# MECHANISM DESIGN BY AN INFORMED PRINCIPAL: THE QUASI-LINEAR PRIVATE-VALUES CASE

TYMOFIY MYLOVANOV AND THOMAS TRÖGER

ABSTRACT. In environments with independent private values and transferable utility, we provide a characterization of a mechanism selected by a privately informed principal. Our characterization implies that a privately informed principal will implement a mechanism that is ex-ante optimal for her. Furthermore, the characterization indicates conditions under which the asymmetry of information about the principal's preferences a ects mechanism design. As an application of our characterization result, we consider a bilateral exchange environment. We show that if the property rights over the good are dispersed among the parties, the principal whose valuation is uncertain will implement an allocation in which she is almost surely strictly better o than if her type is commonly known. The optimal mechanism is a combination of a participation fee for the agent, a buyout option for the principal, and a resale stage with posted prices and, hence, is a generalization of the posted price that would optimal if the principal's valuation were commonly known.

## 1. INTRODUCTION

The optimal design of contracts and institutions in the presence of privately informed market participants is central to economics, with applications including auctions, procurement, public good provision, organizational contract design, legislative bargaining, etc. In many of these models, transferability of utility serves as a convenient assumption that makes problems tractable and allows for a clean welfare analysis.

A restriction in much of this theory is that a contract or a mechanism is either designed by a third party, e.g., a benevolent planner, or is proposed by a party who has no private information. As such, the theory is not applicable to a large set of environments in which contracts or institutions arise endogenously as a choice of privately informed agents such as in, e.g., collusion, resale, contract renegotiation, bargaining over arbitration procedures, design of international agreements, etc.

Furthermore, the assumption that the designer does not have private information, even when it is a benevolent planner, is a tractability-driven simpli cation. As such, it is useful to understand under what conditions this assumption is *with* loss of generality, whether uncertainty about the designer's information advances or hinders the design objectives, and how the qualitative structure of optimal institutions can be a ected by this uncertainty.

---

*Date*: June 11, 2012.

Mylovanov: University of Pennsylvania, Department of Economics, 3718 Locust Walk, Philadelphia, PA 19104, USA, mylovanov@gmail.com. Troeger: University of Mannheim, Department of Economics, L7, 3-5 68131 Mannheim, Germany, troeger@uni-mannheim.de. We are deeply grateful to Simon Board, Yeon-Koo Che, Marina Halac, Benny Moldovanu, Roger Myerson, Rakesh Vohra, and Charles Zheng as well as the seminar audiences at Northwestern University, University of Pennsylvania, and University of Bonn for the very helpful comments. We gratefully acknowledge nancial support from the National Science Foundation, grant 1024683, and from the German Science Foundation (DFG) through SFB/TR 15 \Governance and the E ciency of Economic Systems."

The problem of mechanism design by a privately informed party was first considered by Myerson (1983), Maskin and Tirole (1990), and Maskin and Tirole (1992). They observe that the informed principal's preferences over mechanisms might depend on her information. The mechanism proposal may signal the principal's information to the other participants so that the value of the mechanism is determined endogenously in equilibrium.

In this paper, we provide a characterization of a mechanism selected by the informed principal in independent private value environments with transferable utility. In these environments, the agents' evaluation of the outcomes is not affected by the principal's information; the only inference that the agents can make from the principal's choice of a mechanism is about the principal's course of action within the mechanism. Thus, the assumption of independent private values allows focusing on the issue of signaling the principal's strategic position and abstracting from other signaling concerns.

We consider a Bayesian mechanism selection game in which the principal and the agents have private information about their payoff-type. The beliefs about the distribution of the payoff types are common knowledge. The principal can propose a mechanism; an outcome of a mechanism is a public action and a set of transfers among the parties. If the agents unanimously accept the proposal, the mechanism is played. Otherwise, an exogenously specified disagreement outcome takes place. The parties have quasilinear payoff functions over the public action and transfers.

Our main result is the following characterization of an allocation (mechanism) $\rho$ implemented by the informed principal. Let $t_0 \in T_0$ denote the type of the principal and $U_0^\rho(t_0)$ denote the expected payoff of principal type $t_0$ in allocation $\rho$. Let $p_0$ be the prior belief about the principal type. An allocation $\rho$ is implemented by the principal if the vector of principal type payoffs $U_0^\rho(t_0)$ satisfies

$$\eta(q_0) \;\leq\; \int_{T_0} U_0^\rho(t_0)\mathrm{d}q_0(t_0) \quad \text{for all } q_0 \text{ absolutely continuous rel. to } p_0, \tag{1}$$

where $\eta(q_0)$ is the principal's ex-ante optimal payoff given beliefs $q_0$ about the principal, i.e., the maximal expected payoff on the set of interim incentive compatible and individually rational mechanisms, where the expected payoffs and incentive compatibility and individual rationality constraints are defined with respect to $q_0$ and prior beliefs about the agents' types.

This is a rather restrictive condition that requires the principal's expected payoff in the allocation corresponding to $\rho$, when weighed according to $q_0$, to be not less than the total expected surplus available to the principal if $q_0$ reflects the agents' beliefs about the principal, and this condition must hold for all $q_0$ absolutely continuous with respect to prior beliefs about the principal $p_0$. Proposition 1 shows that an allocation satisfying this condition exists essentially in any finite type space environment. In Section 5.3, we study an application with continuous type spaces in which such an allocation exists.

The mechanism satisfying (1) is strongly-neologism-proof; this solution concept was developed by (Mylovanov and Tröger forthcoming), building on Maskin and Tirole (1990, 1992). An allocation $\rho$ is strongly-neologism-proof if there does not exist another allocation $\rho'$ that Pareto dominates $\rho$, in terms of the principal type payoffs, and that this allocation is incentive compatible and individually rational given any belief that puts the entire mass on principal types who are at least as well off in $\rho'$.

Our main result is, then, that in environments with transferable utility an allocation is strongly-neologism-proof if and only if it satisfies (1) (Proposition 2). The sufficiency of this characterization follows directly by definition of strong neologism proofness. The

substantive part of the result is the necessity: Instead of verifying the (non-)existence of a Pareto dominating allocation pointwise for each type of the principal, it is sufficient to compare the expected surplus in these allocations. The assumption of transferable utility is important. In Section 3, we provide an example of an environment without transfers in which a strongly-neologism-proof allocation does not satisfy (1). The proof of the result is subtle: the main difficulty is to demonstrate that if (1) is violated then there exists a dominating allocation that *satisfies* the incentive constraints of the principal.

If the type spaces are finite, the Pareto properties of a strongly-neologism-proof allocation $\rho$ ensure that it can be supported as an equilibrium outcome of the informed principal game.[1] If $\rho$ is offered on the equilibrium path, then for any arbitrary mechanism $\rho'$ there exist beliefs which make deviation to $\rho'$ unprofitable for all types of the principal. If this were not the case, we would be able to construct a dominating allocation, contradicting strong neologism proofness of $\rho$ (Mylovanov and Tröger forthcoming).

We do not know if any perfect Bayesian equilibrium outcome is strongly-neologism-proof.[2] The property of a strongly-neologism proof equilibrium outcome is that no information type of the principal can strictly gain from proposing an alternative allocation given any belief that puts probability 0 on types that would strictly lose.[3] Neologism-proof equilibria satisfy the intuitive criterion (Cho and Kreps 1987), are announcement-proof (Matthews, Okuno-Fujiwara, and Postlewaite 1991), undefeated (Mailath, Okuno-Fujiwara, and Postlewaite 1993), and immune to credible deviations (Esö and Schummer 2009).

Characterization (1) shows that any strongly neologism-proof allocation is ex-ante optimal for the principal, i.e., the principal's expected surplus is equal to $\eta(p_0)$. This result is most convenient in environments with continuous type spaces where the ex-ante optimal payoffs are unique: in such environments there is an essentially unique candidate for a strongly neologism-proof allocation. On the other hand, in finite-type examples it can be seen that not every ex-ante optimal allocation is strongly neologism proof.[4] Also, simple examples illustrate that ex-ante optimality fails in general without transferable utility (see Section 3).

Thus, the issue of the principal's potential information leakage through the choice of the mechanism imposes no cost on the principal in terms of the total surplus she realizes in equilibrium: Even though the principal's information is realized and the principal's preferences over how to distribute the available surplus among her types has changed, the principal nevertheless implements an allocation that would be optimal for her ex ante before she learns her type and no surplus is lost as long as the agents are still uncertain about the principal's type.

A further implication of this result is that in the environments in which the principal learns her type over time, the principal is indifferent about whether to write an ex-ante (long-term) contract or offer a (short-term) contract after her information is realized; this might explain

[1]C.f. footnote 11 for infinite type spaces.

[2]A proof may be achieved along the following lines. Consider an allocation that is not strongly neologism-proof. Construct a mechanism that virtually implements (Abreu and Matsushima 1992) a strongly neologism-proof allocation for every belief about the principal. Then at least one type of the principal has an incentive to deviate to this mechanism (cf. Maskin and Tirole, 1990, proof of Proposition 7). With multiple agents, one also needs to exclude coordinated rejection of off-equilibrium mechanisms. Coordinated rejection makes every disagreement-outcome-dominating allocation an equilibrium allocation.

[3]Neologism-proofness was introduced by Farrell (1993); it is also related to the concept of perfect sequential equilibrium by Grossman and Perry (1986).

[4]C.f. footnote 9.

why sometimes we do not observe complete long-term contracts. Finally, from a technical perspective the result connects the informed principal problem to the standard mechanism design approach that can be used to characterize ex-ante optimal mechanisms.

Characterization (1) also shows that the question of whether or not the principal benefits from the uncertainty about her information or, equivalently, offers a different allocation that she would when her information is commonly known ("best separable allocation") boils down to the question of whether or not a best separable allocation is ex-ante optimal for various beliefs about the principal's type. In Mylovanov and Tröger (2012), we study independent private value environments with linear payoff functions (Ledyard and Palfrey 2007) and use characterization (1) to derive conditions under which a best separable allocation is ex-ante optimal and, thus, the privacy of the principal's information does not affect the implemented allocation. In that paper, we show that the irrelevance result holds in linear environments if the parties' payoffs, net of disagreement payoffs, are monotonic in their type for each outcome.

Our characterization of strongly neologism-proof allocations can be also be useful in answering a number of other questions. For instance, it can be used to understand when restrictions on the class of mechanisms available to the principal, often made in applications, is without loss of generality. Characterization of strongly neologism-proof allocations can be helpful for constructing equilibria in dynamic environments, in which the parties bargain over the mechanisms or contract over allocations sequentially. It can also be applied to derive the default outcomes, such as assignment of property rights or legal regimes governing allocations in the absence of contracts, that would induce the principal to implement socially efficient allocations.

As a specific application of our characterization result, we consider a standard bilateral exchange environment a la Myerson and Satterthwaite (1983). We show that if the property rights over canappr-3cterisag-27(ersp)apnthee principal,ailluiono
1irthllo342(TRim(e)emen)28(t)-344(and)342(TRlo)-27(cation)-348(in)-361(TRw(hat)27(h)-303(TRs)-345(1s

and $t_1$ is the good's valuation. This outside option can be equivalently interpreted as the probability of getting the good if the parties do not agree on the mechanism or the value of the substitute good available to the agent. The optimal mechanism for the informed principal in this environment is the multistage mechanism described above provided his outside option is given by $\alpha_0 t_0$, where $\alpha_0 \geq 0$, $t_0$ is the principal's valuation of the good, and $\alpha_0 + \alpha_1 \leq 1$.

The intuition for the result is as follows. Let us normalize disagreement payoffs to be 0. In this bilateral trade environment with type-dependent outside option of the agent, if the principal's type is commonly known, then it is optimal for her to set a bid price (at which she is willing to buy) and an ask price (at which she is willing to sell). The best separable allocation consists of a collection of the optimal bid and ask prices. A low-type principal will set low prices. Hence, when dealing with a low-type principal, many agent-types will get the good, implying that the agent's payoff will be increasing over a relatively large range of her type space, implying that the agent's participation constraint will be binding for relatively low agent types. Vice versa, when dealing with a high-type principal, the agent's participation constraint will be binding for relatively high types. In summary, due to the failure of payoff monotonicity, the agent's participation constraint will be binding for different types, depending on which principal type the agent is dealing with. In an ex-ante optimal allocation, the agent's participation (and incentive) constraints are only required to hold in expectation over the principal's types. As a result, in the ex-ante optimal allocation the principal can extract more rents than in the best separable allocation. The ex-ante optimal allocation is strongly neologism-proof. In the multi-stage mechanism implementing the ex-ante optimal allocation, at the moment of accepting the mechanism and paying the participation fee, the agent is kept in the dark about the principal's type and is uncertain whether the principal will exercise her buy-out option. The agent's participation constraint can be violated conditional on a particular type of the principal, but is satisfied in expectation.

## 2. Literature review

*Early literature.* Myerson (1983) and Maskin and Tirole (1990, 1992) were the first to consider the problem of mechanism selection by an informed principal. The analysis in Myerson (1983) and Maskin and Tirole (1990) applies to independent private value environments; while they allow for quasilinear preferences, the papers focus on environments with non-transferable utility. Myerson's elegant analysis is based on an axiomatic approach; the paper introduces a neutral optimal solution, provides its algebraic characterization and, in particular, shows that it is a sequential equilibrium of an informed principal game. Myerson's results apply to general environments with finite outcome and type spaces. Maskin and Tirole characterize a perfect Bayesian equilibrium outcome of the informed principal game in a class of single-agent environments with two possible types of the agent under particular structural assumptions about the outcome space and the players' payoff functions.

*Mylovanov and Tröger (forthcoming).* The concept of a strongly-neologism-proof allocation was introduced in Mylovanov and Tröger (forthcoming) for private value environments with finite type spaces and general preferences. In this paper, the definition is extended to continuous type spaces. The existence result for environments with transferable utility and finite type spaces in Proposition 1 in this paper is based on the existence result in Mylovanov and Tröger (forthcoming) extended to environments with non-compact outcome spaces (the

result cannot be applied directly to quasilinear environments in this paper because utility transfers are not bounded.)

Strong-neologism-proofness can be seen as extending ideas of Maskin and Tirole (1990) to general settings.[6] Maskin and Tirole's study the concept of a \strong unconstrained Pareto optimum" (SUPO), which can be recast in terms of Farrell's neologism-proofness (Maskin and Tirole 1990, Mylovanov and Tröger forthcoming). In more general settings, SUPO may not exist. By contrast, a strongly neologism-proof equilibrium essentially always exists.

*Quasilinear private values environments.* A few papers have studied independent private value environments with transferable utility. The recurrent theme in the literature is that the principal's private information does not a ect the choice of the mechanism. In particular, Maskin and Tirole (1990) show the informed principal will o er the same mechanism as when her information is commonly known in their environment if utility is transferable.[7] Similar results have subsequently been obtained for other private value environments with transferable utility (Tan 1996, Yilankaya 1999, Balestrieri 2008, Skreta 2009).

There are several independent private value environments in which this irrelevance result does not hold. In an environment close to Maskin and Tirole's, Fleckinger (2007) shows that a privately informed principal can extract the entire surplus while this is not feasible for the principal whose valuation is commonly known to the agents. The uncertainty about the principal's preferences a ect the mechanism selected by the principal the leading example in Mylovanov and Tröger (forthcoming) and Mylovanov and Tröger (2008), as well as in in the bilateral exchange environment with type-dependent outside option of the agent considered in Section 5.3 in this paper. In Mylovanov and Tröger (2012), we apply results in this paper to show that in any linear environment with monotonic payo s and disagreement payo s normalized to be 0, any best separable allocation is strongly neologism-proof.

*Recent literature.* A number of recent papers study the informed principal problem in other environments. In environments with correlated types and multiple agents, Severinov (2008) provides a construction that allows the informed principal to extract the entire surplus. In an environment with correlated types and a single agent, Cella (2008) shows that the principal bene ts from privacy of her information. Skreta (2009) focuses on the optimal disclosure policy for the principal who has a private signal correlated with the valuations of the agents. Balkenborg and Makris (2010) look at common value environments and provide a novel characterization of a solution to the informed principal problem. Izmalkov and Balestrieri (2012) study the problem of the informed principal in an environment with horizontally di erentiated goods, where the principal is privately informed about the characteristic of the good. Halac (2012) considers optimal relational contracts in a repeated setting where the principal has persistent private information about her outside option.

Finally, there exists a separate literature that studies the informed-principal problem in moral-hazard environments, rather than in adverse-selection environments considered here (see, for example, Beaudry (1994), Jost (1996), Bond and Gresik (1997), Mezzetti and Tsoulouhas (2000), Chade and Silvers (2002), and Kaya (2010)).

---

[6] Quesada (2010) provides conditions for equilibrium allocations in Maskin and Tirole (1990) to be deterministic and shows that their characterization continues to hold in a less restrictive environment.

[7] The main focus of Maskin and Tirole (1990) is on non-transferable utility, in which case the result is di erent.

*Application: bilateral exchange.* In Section 5.3, we apply our characterization of strongly-neologism-proof allocations to the bilateral trade environment in which a single unit is traded (Myerson and Satterthwaite 1983). We assume that one of the parties, $i = 0$, is a principal proposing a mechanism and that the other party decides whether to accept it. If the proposal is rejected, each party's payoff is $\alpha_i t_i$, where $\alpha_i \in (0,1), \sum \alpha_i \leq 1$, and $t_i$ is the party's valuation of the good. If $\alpha_i$ capture the property rights on the share of the good and $\sum \alpha_i = 1$, the payoff environment is formally equivalent to a partnership dissolution problem (Cramton, Gibbons, and Klemperer 1987). The disagreement outcome could be interpreted as a lottery determining who gets the good or as an outcome of bargaining with another player or over a substitute good.

The informed principal problem is well understood in the environments with extreme allocation of property rights $(\alpha_0, \alpha_1) \in \{(0,1),(1,0)\}$. There, the principal is either a buyer or a seller and the informed principal implements a collection of posted prices, conditional on her valuation (Yilankaya 1999). Posted prices also maximize the ex-ante expected payoff of the principal and are optimal if the principal's value is commonly known (Riley and Zeckhauser 1983, Williams 1987). As we show in Section 5.3, the optimal mechanism for the informed principal is not a posted price and the principal strictly benefits from keeping her valuation private if the property rights are not extreme.

Cramton, Gibbons, and Klemperer (1987) have characterized conditions under which there exists an ex-post efficient allocation in the environment with $\sum \alpha_i = 1$.[8] The informed principal, however, will not implement the ex-post efficient allocation because she can extract additional surplus from the agent by distorting the ex-post efficient allocation.

## 3. Example

In this section, we provide an example that illustrates the importance of transferable utility, the conflict of preference between different principal types, and why concealing the principal's type might increase the total surplus available to principal types.

There are a principal (player 0) and an agent (player 1). The parties can choose one of three actions in $A = \{a^H, a^D, a'\}$ or a disagreement outcome $z_0$. Each player has two payoff types, $t_0 \in \{H, D\}$ and $t_1 \in \{h, d\}$; the players' payoffs are depicted in Table 1. The players get positive utility from actions $a^H$ and $a^D$ if the action matches the type and negative utility otherwise. The disagreement payoff is normalized to 0. In addition, action $a'$ gives principal type $t_0 = H$ a higher utility than any other action. The other principal type gets negative utility from this action, while the agent slightly prefers this action to disagreement.

| | $H$ | $D$ |
|---|---|---|
| $a^H$ | 1 | -1 |
| $a^D$ | -1 | 1 |
| $a'$ | $1 + \epsilon$ | -1 |
| $z_0$ | 0 | 0 |

| | $h$ | $d$ |
|---|---|---|
| $a^H$ | 1 | $-y$ |
| $a^D$ | $-y$ | 1 |
| $a'$ | $\epsilon$ | $\epsilon$ |
| $z_0$ | 0 | 0 |

TABLE 1. Players' payoffs; $y, \epsilon > 0$.

[8] For the partnership dissolution problem in the environments with interdependent values see, for example, Fieseler, Kittsteiner, and Moldovanu (2003) and Jehiel and Pauzner (2006).

All type profiles $(t_0, t_1)$ are equally likely. After the types are realized and privately observed, the principal proposes a contract, a finite game form with perfect recall, and the agent decides whether to accept it. If the agent rejects the proposal, the disagreement outcome is realized. The solution concept is Perfect Bayesian equilibrium.

If utility is not transferable, there exists a separating equilibrium in which each principal type offers her most preferred action, and this offer is accepted by the agent unless $t_1 = h$ and $t_0 = D$:

- $t_0 = H \to a'$, accepted;
- $t_0 = D \to a^D$, accepted iff $t_1 = d$.

Note, that this outcome would also be an equilibrium outcome if the principal's type were common knowledge. Furthermore, this is the *unique* equilibrium outcome. This is implied by the fact that principal type $H$ can always offer action $a'$ and this offer will be accepted by the agent.

Per Inscrutability Principle of Myerson (1983), this outcome can be implemented in a pooling equilibrium in which both types of the principal offer the following direct mechanism:

- $(H, h), (H, d) \to a'$;
- $(D, d) \to a^D$, $(D, h) \to z_0$.

Clearly, this mechanism is individually rational and incentive compatible for each player and will be accepted by the agent. This equilibrium can be supported by multiple beliefs; perhaps, the easiest construction is to assign agent's belief $t_0 = D$ for each alternative offer of the mechanism.

Nevertheless, if $y < 1$ and $\epsilon < 1/2$, this outcome does not maximize the ex-ante expected payoff of the principal and would not be implemented by the principal if she were to make her choice of a mechanism before learning her type. The outcome is dominated by the following individually rational and incentive compatible outcome:

- $H \to a^H$;
- $D \to a^D$.

(The principal's equilibrium expected payoff is $\frac{1}{2}(1 + \epsilon) + \frac{1}{4}$, whereas the principal's payoff given this outcome is 1.) Thus, in the environment without transferable utility, realization of private information of the principal changes the choice of the mechanism and destroys some of the expected surplus available to the principal. In equilibrium, optimal behavior of type $H$ requires implementing action $a'$, which, however, imposes a negative externality on the ability of the other type to implement action $a^D$ by limiting the amount of surplus left to the agent.

Let us now consider the environment in which the mechanism can execute utility transfers between the players. As a benchmark, assume that the principal's type is commonly known. In the optimal mechanism in this environment,

- principal type $H$ gets $\frac{3}{2} + \epsilon$ by choosing actions $a^H$ if $t_1 = h$ and $a'$ otherwise and charging the agent, correspondingly, 1 and $\epsilon$;
- whereas principal type $D$ gets 1 by choosing actions $a^D$ and charging the agent 1 if $t_1 = d$ and implementing $z_0$ otherwise.

In this mechanism, $H$ implements the allocation that maximizes the total surplus conditional on her type, and extracts the entire surplus from the agent. By contrast, type $D$ of the principal leaves some surplus on the table whenever $t_1 = H$ by choosing the disagreement outcome $z_0$ rather than action $a^D$ that would generate the total surplus of $1 - y > 0$. This

is optimal for type $D$ because of the agent's incentive constraints; implementing action $a^D$ would decrease the price that can be charged to the agent to $y$, giving the principal the payo of $1 - y < 1$.

This amount of surplus left on the table in this environment can be picked up by the principal if her type is not known to the agent. The following direct mechanism is incentive compatible, individually rational, maximizes the total surplus of the parties, and allows the principal to extract the entire surplus from the agent (we assume that $y + \epsilon < \frac{1}{2}$):

- $(H, h) \to a^H$, $(H, d) \to a'$, and the agent is charged his surplus from the action, 1 and $\epsilon$ respectively,
- $(D, d), (D, h) \to a^D$ and the agent is charged his surplus from the action, 1 and $-y$ respectively.

| | $H$ | $D$ |
|---|---|---|
| $a^H$ | $\frac{1}{2}2$ | |
| $a^D$ | | $2 - y$ |
| $a'$ | $\frac{1}{2}(1 + 2\epsilon)$ | |
| $z_0$ | 0 | 0 |

| | $h$ | $d$ |
|---|---|---|
| $a^H$ | 0 | |
| $a^D$ | 0 | 0 |
| $a'$ | | 0 |
| $z_0$ | 0 | 0 |

TABLE 2. Players' payo s in equilibrium with transferable utility

In this mechanism, the agent's incentive constraints are strictly satis ed conditional on principal type $H$; this allows the other type of the principal to extract higher surplus by violating the agent's incentive constraints conditional on her type. The mechanism is an optimal choice for the principal if she selects a mechanism before learning her type. It is also an equilibrium of the mechanism selection game by the privately informed principal.[9] As in the case without transfers, the easiest way to support this equilibrium is to assign belief $t_0 = D$ to each deviation.

Let *ex-post, interim*, and *ex-ante* denote correspondingly the mechanism selected by the principal before learning her type, after privately learning her type, and after her type becomes common knowledge. In the example in this section, we have shown that without transferable utility

- ex-post = interim $\neq$ ex-ante

while with transferable utility

- ex-post $\neq$ interim $\subseteq$ ex-ante.

The main result in this paper is to demonstrate that in general private value environments there exists an ex-ante optimal allocation that allocates the surplus across di erent types of the principal in such a manner that it prevents all possible deviations of all possible types of the principal. This is particularly striking in continuous types environments where often an ex-ante optimal allocation is unique and there is no freedom in allocating the surplus.

[9]There exists multiple incentive compatible and individually rational allocations that generate the same total expected payo for the principal but di erent principal type payo vectors. Some of them cannot be implemented in equilibrium by the informed principal because they assign a su ciently low payo to one of the principal types.

## 4. Model

4.1. **Environment.** Consider players $i = 0, \ldots, n$ who have to collectively choose from a space of basic outcomes

$$Z = A \times \mathbb{R}^n,$$

where the measurable space $A$ represents a set of veri able collective actions, and $\mathbb{R}^n$ is the set of vectors of agents' payments. For example, in an environment where the collective action is the allocation of a single unit of a private good among the players, $A = \{0, \ldots, n\}$, indicating who obtains the good.

Every player $i$ has a type $t_i \in T_i$ that captures her private information. A player's type space $T_i$ may be any compact metric space. The product of players' type spaces is denoted $\mathbf{T} = T_0 \times \cdots \times T_n$. The types $t_0, \ldots, t_n$ are realizations of stochastically independent Borel probability measures $p_0, \ldots, p_n$ with supp($p_i$) = $T_i$ for all $i$. The probability of any Borel set $B \subseteq T_i$ of player-$i$ types is denoted $p_i(B)$.

Player $i$'s payo function is denoted

$$u_i : Z \times T_i \to \mathbb{R}.$$

We consider private-value environments with quasi-linear payo functions,

$$\begin{aligned} u_0(a, \mathbf{x}, t_0) &= v_0(a, t_0) + x_1 + \cdots + x_n, \\ u_i(a, \mathbf{x}, t_i) &= v_i(a, t_i) - x_i, \end{aligned}$$

where $\mathbf{x} = (x_1, \ldots, x_n)$, and $v_0, \ldots, v_n$ are called *valuation functions*. We assume that the family of functions $(v_i(a, \cdot))_{a \in A}$ is equi-continuous for all $i$ (observe that this assumption is void if type spaces are nite).

The players' interaction results in an outcome that is a probability measure on the set of basic outcomes; the set of outcomes is denoted

$$\mathcal{Z} = \mathcal{A} \times \mathbb{R}^n,$$

where $\mathcal{A}$ denotes the set of probability measures on $A$, and $\mathbb{R}^n$ is the vector of the agents' expected payments.

If the players cannot agree on an outcome, some exogenously given disagreement outcome $\underline{z}$ obtains. The disagreement outcome $\underline{z} = (\underline{\alpha}, 0, \ldots, 0)$ for some (possibly random) collective action $\underline{\alpha} \in \mathcal{A}$. We normalize the valuation functions such that each player's expected valuation from the disagreement outcome equals 0, that is, $\int_A v_i(a, t_i) \mathrm{d}\underline{\alpha}(a) = 0$ for all $i$ and $t_i$.

A player's (expected) payo from any outcome $\zeta = (\alpha, \mathbf{x}) \in \mathcal{Z}$ is denoted

$$u_i(\zeta, t_i) = \int_A v_i(a, t_i) \mathrm{d}\alpha(a) - x_i \stackrel{\text{def}}{=} \int_Z u_i(z, t_i) \mathrm{d}\zeta(z),$$

where $x_0 = -x_1 - \cdots - x_n$.

An *allocation* is a complete type-dependent description of the result of the players' interaction; it is described by a map

$$\rho(\cdot) = (\alpha(\cdot), \mathbf{x}(\cdot)) : \mathbf{T} \to \mathcal{Z}$$

such that payments are uniformly bounded (that is, $\sup_{\mathbf{t} \in \mathbf{T}} ||\mathbf{x}(\mathbf{t})|| < \infty$, to guarantee integrability) and such that the appropriate measurability restrictions are satis ed (that is, for

any measurable set $B \subseteq A$, the map $\mathbf{T} \to \mathbb{R}$, $\mathbf{t} \mapsto \alpha(\mathbf{t})(B)$ is Borel measurable, and $\mathbf{x}(\cdot)$ is Borel measurable).

This class of quasilinear environments is prominent in the literature on mechanism design and includes bilateral exchange environments, single and multi-unit auctions and procurement, public good provision, non-linear pricing, regulation, taxation, franchise, as well as delegation, legislative bargaining, and assignment problems in environments with transferable utility.

4.2. **Strongly neologism-proof allocation.** We are interested in the problem of optimally selecting an allocation in the absence of a disinterested outsider. Rather, one of the players is designated as the proposer of the allocation. We will assume from now on that the proposer is player 0. We call her the principal; the other players are called agents.

Given the presence of private information, incentive and participation constraints will play a major role in our analysis. Expected payo s are computed with respect to the prior beliefs $p_1, \ldots, p_n$ about the agents' types. However, during the interaction the agents may update their belief about the principal's type, away from the prior $p_0$. Let $q_0$ denote a Borel probability measure on $T_0$ that represents the agents' belief about the principal's type. For most of our purposes it is enough work with a belief $q_0$ that is absolutely continuous relative to $p_0$.

Given an allocation $\rho$ and a belief $q_0$, the expected payo of type $t_i$ of player $i$ if she announces type $\hat{t}_i$ is denoted

$$U_i^{\rho,q_0}(\hat{t}_i, t_i) = \int_{\mathbf{T}_{-i}} u_i(\rho(\hat{t}_i, \mathbf{t}_{-i}), t_i) \mathrm{d}\mathbf{q}_{-i}(\mathbf{t}_{-i}),$$

where $\mathbf{q}_{-i}$ denotes the product measure obtained from deleting dimension $i$ of $q_0, p_1, \ldots, p_n$. The expected payo of type $t_i$ of player $i$ from allocation $\rho$ is

$$U_i^{\rho,q_0}(t_i) = U_i^{\rho,q_0}(t_i, t_i).$$

We will use the shortcut $U_0^{\rho}(t_0) = U_0^{\rho,q_0}(t_0)$, which is justi ed by the fact that the principal's expected payo is independent of $q_0$.

An allocation $\rho$ is called *$q_0$-feasible* if, for all players $i$, the $q_0$-incentive constraints (2) and the $q_0$-participation constraints (3) are satis ed,

$$\forall t_i, \hat{t}_i \in T_i: \quad U_i^{\rho,q_0}(t_i) \geq U_i^{\rho,q_0}(\hat{t}_i, t_i), \tag{2}$$

$$\forall t_i \in T_i: \quad U_i^{\rho,q_0}(t_i) \geq 0. \tag{3}$$

Given allocations $\rho$ and $\rho'$ and a belief $q_0$, we say that $\rho$ is *$q_0$-dominated* by $\rho'$ if $\rho'$ is $q_0$-feasible and

$$\begin{aligned} \forall t_0 \in \mathrm{supp}(q_0): \quad & U_0^{\rho'}(t_0) \geq U_0^{\rho}(t_0), \\ \exists B \subseteq \mathrm{supp}(q_0),\ q_0(B) > 0\ \forall t_0 \in B: \quad & U_0^{\rho'}(t_0) > U_0^{\rho}(t_0). \end{aligned}$$

The domination is *strict* if \>" holds for all $t_0 \in \mathrm{supp}(q_0)$.

Our notion of domination refers to the principal's payo . If some types of the principal have an incentive to deviate to a dominating allocation, and the dominating allocation is feasible given a belief that excludes all the principal-types who would su er from the deviation, then

we may not expect the original allocation to persist. This idea is behind our concept of a strongly neologism-proof allocation (Mylovanov and Tröger (forthcoming)).[10]

**Definition 1.** *An allocation $\rho$ is* strongly neologism-proof *if (i) $\rho$ is $p_0$-feasible and (ii) $\rho$ is not $q_0$-dominated for any belief $q_0$ that is absolutely continuous relative to $p_0$.*

In finite-type-space environments, any strong neologism-proofness is a perfect-Bayesian equilibrium in a mechanism-selection game in which any finite game form with perfect recall may be proposed as a mechanism (Mylovanov and Tröger (forthcoming)).[11]

4.3. **Existence.** In Mylovanov and Tröger (forthcoming) we prove the existence of a strongly neologism-proof allocation for finite-type-space environments under otherwise rather weak assumptions. However, the proof relies on the compactness of the outcome space, which is violated in quasi-linear environments because arbitrarily large payments are possible. Nevertheless, we can extend the existence proof to quasi-linear environments considered in this paper. We assume here that the space of collective actions $A$ is a compact metric space and make the technical assumption of separability that was introduced in Mylovanov and Tröger (forthcoming).

**Proposition 1.** *Suppose that the type spaces $T_0, \ldots, T_n$ are finite, that $A$ is a compact metric space, the valuation functions $v_0, \ldots, v_n$ are continuous, and separability holds. Then a strongly neologism-proof allocation exists.*

4.4. **Ex-ante optimal and best separable allocations.** A core point of our paper will be that strong neologism-proofness is closely related to the ex-ante optimality of an allocation. For any belief $q_0$, the problem of maximizing the principal's $q_0$-ex-ante expected payoff across all allocations that are $q_0$-feasible is

$$\max_{\rho \; q_0\text{-feasible}} \int_{T_0} U_0^{\rho}(t_0) \mathrm{d}q_0(t_0). \tag{4}$$

Let $\eta(q_0)$ denote the supremum value of the problem. In general, a maximum may fail to exist. This may be because arbitrarily high payoffs can be achieved ($\eta(q_0) = \infty$), or because the supremum cannot be achieved exactly.

**Definition 2.** *An allocation $\rho$ is* ex-ante optimal *if it solves problem (4) with $q_0 = p_0$.*

An important benchmark is the *best separable*[12] allocation — the allocation that the principal would optimally propose if her type were commonly known, that is, if the agents did have a point belief about the principal's type. Equivalently, a best separable allocation will be selected if the principal is restricted to offer a mechanism in which she is not a player herself.[13]

[10] The definition in Mylovanov and Tröger (TE) includes provisions about "happy types" who obtain the highest feasible payoff. In quasilinear environments, there are no happy types because payments can be arbitrarily high.

[11] In environments with infinite type spaces, there is no "natural" set of feasible mechanisms, nor is there an obvious choice for the definition of equilibrium. Any definition would have to deal with exempting deviations to mechanisms such that in the resulting continuation games no equilibrium exists, or in which the belief-equilibrium correspondence does not have the requisite continuity properties (c.f., Zheng (2002)). On the other hand, the idea behind our arguments in the earlier paper (Mylovanov and Tröger, forthcoming, proof of Proposition 1) is simple and appears robust.

[12] Maskin and Tirole (1990) use the term *full-information optimal allocation* instead.

[13] Zheng (2002) calls such mechanisms "transparent".

**Definition 3.** *An allocation $\rho$ is* best separable *if, for all point beliefs $q_0$, $\rho$ is $q_0$-feasible and $\rho$ is not $q_0$-dominated.*

Observe that the concept of a best separable allocation is entirely independent of the agent's prior belief $p_0$.

## 5. Results

5.1. **Characterization.** The main result in this section is a characterization of strong neologism-proofness in quasi-linear environments. We show that strong neologism-proofness requires, for all beliefs $q_0$ that are absolutely continuous with respect to the prior $p_0$, that the principal's highest possible $q_0$-ex-ante expected payoff cannot exceed the $q_0$-ex-ante expectation of the vector of her strongly neologism-proof payoffs.

**Proposition 2.** *A $p_0$-feasible allocation $\rho$ is strongly neologism-proof if and only if*

$$\eta(q_0) \leq \int_{T_0} U_0^{\rho}(t_0)\,dq_0(t_0) \quad \textit{for all } q_0 \textit{ absolutely continuous rel. to } p_0. \tag{5}$$

We prove the "if" part by showing the counterfactual, which is simple: an allocation that $q_0$-dominates $\rho$ also yields a strictly higher $q_0$-ex-ante-expected payoff, and $\eta(q_0)$ is, by definition, not smaller than this payoff.

To prove "only if", we again show the counterfactual. That is, we suppose that, given a strongly neologism-proof allocation $\rho$, there exists a belief $q_0$ such that (5) fails. By definition of $\eta(q_0)$, there exists a $q_0$-feasible allocation $\rho'$ with a strictly higher $q_0$-ex-ante-expected payoff than $\rho$. Starting with $\rho'$, by redistributing payments between principal-types we can construct an allocation $\rho''$ such that each principal-type is strictly better off than in $\rho$. This may lead, however, to a violation of a principal-type's incentive constraint in $\rho''$. The remaining, more difficult, part of the proof consists in resurrecting the principal's incentive constraints.

We find a belief $r_0$ and an allocation $\sigma$ that $r_0$-dominates $\rho$, thereby showing that $\rho$ is not strongly neologism-proof. Starting with the belief $q_0$ and the allocation $\rho''$, this can be imagined as being achieved by altering the allocation and the belief multiple times in a procedure that ends with $r_0$ and $\sigma$ after finitely many steps.

In environments with finite type spaces, the procedure can be imagined as follows. Suppose $\rho''$ violates the incentive constraint of some principal-type. We may restrict attention here to types in the support of $q_0$ (all other types may be assumed to announce whatever type is optimal among the types in the support of $q_0$). Alter $\rho''$ by giving the type with the violated constraint a different allocation: the average over what she had and what she is attracted to. Alter $q_0$ by adding to her previous probability the probability of the type that she was attracted to, and assign this type probability 0. From the viewpoint of the agents (i.e., in expectation over the principal's types), the new allocation together with the new belief is indistinguishable from the old one together with the old belief. Moreover, the new belief has a smaller support. Repeating this procedure leads to smaller and smaller supports, until incentive compatibility is satisfied.

The procedure is more complicated in environments with non-finite type spaces. First, we partition the principal's type space into a finite number of small cells such that when we replace in each cell the allocation by its average across the cell, then the new allocation $\rho'''$ is $q_0$-almost surely better than $\rho$. The crucial property of the new allocation is that, in the direct-mechanism interpretation, there exist only *finitely* many essentially different

announcements of principal-types. In summary, $\rho'''$ belongs to the set $\mathcal{E}$ of all allocations that (i) have this finiteness property, and (ii) are $r_0$-almost surely better for the principal than $\rho$, where (iii) $r_0$ is any belief such that the agents' $r_0$-incentive and participation constraints are satisfied (while the principal's constraints are not necessarily satisfied). We consider an allocation $\sigma^*$ in $\mathcal{E}$ that is minimal with respect to the finiteness property (that is, it is not possible to further reduce the number of essentially different principal-type announcements with violating (ii) or (iii)). Using the averaging idea from the finite-type world, we show that $\sigma^*$ satisfies the principal's incentive constraints $r_0$-almost surely. Hence, we can construct an $r_0$-feasible allocation $\sigma$ by altering $\sigma^*$ on an $r_0$-probability-0 set. Using continuity and the fact that property (ii) holds for $\sigma^*$, we conclude that $\rho$ is $r_0$-dominated by $\sigma$.

The complete proof is in the appendix.

5.2. **Applications of characterization.** Characterization in Proposition 2 has several direct implications. Setting $q_0 = p_0$, it follows that any strongly neologism-proof allocation is ex-ante optimal. This result is most convenient in environments with continuous type spaces where the ex-ante optimal payoffs are unique: in such environments there is an essentially unique candidate for a strongly neologism-proof allocation.

**Corollary 1.** *Any strongly neologism-proof allocation is ex-ante optimal.*

This result also implies that the issue of the principal's information leakage through the choice of the mechanism imposes no cost on the principal in terms of the total surplus she realizes in equilibrium: Different principal types, despite their conflict of preference about how to allocate the available surplus, coordinate on a mechanism that maximizes the ex-ante expected total surplus.

A further implication of this result is that in the environments in which the principal learns her type over time, the principal is indifferent about whether to write an ex-ante (long-term) contract or offer a (short-term) contract after her information is realized; this might explain why sometimes we do not observe complete long-term contracts. Finally, from a technical perspective the result connects the informed principal problem to the standard mechanism design approach that can be used to characterize ex-ante optimal mechanisms.

Characterization (1) also shows that the question of whether or not the principal benefits from the uncertainty about her information or, equivalently, offers an allocation that differs from what she would if her information were commonly known ("best separable allocation") boils down to the question of whether or not a best separable allocation is ex-ante optimal for various beliefs about the principal's type, as stated in the corollary below.

**Corollary 2.** *A best separable allocation is strongly neologism-proof if and only if it solves problem (4) for all $q_0$ that are absolutely continuous relative to $p_0$.*

*Proof.* "if" is immediate from Proposition 2. To see "only if", consider a best separable allocation $\rho$ that is strongly neologism-proof. As a best separable allocation, $\rho$ is $q_0$-feasible for *all* beliefs $q_0$. Hence, it solves problem (4) by Proposition 2. *QED*

In Mylovanov and Tröger (2012), we study independent private value environments with linear payoff functions (Ledyard and Palfrey 2007) and use this result to derive conditions under which any best separable allocation is ex-ante optimal and, thus, the privacy of the principal's information does not affect the implemented allocation. In that paper, we show

that the irrelevance result holds in linear environments if the parties' payo s are monotonic in their type for each outcome.

This corollary can also be used to understand when restrictions on the class of mechanisms available to the principal, often made in applications, is with loss of generality. For instance, a best separable allocation will be selected if the principal is restricted to o er a mechanism in which she is not a player herself; if a best separable allocation is dominated given prior or some other beliefs, it is not a strongly neologism-proof. Similarly, if an equilibrium allocation in a semi-separating or a pooling equilibrium of a game in which the principal is restricted in her choice of mechanisms is dominated given some beliefs, e.g., the beliefs that put the entire mass on the set of separating types, it is not a strongly neologism-proof.

Another corollary provides a su cient condition for an allocation to be strongly neologism-proof; it follows from the arguments in the proof of Proposition 2.

**Corollary 3.** *If a $p_0$-feasible allocation is not strongly neologism-proof, then it is* strictly *$q_0$-dominated for some belief $q_0$ that is absolutely continuous with respect to $p_0$.*

This corollary implies that the set of strongly neologism-proof principal-payo vectors is always closed and is helpful in proving the existence of strongly neologism-proof allocations in environments with nite types (Proposition 1).

Finally, Proposition 2 can be applied to design of the disagreement outcomes, e.g., the default allocation of property rights, legal regimes regulating the outcome in the absence of a contract, or default mechanisms, that would induce the principal to implement socially e cient allocations, or to design of collusion-proof mechanisms in the environments where one of the agents of the mechanism can o er a collusion contract.[14]

5.3. **Application: Linear Bilateral Bargaining Environment.** In this section, we use Proposition 2 to characterize a solution to the informed principal problem in a standard two party one good exchange environment a la Myerson and Satterthwaite (1983), where one of the parties is a principal who can propose a mechanism. We assume, however, that if the other party rejects the proposal, the parties' disagreement outcome is for each player to get the good with probability $\underline{\alpha}_i$, where $0 \leq \underline{\alpha}_0 + \underline{\alpha}_1 \leq 1$. This outside option can represent the default property rights, a probability of getting the good outside the current relationship between the players, or the value of an available substitute good. The special case of $\underline{\alpha}_0 + \underline{\alpha}_1 = 1$ corresponds to the model of partnership dissolution (the disagreement outcome may also include a side payment which we normalize to 0), as in Cramton, Gibbons, and Klemperer (1987).

It is known that if the allocation of the property rights over the good is extreme, $(\underline{\alpha}_0, \underline{\alpha}_1) \in \{(0,1),(1,0)\}$, the uncertainty of the principal's valuation plays no role and the informed principal will implement a best separable allocation by making, e.g., a posted price o er (Yilankaya 1999). This environment is a special case of a monotonic linear environment in which any best separable allocation is strongly neologism-proof (Mylovanov and Tr•ger 2012).

We, therefore, are going to assume that the payo s are non-monotonic in types by requiring $\underline{\alpha}_i \in (0,1)$. As it turns out, a best separable allocation is not ex-ante optimal in such

[14]See, e.g., La ont and Martimort (1997), Quesada (2005), Che and Kim (2006), Segal and Whinston (2011).

environments, and, hence, is not strongly neologism-proof. In other words, the privacy of the principal's information matters.

We determine an ex-ante optimal allocation and show that it is strongly neologism-proof by applying the characterization of Proposition 2. In the ex-ante optimal allocation, the principal is $p_0$-almost surely strictly better off than in the best separable allocation.

There are two players $i = 0, 1$. The set of verifiable collective actions is $A = \{0, 1, \underline{\alpha}\}$, indicating who gets assigned one unit of an indivisible good; action $\underline{\alpha}$ is the disagreement action. Hence, any probability distribution over collective actions is described by the probabilities $\alpha_i \in [0,1]$ that the good is assigned to player $i$. The disagreement outcome is that player $i$ obtains the good with probability $\underline{\alpha}_i$, where

$$0 < \underline{\alpha}_i < 1. \tag{6}$$

The type spaces are $T_0 = T_1 = [0,1]$. The environment is linear, with each player's valuation function

$$v_i(a, t_i) = s_i^a t_i,$$

where

$$s_0^a = \mathbf{1}_{a=0} + (\mathbf{1}_{a=\underline{\alpha}} - 1)\underline{\alpha}_0, \quad s_1^a = \mathbf{1}_{a=1} + (\mathbf{1}_{a=\underline{\alpha}} - 1)\underline{\alpha}_1.$$

That is, a player's type represents her valuation of the good. Payoffs are written such that each player's payoff from the disagreement outcome is normalized to 0. As a consequence, the payoff functions are non-monotonic.

The principal's type distribution $p_0$ is described by the continuous c.d.f. $F_0$, that is, we assume that there are no atoms in the type distribution. The agent's type distribution $p_1$ is described by the c.d.f. $F_1$ with strictly positive density $f_1$ and strictly increasing and continuous buyer- and seller virtual valuation functions $\psi^b(t_1) = \psi_1(t_1) = t_1 - \frac{1-F_1(t_1)}{f_1(t_1)}$ and $\psi^s(t_1) = t_1 + \frac{F_1(t_1)}{f_1(t_1)}$. In particular, the environment is regular.

We begin by determining an ex-ante optimal allocation for an arbitrary belief about the principal. Let $\alpha_i(t_0, t_1)$ be the probability that $a = i$ given a type profile.

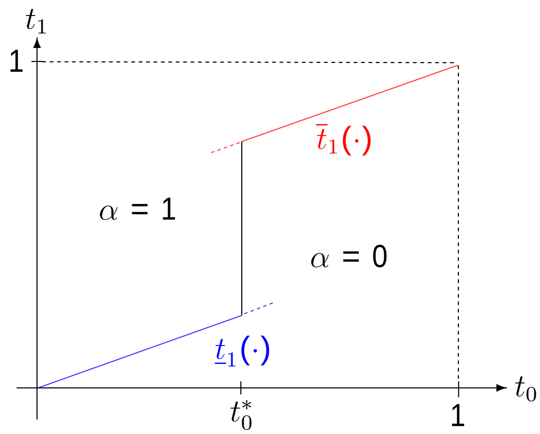
**Proposition 3.** *Consider the bilateral exchange environment in which player 1 gets the good with probability $\underline{\alpha}$ upon disagreement.*

*Let $q_0$ denote a belief that is absolutely continuous relative to $p_0$ and let $G_0$ denote the c.d.f. for $q_0$. Let $t_0^*$ be such that $\underline{\alpha}_1 = G_0(t_0^*)$*

*A solution to problem (4) is given by $\rho(\cdot) = (\alpha(\cdot), \mathbf{x}(\cdot))$, where*

$$(\alpha_0, \alpha_1)(t_0, t_1) = \begin{cases} (1,0) & \text{if } t_0 < t_0^*,\ \psi^s(t_1) < t_0, \\ (0,1) & \text{if } t_0 < t_0^*,\ \psi^s(t_1) > t_0, \\ (1,0) & \text{if } t_0 > t_0^*,\ \psi^b(t_1) < t_0, \\ (0,1) & \text{if } t_0 > t_0^*,\ \psi^b(t_1) > t_0, \end{cases}$$

*and $\mathbf{x}(\cdot)$ is chosen such that $\rho$ is $q_0$-feasible, and such that at least one agent-type's participation constraint is satisfied with equality.*




**Proposition 4.** *The solution given in Proposition 3 is strongly neologism-proof.*

We compare the strongly neologism-proof allocation to the best separable allocation. Using methods very similar to those used in the proof of Proposition 3, we obtain the following result.

**Proposition 5.** *Consider the bilateral exchange environment in which player 1 gets the good with probability $\underline{\alpha}$ upon disagreement.*

*The best separable allocation is given by $\rho(\cdot) = (\alpha(\cdot), \mathbf{x}(\cdot))$, where*

$$\alpha(t_0, t_1) = \begin{cases} 0 & \text{if } \psi^s(t_1) < t_0, \\ 1 & \text{if } \psi^b(t_1) > t_0, \\ \underline{\alpha} & \text{otherwise}, \end{cases}$$

*and $\mathbf{x}(\cdot)$ is chosen such that $\rho$ is $q_0$-feasible, and such that, for each $t_0$, at least one agent-type's participation constraint is satisfied with equality.*

This allocation can be implemented by using, for each type $t_0$, a bid price of $(\psi^s)^{-1}(t_0)$ and an ask price of $(\psi^b)^{-1}(t_0)$. In contract, the allocation described in Proposition 3 is implemented by a multi-stage mechanism involving a combination of a participation fee for the agent, a buyout option for the principal, and a resale stage with posted prices: In the first stage, the agent pays the participation fee and the good is tentatively allocated to the agent. In the second stage, the principal decides whether to exercise a buyout option, in which case the good becomes tentatively allocated to the principal; this option will be exercised by principal types with $t_0 > t_0^*$. In the third stage, given the tentative allocation of property rights, the principal makes a take-it-or-leave-it offer to the agent to sell or buy the good. Hence, the first two stages consolidate the originally dispersed property rights over the good and allocate them either to the principal or the agent determining whether the principal becomes the seller or the buyer in the third stage. This mechanism is a generalization of the bid and ask price mechanism that implements the best separable allocation as well as a generalization of a posted price mechanism that would be optimal in the environments with the extreme property rights allocation in which either the principal or the agent own the good (Williams 1987, Yilankaya 1999).

A direct calculation of the principal type's payoffs implies that almost every type is strictly better off in the strongly neologism-proof allocation and hence the uncertainty about the principal valuation strictly benefits the principal.

**Corollary 4.** *The principal is almost surely strictly better off in strongly neologism-proof allocation than in the best separable allocation.*

## 6. Appendix A

The proof or Proposition 1 relies on two lemmas. Given any allocation $\rho(\cdot) = (\alpha(\cdot), \mathbf{x}(\cdot))$ and any belief $q_0$ about the principal's type, the interim expected payment function of any player $i$ is denoted

$$\underline{x}_i^{\rho,q_0}(t_i) = \int_{\mathbf{T}_{-i}} x_i(t_i, \mathbf{t}_{-i}) \mathrm{d}\mathbf{q}_{-i}(\mathbf{t}_{-i}).$$

**Lemma 1.** *Suppose that $A$ is a compact metric space, and the valuation functions $v_0, \dots, v_n$ are continuous.*

*Then there exists a number $\lambda$ such that, for all beliefs $q_0$, in any $q_0$-feasible allocation, the absolute value of the interim expected payment of any type of any player is smaller than $\lambda$.*

*Proof.* Let $\overline{v}$ denote an upper bound for the absolute value of the valuation of any action for any type of any player.

By (3), each player's $q_0$-ex-ante expected payoff is bounded below by 0. On the other hand, the sum of the players' $q_0$-ex-ante expected payoffs is bounded above by $(n + 1)\overline{v}$ because payments cancel. Hence,

$$0 \leq \int_{T_i} U_i^{\rho,q_0}(t_i) \mathrm{d}q_i(t_i) \leq (n+1)\overline{v} \quad \text{for all } i,$$

where we define $q_i = p_i$ for all $i = 1, \dots, n$.

Turning to interim expected payoffs,

$$|U_i^{\rho,q_0}(t_i, t_i') - U_i^{\rho,q_0}(t_i, t_i)| \leq \max_{a \in A} |v_i(a, t_i') - v_i(a, t_i)| \leq 2\overline{v}. \tag{7}$$

Hence,

$$U_i^{\rho,q_0}(t_i) \leq U_i^{\rho,q_0}(t_i, t_i') + 2\overline{v} \overset{(2)}{\leq} U_i^{\rho,q_0}(t_i') + 2\overline{v}.$$

Thus,

$$U_i^{\rho,q_0}(t_i) \leq \int_{T_i} U_i^{\rho,q_0}(t_i') \mathrm{d}q_i(t_i') + 2\overline{v} \leq (n+3)\overline{v}.$$

Because any player's interim payment can differ from her interim payoff by at most $\overline{v}$, we can set $\lambda = (n + 4)\overline{v}$. This completes the proof.

With finite type spaces, both the space of payment schemes $\mathcal{L} = \mathbb{R}^{|\mathbf{T}|n}$ and the space of interim expected payment schemes $\underline{\mathcal{L}} = \mathbb{R}^{|T_0|+\cdots+|T_n|}$ are finite-dimensional vector spaces. Endow both spaces with the max-norm. We define the linear map

$$\phi^{q_0} : \mathcal{L} \to \underline{\mathcal{L}}, \ \mathbf{x}(\cdot) \mapsto (\underline{x}_0^{\rho,q_0}(\cdot), \dots, \underline{x}_n^{\rho,q_n}(\cdot)).$$

The following lemma says that there exists a number $\kappa$ such that any scheme of interim expected payments that can occur at all can also be obtained from a payment scheme that involves payments at most $\kappa$ times as large (in absolute value) as the largest interim expected payment of any type of any player.

**Lemma 2.** *Suppose that $T_0, \dots, T_n$ are finite. Consider any belief $q_0$. There exists a number $\kappa$ such that, for every $\underline{\mathbf{x}}(\cdot) \in \underline{\mathcal{L}}$, there exists $\mathbf{x}(\cdot) \in \mathcal{L}$ such that $\phi^{q_0}(\mathbf{x}(\cdot)) = \underline{\mathbf{x}}(\cdot)$ and $||\mathbf{x}(\cdot)|| \leq \kappa||\underline{\mathbf{x}}(\cdot)||$.*

*Proof.* The set $\phi^{q_0}(\mathcal{L})$ is a nite-dimensional vector space, hence a Banach space (with the norm induced by the max-norm in $\underline{\mathcal{L}}$), and $\phi^{q_0}$ maps onto that space. Hence, the claim is immediate from the open mapping theorem in functional analysis.

*Proof of Proposition 1.* Consider any sequence of payment bounds $(\lambda_l)$ such that $\lambda_l \to \infty$. From Mylovanov and Tröger (forthcoming), for each $l$, there exists an allocation $\rho_l$ that is strongly neologism-proof in the environment with payment bound $\lambda_l$. By Lemma 2 and Lemma 1 (with $q_0 = p_0$), w.l.o.g., all these allocations use payments that are bounded by the same number $\kappa\lambda$. Hence, the sequence of payment schemes in the sequence $\rho_l$ is bounded in the max-norm. Hence, there exists a convergent subsequence with limit $\rho^*$ (in the dimension of the probability measures on collective actions, the convergence is meant as a weak convergence).

As a limit of $p_0$-feasible allocations, $\rho^*$ is $p_0$-feasible. Suppose that $\rho^*$ is not strongly neologism-proof. By Corollary 3, $\rho^*$ is strictly $q_0$-dominated by some allocation $\rho'$, for some belief $q_0$.

If $l$ is su ciently large, then $\rho'$ (w.l.o.g. by Lemma 2 and Lemma 1) is a feasible allocation in the environment with payment bound $\lambda_l$.

Moreover, if $l$ is su ciently large, then $\rho_l$ is strictly $q_0$-dominated by $\rho'$ because $\rho_l$ approximates $\rho^*$. This contradicts the fact that $\rho_l$ is strongly neologism-proof in the environment with payment bound $\lambda_l$. *QED*

*Proof of Proposition 2.* \if" Suppose that $\rho$ is not strongly neologism-proof. Then there exists a belief $q_0$ and an allocation $\rho'$ that $q_0$-dominates $\rho$. We obtain a contradiction because

$$\eta(q_0) \geq \int_{T_0} U_0^{\rho'}(t_0)\mathrm{d}q_0(t_0) > \int_{T_0} U_0^{\rho}(t_0)\mathrm{d}q_0(t_0).$$

\only if". Consider a strongly neologism-proof allocation $\rho = (\alpha(\cdot), x_1(\cdot), \dots, x_n(\cdot))$.

Suppose there exists a belief $q_0$ such that (5) fails, that is

$$\eta(q_0) \quad > \quad \int_{T_0} U_0^{\rho}(t_0)\mathrm{d}q_0(t_0).$$

By de nition of $\eta(q_0)$, there exists a $q_0$-feasible allocation $\rho' = (\alpha'(\cdot), x_1'(\cdot), \dots, x_n'(\cdot))$ such that

$$\int_{T_0} U_0^{\rho'}(t_0)\mathrm{d}q_0(t_0) - \int_{T_0} U_0^{\rho}(t_0)\mathrm{d}q_0(t_0) \stackrel{\text{def}}{=} \epsilon > 0. \tag{8}$$

Let $\rho'' = (\alpha'(\cdot), x_1''(\cdots), \dots, x_n''(\cdot))$, where

$$\begin{aligned} x_1''(\mathbf{t}) &= x_1'(\mathbf{t}) - (U_0^{\rho}(t_0) - U_0^{\rho'}(t_0) + \epsilon). \\ x_i''(\mathbf{t}) &= x_i'(\mathbf{t}), \quad i = 2, \dots, n. \end{aligned} \tag{9}$$

Then $\rho''$ satisfies the $q_0$-incentive and participation constraints for all $i \notin \{0,1\}$. Also, $\rho''$ satisfies the $q_0$-incentive and participation constraints for $i = 1$ because

$$\begin{aligned} U_1^{\rho'',q_0}(\hat{t}_1, t_1) &= \int_{T_{-1}} \int_A v_1(a, t_1) \mathrm{d}\alpha'(\hat{t}_1, \mathbf{t}_{-1})(a) \mathrm{d}\mathbf{q}_{-1}(\mathbf{t}_{-1}) - \int_{T_{-1}} x_1''(\hat{t}_1, \mathbf{t}_{-1}) \mathrm{d}\mathbf{q}_{-1}(\mathbf{t}_{-1}) \\ &\overset{(9)}{=} U_1^{\rho',q_0}(\hat{t}_1, t_1) + \int_{T_0} (U_0^{\rho}(t_0) - U_0^{\rho'}(t_0)) \mathrm{d}q_0(t_0) + \epsilon \\ &\overset{(8)}{=} U_1^{\rho',q_0}(\hat{t}_1, t_1). \end{aligned}$$

For all $t_0 \in T_0$,

$$U_0^{\rho''}(t_0) - U_0^{\rho}(t_0) \overset{(9)}{=} U_0^{\rho'}(t_0) + (U_0^{\rho}(t_0) - U_0^{\rho'}(t_0) + \epsilon) - U_0^{\rho}(t_0) = \epsilon. \tag{10}$$

In other words, in $\rho''$ every type of the principal is — by the amount $\epsilon$ — better off than in $\rho$. In particular, $\rho''$ satisfies the participation constraints for $i = 0$. However, $\rho''$ may violate a incentive constraint for $i = 0$.

To complete the proof, we show that there exists a belief $r_0$ and an $r_0$-feasible allocation $\sigma$ such that, for all $t_0 \in \mathrm{supp}(r_0)$,

$$U_0^{\sigma}(t_0) \geq U_0^{\rho}(t_0) + \frac{1}{2}\epsilon. \tag{11}$$

It follows that $\rho$ is $r_0$-dominated by $\sigma$; this contradicts the strong neologism-proofness of $\rho$.

Because $v_0$ is equi-continuous and $T_0$ is compact, there exists $\delta > 0$ such that

$$\forall\ t_0, t_0' \in T_0,\ z \in Z: \quad \text{if } |t_0 - t_0'| < \delta \text{ then } |u_0(z, t_0) - u_0(z, t_0')| < \frac{\epsilon}{8}. \tag{12}$$

Similarly, because $\rho$ is $p_0$-feasible, $U_0^{\rho}$ is uniformly continuous. Hence, there exists $\delta' > 0$ such that

$$\forall\ t_0, t_0' \in T_0: \quad \text{if } |t_0 - t_0'| < \delta' \text{ then } |U_0^{\rho}(t_0) - U_0^{\rho}(t_0')| < \frac{\epsilon}{8}. \tag{13}$$

By compactness of $T_0$, there exists a finite partition $\hat{D}_1, \ldots, \hat{D}_{\hat{k}}$ of $T_0$ such that $\mathrm{diam}(\hat{D}_k) < \min\{\delta, \delta'\}$ for all $k = 1, \ldots, \hat{k}$. By dropping any cell $\hat{D}_k$ with $q_0(\hat{D}_k) = 0$, we obtain a partition $D_1, \ldots, D_{\overline{k}}$ of some set $\hat{T}_0 \subseteq T_0$, where $q_0(\hat{T}_0) = 1$ and $q_0(D_k) > 0$ for all $k = 1, \ldots, \overline{k}$.

We construct an allocation $\rho''' = (\alpha'''(\cdot), \mathbf{x}'''(\cdot))$ from $\rho''$ as follows. Given any $\mathbf{t} \in \mathbf{T}$ with $t_0 \in D_k$ for some $k$, we define $\alpha'''(\mathbf{t})$, and $x_i'''(\cdot)$ ($i = 1, \ldots, n$) by taking the average over all types in $D_k$. That is,

$$\begin{aligned} \alpha'''(\mathbf{t})(B) &= \frac{1}{q_0(D_k)} \int_{D_k} \alpha'(t_0', t_{-0})(B) \mathrm{d}q_0(t_0') \quad \text{for all measurable sets } B \subseteq A, \\ x_i'''(\mathbf{t}) &= \frac{1}{q_0(D_k)} \int_{D_k} x_i''(t_0', t_{-0}) \mathrm{d}q_0(t_0'). \end{aligned}$$

Given any $t_0 \in T_0 \setminus \hat{T}_0$, let $\hat{t}_0 \in \hat{T}_0$ be an announcement that is optimal for $t_0$ among all announcements in $\hat{T}_0$ in the direct-mechanism interpretation of $\rho'''$; define $\rho'''(t_0, \mathbf{t}_{-0}) = \rho'''(\hat{t}_0, \mathbf{t}_{-0})$ for all $\mathbf{t}_{-0} \in \mathbf{T}_{-0}$. (By construction of $\rho'''$, there are at most $\overline{k}$ essentially different announcements, so that an optimal one exists.)

By Fubini's Theorem for transition probabilities, for all $k$ and $t_0 \in D_k$,[15]

$$u_0(\rho'''(\mathbf{t}), t_0) = \frac{1}{q_0(D_k)} \int_{D_k} u_0(\rho''(t_0', \mathbf{t}_{-0}), t_0) \mathrm{d}q_0(t_0'). \tag{14}$$

Hence, letting $\mathbf{p}$ denote the product measure of $p_1, \ldots, p_n$,

$$\begin{aligned}
U_0^{\rho'''}(t_0) &= \int_{\mathbf{T}_{-0}} u_0(\rho'''(\mathbf{t}), t_0) \mathrm{d}\mathbf{p}(t_{-0}) \\
&\overset{(14)}{=} \frac{1}{q_0(D_k)} \int_{D_k} \int_{\mathbf{T}_{-0}} u_0(\rho''(t_0', \mathbf{t}_{-0}), t_0) \mathrm{d}\mathbf{p}(t_{-0}) \mathrm{d}q_0(t_0') \\
&\overset{(12)}{>} \frac{1}{q_0(D_k)} \int_{D_k} \int_{\mathbf{T}_{-0}} (u_0(\rho''(t_0', \mathbf{t}_{-0}), t_0') - \frac{\epsilon}{8}) \mathrm{d}\mathbf{p}(t_{-0}) \mathrm{d}q_0(t_0') \\
&= \frac{1}{q_0(D_k)} \int_{D_k} (U_0^{\rho''}(t_0') - \frac{\epsilon}{8}) \mathrm{d}q_0(t_0') \\
&\overset{(10)}{=} \frac{1}{q_0(D_k)} \int_{D_k} (U_0^{\rho}(t_0') + \frac{7}{8}\epsilon) \mathrm{d}q_0(t_0') \\
&\overset{(13)}{>} \frac{1}{q_0(D_k)} \int_{D_k} (U_0^{\rho}(t_0) + \frac{3}{4}\epsilon) \mathrm{d}q_0(t_0') \\
&= U_0^{\rho}(t_0) + \frac{3}{4}\epsilon \quad \text{for all } t_0 \in \hat{T}_0.
\end{aligned}$$

Let $\mathcal{I}(q_0)$ denote the set of allocations that satisfy the agents' (but not necessarily the principal's) $q_0$-incentive and participation constraints.

We show that $\rho''' \in \mathcal{I}(q_0)$. To see this, consider any $i = 1, \ldots, n$ and $\hat{t}_i, t_i \in T_i$. Then

$$\begin{aligned}
U_i^{\rho''', q_0}(\hat{t}_i, t_i) &= \int_{T_{-0-i}} \int_{T_0} u_i(\rho'''(\hat{t}_i, t_{-i}), t_i) \mathrm{d}q_0(t_0) \mathrm{d}p_{-0-i}(t_{-0-i}) \\
&= \int_{T_{-0-i}} \sum_k \int_{D_k} u_i(\rho'''(\hat{t}_i, t_{-i}), t_i) \mathrm{d}q_0(t_0) \mathrm{d}p_{-0-i}(t_{-0-i}) \\
&= \int_{T_{-0-i}} \sum_k q_0(D_k) u_i(\rho'''(\hat{t}_i, t_{-i-0}, t_{0k}), t_i) \mathrm{d}p_{-0-i}(t_{-0-i}),
\end{aligned}$$

where we have selected any $t_{0k} \in D_k$ for all $k$. Applying Fubini's Theorem for transition probabilities, we conclude that

$$\begin{aligned}
U_i^{\rho''', q_0}(\hat{t}_i, t_i) &= \int_{T_{-0-i}} \sum_k \int_{D_k} u_i(\rho''(\hat{t}_i, t_{-i-0}, t_0'), t_i) \mathrm{d}q_0(t_0') \mathrm{d}p_{-0-i}(t_{-0-i}) \\
&= \int_{T_{-0-i}} \int_{T_0} u_i(\rho''(\hat{t}_i, t_{-i-0}, t_0'), t_i) \mathrm{d}q_0(t_0') \mathrm{d}p_{-0-i}(t_{-0-i}) \\
&= U_i^{\rho'', q_0}(\hat{t}_i, t_i).
\end{aligned}$$

Hence, $\rho''' \in \mathcal{I}(q_0)$ because $\rho'' \in \mathcal{I}(q_0)$.

Given $\rho'''$ and any $t_0 \in T_0$, let

$$D^{\rho'''}(t_0) = \{t_0' \in T_0 \mid \forall t_{-0} : \rho'''(t_0', t_{-0}) = \rho'''(t_0, t_{-0})\}.$$

[15] See, e.g., Bauer, Probability Theory, Ch. 36.

By construction, the set

$$\mathcal{D}^{\rho'''} = \{D^{\rho'''}(t_0) \mid t_0 \in T_0\}$$

is a finite partition of $T_0$ (with at most $\overline{k}$ cells).

In summary, $\rho''' \in \mathcal{E}$, where we define

$$\begin{aligned} \mathcal{E} = \{\sigma \mid\ & |\mathcal{D}^{\sigma}| < \infty, \\ & \exists r_0 :\ \sigma \in \mathcal{I}(r_0),\ \exists \hat{T}_0 :\ r_0(\hat{T}_0) = 1, \\ & \forall t_0 \in \hat{T}_0 : U_0^{\sigma}(t_0) - U_0^{\rho}(t_0) > \frac{\epsilon}{2}, \\ & \forall t_0 \in T_0 \setminus \hat{T}_0, t_0' \in T_0 :\ U_0^{\sigma}(t_0) \geq U_0^{\sigma}(t_0', t_0), \\ & \forall t_0 \in \hat{T}_0 :\ \hat{T}_0 \cap \arg\max_{t_0' \in T_0} U_0^{\sigma}(t_0', t_0) \neq \emptyset\}. \end{aligned}$$

Because $\mathcal{E} \neq \emptyset$, there exists $\sigma^* \in \mathcal{E}$ with minimal $|\mathcal{D}^{\sigma^*}|$. Let $r_0$ denote a corresponding belief and let $\hat{T}_0$ a corresponding probability-1 set.

Let $B^*$ denote the set of principal-types for which an incentive constraint is violated in $\sigma^*$. Then $B^* \subseteq \hat{T}_0$ because $\sigma^* \in \mathcal{E}$. We will show that $r_0(B^*) = 0$.

Suppose that $r_0(B^*) > 0$. We will show that this contradicts the minimality of $|\mathcal{D}^{\sigma^*}|$.

Because $|\mathcal{D}^{\sigma^*}| < \infty$, there exists $D' \in \mathcal{D}^{\sigma^*}$ such that $r_0(B^* \cap D') > 0$.

By violation of the incentive constraint, there exists $D'' \in \mathcal{D}^{\sigma^*} \setminus \{D'\}$ such that

$$r_0(B'') > 0, \quad \text{where } B'' = \{t_0 \in B^* \cap D' \mid U_0^{\sigma^*}(\hat{t}_0, t_0) > U_0^{\sigma^*}(t_0) \text{ if } \hat{t}_0 \in D''\}.$$

We construct a new belief $r_0'$ by

$$r_0'(B) = r_0(B \cap B'')\frac{r_0(D' \cup D'')}{r_0(B'')} + r_0(B \setminus \{D' \cup D''\}) \quad \text{for any Borel set } B \subseteq T_0.$$

Clearly, $r_0'$ is absolutely continuous relative to $r_0$ (hence, relative to $p_0$). Also,

$$r_0'(\hat{T}_0') = 1, \quad \text{where } \hat{T}_0' = B'' \cup (\hat{T}_0 \setminus (D' \cup D'')). \tag{15}$$

We construct an allocation $\sigma' = (\beta(\cdot), \mathbf{y}(\cdot))$ from $\sigma^* = (\beta^*(\cdot), \mathbf{y}^*(\cdot))$ as follows.

Given any $\mathbf{t} \in \mathbf{T}$ with $t_0 \in B''$, we define $\beta(\mathbf{t})$, and $y_i(\cdot)$ ($i = 1, \ldots, n$) by taking the average over all types in $D' \cup D''$. That is, for all measurable sets $B \subseteq A$,

$$\begin{aligned} \beta(\mathbf{t})(B) &= \frac{r_0(D')}{r_0(D' \cup D'')}\beta^*(t_0', t_{-0})(B) + \frac{r_0(D'')}{r_0(D' \cup D'')}\beta^*(t_0'', t_{-0})(B), \\ y_i(\mathbf{t}) &= \frac{r_0(D')}{r_0(D' \cup D'')}y_i^*(t_0', t_{-0}) + \frac{r_0(D'')}{r_0(D' \cup D'')}y_i^*(t_0'', t_{-0}), \end{aligned}$$

where we have picked any $t_0' \in D'$ and $t_0'' \in D''$.

Given any $\mathbf{t} \in \mathbf{T}$ with $t_0 \in \hat{T}_0 \setminus (D' \cup D'')$, we define $\sigma'(\mathbf{t}) = \sigma^*(\mathbf{t})$. For all $\mathbf{t} \in \mathbf{T}$ with $t_0 \notin \hat{T}_0'$, define $\sigma'(\mathbf{t})$ by letting type $t_0$ announce, in the direct-mechanism interpretation of $\sigma'$, whatever type she finds optimal in $\hat{T}_0'$. Then

$$|\mathcal{D}^{\sigma'}| \leq |\mathcal{D}^{\sigma^*} \setminus \{D', D''\}| + 1 < |\mathcal{D}^{\sigma^*}|.$$

We will show now that $\sigma' \in \mathcal{E}$, yielding a contradiction to the minimality of $|\mathcal{D}^{\sigma^*}|$.

First we show that

$$\sigma' \in \mathcal{I}(r_0'). \tag{16}$$

Consider any $i = 1, \ldots, n$ and $\hat{t}_i, t_i \in T_i$. Then

$$\begin{aligned} U_i^{\sigma', r_0'}(\hat{t}_i, t_i) &= \int_{T_{-0-i}} \int_{\hat{T}_0} u_i(\sigma'(\hat{t}_i, t_{-i}), t_i) \mathrm{d} r_0'(t_0) \mathrm{d} p_{-0-i}(t_{-0-i}) \\ &= \int_{T_{-0-i}} \int_{\hat{T}_0 \setminus (D' \cup D'')} u_i(\sigma^*(\hat{t}_i, t_{-i}), t_i) \mathrm{d} r_0(t_0) \mathrm{d} p_{-0-i}(t_{-0-i}) \\ &\quad + \int_{T_{-0-i}} \int_{B''} u_i(\sigma'(\hat{t}_i, t_{-i}), t_i) \mathrm{d} r_0'(t_0) \mathrm{d} p_{-0-i}(t_{-0-i}). \end{aligned} \tag{17}$$

Picking any $t_0 \in B''$, and applying Fubini's theorem for transition probabilities,

$$\begin{aligned} \int_{B''} u_i(\sigma'(\hat{t}_i, t_{-i}), t_i) \mathrm{d} r_0'(t_0) &= u_i(\sigma'(\hat{t}_i, t_0, t_{-0-i}), t_i) r_0'(B'') \\ &= \left( \frac{r_0(D')}{r_0(D' \cup D'')} u_i(\sigma^*(\hat{t}_i, t_0', t_{-0-i}), t_i) + \frac{r_0(D'')}{r_0(D' \cup D'')} u_i(\sigma^*(\hat{t}_i, t_0'', t_{-0-i}), t_i) \right) r_0'(B'') \\ &= r_0(D') u_i(\sigma^*(\hat{t}_i, t_0', t_{-0-i}), t_i) + r_0(D'') u_i(\sigma^*(\hat{t}_i, t_0'', t_{-0-i}), t_i) \\ &= \int_{D' \cup D''} u_i(\sigma'(\hat{t}_i, t_{-i}), t_i) \mathrm{d} r_0(t_0). \end{aligned}$$

Plugging this into (17) yields

$$U_i^{\sigma', r_0'}(\hat{t}_i, t_i) = U_i^{\sigma^*, r_0}(\hat{t}_i, t_i).$$

This implies (16) because $\sigma^* \in \mathcal{I}(r_0)$.

Next we show that, for all $t_0 \in \hat{T}_0'$,

$$U_0^{\sigma'}(t_0) - U_0^{\rho}(t_0) > \frac{\epsilon}{2}. \tag{18}$$

First consider $t_0 \in \hat{T}_0 \setminus (D' \cup D'')$. Then $U_0^{\sigma'}(t_0) = U_0^{\sigma^*}(t_0)$, so (18) is immediate from $\sigma^* \in \mathcal{E}$ and from $\hat{T}_0' \subseteq \hat{T}_0$.

For all $t_0 \in B''$, (18) holds because

$$U_0^{\sigma'}(t_0) = \frac{r_0(D')}{r_0(D' \cup D'')} U_0^{\sigma^*}(t_0) + \frac{r_0(D'')}{r_0(D' \cup D'')} U_0^{\sigma^*}(t_0'', t_0) > U_0^{\sigma^*}(t_0).$$

This completes the proof that $\sigma' \in \mathcal{E}$, thereby contradicting the minimality of $|\mathcal{D}^{\sigma^*}|$.

We conclude that $r_0(B^*) = 0$.

Given any $\mathbf{t} \in \mathbf{T}$ with $t_0 \notin B^*$, we define $\sigma(\mathbf{t}) = \sigma^*(\mathbf{t})$. For all $\mathbf{t} \in \mathbf{T}$ with $t_0 \in B^*$, we define $\sigma(\mathbf{t})$ by letting type $t_0$ announce, in the direct-mechanism interpretation of $\sigma^*$, whatever type she finds optimal in $T_0 \setminus B^*$, or assign the disagreement outcome if $t_0$ prefers that.

By construction, the principal's incentive constraints are satisfied for $\sigma$. Also, the agents' $r_0$-incentive and participation constraints are satisfied because $\sigma(\mathbf{t})$ equals $\sigma^*(\mathbf{t})$ for a $r_0$-probability-1 set of principal-types, and because these constraints are satisfied for $\sigma^*$.

By construction, (11) holds for all $t_0 \in T_0 \setminus B^*$. By continuity of $U_0^{\sigma}(\cdot)$, (11) extends to all $t_0 \in \mathrm{supp}(r_0)$. In particular, the principal's participation constraint is satisfied for all types in supp($r_0$). By construction, the same holds for all types not in supp($r_0$). Hence, $\sigma$ is $r_0$-feasible. This completes the proof. $QED$

*Proof or Proposition 3.* By absolute continuity, $G_0$ is continuous. Consider any $q_0$-feasible allocation $\rho(\cdot) = (\alpha(\cdot), \mathrm{x}(\cdot))$. By a standard application of the envelope theorem, the principal's ex-ante expected payo is given by

$$\int_0^1 U_0^\rho(t_0)\mathrm{d}G_0(t_0) \quad = \quad -U_1^{\rho,q_0}(0) + \int_0^1 \int_0^1 \phi(t_0,t_1)\mathrm{d}F_1(t_1)\mathrm{d}G_0(t_0),$$

where

$$\phi(t_0,t_1) = \psi^b(t_1)\left(\alpha_1(t_0,t_1)(1-\underline{\alpha}_1) - \alpha_0(t_0,t_1)\underline{\alpha}_1\right) - t_0\left(\alpha_1(t_0,t_1)\underline{\alpha}_0 - \alpha_0(t_0,t_1)(1-\underline{\alpha}_0)\right)$$

Using the abbreviations $\underline{u} = U_1^{\rho,q_0}(0)$ and

$$y(t_0,t_1) = \alpha_1(t_0,t_1)(1-\underline{\alpha}_1) - \alpha_0(t_0,t_1)\underline{\alpha}_1,$$

the feasible set of problem (4) is equivalently given by the constraints

$$\begin{aligned}
&\underline{u} + \int_0^{t_1}\int_0^1 y(t_0,s)\mathrm{d}G_0(t_0)\mathrm{d}s \geq 0 \quad \forall t_1 \in T_1 \quad (PC),\\
&\int_0^1 y(t_0,t_1)\mathrm{d}G_0(t_0) \text{ is weakly increasing in } t_1 \quad (MC),\\
&0 \leq \alpha_i(t_0,t_1) \leq 1 \quad \forall t_0 \in T_0, t_1 \in T_1, \quad (CC)\\
&\text{(2) and (3) for } i = 0.
\end{aligned}$$

Here, (PC) are the agent's participation constraints written via the envelope formula, (MC) is the monotonicity constraint (see, e.g., Myerson, 1981), (CC) are the capacity constraints. The last group of constraints are the principal's incentive and participation constraints; we will solve the problem by relaxing these constraints and verifying them for the solution of the relaxed problem.

Let

$$= \quad \{(\underline{u}, y(\cdot,\cdot)) \in \mathbb{R} \times \mathbb{R}^{([0,1]^2)} \mid (MC),\ (CC)\}.$$

Observe that is convex.

De ne a function $M$ from into the space $\mathcal{C}([0,1])$ of continuous real-valued functions on $[0,1]$ as follows,

$$M(\underline{u},y)\ (t_1) \quad = \quad \underline{u} + \int_0^{t_1}\int_0^1 y(t_0,s)\mathrm{d}G_0(t_0)\mathrm{d}s \quad (t_1 \in [0,1]).$$

Then we can express our maximization problem as

$$\begin{aligned}
\max_{(\underline{u},y)\in\Omega} \quad & -\underline{u} + \int_0^1\int_0^1 \phi(t_0,t_1)\mathrm{d}F_1(t_1)\mathrm{d}G_0(t_0)\\
\text{s.t.} \quad & M(\underline{u},y) \geq 0.
\end{aligned}$$

Observe that the dual to $\mathcal{C}([0,1])$ is the space $NBV[0,1]$ of normalized functions of bounded variation on $[0,1]$ (see Luenberger, p. 115), and the positive cone in $NBV[0,1]$ that corresponds to the standard positive cone in $\mathcal{C}([0,1])$ is the set of weakly increasing right-continuous functions on $[0,1]$ (cf. Luenberger, p. 215).

From Luenberger (1969, p. 215{221), $(\underline{u}, y) = (\underline{u}^*, y^*)$ solves the original optimization problem if there exists a weakly increasing right-continuous function $z^*$ on $[0,1]$ such that $(\underline{u}, y) = (\underline{u}^*, y^*)$ solves[16]

$$\max_{(\underline{u},y)\in\Omega} \quad -\underline{u} + \int_0^1 \int_0^1 \phi(t_0,t_1)\mathrm{d}F_1(t_1)\mathrm{d}G_0(t_0)$$
$$+ \int_0^1 (\underline{u} + \int_0^{t_1} \int_0^1 y(t_0,s)\mathrm{d}G_0(t_0)\mathrm{d}s)\mathrm{d}z^*(t_1)$$

and

$$\int_0^1 (\underline{u}^* + \int_0^{t_1} \int_0^1 y^*(t_0,s)\mathrm{d}G_0(t_0)\mathrm{d}s)\mathrm{d}z^*(t_1) \quad = \quad 0. \quad (*)$$

Call the above problem the Lagrangian problem. Luenberger also states that the value reached at the maximum of the Lagrangian problem equals the value reached at the optimum of the original problem.

For any $t_0 \in (0,1)$, let

$$\underline{t}_1(t_0) = (\psi^s)^{-1}(t_0) \in (0,t_0), \quad \bar{t}_1(t_0) = (\psi^b)^{-1}(t_0) \in (t_0,1).$$

By the intermediate value theorem, there exists

$$t_1^* \quad \in \quad (\underline{t}_1(t_0^*), \bar{t}_1(t_0^*)) \tag{19}$$

such that

$$\int_{\underline{t}_1(t_0^*)}^{t_1^*} \underbrace{(\psi^s(t_1) - t_0^*)}_{\geq 0} \mathrm{d}F_1(t_1) \quad = \quad -\int_{t_1^*}^{\bar{t}_1(t_0^*)} \underbrace{(\psi^b(t_1) - t_0^*)}_{\leq 0} \mathrm{d}F_1(t_1).$$

We will determine a solution of the Lagrangian problem with

$$z^*(t_1) = \mathbf{1}_{t_1 \geq t_1^*} \quad \text{for all } t_1$$

Then $\underline{u}$ cancels out in the Lagrangian objective. The objective can be rewritten (by changing the order of integration), so that we have to

$$\max_{(\underline{u},y)\in\Omega} \quad -\int_0^1 \int_0^1 \left( \left( t_1 - \frac{z^*(t_1) - F_1(t_1)}{f_1(t_1)} \right) \underline{\alpha}_1 - t_0(1-\underline{\alpha}_0) \right) \alpha_0(t_0,t_1)\mathrm{d}F_1(t_1)\mathrm{d}G_0(t_0)$$
$$+ \quad \int_0^1 \int_0^1 \left( \left( t_1 - \frac{z^*(t_1) - F_1(t_1)}{f_1(t_1)} \right) (1-\underline{\alpha}_1) - t_0\underline{\alpha}_0 \right) \alpha_1(t_0,t_1)\mathrm{d}F_1(t_1)\mathrm{d}G_0(t_0). \quad (**)$$

(Here, $\underline{u}$ has vanished from the objective; it is chosen such that at least one agent's participation constraint is binding.)

It remains to solve (**) and show that the condition (*) is satis ed. The objective of (**) can be written as

$$- \quad (1-\underline{\alpha}_0) \int_0^1 \int_0^1 \left( \frac{\underline{\alpha}_1}{1-\underline{\alpha}_0} h(F_1(t_1)) - t_0 \right) \alpha_0(t_0,t_1)\mathrm{d}F_1(t_1)\mathrm{d}G_0(t_0)$$
$$+ \quad \underline{\alpha}_0 \int_0^1 \int_0^1 \left( \frac{1-\underline{\alpha}_1}{\underline{\alpha}_0} h(F_1(t_1)) - t_0 \right) \alpha_1(t_0,t_1)\mathrm{d}F_1(t_1)\mathrm{d}G_0(t_0),$$

[16] Our integration areas are always closed intervals.

where we use the abbreviation

$$\begin{aligned} h(q) &= F_1^{-1}(q) - \frac{z^*(F_1^{-1}(q)) - q}{f_1(F_1^{-1}(q))} \\ &= \begin{cases} \psi^s(F_1^{-1}(q)) & \text{if } q < F_1(t_1^*), \\ \psi^b(F_1^{-1}(q)) & \text{if } q \geq F_1(t_1^*). \end{cases} \end{aligned}$$

Following Myerson (1981, p. 68), define

$$H(q) = \int_0^q h(r)\mathrm{d}r,$$

let $G$ denote the convex hull of $H$, and define $g(q) = G'(q)$ (extended to $[0,1]$ by right-continuity) and $\overline{c}(t_1) = g(F_1(t_1))$.

Then Myerson's arguments (1981, p. 69-70), applied in expectation over $t_0$, show that a solution to problem (**) is given by

$$(\alpha_0^*, \alpha_1^*)(t_0, t_1) = \begin{cases} (0,1) & \text{if } t_0 \leq \overline{c}(t_1), \\ (1,0) & \text{if } t_0 > \overline{c}(t_1). \end{cases} \tag{20}$$

Using (19), we find

$$\int_0^1 y^*(t_0, t_1^*)\mathrm{d}G_0(t_0) = (1 - \underline{\alpha}_1)G_0(t_0^*) - \underline{\alpha}_1(1 - G_0(t_0^*)) = 0$$

Thus, from the envelope formula and because $t_1 \mapsto \int y^*(t_0, t_1)\mathrm{d}G_0(t_0)$ is a weakly increasing function, the agent's expected payoff $U_1(t_1)$ is minimized at $t_1 = t_1^*$, that is, $U_1(t_1^*) = 0$ if $\underline{u} = \underline{u}^*$ is chosen optimally. This implies (*).

Finally, note that $y^*(t_0, t_1)$ is decreasing in $t_0$, confirming incentive compatibility for the principal. The principal's participation constraints are easily verified. *QED*

Plugging the solution of the ex-ante optimization problem into the principal's objective as given by the formula in the proof, we obtain a formula for the ex-ante expected utility. Here we use the functions $\underline{t}_1 = (\psi^s)^{-1}$ and $\overline{t}_1 = (\psi^b)^{-1}$.

**Corollary 5.** *Let $q_0$ denote a belief that is absolutely continuous relative to $p_0$. Let $t_0^*$ be such that $G_0(t_0^*) = \underline{\alpha}_1$. Then*

$$\begin{aligned} \eta(q_0) = & \int_0^{t_0^*} t_0 F_1(\underline{t}_1(t_0))\, dG_0(t_0) - \int_0^{t_0^*}\int_0^{\underline{t}_1(t_0)} \psi^s(t_1)\, dF_1(t_1)\, dG_0(t_0) \\ & + \int_{t_0^*}^1 t_0 F_1(\overline{t}_1(t_0))\, dG_0(t_0) + \int_{t_0^*}^1 \int_{\overline{t}_1(t_0)}^1 \psi^b(t_1)\, dF_1(t_1)\, dG_0(t_0) \\ & - \underline{\alpha}_0 \int_0^1 t_0\, dG_0(t_0). \end{aligned} \tag{21}$$

*Proof of Proposition 4.* We show this by proving the condition given in Proposition 2.

It is convenient to rearrange the formula from Corollary 5 by first changing the order of integration, then using a transformation of variable, and finally using integration by parts.

Speci cally,

$$
\begin{aligned}
\int_0^{t_0^*}\int_0^{\underline{t}_1(t_0)} \psi^s(t_1)\mathrm{d}F_1(t_1)\mathrm{d}G_0(t_0) &= \int_0^{\underline{t}_1(t_0^*)}\int_{\psi^s(t_1)}^{t_0^*} \mathrm{d}G_0(t_0)\psi^s(t_1)\mathrm{d}F_1(t_1) \\
&= \int_0^{\underline{t}_1(t_0^*)} (\underline{\alpha}_1 - G_0(\psi^s(t_1)))\psi^s(t_1)f_1(t_1)\mathrm{d}t_1 \\
&\overset{\psi^s(t_1)=t_0,\ t_1=\underline{t}_1(t_0)}{=} \int_0^{t_0^*} (\underline{\alpha}_1 - G_0(t_0))t_0 f_1(\underline{t}_1(t_0))\mathrm{d}\underline{t}_1(t_0) \\
&= \int_0^{t_0^*} (\underline{\alpha}_1 - G_0(t_0))t_0 \mathrm{d}F_1(\underline{t}_1(t_0)) \\
&= -\int_0^{t_0^*} (\underline{\alpha}_1 - G_0(t_0) - f_0(t_0)t_0)F_1(\underline{t}_1(t_0))\mathrm{d}t_0 \\
&= -\underline{\alpha}_1\int_0^{t_0^*} F_1(\underline{t}_1(t_0))\mathrm{d}t_0 + \int_0^{t_0^*} t_0F_1(\underline{t}_1(t_0))\mathrm{d}G_0(t_0) \\
&\quad + \int_0^{t_0^*} F_1(\underline{t}_1(t_0))G_0(t_0)\mathrm{d}t_0. \qquad (22)
\end{aligned}
$$

An analogous computation yields

$$
\begin{aligned}
\int_{t_0^*}^1\int_{\bar{t}_1(t_0)}^1 \psi^b(t_1)\mathrm{d}F_1(t_1)\mathrm{d}G_0(t_0) &= \underline{\alpha}_1\int_{t_0^*}^1 F_1(\bar{t}_1(t_0))\mathrm{d}t_0 - \int_{t_0^*}^1 t_0F_1(\bar{t}_1(t_0))\mathrm{d}G_0(t_0) \\
&\quad - \int_{t_0^*}^1 F_1(\bar{t}_1(t_0))G_0(t_0)\mathrm{d}t_0 + (1-\underline{\alpha}_1). \qquad (23)
\end{aligned}
$$

Plugging (22) and (23) into (21) yields

$$
\begin{aligned}
\eta(G_0) &= \underline{\alpha}_1\int_0^{t_0^*} F_1(\underline{t}_1(t_0))\mathrm{d}t_0 - \int_0^{t_0^*} F_1(\underline{t}_1(t_0))G_0(t_0)\mathrm{d}t_0 \\
&\quad +\underline{\alpha}_1\int_{t_0^*}^1 F_1(\bar{t}_1(t_0))\mathrm{d}t_0 - \int_{t_0^*}^1 F_1(\bar{t}_1(t_0))G_0(t_0)\mathrm{d}t_0 + (1-\underline{\alpha}_1) \\
&\quad -\underline{\alpha}_0\int_0^1 t_0\mathrm{d}G_0(t_0). \qquad (24)
\end{aligned}
$$

There is an alternative way to express the principal's ex-ante expected utility,

$$
\eta(G_0) = \int_0^1 U_0(t_0)\mathrm{d}G_0(t_0),
$$

where $U_0(t_0)$ denotes the expected utility of a given type $t_0$.

Using the envelope formula, for all $t_0 < t_0^*$,

$$\begin{aligned} U_0(t_0) &= U_0(t_0^*) - \int_{t_0}^{t_0^*} \int_0^1 (\alpha_0(t_0, t_1) - \underline{\alpha}_0)\mathrm{d}F_1(t_1)\mathrm{d}s \\ &= U_0(t_0^*) - \int_{t_0}^{t_0^*} (F_1(\underline{t}_1(s)) - \underline{\alpha}_0)\mathrm{d}s \\ &= U_0(t_0^*) - \int_{t_0}^{t_0^*} F_1(\underline{t}_1(s))\mathrm{d}s + \underline{\alpha}_0(t_0^* - t_0). \end{aligned} \tag{25}$$

Similarly, for all $t_0 > t_0^*$,

$$\begin{aligned} U_0(t_0) &= U_0(t_0^*) + \int_{t_0^*}^{t_0} \int_0^1 (\alpha_0(t_0, t_1) - \underline{\alpha}_0)\mathrm{d}F_1(t_1)\mathrm{d}s \\ &= U_0(t_0^*) + \int_{t_0^*}^{t_0} (F_1(\bar{t}_1(s)) - \underline{\alpha}_0)\mathrm{d}s \\ &= U_0(t_0^*) + \int_{t_0^*}^{t_0} F_1(\bar{t}_1(s))\mathrm{d}s + \underline{\alpha}_0(t_0^* - t_0). \end{aligned} \tag{26}$$

Hence,

$$\begin{aligned} \eta(G_0) &= U_0(t_0^*) - \int_0^{t_0^*} \int_{t_0}^{t_0^*} F_1(\underline{t}_1(s))\mathrm{d}s\mathrm{d}G_0(t_0) \\ &\quad + \int_{t_0^*}^1 \int_{t_0^*}^{t_0} F_1(\bar{t}_1(s))\mathrm{d}s\mathrm{d}G_0(t_0) - \underline{\alpha}_0 \int_0^1 t_0\mathrm{d}G_0(t_0) + \underline{\alpha}_0 t_0^*. \end{aligned}$$

Using integration by parts,

$$\begin{aligned} \eta(G_0) &= U_0(t_0^*) - \int_0^{t_0^*} F_1(\underline{t}_1(t_0))G_0(t_0)\mathrm{d}t_0 \\ &\quad - \int_{t_0^*}^1 F_1(\bar{t}_1(t_0))G_0(t_0)\mathrm{d}t_0 + \int_{t_0^*}^1 F_1(\bar{t}_1(s))\mathrm{d}s \\ &\quad - \underline{\alpha}_0 \int_0^1 t_0\mathrm{d}G_0(t_0) + \underline{\alpha}_0 t_0^*. \end{aligned} \tag{27}$$

Comparing (27) to (24) yields that

$$U_0(t_0^*) = (1 - \underline{\alpha}_0) \int_0^{t_0^*} F_1(\underline{t}_1(t_0))\mathrm{d}t_0 - \underline{\alpha}_0 \int_{t_0^*}^1 F_1(\bar{t}_1(t_0))\mathrm{d}t_0 + \underline{\alpha}_0 - \underline{\alpha}_0 t_0^*.$$

Plugging this into (25), we find

$$\begin{aligned} U_0^*(t_0) &= (1 - \underline{\alpha}_0) \int_0^{t_0} F_1(\underline{t}_1(s))\mathrm{d}s - \underline{\alpha}_0 \int_{t_0}^{t_0^*} F_1(\underline{t}_1(s))\mathrm{d}s - \underline{\alpha}_0 \int_{t_0^*}^1 F_1(\bar{t}_1(s))\mathrm{d}s \\ &\quad + \underline{\alpha}_0 (1 - t_0) \quad \text{for all } t_0 < t_0^*. \end{aligned}$$

Similarly, using (26),

$$\begin{aligned} U_0^*(t_0) &= (1-\underline{\alpha}_0)\int_0^{t_0^*} F_1(\underline{t}_1(s))\mathrm{d}s + (1-\underline{\alpha}_0)\int_{t_0^*}^{t_0} F_1(\bar{t}_1(s))\mathrm{d}s - \underline{\alpha}_0\int_{t_0}^{1} F_1(\bar{t}_1(s))\mathrm{d}s \\ &\quad +\underline{\alpha}_0\,(1-t_0) \quad \text{for all } t_0 > t_0^*. \end{aligned}$$

Now we show that $U_0^*(\cdot)$ is strongly neologism-proof by verifying the condition given in Proposition 2, that is,

$$\eta(G_0) \;\le\; \int_0^1 U_0^*(t_0)\mathrm{d}G_0(t_0) \quad \text{for all } G_0. \tag{28}$$

Using (28) and (28), the claim (28) can be written as

$$\begin{aligned} \eta(G_0) \;\le\; & (1-\underline{\alpha}_0)\int_0^{t_0^*}\int_0^{t_0} F_1(\underline{t}_1(s))\mathrm{d}s\mathrm{d}G_0(t_0) - \underline{\alpha}_0\int_0^{t_0^*}\int_{t_0}^{t_0^*} F_1(\underline{t}_1(s))\mathrm{d}s\mathrm{d}G_0(t_0) \\ & -\underline{\alpha}_0\int_0^{t_0^*}\int_{t_0^*}^{1} F_1(\bar{t}_1(s))\mathrm{d}s\mathrm{d}G_0(t_0) + (1-\underline{\alpha}_0)\int_{t_0^*}^{1}\int_0^{t_0^*} F_1(\underline{t}_1(s))\mathrm{d}s\mathrm{d}G_0(t_0) \\ & +(1-\underline{\alpha}_0)\int_{t_0^*}^{1}\int_{t_0^*}^{t_0} F_1(\bar{t}_1(s))\mathrm{d}s\mathrm{d}G_0(t_0) - \underline{\alpha}_0\int_{t_0^*}^{1}\int_{t_0}^{1} F_1(\bar{t}_1(s))\mathrm{d}s\mathrm{d}G_0(t_0) \\ & +\underline{\alpha}_0 - \underline{\alpha}_0\int_0^1 t_0\mathrm{d}G_0(t_0). \end{aligned}$$

Using integration by parts, this can be rewritten as

$$\begin{aligned} \eta(G_0) \;\le\; & -\int_0^{t_0^*} F_1(\underline{t}_1(t_0))G_0(t_0)\mathrm{d}t_0 - \int_{t_0^*}^{1} F_1(\bar{t}_1(t_0))G_0(t_0)\mathrm{d}t_0 \\ & +(1-\underline{\alpha}_0)\int_0^{t_0^*} F_1(\underline{t}_1(t_0))\mathrm{d}t_0 + (1-\underline{\alpha}_0)\int_{t_0^*}^{1} F_1(\bar{t}_1(t_0))\mathrm{d}t_0 + \underline{\alpha}_0 - \underline{\alpha}_0\int_0^1 t_0\mathrm{d}G_0(t_0). \end{aligned}$$

Combining this with (24), and using the notation $t_0^{**}$ for an $\underline{\alpha}$-quantile of $G_0$, shows that the claim (28) can be written as

$$\begin{aligned} & -\int_0^{t_0^{**}} F_1(\underline{t}_1(s))G_0(s)\mathrm{d}s - \int_{t_0^{**}}^{1} F_1(\bar{t}_1(s))G_0(s)\mathrm{d}s \\ & +(1-\underline{\alpha}_0)\int_0^{t_0^{**}} F_1(\underline{t}_1(s))\mathrm{d}s + (1-\underline{\alpha}_0)\int_{t_0^{**}}^{1} F_1(\bar{t}_1(s))\mathrm{d}s \\ \le\; & -\int_0^{t_0^{*}} F_1(\underline{t}_1(s))G_0(s)\mathrm{d}s - \int_{t_0^{*}}^{1} F_1(\bar{t}_1(s))G_0(s)\mathrm{d}s \\ & +(1-\underline{\alpha}_0)\int_0^{t_0^{*}} F_1(\underline{t}_1(s))\mathrm{d}s + (1-\underline{\alpha}_0)\int_{t_0^{*}}^{1} F_1(\bar{t}_1(s))\mathrm{d}s. \end{aligned}$$

To verify this, consider the case $t_0^* \ge t_0^{**}$ (\"$\le$" is analogous). Then (28) holds because

$$0 \;\le\; \int_{t_0^{**}}^{t_0^*} (\underbrace{F_1(\bar{t}_1(s)) - F_1(\underline{t}_1(s))}_{>0})(\underbrace{G_0(s) - (1-\underline{\alpha}_0)}_{\ge 0})\mathrm{d}s.$$

*QED*

## References


Abreu, D., and H. Matsushima (1992): \Virtual Implementation in Iteratively Undominated Strategies: Incomplete Information," *mimeo*.

Balestrieri, F. (2008): \A modi ed English auction for an informed buyer," mimeo.

Balkenborg, D., and M. Makris (2010): \An Undominated Mechanism for a Class of Informed Principal Problems with Common Values," *mimeo*.

Beaudry, P. (1994): \Why an Informed Principal May Leave Rents to an Agent," *International Economic Review*, 35(4), 821{32.

Bond, E. W., and T. A. Gresik (1997): \Competition between asymmetrically informed principals," *Economic Theory*, 10(2), 227{240.

Cella, M. (2008): \Informed principal with correlation," *Games and Economic Behavior*, 64(2), 433{456.

Chade, H., and R. Silvers (2002): \Informed principal, moral hazard, and the value of a more informative technology," *Economics Letters*, 74(3), 291{300.

Che, Y.-K., and J. Kim (2006): \Robustly Collusion-Proof Implementation," *Econometrica*, 74(4), 1063{1107.

Cho, I.-K., and D. M. Kreps (1987): \Signaling Games and Stable Equilibria," *The Quarterly Journal of Economics*, 102(2), 179{221.

Cramton, P., R. Gibbons, and P. Klemperer (1987): \Dissolving a Partnership E ciently," *Econometrica*, 55(3), 615{32.

Esö, P., and J. Schummer (2009): \Credible deviations from signaling equilibria," *International Journal of Game Theory*, 38(3), 411{430.

Farrell, J. (1993): \Meaning and Credibility in Cheap-Talk Games," *Games and Economic Behavior*, 5, 514{531.

Fieseler, K., T. Kittsteiner, and B. Moldovanu (2003): \Partnerships, lemons, and e cient trade," *Journal of Economic Theory*, 113(2), 223{234.

Fleckinger, P. (2007): \Informed principal and countervailing incentives," *Economics Letters*, 94(2), 240{244.

Grossman, S. J., and M. Perry (1986): \Perfect Sequential Equilibrium," *Journal of Economic Theory*, 39, 97{119.

Halac, M. (2012): \Relational Contracts and the Value of Relationships," *American Economic Review*, 102(2), 750{79.

Izmalkov, S., and F. Balestrieri (2012): \The informed seller problem: The case of horizontal di erentiation," *mimeo*.

Jehiel, P., and A. Pauzner (2006): \Partnership Dissolution with Interdependent Values," *RAND Journal of Economics*, 37(1), 1{22.

Jost, P.-J. (1996): \On the Role of Commitment in a Principal-Agent Relationship with an Informed Principal," *Journal of Economic Theory*, 68(2), 510{530.

Kaya, A. (2010): \When Does It Pay To Get Informed?," *International Economic Review*, 51(2), 533{551.

Laffont, J.-J., and D. Martimort (1997): \Collusion under Asymmetric Information," *Econometrica*, 65(4), 875{912.

Ledyard, J. O., and T. R. Palfrey (2007): \A general characterization of interim e cient mechanisms for independent linear environments," *Journal of Economic Theory*, 133(1), 441{466.

Mailath, G. J., M. Okuno-Fujiwara, and A. Postlewaite (1993): \Belief-Based Re nements in Signalling Games," *Journal of Economic Theory*, 60(2), 241{276.

Maskin, E., and J. Tirole (1990): \The principal-agent relationship with an informed principal: The case of private values," *Econometrica*, 58(2), 379{409.

——— (1992): \The principal-agent relationship with an informed principal, II: Common values," *Econometrica*, 60(1), 1{42.

Matthews, S. A., M. Okuno-Fujiwara, and A. Postlewaite (1991): \Re ning cheap-talk equilibria," *Journal of Economic Theory*, 55(2), 247{273.

Mezzetti, C., and T. Tsoulouhas (2000): \Gathering information before signing a contract with a privately informed principal," *International Journal of Industrial Organization*, 18(4), 667{689.

Myerson, R. B. (1983): \Mechanism design by an informed principal," *Econometrica*, 51(6), 1767{1798.

Myerson, R. B., and M. A. Satterthwaite (1983): \E cient mechanisms for bilateral trading," *Journal of Economic Theory*, 29(2), 265{281.

Mylovanov, T., and T. Tröger (2008): \Optimal Auction Design and Irrelevance of Privacy of Information," *mimeo*.

——— (2012): \Irrelevancy of the privacy of the principal's information in independent private values environments," *mimeo*.

——— (forthcoming): \Informed-principal problems in environments with generalized private values," *Theoretical Economics*.

Quesada, L. (2005): \Collusion as an informed principal problem," *mimeo*.

——— (2010): \A Comprehensive Note on the Informed Principal with Private Values and Independent Types," *mimeo*.

Riley, J., and R. J. Zeckhauser (1983): \Optimal selling strategies: when to haggle, when to hold rm," *Quarterly Journal of Economics*, 98, 267{289.

Segal, I. R., and M. D. Whinston (2011): \A simple status quo that ensures participation (with application to e cient bargaining)," *Theoretical Economics*, 6(1).

Severinov, S. (2008): \An e cient solution to the informed principal problem," *Journal of Economic Theory*, 141(1), 114{133.

Skreta, V. (2009): \On the informed seller problem: optimal information disclosure," *Review of Economic Design*, 15(1), 1{36.

Tan, G. (1996): \Optimal Procurement Mechanisms for an Informed Buyer," *Canadian Journal of Economics*, 29(3), 699{716.

Williams, S. R. (1987): \E cient performance in two-agent bargaining," *Journal of economic theory*, 41(1), 154{172.

Yilankaya, O. (1999): \A note on the seller's optimal mechanism in bilateral trade with two-sided incomplete information," *Journal of Economic Theory*, 87(1), 125{143.

Zheng, C. Z. (2002): \Optimal auction with resale," *Econometrica*, 70(6), 2197{2224.