The best contract, denoted by $(-U_P^*(\beta_2), \mathbf{q}^*)$, is the (unique) solution to

$$\mathscr{P}: \quad \max_{(-U_p(\beta_2), \mathbf{q})} \Delta \quad s.t. \quad (12)\text{–}(14).$$

I can now give the characterization. Clearly, the principal will choose the participation fee as large as the participation condition, (13), admits, and thus fully extract the expected rent to the pessimistic posterior type:

$$-U_P^*(\beta_2) = \gamma_P(\beta_1)(\beta_2 - \beta_1)q_P^*(\beta_2).$$

The optimal output schedules are characterized in the following proposition.

Proposition 2. *Let λ be the nonnegative Lagrange multiplier associated to condition (14) in $\mathscr{P}$. Then,*

$$q_O^*(\beta_1) = V'^{-1}\left(\beta_1 - \frac{\lambda\pi_p\left[\gamma_O(\beta_1) - \gamma_P(\beta_1)\right](\beta_2 - \beta_1)}{\gamma_O(\beta_1)}\right)$$

$$q_O^*(\beta_2) = \hat{q}(\beta_2)$$

$$q_P^*(\beta_1) = \hat{q}(\beta_1)$$

$$q_P^*(\beta_2) = V'^{-1}\left(\beta_2 + \frac{\pi_O(1 + \lambda\pi_P)\left[\gamma_O(\beta_1) - \gamma_P(\beta_1)\right](\beta_2 - \beta_1)}{\pi_P\gamma_P(\beta_2)}\right).$$

There exists a cutoff level of investigation costs $e^* > 0$ such that $\lambda > 0$ if and only if $e > e^*$.

Since the principal can fully extract any extra expected rent that the optimistic type receives with his terms of trade, the surplus that she can *not* extract only depends on the terms of trade for the pessimistic type, precisely, on the output level $q_P(\beta_2)$. Now, if the investigation costs are low, the principal need not distort the contract in order to induce information acquisition. Therefore, she stipulates an inefficiently low output level $q_P(\beta_2)$ and otherwise requires efficient production. For high e, on the other hand, the information-acquisition condition (14) binds. The principal will then increase $q_O(\beta_1)$ above the efficient level and further reduce $q_P(\beta_2)$, to implement the extra output that (14) requires.

5.6 Comparative static

According to the present paper, the principal may induce information acquisition if the information is imperfect. CK, on the other hand, demonstrate that the acquisition of perfect information will always be deterred. In view of this contrast, I now study how the gain from inducing information acquisition changes if the information becomes more precise.

Let $\mathcal{I} := \langle(\pi_j)_j, ((\gamma_j(\beta_i))_{i,j}\rangle$ be an *information structure*. I define increased precision in terms of Blackwell sufficiency (see Blackwell and Girshick 1954, Ch. 12.5). Specifically, given two information structures $\mathcal{I}', \mathcal{I}''$, I say that precontractual information is *more precise* under $\mathcal{I}''$ than under $\mathcal{I}'$ if and only if

$$\gamma_O''(\beta_1) \geq \gamma_O'(\beta_1), \quad \gamma_P''(\beta_2) \geq \gamma_P'(\beta_2),$$

and, for both cost types $i = 1, 2$,

$$\sum_j \pi_j'' \gamma_j''(\beta_i) = \sum_j \pi_j' \gamma_j'(\beta_i).$$

With more precise information, the posterior probabilities of the two cost types are thus more extreme, whereas the prior probabilities remain unchanged.

Let the *gain* from inducing information acquisition be $W(0) - \bar{W}(0)$, the difference between the maximum expected payoffs with contracts that induce and deter information acquisition, respectively, evaluated at zero investigation costs, where this difference is largest (by Lemma 2). As shown in the appendix (see the proof of Lemma A3),

$$\bar{W}(0) = \max_{\bar{\mathbf{q}}} \sum_j \pi_j \sum_i \gamma_j(\beta_i) \left[V(\bar{q}(\beta_i)) - \beta_i \bar{q}(\beta_i) \right]$$
$$- \pi_O \left[\gamma_O(\beta_1) - \gamma_P(\beta_2) \right] (\beta_2 - \beta_1) \bar{q}(\beta_2),$$

and according to the previous section,

$$W(0) = \max_{\mathbf{q}} \sum_j \pi_j \sum_i \gamma_j(\beta_i) \left[V(q_j(\beta_i)) - \beta_i q_j(\beta_i) \right]$$
$$- \pi_O \left[\gamma_O(\beta_1) - \gamma_P(\beta_1) \right] (\beta_2 - \beta_1) q_P(\beta_2).$$

Note that if precontractual information gets more precise, *both* payoffs decrease. Specifically, the principal has to leave more surplus to the agent per unit $\bar{q}(\beta_2)$ and $q_P(\beta_2)$, respectively, as the term $\pi_O[\gamma_O(\beta_1) - \gamma_P(\beta_1)]$ increases. This is intuitive: if precontractual information was not available—or pure noise—the principal could fully extract all gains from trade.

As pointed out in the previous section, with contracts that induce information acquisition the surplus that the principal cannot extract only depends on the terms of trade for the pessimistic posterior type, precisely, on the output level $q_P(\beta_2)$. With contracts that *deter* information acquisition, in contrast, it depends on the output level $\bar{q}(\beta_2)$—which applies to both posterior types. In the limit, when precontractual information becomes arbitrarily precise (i.e., perfect), this difference becomes irrelevant, as $\gamma_O(\beta_2) \to 0$. The gain from inducing information acquisition vanishes then:

Proposition 3. *Let* $(\mathcal{I}_\tau)_{\tau \geq 0}$ *be a sequence of information structures such that if* $\tau'' > \tau'$, *precontractual information is more precise under* $\mathcal{I}_{\tau''}$ *than under* $\mathcal{I}_{\tau'}$, *and* $\lim_\tau \gamma_{O,\tau}(\beta_1) = \lim_\tau \gamma_{P,\tau}(\beta_2) = 1$. *Then,* $\lim_\tau W_\tau(0) - \bar{W}_\tau(0) = 0$.

Apart from the limit, increased precision may raise the gain from inducing information acquisition, as the example below demonstrates. This is because, with increased precision, the principal must leave more surplus to the agent per unit $\bar{q}(\beta_2)$ and $q_P(\beta_2)$, respectively, and the latter output level is at the optimum smaller than the former one, since it only applies to the pessimistic posterior type.

Example 1. For $V(q) = 100\sqrt{q}$, $\beta_1 = 1$, and $\beta_2 = 2$, let $\mathcal{I}'$ be characterized by $(\pi_O', \gamma_O'(\beta_1), \gamma_P'(\beta_2)) = (0.5, 0.6, 0.6)$ and $\mathcal{I}''$ by $(\pi_O'', \gamma_O''(\beta_1), \gamma_P''(\beta_2)) = (0.5, 0.8, 0.8)$. Then, precontractual information is more precise under $\mathcal{I}''$ than under $\mathcal{I}'$, but $W'(0) - \bar{W}'(0) \approx 3$ whereas $W''(0) - \bar{W}''(0) \approx 8$.

6 Conclusion

The paper offers a new perspective on information gathering for rent-seeking purposes in contracting. Of course, it is best for a principal if the agent obtains all private information only after accepting a contract. But if the agent *can* acquire private information in advance, while deliberating whether to accept, the principal may not deter the acquisition. A key insight is that situations with pre-and postcontractual private information are generally instances of sequential-screening problems.

In this paper, the agent learns his payoff type exogenously after the signing of the contract. In many situations, a principal can *disclose* some source of private information to an agent. The recent literature on disclosure rules in optimal auctions, for example, considers auction settings in which an auctioneer can disclose, without observing, private information to bidders that is relevant for their valuations (see, e.g., Bergemann and Pesendorfer 2007; Esö and Szentes 2007; Li and Shi 2015). In such a situation, an agent's possibility to gather information for rent-seeking purposes might affect both the design of the contract and the disclosure rule. This could be a topic for future research.

Appendix

Proof of Lemma 1

By standard arguments (see, e.g., Laffont and Martimort 2002, Ch. 3.1), (1) is equivalent to

$$\bar{U}(\beta_i) - \bar{U}(\beta_{i+1}) \in [(\beta_{i+1} - \beta_i)\,\bar{q}(\beta_{i+1}), (\beta_{i+1} - \beta_i)\,\bar{q}(\beta_i)] \quad \forall i < n \qquad \text{(A.1)}$$

together with the monotonicity condition (8). To proceed with the proof, I establish the following two lemmas. The first one allows to simplify (3).

Lemma A1. *Given (A.1), (8), and (2), condition (3) holds if and only if*

$$\sum_{j=l}^{m} \pi_j \sum_i \gamma_j(\beta_i)\bar{U}(\beta_i) + e \geq 0 \quad \forall l \in \{2,\dots,m\}. \qquad \text{(A.2)}$$

Proof. Note first that (A.1) and (8) together imply that $\bar{U}(\beta_i)$ is decreasing in i and hence, by the first-order stochastic dominance ordering of the posteriors, that $\sum_i \gamma_j(\beta_i)\bar{U}(\beta_i)$ is decreasing in j. Note also that, thus, a necessary condition for (2) is $\sum_i \gamma_1 \bar{U}(\beta_i) \geq 0$.

Now, suppose (A.1), (8), and (2) hold but (3) not. Then, there is a posterior type $l \in \{2,\dots,m\}$ such that

$$0 > \sum_j \pi_j \min\left\{0, \sum_i \gamma_j(\beta_i)\bar{U}(\beta_i)\right\} + e = \sum_{j=l}^{m} \pi_j \sum_i \gamma_j(\beta_i)\bar{U}(\beta_i) + e.$$

Hence, (A.2) does not hold. Suppose next that (A.1), (8), and (2) hold but (A.2) not. Let $l \in \{2,\dots,m\}$ be the smallest posterior type such that $\sum_i \gamma_j \bar{U}(\beta_i) < 0$. Then,

$$0 > \sum_{j=l}^{m} \pi_j \sum_i \gamma_j(\beta_i)\bar{U}_i + e = \sum_j \pi_j \min\left\{0, \sum_i \gamma_j(\beta_i)\bar{U}(\beta_i)\right\} + e.$$

Hence, condition (3) does not hold. □

Next, I show that (A.1) can without loss of generality be replaced by the stronger condition (11), and consequently (2) by (9) and (A.2) by (10). This will conclude the proof of Lemma 1.

Lemma A2. *Among the contracts $(\bar{\mathbf{t}}, \bar{\mathbf{q}})$, for each contract that satisfies (A.1), (8), (2), and (A.2) there is a contract with identical expected payoffs for both parties that satisfies (8)–(11).*

Proof. Let $(\bar{\mathbf{t}}', \bar{\mathbf{q}}')$ be any contract that satisfies (A.1), (8), (2), and (A.2). For a given cost type $k < n$, define

$$\epsilon := \bar{U}'(\beta_k) - \bar{U}'(\beta_{k+1}) - (\beta_{k+1} - \beta_k)\,\bar{q}'(\beta_{k+1}).$$

Consider the alternative contract $(\bar{\mathbf{t}}'', \bar{\mathbf{q}}'')$, which differs from $(\bar{\mathbf{t}}', \bar{\mathbf{q}}')$ only with respect to transfers, namely such that

$$\bar{t}''(\beta_i) = \begin{cases} \bar{t}'(\beta_i) + \sum_j \pi_j \Gamma_j(\beta_k)\epsilon - \epsilon & \text{if } i \in \{1, \ldots, k\} \\ \bar{t}'(\beta_i) + \sum_j \pi_j \Gamma_j(\beta_k)\epsilon & \text{if } i \in \{k+1, \ldots, n\}. \end{cases}$$

Given these transfers, the alternative contract satisfies

$$\bar{U}''(\beta_k) - \bar{U}''(\beta_{k+1}) = (\beta_{k+1} - \beta_k)\bar{q}''(\beta_{k+1}),$$

and the other payoff differences are as under the original contract,

$$\bar{U}''(\beta_i) - \bar{U}''(\beta_{i+1}) = \bar{U}'(\beta_i) - \bar{U}'(\beta_{i+1}) \quad \forall i \in \{1, \ldots, n-1\} \setminus k.$$

Moreover,

$$\sum_j \pi_j \sum_i \gamma_j(\beta_i)\bar{t}''(\beta_i) = \sum_j \pi_j \sum_i \gamma_j(\beta_i)\bar{t}'(\beta_i),$$

so the alternative contract results in identical expected payoffs for both parties as the original one. Note that this implies that the alternative contract satisfies condition (2). Finally, for any posterior type $l \in \{2, \ldots, m\}$ it holds that

$$
\begin{aligned}
& \sum_{j=l}^m \pi_j \sum_i \gamma_j(\beta_i)\bar{U}''(\beta_i) + e \\
=\ & \sum_{j=l}^m \pi_j \left[\sum_i \gamma_j(\beta_i)\bar{U}'(\beta_i) + \sum_j \pi_j \Gamma_j(\beta_k)\epsilon - \Gamma_j(\beta_k)\epsilon \right] + e \\
=\ & \sum_{j=l}^m \pi_j \sum_i \gamma_j(\beta_i)\bar{U}'(\beta_i) + \sum_{j=l}^m \pi_j \sum_j \pi_j \Gamma_j(\beta_k)\epsilon - \sum_{j=l}^m \pi_j \Gamma_j(\beta_k)\epsilon + e \\
\geq\ & \sum_{j=l}^m \pi_j \sum_i \gamma_j(\beta_i)\bar{U}'(\beta_i) + e \\
\geq\ & 0,
\end{aligned}
$$

where the first inequality follows from the first-order stochastic dominance ranking of the posteriors and the second one from the hypothesis that the original contract satisfies (A.2). Thus, the alternative contract also satisfies (A.2).

Repeating the argument for each $k < n$ establishes Lemma A2. $\qquad \square$

Proof of Proposition 1

The first-order stochastic dominance ranking of the posteriors implies that

$$\sum_{i=1}^{n-1} \Gamma_j(\beta_i)(\beta_{i+1} - \beta_i)\hat{q}(\beta_{i+1})$$

strictly decreases in j. Hence, there are posterior types $l \in \{2, \ldots, m\}$ with

$$\sum_{j=l}^{m} \pi_j \left[\hat{U}(\beta_n) + \sum_{i=1}^{n-1} \Gamma_j(\beta_i)(\beta_{i+1} - \beta_i)\hat{q}(\beta_{i+1}) \right] < 0.$$

Let l^* be the smallest such l. Then, (10) holds if and only if

$$e \geq \hat{e} := -\sum_{j=l^*}^{m} \pi_j \left[\hat{U}(\beta_n) + \sum_{i=1}^{n-1} \Gamma_j(\beta_i)(\beta_{i+1} - \beta_i)\hat{q}(\beta_{i+1}) \right] > 0.$$

Proof of Lemma 2

Formally, $\bar{W}$ and W are the value functions of the optimization problems $\bar{\mathscr{P}}$ (or $\bar{\mathcal{P}}$) and $\mathcal{P}$, respectively, for the parameter $e \geq 0$. That these functions are nondecreasing and nonincreasing, respectively, follows from the fact that if e increases, the choice sets in $\bar{\mathscr{P}}$ and $\mathcal{P}$ become larger and smaller, respectively, whereas the objective functions do not vary with e.

To prove continuity, I use the fact that if a problem has a concave objective function and quasiconcave constraint functions, then its value function is concave (e.g., de la Fuente 2000, Thm. 2.12, p. 313). A concave function is continuous on the interior of its domain. Now, $\bar{\mathscr{P}}$ and $\mathcal{P}$ have concave objective and affine, hence quasiconcave, constraint functions. Note that this would also be the case if $e \in \mathbb{R}$ rather than $e \geq 0$. The such extended value functions would thus be concave, hence continuous, on $\mathbb{R}$. So $\bar{W}$ and W are continuous on their domain, $\mathbb{R}_{\geq 0}$.

To prove that these functions have a unique intersection, I first show that intersections exist. At $e = 0$, to any contract $(\bar{\mathbf{t}}, \bar{\mathbf{q}})$ that satisfies (1)–(3) there corresponds a contract $(\mathbf{t}, \mathbf{q})$ that satisfies (4)–(7) with

$$(t_l(\beta_k), q_l(\beta_k))_k = (\bar{t}(\beta_k), \bar{q}(\beta_k))_k \quad \forall l \in J.$$

Thus, $W(0) \geq \bar{W}(0)$. Define $\tilde{e} := \hat{W} - \bar{W}(0)$, and note that $W(e) \leq \hat{W} - e$ for all e. Since $\bar{W}$ is nondecreasing, $W(\tilde{e}) \leq \bar{W}(\tilde{e})$. Since W and $\bar{W}$ are furthermore continuous, the intermediate value theorem implies that intersections exist.

To see that there is just one intersection, note first that $\bar{W}(e) = \hat{W}$ for $e \geq \hat{e}$ by Proposition 1, so any intersection must lie in $[0, \hat{e})$. Now, since $\bar{W}$ is concave, it is differentiable almost everywhere. At points where the derivative exists, it holds that

$$\frac{\mathrm{d}\bar{W}(e)}{\mathrm{d}e} = \sum_{j=2}^{m} \kappa_j,$$

where κ_j, $j = 2, \ldots, m$, are nonnegative Lagrange multipliers associated to condition (10). Again by Proposition 1, for $e \in [0, \hat{e})$ at least one multiplier is strictly positive. Being continuous, $\bar{W}$ must hence be strictly increasing on $[0, \hat{e})$. This implies that there is just one intersection, given that W is nonincreasing.

Proof of Lemma 3

Consider first the contracts that deter information acquisition. The following lemma says that the best contract—the solution $(-U^*(\beta_n), \bar{\mathbf{q}}^*)$ to $\bar{\mathscr{P}}$—exhibits downward distortions for all cost types except the lowest one.

Lemma A3. *If $e = 0$, then $\bar{q}^*(\beta_i) < \hat{q}(\beta_i)$ for all $i \in \{2, \ldots, m\}$ and $q^*(\beta_1) = \hat{q}(\beta_1)$.*

Proof. If $e = 0$, then, by the first-order stochastic dominance ordering of the posteriors, (10) holds if and only if

$$\bar{U}(\beta_n) + \sum_{i=1}^{n-1} \Gamma_m(\beta_i)(\beta_{i+1} - \beta_i)\bar{q}(\beta_{i+1}) \geq 0.$$

Clearly, this condition implies (9) and binds. $\bar{\mathscr{P}}$ can thus be restated as

$$\max_{\bar{\mathbf{q}}} \sum_{j} \pi_j \sum_{i} \gamma_j(\beta_i)[V(\bar{q}(\beta_i)) - \beta_i \bar{q}(\beta_i)]$$

$$- \sum_{j=1}^{m-1} \pi_j \sum_{i=1}^{n-1} [\Gamma_j(\beta_i) - \Gamma_m(\beta_i)] (\beta_{i+1} - \beta_i)\bar{q}(\beta_{i+1}) \quad s.t. \quad (8).$$

Let $\mu(\beta_i)$, $i < n$, be nonnegative Lagrange multipliers associated to (8). Then,

$$\bar{q}^*(\beta_n) \in \arg\max_{\bar{q}} \sum_{j} \pi_j \gamma_j(\beta_n) \left[V(\bar{q}) - \beta_n \bar{q}\right]$$

$$- \sum_{j=1}^{m-1} \pi_j \left[\Gamma_j(\beta_{n-1}) - \Gamma_m(\beta_{n-1})\right] (\beta_n - \beta_{n-1})\bar{q}$$

$$+ \mu(\beta_{n-1}) \left[\bar{q}^*(\beta_{n-1}) - \bar{q}\right].$$

For all cost types $i \in \{2, \ldots, n-1\}$,

$$\bar{q}^*(\beta_i) \in \arg\max_{\bar{q}} \sum_j \pi_j \gamma_j(\beta_i) \left[V(\bar{q}) - \beta_i \bar{q} \right]$$

$$- \sum_{j=1}^{m-1} \pi_j \left[\Gamma_j(\beta_{i-1}) - \Gamma_m(\beta_{i-1}) \right] (\beta_i - \beta_{i-1}) \bar{q}$$

$$+ \mu(\beta_i) \left[\bar{q} - \bar{q}^*(\beta_{i+1}) \right] + \mu(\beta_{i-1}) \left[\bar{q}^*(\beta_{i-1}) - \bar{q} \right]).$$

Finally,

$$\bar{q}^*(\beta_1) \in \arg\max_{\bar{q}} \sum_j \pi_j \gamma_j(\beta_1) \left[V(\bar{q}) - \beta_1 \bar{q} \right] + \mu(\beta_1) \left[\bar{q} - \bar{q}^*(\beta_2) \right].$$

It follows that $\bar{q}^*(\beta_n) < \hat{q}(\beta_n)$. For $i \in \{2, \ldots, n-1\}$, suppose $\bar{q}^*(\beta_{i+1}) < \hat{q}(\beta_{i+1})$. If $\mu(\beta_i) = 0$, then $\bar{q}^*(\beta_i) < \hat{q}(\beta_i)$. If $\mu(\beta_i) > 0$, complementary slackness implies $\bar{q}^*(\beta_i) = \bar{q}^*(\beta_{i+1})$, so $\bar{q}^*(\beta_i) < \hat{q}(\beta_i)$ by the induction hypothesis. Finally, $\bar{q}^*(\beta_2) < \hat{q}(\beta_2)$ implies $\mu(\beta_1) = 0$, so $\bar{q}^*(\beta_1) = \hat{q}(\beta_1)$. $\qquad\square$

Turn now to the contracts that induce information acquisition. I first replace (4)–(7) by *sufficient* alternative conditions. Afterwards, I construct a contract that satisfies them and improves over the best contract that deters information acquisition.

By standard arguments, (4) is equivalent to

$$q_j(\beta_i) - q_j(\beta_{i+1}) \geq 0 \quad \forall j \in J; i < n \tag{A.3}$$

together with the monotonicity condition

$$U_j(\beta_i) - U_j(\beta_{i+1}) \in \left[(\beta_{i+1} - \beta_i) q_j(\beta_{i+1}), (\beta_{i+1} - \beta_i) q_j(\beta_i) \right] \quad \forall j \in J; i < n.$$

I replace (4) by (A.3) and

$$U_j(\beta_i) - U_j(\beta_{i+1}) = (\beta_{i+1} - \beta_i) q_j(\beta_{i+1}) \quad \forall j \in J; i < n. \tag{A.4}$$

The truthtelling condition (5) then reads

$$U_j(\beta_n) + \sum_{i=1}^{n-1} \Gamma_j(\beta_i)(\beta_{i+1} - \beta_i) q_j(\beta_i)$$

$$\geq U_l(\beta_n) + \sum_{i=1}^{n-1} \Gamma_j(\beta_i)(\beta_{i+1} - \beta_i) q_l(\beta_i) \quad \forall j, l \in J, \tag{A.5}$$

26

and the participation condition, (6),

$$U_j(\beta_n) + \sum_{i=1}^{n-1} \Gamma_j(\beta_i)(\beta_{i+1} - \beta_i)q_j(\beta_i) \geq 0 \quad \forall j \in J. \tag{A.6}$$

The information-acquisition condition, (7), is implied by (5) and (6) at $e = 0$.

Note that by (A.4), the transfers are for each posterior type j pinned down by that type's output schedule up to the constant $-U_j(\beta_n)$. I therefore regard contracts that induce information acquisition in this proof as combinations $(-U_1(\beta_n), \ldots, -U_m(\beta_n), \mathbf{q})$, the relevant conditions being (A.3), (A.5), and (A.6).

Consider the contract $(-U_1'(\beta_n), \ldots, -U_m'(\beta_n), \mathbf{q}')$, designed as follows. For all posterior types $j \in \{2, \ldots, m\}$,

$$(q_j'(\beta_i))_i = (\bar{q}^*(\beta_i))_i$$
$$-U_j'(\beta_n) = -\bar{U}^*(\beta_n),$$

that is, the agent produces according to the same output schedule as under the best contract that deters information acquisition, $(-\bar{U}^*(\beta_n), \bar{\mathbf{q}}^*)$, and also gets the same expected payoff. For the posterior type $j = 1$, on the other hand,

$$(q_1'(\beta_i))_i = (\hat{q}(\beta_i))_i$$
$$-U_1'(\beta_n) = -\bar{U}^*(\beta_n) + \sum_{i=1}^{n+1} \Gamma_1(\beta_i)(\beta_{i+1} - \beta_i) \left[\hat{q}(\beta_i) - \bar{q}^*(\beta_i)\right],$$

that is, the agent again gets the same expected payoff as with $(-\bar{U}^*(\beta_n), \bar{\mathbf{q}}^*)$, but produces efficiently.

Clearly, this contract satisfies (A.3). To check (A.5), note that for $j = 1$ the agent is indifferent, and thus willing to report truthfully. For all $j \in \{2, \ldots, m\}$, there is no incentive to deviate either:

$$U_1'(\beta_n) + \sum_{i=1}^{n-1} \Gamma_j(\beta_i)(\beta_{i+1} - \beta_i)q_1'(\beta_{i+1})$$

$$= \bar{U}^*(\beta_n) - \sum_{i=1}^{n+1} \Gamma_1(\beta_i)(\beta_{i+1} - \beta_i)\left[\hat{q}(\beta_i) - \bar{q}^*(\beta_i)\right] + \sum_{i=1}^{n-1} \Gamma_j(\beta_i)(\beta_{i+1} - \beta_i)\hat{q}(\beta_{i+1})$$

$$< \bar{U}^*(\beta_n) - \sum_{i=1}^{n+1} \Gamma_j(\beta_i)(\beta_{i+1} - \beta_i)\left[\hat{q}(\beta_i) - \bar{q}^*(\beta_i)\right] + \sum_{i=1}^{n-1} \Gamma_j(\beta_i)(\beta_{i+1} - \beta_i)\hat{q}(\beta_{i+1})$$

$$= U_j'(\beta_n) + \sum_{i=1}^{n-1} \Gamma_j(\beta_i)(\beta_{i+1} - \beta_i)q_j'(\beta_{i+1})$$

where the inequality follows from the first-order stochastic dominance ordering of the posteriors and Lemma A3. Finally, (A.6) holds by (A.5) and the fact that $(-\bar{U}^*(\beta_n), \mathbf{q}^*)$ satisfies (10).

By Lemma A3, $(-U_1'(\beta_n), \ldots, -U_m'(\beta_n), \mathbf{q}')$ generates a strictly larger expected surplus than the best contract that deters information acquisition, $(-\bar{U}^*(\beta_n), \bar{\mathbf{q}}^*)$. On the other hand, it provides the agent with the same expected payoff. This observation establishes $W(0) > \bar{W}(0)$ and thus concludes the proof.

Proof of Theorem 1

Follows directly from Lemmas 2 and 3.

Proof of Lemma 4

By standard arguments, (4) is equivalent to

$$U_j(\beta_1) - U_j(\beta_2) \in [(\beta_2 - \beta_1)q_j(\beta_2), (\beta_2 - \beta_1)q_j(\beta_1)] \quad \forall j = O, P \qquad (A.7)$$

together with the monotonicity condition (12). As thus $U_j(\beta_1) - U_j(\beta_2) \geq 0$, the first order stochastic dominance ranking of the posteriors (i.e., $\gamma_O(\beta_1) > \gamma_P(\beta_1)$) and (5) imply that (6) holds if and only if

$$U_P(\beta_1) + \gamma_P(\beta_1)\left[U_P(\beta_1) - U_P(\beta_2)\right] \geq 0, \qquad (A.8)$$

and that in (7), the double sum on the right hand side is nonnegative. By the latter fact, (7) holds if and only if

$$\begin{aligned}
&U_j(\beta_2) + \gamma_j(\beta_1)\left[U_j(\beta_1) - U_j(\beta_2)\right] - \frac{e}{\pi_j} \\
&\geq \quad U_l(\beta_2) + \gamma_j(\beta_1)\left[U_l(\beta_1) - U_l(\beta_2)\right] \quad \forall j, l \in \{O, P\}.
\end{aligned} \qquad (A.9)$$

Note that this condition implies (5). Thus, (4)–(7) hold if and only if (12) and (A.7)–(A.9) hold.

Next, I show that (A.7) can without loss of generality be replaced by the stronger conditions (15) and (16). Let $(\mathbf{t}', \mathbf{q}')$ be any contract that satisfies (12) and (A.7)–(A.9), and define

$$\epsilon_O := (\beta_2 - \beta_1)q_O'(\beta_1) - \left[U_O'(\beta_1) - U_O'(\beta_2)\right]$$

28

and

$$\epsilon_P := U_P'(\beta_1) - U_P'(\beta_2) - (\beta_2 - \beta_1)q_P'(\beta_2).$$

Consider the alternative contract $(\mathbf{t}'', \mathbf{q}'')$, which differs from $(\mathbf{t}', \mathbf{q}')$ only with respect to transfers, namely such that

$$t_O''(\beta_1) = t_O'(\beta_1) + \gamma_O(\beta_2)\epsilon_O \qquad t_O''(\beta_2) = t_O'(\beta_2) - \gamma_O(\beta_1)\epsilon_O$$

and

$$t_P''(\beta_1) = t_P'(\beta_1) - \gamma_P(\beta_2)\epsilon_P \qquad t_P''(\beta_2) = t_P'(\beta_2) + \gamma_P(\beta_1)\epsilon_P.$$

Given these transfers, the alternative contract satisfies (15) and (16). Moreover, conditional on the posterior type, it results in identical expected payoffs for both parties as the original contract. Note that the alternative contract therefore satisfies (A.8). As to (A.9), it holds that

$$
\begin{aligned}
& U_O''(\beta_2) + \gamma_O(\beta_1)\left[U_O''(\beta_1) - U_O''(\beta_2)\right] - \frac{e}{\pi_O} \\
= \ & U_O'(\beta_2) + \gamma_O(\beta_1)\left[U_O'(\beta_1) - U_O'(\beta_2)\right] - \frac{e}{\pi_O} \\
\geq \ & U_P'(\beta_2) + \gamma_O(\beta_1)\left[U_P'(\beta_1) - U_P'(\beta_2)\right] \\
\geq \ & U_P'(\beta_2) + \gamma_P(\beta_1)\epsilon_P + \gamma_O(\beta_1)\left[U_P'(\beta_1) - U_P'(\beta_2) - \epsilon_P\right] \\
= \ & U_P''(\beta_2) + \gamma_O(\beta_1)\left[U_P''(\beta_1) - U_P''(\beta_2)\right],
\end{aligned}
$$

where the first inequality follows from the fact that $(\mathbf{t}', \mathbf{q}')$ satisfies (A.9) and the second one from the first-order stochastic dominance ordering of the posteriors. An analogous comparison shows that the other inequality in (A.9) is met as well.

Given (15) and (16), condition (A.8) reads (13), and (A.9) is by standard arguments equivalent to

$$
\begin{aligned}
U_O(\beta_2) - U_P(\beta_2) \in \Big[&-\gamma_O(\beta_1)(\beta_2 - \beta_1)\left[q_O(\beta_1) - q_P(\beta_2)\right] + \frac{e}{\pi_O}, \\
&-\gamma_P(\beta_1)(\beta_2 - \beta_1)\left[q_O(\beta_1) - q_P(\beta_2)\right] - \frac{e}{\pi_P}\Big]
\end{aligned}
\tag{A.10}
$$

together with the monotonicity condition (14).

To conclude the proof, I show that (A.10) can without loss of generality be replaced by the stronger condition (17). Let $(\mathbf{t}^\dagger, \mathbf{q}^\dagger)$ be any contract that satisfies (12)–(14) as

well as (15), (16), and (A.10). Define

$$\eta := U_O^\dagger(\beta_2) - U_P^\dagger(\beta_2) + \gamma_O(\beta_1)(\beta_2 - \beta_1)\left[q_O^\dagger(\beta_1) - q_P^\dagger(\beta_2)\right] - \frac{e}{\pi_O}$$

Consider the alternative contract $(\mathbf{t}^\ddagger, \mathbf{q}^\ddagger)$, which differs from $(\mathbf{t}^\dagger, \mathbf{q}^\dagger)$ only with respect to transfers, namely such that for each cost type $i = 1, 2$

$$t_O^\ddagger(\beta_i) = t_O^\dagger(\beta_i) - \pi_P\eta \qquad t_P^\ddagger(\beta_i) = t_P^\dagger(\beta_i) + \pi_O\eta.$$

Given these transfers, the alternative contract satisfies (17). Moreover, conditional on the prior, it results in identical expected payoffs for both parties as the original contract. Finally, (13), (15), and (16) clearly hold as well.

Proof of Proposition 2

For $e = 0$, it holds that $q_O^*(\beta_1) = \hat{q}(\beta_1)$ and

$$q_P^*(\beta_2) = V'^{-1}\left(\beta_2 + \frac{\pi_O\left[\gamma_O(\beta_1) - \gamma_P(\beta_1)\right](\beta_2 - \beta_1)}{\pi_P\gamma_P(\beta_2)}\right).$$

Thus,

$$
\begin{aligned}
e^* :=& \pi_O\pi_P\left[\gamma_O(\beta_1) - \gamma_P(\beta_1)\right](\beta_2 - \beta_1) \\
&\times \quad \left[\hat{q}(\beta_1) - V'^{-1}\left(\beta_2 + \frac{\pi_O\left[\gamma_O(\beta_1) - \gamma_P(\beta_1)\right](\beta_2 - \beta_1)}{\pi_P\gamma_P(\beta_2)}\right)\right] > 0.
\end{aligned}
$$

Proof of Proposition 3

First omitting the index τ, note that

$$
\begin{aligned}
W(0) - \bar{W}(0) =& \ \pi_O\gamma_O(\beta_2)\left[V(q_O^*(\beta_2)) - \beta_2 q_O^*(\beta_2) - (V(\bar{q}^*(\beta_2)) - \beta_2\bar{q}^*(\beta_2))\right] \\
& + \pi_P\gamma_P(\beta_2)\left[V(q_P^*(\beta_2)) - \beta_2 q_P^*(\beta_2) - (V(\bar{q}^*(\beta_2)) - \beta_2\bar{q}^*(\beta_2))\right] \\
& - \pi_O\left[\gamma_O(\beta_1) - \gamma_P(\beta_1)\right](\beta_2 - \beta_1)\left[\bar{q}^*(\beta_2) - q_P^*(\beta_2)\right],
\end{aligned}
$$

where $q_O^*(\beta_2) = \hat{q}(\beta_2)$,

$$q_P^*(\beta_2) = V'^{-1}\left(\beta_2 + \frac{\pi_O\left[\gamma_O(\beta_1) - \gamma_P(\beta_1)\right](\beta_2 - \beta_1)}{\pi_P\gamma_P(\beta_2)}\right),$$

and

$$\bar{q}^*(\beta_2) = V'^{-1}\left(\beta_2 + \frac{\pi_O\left[\gamma_O(\beta_1) - \gamma_P(\beta_1)\right](\beta_2 - \beta_1)}{\pi_O\gamma_O(\beta_2) + \pi_P\gamma_P(\beta_2)}\right).$$

Now, as V' is continuous, one-to-one, and has an open domain, V'^{-1} is continuous by Brouwer's Domain Invariance Theorem (see, e.g., Munkres 2000, §62, p. 381–385). Since $\gamma_O(\beta_2) \to 0$, this implies that $\lim_\tau q^*_{P,\tau}(\beta_2) = \lim_\tau \bar{q}^*_\tau(\beta_2)$. The result then follows from standard limit laws for convergent sequences (e.g., Rudin 1976, Thm. 3.3, p. 49).

References

Battaglini, M. (2005), "Long-term contracting with markovian consumers." *American Economic Review*, 95, 637–658.

Battaglini, M. and R. Lamba (2014), "Optimal dynamic contracting: The first-order approach and beyond." mimeo.

Bergemann, D. and M. Pesendorfer (2007), "Information structures in optimal auctions." *Journal of Economic Theory*, 137, 580–609.

Bergemann, D. and J. Välimäki (2002), "Information acquisition and efficient mechanism design." *Econometrica*, 70, 1007–1033.

Blackwell, D. and M.A. Girshick (1954), *Theory of Games and Statistical Decisions*. John Wiley & Sons, New York.

Boleslavsky, R. and M. Said (2013), "Progressive screening: Long-term contracting with a privately known stochastic process." *Review of Economic Studies*, 80, 1–34.

Compte, O. and P. Jehiel (2008), "Gathering information before signing a contract: A screening perspective." *International Journal of Industrial Organization*, 26, 206–212.

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

Crémer, J. and F. Khalil (1992), "Gathering information before signing a contract." *American Economic Review*, 82, 566–578.

Crémer, J. and F. Khalil (1994), "Gathering information before the contract is offered: The case with two states of nature." *European Economic Review*, 38, 675–682.

Crémer, J., F. Khalil, and J.-C. Rochet (1998a), "Contracts and productive information gathering." *Games and Economic Behavior*, 25, 174–193.

Crémer, J., F. Khalil, and J.-C. Rochet (1998b), "Strategic information gathering before a contract is offered." *Journal of Economic Theory*, 81, 163–200.

de la Fuente, A. (2000), *Mathematical methods and models for economists*. Cambridge University Press, New York.

Esö, P. and B. Szentes (2007), "Optimal information disclosure in auctions and the handicap auction." *Review of Economic Studies*, 74, 705–731.

Krähmer, D. and R. Strausz (2011), "Optimal procurement contracts with pre-project planning." *Review of Economic Studies*, 78, 1015–1041.

Krähmer, D. and R. Strausz (2015), "Optimal sales contracts with withdrawal rights." *Review of Economics Studies*, 82, 762–790.

Laffont, J.-J. and D. Martimort (2002), *The theory of incentives: the principal-agent model*. Princeton University Press, Princeton.

Lewis, T.R. and D.E.M. Sappington (1997), "Information management in incentive problems." *Journal of Political Economy*, 105, 796–821.

Li, H. and X. Shi (2015), "Discriminatory information disclosure." mimeo.

Munkres, J.R. (2000), *Topology*, second edition. Prentice Hall, Upper Saddle River.

Myerson, R.B. (1986), "Multistage games with communication." *Econometrica*, 54, 323–358.

Pavan, A., I. Segal, and J. Toikka (2014), "Dynamic mechanism design: A Myersonian approach." *Econometrica*, 82, 601–653.

Rudin, W. (1976), *Principles of Mathematical Analysis*, third edition. McGraw-Hill, New York.

Shi, X. (2012), "Optimal auctions with information acquisition." *Games and Economic Behavior*, 74, 666–686.

Szalay, D. (2009), "Contracts with endogenous information." *Games and Economic Behavior*, 65, 586–625.