Informed Seller with Bayesian Persuasion and Mechanism Design

Yanlin Chen

September 15, 2018

Abstract

This paper investigates how a privately informed seller can signal her type through Bayesian persuasion and mechanism design. We first fully characterize the RSW (Rothschild-Stiglitz-Wilson) mechanism, which is the least cost separating equilibrium. In this mechanism, the low-type seller acts as if her type was known and the high-type seller tries to prevent the low type from mimicking. The high type discloses to the buyer whether his value is above a cutoff, subsidizes the buyer before information disclosure and sets price equal to expected value conditional on that value is above the chosen cutoff and the type is high. She randomizes on one or two packages of different cutoffs and prices. The RSW mechanism is the unique equilibrium mechanism if and only if the low type is relatively close to the high type and is with relatively high proportion. Finally, when there are multiple equilibria, the RSW mechanism survives intuitive criterion and is the unique one when the low type is relatively close to the high type.

Keywords: Bayesian persuasion, mechanism design, informed principal.
JEL Classifications: D11, D82, D83, L12

1 Introduction

In many selling situations, buyers (he) are initially uncertain about a product's value and sellers (she) often control how much new information to disclose to buyers. For example, the value of tutorial lesson to a student depends on how his preference matches with the tutor's teaching style. How much a consumer likes a new snack depends on how his taste matches with the snack's characteristics. Sellers can offer free courses and samples so that consumers can learn through trial. Meanwhile, sellers are likely to have an information advantage over buyers. For instance, the tutor's ability is private information. The firm produces the snack and therefore is better informed. The same situation applies to software, games, movie, vehicle, technology, insurances, financial products, etc. [1]

This paper investigates how an informed seller signals her type through information design and mechanism design simultaneously. In this paper, an informed seller sells one unit of indivisible object to a buyer. The quality of the product is either high or low. If the quality is high, the seller is more likely to provide the buyer with a higher value. In a sense, the product is both vertical and horizontally differentiated: different type of seller has different quality and given the same quality the buyer may have different value for the product. The seller can disclose information about the buyer's value through Bayesian persuasion as in Kamenica and Gentzkow [12], and can adopt any selling scheme to sell the product.

Following the literature on informed principal problem, the seller proposes a contract on information disclosure and selling scheme, which the buyer accepts or refuses. The contract is executed if and only if the buyer accepts. Any equilibrium of this game is a feasible direct mechanism which is incentive compatible and individual rational for both types of the seller and the buyer. Since both the information disclosure and selling scheme are complex, our first result is to simplify the problem by showing that any feasible mechanism can be replicated by a linear combination of different packages of disclosure policy and deterministic selling scheme. As a result, it is without loss of generality to focus on disclosure policies with at most two signals.

We then characterize RSW (Rothschild-Stiglitz-Wilson) mechanism introduced by Maskin and Tirole [17]. Among all safe mechanisms that the buyer would participate independent of his belief, it achieves the maximal revenue for each type of the seller. It is the least cost separating equilibrium. In the RSW mechanism, the low type sells to the buyer with certainty as if her type was known. The high type discloses only to the buyer whether the value is above a cutoff by monotone binary partition. She pays the buyer before the signal is realized and sets a price equal to expected value provided that value is above the chosen cutoff and that the seller is of high type. She randomizes on different packages of monotone binary partition, posted price and information fee to deter the low type from mimicking.

Depending on the distinction in types of the seller, the high type has two strategies. When

[1] See Li and Shi [15], Chen and Zhang (2018) for more detailed examples.

the low type generates low value for the buyer, the high type would pay the buyer with a positive information subsidy. When the low type is close to the high type, high-type seller could assign positive probability to one or two packages. She charges zero information fee and leaves zero surplus to the buyer. These two cases resemble real-life examples. Open course to the public is common in tutorial market to attract new students, frequently accompanied by small free gifts like a pen or lecture notes. They encourage the buyer with subsidy to attend open courses before making purchase decision. In this market, quality of the product, tutorial, varies. In a more regulated and homogeneous market, for instance, snack, firms often compete with price and give out free samples. Firms mainly stick to one selling price and sample distribution policy. Occasionally they offer limited-time discounts and leave the buyer with less time to make decision.

We conduct comparative statics. We fix the high type and increase the low type. High-type seller's revenue and social welfare increases in the RSW mechanism. By contrast, the buyer's exante payoff decreases as information subsidy from the high type decreases. This implies that a restriction on the lowest quality faces a trade-off between social welfare and buyers' payoff.

The RSW mechanism is important and intuitive. It provides the lower bound of payoff for each type of the seller. Therefore it equals to the set of equilibrium mechanism if and only if it is not Pareto dominated. We provide sufficient and necessary condition for the RSW mechanism to be the unique equilibrium mechanism such that the low type is relatively close to the high type and the proportion of the low type is relatively high. When the RSW mechanism is not unique equilibrium mechanism, we impose intuitive criterion to select equilibria. We show that the RSW mechanism survives the refinement. When the low type is relatively close to the high type, it is the unique surviving mechanism.

The rest of the paper is organized as follows. In Section 2, we conduct a literature review. In Section 3, we describe the model. In Section 4, we simplify the analysis by restricting to binary disclosure and deterministic selling scheme. In Section 5, we characterize the RSW mechanism. In Section 6, we provide sufficient and necessary condition under which the RSW mechanism is the unique equilibrium mechanism. In section 7, we impose intuitive criterion to select equilibria. In section 8, we conclude.

2 Literature review

A companion work by Chen and Zhang (2018) analyzes signalling consideration of a privately informed seller through Bayesian persuasion and a take-it-or-leave-it price. We show that there is a unique intuitive equilibrium which is the least cost separating equilibrium. By contrast, in this paper, the seller is free to choose any selling scheme. Since the seller's action space is extended, previous result of intuitive equilibrium does not apply in general. The high type makes higher revenue in intuitive equilibrium and the intuitive equilibrium may no longer be unique. We provide sufficient condition under which the setting in Chen and Zhang (2018) is without loss of generality.

This paper is mainly related to two strands of literature: Information Disclosure of a privately informed sender and Mechanism Design by an informed principal. In intersection of this two literature, Skreta [25] studies an informed seller who observes a signal of each buyer's value. Before choosing the selling scheme, the seller decides whether to inform each buyer of his competitors' signal. In this paper, the seller design information disclosure about the buyer's value and selling scheme simultaneously to signal her private type. Koessler and Skreta [13], [14] show that an informed principal does not benefit from certification technology with respect to her own type and exante optimal allocation can be supported as an equilibrium. In their setting, information disclosure is about the seller's type and is contractible. In contrast, we allow the seller to disclose information about the buyer's value. Therefore, the buyer's private information is endogenously determined by information disclosure policy while in above two papers, the buyer exogenously learns his private type. Moreover, instead of exante optimization, we focus on the RSW mechanism, which is optimal for each type of the seller conditional on buyer's full-information participation constraint. Exante efficient mechanism may not be an equilibrium in our setting.

First, this paper models information disclosure as Bayesian persuasion introduced by Kamenica and Gentzkow [12] which is then extended to several directions. This paper is mostly related to the literature on Bayesian persuasion with informed sender. Perez-Richet [22] considers a perfectly informed sender and demonstrates that it is without loss of generality to focus on pooling equilibria. Alonso and Camara [1] compare revenue of an informed sender and argue that the sender does not benefit from private information. Hedlund [10] studies an imperfectly informed sender who signals her type through Bayesian persuasion and selects equilibria with D1 criterion. These papers discuss Bayesian persuasion as the main signaling tool. In this paper, the action space is two-dimensional: Bayesian persuasion and selling scheme. After information is disclosed to the buyer, the seller provides incentive to him to report realized signal truthfully. This paper adopts an approach of informed principal. This approach restricts the equilibrium set since the seller has more commitment power.

Second, this paper is related to the literature on Mechanism design by an informed principal following Myerson [18], Maskin and Tirole [16], [17]. Myerson [18] introduces the inscrutability principle which states that it is without loss of generality to focus on the pooling equilibria. He also introduces the concept of safe allocations which is incentive compatible and individually rational for any belief about the principal's type. Maskin and Tirole [16], [17] model mechanism selection by a privately informed principal as a three-stage noncooperative game with private and common value, respectively. In set up of private value where the seller's type does not directly affect the buyer's utility, Maskin and Tirole [16] discuss whether the principal benefits from private information. Mylovanov and Troger [19], [20] discuss existence and properties of strongly neologism-proof allocations. This paper considers a selling problem which shares a feature of common value since the seller's private information is correlated with the buyer's value. Maskin and Tirole [17] introduce Rothschild-Stiglitz-Wilson (RSW) allocation which yields the optimal revenue for each type of principal among safe allocations. They demonstrate that any equilibrium (weakly) Pareto dominates the RSW allocation. Therefore, RSW allocation is the unique equilibrium outcome if and only if it is interim efficient. Cella [6] shows that the principal can extract higher surplus from the agent when the private information of both types are correlated. Balkenborg and Makrisz [4] propose assured allocation and discuss its relation with RSW allocation. Bedard [5] discusses whether and when the principal benefits from private information. The informed principal problem is also considered in a variety of settings, such as bilateral trading (Yilankaya [27], Nishimura [21]), procurement contracting (Tan [26], Balestrieri [2]), social decision making (Severinov [23]), collusion (Francetich and Troyan [9]) and Hotelling market (Balestrieri and Izmalkov [3]). This paper studies an informed seller who design both information disclosure about the buyer's value and selling scheme. We focus on the RSW mechanism and show that it can be supported as an equilibrium. Then we provide sufficient and necessary condition under which any equilibrium is the RSW mechanism.

This paper is also related to the literature that the seller designs information disclosure and selling scheme jointly. Eso and Szentes [8] develop an orthogonal decomposition technique to transform any additional information into shocks independent from the buyer's original belief. They identify conditions under which the optimal mechanism can be implemented by full disclosure in a handicap auction. Hoffmann and Inderst [11] discuss the seller's incentive to choose costly information disclosure among an ordered set of signal technologies. The seller discriminates the buyer with price as well as the quality of information. Li and Shi [15] investigate the situations in which discriminatory information disclosure generates higher revenue than full disclosure. These papers consider a privately informed buyer in screening setup while we study a signalling problem with an informed seller. In this paper, the buyer has no initial information but endogenously learn signals about his value. In contrast to maximizing the revenue of the seller, we characterize the RSW mechanism. In other words, we solve the optimal information design and mechanism design when the seller fully separates.

3 The model

There is a risk neutral seller (she) and a risk neutral buyer (he) for one unit of indivisible product. The value of the product to the seller is normalized to 0. The seller has private and unverifiable information θ that affects the distribution of the buyer's value of the product. This private information is characterized by a binary distribution on $\{L, H\}$, with probability μ_H^0 and $\mu_L^0 = 1 - \mu_H^0$, respectively. To focus on the informed seller problem, we model that the buyer does not have private information at the beginning. The buyer values the product at v. When the seller's type is θ, the buyer's value v follows an atomless distribution with $c.d.f$ $F_\theta(v)$ and $p.d.f$ $f_\theta(v)$ on the common support $[0, \overline{v}]$. We assume that $F_\theta(v)$ satisfies the monotone likelihood ratio property: $\frac{f_H(v)}{f_L(v)} > \frac{f_H(v')}{f_L(v')}, \forall v > v'$. This implies that the buyer is willing to pay more if the seller is of type H; thus, we can refer to type θ as the quality of the seller. To summarize, the product exhibits both horizontal and vertical differentiation. Bayes' rule implies that the buyer's prior belief about the value of the product is $f(v) = \mu_H^0 f_H(v) + \mu_L^0 f_L(v)$.

The grand game and equilibrium concept

We adopt the informed principal methodology and model the game as a grand game. The seller proposes a contract. Then the buyer decides whether to participate. If the buyer does not participate, the game ends and both the seller and the buyer get the outside option payoff of 0; otherwise, the contract is executed. A contract, M, is a continuation game that specifies message space for both the buyer and the seller, and implements actions based on their announcements. The action space of the seller is the set of packages of information disclosure, $\pi \in \Pi$, and selling scheme, Γ.

This paper studies **Perfect Bayesian Equilibrium (PBE)**. A PBE of the game specifies (i) for each type of the seller, an optimal contract, (ii) for each contract, a belief about the seller's type using Bayes's rule.

By revelation principle, we focus on direct mechanism in which the seller reports her type and the buyer reports her observed signal. Following inscrutability principle by Myerson [18], we focus on pooling equilibria where each type of the seller proposes the same mechanism.

The seller's action

We model information disclosure as Bayesian persuasion following Kamenica and Gentzkow [12]. A disclosure policy is a costless signal technology π which is a family of conditional distribution $\pi(s|v)$ over a finite set of realization space S such that $\sum_{s\in S}\pi(s|v) = 1, \forall v$.[2] We model that when the seller discloses information through Bayesian persuasion, the realized signal is only observable to the buyer. For instance, after buyers experience the trial course, sellers can hardly tell whether they likes it or not.[3] Note that, we do not distinguish disclosure policies that are different in zero measure.

For the selling scheme, as standard, we focus on direct selling scheme in which the buyer reports observed signal to the seller. Given the report, the seller assigns the product to the buyer with probability of $q(s)$ and charges $t(s)$. Thus, the selling scheme can be represented by $\Gamma = (q(s), t(s))_{s\in S}$.To summarize, a mechanism is a distribution of packages of information disclosure policy and selling scheme, $\sigma : \Theta \to \Delta(\Pi \times \Gamma)$, such that

$$\sum_{(\pi,\Gamma)} \sigma(\pi, \Gamma|\theta) = 1, \forall\theta. \tag{1}$$

There are three stages in this mechanism. In the first stage, the seller reports her type $\widetilde{\theta}$ and (π, Γ) is implemented with probability $\sigma(\pi, \Gamma|\widetilde{\theta})$. In the second stage, a signal s is generated by π and is privately observed by the buyer. In the third stage, the buyer reports his realization of the signal $\widetilde{s}$ and gets the object with probability $q(\widetilde{s})$ and pays $t(\widetilde{s})$.

The buyer's belief and action

The buyer's action is constituted of two parts. First, he decides whether to participate. Second, he makes report given realized signal s.

Observing a mechanism σ, the buyer believes that the seller is type-θ with probability

[2] Here we assume that the seller's type is not verifiable and she cannot disclose her private information directly from information disclosure. Otherwise, separating arrives trivially.

[3] In the standard Bayesian persuasion literature, whether or the sender learns the realization of the signal or not does not matter since the sender does not make other decisions in the model.

μ_θ. In pooling equilibrium, the buyer believes that the seller is high type with probability $\mu_H = \mu_H^0$. The buyer decides whither to participate in the mechanism. Denote $\varphi(\sigma, \mu) = 1$ if the buyer participates; otherwise $\varphi(\sigma, \mu) = 0$. If the buyer does not participate, the game ends.

If the buyer participates, the seller implements (π, Γ) given her own report. The buyer updates his interim belief conditional on (π, Γ) such that the seller is high type with probability

$$\lambda(\pi, \Gamma|\sigma) = \frac{\mu_H \sigma(\pi, \Gamma|H)}{\sum_\theta \mu_\theta \sigma(\pi, \Gamma|\theta)}. \tag{2}$$

Therefore, the buyer believes that the value follows distribution

$$F_{\lambda(\pi,\Gamma|\sigma)}(v) = \lambda(\pi, \Gamma|\sigma)F_H(v) + (1 - \lambda(\pi, \Gamma|\sigma))F_L(v). \tag{3}$$

Interim belief λ matters in two channels. In the first channel, it affects the belief about distribution of realized signal. Denote the probability of generating signal s from the buyer's perspective with interim belief λ,as

$$p_\lambda(s) = \int_0^{\overline{v}} \pi(s|v) f_\lambda(v) dv. \tag{4}$$

Similarly, from type θ's point of view,

$$p_\theta(s) = \int_0^{\overline{v}} \pi(s|v) f_\theta(v) dv. \tag{5}$$

Note that $p_H(s) = p_1(s), p_L(s) = p_0(s)$. In the second channel, it affects posterior expected value given realized signal. Denote the posterior expected value as

$$r(s, \lambda) = \frac{\int_0^{\overline{v}} \pi(s|v) v f_\lambda(v) dv}{\int_0^{\overline{v}} \pi(s|v) f_\lambda(v) dv}. \tag{6}$$

By abuse of notation, denote $s < s'$ if $r(s, \lambda) < r(s', \lambda)$ or $r(s, \lambda) = r(s', \lambda)$, $q(s) < q(s')$. From a player's perspective who has interim belief λ'[4], denote the distribution of s be

$$P_{\lambda'}(s, \lambda) = \int_{\tilde{s} \leq s} p_{\lambda'}(s) ds. \tag{7}$$

[4]Since we allow both continuous and discrete disclosure where there are finite number of signals, the distribution of signal s, $P_{\lambda'}(s, \lambda)$ may have mass points and is right continuous in s. Note that $P_{\lambda'}(s, \lambda)$ have the same mass points regardless of λ'.

Payoff and feasibility

There are two types of feasibility constraints for mechanism: incentive compatibility and participation constraints. A mechanism σ is feasible with belief μ if and only if it is incentive compatible for both the seller and the buyer, and individual rational for the buyer.

In the third stage, given $\Gamma = (q(s), t(s))_{s\in S}$ and on-path interim belief λ, a buyer with expected value s who reports $\widetilde{s}$ gets utility $q(\widetilde{s})r(s,\lambda) - t(\widetilde{s})$. Particularly, a truthful buyer in the second stage gets $U(s,\lambda|\Gamma) = q(s)r(s,\lambda) - t(s)$. Then IC constraints require that

$$U(s,\lambda|\Gamma) = q(s)r(s,\lambda) - t(s) \geq q(\widetilde{s})r(s,\lambda) - t(\widetilde{s}), \forall s, \widetilde{s}. \qquad (IC_A)$$

$\forall s$, $\{r(s,\lambda)\}_{s\in S} \subset [0,\overline{v}]$. By Skreta [24], we can always construct a direct selling scheme that extends the posterior expected value space to $[0,\overline{v}]$. Denote $U(0|\Gamma)$ be the utility when a signal induces expected value of 0. Therefore, by standard arguments,

$$q(s) \text{ is weakly increasing in } r(s,\lambda), \qquad (IC_{A1})$$

$$U(s,\lambda|\Gamma) = \int_{\widetilde{s}\leq s} q(s)dr(\widetilde{s},\lambda) + U(0|\Gamma).^5 \qquad (IC_{A2})$$

Denote $w(\Gamma) = -U(0|\Gamma)$. Then $w(\Gamma) \geq 0$ if and only if $U(0|\Gamma) \leq 0$. We can interpret w as information fee (subsidy) when $w > 0$ (≤ 0) since regardless of realized signal the buyer bears the burden of w. When the seller designs $U(0|\Gamma)$, she is choosing w. The buyer's interim payoff of the second stage with on-path interim belief λ is

$$\begin{aligned} U(\lambda|\pi,\Gamma) &= \int_{s\in S} U(s,\lambda|\Gamma)dP_\lambda(s,\lambda) \\ &= \int_{s\in S} q(s)(1-P_\lambda(s,\lambda))dr(s,\lambda) - w(\Gamma). \end{aligned} \qquad (8)$$

The payoff of a buyer who does not participate is normalized to 0. Therefore, exante IR constraint with belief μ requires that $\varphi(\sigma,\mu) = 1$ if and only if

$$U(\sigma,\mu) = \sum_{\theta} \mu_\theta \sum_{(\pi,\Gamma)} \sigma(\pi,\Gamma|\theta)U(\lambda(\pi,\Gamma|\sigma)|\pi,\Gamma) \geq 0. \qquad (IR_A^\mu)$$

Full-information IR constraint requires that when θ is observable, the buyer makes nonnegative payoff: $\forall\theta$,

$$\sum_{(\pi,\Gamma)} \sigma(\pi,\Gamma|\theta)U(\theta|\pi,\Gamma) \geq 0. \qquad (IR_A^\theta)$$

A mechanism is safe if IC and full-information IR constrains hold.

Given (π, Γ) and λ, the seller makes revenue of

$$\begin{aligned} & R_\theta(\pi, \Gamma, \lambda) \\ = & \int_{s\in S} t(s)dP_\theta(s,\lambda) \\ = & \int_{s\in S} q(s)(r(s,\lambda)dP_\theta(s,\lambda) - (1-P_\theta(s,\lambda))dr(s,\lambda)) + w(\Gamma). \end{aligned} \tag{9}$$

Denote $R_\theta(\sigma, \theta', \mu) = \sum_{(\pi,\Gamma)} \sigma(\pi, \Gamma|\theta')R_\theta(\pi, \Gamma, \lambda(\pi, \Gamma|\sigma))$, which is the expected revenue of θ who reports θ'. IC constraint requires that

$$R_\theta(\sigma, \theta, \mu) \geq R_\theta(\sigma, \theta', \mu), \forall \theta, \theta'. \tag{$IC_{\theta,\theta'}$}$$

The seller gets 0 if the buyer does not participate. Then the seller with type θ makes revenue of

$$R_\theta(\sigma, \mu) = \varphi(\sigma, \mu)R_\theta(\sigma, \theta). \tag{10}$$

By abuse of notation, with the prior μ^0, denote $R_\theta(\sigma, \theta) = R_\theta(\sigma, \theta, \mu^0), R_\theta(\sigma) = R_\theta(\sigma, \mu^0), U(\sigma) = U(\sigma, \mu^0)$.

Timing

The timing of the game is as follows:

- The nature draws a private type for the seller.
- The seller chooses a mechanism σ.
- The buyer observes the seller's choice, and decides whether to participate.
- If not, the game ends; otherwise, the mechanism is implemented and there are three stages.
- In the first stage, the seller reports her type θ and (π, Γ) is adopted with probability $\sigma(\pi, \Gamma|\theta)$.
- In the second stage, s is generated according to π. The buyer observes s.
- In the third stage, the buyer reports s and is assigned the product with probability of $q(s)$ at price of $t(s)$.

4 Simplifying analysis

Although we focus on direct mechanism, the seller's action space is very large. However, many mechanisms are outcome equivalent. In this paper we do not distinguish those mechanisms that generate the same outcome.

Definition 1 *Two feasible mechanisms with belief μ are outcome equivalent if they generate the same expected revenue for each type of the seller, the same expected payoff and expected allocation probability for each expost value of the buyer.*

In general, a seller could assign a product to a buyer with any probability between zero and one. A deterministic selling scheme, Γ^d is where for any signal s, $q(s) \in \{0, 1\}$. There is no uncertainty after the buyer reports a signal. We will show that it is without loss of generality to focus on deterministic selling scheme. Given disclosure policy and interim belief, for any selling scheme, there exists a mixture of deterministic selling scheme which generates the same payoff for each type of the seller. Moreover, it is outcome equivalent in the sense that exante allocation and payoff of the buyer given expost value remains the same. Instead of randomizing in a selling scheme, we can randomize in mechanism with respect to deterministic selling scheme. Consequently, binary partition with $s \in \{s_1, s_2\}$ is sufficient in terms of information disclosure. By notation, we refer s_1 (s_2) to be the signal such that $q(s_1) = 1$ $(q(s_2) = 0)$.

Proposition 1 *For any mechanism σ that is feasible with belief μ, there exists an outcome-equivalent mechanism σ^d, that is feasible with belief μ, whose support is a subset of deterministic selling scheme and binary partition.*

From now on, we focus on deterministic selling scheme and binary partition with $s \in \{s_1, s_2\}$. Then for each type of the seller,

$$R_\theta(\pi, \Gamma, \lambda) = p_\theta(s_1)\, t(\Gamma) + w(\Gamma), \tag{11}$$

where $r(s_2, \lambda) \leq t(\Gamma) \leq r(s_1, \lambda)$. Note that we distinct deterministic selling scheme from a (t, w), in which t is a posted price and w is an information fee. The buyer pay the information fee to learn a signal, which could be either negative or positive. When it is negative, the seller subsidizes the buyer to learn realized signal. After the buyer observes the signal, he decides whether to purchase the product at price t. Purchase occurs if and only if realized signal leads to interim expected value that is higher than price t, that is, $r(s,\lambda) \geq t$. For a (t,w), we assume the buyer purchases the product whenever he is indifferent while in deterministic mechanism, the principal decides whether to allocate the product or not. This assumption has no impact on the buyer's utility but generates higher revenue for the seller. A deterministic selling scheme is equivalent to a (t,w) if and only if either there is only one signal or the expected value of two signals are different with $r(s_1,\lambda) > r(s_2,\lambda)$.

We introduce a class of disclosure policy that is of significance: monotone binary partition with threshold y, denoted by $\pi^M(y)$. The seller only informs the buyer whether his value is above the threshold or not with

$$\pi(s_1|v) = \begin{cases} 1, & \text{if } v \geq y; \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

When $y = 0$, it reduces to no disclosure which generates the same signal for each value. Denote $\pi^N(s_1|v) = 1, \forall v$. The buyer cannot induce any information from realized signal. For any interim belief λ, any (t,w) is equivalent as long as $t+w$ is a constant and $t \leq r(s_1,\lambda)$ such that the buyer purchases if and only if signal s_1 is realized. It is without loss of generality to focus on $(r(s_1,\lambda), w)$ when $\pi = \pi^N$.

5 The RSW mechanism

In this section, we characterize properties of the RSW (Rothschild-Stiglitz-Wilson) mechanism introduced by Maskin and Tirole [17]. It corresponds to Riley outcome in signalling literature. It is the maximal revenue for each type among safe mechanisms given the buyer's full-information IR constraint and IC constraints of both the buyer and the seller.

Definition 2 *A mechanism σ with associated payoffs $\{R^W_\theta\}_{\theta\in\Theta}$ is a RSW mechanism if and only if*

$$R^W_\theta = R_\theta(\sigma^W) = \max_{\widetilde{\sigma}} \sum_{(\pi,\Gamma)} \widetilde{\sigma}(\pi,\Gamma|\theta) R_\theta(\pi,\Gamma,I_{\theta=H}), \tag{R^W_θ}$$

$$s.t. IC_A \quad , IC_{\theta',\theta''}, IR^{\theta'}_A, \forall\theta',\theta''.$$

6

[6] An alternative definition is R^W_θ

By continuity in σ, π and Γ, a RSW mechanism exists. Note that we do not exclude the possibility of mixed action for any type of the seller. Best safe mechanism is not necessarily unique although the payoff is unique for each type of the seller. Now we characterize revenue and strategy for each type separately.

5.1 The low type

Under full-information IR constraint, the low type cannot do better than the case when her type is public information. It is optimal to maximize welfare and extract the full surplus. She would sell the product to the buyer regardless of his value. Thus, the outcome is efficient.

Proposition 2 *In a RSW mechanism, the low type makes revenue* $R_L^W = \int_0^{\overline{v}} v f_L(v)dv$, *and implements* (π, Γ) *with positive probability, that is,* $\sigma^L(\pi, \Gamma|L) > 0$, *only if*

$$\pi = \pi^N, \tag{13}$$

$$\Gamma = \left(r\left(s_1, 0\right), w\right), \tag{14}$$

where $\sum_{(\pi,\Gamma)} \sigma^L(\pi, \Gamma|L)w = 0$.

We construct one package, $(\pi, (t, w))$, that yields revenue for low-type-seller that equals R_L^W. The seller chooses no disclosure and sets a posted price equal to expected value provided that type is low. There are linear combinations of packages that can achieve the same goal. However, all these linear combinations are equivalent in the sense that, for each ex-post value v, the buyer gets the product with certainty and pays the same amount in total.

5.2 The high type

For the high type, analysis is much more complicated. The main challenge stems from $IC_{L,H}$. The low type has incentive to pretend as the high type to charge a higher payment. The high type must give up some efficiency to signal her type. First, we show that in a RSW mechanism, the high type adopts monotone binary partition, sets price equal to expected

$= \max_{\widetilde{\sigma}} \sum_{(\pi,\Gamma)} \widetilde{\sigma}(\pi, \Gamma|\theta) R_\theta(\pi, \Gamma, \lambda(\pi, \Gamma|\widetilde{\sigma}))$, $s.t. IR_A^{\theta'}, IC_{\theta',\theta''}, \forall \theta', \theta''$. It generates the same mechanism. Note that our definition is different from that in Maskin and Tirole [17]. In their setup, the buyer's belief only matters in IR constrain, while in our setting, we also have IC constraint for the buyer in the selling mechanism.

value and information subsidy. Second, we characterize high-type seller's revenue in a RSW mechanism. Then we provide strategies that guarantee her this revenue. Third, we discuss uniqueness of strategy.

Lemma 1 *Consider any feasible mechanism σ with belief μ such that $R_L(\sigma^o, \mu) \leq \int_0^{\overline{v}} v f_{\mu_H}(v)dv$. There exists a feasible mechanism σ^o with belief u such that*

(a) $\sigma^o(\pi, \Gamma|L) > 0$ *only if* $\pi = \pi^N$, $\Gamma = (r(s_1, 0), R_L(\sigma, \mu) - r(s_1, 0))$;

(b) $\sigma^o(\pi, \Gamma|H) > 0$ *only if* $\pi = \pi^M(y)$ *for some threshold* y, $\Gamma = (r(s_1, 1), w)$ *for some* w;

(c) $R_L(\sigma^o, \mu) = R_L(\sigma, \mu)$, $R_H(\sigma^o, \mu) \geq R_H(\sigma, \mu), U(\sigma^o, \mu) \geq U(\sigma, \mu)$.

Moreover, if condition (b) *does not hold for* σ, *there exists* σ^o *that satisfies* $(a - c)$ *such that* $R_H(\sigma^o, \mu) > R_H(\sigma, \mu)$.

This lemma shows that the high type strictly benefits from monotone binary partition when the low type is indifferent. For any monotone binary partition, $\pi^M(y)$, the updated expected value of signal s_1 is higher than that of s_2, that is, $r(s_1, \lambda) > r(s_2, \lambda)$. Therefore, any deterministic selling scheme can be interpreted as a package of posted price and information fee, (t, w). Moreover, she prefers to separate from the low type to set price equal to expected value given that the type is high and the value is above the chosen cutoff.

By Proposition 2, in the RSW mechanism, the low type makes revenue of $\int_0^{\overline{v}} v f_L(v)dv$, which is lower than $\int_0^{\overline{v}} v f(v)dv$. Therefore, we can apply Lemma 1 to show the following proposition that in a RSW mechanism, to maximize profit, the high type randomizes on packages of monotone binary partition with different cutoffs, price equal to expected value given signal s_1 and that type is high. Otherwise, condition (b) does not hold, and the high type can make a strictly higher payoff, which contradicts the maximization of high-type seller's revenue in the RSW mechanism. Since the high type extracts full surplus with price, by full-information participation constraint, she offers information subsidy to the buyer.

Proposition 3 *In a RSW mechanism, the high type implements* (π, Γ) *with positive probability, that is,* $\sigma^W(\pi, \Gamma|H) > 0$, *only if for some threshold* y,

$$\pi = \pi^M(y), \tag{15}$$

$$\Gamma = (r(s_1, 1), w). \tag{16}$$

Moreover, $\sum_{\pi,\Gamma} \sigma^W(\pi, \Gamma|H)w(\Gamma) \leq 0$.

This proposition provides necessary conditions of high-type seller's strategy in a RSW mechanism. For any package that the high type implements with positive probability, that is, $\sigma^W(\pi,\Gamma|H)>0$, only two variables matter: cutoff of monotone binary partition, y, and information fee, w. By abuse of notation, we can write $\sigma^W(\pi,\Gamma|H)=\sigma^W(y,w|H)$. Now we can rewrite the problem for the high type:

$$R_H^W = R_H(\sigma^W) = \max_{\widetilde{\sigma}} \sum_{(y,w)} \widetilde{\sigma}(y,w|\theta)\left(\int_y^{\overline{v}} v f_H(v)dv + w\right), \tag{17}$$

$$s.t. \quad \sum_{(y,w)} \widetilde{\sigma}(y,w|\theta)\left(R_L(y)+w\right) \le R_L^W, \tag{18}$$

$$\sum_{(y,w)} \widetilde{\sigma}(y,w|\theta) w \le 0, \tag{19}$$

where $R_L(y) = \frac{\int_y^{\overline{v}} v dF_H(v)}{\int_y^{\overline{v}} dF_H(v)} \int_y^{\overline{v}} dF_L(v)$. It is low-type seller's revenue if she pretends to be high type who adopts a monotone binary partition with threshold y, a price equal to expected value conditional on known as the high type and threshold is above chosen threshold. It is strictly decreasing in y. Since the objective function and both constraints are linear in w, it is equivalent to

$$R_H^W = R_H(\sigma^W) = \max_{\widetilde{\sigma},\widetilde{w}} \sum_{y} \widetilde{\sigma}(y|\theta) \int_y^{\overline{v}} v f_H(v)dv + \widetilde{w}, \tag{20}$$

$$s.t. \quad \sum_{y} \widetilde{\sigma}(y|\theta) R_L(y) + \widetilde{w} \le R_L^W, \tag{21}$$

$$\widetilde{w} \le 0. \tag{22}$$

The rest is to characterize the support of threshold and exante information fee. Targeting at maximal revenue, high-type seller maximizes revenue from each action with positive probability. This problem can be separated into point maximization problems with R_L:

$$\max_{\widetilde{y},\widetilde{w}} \int_{\widetilde{y}}^{\overline{v}} v f_H(v)dv + \widetilde{w}, \tag{23}$$

$$s.t. \qquad R_L(\widetilde{y}) + \widetilde{w} \le R_L, \tag{24}$$

$$\widetilde{w} \le 0. \tag{25}$$

It can be interpreted as maximization of revenue difference to low-type seller with existence of information subsidy. Whenever the low type makes higher revenue by mimicking than her original one, R_L, the high type can increase information subsidy to reduce mimicking revenue. This further relaxes the buyer's participation constraint. Therefore, we can combine two constraints into one and the problem is equivalent to:

$$D_H(R_L) = \max_{\widetilde{y}} \int_{\widetilde{y}}^{\overline{v}} v f_H(v)dv - R_L(\widetilde{y}) \qquad \text{(Problem I)}$$

$$s.t. \qquad R_L(\widetilde{y}) \geq R_L.$$

Denote the solution be $\underline{y(R_L)}$.[7] By continuity, for any $R_L \in [0, \int_0^{\overline{v}} v f_H(v)dv]$, such $\underline{y(R_L)}$ exists and is unique. We introduce an important indicator. Denote $y_H^* = \underline{y(0)}$. It is the threshold that maximizes difference in revenue without constraint on low-type seller's mimicking revenue. We assume y_H^* uniquely maximizes $\int_y^{\overline{v}} v f_H(v)dv - R_L(y)$.[8] That is, for any $y \neq y_H^*$,

$$\int_y^{\overline{v}} v f_H(v)dv - R_L(y) < \int_{y_H^*}^{\overline{v}} v f_H(v)dv - R_L(y_H^*). \tag{26}$$

Since $\int_y^{\overline{v}} v f_H(v)dv - R_L(y) = 0$ at $y = 0, \overline{v}$, there exists an interior solution. Therefore, first-order condition holds:

$$(f_L(y_H^*) - f_H(y_H^*)) \frac{\int_{y_H^*}^{\overline{v}} v f_H(v)dv}{\int_{y_H^*}^{\overline{v}} f_H(v)dv} + \frac{(F_L(y_H^*) - F_H(y_H^*))f_H(y_H^*)}{\int_{y_H^*}^{\overline{v}} f_H(v)dv} \left(\frac{\int_{y_H^*}^{\overline{v}} v f_H(v)dv}{\int_{y_H^*}^{\overline{v}} f_H(v)dv} - y_H^* \right) = 0. \tag{27}$$

It is necessary and may not be sufficient.

Remark 1: For $\frac{f_H(y)}{f_L(y)} \leq 1$, $(\int_y^{\overline{v}} v f_H(v)dv - \int_y^{\overline{v}} v f_L(v)dv)$ is increasing. Since $\frac{\int_y^{\overline{v}} v f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv}$ is increasing in y, $\frac{\int_y^{\overline{v}} v f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv}(\int_y^{\overline{v}} v f_H(v)dv - \int_y^{\overline{v}} v f_L(v)dv)$ is increasing for $\frac{f_H(y)}{f_L(y)} \leq 1$. Therefore, $\frac{f_H(y_H^*)}{f_L(y_H^*)} > 1$.

Remark 2: It is an indicator to measure difference of two types. If $R_L^W \leq R_L(y_H^*)$, the low type generates low value for the buyer. Her quality is relatively low compared with the high type. If $R_L^W > R_L(y_H^*)$, the low type is relatively close to the high type. We will provide more explanations with Proposition 9.

Starting with any threshold, the high type potentially can decrease the threshold of monotone binary partition combined with a lower information fee to generate higher revenue

[7] If there are more than one solution, we denote it to be the inferior of the solution.

[8] We would have similar results if uniqueness does not hold. We need it only to get neat results.

while maintaining low-type seller's mimicking revenue and relaxing the buyer's participation constraint. This has three implications. First, as the following proposition shows, the support of high-type seller's action in a RSW mechanism must be a subset of thresholds where she cannot do better by decreasing threshold and information fee at the same time.

Proposition 4 *In a RSW mechanism, the high type implements* (π, Γ) *with positive probability, that is,* $\sigma^W(\pi, \Gamma|H) > 0$*, only if*

$$\pi = \pi^M(y), \tag{28}$$

$$\Gamma = (r(s_1, 1), w), \tag{29}$$

where $y = \underline{y(R_L(y))}$.

Second, given R_L, $R_L + D_H(R_L)$ is the maximal revenue the high type can achieve by deterministic mechanism with IC and IR constraints. If we restrict to deterministic mechanism, then $\underline{y(R_L^W)}$ is the threshold of the high type in the RSW mechanism. Moreover, she would charge an information fee equal to $R_L^W - R_L(\underline{y(R_L^W)})$, and make revenue of $R_L^W + D_H(R_L^W)$. By construction, the function $D_H(R_L)$ has the following properties:

Lemma 2 $D_H(R_L)$ *is continuous and decreasing in* R_L*. Moreover,* $D_H(R_L) \geq 0$ *with equality satisfied if and only if* $R_L = \int_0^{\overline{v}} v f_H(v) dv$.

Third, Problem I provides an insight on characterizing high-type seller's revenue in the RSW mechanism. We can use concavification method to find it. Let $Z_H(R_L)$ be the concave closure of $D_H(R_L)$:

$$Z_H(R_L) = \sup\{z | (R_L, z) \in co(D_H)\}, \tag{30}$$

where $co(R_H)$ denotes the convex hull of the graph of R_H.

Proposition 5 *In the RSW mechanism, the high type makes revenue* $R_H^W = R_L^W + Z_H(R_L^W)$*. She could achieve this revenue by setting* $\sigma^W(\pi, \Gamma|H) > 0$ *only if for some threshold* y,

$$\pi = \pi^M(y), \tag{31}$$

$$\Gamma = (r(s_1, 1), w), \tag{32}$$

and $\sum_{\pi,\Gamma} \sigma^W(\pi, \Gamma|H) w(\Gamma) \leq 0$.

This proposition characterizes revenue for the high type and sufficient strategies that guarantees it. She can achieve it by randomizing on monotone binary partition combined with posted price and an information fee. While the posted price is set to extract full surplus, the high type rewards the buyer of all values with an information subsidy. Potentially, the high type pays the buyer to learn further information.

Note that the strategy need not to be unique although the revenue is unique. We discuss necessary part of high-type-seller's strategy in a RSW mechanism in the following proposition. This is equivalent to solve the concave closure.

Proposition 6 *In the RSW mechanism, high-type seller's strategy is as follows:*

(1) If $R_L^W \leq R_L(y_H^*)$, $\sigma^W(\pi^M(y_H^*), (r(s_1, 1), R_L^W - R_L(y_H^*))|H) = 1$*;*

(2) If $R_L^W > R_L(y_H^*)$, *there exists* $(\pi^M(y^1), (r(s_1^1, 1), 0))$ *and* $(\pi^M(y^2), (r(s_1^1, 1), 0))$ *where* $R_L(y^2) \geq R_L^W \geq R_L(y^1) \geq R_L(y_H^*)$ *such that*

$$\sum_{i=1,2} \sigma^W(\pi^i, \Gamma^i|H) = 1, \tag{33}$$

$$\sum_{i=1,2} \sigma^W(\pi^i, \Gamma^i|H) R_L(y^i) = R_L^W. \tag{34}$$

Moreover, the buyer has exante utility,

$$U(\sigma^W) \begin{cases} > 0 \text{ if } R_L^W < R_L(y_H^*), \\ = 0, \text{ otherwise.} \end{cases} \tag{35}$$

When $R_L^W \leq R_L(y_H^*)$, the low type is very distinct from the high type and generates low value for the buyer. The high type maximizes revenue difference with the unique threshold, y_H^* and transfers to the buyer a positive information subsidy. In the other case, $R_L^W > R_L(y_H^*)$, and the low type has lower mimicking incentive. The high type can do better and adopts lower thresholds. She randomizes on one or two different thresholds. Although the high type has incentive to subsidize the buyer, she sets it to be zero. This is because for any package with positive information subsidy, there exists a linear combination of packages with zero information subsidy that generates higher revenue for the high type and the same revenue for the low type. In this case, monotone binary partition is more efficient than the subsidy in separating.

Denote

$$DE_H(R_L) = \int_{y(R_L)}^{\overline{v}} v f_H(v) dv - R_L.$$

It is the difference in revenue of the high type from the low type with monotone binary partition and posted price only provided that low-type seller's mimicking revenue is not higher than R_L. Note that $DE_H(R_L)$ is a simple calculation form R_L compared with $D_H(R_L)$, which is a solution to an maximization problem. Therefore, $DE_H(R_L) \leq D_H(R_L)$ with equality holds if and only if $DE_H(R'_L) \leq DE_H(R_L), \forall R'_L > R_L$. However, when $R_L \geq R_L(y^*_H)$, concave closure of these two functions coincides.

Corollary 1 $Z_H(R_L) = \begin{cases} DE_H(R_L(y^*_H)), & \text{if } R_L \leq R_L(y^*_H), \\ \sup\{z|(R_L, z) \in co(DE_H)\}, & \text{otherwise.} \end{cases}$ *Moreover,* $Z_H(R_L)$ *is strictly decreasing in* R_L *when* $R_L > R_L(y^*_H)$.

We can find high-type seller's revenue as well as strategy in the RSW mechanism through the following three steps. In step 1, we solve the maximal difference in revenue and find y^*_H. In step 2, we calculate $R_L(y^*_H)$. If $R^W_L \leq R_L(y^*_H)$, it is straightforward that the high type makes revenue of $R^W_H = R^W_L + DE_H(R_L(y^*_H))$. She achieves it by setting threshold of monotone binary partition to be y^*_H, posted price equal to expected value of signal s_1 and an information subsidy just high enough to deter the low type from mimicking such that

$$w = R^W_L - R_L(y^*_H).$$

Information fee $w \leq 0$ with equality satisfied if and only if $R^W_L = R_L(y^*_H)$. In step 3, if $R^W_L > R_L(y^*_H)$, we characterize $DE_H(R_L)$ and find the concave closure. By Caratheodory's theorem, it can be written as a convex combination of at most 2 points. By definition of $DE_H(R_L)$, the seller charges zero information fee.

When $R^W_L > R_L(y^*_H)$, it is demanding to find the convex hull. We provide sufficient conditions under which $DE_H(R_L)$ is a concave function and therefore concave closure is itself. Consider following conditions: $(A1)$ $f_H(v)$ is increasing in v; $(A2)$ $f_H(v)$ is decreasing in v; $(A3)$ $v f_H(v)$ is increasing; $(A4)$ $\frac{\frac{\partial \frac{f_L(v)}{f_H(v)}}{\partial v}}{f_H(v)}$ is increasing. .

Lemma 3 *If* $A1$ *holds,* $DE_H(R_L)$ *is a strictly concave function of* R_L. *Moreover,* y^*_H *is uniquely determined by first-order condition,* (27).

Lemma 4 *If any two of $A2 - A4$ conditions hold, $DE_H(R_L)$ is a strictly concave function of R_L. Moreover, y_H^* is uniquely determined by first-order condition, (27).*

These two lemmas show that if condition $(A1)$ or any two of conditions $(A2)$, $(A3)$, $(A4)$ hold, $DE_H(R_L)$ is a concave function, whose concave closure is itself. Therefore,

$$Z_H(R_L) = \begin{cases} DE_H(R_L(y_H^*)), \text{ if } R_L \leq R_L(y_H^*), \\ DE_H(R_L), \text{ otherwise.} \end{cases} \tag{36}$$

The high type cannot be better to randomize on different packages than to take a singleton action.

Proposition 7 *If condition (A1) or any two of conditions (A2), (A3), (A4) hold, in RSW mechanism, high-type seller's action is as follows:*

(1) If $R_L^W \leq R_L(y_H^*)$, $\sigma^W(\pi^M(y_H^*), (r(s_1, 1), R_L^W - R_L(y_H^*))|H) = 1$;

(2) If $R_L^W > R_L(y_H^*)$, $\sigma^W(\pi^M(y(R_L^W)), (r(s_1, 1), 0)|H) = 1$.

Moreover, the buyer has exante utility,

$$U(\sigma^W) \begin{cases} > 0 \ \textit{if } R_L^W < R_L(y_H^*), \\ = 0, \ \textit{otherwise.} \end{cases} \tag{37}$$

This proposition provides situations under which result in Chen and Zhang (2018) still holds in this more general setting. When $R_L^W \geq R_L(y_H^*)$, it adds no value to the high type to set an information fee or to have more committing power with a mechanism. The high type would adopt monotone binary partition and set take-it-or-leave-it price equal to expected value given that value is above the threshold and the seller is the high type. The threshold is uniquely determined by the equation: $R_L(y) = R_L^W$.

We provide some examples under which some of above conditions hold and therefore $DE_H(R_L)$ is concave. First, consider $F_\theta(v) = v^\theta$ on $[0, 1]$. If $\theta \geq 1$, $(A1)$ holds. If $\theta < 1$, $(A2)$ and $(A4)$ hold. Second, consider $F_\theta(v) = 1 - e^{-\frac{1}{\theta}v}$ on $[0, \infty)$. Then condition $(A2)$ holds. Condition $(A4)$ holds if and only if $2\theta_L \leq \theta_H$.

5.3 The RSW mechanism

To summarize, we have the following theorem that characterizes both types in the RSW mechanism:

Theorem 1 *For any RSW mechanism, it has the following properties:*

(1) The low type makes revenue $R_L^W = \int_0^{\overline{v}} v f_L(v)dv$. She implements no disclosure, sets price equal to expected value, $r(s_1, 0) = \int_0^{\overline{v}} v f_L(v)dv$, exante zero information fee.

(2) The high type makes revenue $R_H^W = R_L^W + Z_H(R_L^W)$. There are two cases:

(i) If $R_L^W \leq R_L(y_H^*)$, she adopts monotone binary partition with threshold y_H^*, sets price equal to $r(s_1, 1) = \frac{\int_{y_H^*}^{\overline{v}} v f_H(v)dv}{\int_{y_H^*}^{\overline{v}} f_H(v)dv}$, and information fee equal to $R_L^W - R_L(y_H^*)$;

(ii) If $R_L^W > R_L(y_H^*)$, she sets zero information fee. She randomizes on one or two packages of monotone binary partition with thresholds y^1, y^2, and price equal to $\frac{\int_{y^1}^{\overline{v}} v f_H(v)dv}{\int_{y^1}^{\overline{v}} f_H(v)dv}$, $\frac{\int_{y^2}^{\overline{v}} v f_H(v)dv}{\int_{y^2}^{\overline{v}} f_H(v)dv}$ respectively, where $R_L(y^2) \geq R_L^W \geq R_L(y^1) \geq R_L(y_H^*)$ such that

$$\sum_{i=1,2} \sigma^W(\pi^i, \Gamma^i | H) = 1, \tag{38}$$

$$\sum_{i=1,2} \sigma^W(\pi^i, \Gamma^i | H) R_L(y^i) = R_L^W. \tag{39}$$

Moreover, it is feasible and can be supported as PBE.

This theorem fully characterizes properties of the RSW mechanism. First, the low type extracts full surplus as if her type was known. In contrast, the high type is forced to exclude the buyer who has relatively low value to deter the low type from mimicking. Second, each type implements a selling scheme featured by a take-it-or-leave-it price and an information fee. Third, each type implements monotone binary partition but with different cutoffs driven by different causes. The low type adopts no disclosure to extract full surplus while the high type tries to minimize low-type seller's mimicking incentive. The high type also sets price equal to expected value provided that type is known and value is above the cutoff. Forth, both type possibly randomize on different packages. For the low type, each package sells with certainty and differs only in information fee. The high type randomizes on different cutoffs and therefore different selling probabilities. Fifth, a positive information subsidy may help. In this case, the buyer makes positive utility transacting with the high type. Moreover, the subsidy is small compared with price.

This theorem shows that the RSW mechanism is feasible and an equilibrium. It is the least cost separating equilibrium in which each type separates in implementation of information disclosure and selling scheme. Among all the information disclosure policies, monotone binary partition is the most effective in separating from the low type. For the selling scheme, it is effective to set price at expected value and reward the buyer with information subsidy. The high type weights these two strategies. As shown by Proposition 4, she can decrease cutoff of monotone binary partition and meanwhile increase the information subsidy to offset the effect on the low-type seller's mimicking revenue. When the low type is distinct from the high type, it is more effective to subsidize the buyer; When the low type is close to the high type, monotone binary partition is more effective and the high type offers zero information subsidy.

Moreover, it provides the lower bound of equilibrium payoff. Maskin and Tirole [17] show that $\{R_\theta\}_{\theta\in\Theta}$ is achievable in a PBE only if it weakly Pareto dominates the RSW. We have the same result here. Since it satisfies full-information participation constraint, any type of seller can always adopt a RSW mechanism and therefore guarantees herself the corresponding payoff.

Proposition 8 *$\{R_\theta\}_{\theta\in\Theta}$ is payoff of the seller in a PBE only if $R_\theta \geq R_\theta^W, \forall\theta$.*

5.4 Comparative statics

Consider any $F_H(v), F_L(v)$ that satisfies the monotone likelihood ratio property and $F_\alpha(v) = \alpha F_H(v) + (1-\alpha) F_L(v)$ for any $\alpha \in [0,1]$. Then for any $\alpha' > \alpha$, $F_{\alpha'}(v)$ dominates $F_\alpha(v)$ in terms of likelihood ratio dominance. Particularly, $F_H(v)$ dominates $F_\alpha(v)$. We increase the low type by increasing α and conduct comparative statics with respect to it. Denote $y_H^*(\alpha)$ be the solution to the following problem:

$$\max_{\widetilde{y}} \int_{\widetilde{y}}^{\overline{v}} v f_H(v) dv - R_\alpha(\widetilde{y}). \tag{40}$$

It maximizes revenue difference when the low type is of type α.

Proposition 9 *For any $F_H(v), F_L(v)$ that satisfies the monotone likelihood ratio property and $F_\alpha(v) = \alpha F_H(v) + (1-\alpha) F_L(v)$, there exists a lower bound $\underline{\alpha} \in [0,1)$ such that $R_\alpha^W > R_\alpha(y_H^*(\alpha))$ if and only if $\alpha > \underline{\alpha}$. Moreover, in the RSW mechanism, for the high type, the information subsidy is decreasing in α and is strictly decreasing when $\alpha < \underline{\alpha}$; the revenue is strictly increasing in α. The social welfare is also strictly increasing in α. The buyer's exante payoff is decreasing in α and is strictly decreasing when $\alpha < \underline{\alpha}$.*

When we increase α, the low type gets closer to the high type. This proposition shows that if we increase the quality of low type, the condition $R_{\alpha}^{W} > R_{\alpha}(y_H^*(\alpha))$ is single crossing. Therefore, y_H^* measures the difference of low type to the high type. If the condition does not hold, the low type is relatively far from the high type. In the RSW mechanism, the high type subsidizes the buyer to learn information. In this case, when the low type increases, the information subsidy strictly decreases. If the condition holds, the high type charges zero information fee and allows the buyer to learn information for free. As we can observe in real life, a less regulated market like tutorial market, widely adopts information subsidy; in contrast, a highly regulated market like snack market that restricts the lowest quality, free information disclosure is more common.

This proposition also provides insights on policy implication. High-type seller's revenue and social welfare are both strictly increasing in the low type. Therefore, regulation on the lowest quality would benefits both the high type and the whole society. However, the buyer may be worse. When $\alpha < \underline{\alpha}$, the lower bound that condition $R_{\alpha}^{W} > R_{\alpha}(y_H^*(\alpha))$ holds, the buyer receives positive exante payoff from an information subsidy in the RSW mechanism. It strictly decreases in the low type and is zero when the low type reaches and crosses the lower bound. A regulator who concerns more about the buyer's payoff should think twice before restricting the lowest quality.

6 Equilibrium

In this section, we discuss more properties of equilibrium. First, we characterize the revenue frontier. As we have discussed, the RSW mechanism provides lower bounds of revenue. We discuss upper bounds in this section. Second, if lower bound and upper bound coincides for any type, the RSW mechanism is the unique equilibrium mechanism, for which we provide sufficient and necessary condition.

6.1 Equilibrium revenue

Now we are ready to characterize the revenue frontier. Consider any vector of revenue for each type of the seller, $[R_L, R_H]$. Proposition 8 demonstrates that there exists a lower bound for each type of the seller. The following proposition characterizes the upper bound. There are two types of feasible constraint, which impose limits on feasible equilibrium revenue of the seller. First, the buyer's participation constraint requires that the exante expected revenue cannot exceed total social surplus:

$$\mu_H^0 R_H + \mu_L^0 R_L \leq \int_0^{\overline{v}} v f(v) dv. \tag{41}$$

Second, for the high type, $IC_{L,H}$ restricts feasible strategies and therefore revenue of the high type. There are two approaches to relax $IC_{L,H}$. One is to adopt actions of least indirect separating cost. It is advantageous for the high type to adopt monotone binary partition, set price equal to expected value. Another approach is to compensate the low type to reduce her deviating incentive. To accomplish this target, the high type leaves surplus to the buyer so that the low type could charge more than the welfare she generates for the buyer. The high type could charge a negative information fee which functions as a subsidy. The buyer would be willing to participate even though she makes negative payoff if met with the low type. When $R_L \leq \int_0^{\overline{v}} v f(v) dv$, it is without loss of generality to consider any feasible mechanism σ^o that satisfies conditions (1) and (2) in Lemma 1.

Proposition 10 *$[R_L, R_H]$ is a vector of revenue for each type of the seller in a PBE only if*

$$R_L^W \leq R_L, \tag{42}$$

$$R_H^W \leq R_H \leq \begin{cases} \int_0^{\overline{v}} v f(v) dv, & \text{if } R_L > \int_0^{\overline{v}} v f(v) dv, \\ R_L + Z_H\left(\frac{R_L - \mu_L^0 R_L^W}{\mu_H^0}\right), & \text{otherwise.} \end{cases} \tag{43}$$

In generic, participation constraint and incentive compatible constraint of the low type do not bind at the same time. When $R_L = \int_0^{\overline{v}} v f(v) dv$, two upper bounds of the high type coincide. Moreover, these two limits may not bind as there are also other feasible constraints, like $IC_{H,L}$. Therefore, the previous proposition only provides necessary condition on revenue bounds. The following proposition discusses sufficient condition. There is a gap between sufficient and necessary conditions in terms of upper bound for the low type and lower bound for the high type. It is possible that the high type makes lower revenue than the low type in equilibrium. It is difficult to reduce the gap without discussing the structure of distribution function, $F_\theta(v)$.

Proposition 11 *$[R_L, R_H]$ is a vector of revenue for each type of the seller in a PBE if*

$$R_L^W \leq R_L \leq \int_0^{\overline{v}} v f(v) dv, \tag{44}$$

$$\max\{R_L, R_H^W\} \leq R_H \leq R_L + Z_H\left(\frac{R_L - \mu_L^0 R_L^W}{\mu_H^0}\right). \tag{45}$$

Moreover, there exists a feasible mechanism σ *with belief* μ^0 *that can be supported as PBE such that*

(*a*) $\sigma(\pi, \Gamma|L) > 0$ *only if* $\pi = \pi^N$, $\Gamma = (r(s_1, 0), R_L - r(s_1, 0))$;

(*b*) $\sigma(\pi, \Gamma|H) > 0$ *only if* $\pi = \pi^M(y)$ *for some threshold* y, $\Gamma = (r(s_1, 1), w)$ *for some* $w \leq 0$;

(*c*) $R_\theta(\sigma) = R_\theta$.

This proposition shows sufficient condition for a vector of revenue to be PBE revenue for the seller. The proof is by construction of a feasible mechanism on monotone binary partition, posted price and information fee. Each type of the seller sets price equal to expected value given that type is known. The low type charges positive information fee while the high type subsidizes the buyer with negative information fee. The low type benefits from private information and makes revenue weakly higher than her total surplus. It is feasible only because the high type offers an information subsidy to stimulate the buyer to participate and therefore indirectly compensates the low type.

Another property is that each type of the seller separates on packages of disclosure policy and selling scheme. Generally, the low type benefits from bunching with the high type as she can charge a higher price. Now she can achieve the same goal by asking for a higher information fee. The high type randomizes on packages of monotone binary partition, posted price and information subsidy to deter the low type from mimicking. It is the least cost separating strategy.

Notably, in equilibrium, the high type may pay the buyer to observe realized signal. It is driven by two motivations. First, indirectly, it relaxes low-type seller's incentive-compatible constraint. If the buyer has higher payoff when met with the high type, she is more willing to participate in the mechanism. Consequently, the low type can make a higher revenue with a higher information fee and her incentive to mimic the high type decreases. Second, disclosure policy is the main signalling tool and a subsidy stimulates the buyer to learn more information. Since the low type has relatively low value to the buyer, it is too costly for her to reveal information. The low type loses more from subsidizing the buyer to learn realized signal than she benefits from bunching with the high type to charge a higher price.

6.2 Uniqueness of PBE

Since the RSW mechanism provides lower bound of equilibrium revenue for each type, it will be the only equilibrium outcome if it is undominated. By proposition 10, if there exists any PBE other than RSW mechanism, there will exists $[R_L, R_H] \neq \left[R_L^W, R_H^W\right]$ such that conditions (42, 43) hold. To prove the contrary, it is sufficient to show that for any $R_L^W < R_L \leq \int_0^{\overline{v}} vf(v)dv$,

$$R_H^W > R_L + Z_H(\frac{R_L - \mu_L^0 R_L^W}{\mu_H^0}). \tag{46}$$

Since $R_H^W = R_L^W + Z_H(R_L^W)$, it is equivalent to show that the RHS of (46) achieves maximal only at $R_L = R_L^W$. It is also necessary by Proposition 11.

Theorem 2 *Any equilibrium can be achieved by the RSW mechanism if and only if (1) $R_L^W \geq R_L(y_H^*)$, (2) $\mu_H^0 \leq -\frac{\partial Z_H(R_L^W)}{\partial R_L}$ with inequality strictly holds if $\frac{\partial^2 Z_H(R_L^W)}{\partial (R_L)^2} = 0$.*

This theorem provides sufficient and necessary conditions under which any PBE can be achieved by the RSW mechanism. For the high type in any PBE, there is trade-off between the buyer's IR constraint and the low-type seller's IC constraint. Condition (1) requires that quality of the low type is not too low. Otherwise, the low type has strong incentive to mimic the high type, forcing the high type to give up too much surplus in the RSW mechanism. To relax IC constraint of the low type, the high type benefits from compensating the low type indirectly through subsidizing the buyer. Condition (2) requires that the proportion of the high type is lower a threshold. In this case, the loss from bearing a larger burden of IR constraint to transfer surplus indirectly to the low type is higher than the cost of satisfying IC constraint. The high type prefer to separate from the low type in participation constraint.

7 Intuitive equilibrium

Although we impose inscrutability principle which restricts the buyer's on-equilibrium-path belief, the set of equilibrium outcome is not reduced. Now we further impose intuitive criterion introduced by Cho and Kreps [7] directly to select equilibrium outcome. For any off-equilibrium-path action $\widetilde{M}$, it puts a restriction on the receiver's off-equilibrium-path belief, $\mu\left(\widetilde{M}\right)$, based on payoff dominance. The buyer decides whether to participate given this belief and updates his belief in the following stages according to Bayes rule. Denote

$R_\theta(\widetilde{M}, \widetilde{u}))$ be the seller's revenue from $\widetilde{M}$ given the buyer's belief, $\widetilde{\mu}$. Denote $\mu\left(\widetilde{\Theta}\right)$ be the set that contains all possible beliefs that only put positive weights to type $\theta \in \widetilde{\Theta}$. Define for each action the set of types $\Theta^{**}(\widetilde{M}) = \{\theta \in \Theta | R_\theta(\sigma) \leq \max_{\widetilde{\mu} \in \mu(\Theta)} R_\theta(\widetilde{M}, \widetilde{\mu})\}$.

Definition 3 *A PBE with mechanism σ violates the intuitive criterion if there exists a type θ' and M' such that*

$$\min_{\widetilde{\mu} \in \mu(\Theta^{**}(M'))} R_{\theta'}(M', \widetilde{\mu}) > R_{\theta'}(\sigma). \tag{47}$$

Intuitive criterion indicates that a reasonable off-equilibrium-path belief assigns zero probability to those types who are strictly worse off than their equilibrium revenue. Any PBE where there exists some type having incentive to deviate provided that the buyer has reasonable belief will be eliminated. The PBEs that survives the intuitive criterion are called **intuitive equilibria**.

Proposition 12 *The RSW mechanism survives intuitive criterion.*

This proposition justifies why we focus on the RSW mechanism. It is a reasonable equilibrium that survives refinement of intuitive criterion. The proof is by construction of off-equilibrium-path belief. We assign the off-equilibrium-path belief to be either the low type or the high type. Moreover, it is the unique equilibrium mechanism that survives intuitive criterion when the low type is relatively close to the high type as is shown by the following proposition.

Proposition 13 *When $R_L^W \geq R_L(y_H^*)$, a PBE with mechanism σ survives Intuitive criterion if and only if σ is the RSW mechanism.*

This proposition characterizes the set of intuitive equilibrium. When the low type is close to the high type, that is, $R_L^W \geq R_L(y_H^*)$, any other PBE is vulnerable to intuitive criterion. We first show that there exists a lower bound of high-type seller's revenue in intuitive equilibria and it is weakly higher than the upper bound of PBE revenue characterized in Proposition 10. Therefore, an equilibrium survives intuitive criterion only if these two bounds coincide for the high type, which is equivalent to that an equilibrium is achieved by the RSW mechanism.

However, the RSW mechanism may not be the unique intuitive equilibrium mechanism. It depends on the difference of the high type to the low type. When $R_L^W < R_L(y_H^*)$, that is, the low type is distinct from the high type, the set of mechanism of intuitive equilibrium is larger than the RSW mechanism.

8 Conclusion and discussion

This paper considers an informed principal problem in which the seller is privately and imperfectly informed of the buyer's value. The seller designs a mechanism where conditional on the seller's report of type, a package of disclosure policy and selling scheme is implemented. This paper characterizes the RSW mechanism and shows that any equilibrium Pareto dominates the RSW mechanism following the literature on informed principal problem. This paper discusses sufficient and necessary condition under which the RSW mechanism is the unique equilibrium mechanism. When we impose intuitive criterion to select equilibrium by restricting off-equilibrium-path belief, the RSW mechanism survives intuitive criterion.

In the RSW mechanism, the low type cannot do better than the case when her type is commonly observable. She extracts the full surplus. The high type tries to separate from the low type by randomizing on packages of monotone binary partition and deterministic selling scheme such that the low type is indifferent between mimicking and choosing her own action. The high type sets price equal to expected value provided that it is above the threshold. Information subsidy can be either positive or zero.

8.1 Comparison with Chen an Zhang (2018)

As discussed in Section 2, this paper extends action space of the seller in Chen and Zhang (2018) . In the RSW mechanism, which corresponds to the Riley outcome in Chen and Zhang (2018) . the low type is the same in expected revenue and strategy; the high type is better in general with a higher revenue and generates higher social surplus. In both setups, it is the least cost in separating for the high to choose monotone binary partition and set price equal to expected value. However, the high type may set different cutoff in this paper from that in Chen and Zhang (2018). If $R_L^W < R_L(y_H^*)$, the high type chooses a lower cutoff and offers an information subsidy. If $R_L^W \geq R_L(y_H^*)$, the high type may randomize on two cutoffs. When the high type does not randomize, the cutoff would be the same as in Chen and Zhang (2018) .

Extension of action space does not extend the set of equilibrium. As in informed principle problem, any PBE must weakly dominates the RSW mechanism in terms of the seller's revenue. We are able to exclude many unreasonable equilibria. Under a sufficient and necessary condition, any equilibrium can be achieved by the RSW mechanism. It is different from a signalling setup where many separating equilibria with high-type seller's payoff lower than that in Riley outcome exist.

We also impose intuitive criterion in this paper. Notably, the RSW mechanism may not be the unique intuitive equilibrium, which contrasts the result in Chen and Zhang (2018). If the low type is relatively distinct from the high type, there exists an infinite number of intuitive equilibria.

8.2 Benchmark

We discuss two benchmarks: full-information optimal mechanism and exante optimal mechanism. First, suppose the seller's type is not private but public information. Each type of seller makes full-information optimal revenue

$$R_\theta^F = \int_0^{\overline{v}} v f_\theta(v) dv, \tag{48}$$

with no disclosure, π^N and a posted price, $p_\theta = \int_0^{\overline{v}} v f_\theta(v) dv$, to extract full surplus. Since $R_L^F < R_L(\pi_H^F, \Gamma_H^F)$, the low type has incentive to mimic the high type and therefore it is not feasible. In the RSW mechanism, the low type makes the same revenue as the full-information optimal revenue and the high type makes a lower revenue than her full-information optimal revenue. The high type suffers from private information and is forced to give up some surplus to separate from the low type. By contrast, the low type benefits from private information and in any PBE makes a revenue higher than her full-information benchmark.

Second, a mechanism is exante optimal if it maximizes exante revenue. The seller extracts the full exante surplus and each type makes revenue

$$R_\theta^{EO} = \int_0^{\overline{v}} v f(v) dv, \tag{49}$$

with no disclosure and a posted price equal to exante expected value. It is feasible. However, it may not be an equilibrium as we discuss in section 6. If the high type makes a strictly higher revenue in the RSW mechanism than the exante optimal one, the high type prefers to separate and therefore exante optimal one cannot be supported as PBE. Therefore, the signalling consideration hurts efficiency. By contrast, the buyer could be better off as he makes positive exante payoff with information subsidy.

8.3 Another interpretation

This setup also applies to managerial compensation. The employee has a mission to complete which requires effort e from the employer. Suppose e follows a distribution on $[0, \overline{e}]$. The mission is either difficult (L) or easy (H). Employee can disclose some information about e and designs a mechanism in which specifies the probability of assigning the mission to an employer at compensation. Then denote $v = \overline{e} - e$, $t = \overline{e}-$compensation, our analysis applies.

8.4 Extension

We can extend to the case when type is continuous and ordered in terms of likelihood ratio dominance. In the RSW mechanism, each type of the seller randomizes on different packages of monotone binary partition and deterministic selling scheme with nonpositive information fee. Under some conditions, local incentive constraint is sufficient and necessary. We can adopt the same concavification method. And the revenue of each type is determined by a differentiation function. We provide sufficient condition under which the convex hull is the function itself. For each type of seller, it is sufficient to randomize on at most two actions.

9 Appendix

Proof for Proposition 1 :

For the buyer with expost value v, denote $Q(v|\sigma), U(v|\sigma)$ be exante expected probability of getting the product and exante utility, respectively.

$$Q(v|\sigma) = \sum_{\theta} \mu(\theta) \sum_{(\pi,\Gamma)} \sigma(\pi, \Gamma|\theta) \int_{s \in S} \int_{0}^{\overline{v}} \pi(s|v) q(s) f_{\theta}(v) dv ds, \tag{50}$$

$$U(v|\sigma) = \sum_{\theta} \mu(\theta) \sum_{(\pi,\Gamma)} \sigma(\pi, \Gamma|\theta) \int_{s \in S} \int_{0}^{\overline{v}} \pi(s|v) U(s, \lambda|\Gamma) f_{\theta}(v) dv ds. \tag{51}$$

There are three steps.

In step 1, we will show that for any (π, Γ), given interim belief λ, there exists a mixture of deterministic selling scheme such that $q(\widetilde{s}) = \int q^s(\widetilde{s}) d\rho(\Gamma^s)$ almost everywhere. Moreover, $R_\theta(\pi, \Gamma, \lambda) = \int R_\theta(\pi, \Gamma^s, \lambda) d\rho(\Gamma^s)$, $U(\lambda|\pi, \Gamma) = \int U(\lambda|\pi, \Gamma^s) d\rho(\Gamma^s)$. Since $P_\lambda(s, \lambda) \leq 1$, there are at most finite mass points. Suppose $\{s_i\}_{i=1}^m$ are mass points and $s_i < s_{i+1}$. Denote $r(s_0,\lambda) = 0, r(s_{m+1},\lambda) = \overline{v}$. Then $p_\lambda(s,\lambda)$ is continuous in (s_i, s_{i+1}) $\forall i = 0,\cdots m$. Since q is monotone by IC constraints, the set of points where q is discontinuous is at most countable and the measure of discontinuous points is 0. Therefore, changing q to right continuous q' that is different from q only at the set of discontinuous points in (s_i, s_{i+1}) does not affect $R_\theta(\pi,\Gamma,\lambda)$ or $U(\lambda|\pi,\Gamma)$. Denote $\Gamma^s = \{q(\widetilde{s}) = I_{\widetilde{s}\geq s}, w(\Gamma)\}$. It satisfies IC_A by definition of s. Let $\rho(\Gamma^s) = \begin{cases} q'(s), \forall r(s,\lambda) < \overline{v}, \\ 1 - q'(s), r(s,\lambda) = \overline{v}. \end{cases}$ Then we will show that it is the objective mixture. First,

$$\begin{aligned} & \int_{s\in S} I_{\widetilde{s}\geq s} d\rho(\Gamma^s) \\ = & \int_{s\in S} I_{\widetilde{s}\geq s} dq'(s) \\ = & q'(\widetilde{s}). \end{aligned} \tag{52}$$

Since $q'(\widetilde{s}) = q(\widetilde{s})$ almost everywhere, $q(\widetilde{s}) = \int_{s\in S} I_{\widetilde{s}\geq s} d\rho(\Gamma^s)$ almost everywhere. Second,

$$\begin{aligned} & U(\widetilde{s},\lambda|\Gamma) \\ = & \int_{r(s,\lambda)\leq r(\widetilde{s},\lambda)} q(s) dr(s,\lambda) - w(\Gamma) \\ = & \int_{r(s,\lambda)\leq r(\widetilde{s},\lambda)} \int_{\widehat{s}\in S} I_{s\geq\widehat{s}} d\rho(\Gamma^{\widehat{s}}) dr(s,\lambda) - w(\Gamma) \\ = & \int_{\widehat{s}\in S} I_{s\geq\widehat{s}} \int_{r(s,\lambda)\leq r(\widetilde{s},\lambda)} dr(s,\lambda) d\rho(\Gamma^{\widehat{s}}) - w(\Gamma) \\ = & \int_{\widehat{s}\in S} I_{\widetilde{s}\geq\widehat{s}} \int_{r(\widehat{s},\lambda)\leq r(s,\lambda)\leq r(\widetilde{s},\lambda)} dr(s,\lambda) d\rho(\Gamma^{\widehat{s}}) - w(\Gamma) \\ = & \int_{\widehat{s}\in S} I_{\widetilde{s}\geq\widehat{s}} \int_{r(s,\lambda)\leq r(\widetilde{s},\lambda)} I_{s\geq\widehat{s}} dr(s,\lambda) d\rho(\Gamma^{\widehat{s}}) - w(\Gamma) \\ = & \int_{\widehat{s}\in S} U(\widetilde{s},\lambda|\Gamma^{\widehat{s}}) d\rho(\Gamma^{\widehat{s}}). \end{aligned} \tag{53}$$

Therefore, $U(\lambda|\pi,\Gamma) = \int U(\lambda|\pi,\Gamma^d)d\rho(\Gamma^d)$. Third,

$$\begin{aligned}
& R_\theta(\pi,\Gamma,\lambda) \\
= & -\int_{s\in S} q'(s)d(r(s,\lambda)(1-P_\theta(s,\lambda))) + w(\Gamma) \\
= & \int_{s\in S} r(s,\lambda)(1-P_\theta(s,\lambda))dq'(s) + w(\Gamma) \\
= & \int_{s\in S} R_\theta(\pi,\Gamma^s,\lambda)d\rho(\Gamma^s). \qquad (54)
\end{aligned}$$

In step 2, we consider any deterministic mechanism $\Gamma^s = (q(\widetilde{s}) = I_{\widetilde{s}\geq s}, w)$. Given λ, combine signals with $q(s) = 1$ into one signal s_1 and the rest into s_2, that is,

$$\pi^1(s_1|v) = \sum_{\widetilde{s}\geq s} \pi(\widetilde{s}|v), \qquad (55)$$

$$\pi^1(s_2|v) = 1 - \pi^1(s_1|v). \qquad (56)$$

Set $\chi(\pi^1,\Gamma^s|\pi,\Gamma,\theta) = 1$, and $\chi(\widetilde{\pi},\widetilde{\Gamma}|\pi,\Gamma,\theta) = 0$ for any other $\left(\widetilde{\pi},\widetilde{\Gamma}\right)$.

In step 3, construct $\sigma^1(\pi^1,\Gamma^s|\theta) = \int_{\pi,\Gamma} \chi(\pi^1,\Gamma^s|\pi,\Gamma,\theta)\sigma(\pi,\Gamma|\theta)d\rho(\Gamma^s)$. First, by construction, for any (π^1,Γ^s),

$$U(s_1,\lambda|\Gamma^s) = r(s_1,\lambda) - r^s - w^s, \qquad (57)$$

$$U(s_2,\lambda|\Gamma^s) = -w^s. \qquad (58)$$

If $s_1 > s_2$ for λ_1, λ_2, it also holds for any $\alpha\lambda_1 + (1-\alpha)\lambda_2$, $\forall\alpha \in (0,1)$. Second,

$$
\begin{aligned}
& Q(v|\sigma) \\
= & \sum_{\theta} \mu(\theta) \sum_{(\pi,\Gamma)} \sigma(\pi,\Gamma|\theta) \int_{s\in S} \int_0^{\overline{v}} \pi(s|v) q(s) f_\theta(v) dv ds \\
= & \sum_{\theta} \mu(\theta) \sum_{(\pi,\Gamma)} \sigma(\pi,\Gamma|\theta) \int_{s\in S} \int_0^{\overline{v}} \pi(s|v) \int_{\widetilde{s}\in S} I_{s\geq\widetilde{s}} d\rho(\Gamma^{\widetilde{s}}) f_\theta(v) dv ds \\
= & \sum_{\theta} \mu(\theta) \sum_{(\pi,\Gamma)} \sigma(\pi,\Gamma|\theta) \sum_{(\pi^1,\Gamma^{\widetilde{s}})} \chi(\pi^1,\Gamma^s|\pi,\Gamma,\theta) \int_0^{\overline{v}} \pi^1(s_1|v) d\rho(\Gamma^{\widetilde{s}}) f_\theta(v) dv \\
= & \sum_{\theta} \mu(\theta) \sum_{(\pi^1,\Gamma^{\widetilde{s}})} \sum_{(\pi,\Gamma)} \sigma(\pi,\Gamma|\theta) \chi(\pi^1,\Gamma^s|\pi,\Gamma,\theta) \int_0^{\overline{v}} \pi^1(s_1|v) f_\theta(v) dv d\rho(\Gamma^{\widetilde{s}}) \\
= & \sum_{\theta} \mu(\theta) \sum_{(\pi^1,\Gamma^{\widetilde{s}})} \sigma(\pi^1,\Gamma^{\widetilde{s}}|\theta) \int_0^{\overline{v}} \pi^1(s_1|v) f_\theta(v) dv \\
= & Q(v|\sigma^1).
\end{aligned}
\tag{59}
$$

And

$$
\begin{aligned}
& U(v|\sigma) \\
= & \sum_{\theta}\mu(\theta)\sum_{(\pi,\Gamma)}\sigma(\pi,\Gamma|\theta)\int_{s\in S}\int_0^{\overline{v}}\pi(s|v)U(s,\lambda(\pi,\Gamma|\sigma)|\Gamma)f_\theta(v)dvds \\
= & \sum_{\theta}\mu(\theta)\sum_{(\pi,\Gamma)}\sigma(\pi,\Gamma|\theta)\int_{s\in S}\int_0^{\overline{v}}\pi(s|v)\int_{\widetilde{s}\in S}U(s,\lambda(\pi,\Gamma|\sigma)|\Gamma^{\widetilde{s}})d\rho(\Gamma^{\widetilde{s}})f_\theta(v)dvds \\
= & \sum_{\theta}\mu(\theta)\sum_{(\pi,\Gamma)}\sigma(\pi,\Gamma|\theta)\int_{s\in S}\int_0^{\overline{v}}\pi(s|v)I_{s\geq\widetilde{s}}\left(r(s,\lambda(\pi,\Gamma|\sigma))-r^{\widetilde{s}}\right)-w^{\widetilde{s}}d\rho(\Gamma^{\widetilde{s}})f_\theta(v)dvds \\
= & \sum_{\theta}\mu(\theta)\sum_{(\pi,\Gamma)}\sigma(\pi,\Gamma|\theta)\sum_{(\pi^1,\Gamma^{\widetilde{s}})}\chi(\pi^1,\Gamma^s|\pi,\Gamma,\theta)* \\
& \int_0^{\overline{v}}\pi(s_1|v)\left(r(s_1,\lambda(\pi^1,\Gamma^{\widetilde{s}}|\sigma))-r^{\widetilde{s}}\right)-w^{\widetilde{s}}d\rho(\Gamma^{\widetilde{s}})f_\theta(v)dv \qquad (60) \\
= & \sum_{\theta}\mu(\theta)\sum_{(\pi^1,\Gamma^{\widetilde{s}})}\sigma(\pi^1,\Gamma^{\widetilde{s}}|\theta)\int_0^{\overline{v}}\left(\pi(s_1|v)U(s_1,\lambda(\pi^1,\Gamma^{\widetilde{s}}|\sigma))+\pi(s_2|v)U(s_2,\lambda(\pi^1,\Gamma^{\widetilde{s}}|\sigma))\right)f_\theta(v)dv \\
= & U(v|\sigma^1). \qquad (61)
\end{aligned}
$$

Similarly, $R_\theta(\sigma,\theta') = R_\theta(\sigma^1,\theta'), \forall\theta,\theta'$.

Q.E.D.

Proof for Proposition 2:

First, we show that $R_L^W \leq \int_0^{\overline{v}} vf_L(v)dv$. By full-information participation constraint of the buyer,

$$
\begin{aligned}
& \sum_{(\pi,\Gamma)}\sigma(\pi,\Gamma|L)U(0|\pi,\Gamma) \\
= & \sum_{(\pi,\Gamma)}\sigma(\pi,\Gamma|L)(\int_{s\in S}q(s)(r(s,0)-t(s))dP_L(s,0)+w) \\
\geq & \ 0. \qquad (62)
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
& \sum_{(\pi,\Gamma)} \sigma(\pi,\Gamma|L) R_L(\pi,\Gamma,0) \\
= & \sum_{(\pi,\Gamma)} \sigma(\pi,\Gamma|L) (\int_{s\in S} q(s)t(s)dP_L(s,0) + w) \\
\leq & \int_{s\in S} q(s)r(s,0)dP_L(s,0) \\
\leq & \int_{s\in S} r(s,0)dP_L(s,0) \ (q(s) \leq 1) \\
= & \int_0^{\overline{v}} v f_L(v)dv \text{ (by Bayes plausible),} \qquad (63)
\end{aligned}
$$

with equality holds only if for any $\sigma(\pi,\Gamma|L) > 0$, $q(s) = 1, \forall s$.

Second, consider (π,Γ) with

$$\pi = \pi^N, \qquad (64)$$

$$w = 0, \qquad (65)$$

$$t = \int_0^{\overline{v}} v f_L(v)dv. \qquad (66)$$

Construct $\sigma(\pi,\Gamma|L) = 1$. Since $U(0|\pi,\Gamma) \geq 0$,

$$R_L(\sigma) = R_L(\pi,\Gamma,0) = \int_0^{\overline{v}} v f_L(v)dv. \qquad (67)$$

Third, since $\sum_{(\pi,\Gamma)} \sigma(\pi,\Gamma|L) R_L(\pi,\Gamma,0) = \int_0^{\overline{v}} v f_L(v)dv$ only if for any $\sigma(\pi,\Gamma|L) > 0$, $q(s) = 1, \forall s$ and we focus on binary partition, $\sigma(\pi,\Gamma|L) > 0$ only if $\pi\,(s_1|v) = 1, \forall v$. Since we focus on deterministic selling scheme, with only one signal, it is without loss of generality to focus on $(\int_0^{\overline{v}} v f_L(v)dv, w)$. Then

$$
\begin{aligned}
& \sum_{(\pi,\Gamma)} \sigma(\pi,\Gamma|L) R_L(\pi,\Gamma,0) = \sum_{(\pi,\Gamma)} \sigma(\pi,\Gamma|L) (\int_0^{\overline{v}} v f_L(v)dv + w) \\
\Leftrightarrow & \sum_{(\pi,\Gamma)} \sigma(\pi,\Gamma|L) w = 0. \qquad (68)
\end{aligned}
$$

The claim is verified.

Q.E.D.

Proof for Lemma 1 :

Given interim belief λ, denote

$$U(\theta,\lambda|\pi,\Gamma)=\int_{s\in S}q(s)\left(r\left(s_1,I_{\theta=H}\right)-t(s)\right)dP_\theta(s,\lambda). \tag{69}$$

Therefore,

$$U(\lambda|\pi,\Gamma)=\lambda\int_{s\in S}U(s,\lambda|\Gamma)dP_H(s,\lambda)+(1-\lambda)\int_{s\in S}U(s,\lambda|\Gamma)dP_L(s,\lambda). \tag{70}$$

The proof is by construction. There are three steps.

In step 1, consider (π,Γ) with interim belief λ such that $\sigma(\pi,\Gamma|H)>0$. Therefore, $0<\lambda\leq 1$. There are three cases.

In case 1, $S=\{s_1\}$. There exists an equivalent $(\pi',\Gamma')=\left(\pi^N,\left(r\left(s_1',1\right),-U(\lambda|\pi',\Gamma')\right)\right)$ such that

$$R_L(\pi',\Gamma',1)=R_L(\pi,\Gamma,\lambda), \tag{71}$$

$$R_H(\pi',\Gamma',1)=R_H(\pi,\Gamma,\lambda), \tag{72}$$

$$U(1|\pi',\Gamma')=U(H,\lambda|\pi,\Gamma). \tag{73}$$

Construct $\chi(\pi',\Gamma'|\pi,\Gamma,H)=1$, and $\chi(\widetilde{\pi},\widetilde{\Gamma}|\pi,\Gamma,H)=0$ for any other $\left(\widetilde{\pi},\widetilde{\Gamma}\right)$.

In case 2, $S=\{s_2\}$, and therefore, $R_\theta(\pi,\Gamma,\lambda)=-U(\lambda|\pi',\Gamma')$. There exists $y'\in(0,\overline{v})$ and $w'=-\int_{y'}^{\overline{v}}f_L(v)dv\frac{\int_{y'}^{\overline{v}}vf_H(v)dv}{\int_{y'}^{\overline{v}}f_H(v)dv}-U(\lambda|\pi',\Gamma')$ such that

$$R_L(\pi',\Gamma',1)=-U(\lambda|\pi',\Gamma'), \tag{74}$$

$$U(1|\pi',\Gamma')>U(\lambda|\pi,\Gamma), \tag{75}$$

where $(\pi', \Gamma') = \left(\pi^M(y'), (r(s_1', 1), w')\right)$. Since $\int_{y'}^{\overline{v}} f_L(v)dv < \int_{y'}^{\overline{v}} f_H(v)dv$,

$$\begin{aligned}
& R_H(\pi', \Gamma', 1) \\
= \; & \int_{y'}^{\overline{v}} v f_H(v)dv - \int_{y'}^{\overline{v}} f_L(v)dv \frac{\int_{y'}^{\overline{v}} v f_H(v)dv}{\int_{y'}^{\overline{v}} f_H(v)dv} - U(\lambda|\pi', \Gamma') \\
> \; & \int_{y'}^{\overline{v}} v f_H(v)dv - \int_{y'}^{\overline{v}} v f_H(v)dv - U(\lambda|\pi', \Gamma') \\
= \; & -U(\lambda|\pi', \Gamma'). \qquad (76)
\end{aligned}$$

Therefore,

$$R_L(\pi', \Gamma', 1) = R_L(\pi, \Gamma, \lambda), \qquad (77)$$

$$R_H(\pi', \Gamma', 1) > R_H(\pi, \Gamma, \lambda), \qquad (78)$$

$$U(1|\pi', \Gamma') > U(H, \lambda|\pi, \Gamma). \qquad (79)$$

Construct $\chi(\pi', \Gamma'|\pi, \Gamma, H) = 1$, and $\chi(\widetilde{\pi}, \widetilde{\Gamma}|\pi, \Gamma, H) = 0$ for any other $\left(\widetilde{\pi}, \widetilde{\Gamma}\right)$.

In case 3, $S = \{s_1, s_2\}$ and therefore

$$0 < \int_0^{\overline{v}} \pi(s_1|v) v f_H(v)dv < \int_0^{\overline{v}} v f_H(v)dv. \qquad (80)$$

Since $\int_y^{\overline{v}} v f_H(v)dv$ is strictly decreasing in y, there exists a unique y' such that

$$\int_{y'}^{\overline{v}} v f_H(v)dv = \int_0^{\overline{v}} \pi(s_1|v) v f_H(v)dv. \qquad (81)$$

Construct $(\pi', \Gamma') = \left(\pi^M(y'), (r(s_1', 1), -U(H, \lambda|\pi, \Gamma))\right)$. Therefore,

$$\begin{aligned}
& R_H(\pi', \Gamma', 1) \\
= \; & \int_{y'}^{\overline{v}} v f_H(v)dv - U(H, \lambda|\pi, \Gamma) \\
= \; & \int_0^{\overline{v}} \pi(s_1|v) v f_H(v)dv - U(H, \lambda|\pi, \Gamma) \\
= \; & R_H(\pi, \Gamma, \lambda), \qquad (82)
\end{aligned}$$

$$U\left(1|\pi',\Gamma'\right)=U\left(H,\lambda|\pi,\Gamma\right). \tag{83}$$

We want to show that

$$R_L\left(\pi',\Gamma',1\right)\leq R_L\left(\pi,\Gamma,\lambda\right), \tag{84}$$

with inequality strictly holds if $(\pi,\Gamma)\neq(\pi',\Gamma')$. First, we will show that if $\pi\neq\pi^M(y)$,

$$\int_{y'}^{\overline{v}} f_H(v)dv<\int_0^{\overline{v}}\pi\left(s_1|v\right)f_H(v)dv, \tag{85}$$

$$\frac{\int_{y'}^{\overline{v}} vf_H(v)dv}{\int_{y'}^{\overline{v}} f_H(v)dv}>\frac{\int_0^{\overline{v}}\pi\left(s_1|v\right)vf_H(v)dv}{\int_0^{\overline{v}}\pi\left(s_1|v\right)f_H(v)dv}, \tag{86}$$

$$\frac{\int_{y'}^{\overline{v}} f_L(v)dv}{\int_{y'}^{\overline{v}} f_H(v)dv}<\frac{p_L(s_1)}{p_H(s_1)}. \tag{87}$$

Suppose $\int_0^{y'}\pi(s_1|v)f_H(v)dv=0$.

$$\begin{aligned}\int_0^{\overline{v}}\pi(s_1|v)vf_H(v)dv &= \int_{y'}^{\overline{v}}\pi(s_1|v)vf_H(v)dv\\ &< \int_{y'}^{\overline{v}} vf_H(v)dv.\end{aligned} \tag{88}$$

The last inequality follows from that $\pi\neq\pi^M(y)$. Contradiction. Therefore, $\int_0^{y'}\pi(s_1|v)f_H(v)dv>0$. Since

$$\int_0^{y'}\pi(s_1|v)vf_H(v)dv<y'\int_0^{y'}\pi(s_1|v)f_H(v)dv, \tag{89}$$

$$\int_{y'}^{\overline{v}}\left(1-\pi\left(s_1|v\right)\right)vf_H(v)dv\geq y'\int_{y'}^{\overline{v}}\left(1-\pi\left(s_1|v\right)\right)f_H(v)dv, \tag{90}$$

$$\int_{y'}^{\overline{v}} v f_H(v)dv = \int_0^{\overline{v}} \pi(s_1|v) v f_H(v)dv$$

$$\Leftrightarrow \int_{y'}^{\overline{v}} \left(1-\pi\left(s_1|v\right)\right) v f_H(v)dv = \int_0^{y'} \pi(s_1|v) v f_H(v)dv$$

$$\Leftrightarrow y' \int_{y'}^{\overline{v}} \left(1-\pi\left(s_1|v\right)\right) f_H(v)dv < y' \int_0^{y'} \pi(s_1|v) f_H(v)dv$$

$$\Leftrightarrow \int_{y'}^{\overline{v}} \left(1-\pi\left(s_1|v\right)\right) f_H(v)dv < \int_0^{y'} \pi(s_1|v) f_H(v)dv$$

$$\Leftrightarrow \int_{y'}^{\overline{v}} f_H(v)dv < \int_0^{\overline{v}} \pi\left(s_1|v\right) f_H(v)dv. \tag{91}$$

Therefore, $(85), (86)$ hold. Denote

$$\int_{y'}^{\overline{v}} \pi(s_1|v) f_H(v)dv = a, \tag{92}$$

$$\int_0^{y'} \pi(s_1|v) f_H(v)dv = b, \tag{93}$$

$$\int_{y'}^{\overline{v}} (1-\pi(s_1|v)) f_H(v)dv = c, \tag{94}$$

$$\int_{y'}^{\overline{v}} \frac{f_L(v)}{f_H(v)} \pi(s_1|v) f_H(v)dv = a x_1, \tag{95}$$

$$\int_0^{y'} \frac{f_L(v)}{f_H(v)} \pi(s_1|v) f_H(v)dv = b x_2, \tag{96}$$

$$\int_{y'}^{\overline{v}} \frac{f_L(v)}{f_H(v)} \left(1-\pi(s_1|v)\right) f_H(v)dv = c x_3. \tag{97}$$

Note that we have $b > 0$ and $b > c$. Furthermore, Since $a + c = \int_{y'}^{\overline{v}} f_H(v)dv > 0$, either $a > 0$ or $c > 0$. Since $\frac{f_L(v)}{fH(v)}$ is decreasing in v, we have $bx_2 > b\frac{f_L(y')}{f_H(y')}$, and $ax_1 \leq a\frac{f_L(y')}{f_H(y')}$ with equality holds if and only if $a = 0$, and $cx_3 \leq c\frac{f_L(y')}{f_H(y')}$ with equality holds if and only if $c = 0$. Set $x_1 = \frac{f_L(y')}{f_H(y')}$ when $a = 0$, $x_3 = \frac{f_L(y')}{f_H(y')}$when $c = 0$. Then $x_2 > x_1, x_2 > x_3$. Given those relationships, we have

$$
\begin{aligned}
& \frac{\int_0^{\overline{v}} \pi(s_1|v) f_L(v)dv}{\int_0^{\overline{v}} \pi(s_1|v) f_H(v)dv} - \frac{\int_{y'}^{\overline{v}} f_L(v)dv}{\int_{y'}^{\overline{v}} f_H(v)dv} \\
= \;& \frac{ax_1 + bx_2}{a+b} - \frac{ax_1 + cx_3}{a+c} \\
= \;& \frac{acx_1 + abx_2 + bcx_2 - abx_1 - acx_3 - bcx_3}{(a+b)(a+c)} \\
\geq \;& \frac{acx_1 + abx_2 + bcx_2 - abx_1 - acx_2 - bcx_3}{(a+b)(a+c)} \quad (since\ x_2 > x_3) \\
= \;& \frac{a(b-c)(x_2 - x_1) + bc(x_2 - x_3)}{(a+b)(a+c)} \\
> \;& 0 \quad (since\ b > c, x_2 > x_1, x_2 > x_3, b > 0,\ and\ either\ a > 0\ or\ c > 0).
\end{aligned} \tag{98}
$$

Therefore, (87) holds. Second, there exists t, w such that $R_\theta(\pi, \Gamma, \lambda) = p_\theta(s_1) t + w$. Since

$$t \leq r(s_1, \lambda) \leq r(s_1, 1), \tag{99}$$

where the second inequality follows from Lemma 2 in Chen and Zhang (2018), and strictly holds if $\lambda < 1$,

$$p_H(s_1) t \leq \int_0^{\overline{v}} \pi(s_1|v) v f_H(v)dv = \int_{y'}^{\overline{v}} v f_H(v)dv, \tag{100}$$

with inequality strictly holds if $\lambda < 1$. Third, since

$$R_\theta(\pi', \Gamma', 1) = p_\theta(s_1') r(s_1', 1) - U(H, \lambda|\pi, \Gamma), \tag{101}$$

$$
\begin{aligned}
& R_L(\pi', \Gamma', 1) - R_H(\pi', \Gamma', 1) \\
= \;& \int_{y'}^{\overline{v}} f_L(v)dv \frac{\int_{y'}^{\overline{v}} v f_H(v)dv}{\int_{y'}^{\overline{v}} f_H(v)dv} - \int_{y'}^{\overline{v}} v f_H(v)dv \\
= \;& \left(\frac{\int_{y'}^{\overline{v}} f_L(v)dv}{\int_{y'}^{\overline{v}} f_H(v)dv} - 1 \right) \int_{y'}^{\overline{v}} v f_H(v)dv.
\end{aligned} \tag{102}
$$

Similarly,

$$R_L(\pi, \Gamma, \lambda) - R_H(\pi, \Gamma, 1) = \left(\frac{p_L(s_1)}{p_H(s_1)} - 1 \right) p_H(s_1) t. \tag{103}$$

Therefore

$$
\begin{aligned}
& R_L\left(\pi', \Gamma', 1\right) - R_H\left(\pi', \Gamma', 1\right) \leq R_L\left(\pi, \Gamma, \lambda\right) - R_H\left(\pi, \Gamma, 1\right) \\
\Leftrightarrow \quad & R_L\left(\pi', \Gamma', 1\right) \leq R_L\left(\pi, \Gamma, \lambda\right), \qquad (104)
\end{aligned}
$$

with inequality strictly holds if $\pi \neq \pi'$, or $\lambda < 1$, following (**??**), (100). If $(\pi, \Gamma) \neq (\pi', \Gamma')$, since both $\int_y^{\overline{v}} v f_H(v)dv, \int_y^{\overline{v}} f_L(v)dv \frac{\int_y^{\overline{v}} v f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv}$ are strictly decreasing in y, there exists $\epsilon > 0$ such that

$$R_L(\pi'', \Gamma'', 1) < R_L(\pi, \Gamma, \lambda), \qquad (105)$$

$$R_H(\pi'', \Gamma'', 1) > R_H(\pi, \Gamma, \lambda), \qquad (106)$$

$$U\left(1|\pi'', \Gamma''\right) = U\left(H, \lambda|\pi, \Gamma\right), \qquad (107)$$

where $(\pi'', \Gamma'') = \left(\pi^M\left(y' - \epsilon\right), \left(r\left(s_1'', 1\right), -U\left(H, \lambda|\pi, \Gamma\right)\right)\right)$. Construct $\chi(\pi', \Gamma'|\pi, \Gamma, H) = 1$, and $\chi(\widetilde{\pi}, \widetilde{\Gamma}|\pi, \Gamma, H) = 0$ for any other $\left(\widetilde{\pi}, \widetilde{\Gamma}\right)$.

In step 2, construct

$$\sigma^o\left(\pi^N, \left(r(s_1, 0), R_L(\sigma, \mu) - r(s_1, 0)\right)\right) = 1. \qquad (108)$$

Therefore,

$$R_L\left(\pi^N, \left(r(s_1, 0), R_L(\sigma, \mu) - r(s_1, 0)\right), 0\right) = R_L(\sigma, \mu), \qquad (109)$$

$$R_H\left(\pi^N, \left(r(s_1, 0), R_L(\sigma, \mu) - r(s_1, 0)\right), 0\right) = R_L(\sigma, \mu), \qquad (110)$$

$$U\left(0|\pi^N, \left(r(s_1, 0), R_L(\sigma, \mu) - r(s_1, 0)\right)\right) \geq \sum_{(\pi, \Gamma)} \sigma\left(\pi, \Gamma|L\right) U\left(L, \lambda|\pi, \Gamma\right). \qquad (111)$$

Construct

$$\sigma^o(\pi, \Gamma|\theta) = \sum_{(\widetilde{\pi}, \widetilde{\Gamma})} \chi(\pi, \Gamma|\widetilde{\pi}, \widetilde{\Gamma}, L)\sigma(\widetilde{\pi}, \widetilde{\Gamma}|H). \qquad (112)$$

By (73), (79), (107),

$$\sum_{(\pi, \Gamma)} \sigma^o\left(\pi, \Gamma|H\right) U\left(H, \lambda|\pi, \Gamma\right) \geq \sum_{(\pi, \Gamma)} \sigma\left(\pi, \Gamma|H\right) U\left(H, \lambda|\pi, \Gamma\right). \qquad (113)$$

Combined with (111), IR_A^μ holds. By (72), (78), (106)

$$R_H(\sigma^o, \mu) \geq R_H(\sigma, \mu), \qquad (114)$$

with inequality strictly holds if condition (b) does not hold for σ. By $(71), (77), (105)$,

$$R_L(\sigma^o, H, \mu) \leq R_L(\sigma, H, \mu), \tag{115}$$

and therefore $IC_{L,H}$ holds.

In step 3, the rest is to show $IC_{H,L}$. If it holds, then σ^o is the objective feasible mechanism. If it does not hold, then $R_H(\sigma^o, \mu) < R_H(\sigma^o, L, \mu) = R_L(\sigma^o, \mu)$. Construct

$$\sigma\prime \left(\pi^N, (r(s_1, 0), R_L(\sigma, \mu) - r(s_1, 0))\right) = 1, \tag{116}$$

$$\sigma\prime \left(\pi^N, (r(s_1, 1), R_L(\sigma, \mu) - r(s_1, 1))\right) = 1. \tag{117}$$

Since $R_L(\sigma^o, \mu) \leq \int_0^{\overline{v}} v f_{\mu_H}(v) dv$, IR_A^μ holds. It is feasible and

$$R_H(\sigma\prime, \mu) > R_H(\sigma^o, \mu) > R_H(\sigma, \mu). \tag{118}$$

Q.E.D.

Proof for Proposition 3 :

Consider a RSW mechanism σ. By Proposition 2,

$$R_L(\sigma) = \int_0^{\overline{v}} v f_L(v) dv \leq \int_0^{\overline{v}} v f(v) dv. \tag{119}$$

By Lemma 1, in a RSW mechanism, if condition (b) does not hold, there exists a feasible mechanism σ' such that $R_H(\sigma\prime) > R_H(\sigma, \mu)$, which is contradiction. Therefore, $\sigma(\pi, \Gamma|H) > 0$, only if for some threshold y,

$$\pi = \pi^M(y), \tag{120}$$

$$\Gamma = (r(s_1, 1), w). \tag{121}$$

By IC_A^H, $\sum_{\pi,\Gamma} \sigma(\pi, \Gamma|H) w(\Gamma) \leq 0$. The claim is verified.

Q.E.D.

Proof for Proposition 4:

We prove it by contradiction that if $\sigma^W(\pi^M(y), (r(s_1, 1), w))|H) > 0$, and $y > \underline{y(R_L(y))}$, there exists a σ' such that

$$R_H(\sigma') > R_H(\sigma), \tag{122}$$

$$R_L(\sigma') = R_L(\sigma), \tag{123}$$

$$U(\sigma') > U(\sigma). \tag{124}$$

Construct

$$\pi' = \pi^M\left(\underline{y(R_L(y))}\right), \tag{125}$$

$$w' = w + R_L(y) - R_L(\underline{y(R_L(y))}). \tag{126}$$

Therefore, $w' < w$, and

$$R_L(\pi', (r(s_1', 1), w'), 1) = R_L\left(\pi^M(y), (r(s_1, 1), w)), 1\right). \tag{127}$$

By definition of $\underline{y(R_L(y))}$,

$$R_H(\pi', (r(s_1', 1), w'), 1) - R_L(\pi', (r(s_1', 1), w'), 1) \tag{128}$$

$$> R_H\left(\pi^M(y), (r(s_1, 1), w)), 1\right) - R_L\left(\pi^M(y), (r(s_1, 1), w)), 1\right) R_H(y) \tag{129}$$

$$\Leftrightarrow R_H(\pi', (r(s_1', 1), w'), 1) > R_H\left(\pi^M(y), (r(s_1, 1), w)), 1\right). \tag{130}$$

Set

$$\chi\left(\widetilde{\pi}, \widetilde{\Gamma}|\pi^M(y), (r(s_1, 1), w)), H\right) = \begin{cases} 1, \text{ if } \left(\widetilde{\pi}, \widetilde{\Gamma}\right) = (\pi', (r(s_1', 1), w')), \\ \quad 0, \text{ otherwise}, \end{cases} \tag{131}$$

$$\chi\left(\widetilde{\pi}, \widetilde{\Gamma}|\pi, \Gamma\theta\right) = I_{\left(\widetilde{\pi}, \widetilde{\Gamma}\right)=(\pi,\Gamma)} \text{ if } (\pi, \Gamma) \neq \left(\pi^M(y), (r(s_1, 1), w))\right). \tag{132}$$

Construct

$$\sigma'(\pi, \Gamma|\theta) = \sum_{\widetilde{\pi}, \widetilde{\Gamma}} \chi\left(\widetilde{\pi}, \widetilde{\Gamma}|\pi, \Gamma, \theta\right) \sigma\left(\widetilde{\pi}, \widetilde{\Gamma}|\theta\right). \tag{133}$$

σ' is the objective mechanism.

Q.E.D.

Proof for Lemma 2 :

By Maximization theorem, $D_H(R_L)$ is continuous. Moreover, by definition of $\underline{y(R_L)}$, $D_H(R_L)$ is decreasing in R_L. Since $F_H(v) \succsim_{LR} F_L(v)$, $\int_y^{\overline{v}} f_H(v)dv - \int_y^{\overline{v}} f_L(v)dv \geq 0$ and equality holds if and only if $y = 0$ or $\overline{v}$. Since the solution to Problem I, $y_H^* < \overline{v}$, $D_H(R_L) \geq 0$ with equality satisfied if and only if $R_L = \int_0^{\overline{v}} v f_H(v)dv$.

Q.E.D.

Proof for Proposition 5:

By construction of Z_H, there exists a feasible mechanism σ^W with $\sigma^W(\pi, \Gamma|H) > 0$ only if for some threshold y, $\pi = \pi^M(y)$, $t = r(s_1, 1)$, $w \leq 0$ such that $R_H(\sigma^W) = R_H^W$. The rest is to show that for any σ, given $IC_{L,H}$, IR_A^H, $R_H(\sigma) \leq R_H^W$. By Proposition 3, in a RSW mechanism, $\sigma(\pi, \Gamma|H) > 0$ only if for some threshold y, $\pi = \pi^M(y), \Gamma = (r(s_1, 1), w)$. Denote $w^W = \sum_{(\pi,\Gamma)} \sigma(\pi, \Gamma|H)w$. Therefore, $w^W \leq 0$ and

$$R_H(\sigma) = w^W + \sum_{(y,w)} \sigma(y, w)|H) R_H(\pi^M(y), (r(s_1, 1), 0), 1), \tag{134}$$

$$R_L(\sigma, H) = w^W + \sum_{(y,w)} \sigma(y, w)|H) R_L(y). \tag{135}$$

By $\text{IC}_{L,H}$,

$$\sum_{(y,w)} \sigma(y, w|H) R_L(y) + w^W \leq R_L^W. \tag{136}$$

Therefore,

$$\begin{aligned}
& R_H(\sigma) \\
= \;& w^W + \sum_{(y,w)} \sigma(y, w|H) R_H(\pi^M(y), (r(s_1, 1), 0), 1) \\
= \;& w^W + \sum_{(y,w)} \sigma(y, w|H) R_L(y) + \sum_{(y,w)} \sigma(y, w|H)(R_H(\pi^M(y), (r(s_1, 1), 0), 1) - R_L(y)) \\
\leq \;& R_L^W + Z_H(R_L^W - w^W) \\
\leq \;& R_L^W + Z_H(R_L^W).
\end{aligned} \tag{137}$$

Q.E.D.

Proof for Proposition 6 :

By Proposition 3, for any $\sigma^W(\pi, \Gamma|H) > 0$ only two variables matter: y, w. By abuse of notation, we can write $\sigma^W(\pi, \Gamma|H) = \sigma^W(y, w|H)$. Now we can rewrite the problem for the high type:

$$Z_H(R_L^W) + R_L^W = \max_{\widetilde{\sigma}} \sum_{(y,w)} \widetilde{\sigma}(y, w|\theta) \left(\int_y^{\overline{v}} v f_H(v) dv + w \right), \text{(Problem II)}$$

$$s.t. \quad \sum_{(y,w)} \widetilde{\sigma}(y, w|\theta) \left(R_L(y) + w \right) = R_L^W, \tag{138}$$

$$\sum_{(y,w)} \widetilde{\sigma}(y, w|\theta) w \leq 0. \tag{139}$$

There are two cases. In case 1, $R_L^W \leq R_L(y_H^*)$. By definition of y_H^*, Problem II is uniquely solved by $\sigma^W(\pi^M(y_H^*), (r(s_1, 1), R_L^W - R_L(y_H^*))) = 1$. In case 2, $R_L^W > R_L(y_H^*)$. By Caratheodory's theorem, there exists $\sigma^W(\cdot)$ such that $\sigma^W(y, w|R_L) > 0$ for at most two (y, w). The rest is to show $\sum_{y,w} \sigma^W(y, w)w = 0$. Consider any $\left(\pi^M(y), (r(s_1, 1), w)\right)$ with $w < 0$. Since

$$\int_y^{\overline{v}} v f_H(v)dv - R_L(y) < \int_{y_H^*}^{\overline{v}} v f_H(v)dv - R_L(y_H^*), \tag{140}$$

$$\begin{aligned} & R_H\left(\pi^M(y), (r(s_1, 1), w), 1\right) \\ < \ & \alpha R_H(\pi^M(y), (r(s_1, 1), 0), 1) + (1 - \alpha) R_H(\pi^M(y_H^*), (r(s_1, 1), 0), 1), \end{aligned} \tag{141}$$

where

$$R_L\left(\pi^M(y), (r(s_1, 1), w), 1\right) = \alpha R_L(\pi^M(y), (r(s_1, 1), 0), 1) + (1 - \alpha) R_L(y_H^*). \tag{142}$$

The claim is verified.

Q.E.D.

Proof for Corollary 1 :

When $R_L > R_L(y_H^*)$, since $D_H(R_L)$ is decreasing in R_L, $Z_H(R_L)$ is strictly decreasing in R_L.

Q.E.D.

Proof for Lemma 3:

Take derivative of $DE_H(R_L)$ with respect to R_L,

$$\begin{aligned} \frac{\partial DE_H(R_L)}{\partial R_L} &= \frac{\partial DE_H(R_L)}{\partial y} \frac{\partial y}{\partial R_L} \\ &= \frac{y f_H(y)}{f_L(y) \frac{\int_y^{\overline{v}} v f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv} + \frac{f_H(y) \int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv} (y - \frac{\int_y^{\overline{v}} v f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv})} - 1 \\ &= \frac{1}{\frac{f_L(y)}{f_H(y)} \frac{\int_y^{\overline{v}} v f_H(v)dv}{y \int_y^{\overline{v}} f_H(v)dv} + \frac{\int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv} (1 - \frac{\int_y^{\overline{v}} v f_H(v)dv}{y \int_y^{\overline{v}} f_H(v)dv})} - 1. \end{aligned} \tag{143}$$

Denote

$$T(y) = \frac{f_L(y)}{f_H(y)} \frac{\int_y^{\overline{v}} v f_H(v)dv}{y \int_y^{\overline{v}} f_H(v)dv} + \frac{\int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv} (1 - \frac{\int_y^{\overline{v}} v f_H(v)dv}{y \int_y^{\overline{v}} f_H(v)dv}). \tag{144}$$

Note that

$$
\begin{aligned}
& T(y) \\
= \; & \left(\frac{f_L(y)}{f_H(y)} - \frac{\int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv}\right)\frac{\int_y^{\overline{v}} v f_H(v)dv}{y\int_y^{\overline{v}} f_H(v)dv} + \frac{\int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv} \\
\geq \; & \frac{\int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv} \\
\geq \; & 0. \qquad (145)
\end{aligned}
$$

Then $\frac{\partial DE_H(R_L)}{\partial R_L} = \frac{1}{T(y)} - 1$, $\frac{\partial y}{\partial R_L} = -\frac{1}{T(y)yf_H(y)}$. Therefore,

$$
\begin{aligned}
& \frac{\partial^2 DE_H(R_L)}{(\partial R_L)^2} \\
= \; & \frac{\partial \frac{\partial DE_H(R_L)}{\partial R_L}}{\partial y}\frac{\partial y}{\partial R_L} \\
= \; & \frac{1}{(T(y))^3 y f_H(y)}\frac{\partial T(y)}{\partial y}, \qquad (146)
\end{aligned}
$$

where

$$
\begin{aligned}
& \frac{\partial T(y)}{\partial y} \\
= \; & \frac{\partial \frac{f_L(y)}{f_H(y)}}{\partial y}\frac{\int_y^{\overline{v}} v f_H(v)dv}{y\int_y^{\overline{v}} f_H(v)dv} \qquad (147) \\
& -\left(\frac{f_L(y)}{f_H(y)} - \frac{\int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv}\right)\left(\frac{\int_y^{\overline{v}} v f_H(v)dv}{y^2\int_y^{\overline{v}} f_H(v)dv} + \frac{2f_H(y)\left(y - \frac{\int_y^{\overline{v}} v f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv}\right)}{y\int_y^{\overline{v}} f_H(v)dv}\right). \qquad (148)
\end{aligned}
$$

Note that $\frac{\partial \frac{f_L(y)}{f_H(y)}}{\partial y} < 0$, $\frac{f_L(y)}{f_H(y)} \leq \frac{\int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv}$. The rest is to show

$$\begin{aligned}
& \frac{\int_y^{\overline{v}} v f_H(v)dv}{y^2 \int_y^{\overline{v}} f_H(v)dv} + \frac{2f_H(y)(y - \frac{\int_y^{\overline{v}} v f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv})}{y \int_y^{\overline{v}} f_H(v)dv} \\
= & \frac{\int_y^{\overline{v}} v f_H(v)dv - 2y f_H(y)(\frac{\int_y^{\overline{v}} v f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv} - y)}{y^2 \int_y^{\overline{v}} f_H(v)dv} \\
\geq & \ 0. \qquad (149)
\end{aligned}$$

Denote $M(y) = \int_y^{\overline{v}} v f_H(v)dv - 2y f_H(y)(\frac{\int_y^{\overline{v}} v f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv} - y)$. Take derivative of $M(y)$ with respect y,

$$\begin{aligned}
& -y f_H(y) + 4y f_H(y) - 2y(\frac{\int_y^{\overline{v}} v f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv} - y)\frac{\partial f_H(y)}{\partial y} \\
& -2f_H(y)\frac{\int_y^{\overline{v}} v f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv} - 2\frac{y(f_H(y))^2}{\int_y^{\overline{v}} f_H(v)dv}(\frac{\int_y^{\overline{v}} v f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv} - y) \\
= & -2y(\frac{\int_y^{\overline{v}} v f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv} - y)\frac{\partial f_H(y)}{\partial y} - 2(\frac{\int_y^{\overline{v}} v f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv} - y) f_H(y) \\
& + y f_H(y)(1 - \frac{2f_H(y)}{\int_y^{\overline{v}} f_H(v)dv}(\frac{\int_y^{\overline{v}} v f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv} - y)) \\
\leq & \ y f_H(y)(1 - \frac{2f_H(y)}{\int_y^{\overline{v}} f_H(v)dv}(\frac{\int_y^{\overline{v}} v f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv} - y)) \\
\leq & \ y f_H(y)(1 - \frac{2y f_H(y)}{\int_y^{\overline{v}} v f_H(v)dv}(\frac{\int_y^{\overline{v}} v f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv} - y)) \\
= & \ \frac{y f_H(y) M(y)}{\int_y^{\overline{v}} v f_H(v)dv}. \qquad (150)
\end{aligned}$$

The second inequality strictly holds if $y < \overline{v}$. Therefore, if there exists some $\widetilde{y} < \overline{v}$ such that $M(\widetilde{y}) \leq 0$, $M(y) < 0, \forall y > \widetilde{y}$. Since $M(\overline{v}) = 0$, there is contradiction. Therefore, $M(y) \geq 0$.

Now we will show that $\frac{\partial^2 DE_H(R_L)}{(\partial R_L)^2} < 0$, with equality holds only if $y = \overline{v}$, implies that first-order condition (27) is necessary and sufficient. First-order derivative is

$$\begin{aligned}
& (f_L(y) - f_H(y))\frac{\int_y^{\overline{v}} vf_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv} + \frac{(F_L(y) - F_H(y))f_H(y)}{\int_y^{\overline{v}} f_H(v)dv}(\frac{\int_y^{\overline{v}} vf_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv} - y) \\
= \; & yf_H(y)[(\frac{f_L(y)}{f_H(y)} - 1)\frac{\int_y^{\overline{v}} vf_H(v)dv}{y\int_y^{\overline{v}} f_H(v)dv} + \frac{F_L(y) - F_H(y)}{\int_y^{\overline{v}} f_H(v)dv}(\frac{\int_y^{\overline{v}} vf_H(v)dv}{y\int_y^{\overline{v}} f_H(v)dv} - 1)] \\
= \; & yf_H(y)[(\frac{f_L(y)}{f_H(y)} - \frac{1 - F_L(y)}{\int_y^{\overline{v}} f_H(v)dv})\frac{\int_y^{\overline{v}} vf_H(v)dv}{y\int_y^{\overline{v}} f_H(v)dv} - \frac{F_L(y) - F_H(y)}{\int_y^{\overline{v}} f_H(v)dv}] \\
= \; & yf_H(y)(T(y) - 1). \qquad (151)
\end{aligned}$$

Since $\frac{\partial T(y)}{\partial y} \leq 0$, given $yf_H(y) > 0$ for any $y > 0$, first-order derivative crosses 0 once and only once. Therefore, first-order condition (27) is sufficient.

Q.E.D.

Proof for Lemma 4**:**

Take derivative of $DE_H(R_L)$ with respect to R_L,

$$\begin{aligned}
\frac{\partial DE_H(R_L)}{\partial R_L} &= \frac{\partial DE_H(R_L)}{\partial y}\frac{\partial y}{\partial R_L} \\
&= \frac{yf_H(y)}{f_L(y)\frac{\int_y^{\overline{v}} vf_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv} + \frac{f_H(y)\int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv}(y - \frac{\int_y^{\overline{v}} vf_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv})} - 1 \\
&= \frac{1}{\frac{f_L(y)}{f_H(y)}\frac{\int_y^{\overline{v}} vf_H(v)dv}{y\int_y^{\overline{v}} f_H(v)dv} + \frac{\int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv}(1 - \frac{\int_y^{\overline{v}} vf_H(v)dv}{y\int_y^{\overline{v}} f_H(v)dv})} - 1. \qquad (152)
\end{aligned}$$

As in Lemma 3, denote

$$T(y) = \frac{f_L(y)}{f_H(y)}\frac{\int_y^{\overline{v}} vf_H(v)dv}{y\int_y^{\overline{v}} f_H(v)dv} + \frac{\int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv}(1 - \frac{\int_y^{\overline{v}} vf_H(v)dv}{y\int_y^{\overline{v}} f_H(v)dv}). \qquad (153)$$

Note that $T(y) \geq 0$. Then $\frac{\partial DE_H(R_L)}{\partial R_L} = \frac{1}{T(y)} - 1$, $\frac{\partial y}{\partial R_L} = -\frac{1}{T(y)yf_H(y)}$. Therefore,

$$\begin{aligned} & \frac{\partial^2 DE_H(R_L)}{(\partial R_L)^2} \\ = \; & \frac{\partial \frac{\partial DE_H(R_L)}{\partial R_L}}{\partial y} \frac{\partial y}{\partial R_L} \\ = \; & \frac{1}{(T(y))^3 y f_H(y)} \frac{\partial T(y)}{\partial y}, \end{aligned} \tag{154}$$

where

$$\begin{aligned} & \frac{\partial T(y)}{\partial y} \\ = \; & \frac{\partial \frac{f_L(y)}{f_H(y)}}{\partial y} \frac{\int_y^{\overline{v}} v f_H(v) dv}{y \int_y^{\overline{v}} f_H(v) dv} - \left(\frac{f_L(y)}{f_H(y)} - \frac{\int_y^{\overline{v}} f_L(v) dv}{\int_y^{\overline{v}} f_H(v) dv}\right)\left(\frac{\int_y^{\overline{v}} v f_H(v) dv}{y^2 \int_y^{\overline{v}} f_H(v) dv} + \frac{2 f_H(y)\left(y - \frac{\int_y^{\overline{v}} v f_H(v) dv}{\int_y^{\overline{v}} f_H(v) dv}\right)}{y \int_y^{\overline{v}} f_H(v) dv}\right) \end{aligned} \tag{155}$$

If condition $(A2)$ holds,

$$\begin{aligned} & \frac{\int_y^{\overline{v}} v f_H(v) dv}{\int_y^{\overline{v}} f_H(v) dv} - \frac{\overline{v} + y}{2} \\ = \; & \frac{\int_y^{\overline{v}} v - \frac{\overline{v}+y}{2} f_H(v) dv}{\int_y^{\overline{v}} f_H(v) dv} \\ = \; & \frac{\int_y^{\frac{\overline{v}+y}{2}} v - \frac{\overline{v}+y}{2} f_H(v) dv + \int_{\frac{\overline{v}+y}{2}}^{\overline{v}} v - \frac{\overline{v}+y}{2} f_H(v) dv}{\int_y^{\overline{v}} f_H(v) dv} \\ \leq \; & \frac{f_H(\frac{\overline{v}+y}{2})\left(\int_y^{\frac{\overline{v}+y}{2}} v - \frac{\overline{v}+y}{2} dv + \int_{\frac{\overline{v}+y}{2}}^{\overline{v}} v - \frac{\overline{v}+y}{2} dv\right)}{\int_y^{\overline{v}} f_H(v) dv} \\ = \; & \frac{f_H(\frac{\overline{v}+y}{2}) \int_y^{\overline{v}} v - \frac{\overline{v}+y}{2} dv}{\int_y^{\overline{v}} f_H(v) dv} \\ = \; & 0, \\ \Leftrightarrow \; & \overline{v} - y \geq 2\left(\frac{\int_y^{\overline{v}} v f_H(v) dv}{\int_y^{\overline{v}} f_H(v) dv} - y\right). \end{aligned} \tag{156}$$

Similarly, if condition $(A2)$ holds, $\frac{f_L(y)}{f_H(y)} - \frac{f_L(\overline{v})}{f_H(\overline{v})} \geq 2(\frac{f_L(y)}{f_H(y)} - \frac{\int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv})$. If condition $(A3)$ holds,

$$
\begin{aligned}
& \int_y^{\overline{v}} v f_H(v)dv \geq (\overline{v} - y) y f_H(v) \\
\Leftrightarrow \quad & \frac{\int_y^{\overline{v}} v f_H(v)dv}{y^2 \int_y^{\overline{v}} f_H(v)dv} \geq \frac{(\overline{v} - y) f_H(y)}{y \int_y^{\overline{v}} f_H(v)dv}.
\end{aligned} \tag{157}
$$

If condition $(A4)$ holds,

$$
\begin{aligned}
& \frac{\frac{\partial \frac{f_L(v)}{f_H(v)}}{\partial v}}{f_H(v)} \geq \frac{\frac{\partial \frac{f_L(y)}{f_H(y)}}{\partial y}}{f_H(y)}, \forall v \geq y, \\
\Leftrightarrow \quad & \frac{\int_y^{\overline{v}} \frac{\frac{\partial \frac{f_L(v)}{f_H(v)}}{\partial v}}{f_H(v)} f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv} \geq \frac{\frac{\partial \frac{f_L(y)}{f_H(y)}}{\partial y}}{f_H(y)}, \\
\Leftrightarrow \quad & \frac{\int_y^{\overline{v}} \frac{\partial \frac{f_L(v)}{f_H(v)}}{\partial v} dv}{\int_y^{\overline{v}} f_H(v)dv} \geq \frac{\frac{\partial \frac{f_L(y)}{f_H(y)}}{\partial y}}{f_H(y)}, \\
\Leftrightarrow \quad & \frac{\frac{f_L(\overline{v})}{f_H(\overline{v})} - \frac{f_L(y)}{f_H(y)}}{\int_y^{\overline{v}} f_H(v)dv} \geq \frac{\frac{\partial \frac{f_L(y)}{f_H(y)}}{\partial y}}{f_H(y)}.
\end{aligned} \tag{158}
$$

Then if condition $(A2)$ and $(A3)$ hold,

$$
\begin{aligned}
& \frac{\partial \frac{f_L(y)}{f_H(y)}}{\partial y} \frac{\int_y^{\overline{v}} v f_H(v)dv}{y \int_y^{\overline{v}} f_H(v)dv} - (\frac{f_L(y)}{f_H(y)} - \frac{\int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv})(\frac{\int_y^{\overline{v}} v f_H(v)dv}{y^2 \int_y^{\overline{v}} f_H(v)dv} + \frac{2 f_H(y)(y - \frac{\int_y^{\overline{v}} v f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv})}{y \int_y^{\overline{v}} f_H(v)dv}) \\
< \quad & -(\frac{f_L(y)}{f_H(y)} - \frac{\int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv})(\frac{\int_y^{\overline{v}} v f_H(v)dv}{y^2 \int_y^{\overline{v}} f_H(v)dv} + \frac{2 f_H(y)(y - \frac{\int_y^{\overline{v}} v f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv})}{y \int_y^{\overline{v}} f_H(v)dv}) \\
\leq \quad & -(\frac{f_L(y)}{f_H(y)} - \frac{\int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv}) \frac{f_H(y)(\overline{v} - y + 2(y - \frac{\int_y^{\overline{v}} v f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv}))}{y \int_y^{\overline{v}} f_H(v)dv} \\
\leq \quad & 0.
\end{aligned} \tag{159}
$$

Similarly, if condition $(A2)$ and $(A4)$ hold,

$$\frac{\partial \frac{f_L(y)}{f_H(y)}}{\partial y} \frac{\int_y^{\overline{v}} v f_H(v)dv}{y\int_y^{\overline{v}} f_H(v)dv} - (\frac{f_L(y)}{f_H(y)} - \frac{\int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv})(\frac{\int_y^{\overline{v}} v f_H(v)dv}{y^2\int_y^{\overline{v}} f_H(v)dv} + \frac{2f_H(y)(y - \frac{\int_y^{\overline{v}} v f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv})}{y\int_y^{\overline{v}} f_H(v)dv})$$

$$\leq \quad 0.$$

If condition $(A3)$ and $(A4)$ hold,

$$\begin{aligned}
&\frac{\partial \frac{f_L(y)}{f_H(y)}}{\partial y} \frac{\int_y^{\overline{v}} v f_H(v)dv}{y\int_y^{\overline{v}} f_H(v)dv} - (\frac{f_L(y)}{f_H(y)} - \frac{\int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv})(\frac{\int_y^{\overline{v}} v f_H(v)dv}{y^2\int_y^{\overline{v}} f_H(v)dv} + \frac{2f_H(y)(y - \frac{\int_y^{\overline{v}} v f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv})}{y\int_y^{\overline{v}} f_H(v)dv}) \\
= \quad & \frac{\int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv} - \frac{f_L(y)}{f_H(y)} + \frac{\int_y^{\overline{v}} v f_H(v)dv}{y\int_y^{\overline{v}} f_H(v)dv}(\frac{\partial \frac{f_L(y)}{f_H(y)}}{\partial y} + (\frac{f_L(y)}{f_H(y)} - \frac{\int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv})\frac{f_H(y)}{\int_y^{\overline{v}} f_H(v)dv}) \\
& -(\frac{f_L(y)}{f_H(y)} - \frac{\int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv})(\frac{\int_y^{\overline{v}} v f_H(v)dv}{y^2\int_y^{\overline{v}} f_H(v)dv} + \frac{f_H(y)}{\int_y^{\overline{v}} f_H(v)dv}(1 - \frac{\int_y^{\overline{v}} v f_H(v)dv}{y\int_y^{\overline{v}} f_H(v)dv})) \\
\leq \quad & \frac{\int_y^{\overline{v}} v f_H(v)dv}{y\int_y^{\overline{v}} f_H(v)dv}(\frac{\partial \frac{f_L(y)}{f_H(y)}}{\partial y} + (\frac{f_L(y)}{f_H(y)} - \frac{\int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv})\frac{f_H(y)}{\int_y^{\overline{v}} f_H(v)dv}) \\
& -(\frac{f_L(y)}{f_H(y)} - \frac{\int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv})(\frac{\int_y^{\overline{v}} v f_H(v)dv}{y^2\int_y^{\overline{v}} f_H(v)dv} + \frac{f_H(y)}{\int_y^{\overline{v}} f_H(v)dv}(1 - \frac{\int_y^{\overline{v}} v f_H(v)dv}{y\int_y^{\overline{v}} f_H(v)dv})) \\
\leq \quad & \frac{\int_y^{\overline{v}} v f_H(v)dv}{y\int_y^{\overline{v}} f_H(v)dv}(\frac{f_L(\overline{v})}{f_H(\overline{v})} - \frac{f_L(y)}{f_H(y)} + \frac{f_L(y)}{f_H(y)} - \frac{\int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv})\frac{f_H(y)}{\int_y^{\overline{v}} f_H(v)dv} \\
& -(\frac{f_L(y)}{f_H(y)} - \frac{\int_y^{\overline{v}} f_L(v)dv}{\int_y^{\overline{v}} f_H(v)dv})(\frac{\overline{v} - y + y - \frac{\int_y^{\overline{v}} v f_H(v)dv}{\int_y^{\overline{v}} f_H(v)dv}}{y\int_y^{\overline{v}} f_H(v)dv}\frac{f_H(y)}{\int_y^{\overline{v}} f_H(v)dv}) \\
\leq \quad & 0.
\end{aligned} \tag{160}$$

The second inequality strictly holds if $y < \overline{v}$.

With the same argument in proof of Lemma 3, $\frac{\partial^2 DE_H(R_L)}{(\partial R_L)^2} \leq 0$ with equality holds only if $y = \overline{v}$. First-order condition (27) is necessary and sufficient. The claim is verified.

Q.E.D.

Proof for Theorem 1 :

Consider any σ^W in a RSW mechanism. By Proposition 2, 6, $(1,2)$ hold. By construction, σ^W is feasible.

Now we will show that it can be supported as PBE. Consider any off-equilibrium-path contract M. It corresponds to a direct mechanism σ. Construct the following belief:

$$\mu_H(M) = \begin{cases} 1, \text{ if } IR_A^H < 0, \\ 0, \text{ otherwise,} \end{cases} \tag{161}$$

There are two cases. In case 1, $IR_A^H < 0$. The buyer would not participate and therefore each type makes zero revenue by deviating to M. In case 2, $IR_A^H \geq 0$. If $IR_A^L < 0$, the buyer would not participate and therefore each type makes zero revenue by deviating to M. Otherwise, full-information participation constraint hold. Therefore, for any belief u,

$$R_L(\sigma, \mu) \leq \int_0^{\overline{v}} v f_L(v) dv \leq \int_0^{\overline{v}} v f_{\mu_H}(v) dv. \tag{162}$$

By Lemma 1, there exists a feasible mechanism, σ^o with conditions $(1-3)$ satisfied. σ^o is a candidate to the problem in definition of the RSW mechanism. By maximization, $R_\theta(\sigma, \mu(\sigma)) \leq R_\theta(\sigma^o, \mu(\sigma)) \leq R_\theta^W$. The claim is verified.

Q.E.D.

Proof for Proposition 9 :

First, denote $y(\alpha)$ be the cutoff such that $R_\alpha^W = R_\alpha(y(\alpha))$. Since $R_\alpha(y)$ is strictly decreasing in y, $y(\alpha)$ exists and is unique. When $y(\alpha) = 0$, $R_\alpha(y(\alpha)) = \int_0^{\overline{v}} v f_H(v) dv$. Therefore, $y(\alpha) \geq 0$ with equality holds only if $\alpha = 1$. Since

$$\begin{aligned} 0 &= R_\alpha^W - R_\alpha(y(\alpha)) \\ &= \alpha \int_0^{\overline{v}} v f_H(v) dv + (1-\alpha) R_L^W - \left(\alpha \int_{y(\alpha)}^{\overline{v}} v f_H(v) dv + (1-\alpha) R_L(y(\alpha)) \right) \\ &\geq (1-\alpha)\left(R_L^W - R_L(y(\alpha))\right), \end{aligned} \tag{163}$$

with equality holds only if $\alpha = 0$ or $y(\alpha) = 0$. Therefore, $R_L^W \leq R_L(y(\alpha))$ with inequality strictly holds if $\alpha \in (0,1)$. For $\alpha' > \alpha$,

$$
\begin{aligned}
& R^W_{\alpha'} - R_{\alpha'}(y(\alpha)) \\
= \; & \alpha' \int_0^{\overline{v}} v f_H(v) dv + (1-\alpha') R^W_L - \left(\alpha' \int_{y(\alpha)}^{\overline{v}} v f_H(v) dv + (1-\alpha') R_L(y(\alpha)) \right) \\
= \; & R^W_\alpha - R_\alpha(y(\alpha)) + (\alpha' - \alpha) \left(\int_0^{\overline{v}} v f_H(v) dv - R^W_L - \int_{y(\alpha)}^{\overline{v}} v f_H(v) dv + R_L(y(\alpha)) \right) \\
= \; & (\alpha' - \alpha) \left(\int_0^{\overline{v}} v f_H(v) dv - R^W_L - \int_{y(\alpha)}^{\overline{v}} v f_H(v) dv + R_L(y(\alpha)) \right) \\
> \; & 0, \qquad (164)
\end{aligned}
$$

Therefore, $y(\alpha') < y(\alpha)$. $y^*_H(\alpha) = y^*_H$ since

$$
\int_{\widetilde{y}}^{\overline{v}} v f_H(v) dv - R_\alpha(\widetilde{y}) = (1-\alpha) \left(\int_{\widetilde{y}}^{\overline{v}} v f_H(v) dv - R_L(\widetilde{y}) \right). \qquad (165)
$$

Then there exists $\underline{\alpha} \in [0,1]$ such that $R^W_\alpha > R_\alpha(y^*_H(\alpha))$ if and only if $\alpha > \underline{\alpha}$. $\underline{\alpha} < 1$ since $y(1) = 0$, $y^*_H > 0$.

Second, denote σ^W_α be a RSW mechanism when the low type is of type α and expected information fee

$$
w_H(\alpha) = \sum_{(y,w)} \sigma^W_\alpha (y, w|H) w. \qquad (166)
$$

When $\alpha > \underline{\alpha}$, $w_H(\alpha) = 0$. When $\alpha \leq \underline{\alpha}$,

$$
\begin{aligned}
& w_H(\alpha) \\
= \; & R^W_\alpha - R_\alpha(y^*_H(\alpha)) \\
= \; & R^W_\alpha - R_\alpha(y^*_H) \\
= \; & \alpha \int_0^{\overline{v}} v f_H(v) dv + (1-\alpha) R^W_L - \left(\alpha \int_{y^*_H}^{\overline{v}} v f_H(v) dv + (1-\alpha) R_L(y^*_H) \right) \\
\leq \; & 0, \qquad (167)
\end{aligned}
$$

and

$$
\begin{aligned}
& \frac{\partial w_H(\alpha)}{\partial \alpha} \\
= & \int_0^{\overline{v}} v f_H(v) dv - R_L^W - \left(\int_{y_H^*}^{\overline{v}} v f_H(v) dv - R_L(y_H^*) \right) \\
\geq & R_L(y_H^*) - R_L^W \\
\geq & 0. \qquad (168)
\end{aligned}
$$

Therefore, $w_H(\alpha)$ is strictly increasing when $\alpha \leq \underline{\alpha}$.

Third, denote the revenue of the high type in the RSW mechanism be $R_H^W(\alpha)$ when the low type is of type α. $R_H^W(\alpha)$ is continuous in α. When $\alpha \leq \underline{\alpha}$, $R_H^W(\alpha) = \int_{y_H^*}^{\overline{v}} v f_H(v) dv + w_H(\alpha)$. Since $w_H(\alpha)$ is strictly increasing, $R_H^W(\alpha)$ is strictly increasing. When $\alpha > \underline{\alpha}$, $w_H(\alpha) = 0$, and therefore by abuse of notation, write $\sigma_\alpha^W(y|H) = \sum_w \sigma_\alpha^W(y, w|H)$. Suppose $\alpha'' > \alpha' > \underline{\alpha}$.

$$
\begin{aligned}
0 &= R_{\alpha'}^W - \sum_y \sigma_{\alpha'}^W(y|H) R_{\alpha'}(y) \\
&\geq (1-\alpha') \left(R_L^W - \sum_y \sigma_{\alpha'}^W(y|H) R_L(y) \right), \qquad (169)
\end{aligned}
$$

with equality holds only if $\alpha' = 0$ or 1. Therefore, $R_L^W \leq \sum_y \sigma_{\alpha'}^W(y|H) R_L(y)$ with inequality strictly holds if $\alpha \in (0,1)$. Since

$$
\begin{aligned}
& R_{\alpha''}^W - \sum_y \sigma_{\alpha'}^W(y|H) R_{\alpha''}(y) \\
= & R_\alpha^W - \sum_y \sigma_{\alpha'}^W(y|H) R_{\alpha'}(y) + (\alpha' - \alpha) * \\
& \left(\int_0^{\overline{v}} v f_H(v) dv - R_L^W + \sum_y \sigma_{\alpha'}^W(y|H) \left(R_{\alpha'}(y) - \int_y^{\overline{v}} v f_H(v) dv \right) \right) \qquad (170) \\
= & (\alpha' - \alpha) \left(\int_0^{\overline{v}} v f_H(v) dv - R_L^W + \sum_y \sigma_{\alpha'}^W(y|H) \left(R_{\alpha'}(y) - \int_y^{\overline{v}} v f_H(v) dv \right) \right) \\
> & 0, \qquad (171)
\end{aligned}
$$

$\sigma^W_{\alpha'}(y|H)$ is feasible for the high type when the low type is of type α''. Since $R^W_{\alpha''} = \sum_y \sigma^W_{\alpha''}(y|H) R_{\alpha''}(y)$,

$$\begin{aligned} R^W_H(\alpha'') &= \sum_y \sigma^W_{\alpha''}(y|H) \int_y^{\overline{v}} v f_H(v) dv \\ &> \sum_y \sigma^W_{\alpha'}(y|H) \int_y^{\overline{v}} v f_H(v) dv \\ &= R^W_H(\alpha''). \end{aligned} \tag{172}$$

Forth, denote the social welfare be $SW(\alpha)$ when the low type is of type α. $SW(\alpha)$ is continuous in α. When $\alpha \leq \underline{\alpha}$,

$$SW(\alpha) = \mu^0_H \int_{y^*_H}^{\overline{v}} v f_H(v) dv + \left(1 - \mu^0_H\right) \int_0^{\overline{v}} v f_\alpha(v) dv. \tag{173}$$

Since $\int_0^{\overline{v}} v f_\alpha(v) dv$ strictly increases in α, $SW(\alpha)$ strictly increases. When $\alpha > \underline{\alpha}$,

$$SW(\alpha) = \mu^0_H R^W_H(\alpha) + \left(1 - \mu^0_H\right) \int_0^{\overline{v}} v f_\alpha(v) dv. \tag{174}$$

Since $R^W_H(\alpha)$ and $\int_0^{\overline{v}} v f_\alpha(v) dv$ strictly increase in α, $SW(\alpha)$ strictly increases.

Fifth, the buyer's exante payoff is $-\mu^0_H w_H(\alpha)$. Therefore, it is decreasing in α and is strictly decreasing when $\alpha < \underline{\alpha}$.

Q.E.D.

Proof for Proposition 10:

By Proposition 8, $R_\theta \geq R^W_\theta$. The rest is to show the upper bound of the high type. First, by $IR^{\mu^0}_A$,

$$\begin{aligned} \mu^0_H R_H + \mu^0_L R_L &\leq \int_0^{\overline{v}} v f(v) dv \\ \Leftrightarrow \quad R_H &\leq \frac{\int_0^{\overline{v}} v f(v) dv - \mu^0_L R_L}{\mu^0_H}. \end{aligned} \tag{175}$$

If $R_L > \int_0^{\overline{v}} v f(v) dv$,

$$R_H < \int_0^{\overline{v}} v f(v) dv. \tag{176}$$

Second, we consider $IC_{L,H}$. If $R_L \leq \int_0^{\overline{v}} vf(v)dv$,

$$R_L \leq \int_0^{\overline{v}} vf_H(v)dv$$

$$\Leftrightarrow \quad \frac{R_L - \mu_L^0 R_L^W}{\mu_H^0} \leq \int_0^{\overline{v}} vf_H(v)dv. \tag{177}$$

Therefore, $Z_H(\frac{R_L-\mu_L^0 R_L^W}{\mu_H^0})$ is valid. Since $Z_H(\frac{R_L-\mu_L^0 R_L^W}{\mu_H^0}) \geq 0$, $R_L + Z_H(\frac{R_L-\mu_L^0 R_L^W}{\mu_H^0}) \geq R_L$. By Lemma 1, we consider any feasible mechanism σ^o that satisfies conditions (1) and (2) in Lemma 1. Denote $w_H = \sum_{\pi,\Gamma} \sigma^o(\pi, \Gamma|H)w(\Gamma)$. By $IR_A^{\mu^0}$, since $R_L^W \leq R_L$, $w_H \leq 0$ and

$$R_L - R_L^W \leq -\frac{\mu_H^0}{\mu_L^0} w_H. \tag{178}$$

By abuse of notation, we can write $\sigma^o(\pi, \Gamma|H) = \sigma^o(y, w|H)$. Therefore, for any R_L, the high type makes maximal revenue R_H:

$$\max R_H = \sum_{\widetilde{y},\widetilde{w}} \sigma^o(\widetilde{y}, \widetilde{w}|H) \int_{\widetilde{y}}^{\overline{v}} vf_H\left(v\right)dv + w_H \tag{179}$$

$$s.t. \quad R_L \geq \sum_{\widetilde{y},\widetilde{w}} \sigma^o(\widetilde{y}, \widetilde{w}|H) R_L\left(\widetilde{y}\right) + w_H, \tag{180}$$

$$R_L - R_L^W \leq -\frac{\mu_H^0}{\mu_L^0} w_H, \tag{181}$$

$$w_H \leq 0. \tag{182}$$

By definition of Z_H,

$$R_H \leq R_L + Z_H(R_L - w_H). \tag{183}$$

Since $Z_H(R_L - w_H)$ is increasing in w_H, $R_H \leq R_L + Z_H(\frac{R_L-\mu_L^0 R_L^W}{\mu_H^0})$. Moreover, by definition of Z_H,

$$R_L + Z_H(\frac{R_L - \mu_L^0 R_L^W}{\mu_H^0}) \leq \frac{\int_0^{\overline{v}} vf(v)dv - \mu_L^0 R_L}{\mu_H^0}. \tag{184}$$

Q.E.D.

Proof for Proposition 11 :

We prove by construction. Consider any $[R_L, R_H]$ such that $(44, 45)$ holds. Since $R_L \leq \int_0^{\overline{v}} vf(v)dv$, by definition of Z_H, there exists a feasible σ^o with $\sigma^o(\pi^N, (\int_0^{\overline{v}} vf(v)dv, R_L -$

$\int_0^{\overline{v}} vf(v)dv)|L) = 1$ such that

$$R_L(\sigma^o) = R_L, \tag{185}$$

$$\sum_{\widetilde{y},\widetilde{w}} \sigma^o(\widetilde{y},\widetilde{w}|H)\widetilde{w} = -\frac{\mu_L^0\left(R_L - R_L^W\right)}{\mu_H^0} \tag{186}$$

$$\sum_{\widetilde{y},\widetilde{w}} \sigma^o(\widetilde{y},\widetilde{w}|H)\left(R_L\left(\widetilde{y}\right) + \widetilde{w}\right) = R_L, \tag{187}$$

$$\sum_{\widetilde{y},\widetilde{w}} \sigma^o(\widetilde{y},\widetilde{w}|H)\left(\int_{\widetilde{y}}^{\overline{v}} vf_H\left(v\right)dv + \widetilde{w}\right) = R_L + Z_H(\frac{R_L - \mu_L^0 R_L^W}{\mu_H^0}). \tag{188}$$

Construct

$$\sigma(\pi^N, (\int_0^{\overline{v}} vf(v)dv, R_L - \int_0^{\overline{v}} vf(v)dv)|L) = 1, \tag{189}$$

$$\sigma(y, -\frac{\mu_L^0\left(R_L - R_L^W\right)}{\mu_H^0} - R_H(\sigma^o) + R_H|H) = \sum_{\widetilde{w}} \sigma^o(y,\widetilde{w}|H). \tag{190}$$

Therefore, $R_\theta(\sigma) = R_\theta$. For any $\sigma^o(y,w|H) > 0$, since $R_H \le R_L + Z_H(\frac{R_L - \mu_L^0 R_L^W}{\mu_H^0})$,

$$\begin{aligned} & w \\ = \; & -\frac{\mu_L^0\left(R_L - R_L^W\right)}{\mu_H^0} - R_H(\sigma^o) + R_H \\ \le \; & -\frac{\mu_L^0\left(R_L - R_L^W\right)}{\mu_H^0} \\ \le \; & 0. \end{aligned} \tag{191}$$

Therefore $IR_A^{\mu^0}$, $IC_{L,H}$ hold. Since

$$R_H(\sigma, L, \mu^0) = R_L \le R_H, \tag{192}$$

$IC_{H,L}$ holds. Therefore, σ is feasible and the objective mechanism. Since $R_\theta \ge R_\theta^W$, σ can be supported as PBE with off-equilibrium-path belief constructed in Theorem 1.

Q.E.D.

Proof for Theorem 2 :

First, we show that condition (1) is necessary. Suppose in contrary, $R_L^W < R_L(y_H^*)$. In the RSW mechanism,

$$\sigma^W(\pi^M(y_H^*), (r(s_1, 1), R_L^W - R_L(y_H^*))|H) = 1 \tag{193}$$

and $R_L^W - R_L(y_H^*) < 0$. There exists $\epsilon > 0$ such that $\epsilon < \mu_H^0 \left(R_L(y_H^*) - R_L^W\right)$. Construct

$$\sigma(\pi^N, (r(s_1, 1), \epsilon)|L) = 1, \tag{194}$$

$$\sigma(\pi^M(y_H^*), (r(s_1, 1), R_L^W - R_L(y_H^*) + \epsilon)|H) = 1. \tag{195}$$

It is feasible and strictly Pareto dominates the RSW mechanism.

Second, suppose condition (1) holds. By Proposition 10, the RSW mechanism is undominated if

$$R_H^W \geq \begin{cases} \int_0^{\overline{v}} vf(v)dv, \text{ if } R_L > \int_0^{\overline{v}} vf(v)dv, \\ R_L + Z_H(\frac{R_L - \mu_L^0 R_L^W}{\mu_H^0}), \text{ otherwise.} \end{cases} \tag{196}$$

with equality holds only if $R_L = R_L^W$. Since $R_L + Z_H(\frac{R_L - \mu_L^0 R_L^W}{\mu_H^0}) = \int_0^{\overline{v}} vf(v)dv$ at $R_L = \int_0^{\overline{v}} vf(v)dv$. It is sufficient to show that for any $R_L^W < R_L \leq \int_0^{\overline{v}} vf(v)dv$,

$$R_H^W > R_L + Z_H(\frac{R_L - \mu_L^0 R_L^W}{\mu_H^0}). \tag{197}$$

It is equivalent to show that $R_L + Z_H(\frac{R_L - \mu_L^0 R_L^W}{\mu_H^0})$ achieves maximal only at $R_L = R_L^W$. Take first-order and second-order derivative with respect to R_L where it is differentiable,

$$FOC(R_L) = 1 + \frac{1}{\mu_H^0} \frac{\partial Z_H(x)}{\partial x}|_{x = \frac{R_L - \mu_L^0 R_L^W}{\mu_H^0}}, \tag{198}$$

$$SOC(R_L) = \frac{1}{(\mu_H^0)^2} \frac{\partial^2 Z_H(x)}{\partial x^2}|_{x = \frac{R_L - \mu_L^0 R_L^W}{\mu_H^0}} \leq 0 \tag{199}$$

Therefore, condition (2) is sufficient. The rest is to show condition (2) is necessary. If $FOC(R_L) = 0$ when $SOC(R_L) = 0$, or $FOC(R_L) > 0$ at $R_L = R_L^W$, there exists $\varepsilon > 0$ such that

$$R_L^W + \varepsilon + Z_H(\frac{R_L^W + \varepsilon - \mu_L^0 R_L^W}{\mu_H^0}) \geq R_H^W \tag{200}$$

and therefore there exists a PBE that generates revenue $\left[R_L^W + \varepsilon,_L^W + \varepsilon + Z_H(\frac{R_L^W + \varepsilon - \mu_L^0 R_L^W}{\mu_H^0})\right]$ by Proposition 11. Thus condition (2) holds necessarily.

Q.E.D.

Proof for Proposition 12 :

Consider any off-equilibrium-path contract M, it corresponds to a direct mechanism σ. Construct off-equilibrium-path belief

$$\mu(M) = \begin{cases} 1, \text{ if } \max_{\widetilde{\mu}} R_H(\sigma, \widetilde{\mu}) \geq R_H^W \text{ and } \max_{\widetilde{\mu}} R_L(\sigma, \widetilde{\mu}) \leq R_L^W, \\ 0, \text{ otherwise.} \end{cases} \tag{201}$$

The constructed off-equilibrium-path belief satisfies the intuitive criterion. For any σ, there are four cases. In case 1, $IR_A^\theta < 0, \forall\theta$, and therefore, $R_\theta(\sigma, \mu) = 0$ for any μ. In case 2, $IR_A^\theta \geq 0, \forall\theta$. Therefore, for any belief u,

$$R_L(\sigma, \mu) \leq \int_0^{\overline{v}} v f_L(v) dv \leq \int_0^{\overline{v}} v f_{\mu_H}(v) dv. \tag{202}$$

By Lemma 1, there exists a feasible σ^o satisfies conditions $(1-3)$. Therefore by definition of RSW mechanism, $R_\theta(\sigma, \mu(\sigma)) \leq R_\theta(\sigma^o, \mu(\sigma)) \leq R_\theta^W$. In case 3, $IR_A^H < 0$, $IR_A^L \geq 0$. If $\mu(\sigma) = 1$, $R_\theta(\sigma, \mu(\sigma)) = 0$. If $\mu(\sigma) = 0$, then $R_L(\sigma, 0) \leq R_L^W$. By Lemma 1, there exists a feasible σ^o satisfies conditions $(1-3)$. By $IC_{L,H}$ of σ^o, $R_L(\sigma, H, 0) \leq R_L(\sigma, 0) \leq R_L^W$. Therefore, by definition of RSW mechanism, $R_H(\sigma, 0) \leq R_H^W$. In case 4, $IR_A^L < 0$, $IR_A^H \geq 0$. If $\mu(\sigma) = 0$, $R_\theta(\sigma, \mu(\sigma)) = 0$. If $\mu(\sigma) = 1$, then $\max_{\widetilde{\mu}} R_H(\sigma, \widetilde{\mu}) \geq R_H^W$ and $\max_{\widetilde{\mu}} R_L(\sigma, \widetilde{\mu}) \leq R_L^W$. Therefore, $R_L(\sigma, 1) \leq R_L^W$. Similar to arguments in case 3, $R_H(\sigma, \mu) \leq R_H^W$. The claim is verified.

Q.E.D.

Proof for Proposition 13**:**

Sufficiency follows from Proposition 12. The rest is to show necessary part. Consider any σ with $R_H(\sigma) < R_L(\sigma) + Z_H(R_L(\sigma))$, we will show that there exists some other mechanism σ' that blocks it for the high type with intuitive criterion. By definition of $Z_H(R_L)$, there exists a σ' that satisfies conditions (1) and (2) in Lemma 1 such that

$$\sigma'(\pi^N, (0, R_L(\sigma) - \epsilon)|L) = 1, \tag{203}$$

$$\sum_{\widetilde{y},\widetilde{w}} \sigma'(\widetilde{y}, \widetilde{w}|H) \left(R_L\left(\widetilde{y}\right) + \widetilde{w}\right) = R_L(\sigma) - \epsilon, \tag{204}$$

$$\sum_{\widetilde{y},\widetilde{w}} \sigma'(\widetilde{y}, \widetilde{w}|H) \left(\int_{\widetilde{y}}^{\overline{v}} v f_H\left(v\right) dv + \widetilde{w}\right) = R_L(\sigma) - \epsilon + Z_H(R_L(\sigma) - \epsilon), \tag{205}$$

$$\sum_{\widetilde{y},\widetilde{w}} \sigma'(\widetilde{y}, \widetilde{w}|H)\widetilde{w} \leq 0, \tag{206}$$

where by abuse of notation, $\sigma\prime(\pi, \Gamma|H) = \sigma\prime(y, w|H)$. By (206) , $U(\sigma', 1) \geq 0$. By (204) , $IC_{L,H}$ hold. Since $Z_H(R_L(\sigma) - \epsilon) \geq 0$, $IC_{H,L}$ holds. By continuity, there exists $\epsilon > 0$ such that $R_H(\sigma', 1) > R_H(\sigma)$. Since $\forall\mu$,

$$R_L(\sigma', \mu) \leq R_L(\sigma) - \epsilon, \tag{207}$$

by intuitive criterion, $\mu(\sigma) = 1$. Since $R_H(\sigma', 1) > R_H(\sigma)$, the high type deviates. By Proposition 10,

$$R_H(\sigma) \leq \begin{cases} \int_0^{\overline{v}} v f(v) dv, \text{ if } R_L > \int_0^{\overline{v}} v f(v) dv, \\ R_L + Z_H(\frac{R_L - \mu_L^0 R_L^W}{\mu_H^0}), \text{ otherwise.} \end{cases} . \tag{208}$$

Therefore, $R_L(\sigma) = R_L^W$, $R_H(\sigma) = R_H^W$ by Corollary 1. The rest to show that σ is the RSW mechanism. It follows from Lemma 1 and definition of RSW mechanism.

References

[1] Alonso, R., and Câmara, O. On the value of persuasion by experts.

[2] Balestrieri, F. A modified english auction for an informed buyer. In *GAME THEORY AND MANAGEMENT. Collected abstracts of papers presented on the Second International Conference Game Theory and Management/Editors Leon A. Petrosjan, Nikolay A. Zenkevich.–SPb.: Graduate School of Management SPbU, 2008.–236 p. The collection contains abstracts of papers accepted for the International* (2008), p. 14.

[3] Balestrieri, F., and Izmalkov, S. Informed seller in a hotelling market.

[4] Balkenborg, D., and Makris, M. An undominated mechanism for a class of informed principal problems with common values. *Journal of Economic Theory 157* (2015), 918–958.

[5] Bedard, N. C. The strategically ignorant principal. *Games and Economic Behavior 102* (2017), 548–561.

[6] Cella, M. Informed principal with correlation. *Games and Economic Behavior 64*, 2 (2008), 433–456.

[7] Cho, I.-K., and Kreps, D. M. Signaling games and stable equilibria. *Quarterly Journal of Economics 102* (1987), 179–221.

[8] Eso, P., and Szentes, B. Optimal information disclosure in auctions and the handicap auction. *Review of Economic Studies 74*, 3 (July 2007), 705–731.

[9] Francetich, A., and Troyan, P. Collusion agreements in auctions: Design and execution by an informed principal. Tech. rep., Mimeo, 2012.

[10] Hedlund, J. Bayesian persuasion by a privately informed sender. Tech. rep., Working paper, 2015.

[11] Hoffmann, F., and Inderst, R. Pre-sale information. *Journal of Economic Theory 146*, 6 (2011), 2333–2355.

[12] Kamenica, E., and Gentzkow, M. Bayesian persuasion. *The American Economic Review 101*, 6 (2011), pp. 2590–2615.

[13] Koessler, F., and Skreta, V. Informed seller with taste heterogeneity. *Journal of Economic Theory 165* (2016), 456–471.

[14] Koessler, F., and Skreta, V. Selling with evidence.

[15] Li, H., and Shi, X. Discriminatory information disclosure. *American Economic Review 107*, 11 (2017), 3363–85.

[16] Maskin, E., and Tirole, J. The principal-agent relationship with an informed principal: The case of private values. *Econometrica: Journal of the Econometric Society* (1990), 379–409.

[17] Maskin, E., and Tirole, J. The principal-agent relationship with an informed principal, ii: Common values. *Econometrica 60*, 1 (January 1992), 1–42.

[18] Myerson, R. B. Mechanism design by an informed principal. *Econometrica 51*, 6 (November 1983), 1767–97.

[19] Mylovanov, T., and Tröger, T. Informed-principal problems in environments with generalized private values. *Theoretical Economics 7*, 3 (2012), 465–488.

[20] Mylovanov, T., and Tröger, T. Mechanism design by an informed principal: Private values with transferable utility. *The Review of Economic Studies 81*, 4 (2014), 1668–1707.

[21] Nishimura, T. Informed principal problems in bilateral trading.

[22] Perez-Richet, E. Interim bayesian persuasion: First steps. *The American Economic Review 104*, 5 (2014), 469–474.

[23] Severinov, S. An efficient solution to the informed principal problem. *Journal of Economic Theory 141*, 1 (July 2008), 114–133.

[24] Skreta, V. Mechanism design for arbitrary type spaces. *Economics Letters 91*, 2 (2006), 293–299.

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

[26] Tan, G. Optimal procurement mechanisms for an informed buyer. *Canadian Journal of Economics* (1996), 699–716.

[27] Yilankaya, O. A note on the seller's optimal mechanism in bilateral trade with two-sided incomplete information. *Journal of Economic Theory 87*, 1 (1999), 267–271.