On the Informed Principal Model with Common Values *

Anastasios Dosis

ESSEC BUSINESS SCHOOL

The pioneering spirit

ON THE INFORMED PRINCIPAL MODEL WITH COMMON VALUES

RESEARCH CENTER

ANASTASIOS DOSIS

ESSEC WORKING PAPER 1905

APRIL 2019

On the Informed Principal Model with Common Values*

Anastasios Dosis †

April 5, 2019

Abstract

This paper reconsiders the general informed principal model with unilateral private information and common values. First, it identifies some fundamental properties of the Rothschild-Stiglitz-Wilson (RSW) allocation (i.e., the undominated for the principal allocation within the set of incentive compatible and individually rational for the agent type by type allocations). Based on these properties: (i) it constructs a more robust, and perhaps simpler, proof of the "if part" of Theorem 1 (i.e., the main result) of Maskin and Tirole (1992), and, (ii) it establishes that if the principal is restricted to offering mechanisms in which only she makes announcements (e.g., direct revelation mechanisms), then the conclusion of that theorem holds even in environments in which the RSW allocation is not interim efficient relative to any non-degenerate beliefs. Second, it provides a sufficient condition that allows for the complete characterisation of the set of equilibrium allocations even in environments in which single-crossing is not satisfied.

KEYWORDS: Mechanism design, informed principal, Rothschild-Stiglitz-Wilson allocation, perfect Bayesian equilibrium

JEL CLASSIFICATION: C72, D82

1 INTRODUCTION

In this paper, I reconsider the general principal-agent model with an informed principal. In particular, I study the set of equilibrium allocations of the standard three-stage game studied in Maskin and Tirole (1990, 1992): the principal offers a mechanism, and

*I thank Gorkem Celik, Eric Danan, Ani Guerjikova, Peter Hammond, Johannes Horner, Phil Reny, Tristan Tomala, Peter Vida, and Nicolas Vielle, as well as participants at the Conference for Research in Economic Theory and Econometrics, Greece, and seminar attendees at the University of Warwick, HEC Paris and CREST/Polytechnique for useful comments. This research has been conducted as part of the project Labex MME-DII (ANR11-LBX-0023-01).

†Department of Economics - ESSEC Business School and THEMA, 3 Av. Bernard Hirsch, B.P. – 50105, Cergy, 95021, France, Email: dosis@essec.com.

the agent accepts or rejects; upon acceptance, the mechanism is executed, whereas rejection leads to the termination of negotiations, and the reservation allocation enters into force. I concentrate on environments in which only the principal has private information, and this information could directly affect the payoff of the agent, i.e., unilateral private information and common values. Hence, the analysis is closer to Maskin and Tirole (1992) (hereafter, MT92) than to Maskin and Tirole (1990).

As shown in MT92, the set of Rothschild-Stiglitz-Wilson (RSW) allocations is key to the characterisation of the set of equilibrium allocations in the three-stage game. An RSW allocation is one that is undominated for the principal within the set of incentive compatible for the principal and individually rational for the agent type by type allocations.[1] The main result in MT92 is that when the principal is allowed to offer a mechanism from an arbitrary set of mechanisms, if the RSW allocation is interim efficient relative to some non-degenerate beliefs, then the set of equilibrium allocations consists of the RSW allocation as well as every incentive compatible allocation that dominates the RSW allocation and is individually rational for the agent relative to the prior beliefs (if such an allocation exists). I provide an alternative, perhaps more intuitive, proof of this result. The proof is decomposed in two steps. The first step highlights one of the principal properties of the RSW allocation. In particular, I demonstrate that if the RSW allocation is interim efficient relative to some non-degenerate beliefs, then for every arbitrary mechanism, there exist beliefs such that in every equilibrium of this mechanism under these beliefs, either every type is worse off relative to the RSW allocation or the agent is worse off relative to the reservation allocation. In light of this result, Theorem 1 of MT92 can be simply proven by constructing equilibrium strategies and beliefs.

It is noteworthy that the proof provided in this paper establishes that for every mechanism different from the RSW allocation, there exist beliefs such that in *every* equilibrium of the mechanism, either every type is worse off relative to the RSW allocation or the agent is worse off relative to the reservation allocation. MT92 prove instead that there exists *an* equilibrium with these properties. Therefore, part of the contribution of the paper is that it constitutes an extra robustness check on the main result in MT92.

I then restrict attention to environments in which the principal is restricted to offering mechanisms in which only she can make announcements. An important subset within

[1] Dominance here is defined with respect to the principal's preferences only; an incentive compatible allocation dominates another incentive compatible allocation if it is weakly prefered by every type of the principal and strictly prefered by at least one type.

this class of mechanisms is the set of direct revelation mechanisms (DRMs) in which the principal simply announces a type and an outcome is instructed based on his announcement. Under this restriction, I show that for every DRM different from the RSW allocation, there exist beliefs such that either every type is worse off relative to the RSW allocation or the agent is worse off relative to the reservation allocation. I provide a simple example that highlights the discrepancy in the sufficient conditions between the two cases (i.e., arbitrary mechanisms *vs* DRMs). In this example, the RSW allocation is not interim efficient relative to any non-degenerate beliefs. Although there exists a non-DRM (i.e., a mechanism in which the agent makes announcements) that dominates the RSW allocation and provides the agent a payoff that is strictly greater than his payoff in the reservation allocation relative to all possible beliefs, for every DRM that dominates the RSW allocation, there exist beliefs for which the agent is worse off relative to the reservation allocation. Therefore, by restricting attention to DRMs, one can establish that the set of equilibrium allocations consists of the RSW allocation as well as every allocation that dominates it and is individually rational for the agent relative to the prior beliefs (if such an allocation exists) even if the RSW allocation is not interim efficient relative to any non-degenerate beliefs.

To proceed to a complete characterisation of the set of equilibrium allocations, MT92 impose a no-tangency (i.e., single-crossing) condition. In particular, they show that if the indifference curves of different types are nowhere tangent, then every equilibrium allocation guarantees every type of the principal a payoff that is weakly greater than her payoff in the RSW allocation.[2] I provide a weaker condition according to which there exists a sequence of strictly incentive compatible for the principal and strictly individually rational for the agent type by type allocations such that the payoff of every type converges to her RSW allocation payoff. By means of an example, I show that this condition is less stringent than the usual single-crossing condition imposed in MT92.

In addition to MT92, the paper is related to several other works in the literature. The seminal paper is Myerson (1983), who introduces the general collective Bayesian incentive problem with multiple agents and interdependent values and axiomatically characterises reasonable solutions for the selection of a mechanism by the informed principal. Although the analysis is mainly cooperative, it touches on the non-cooperative equilibria of a game similar to that studied in MT92. Myerson (1983) shows that the axiomatically char-

[2]Proposition 5 (p. 18) and Theorem 1 (p. 19) in MT92.

acterised mechanisms can be supported as non-cooperative equilibrium mechanisms.

Maskin and Tirole (1990) consider the problem of mechanism design by an informed principal in a non-cooperative environment with private values, i.e., the private information of the principal does not directly affect the payoff of the agent. In non-quasilinear environments, the principal is strictly better off relative to her information being public. In quasilinear environments, the ex ante, interim and ex post optimal mechanisms are equivalent. Similar results are obtained in Myerson (1985), Tan (1996), Yilankaya (1999) and Skreta (2011). However, this equivalence result is challenged in Fleckinger (2007) and more recently in Mylovanov and Tröger (2014), who consider countervailing incentives, and in Cella (2008), who allows the principal's and agent's types to be correlated.

Perhaps surprisingly, the literature on informed principal models with common values is less developed than its private-values counterpart. Severinov (2008) analyses a general multi-agent, informed principal environment with interdependent values and establishes conditions under which a variant of the three-stage game has a perfect sequential equilibrium that is ex post efficient. The result relies on the correlation of types, as in Cremer and McLean (1988). Balkenborg and Makris (2015) analyse an environment with common values and transferable utility. They identify a solution for the informed principal with appealing properties, i.e., the *assured allocation*, and they show that it is sustainable as an equilibrium in the three-stage game.

The present paper is also related to the problem of an informed seller of an indivisible object. Skreta (2011) shows that the optimal information disclosure crucially depends on whether values are private or common. More recently, Koessler and Skreta (2016) and Balestrieri and Izmalkov (2016) study optimal mechanisms for an informed seller with common values when the seller is solely interested in maximising her revenue (i.e., she has no reservation value for the object).

The remainder of this paper is organised as follows. In Section 2, I present the general informed principal model and define allocations, incentive compatibility, dominance, interim efficiency, individual rationality and the RSW allocation. I also define mechanisms. In Section 3, I study the properties of the RSW allocation. In Section 4, I characterise the set of equilibrium allocations of the three-stage game. Section 5 concludes the paper.

2 THE MODEL

■ **Agents and Payoffs.** Consider two players: a principal (P) and an agent (A).[3] The principal has a finite number of possible types $i = 1, ..., n$, which is her private information. Let $\Pi^{\cdot} = (\Pi^i)_i$ denote the *prior beliefs* of the agent about the principal's type, where $\Pi^i > 0$ for every i and $\sum \Pi^i = 1$. Let also X be a compact subset of $\mathbb{R}^K$ with representative element $x \in X$.[4] As a concrete example, one can consider the principal as a seller of a product and the agent as a buyer. In this case, $x = (t, q)$, where t denotes a transfer from the buyer to the seller and q denotes the quantity or quality produced by the seller.

The preferences of all players and types over X admit an expected utility representation, with the (VNM) utility index of the principal and the agent denoted by $V^i(x)$ and $U^i(x)$, respectively. Both functions are continuous for every i.[5] The two players wish to select a (possibly random) *contractible outcome* μ from $\mathcal{M} = \Delta(X)$, where $\Delta(X)$ denotes the set of probability distributions over X.[6] Let $V^i(\mu) = \int V^i(x)d\mu$ and $U^i(\mu) = \int U^i(x)d\mu$.

□ **Allocations.** Following MT92, the study of mechanism selection by the informed principal is usually performed through the characterisation of equilibrium allocations. Some standard definitions follow.

Definition 1. *An allocation is a menu of outcomes $\mu^{\cdot} = (\mu^i)_i$, one for each type of principal.*

Of prominent importance are the concepts of *incentive compatibility*, *dominance* and *interim efficiency*.

Definition 2. *An allocation $\mu^{\cdot}$ is incentive compatible if $V^i(\mu^i) \geq V^i(\mu^j)$ for every i, j.*

Definition 3. *An incentive compatible allocation $\bar{\mu}^{\cdot}$ dominates an incentive compatible allocation $\mu^{\cdot}$ if $V^i(\bar{\mu}^i) \geq V^i(\mu^i)$ for every i, with the inequality being strict for at least one i.*

Note that incentive compatibility and dominance are defined with respect to the principal's preferences only.

[3] To the extent possible and to facilitate direct comparison, I follow the notation of MT92. Moreover, for simplicity, as in MT92, I refer to the principal as feminine (she or her) and the agent as masculine (he or him). When I refer to any player (either the principal or the agent), I will use the feminine pronoun.

[4] Note that more general (compact) spaces can be considered. Restriction to real spaces is to remain as close as possible to the environment studied in MT92.

[5] Compactness and continuity are sufficient for optimal contracts to exist. In fact, the compactness of X might be stronger than is required; it suffices that for every real number u and every type i, $U^i(x) \geq u$ defines a compact set.

[6] Random outcomes (and consequently random allocations) are only used in the proofs of Proposition 1 and Theorem 2. They are not required in the proofs of Proposition 2 and Theorem 2.

Definition 4. *An allocation $\bar{\mu}^{\cdot}$ is interim efficient relative to beliefs $\bar{\Pi}^{\cdot}$ (not necessarily the prior beliefs) if (i) it is incentive compatible, and (ii) there exists no allocation $\mu^{\cdot} \neq \bar{\mu}^{\cdot}$ that is incentive compatible, satisfies $\sum \bar{\Pi}^i U^i(\mu^i) \geq \sum \bar{\Pi}^i U^i(\bar{\mu}^i)$, and dominates $\bar{\mu}^{\cdot}$.*

A weaker notion of efficiency is that of weak interim efficiency. A formal definition is provided below.

Definition 5. *An allocation $\bar{\mu}^{\cdot}$ is weakly interim efficient (or interim efficient type by type) if (i) it is incentive compatible, and (ii) there exists no allocation $\mu^{\cdot} \neq \bar{\mu}^{\cdot}$ that is incentive compatible, satisfies $U^i(\mu^i) \geq U^i(\bar{\mu}^i)$ for every i, and dominates $\bar{\mu}^{\cdot}$.*

Because in the game played between the principal and the agent (to be described below), the agent has the right to reject the mechanism proposed by the principal, it becomes indispensable to define the payoff of the agent in a possible disagreement of the two players. MT92 assume that there exists a *reservation allocation* $\mu_0^{\cdot}$, which can be regarded as either an outside option or a prior contract that binds the two players and they wish to renegotiate.[7] Without loss of generality, let $\mu_0^{\cdot}$ be incentive compatible. As is well-known, individually rational allocations take into consideration the reservation allocation.

Definition 6. *An incentive compatible allocation $\mu^{\cdot}$ is individually rational relative to beliefs $\bar{\Pi}^{\cdot}$ (not necessarily the prior beliefs) if $\sum \bar{\Pi}^i U^i(\mu^i) \geq \sum \bar{\Pi}^i U^i(\mu_0^i)$.*

A weaker notion of individual rationality is individual rationality type by type. A formal definition of individually rational type by type allocations is provided below.

Definition 7. *An incentive compatible allocation $\mu^{\cdot}$ is individually rational type by type if $U^i(\mu^i) \geq U^i(\mu_0^i)$ for every i.*

A subset of the set of weakly interim efficient allocations is the set of *Rothschild-Stiglitz-Wilson* (RSW) allocations. RSW allocations are weakly interim efficient allocations that are, furthermore, individually rational type by type. RSW allocations have a central role in the characterisation of the set of equilibrium allocations.

Definition 8. *An allocation $\mu^{\cdot}$ is an RSW allocation (relative to the reservation allocation $\mu_0^{\cdot}$) if (i) it is incentive compatible and individually rational type by type, and (ii) there exists no allocation $\mu^{\cdot} \neq \bar{\mu}^{\cdot}$ that is incentive compatible, individually rational type by type, and dominates $\mu^{\cdot}$.*

[7] See p. 9 in MT92 for further details.

In other words, the set of RSW allocations includes all the undominated allocations within the set of incentive compatible and individually rational type by type allocations. Because all RSW allocations are payoff equivalent for the principal, as in MT92, I assume that there exists a unique RSW allocation and refer to this as *the* RSW allocation instead of *an* RSW allocation.[8] Therefore, let $\hat{\mu}^{\cdot}(\mu_0^{\cdot})$ denote the RSW allocation.

□ **Mechanisms.** A general *finite mechanism* is denoted by $m = (S, g)$ and consists of two finite sets of payoff-irrelevant messages, one for each player, $S = S_P \times S_A$, and a mapping from the set of messages to the set of contractible outcomes, $g : S \to \mathcal{M}$. A mechanism defines a game of incomplete information in which the two players simultaneously and independently select a message, and an outcome is determined based on the two messages and the protocol of the mechanism.[9]

A strategy for the principal in mechanism m specifies a probability distribution over the set of her possible messages for every possible type. A strategy for the agent in mechanism m specifies a probability distribution over the set of his possible messages.[10] For given beliefs $\bar{\Pi}^{\cdot}$, a Bayesian Nash equilibrium in mechanism m consists of a profile of strategies, one for each player, such that, conditional on the strategy of the other player, no player has a unilateral profitable deviation. Every equilibrium is associated with an (expected) ex post equilibrium payoff profile. For notational convenience, let $(\bar{V}^{\cdot}(m, \bar{\Pi}^{\cdot}), \bar{U}^{\cdot}(m, \bar{\Pi}^{\cdot}))$, where $\bar{V}^{\cdot}(m, \bar{\Pi}^{\cdot}) = (\bar{V}^i(m, \bar{\Pi}^{\cdot}))_i$ and $\bar{U}^{\cdot}(m, \bar{\Pi}^{\cdot}) = (\bar{U}^i(m, \bar{\Pi}^{\cdot}))_i$, denote an equilibrium payoff profile. Clearly, multiple Bayesian Nash equilibria might exist in mechanism m under beliefs $\bar{\Pi}^{\cdot}$. For now, no equilibrium selection is implicitly made. When an equilibrium selection is made, e.g., in the proofs of Proposition 1 and Theorem 1, this will be clearly stated.

A simpler class of mechanisms is the class of *direct revelation mechanisms* (DRMs), in which the principal simply announces a type (not necessarily the true type), and the agent makes no announcement.[11] Hence, $S_P = \{1, ..., n\}$ and $S_A = \emptyset$. I will freqently focus on DRMs in which the principal truthfully announces her type by invoking the well-known

[8] See the appendix for a formal proof of this statement.

[9] One can consider more general mechanisms as in Mylovanov and Tröger (2014). Restriction to finite, simultaneous-move mechanisms suffices for the existence of a Bayesian Nash equilibrium in every mechanism.

[10] Note that the strategies specified here are with no reference to any mechanism-proposal game; hence, they should not be confused with the strategies in the three-stage game defined below (i.e., mechanism proposal/acceptance-rejection/mechanism execution).

[11] An allocation is a DRM.

revelation principle. Moreover, I will examine the case in which the principal is restricted to offering only DRMs.

3 Properties of the RSW Allocation

■ **General Properties.** The significance of the RSW allocation arises from the following two propositions:

Proposition 1. *Suppose that $\hat{\mu}^{\cdot}(\mu_0^{\cdot})$ (i.e., the RSW allocation) is interim efficient relative to beliefs $\hat{\Pi}^{\cdot}$ (not necessarily the prior beliefs), where $\hat{\Pi}^i > 0$ for every i; then, for every mechanism $m \neq \hat{\mu}^{\cdot}(\mu_0^{\cdot})$, there exist beliefs $\bar{\Pi}^{\cdot}$ such that in every equilibrium of m under $\bar{\Pi}^{\cdot}$ with an associated equilbrium payoff profile $(\bar{V}^{\cdot}(m, \bar{\Pi}^{\cdot}), \bar{U}^{\cdot}(m, \bar{\Pi}^{\cdot}))$, either*

$$\bar{V}^i(m, \bar{\Pi}^{\cdot}) \leq V^i(\hat{\mu}^i(\mu_0^{\cdot})) \text{ for every } i \tag{1}$$

or

$$\sum \bar{\Pi}^i \bar{U}^i(m, \bar{\Pi}^{\cdot}) < \sum \bar{\Pi}^i U^i(\mu_0^i) \tag{2}$$

Proof. I prove the result by contradiction. Suppose that the RSW allocation $\hat{\mu}^{\cdot}(\mu_0^{\cdot})$ is interim efficient relative to beliefs $\hat{\Pi}^{\cdot}$, where $\hat{\Pi}^i > 0$ for every i, and there exists a mechanism $m \neq \hat{\mu}^{\cdot}(\mu_0^{\cdot})$ and a subset of types $I \subseteq \{1, ..., n\}$ such that for every $\bar{\Pi}^{\cdot}$, there exists an equilibrium with an associated equilbrium payoff profile $(\bar{V}^{\cdot}(m, \bar{\Pi}^{\cdot}), \bar{U}^{\cdot}(m, \bar{\Pi}^{\cdot}))$, such that:

(i) $\bar{V}^i(m, \bar{\Pi}^{\cdot}) > V^i(\hat{\mu}^i(\mu_0^{\cdot}))$ for every $i \in I$, and $\bar{V}^i(m, \bar{\Pi}^{\cdot}) \leq V^i(\hat{\mu}^i(\mu_0^{\cdot}))$ for every $i \notin I$

and

(ii) $\sum \bar{\Pi}^i \bar{U}^i(m, \bar{\Pi}^{\cdot}) \geq \sum \bar{\Pi}^i U^i(\mu_0^i)$

Consider $\Pi_I^{\cdot}$, where $\Pi_I^i = \hat{\Pi}^i / \sum_{j \in I} \hat{\Pi}^j$ for every $i \in I$, and $\Pi_I^i = 0$ for every $i \notin I$. Note that $\Pi_I^i > 0$ for every $i \in I$ because $\hat{\Pi}^i > 0$ for every i. Consider allocation $\tilde{\mu}^{\cdot}$, where $V^i(\tilde{\mu}^i) = \bar{V}^i(m, \Pi_I^{\cdot})$ for every $i \in I$ and $\tilde{\mu}^i = \hat{\mu}^i(\mu_0^{\cdot})$ for every $i \notin I$. Allocation $\tilde{\mu}^{\cdot}$ is incentive compatible because of (i) above, and $\hat{\mu}^{\cdot}(\mu_0^{\cdot})$ is incentive compatible by definition. Moreover,

$$\sum \hat{\Pi}^i U^i(\tilde{\mu}^i) = \sum_{i \in I} \hat{\Pi}^i \bar{U}^i(m, \Pi_I^{\cdot}) + \sum_{i \notin I} \hat{\Pi}^i U^i(\hat{\mu}^i(\mu_0^{\cdot})) \geq \sum \hat{\Pi}^i U^i(\mu_0^i) \tag{3}$$

because

$$\sum_{i\in I}\frac{\hat{\Pi}^i}{\sum_{j\in I}\hat{\Pi}^j}\bar{U}^i(m,\Pi_I^{\cdot})\geq\sum_{i\in I}\frac{\hat{\Pi}^i}{\sum_{j\in I}\hat{\Pi}^j}U^i(\mu_0^i) \tag{4}$$

from (ii) above and due to the fact that the RSW allocation is individually rational type by type. Therefore, $\tilde{\mu}^{\cdot}$ is individually rational relative to beliefs $\hat{\Pi}^{\cdot}$. Because $\tilde{\mu}^{\cdot}$ dominates $\hat{\mu}^i(\mu_0^{\cdot})$ and is individually rational relative to beliefs $\hat{\Pi}^{\cdot}$, there is a contradiction with $\hat{\mu}^i(\mu_0^{\cdot})$ being interim efficient relative to beliefs $\hat{\Pi}^{\cdot}$. □

Put differently, Proposition 1 states that for every mechanism for which, for all beliefs and some equilibrium in the mechanism, the agent is not worse off relative to the reservation allocation, there exist beliefs for which, in every equilibrium of this mechanism given these beliefs, every type is worse off relative to the RSW allocation. The result exploits the definition of the RSW allocation and the fact that this is interim efficient relative to some non-degenerate beliefs. It shows by contradiction that if there were a mechanism that satisfied (in some equilibrium of the mechanism) the individual rationality constraint of the agent and improved the payoff of a subset of types for all possible beliefs, one could construct a new allocation that would dominate the RSW allocation and be individually rational relative to the beliefs for which this is interim efficient. The significance of the hypothesis that the RSW allocation be efficient relative to some non-degenerate beliefs rests on the possibility of applying Bayes' rule to construct the required beliefs and Eqs. (3) and (4).

Remark. MT92 show that in standard two-dimensional environments, i.e., environments with transferable utility in which the principal's preferences satisfy sorting and the agent's utility (from every outcome) and reservation utility are increasing in the type of principal, the RSW allocation is indeed interim efficient relative to some non-degenerate beliefs.

I now restrict the set of feasible mechanisms to mechanisms in which only the principal can make meaningful announcements. Under this restriction, there is no loss of generality in considering only DRMs, and specifically, incentive compatible DRMs. The following proposition shows how the results change under such a restriction.

Proposition 2. *Suppose that the principal is restricted to offering only DRMs; then, for every incentive compatible allocation* $\mu^{\cdot}\neq\hat{\mu}^{\cdot}(\mu_0^{\cdot})$ *(where* $\hat{\mu}^{\cdot}(\mu_0^{\cdot})$ *is the RSW allocation), there exist*

beliefs $\bar{\Pi}^{\cdot}$ *such that either*

$$V^i(\mu^i) \leq V^i(\hat{\mu}^i(\mu_0^{\cdot})) \text{ for every } i \tag{5}$$

or

$$\sum \bar{\Pi}^i U^i(\mu^i) < \sum \bar{\Pi}^i U^i(\mu_0^i) \tag{6}$$

Proof. Suppose that the principal is restricted to offering only DRMs and there exists an incentive compatible allocation $\mu^{\cdot} \neq \hat{\mu}^{\cdot}(\mu_0^{\cdot})$ and subset of types $I \subseteq \{1, ..., n\}$ such that for every $\bar{\Pi}^{\cdot}$:

(i) $V^i(\mu^i) > V^i(\hat{\mu}^i(\mu_0^{\cdot}))$ for every $i \in I$ and $V^i(\mu^i) \leq V^i(\hat{\mu}^i(\mu_0^{\cdot}))$ for every $i \notin I$

and

(ii) $\sum \bar{\Pi}^i U^i(\mu^i) \geq \sum \bar{\Pi}^i U^i(\mu_0^i)$

Because $\mu^{\cdot}$ satisfies (ii) for every $\bar{\Pi}^{\cdot}$, it is individually rational type by type. Consider allocation $\tilde{\mu}^{\cdot}$, where

$$\tilde{\mu}^i = \begin{cases} \mu^i, & \text{if } i \in I \\ \hat{\mu}^i(\mu_0^{\cdot}), & \text{otherwise} \end{cases}$$

This allocation is incentive compatible because $\mu^{\cdot}$ and $\hat{\mu}^{\cdot}(\mu_0^{\cdot})$ are incentive compatible and (i) above. Moreover, $\tilde{\mu}^{\cdot}$ is individually rational type by type because $\hat{\mu}^{\cdot}(\mu_0^{\cdot})$ and $\mu^{\cdot}$ are individually rational type by type by the definition of the RSW allocation and the fact that $\mu^{\cdot}$ is individually rational type by type respectively. Therefore, $\mu^{\cdot}$ dominates $\hat{\mu}^{\cdot}(\mu_0^{\cdot})$ and is individually rational type by type, which contradicts $\hat{\mu}^{\cdot}(\mu_0^{\cdot})$ being an RSW allocation. □

Proposition 2 states that if one considers only DRMs, then for any DRM that improves the payoff of at least one of the principal's types relative to the RSW allocation, there exist beliefs such that the agent is worse off relative to the reservation allocation. The idea behind the proof relies on the main property of the RSW allocation: the RSW allocation is individually rational type by type. Due to this property, if there were a DRM that improved the payoff of at least one type relative to all possible beliefs, then the composite allocation constructed by the RSW allocation and this allocation would be incentive compatible and individually rational type by type and would dominate the RSW allocation. That would contradict the definition of the RSW allocation.

μ	$U^1(\mu)$	$U^2(\mu)$	$V^1(\mu)$	$V^2(\mu)$
-1	2	-1	2	2
0	0	0	1	1
1	-1	1	3	3
μ_0	0	0	0	0

Table 1

This last point also explains the discrepancy between Propositions 1 and 2. Note that when arbitrary mechanisms are allowed, the payoffs of both players in the mechanism depend on the beliefs that the agent entertains about the principal's type, as it is the agent alongside the principal who makes announcements in the mechanism.[12] Therefore, the logic of the proof of Proposition 2 does not extend to the case in which the agent makes announcements, as the composite DRM constructed by combining the RSW allocation and the mechanism for all possible degenerate beliefs might not be incentive compatible (as it is when only DRMs are allowed).

Propositions 1 and 2 are key to the characterisation of the set of equilibrium allocations, as will become clear in Section 4. In what follows, I provide a simple example that sheds light on the discrepancy between the sufficient conditions in the two propositions.

Example 1. Suppose that $n = 2$, $\Pi^i > 0$ for every i and $\mathcal{M} = \{-1, 0, 1\}$.[13] The payoffs of the two players are given in Table 1.

In this example, the set of incentive compatible allocations is $\{(-1,-1),(0,0),(1,1)\}$, and the RSW allocation is $\hat{\mu}^{\cdot}(\mu_0^{\cdot}) = (0,0)$. The expected payoff of the agent from the RSW allocation is zero for any beliefs. Moreover, for given beliefs $\bar{\Pi}^{\cdot}$, the respective expected payoffs of the agent from allocations $(-1,-1)$ and $(1,1)$ are

$$\sum \bar{\Pi}^i U^i(-1) = 2\bar{\Pi}^1 - (1 - \bar{\Pi}^1) = 3\bar{\Pi}^1 - 1,$$

which is strictly positive if and only if $\bar{\Pi}^1 > 1/3$, and

$$\sum \bar{\Pi}^i U^i(1) = -\bar{\Pi}^1 + (1 - \bar{\Pi}^1) = 1 - 2\bar{\Pi}^1,$$

[12] Evidently, even when only DRMs are allowed, the payoff of the agent depends on his beliefs. What is critical is that the payoff of the principal depends on the beliefs of the agent about her type, which is not true in a DRM.

[13] For the sake of the example, I drop the assumption of random outcomes and allow only for deterministic outcomes and allocations. Given that, as I show below, the RSW allocation is not interim efficient relative to any non-degenerate beliefs, Proposition 1 does not apply and as mentioned above, random outcomes are only required in the proof of Proposition 1.

which is strictly positive if and only if $\bar{\Pi}^1 < 1/2$.

The RSW allocation is not interim efficient relative to any non degenerate beliefs. Indeed, both $(-1,-1)$ and $(1,1)$ dominate the RSW allocation, and for $\bar{\Pi}^1 \geq 1/3$, the expected payoff of the agent is greater in $(-1,-1)$ than in the RSW allocation, whereas, for $\bar{\Pi}^1 \leq 1/2$, the expected payoff of the agent is greater in $(1,1)$ than in the RSW allocation.

Note also that

$$\sum \bar{\Pi}^i U^i(-1) > \sum \bar{\Pi}^i U^i(1)$$

if and only if $\bar{\Pi}^1 > 2/5$.

Now consider mechanism $m^d = (S^d, g^d)$, where $S_P^d = \emptyset$, $S_A^d = \{-1, 1\}$ and $g^d(s) = s$ for every $s \in S_A^d$. In this mechanism, the principal has no available announcements to make, whereas the agent makes an announcement by selecting -1 or 1 with the outcome of the mechanism coinciding with the announcement of the agent. The payoff maximising announcement for the agent depends on his beliefs regarding the principal's type. If $\bar{\Pi}^1 > 2/5$, the agent strictly prefers -1 over 1; if $\bar{\Pi}^1 < 2/5$, he strictly prefers 1 over -1; and if $\bar{\Pi}^1 = 2/5$, he is indifferent between -1 and 1. Note, however, that regardless of his beliefs, this mechanism provides the agent with an equilibrium payoff that is strictly greater than his payoff in the reservation allocation, which equals zero. Therefore, for mechanism m^d, there exists no beliefs for which either (1) or (2) is satisfied.

Instead, suppose that the principal is restricted to offering only DRMs. As argued above, by the revelation principle, it suffices to consider solely incentive compatible DRMs, and as mentioned above, there are three incentive compatible DRMs. In each of these DRMs, both types truthfully announce their type: in the first, the outcome for both is -1; in the second, the outcome for both is 0; and in the third, the outcome for both is 1. Note, however, that for allocation $(-1,-1)$ and $\bar{\Pi}^1 < 1/3$, the payoff of the agent is strictly negative and hence lower than his payoff in the reservation allocation, whereas for allocation $(1,1)$ and $\bar{\Pi}^1 > 1/2$, the payoff of the agent is strictly negative and hence lower than his payoff in the reservation allocation. Hence, for every incentive compatible allocation, there exist beliefs for which either (5) or (6) is satisfied.

4 Equilibrium Allocations

■ **The Game.** The goal of the paper is to characterise the set of allocations that can be sustained as perfect Bayesian equilibrium allocations in the extensive-form game studied

in MT92. This extensive-form game has three stages specified below.

- *Stage 1:* The principal proposes a mechanism.
- *Stage 2:* The agent accepts or rejects the proposal. If the agent rejects, the game ends. If the agent accepts, the game moves to the next stage.
- *Stage 3:* The mechanism is executed.

A strategy for the principal consists of a proposal of a mechanism in *Stage 1* for every possible type, and a probability distribution over the set of her possible messages in *Stage 3* if the agent accepts the mechanism for every possible mechanism proposed. A strategy for the agent consists of an acceptance or rejection decision in *Stage 2* for every possible proposal and, if arbitrary mechanisms are allowed, a probability distribution over the set of his possible messages in the mechanism for every possible history of play in *Stage 3*.[14]

A system of posterior beliefs for the agent in the beginning of *Stage 2* is a probability distribution over the set of types of the principal for every possible proposed mechanism in *Stage 1*.

A perfect Bayesian equilibrium is a profile of strategies and beliefs such that (i) the strategy of every player is optimal given the strategy of the other player and her beliefs at every node of the game tree (sequential rationality); (ii) beliefs are updated by Bayes rule whenever a node is reached given the equilibrium strategies; and (iii) beliefs are arbitrarily determined in nodes that are not reached given the equilibrium strategies but are consistent with the equilibrium strategies.

□ **Equilibrium Allocations – Partial Characterisation.** Myerson (1983) shows that for every possible equilibrium where different subsets of types in a partition of the type space offer distinct mechanisms, there is a payoff-equivalent equilibrium in which all types offer the same mechanism. This illuminating observation is known as the *inscrutability principle* and holds considerable value when studying informed principal problems. Thanks

[14]Note that I concentrate on equilibria in which, on-the-equilibrium path, the principal does not randomise over the set of possible mechanisms although she might randomise over the set of possible messages in a mechanism. For, if the principal offers a mechanism that entails the agent making an announcement, given that the players take actions simultaneously, randomisation over the set of possible messages is required to ensure the existence of an equilibrium in every continuation of the game (i.e., on- or off-the equilibrium path). In contrast, when the principal is restricted to offering only DRMs, one can focus on equilibria in which the principal does not even randomise over the set of her possible messages on- or off-the equilibrium path.

to this principle, there is no loss of generality when focusing on equilibria in which (in equilibrium) all types offer the same DRM (i.e., allocation).

The following theorem is one part of the main result in MT92 (i.e., Theorem 1, p. 19). It characterises the set of allocations that can be sustained as equilibrium allocations in the three-stage game.

Theorem 1 (MT92). *Suppose that $\hat{\mu}^{\cdot}(\mu_0^{\cdot})$ (i.e., the RSW allocation) is interim efficient relative to beliefs $\hat{\Pi}^{\cdot}$ (not necessarily the prior beliefs), where $\hat{\Pi}^i > 0$ for every i; then, every allocation $\bar{\mu}^{\cdot}$ that is incentive compatible and satisfies*

$$V^i(\bar{\mu}^i) \geq V^i(\hat{\mu}^i(\mu_0^{\cdot})) \; \forall \; i \tag{7}$$

$$\sum \Pi^i U^i(\bar{\mu}^i) \geq \sum \Pi^i U^i(\mu_0^i) \tag{8}$$

is an equilibrium allocation of the three-stage game.

Proof. Let $\bar{\mu}^{\cdot}$ be an incentive compatible allocation that satisfies (7) and (8) and consider the following profile of strategies and beliefs:

- *Principal's strategy:* Every type i proposes $\bar{\mu}^{\cdot}$ in *Stage 1* and truthfully reveals her type in *Stage 3*. For every other mechanism m, the principal selects a probability distribution over the set of her possible messages (in the mechanism) in *Stage 3* that is optimal given the strategy of the agent and the posterior beliefs.

- *Agent's strategy:* If the principal offers $\bar{\mu}^{\cdot}$ or any mechanism $m \neq \bar{\mu}^{\cdot}$, where, for every $\bar{\Pi}^{\cdot}$ and every equilibrium in m under $\bar{\Pi}^{\cdot}$ with an associated equilibrium payoff profile $(\bar{V}^{\cdot}(m, \bar{\Pi}^{\cdot}), \bar{U}^{\cdot}(m, \bar{\Pi}^{\cdot}))$, it is true that:

$$\sum \bar{\Pi}^i \bar{U}^i(m, \bar{\Pi}^{\cdot}) \geq \sum \bar{\Pi}^i U^i(\mu_0^i) \tag{9}$$

then the agent "accepts" in *Stage 2*, and if the mechanism proposed entails his making an announcement, he selects a probability distribution over the set of his possible messages (in the mechanism) in *Stage 3* that is optimal given the strategy of the principal and his posterior beliefs. If the principal offers any other mechanism, then the agent "rejects" in *Stage 2*.

- *Beliefs:* If the principal offers $\bar{\mu}^{\cdot}$, then the beliefs remain equal to the prior beliefs. If the principal offers $m \neq \bar{\mu}^{\cdot}$ such that, for every $\bar{\Pi}^{\cdot}$ and in some equilibrium in mechanism m given beliefs $\bar{\Pi}^{\cdot}$ with an associated equilibrium payoff profile $(\bar{V}^{\cdot}(m, \bar{\Pi}^{\cdot}), \bar{U}^{\cdot}(m, \bar{\Pi}^{\cdot}))$, (9)

is satisfied, then the beliefs are updated to $\bar{\bar{\Pi}}^{\cdot}$ such that in every equilibrium in m given beliefs $\bar{\bar{\Pi}}^{\cdot}$ with an associated equilibrium payoff profile $(\bar{V}^{\cdot}(m, \bar{\bar{\Pi}}^{\cdot}), \bar{U}^{\cdot}(m, \bar{\bar{\Pi}}^{\cdot}))$, $\bar{V}^{i}(m, \bar{\bar{\Pi}}^{\cdot}) \leq V^{i}(\hat{\mu}^{i}(\mu_0^{\cdot}))$ for every i. Proposition 1 ensures that $\bar{\bar{\Pi}}^{\cdot}$ always exists. If the principal offers any other mechanism, then the beliefs are determined such that (9) is not satisfied in any equilibrium of the mechanism.

Both players' strategies are sequentially rational given the beliefs of the agent, and the beliefs are determined by Bayes' rule on the equilibrium path and are determined in such a way that no player has a unilateral profitable deviation off the equilibrium path. To see this, consider first the agent's strategy. Given the belief system, in *Stage 2*, the agent rejects any proposed mechanism that, in the continuation of the game (i.e., in some equilibrium of the mechanism), will result in an expected payoff lower than that he can attain in the reservation allocation and accepts every mechanism that, regardless of his beliefs, results in an expected payoff that, in every continuation of the game (i.e., in every equilibrium of the mechanism), will result in a payoff higher than that he can attain in the reservation allocation. Note also that his strategy remains optimal in *Stage 3* given his beliefs and the principal's strategy. Given the agent's specified strategy, the principal cannot unilaterally deviate and increase her expected payoff in *Stage 1*. This is because for every other mechanism, in the continuation of the game, the agent will either reject in *Stage 2* or accept in *Stage 2*, but the resulting equilibrium payoff in the mechanism in *Stage 3* will be lower than that in the RSW allocation for every type, which from (7) is weakly lower than her payoff in the DRM $\bar{\mu}^{\cdot}$. □

Consider the RSW allocation or any other incentive compatible allocation that dominates the RSW allocation (if such an allocation exists). Suppose that all principal types offer this allocation, and the agent upon observing such an offer does not update his beliefs, i.e., continues to hold his prior beliefs. The critical question is whether some [principal] type can benefit by offering some other mechanism. The key in answering this question is the determination of posterior beliefs for every mechanism other than the one that the agent expects the principal to offer (hereafter, the "on-the-equilibrium path" mechanism). In fact, one can decompose the set of all feasible mechanisms in two mutually exclusive subsets (say, A and A^c). Subset A includes those mechanisms that are individually rational for the agent relative to all possible beliefs, i.e., for all possible beliefs and every equilibrium in the mechanism, the agent earns an expected payoff higher than that in the reservation allocation. Proposition 1 establishes that for every such mechanism, there

exist beliefs such that in every equilibrium of the mechanism, all principal types earn a payoff lower than the maximum payoff they can obtain in the RSW allocation. Therefore, if the beliefs are determined in such a way, the principal cannot benefit by proposing any of the mechanisms in subset A, as she expects to earn a payoff in the continuation of the game that is lower than her payoff in the on-the-equilibrium path mechanism. Subset A^c is the complement of subset A, i.e., it includes every other feasible mechanism. For every mechanism in A^c, there exist beliefs such that the mechanism is not individually rational for the agent, i.e., in every equilibrium of the mechanism, the agent earns an expected payoff strictly lower than the expected payoff he can attain in the reservation allocation. Because beliefs are arbitrarily determined off-the-equilibrium path, if they are determined such that the mechanism is not individually rational for the agent, the agent will reject the offer, and hence no type can profitably deviate by offering any mechanism in subset A^c either.

Remark. In the proof of Proposition 1 (and consequently Theorem 1), it is shown that for every mechanism (other than the RSW allocation), there exist beliefs such that in *every* equilibrium of this mechanism under these beliefs, either the principal is worse off relative to the RSW allocation or the agent is worse off relative to the reservation allocation. MT92 prove that for every mechanism (other than the RSW allocation), there exist beliefs such that in *some* equilibrium of the mechanism, either the principal is worse off relative to the RSW allocation or the agent is worse off relative to the reservation allocation. Therefore, the proof provided in this paper enhances the robustness of the main result in MT92.

I now examine the equilibrium allocations of the three-stage game if the principal is restricted to offering only DRMs.

Theorem 2. *Suppose that the principal is restricted to offering only DRMs; then, every allocation $\bar{\mu}(\cdot)$ that satisfies (7) and (8) is an equilibrium allocation of the three-stage game.*

The RSW allocation and every allocation that dominates it relative to the prior beliefs (if such an allocation exists) can be sustained as an equilibrium allocation in the three-stage game if the principal can propose only DRMs. This is true even if the RSW allocation is not interim efficient relative to some non-degenerate beliefs. The proof of this result

is almost identical to that of Theorem 1 but exploits Proposition 2 instead of Proposition 1.

■ **Equilibrium Allocations – Complete Characterisation.** As mentioned above, Theorems 1 and 2 provide only a partial characterisation of the set of equilibrium allocations. This is because they are mute regarding allocations in which the payoff of some [principal] type(s) is strictly lower than her payoff in the RSW allocation. MT92 assume that X is convex and that a type i indifference curve is nowhere tangent to a type j indifference curve ($i \neq j$). Under this condition, they show that the payoff of every type in every equilibrium allocation of the three-stage game is weakly greater than her payoff in the RSW allocation (Proposition 5, p. 18 in MT92). Evidently, by imposing the same condition, one could also fully characterise the set of equilibrium allocations. In fact, if one is interested in a complete characterisation of the set of equilibrium allocations, then one can provide a more general sufficient condition than the no-tangency condition imposed in MT92. Doing so requires two additional definitions.

Definition 9. *An allocation $\mu^{\cdot}$ is strictly incentive compatible (SIC) if $V^i(\mu^i) > V^i(\mu^j)$ for every i, j.*

Definition 10. *An IC allocation $\mu^{\cdot}$ is strictly individually rational (SIR) type by type if $U^i(\mu^i) > U^i(\mu_0^i)$.*

Assumption 1. *There exists a sequence of SIC and SIR type by type allocations $\{\mu_p^{\cdot}\}_{p=1}^{\infty}$, i.e., $V^i(\mu_p^i) > V^i(\mu_p^j)$ for every i, j, p and $U^i(\mu_p^i) > U^i(\mu_0^i)$ for every i, p, such that $\{V^i(\mu_p^i)\}_{p=1}^{\infty}$ converges to $V^i(\hat{\mu}^i(\mu_0^{\cdot}))$ for every i.*

Proposition 3. *Suppose that Assumption 1 is satisfied; then, for every equilibrium allocation of the three-stage game $\bar{\mu}^{\cdot}$, (7) is satisfied.*

Proof. I prove the result by contraposition. Suppose that Assumption 1 is satisfied. Consider an IC allocation $\bar{\mu}^{\cdot}$ in which for some type j, $V^j(\bar{\mu}^j) < V^j(\hat{\mu}^j(\mu_0^{\cdot}))$. Let $V^j(\hat{\mu}^j(\mu_0^{\cdot})) - V^j(\bar{\mu}^j) = \delta > 0$. The following lemma facilitates the proof.

Lemma 1. *There exists p_δ such that $V^j(\mu_p^j) \geq V^j(\bar{\mu}^j)$ for every $p \geq p_\delta$.*

Proof. Because there exists a sequence of SIC and SIR type by type allocations $\{\mu_p^{\cdot}\}_{p=1}^{\infty}$ such that $\{V^i(\mu_p^i)\}_{p=1}^{\infty}$ converges to $V^i(\hat{\mu}^i(\mu_0^{\cdot}))$ for every i, there exists p_δ such that $|V^j(\mu_p^j) - V^j(\hat{\mu}^j(\mu_0^{\cdot}))| < \delta$ for every $p \geq p_\delta$. Suppose that $V^j(\mu_p^j) > V^j(\hat{\mu}^j(\mu_0^{\cdot}))$ for some $p \geq p_\delta$.

Consider $\tilde{\mu}^{\cdot}$, where $\tilde{\mu}^i = \mu_p^i$ for $i = j$ and $\tilde{\mu}^i = \hat{\mu}^i(\mu_0^{\cdot})$ for $i \neq j$. Allocation $\tilde{\mu}^{\cdot}$ is IC and IR type by type which contradicts the definition of the RSW allocation. Therefore, $V^j(\mu_p^j) \leq V^j(\hat{\mu}^j(\mu_0^{\cdot}))$ for every $p \geq p_\delta$. Then, $V^j(\hat{\mu}^j(\mu_0^{\cdot})) - V^j(\mu_p^j) < \delta$ and hence $V^j(\hat{\mu}^j(\mu_0^{\cdot})) - V^j(\mu_p^j) < V^j(\hat{\mu}^j(\mu_0^{\cdot})) - V^j(\bar{\mu}^j)$, which is equivalent to $V^j(\mu_p^j) > V^j(\bar{\mu}^j)$. □

Consider mechanism $\mu_{p'_\delta}^{\cdot}$, where $p'_\delta > p_\delta$. Because this mechanism is SIC and SIR type by type, it provides the agent with a payoff strictly greater than the payoff he can obtain in the reservation allocation regardless of his beliefs. Therefore, if this mechanism is proposed by type j, it should be accepted by the agent; otherwise the equilibrium fails to be sequentially rational. Type j can achieve a higher payoff by proposing mechanism $\mu_{p'_\delta}^{\cdot}$ than by proposing $\bar{\mu}^{\cdot}$, which means that allocation $\bar{\mu}^{\cdot}$ cannot constitute an equilibrium allocation. □

The potential difficulty for the principal in obtaining the RSW allocation payoff arises from the specification of the posterior beliefs of the agent: if the agent holds unduly pessimistic beliefs about the principal's type, when he observes a deviation from some expected proposal of a mechanism, he rejects it. The sufficient condition of Proposition 3 allows the principal to overcome this obstacle. To see this, suppose that the agent expects the principal to propose a mechanism that provides a payoff to at least one type that is strictly lower than that in the RSW allocation. Suppose that this type proposes a mechanism with a strictly higher payoff. For the agent not to reject this mechanism, it must be that he is not worse off for all possible beliefs and for every possible continuation of the game. This is ensured if there is a sequence of SIC and SIR type by type allocations with payoffs converging to the payoffs in the RSW allocation. Note also that if the no-tangency condition is satisfied, as in MT92, then the sufficient condition provided in Proposition 3 is satisfied.[15] However, the sufficient condition of Proposition 3 might be satisfied even if the no-tangency condition is not satisfied, as the following example demonstrates.

Example 2. Suppose that $n = 2$ and $X = \mathbb{R}_+^2$. The principal is a seller and the agent is a buyer. Suppose that the outside option is zero. Let q denote the quantity (or quality) of the product and t the transfer from the agent to the principal. The cost of production for types 1 and 2 are $q(1+q)/3$ and q^2 respectively. The value of the product for the agent is

[15] Indeed, in the proof of Proposition 5 (p. 18), MT92 show that for every $\epsilon > 0$ small enough, there exists a "ϵ-perturbed RSW allocation", where "ϵ-perturbed" means that the allocation is SIC and SIR type by type and is ϵ-close to the RSW allocation.

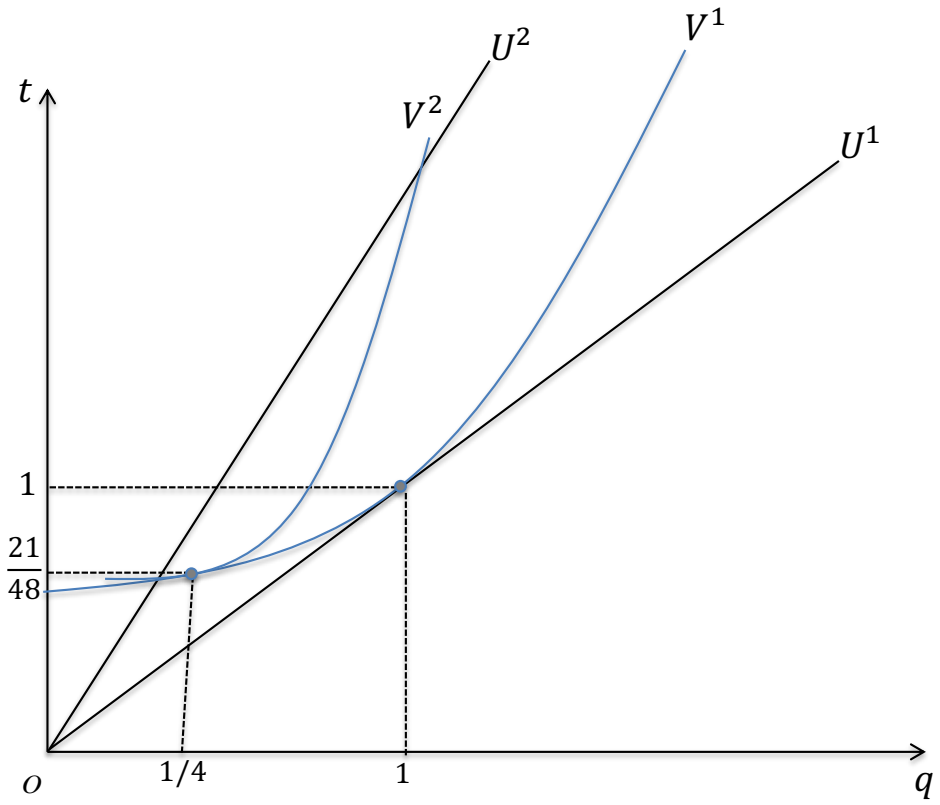

Figure 1

iq; hence, the agent values more a product from type 2. The payoffs of the two players are given below

$$V^1(t,q) = t - q(1+q)/3, \qquad V^2(t,q) = t - q^2$$

$$U^1(t,q) = q - t, \qquad U^2(t,q) = 2q - t$$

The payoffs in this example clearly violate the no-tangency condition. To see this, note that the marginal rates of substitution for the two types are $-V_q^1/V_t^1 = (1+2q)/3$ and $-V_q^2/V_t^2 = 2q$; hence the indifference curves are tangent at $q = 1/4$.

The RSW allocation is $\big((1,1),(21/48,1/4)\big)$. In the RSW allocation, type 1 suffers no distortion but type 2 underproduces relative to the case in which players are symmetrically informed.[16] The payoffs for type 1 and type 2 in the RSW allocation are $1/3$ and $9/24$. The RSW allocation is depicted in Figure 1, in which t is depicted in the vertical axis and q in the horizontal axis. Consider the sequence $\{\mu_p^{\cdot}\}_{p=1}^{\infty}$, where $\mu_p^1 = (1 - 1/p, 1)$

[16]To find the RSW allocation, one needs to first maximise $V^1(t,q)$ subject to $U^1(t,q) = 0$ and then maximise $V^2(t,q)$ subject to $U^2(t,q) \geq 0$ and $1/3 = V^1(t,q)$. The last constraint is the IC constraint satisfied as an equality.

and $\mu_p^2 = (21/48 - 2/p, 2)$. The payoffs for the two types of principal and the agent are, respectively

$$V^1(\mu_p^1) = 1/3 - 1/p, \qquad V^2(\mu_p^2) = 9/24 - 2/p$$
$$U^1(\mu_p^1) = 1/p, \qquad U^2(\mu_p^2) = 1/16 + 2/p$$

Note that $\{V^1(\mu_p^i)\}_{p=1}^{\infty}$ and $\{V^2(\mu_p^i)\}_{p=1}^{\infty}$ are both strictly increasing and converge to $1/3$ and $9/24$ respectively. Moreover, $\{U^1(\mu_p^i)\}_{p=1}^{\infty}$ and $\{U^2(\mu_p^i)\}_{p=1}^{\infty}$ are both strictly positive, strictly decreasing and converge to zero and $1/16$ respectively. Last, consider the payoff of type 1 from μ_p^2

$$V^1(\mu_p^2) = 1/3 - 2/p$$

and the payoff of type 2 from μ_p^1

$$V^2(\mu_p^2) = -1/p$$

It is only straightforward that $V^1(\mu_p^1) > V^1(\mu_p^2)$ and $V^2(\mu_p^2) > V^2(\mu_p^1)$ for every p; that is $\{\mu_p^{\cdot}\}_{p=1}^{\infty}$ is SIC and SIR type by type.

Theorems 1 and 2 alongside Proposition 3 provide a complete characterisation of the set of equilibrium allocations.

5 Conclusion

This paper reconsidered the general informed principal model with unilateral private information and common values.

Two implications arise from the analysis. First, in general environments, restriction to DRMs restricts the set of profitable deviations and therefore allows one to establish the existence of equilibrium, and characterise the properties of equilibrium allocations, under fairly general conditions. Second, in environments in which the RSW allocation is interim efficient relative to some non-degenerate beliefs, restriction to DRMs comes without loss of generality. This is for instance the case in stylised environments such as seller-buyer relationships, insurance environments, employer-employee relationships, etc.

A fruitful avenue for future research is to study environments with bilateral private information and interdependent values, i.e., both the principal and the agent possess payoff-relevant information. For instance, it might be that the buyer of an indivisible object has private information about his marginal valuation of the object as in Koessler and Skreta (2016), Balestrieri and Izmalkov (2016) and Koessler and Skreta (2017). One

question is how the results established in this paper extend to such bilateral private information settings.

References

Balestrieri, Filippo, and Sergei Izmalkov, "Informed seller in a Hotelling market," *Available at SSRN 2398258*, (2016).

Balkenborg, Dieter, and Miltiadis Makris, "An undominated mechanism for a class of informed principal problems with common values," *Journal of Economic Theory*, 157 (2015), 918–958.

Cella, Michela, "Informed principal with correlation," *Games and Economic Behavior*, 64 (2008), 433–456.

Cremer, Jaques, and Richard P. McLean, "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica*, 56 (1988), 1247–1257.

Fleckinger, Pierre, "Informed principal and countervailing incentives," *Economics Letters*, 94 (2007), 240–244.

Koessler, Frédéric, and Vasiliki Skreta, "Informed seller with taste heterogeneity," *Journal of Economic Theory*, 165 (2016), 456–471.

——, "Selling with Evidence," (2017).

Maskin, Eric, and Jean Tirole, "The principal-agent relationship with an informed principal: The case of private values," *Econometrica*, 58 (1990), 379–409.

——, "The Principal-Agent Relationship with an Informed Principal, II: Common Values," *Econometrica*, 60 (1992), 1–42.

Myerson, Roger, "Analysis of two bargaining problems with incomplete information," *Game-theoretic models of bargaining*, 73 (1985).

Myerson, Roger B., "Mechanism Design by an Informed Principal," *Econometrica*, 51 (1983), 1767–1797.

Mylovanov, Tymofiy, and Thomas Tröger, "Mechanism design by an informed principal: Private values with transferable utility," *The Review of Economic Studies*, 81 (2014), 1668–1707.

Severinov, Sergei, "An efficient solution to the informed principal problem," *Journal of Economic Theory*, 141 (2008), 114–133.

Skreta, Vasiliki, "On the informed seller problem: optimal information disclosure," *Review of Economic Design*, 15 (2011), 1–36.

Tan, Guofu, "Optimal procurement mechanisms for an informed buyer," *Canadian Journal of Economics*, 29 (1996), 699–716.

Yilankaya, Okan, "A note on the seller's optimal mechanism in bilateral trade with two-sided incomplete information," *Journal of Economic Theory*, 87 (1999), 267–271.

6 Appendix

Lemma 2. *If $\hat{\mu}_1^{\cdot}(\mu_0^{\cdot})$ and $\hat{\mu}_2^{\cdot}(\mu_0^{\cdot})$ are RSW allocations, where $\hat{\mu}_1^{\cdot}(\mu_0^{\cdot}) \neq \hat{\mu}_2^{\cdot}(\mu_0^{\cdot})$, then $V^i(\hat{\mu}_1^i(\mu_0^{\cdot})) = V^i(\hat{\mu}_2^i(\mu_0^{\cdot}))$ for every i.*

Proof. Suppose that $\hat{\mu}_1^{\cdot}(\mu_0^{\cdot})$ and $\hat{\mu}_2^{\cdot}(\mu_0^{\cdot})$ are RSW allocations, where $\hat{\mu}_1^{\cdot}(\mu_0^{\cdot}) \neq \hat{\mu}_2^{\cdot}(\mu_0^{\cdot})$ and $V^i(\hat{\mu}_1^i(\mu_0^{\cdot})) \neq V^i(\hat{\mu}_2^i(\mu_0^{\cdot}))$ for some i. Let $I_1 = \{i : V^i(\hat{\mu}_1^i(\mu_0^{\cdot})) \geq V^i(\hat{\mu}_2^i(\mu_0^{\cdot}))\}$ and $I_2 = \{i : V^i(\hat{\mu}_1^i(\mu_0^{\cdot})) < V^i(\hat{\mu}_2^i(\mu_0^{\cdot}))\}$. Because $\hat{\mu}_1^{\cdot}(\mu_0^{\cdot})$ and $\hat{\mu}_2^{\cdot}(\mu_0^{\cdot})$ are incentive compatible and individually rational type by type, allocation $\tilde{\mu}^{\cdot}$, which maps each type from I_1 to her outcome in allocation $\hat{\mu}_1^{\cdot}(\mu_0^{\cdot})$ and each type from I_2 to her outcome in allocation $\hat{\mu}_2^{\cdot}(\mu_0^{\cdot})$, is also incentive compatible and individually rational type by type. Allocation $\tilde{\mu}^{\cdot}$ dominates both $\hat{\mu}_1^{\cdot}(\mu_0^{\cdot})$ and $\hat{\mu}_2^{\cdot}(\mu_0^{\cdot})$, which contradicts the definition of an RSW allocation. □

ESSEC
BUSINESS SCHOOL

ESSEC Business School
3 avenue Bernard-Hirsch
CS 50105 Cergy
95021 Cergy-Pontoise Cedex
France
Tel. +33 (0)1 34 43 30 00
www.essec.edu

ESSEC Executive Education
CNIT BP 230
92053 Paris-La Défense
France
Tel. +33 (0)1 46 92 49 00
www.executive-education.essec.edu

ESSEC Asia-Pacific
5 Nepal Park
Singapore 139408
Tel. +65 6884 9780
www.essec.edu/asia

ESSEC | CPE Registration number 200511927D
Period of registration: 30 June 2017 - 29 June 2023
Committee of Private Education (CPE) is part of SkillsFuture Singapore (SSG)

ESSEC Africa
Plage des Nations - Golf City
Route de Kénitra - Sidi Bouknadel (Rabat-Salé)
Morocco
Tel. +212 (0)5 37 82 40 00
www.essec.edu

CONTACT

Centre de Recherche
Tel. +33 (0)1 34 43 30 91
research.center@essec.fr