A Necessary Condition for Robust Implementation: Theory and Applications*

Takuro Yamashita†

September 26, 2011

Abstract

We derive a necessary condition, called the *chain dominance property*, for social choice correspondences to be *admissibly implementable*, i.e., given whatever admissible actions the agents play in each state, the outcome always lies in the correspondence. This requires that the correspondence has a selection that is partially dominant-strategy incentive compatible in a certain sense. Moreover, it can also be sufficient: In worst-case expected welfare maximization in bilateral trading, we show that (i) for a class of priors of the designer, no mechanism can improve over a posted-price mechanism of Hagerty and Rogerson (1987), and (ii) for another class of priors, a non-dominant-strategy mechanism, called a "two-price" mechanism, is optimal.

*I am grateful to Ilya Segal, Matthew O. Jackson, Jonathan D. Levin, Paul Milgrom, Andy Skrzypacz, Koichi Tadenuma, Hideshi Itoh, Eve Ramaekers, Stephen Morris, Alessandro Pavan, Silvio Micali, Muhamet Yildiz, Andrea Prat, Dirk Bergemann, Gabriel D. Carroll, seminar participants at Stanford University, Hitotsubashi University, Yokohama National University, Duke University, University of Rochester, École Polytechnique, California Institute of Technology, Toulouse School of Economics, Boston University, and University of California Davis, for their valuable comments. I gratefully acknowledge financial support from the B.F. Haley and E.S. Shaw Fellowship for Economics through the Stanford Institute for Economic Policy Research.

†Toulouse School of Economics. tytakuroy@gmail.com

1 Introduction

Mechanism design theory examines which social objectives (such as efficiency, fairness, stability, and so on) can be achieved when agents have private information. To predict the possible outcomes of mechanisms, the standard approach is to assume that the agents play a Bayesian-Nash equilibrium (typically with a "common prior").

This Bayesian-Nash approach is often criticized for the sensitivity of the predicted outcomes of mechanisms to the assumptions about the agents' beliefs.[1][2] Namely, this approach relies on the mechanism designer's knowledge of the agents' beliefs about each other's private information, and their (correct) beliefs about each other's strategies. A mechanism that induces "desirable" outcomes (given any objective of the mechanism designer) in a Bayesian Nash equilibrium may induce undesirable outcomes if the agents have different beliefs about each other's private information or strategies.

Given these criticisms, some researchers have investigated more "robust"

[1]For example, in the context of game theory, Wilson (1987) argues:

> Game theory has a great advantage in explicitly analysing the consequences of trading rules that presumably are really common knowledge; it is deficient to the extent it assumes other features to be common knowledge, such as one agent's probability assessment about another's preferences or information. [. . .] I foresee the progress of game theory as depending on successive reduction in the base of common knowledge required to conduct useful analyses of practical problems. Only by repeated weakening of common knowledge assumptions will the theory approximate reality.

See also Neeman (2004) and Bergemann and Morris (2005).

[2]A related problem is that the optimal mechanism is sensitive to the assumptions on the agents' beliefs. For example, Crémer and McLean (1985) show that the first-best efficiency with full-surplus extraction is possible if there is a commonly known correlated prior over the agents' valuations (see d'Aspremont, Cremer, and Gerard-Varet (2004) and Kosenok and Severinov (2008) for similar first-best results under budget balance). Neeman (2004) argues that this result crucially depends on the "beliefs-determine-preferences" assumption, and Heifetz and Neeman (2006) show that this beliefs-determine-preferences property is "non-generic" in a more general type space (see also Barelli (2009)).

mechanisms. The standard approach is to restrict attention to mechanisms that is dominant-strategy incentive compatible. This proves to be restrictive, especially in settings that require a balanced budget.[3]

In this paper, we study *admissible implementation*, as an implementation concept that is robust to the agents' "strategic uncertainty". Namely, we assume that each agent may play any admissible (i.e., not weakly dominated) action given his private information. He may have multiple admissible actions in a mechanism, and therefore, there could be multiple possible outcomes depending on which admissible actions the agents play given their types. We say that the mechanism admissibly implements a social choice correspondence (SCC) if, given whatever admissible actions the agents play, the induced outcome lies in this SCC.

In the literature, Jackson (1992) suggests that we should focus on "bounded" mechanisms to study admissible implementation.[4] He shows that an "unbounded" mechanism can admissibly implement essentially any social choice correspondence, but he argues that implementation by unbounded mechanisms does not seem reasonable, because an unbounded mechanism necessarily has a "tail-chasing" or an "integer-game" structure. Following Jackson (1992), in this paper, we focus on bounded mechanisms.

Jackson (1992) also shows that any social choice *function* that is admissibly implementable (by bounded mechanisms) must be dominant-strategy incentive compatible. Thus, as long as there is a unique desirable allocation rule, our solution concept is equivalent to the dominant-strategy implementation.

However, if the objective of the mechanism designer is implementation of social choice *correspondences*, or similarly, if the objective is maximization of his "utility" (such as welfare or profit), then the restriction to the dominant-strategy mechanisms could be unnecessary. In this problem, the mechanism

[3]See Laffont and Maskin (1980) and Hagerty and Rogerson (1987), for example.

[4]A bounded mechanism is such that, for any action that is weakly dominated for an agent, there is an admissible action that weakly dominates it. For example, a mechanism is bounded if its message spaces are finite.

designer does not care which particular outcome is realized, as long as every possible outcome is desirable. Thus, there may exist a non-dominant-strategy mechanism that "robustly" achieves the objective, even if any dominant-strategy mechanism cannot achieve it.

Indeed, in Section 2, we provide a bilateral-trading example in which a non-dominant-strategy mechanism implements an SCC that is not implementable by any dominant-strategy mechanism.[5] Moreover, if the designer's objective is to maximize expected welfare (or total surplus) based on his prior over the agents' types, then for some priors, we find a non-dominant-strategy mechanism that always attains higher expected welfare than that of any dominant-strategy mechanism, given whatever admissible actions the agents play in the mechanism.

The main objective of the paper is to derive a necessary condition for admissibly implementable SCCs. In Section 3, we show that any implementable SCC must have the "chain dominance property", which is described as follows: First, fix any sequence of types for each agent. If an SCC is implementable, then for any profile of such sequences, we can find a selection of the SCC (i.e., an allocation rule that lies in the SCC) that satisfies dominant-strategy incentive compatibility along the sequences: In this selection, each agent prefers the truth-telling to pretending to be the type that is the immediate predecessor of the true type, given any types of the opponents. Thus, this selection satisfies dominant-strategy incentive compatibility for some pairs of types, but not necessarily for all pairs.[6]

In general, the chain dominance property may not be a sufficient condition for admissible implementation, even if we check all possible sequences of types. However, in some cases, we can guess which incentive constraints implied by the chain dominance property are binding, solve a relaxed problem

[5]By Hagerty and Rogerson (1987), a dominant-strategy mechanism in this example must be a (randomized) posted price mechanism.

[6]This condition generalizes the "strategy resistance" condition shown by Jackson (1992) as a necessary condition on implementable SCCs, which corresponds to the sequences with only two elements.

subject to these constraints, and verify that the allocation rule that solves the relaxed problem implies a "revelation mechanism" that admissibly implements the desired social choice correspondence. This is straightforward if the allocation rule that solves the relaxed problem is dominant-strategy incentive compatible, but this can also be true even if the solution is not dominant-strategy incentive compatible, when any admissible "lies" by the agents in the revelation mechanism induce desirable outcomes in any state.

As an application, in Section 4, we consider a one-dimensional single-crossing environment. The mechanism designer has a prior over the agents' types, and wants to maximize his expected "utility" (such as welfare or profit) given his belief. He does not know which admissible actions the agents play in a mechanism, and therefore, he evaluates a mechanism according to its "worst-case" expected utility among all admissible strategies of the agents. In this situation, we consider the "local downward incentive compatibility" (LDIC) constraints, the incentive constraints implied by the natural chains over the types. We show that, under certain conditions on the environment, the allocation rule that maximizes the designer's expected utility subject to the LDIC constraints implies a revelation mechanism that is optimal among all (bounded) mechanisms in the sense of the worst-case expected utility.

Specifically, we first study (balanced-budget) bilateral trading settings, where there exist a seller and a buyer, and each agent's value for trade is his private information. The designer wants to maximize the worst-case expected welfare based on his prior over the agents' values. We show the following: (i) For a class of priors, the optimal mechanism is a posted-price mechanism. This class of priors includes any prior such that its density function is decreasing in the seller's type, increasing in the buyer's type, and continuous. Because a posted-price mechanism is a dominant-strategy mechanism, this result provides a foundation for dominant-strategy mechanisms. (ii) When each agent's type space is binary, then the optimal mechanism is a "two-price" mechanism (except for trivial cases), which is not a dominant-strategy mechanism. This means that our approach sometimes yields a mechanism that is strictly better than any dominant-strategy mechanism even if we evaluate this mechanism in its worst-case scenario.

In the second application, we consider a quasi-linear environment without balanced budget, which includes expected revenue maximization in an auction setting. We provide sufficient conditions under which the optimal mechanism is a dominant-strategy mechanism.

1.1 Other robust implementation concepts

As a related concept to admissibility, some papers study implementation with iterative elimination of weakly or strictly dominated actions, but in complete-information settings. For example, see Moulin (1979), Srivastava and Trick (1996), Bergemann, Morris, and Tercieux (2010), and Abreu and Matsushima (1992). **?** and Kunimoto and Serrano (2010) study incomplete-information settings, but with a common prior over the agents' types. In this paper, we allow only one round of elimination of weakly dominated actions. This is a more robust concept than theirs in the sense that we do not impose any assumption on the agents beliefs about each other's preference or their mutual knowledge of rationality.

Another branch of the implementation literature studies implementation concepts robust to "structural uncertainty", i.e., agents know each other's strategies (and so they play a Bayesian Nash equilibrium), but they do not know each other's private information and beliefs (and higher-order beliefs) about this information. Bergemann and Morris (2005) and Bergemann and Morris (2010) study Bayesian Nash implementation with arbitrary beliefs in general implementation settings. Bergemann and Morris (2005) show that, in a "separable" environment, robustness to the structural uncertainty implies strategy-proofness. Bergemann and Morris (2010) show that, under certain conditions, their robust implementation concept is equivalent to "rationalizable" implementation, which is based on iterative elimination of strictly dominated actions in incomplete-information settings. Chung and Ely (2007) study the worst-case expected revenue maximization in auction settings, where the worst case is among all beliefs of the agents. They show that no mechanism can attain strictly higher expected revenue than the optimal dominant-strategy mechanism given any beliefs of the agents. Smith (2010) studies (balanced-budget) cost sharing problems in public good provision, and offers a partial ranking of mechanisms based on the notion of improvement given arbitrary beliefs of the agents. He shows that any dominant-strategy mechanism is weakly improvable in his criterion.

2 Example: Bilateral Trading

2.1 Environment

There is a pair of a seller ($i = 1$) and a buyer ($i = 2$). The seller has an object, and $c \in [0, 4]$ denotes his value for the object. $v \in [3, 5]$ denotes the buyer's valuation for the object. We assume that c is the seller's private information, v is the buyer's private information, and $\Theta_i \subset \mathbb{R}_+$ is compact for each i.

An allocation is denoted by $(z, p) \in [0, 1] \times \mathbb{R}$, where z is the probability of trading,[7] and p is the price, or the payment from the buyer to the seller conditional on trading (i.e., in the event that the buyer receives the object from the seller).[8] Let $(0, 0)$ denote the "no-trade" outcome.

The seller's utility and the buyer's utility at state (c, v) are given by $u_1 = (p-c)z$ and $u_2 = (v-p)z$, respectively, and the economic welfare at state (c, v) is $(v-c)z$. The mechanism designer has a prior Φ over $(c, v) \in \Theta$, and wants to maximize the expected welfare. We also assume that, in any mechanism, each agent has a message that corresponds to "non-participation" that always induces the no-trade outcome regardless of the opponent's message.

This problem is studied more extensively in Section 4. In this part, we consider the following specific class of Φ, parametrized by $\varepsilon \in [0, 1]$: There are two states with probability mass: $\Pr\{(c, v) = (1, 3)\} = \Pr\{(c, v) = (4, 5)\} = \frac{1}{2}(1 - \varepsilon)$. All the other $(c, v) \in [0, 4] \times [3, 5]$ are uniformly likely (i.e., the

[7]Another interpretation is that z represents the time of trading in a continuous-time dynamic bargaining setting where the agents' types are persistent. Suppose that the mechanism designer can specify the time of trading τ in a continuous-time model where the agents have the same discount rate r. Then, by setting z so that $z = e^{-r\tau}$, an allocation in this dynamic model is denoted by (z, p), and therefore, we can effectively design the same mechanism. Copic and Ponsati (2008) provide a similar interpretation of (randomized) posted-price mechanisms of Hagerty and Rogerson (1987) in such a dynamic bargaining setting.

[8]Thus, the allocation satisfies balanced budget. There is no payment when they do not trade.

density is $\frac{1}{8}\varepsilon$). Thus, if $\varepsilon = 1$, it is a uniform distribution, and if ε is close to zero, then it is approximately a discrete and perfectly correlated case. The mechanism designer knows the value of ε.

2.2 Dominant-strategy mechanisms

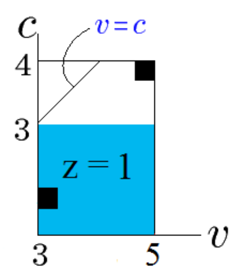
An optimal mechanism among all dominant-strategy mechanisms is a posted-price mechanism (Hagerty and Rogerson (1987)): The mechanism designer first chooses a price p, and a trade occurs (with probability one) if and only if $v > p$ and $c < p$. Thus, the expected welfare of a posted-price mechanism is

$$\int_{(c,v)\in[0,p)\times(p,1]} (v-c)\; d\Phi(c,v).$$

In this example, we observe that the optimal posted-price is $p = 3$ for any ε.[9]

The expected welfare of this mechanism is

$$\frac{15}{4}\varepsilon + (\frac{1}{2}-\varepsilon)\cdot 2 = 1 + \frac{7\varepsilon}{4}.$$



[9]Because the distribution is a convex combination of a uniform distribution on $[0,4]\times[3,5]$ and a discrete, perfectly correlated types (i.e., $\Pr\{(c,v) = (1,3)\} = \Pr\{(c,v) = (4,5)\} = \frac{1}{2}$), it suffices to show that the optimal posted-price is $p = 3$ for each of these distributions. First, with a uniform distribution on $[0,4]\times[3,5]$, the expected welfare with price p is $\frac{5}{16}p(5-p)$, which is maximized at $p = 3$. Second, if $\Pr\{(c,v) = (1,3)\} = \Pr\{(c,v) = (4,5)\} = \frac{1}{2}$, then any $p \in (1,3)$ is optimal. Therefore, for any ε, the supremum of the welfare among all posted-price mechanisms is achieved by a sequence of posted-price mechanisms $p \uparrow 3$. In this section, we informally say that $p = 3$ is the optimal posted price.

Figure 3: The optimal posted-price mechanism

The figure shows when the agents trade in the optimal posted-price mechanism: They trade whenever the seller's cost is lower than 3. Observe that there is a probability mass at $(c, v) = (4, 5)$ who cannot trade in the posted-price mechanism.

2.3 A two-price mechanism

This observation motivates us to consider the following "two-price" mechanism (Table 1). In this mechanism, the seller chooses a price, either $p = 3$ or $p = 4$. Simultaneously, the buyer reports his "highest acceptable price", $\overline{p} = 3$ or $\overline{p} = 4$. If the seller chooses $p = 3$ and $\overline{p} \geq 3$, then they trade with $z = 1$. If the seller chooses $p = 4$ and $\overline{p} = 4$, then they trade with $z = \frac{2}{3}$. If $\overline{p} < p$, then they do not trade.

	$\overline{p} = 3$	$\overline{p} = 4$
$(z_1, p_1) = (\frac{2}{3}, 4)$	$(0, 0)$	$(\frac{2}{3}, 4)$
$(z_2, p_2) = (1, 3)$	$(1, 3)$	$(1, 3)$

Table 1: A two-price mechanism

In this mechanism, the buyer has a dominant strategy: if $v > 4$, then $\overline{p} = 4$, and otherwise, $\overline{p} = 3$. On the other hand, the seller's best action depends on c and "his belief about the buyer's choice of $\overline{p}$". If $c \geq 3$, then it is weakly dominant to choose $p = 4$. If $c \in (1, 3)$, then (i) if he is "optimistic", i.e., if he believes that the buyer chooses $\overline{p} = 4$, then his best action is to choose $p = 4$, because it yields a higher expected profit than choosing $p = 3$ (i.e., $\frac{2}{3}(4 - c) > 3 - c$ for $c \in (1, 3)$). On the other hand, (ii) if he is "pessimistic", i.e., if he believes that the buyer chooses $\overline{p} = 3$, then his best action is to choose $p = 3$, because $p = 4$ would not be acceptable for the buyer.

Finally, if $c \leq 1$, then it is weakly dominant for the seller to choose $p = 3$, because even if he believes that the buyer chooses $\overline{p} = 4$, the expected profit of choosing $p = 3$ is higher than that of choosing $p = 4$ (i.e., $3 - c \geq \frac{2}{3}(4 - c)$ for $c \leq 1$).

We now calculate the level of expected welfare "guaranteed" given whatever admissible strategies the agents play in the mechanism. Observe that the worst-case expected welfare among all admissible strategies is attained when the seller with $c > 1$ chooses $p = 4$, because then, the trade (and hence the welfare) is smaller than when he chooses $p = 3$, regardless of the buyer's behavior (see Figure 4). Therefore, the worst-case expected welfare of this two-price mechanism is

$$\frac{11}{4}\varepsilon + (\frac{1}{2} - \varepsilon) \cdot \frac{8}{3} = \frac{4}{3} + \frac{\varepsilon}{12}.$$

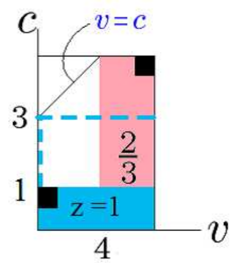


Figure 4: The worst-case welfare in the two-price mechanism

In this two-price mechanism, the seller with cost $c \in (3, 4)$ can trade (if $v > 4$), while he could not in the posted-price mechanism. Hence, the two-price mechanism attains higher welfare in these states. On the other hand, the seller with $c \in (1, 3)$ may deviate to the high-price allocation, which decreases the welfare in these states. Which mechanism guarantees a higher expected welfare depends on the value of ε. Specifically, if $\varepsilon < \frac{1}{5}$, then the two-price mechanism is better than the posted-price mechanism with $p = 3$. If $\varepsilon > \frac{1}{5}$, then the posted-price mechanism is better.

Moreover, as we see in Section 4, if $\varepsilon = 1$ so that the distribution is uniform over $[0, 4] \times [3, 5]$, then the posted-price with $p = 3$ is the optimal mechanism (in the sense of the worst-case expected welfare). Similarly, if $\varepsilon = 0$ so that each agent has a binary type space, then the two-price mechanism we examined is the optimal mechanism.

It is also interesting to compare these worst-case expected welfares with the expected welfare of the optimal Bayesian-Nash mechanism, where we assume that Φ is common knowledge among the agents and the mechanism designer. When $\varepsilon = \frac{1}{2}$, then the optimal Bayesian-Nash mechanism is a double auction mechanism studied by Myerson and Satterthwaite (1983) and Chatterjee and Samuelson (1983). This mechanism attains 96% of the first-best expected welfare,[10] while, the posted-price mechanism with $p = 3$ attains 93% of the first-best expected welfare. This 3% difference can be interpreted as the "price of robustness": To make a mechanism robust to the agents' strategic uncertainty, we lose this amount of expected welfare.

On the other hand, if $\varepsilon = 0$, then the agents have perfectly correlated types, and thus, the optimal Bayesian-Nash mechanism can achieve the first-best welfare, as studied by Crémer and McLean (1985) and Kosenok and Severinov (2008). For example, the following mechanism works.

	$v = 3$	$v = 5$
$c = 4$	$(0, 0)$	$(1, 4)$
$c = 1$	$(1, 3)$	$(1, 3)$

We can interpret this mechanism as a two-price mechanism where the seller chooses between $p = 3$ and $p = 4$, but regardless of the price chosen, the probability of trading is always one. In a common-prior Bayesian-Nash equilibrium, the seller with $c = 1$ reports his cost truthfully, because he believes for sure that the buyer reports $v = 3$.

However, if the mechanism designer is concerned about the worst case when the agents take any admissible strategies, z should be made smaller to have the seller with sufficiently lower costs choose the low price. The highest worst-case expected welfare is 89% of the first-best welfare, and therefore, the "price of robustness" is 11%. Note that it would be 17% if we were restricted

[10]Specifically, we set the probability of trading is $z(c, v) = 1$ if $v > c + 0.89$ and zero otherwise. This is derived by maximizing the weighted virtual surplus as in Myerson and Satterthwaite (1983).

only to dominant-strategy mechanisms. This 6% difference quantifies the welfare loss due to the restriction to dominant-strategy mechanisms.

The discussion is summarized in Table 2.

Robust Welfare Guarantee	Uniform ($\varepsilon = \frac{1}{2}$)	Two-state ($\varepsilon = 0$)
Posted-price	$\underline{93\%}$	83%
Two-price	68%	$\underline{89\%}$
Optimal Mechanism	$\underline{93\%}$	$\underline{89\%}$
Common-prior, Bayesian-Nash	96%*	100%**
Price of Robustness	3%	11%***

* Myerson and Satterthwaite (1983), Chatterjee and Samuelson (1983).

** Kosenok and Severinov (2008), Crémer and McLean (1988).

*** 17% if restricted to dominant-strategy mechanisms

Table 2: Summary of the example

In Section 4, we characterize the optimal mechanisms under more general conditions. First, for a class of priors including uniform distributions, we show that the optimal mechanism is a posted-price mechanism. Second, if each agent has binary types, then the optimal mechanism is a two-price mechanism.

3 Environment

There are N agents. Each agent $i = 1, \dots, N$ has private information $\theta_i \in \Theta_i$, where Θ_i is agent i's type space. Let $\Theta = \prod_i \Theta_i$.

An allocation is denoted by $x \in X$. Agent i's utility function is $u_i : X \times \Theta_i \to \mathbb{R}$. We assume that u_i does not depend on θ_{-i} (private values).

The objective of the mechanism designer is to implement a social choice correspondence (or SCC) $F : \Theta \to 2^X$. For each state θ, $F(\theta) \subseteq X$ is interpreted as the set of desirable outcomes in that state.

A mechanism is denoted by $\Gamma = \langle M, g \rangle$, where $M = \prod_i M_i$, each M_i is a set of messages for agent i, and $g : M \to X$ is called an outcome function.

We say that $m_i \in M_i$ *weakly dominates* $m_i' \in M_i$ for θ_i, if for any $m_{-i} \in M_{-i}$,

$$u_i(g(m_i, m_{-i}), \theta_i) \geq u_i(g(m_i', m_{-i}), \theta_i),$$

and the inequality is strict for at least one m_{-i}. m_i is said to be *admissible* for θ_i if m_i is not weakly dominated for θ_i. Let $M_i^A(\theta_i)$ denote the set of admissible messages for θ_i.

In this paper, we only consider the following class of mechanisms, called "bounded mechanisms" (Jackson (1992)).

Definition 1. Γ is *bounded* if the following is satisfied: For each i and θ_i, if m_i is weakly dominated for θ_i, then there is $m_i' \in M_i^A(\theta_i)$ that weakly dominates m_i (i.e., m_i' itself is not weakly dominated).

Note that, in a bounded mechanism, $M_i^A(\theta_i)$ is nonempty.

The following are some examples of bounded mechanisms. First, a finite mechanism (i.e., a mechanism such that every M_i is finite) is bounded. More generally, a "compact and continuous mechanism" (i.e., a mechanism such that M_i is a compact metric space for each i, and $u_i(g(m), \theta_i)$ is continuous in $m \in M$ for each i and θ_i) is bounded. The third example is a dominant-strategy mechanism, i.e., $M_i = \Theta_i$ for each i, and for each $\theta_i, \theta_i' \in \Theta_i$, $\theta_{-i} \in$

Θ_{-i},

$$u_i(g(\theta_i, \theta_{-i}), \theta_i) \geq u_i(g(\theta_i', \theta_{-i}), \theta_i).$$

We study admissible implementation (by bounded mechanisms), as a robust implementation concept to the agents' strategic uncertainty. Admissible implementation requires that, given whatever admissible actions the agents take in any state, the induced outcome is desirable.

Definition 2. A mechanism Γ *admissibly* implements F if for each θ and each $m \in M^A(\theta)$, $g(m) \in F(\theta)$.

4 The chain dominance property

In this section, we derive a necessary condition on admissibly implementable SCCs, which we call the *chain dominance property*.

A chain on Θ_i is a finite sequence of agent i's types, $C_i = \{\theta_i^t\}_{t=0}^{T_i}$, such that $\theta_i^s \neq \theta_i^t$ for $s \neq t$. Let $C = (C_i)_{i=1}^N$ denote a profile of such chains. An allocation rule $f : \Theta \to X$ is called a selection of an SCC F if for each θ, $f(\theta) \in F(\theta)$.

Definition 3. An SCC F has the *chain dominance property* if, for any profile of chains $C = (C_i)_{i=1}^N$, there exists a selection f of F such that, for each i, $t = 1, \ldots, T_i$, and $\theta_{-i} \in \Theta_{-i}$,

$$u_i(f(\theta_i^t, \theta_{-i}), \theta_i^t) \geq u_i(f(\theta_i^{t-1}, \theta_{-i}), \theta_i^t).$$

The condition means that we have the dominant-strategy incentive compatibility along the chains.[11]

[11] The chain dominance property generalizes the "strategy resistance" of Jackson (1992), which can be interpreted as the chain dominance conditions stated only for the chains with two elements (i.e., $T_i = 1$). As Jackson (1992) has shown, when F is a social choice function (i.e., $F(\theta) = \{f(\theta)\}$ for any θ), then f must be dominant-strategy incentive compatible.

Theorem 1. If an SCC F is admissibly implementable, then F has the chain dominance property.

Proof. We first show the following lemma, proved by Jackson (1992).

Lemma 1. Let $\Gamma = \langle M, g \rangle$ be a bounded mechanism. For any i,θ_i and θ_i', suppose that $m_i \in M_i^A(\theta_i)$. Then, for any $\theta_i' \neq \theta_i$, there exists $m_i' \in M_i^A(\theta_i')$ such that for any $m_{-i} \in M_{-i}$,

$$u_i(g(m_i', m_{-i}), \theta_i') \geq u_i(g(m_i, m_{-i}), \theta_i'),$$

Proof. (of Lemma 1)

For θ_i', either $m_i \in M_i^A(\theta_i')$ or $m_i \notin M_i^A(\theta_i')$.

If $m_i \in M_i^A(\theta_i')$, let $m_i' = m_i$. Then the inequality is satisfied with equality for any $m_{-i} \in M_{-i}$.

If $m_i \notin M_i^A(\theta_i')$, then m_i is weakly dominated by some $m_i' \in M_i^A(\theta_i')$ because Γ is bounded. Thus, m_i' satisfies the inequality for any $m_{-i} \in M_{-i}$.

□

Let $\Gamma = \langle M, g \rangle$ be a mechanism that admissibly implements F. For each i, let $C_i = \{\theta_i^t\}_{t=1}^{T_i}$ be an arbitrary chain on Θ_i.

For each i, we construct $\mu_i : \Theta_i \to M_i$ in the following procedure. For the initial type θ_i^0, let $\mu_i(\theta_i^0)$ be an arbitrary element in $M_i^A(\theta_i^0)$. By induction, for each $t = 1, \ldots, T_i$, given $\mu_i(\theta_i^{t-1}) \in M_i^A(\theta_i^{t-1})$, Lemma 1 implies that there is $\mu_i(\theta_i^t) \in M_i^A(\theta_i^t)$ such that, for any $m_{-i} \in M_{-i}$,

$$u_i(g(\mu_i(\theta_i^t), m_{-i}), \theta_i^t) \geq u_i(g(\mu_i(\theta_i^{t-1}), m_{-i}), \theta_i^{t-1}).$$

Let $\mu = (\mu_i)_{i=1}^N$. Define $f : \Theta \to X$ so that $f(\theta) = g(\mu(\theta))$ for $\theta \in \Theta$. Because each $\mu_i(\theta_i) \in M_i^A(\theta_i)$, we have $f(\theta) \in F(\theta)$. Also, for each i, $t = 1, \ldots, T_i$, and $\theta_{-i} \in \Theta_{-i}$,

$$u_i(f(\theta_i^t, \theta_{-i}), \theta_i^t) \geq u_i(f(\theta_i^{t-1}, \theta_{-i}), \theta_i^t).$$

□

In general, the tree dominance property need not be a sufficient condition. However, as we see in Section 4, we can sometimes "guess" the chain profile that induces a selection f such that a "revelation mechanism" $\langle \Theta, f \rangle$ admissibly implements F. This is straightforward if f is dominant-strategy incentive compatible. However, even if f is not dominant-strategy incentive compatible, if any admissible lies of the agents in $\langle \Theta, f \rangle$ always induce desirable outcomes, then $\langle \Theta, f \rangle$ admissibly implements F.

5 Local downward incentive compatibility

This section applies the findings in the previous section to some economic environments. In the following, let $X \subseteq \prod_{i=1}^{N} X_i$ be the set of allocations, where $(z_i, t_i) \in X_i \subseteq \mathbb{R}^2$ denotes the payoff relevant component for agent i. Also, we assume that, for each i, Θ_i is a compact subset of $\mathbb{R}$, and $u_i = \theta_i z_i + t_i$. For example, some trading settings with or without balanced budget are included, as we see in Section 4.2 and 4.3.

In this one-dimensional, single-crossing environment, we study implications of some "natural" chain dominance conditions.

5.1 Finite type spaces

We first assume that each Θ_i is finite. For each i, consider a chain of types $C_i = (\theta_i^t)_{t=0}^{T_i}$ such that $\theta_i^s < \theta_i^t$ for $s < t$. Theorem 1 implies the following result.

Theorem 2. If a mechanism Γ admissibly implements F, then there is a selection $f : \Theta \to X$ of F such that for each i, $t = 1, \ldots, T_i$, and $\theta_{-i} \in \Theta_{-i}$,

$$\theta_i^t z_i(\theta_i^t, \theta_{-i}) + t_i(\theta_i^t, \theta_{-i}) \geq \theta_i^t z_i(\theta_i^{t-1}, \theta_{-i}) + t_i(\theta_i^{t-1}, \theta_{-i}),$$

where $f(\theta) = (z_i(\theta), t_i(\theta))_{i=1}^N$.

$\langle \Theta, f \rangle$ can be interpreted as a revelation mechanism that satisfies the *local downward incentive compatibility* (LDIC): Each agent of each type has no incentive to pretend to be the "locally" smaller type, because truth-telling is always weakly better than such a deviation. Specifically, the truth-telling either (i) weakly dominates pretending to be the adjacent smaller type, or (ii) he is indifferent between the two.

5.2 Continuous type spaces with finite mechanisms

For simplicity, we assume that $\Theta_i = [0, 1]$ for each i. First, we consider implementation by finite mechanisms. In a finite mechanism, each agent's type space is partitioned into finitely many "strategically equivalent" types in the following sense.

Lemma 2. In a finite mechanism $\Gamma = \langle M, g\rangle$, each agent's type space is partitioned into finitely many connected subsets, $\{\Theta_i^{k_i}\}_{k_i=1}^{T_i}$ for each i, such that any types in the same partition have the same ordinal preference on $g(M) = \{g(m)|m \in M\}$, i.e., for each $x, x' \in g(M)$, $\theta_i, \theta_i' \in \Theta_i^{k_i}$,

$$u_i(x, \theta_i) \geq u_i(x', \theta_i) \Leftrightarrow u_i(x, \theta_i') \geq u_i(x', \theta_i').$$

As a corollary, we obtain $M_i^A(\theta_i) = M_i^A(\theta_i')$ for $\theta_i, \theta_i' \in \Theta_i^{k_i}$. Without loss of generality, we assume $\Theta_i^{k_i} < \Theta_i^{k_i+1}$ in the following. Interpreting each $\Theta_i^{k_i}$ as an ordinary preference type on $g(M)$, we obtain an analogous result as with finite type spaces.

Lemma 3. Suppose that a finite mechanism Γ admissibly implements F, and for each i, let $P_i = \{\Theta_i^{k_i}\}_{k_i=1}^{T_i}$ denote the partitions of strategically equivalent types induced by Γ. Let $P = \prod_i P_i$ and $k = (k_i)_{i=1}^N$. Then, there exist $\tilde{f} : P \to X$ such that (i) for each $\theta \in \Theta^k = \prod_i \Theta_i^{k_i}$, $\tilde{f}(\Theta^k) \in F(\theta)$, and (ii) for each i and $k = (k_i, k_{-i})$,

$$\theta_i^{k_i} z_i^k + t_i^k \geq \theta_i^{k_i} z_i^{k_i-1,k_{-i}} + t_i^{k_i-1,k_{-i}},$$

where $\tilde{f}(\Theta^k) = (z_i^k, t_i^k)_{i=1}^N$.

As in the case with finite type spaces, we can interpret $\langle P, f\rangle$ as a revelation mechanism where each agent reports $\Theta_i^{k_i}$ as the set of equivalent types in which his true type exists, and the inequalities mean the local downward incentive compatibility ("local" in the sense of the equivalent types). We call these inequalities the "ordinal LDIC condition".

The ordinal LDIC condition implies the following, which is proved to be useful in some applications.

Theorem 3. Suppose that an SCC F is admissibly implemented by a finite mechanism. Then, there is a selection $f = (z, t) : \Theta \to X$ of F that satisfies the following. For each i, θ_i, θ_i' and θ_{-i},

$$U_i(\theta) \geq U_i(\theta_i', \theta_{-i}) + \int_{\theta_i'}^{\theta_i} z_i(t, \theta_{-i})\ dt,$$

where $U_i(\theta) = \theta_i z_i(\theta) + t_i(\theta)$.

This is an integral form of the LDIC condition. It is well known that if f is dominant-strategy incentive compatible, then the same condition holds, but with equality (i.e., the change in each agent's utility is exactly pinned down by $z(\cdot)$).[12]

The idea of the proof is the following. Let $P_i = \{\Theta_i^{k_i}\}_{k_i=1}^{T_i}$ denote the partitions of strategically equivalent types induced by Γ, and $\tilde{f} : P \to X$ be the selection of F in Lemma 3.

For each i and k_i, let $\theta_i^{k_i} = \inf \Theta_i^{k_i}$ be the lower limit of the equivalent types $\Theta_i^{k_i}$. In the following, we assume that every $\Theta_i^{k_i}$ is left-closed (i.e., $\theta_i^{k_i} \in \Theta_i^{k_i}$). The proof for the general case is in the appendix.

Proof. By the ordinal LDIC condition:

$$\theta_i^{k_i} z_i^k + t_i^k \geq \theta_i^{k_i} z_i^{k_i-1,k_{-i}} + t_i^{k_i-1,k_{-i}},$$

where $\tilde{f}(\Theta^k) = (z_i^k, t_i^k)_{i=1}^N$.

Define an allocation rule $f = (z_i, t_i)_{i=1}^N$ so that, for each i and $\theta \in \Theta^k$,

$$(z_i(\theta), t_i(\theta)) = (z_i^k, t_i^k).$$

For each i and threshold types θ^k,

$$\theta_i^{k_i}(z_i^k - z_i^{k_i-1,k_{-i}}) + t_i^k - t_i^{k_i-1,k_{-i}} \geq 0.$$

Summing both sides for $j = k_i' + 1, \ldots, k_i$,

$$\sum_{j=k_i'+1}^{k_i} \theta_i^j(z_i^{j,k_{-i}} - z_i^{j-1,k_{-i}}) + t_i^{j,k_{-i}} - t_i^{j-1,k_{-i}} \geq 0.$$

[12] For example, see Hagerty and Rogerson (1987) for bilateral trading cases.

and thus,

$$U_i(\theta^k) \equiv \theta_i^{k_i} z_i^k + t_i^k \geq U_i(\theta_i^{k_i'}, \theta_{-i}^{k_{-i}}) + \sum_{j=k_i'+1}^{k_i} (\theta_i^j - \theta_i^{j-1}) z_i^{j-1,k_{-i}}.$$

Because $(\theta_i^j - \theta_i^{j-1}) z_i^{j-1,k_{-i}} = \int_{\theta_i^{j-1}}^{\theta_i^j} z_i(t, \theta_{-i}^{k_{-i}}) dt$, we obtain

$$U_i(\theta^k) \geq U_i(\theta_i^{k_i'}, \theta_{-i}^{k_{-i}}) + \int_{\theta_i^{k_i'}}^{\theta_i^{k_i}} z_i(t, \theta_{-i}^{k_{-i}}) dt.$$

Now, let $\theta \in \Theta^k$. Because $(z_i(\theta), t_i(\theta)) = (z_i^k, t_i^k)$, we have

$$U_i(\theta) = U_i(\theta^k) + (\theta_i - \theta_i^{k_i}) z_i^k.$$

Therefore, for any θ_i, θ_i' and θ_{-i},

$$U_i(\theta) \geq U_i(\theta_i', \theta_{-i}) + \int_{\theta_i'}^{\theta_i} z_i(t, \theta_{-i}) dt.$$

□

Sometimes, one may want to assume that any mechanism has an "opt-out" or "non-participation" message for each i that assigns $(z_i, t_i) = (0, 0)$ to agent i regardless of the opponents' actions. In that case, we assume that there exists an "opt-out type" who strictly prefers $(0, 0)$ than any other allocations, so that the opt-out message is weakly dominant for this type in any mechanism.

In this case, the LDIC condition obtained by letting this opt-out type to be the initial type of the chain (i.e., θ_i^0 in C_i) implies a lower bound on each agent's information rent: The integral form of the LDIC conditions

$$U_i(\theta) \geq U_i(\theta_i', \theta_{-i}) + \int_{\theta_i'}^{\theta_i} z_i(t, \theta_{-i}) dt,$$

and the LDIC condition for $\theta_i = 0$

$$U_i(0, \theta_{-i}) \geq 0,$$

imply, by letting $\theta_i' = 0$,

$$U_i(\theta) \geq \int_0^{\theta_i} z_i(t, \theta_{-i})dt.$$

We call this inequality the *information rent lower bound* (IRLB). Again, if f is dominant-strategy incentive compatible, then this holds with equality, i.e., the agents' information rents are exactly pinned down by $z(\cdot)$, but an LDIC f bounds the information rents only from below.

5.3 Worst-case maximization problems

In this section, we assume that the mechanism designer has his own utility function $w(x, \theta)$, prior Φ over Θ, and wants to maximize the worst-case expected utility when the agents may play any admissible actions in each state. Specifically, for admissibly implementable F, we define

$$W(F) = \int_\theta \left[\inf_{x \in F(\theta)} w(x, \theta) \right] d\Phi.$$

This $W(F)$ is the "guaranteed" level of the designer's expected utility if F is implemented, given whatever admissible actions the agents play in each state.[13]

If every Θ_i is finite, then for any admissibly implementable F, there is an LDIC selection f of F such that $W(F) \leq \int_\theta w(f(\theta), \theta) d\Phi$. Thus, an upper bound on the highest achievable guarantee of the designer's expected utility

[13] In the following, we assume that $W(F)$ is well defined for any admissibly implementable F. If the worst-case selection of some F is not measurable, the guarantee may be defined as follows, and we obtain the same result: Letting Ω be the set of all measurable functions on Θ,

$$\begin{aligned} \underline{W}(F) \quad = \quad & \sup_{\omega \in \Omega} \int_\theta \omega(\theta) \, d\Phi \\ & \text{sub.to } \omega(\theta) \leq \inf_{x \in F(\theta)} w(x, \theta), \ \forall \theta. \end{aligned}$$

is given by

$$\begin{aligned} \sup\nolimits_f \quad & \int_\theta w(f(\theta),\theta)d\Phi & (1) \\ \text{sub.to} \quad & \text{(LDIC)} & (2) \end{aligned}$$

Even if some Θ_i is infinite, if a finite mechanism admissibly implements F, then F has a selection f with the integral LDIC condition and $W(F) \leq \int_\theta w(f(\theta),\theta)d\Phi$. One may wonder whether we can implement some SCC that does not have an LDIC selection using infinite mechanisms. However, the following result provides a sufficient condition on the environment under which the integral LDIC condition yields a valid upper bound among *all bounded mechanisms* (not only among finite mechanisms).

In the following, let $\Theta_i = [0,1]$ for each i, and we define

$$\begin{aligned} W^* \;=\; & \sup_f \int_\theta w(f(\theta),\theta)d\Phi & (3) \\ & \text{sub.to (integral LDIC).} & (4) \end{aligned}$$

Theorem 4. Suppose that Φ is absolutely continuous with density function ϕ, and there exists a Riemann integrable function $b : \Theta \to \mathbb{R}$ such that, for each θ, θ' and x,

$$|w(x,\theta)\phi(\theta) - w(x,\theta')\phi(\theta')| \leq |b(\theta) - b(\theta')|. \tag{5}$$

Then, for any admissibly implementable F, we have $W(F) \leq W^*$.

Remark 1. In auction settings, the inequality is sometimes violated, because typically, the designer's objective is revenue $\sum_i t_i$ and each t_i can take any real number. However, as we see in Section 4.3, the participation constraints may imply bounds on the transfers, and then, the boundedness of w is satisfied.

Proof. In the proof, we assume $\phi(\theta) \equiv 1$ without loss of generality (otherwise, we redefine w). Fix any mechanism Γ that implements F and $K \in \mathbb{N}$. For each i and $\alpha_i \in [0, \frac{1}{K}]$, consider the chain dominance condition with

$C_i^{\alpha_i} = (\theta_i^{k_i,\alpha_i})_{k_i=0}^K$ where $\theta_i^{k_i,\alpha_i} = \alpha_i + \frac{k_i}{K}$. Theorem 1 implies that there exists a selection $\tilde{f} = (\tilde{z}_i, \tilde{t}_i)_{i=1}^N$ of F such that, for each i, $k_i = 1, \dots, K$, and θ_{-i},

$$\theta_i^{k_i,\alpha_i} \tilde{z}_i(\theta_i^{k_i,\alpha_i}, \theta_{-i}) + \tilde{t}_i(\theta_i^{k_i,\alpha_i}, \theta_{-i}) \geq \theta_i^{k_i,\alpha_i} \tilde{z}_i(\theta_i^{k_i-1,\alpha_i}, \theta_{-i}) + \tilde{t}_i(\theta_i^{k_i-1,\alpha_i}, \theta_{-i}).$$

Denote $\alpha = (\alpha_i)_{i=1}^N$, $k = (k_i)_{i=1}^N$, and $\theta^{k,\alpha} = (\theta_i^{k_i,\alpha_i})_{i=1}^N$. We define the following "Problem (K, α)":

$$\begin{array}{ll} \max_{f=(z_i,t_i)_{i=1}^N} & \dfrac{1}{K^N} \displaystyle\sum_k w(f(\theta^{k,\alpha}), \theta^{k,\alpha}) \\ \text{sub.to} & \theta_i^{k_i,\alpha_i} z_i(\theta_i^{k_i,\alpha_i}, \theta_{-i}) + t_i(\theta_i^{k_i,\alpha_i}, \theta_{-i}) \\ & \geq \theta_i^{k_i,\alpha_i} z_i(\theta_i^{k_i-1,\alpha_i}, \theta_{-i}) + t_i(\theta_i^{k_i-1,\alpha_i}, \theta_{-i}), \ \forall i, k_i, \theta_{-i}. \end{array}$$

Let $W(K, \alpha)$ be the value of this problem. Then,

$$\begin{aligned} W(K, \alpha) &\geq \frac{1}{K^N} \sum_k w(\tilde{f}(\theta^{k,\alpha}), \theta^{k,\alpha}) \\ &\geq \frac{1}{K^N} \sum_k [\inf_{x \in F(\theta^{k,\alpha})} w(x, \theta^{k,\alpha})], \end{aligned}$$

and thus, we obtain $\sup_\alpha W(K, \alpha) \geq W^*$ and $\sup_\alpha W(K, \alpha) \geq W(F)$.

Now we show that, for any $\varepsilon > 0$, there exists $K(\varepsilon)$ such that for any $K \geq K(\varepsilon)$ and $\alpha \in [0, \frac{1}{K}]^N$, $W^* + \varepsilon \geq W(K, \alpha)$. This implies $W^* \geq W(F)$ for any admissibly implementable F, which completes the proof.

In the following, we fix arbitrary $\alpha \in [0, \frac{1}{K}]^N$, and let f^* be the solution to Problem (K, α). Let $\Theta^k = \prod_l [\theta_l^{k_l,\alpha_l}, \theta_l^{k_l+1,\alpha_l})$ and define $\hat{f} : \Theta \to X$ and $\hat{w} : \Theta \to \mathbb{R}$ so that

$$\begin{aligned} \hat{f}(\theta) &= f^*(\theta^{k,\alpha}) \text{ if } \theta \in \Theta^k, \\ \hat{w}(\theta) &= w(f^*(\theta^{k,\alpha}), \theta^{k,\alpha}) \text{ if } \theta \in \Theta^k. \end{aligned}$$

Both are finite step functions (and so they are measurable), and by definition, $\int_\theta \hat{w}(\theta) d\theta = W(K, \alpha)$. Also, because $\hat{f}$ is a finite-step function that satisfies the integral LDIC condition, $\int_\theta w(\hat{f}(\theta), \theta) d\theta \leq W^*$.

Observe that

$$
\begin{aligned}
& \Big| \int_\theta \hat{w}(\theta) d\theta - \int_\theta w(\hat{f}(\theta), \theta) d\theta \Big| \\
\leq \; & \int_\theta |\hat{w}(\theta) - w(\hat{f}(\theta), \theta)| d\theta \\
\leq \; & \frac{1}{K^N} \sum_k | \sup_{\theta \in \Theta^k} w(\hat{f}(\theta), \theta) - \inf_{\theta \in \Theta^k} w(\hat{f}(\theta), \theta)| \\
\leq \; & \frac{1}{K^N} \sum_k | \sup_{\theta \in \Theta^k} b(\theta) - \inf_{\theta \in \Theta^k} b(\theta)|,
\end{aligned}
$$

which is $o(\frac{1}{K})$ because of the Riemann integrability of b. Therefore, $W^* + o(\frac{1}{K}) \geq W(K, \alpha)$. □

5.4 Balanced-budget bilateral trading

5.4.1 Environment

We consider the bilateral trading problem studied in Section 2. Recall that an allocation is a pair (z, p), where z is the probability of trading, and p is the price. The seller's and the buyer's utility in state (c, v) are given by $u_1 = (p - c)z$ and $u_2 = (v - p)z$, respectively, and the designer's utility is the total surplus, $(v - c)z$.[14] We assume that any mechanism has an "opt-out" message for each i, so that whenever agent i chooses the message, $(z, p) = (0, 0)$ is assigned.

The results in the previous section imply the following:

Corollary 1. Suppose $\Theta_1 = \{c_1, \ldots, c_J\}$ and $\Theta_2 = \{v_1, \ldots, v_K\}$. Then, the highest achievable guarantee of the expected welfare is upper bounded by

$$
\begin{aligned}
W^* = \sup_{(z(\cdot),p(\cdot))} \quad & \int_{c,v} (v-c)z(c,v)\ d\Phi \\
\text{sub.to} \quad & (p(c_j, v_k) - c_j)z(c_j, v_k) \geq (p(c_{j+1}, v_k) - c_j)z(c_{j+1}, v_k),\ \forall j, k, \\
& (v_k - p(c_j, v_k))z(c_j, v_k) \geq (v_k - p(c_j, v_{k-1}))z(c_j, v_{k-1}),\ \forall j, k, \\
& (p(c_J, v_k) - c_J)z(c_J, v_k) \geq 0,\ \forall k, \\
& (v_1 - p(c_j, v_1))z(c_j, v_1) \geq 0,\ \forall j.
\end{aligned}
$$

Corollary 2. Suppose that $\Theta_i = [0, 1]$ for each i. Then, the highest achievable guarantee of the expected welfare among all finite mechanisms is upper

[14] In the notation in the previous section, $z_1 = z_2 \equiv z$, $t_1 = -t_2 \equiv pz$, $c = -\theta_1$ and $v = \theta_2$.

bounded by

$$
\begin{aligned}
W^* = \sup_{(z(\cdot),p(\cdot))} \quad & \int_{c,v} (v-c)z(c,v)\ d\Phi \\
\text{sub.to} \quad & (p(c,v)-c)z(c,v) \geq (p(c',v)-c)z(c',v) \int_c^{c'} z(t,v)\ dt,\ \forall c < c', v, \\
& (v-p(c,v))z(c,v) \geq (v-p(c,v'))z(c,v') \int_{v'}^{v} z(c,t)\ dt,\ \forall c, v > v', \\
& (p(1,v)-1)z(1,v) \geq 0,\ \forall v, \\
& (0-p(c,0))z(c,0) \geq 0,\ \forall c.
\end{aligned}
$$

Moreover, if Φ is absolutely continuous with density ϕ, and $(v-c)\phi(c,v)$ is Riemann integral, then, W^* is the upper bound among all bounded mechanisms.

5.4.2 Optimality of posted-price mechanisms

In this section, we use the upper bound to show that, for a class of distributions, no mechanism can improve over the optimal posted-price mechanism. Let $\Theta_i = [0,1]$ and let ϕ be the density of Φ.

Theorem 5. Suppose that $\psi(c,v) \equiv (v-c)\phi(c,v)$ is strictly decreasing in c, strictly increasing in v, and continuous in (c,v), for any $c < v$. Then no mechanism guarantees expected welfare strictly higher than the welfare guarantee of the posted-price mechanism with price p^*, where p^* solves $\int_0^{p^*} \psi(t,p^*)\ dt = \int_{p^*}^1 \psi(p^*,t)\ dt$.[15]

An allocation rule $(z(c,v), p(c,v))_{c,v}$ induces expected welfare $\int_{c,v} \psi(c,v)z(c,v)\ dvdc$, which is a weighted integral of $z(c,v)$, where the weight is $\psi(c,v)$. The monotonicity of the weight function ψ means that more-efficient types have higher weights. This condition is satisfied by independent uniform distributions (i.e., $\phi(c,v) \equiv 1$), and any distribution such that "more efficient types are more likely" (i.e., $\phi(c,v)$ is non-increasing in c, non-decreasing in v).

[15]I thank Gabriel D. Carroll, who pointed out an error in the proof in the previous version.

As is shown in the previous section, no (possibly infinite, but bounded) mechanism can achieve higher expected welfare than W^* under Riemann integrability of ψ, which is satisfied because ψ is assumed to be continuous.

Proof. First, the integral LDIC condition and participation condition imply the agents' information rent lower bounds, as we discussed in the previous section: For each c, v,

$$\begin{aligned}(p(c,v)-c)z(c,v) &\geq \int_c^1 z(t,v)\ dt,\\ (v-p(c,v))z(c,v) &\geq \int_0^v z(c,t)\ dt.\end{aligned}$$

Adding up the agents' information rent lower bounds, and because $z(c,v) \leq 1$, we obtain the following corollary.

Lemma 4. Let $(z(c,v), p(c,v))_{c,v\in[0,1]}$ be an allocation rule with the IRLB condition. Then for any c, v,

$$v - c \geq \int_c^1 z(t,v)\ dt + \int_0^v z(c,t)\ dt \qquad (SC(c,v)).$$

This inequality means that the trading rule of an LDIC revelation mechanism is constrained by the surplus of a trade in state (c,v), i.e., $v-c$. We call this inequality the *surplus constraint* in (c,v) (or $SC(c,v)$).

Obviously, $c > v$ implies $z(c,v) = 0$.

Consider the following relaxed problem for W^*:

$$\begin{aligned}\sup_{z(\cdot)} \quad & \int_{c,v} \psi(c,v)z(c,v)\ dvdc\\ \text{sub.to} \quad & SC(c,v),\ \forall c,v.\end{aligned}$$

We guess which surplus constraints are binding. To give some intuition, we consider a special case with $\phi(c,v) = 1$ for $(c,v) \in [0,1]^2$ (i.e., a bivariate uniform distribution), and hence the theorem yields $p^* = \frac{1}{2}$. See the appendix for the general case. Our guess is that only the surplus constraints $SC(1-q, q)$ for $q \in [\frac{1}{2}, 1]$ are binding:

$$2q - 1 \geq \int_{1-q}^{1} z(t, q)\ dt + \int_{0}^{q} z(1-q, t)\ dt \qquad (SC(1-q, q)),$$

and the other surplus constraints are ignored.

Notice that the objective can be decomposed as follows.

$$\begin{aligned}
& \int_{c,v} \psi(c, v) z(c, v)\ dvdc \\
= & \int_{q=0}^{1} [\int_{1-q}^{1} \psi(t, q) z(t, q)\ dt + \int_{0}^{q} \psi(1-q, t) z(1-q, t)\ dt] dq \\
= & \int_{q=\frac{1}{2}}^{1} [\int_{1-q}^{1} \psi(t, q) z(t, q)\ dt + \int_{0}^{q} \psi(1-q, t) z(1-q, t)\ dt] dq,
\end{aligned}$$

where the last equality obtains because $z(c, v) = 0$ for $c > v$.

For each $q \in [\frac{1}{2}, 1]$, we first solve the following decomposed problem separately, and show that the solutions to them also consist of the solution to the original problem (i.e., W^*):

$$\begin{aligned}
\max_{z(\cdot, q), z(1-q, \cdot) \in [0,1]} \quad & \int_{1-q}^{1} \psi(t, q) z(t, q)\ dt + \int_{0}^{q} \psi(1-q, t) z(1-q, t)\ dt \\
\text{sub.to} \quad & 2q - 1 \geq \int_{1-q}^{1} z(t, q)\ dt + \int_{0}^{q} z(1-q, t)\ dt \qquad (SC(1-q, q)).
\end{aligned}$$

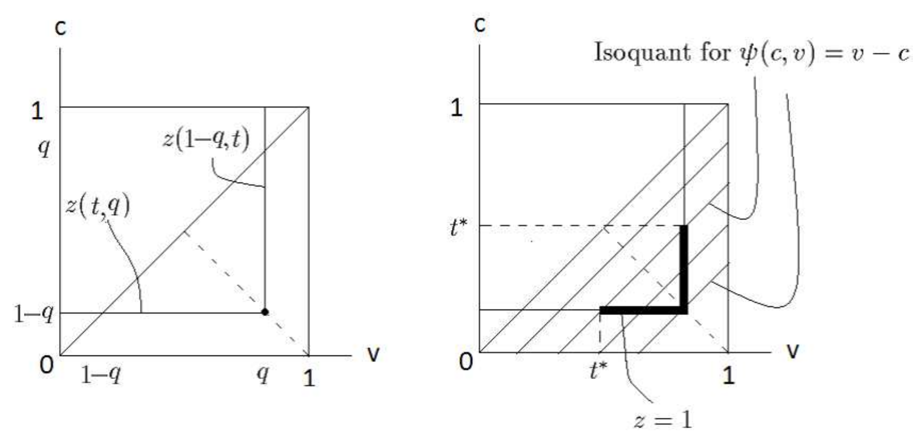


Figure 6: A decomposed problem for q

Because the objective is linear in z, in the solution, there is a value ψ^* such that $z(c, v) = 1$ if and only if $\psi(c, v) \geq \psi^*$, and zero otherwise. Because a uniform distribution implies $\psi(c, v) = v - c$, (i) there is c^* such that $z(t, q) = 1$ if and only if $t \leq c^*$, (ii) there is v^* such that $z(1 - q, t) = 1$ if and only if $t \geq v^*$, (iii) $\psi^* = q - c^* = v^* - (1 - q)$, and (iv) $2q - 1 = [c^* - (1 - q)] + [q - v^*]$ by the surplus constraint at $(1 - q, q)$. These imply $c^* = v^* = \frac{1}{2}$.

Therefore, in the solution to the decomposed problem for any $q \in [\frac{1}{2}, 1]$, the agents trade if and only if $c < \frac{1}{2} < v$. A posted-price mechanism with $p^* = \frac{1}{2}$ induces this allocation rule, and therefore, no mechanism improves over this posted-price mechanism. $\square$

Remark 2. Because of the symmetry of a uniform distribution, all the binding constraints are on the diagonal (i.e., $(1-q, q)$ for $q \in [\frac{1}{2}, 1]$)). For a general distribution, our proof constructs a downward-sloping curve (not necessarily on the diagonal) that connects (p^*, p^*) to $(0, 1)$ such that maximizing the decomposed welfare functions subject to the surplus constraints for the points on this curve yields the posted-price mechanism with price p^*.

Theorem 3 may be interpreted as giving a foundation for the use of dominant-strategy mechanisms as the optimal robust mechanisms for some distributions.

5.4.3 Optimality of two-price mechanisms

In this section, we provide a sufficient condition on the environment where the two-price mechanism we have examined in Section 2 is optimal. Thus, in contrast to the previous section, the optimal mechanism guarantees strictly higher expected welfare than any posted-price mechanism.

Recall that a two-price mechanism where the seller chooses a price is characterized by $(z_1, p_1), (z_2, p_2) \in X$ with $z_1 > z_2$ and $p_1 < p_2$ as follows: The seller chooses $p \in \{p_1, p_2\}$, the buyer chooses $\overline{p} \in \{p_1, p_2\}$, and (z_k, p_k) is assigned if $p = p_k \leq \overline{p}$ (otherwise, no trade). That is, the trading price is chosen by the seller, and the buyer essentially accepts or rejects each price. In the following, this mechanism is called a "two-price-for-seller" mechanism (with $(z_1, p_1), (z_2, p_2)$).

As discussed in Section 2, the buyer has a dominant strategy in this mechanism, while the seller does not. In particular, for the seller with $c \in (\frac{p_1 z_1 - p_2 z_2}{z_1 - z_2}, p_1)$, either price is admissible: If he believes that the buyer chooses $\overline{p} = p_2$, then $p = p_2$ yields is better, while if he believes that the buyer chooses $\overline{p} = p_1$, then $p = p_1$ is better.

Similarly, a "two-price-for-buyer" mechanism with $(z_1, p_1), (z_2, p_2)$ (where $z_1 > z_2$ and $p_1 > p_2$) is such that the buyer chooses $p \in \{p_1, p_2\}$, the seller chooses $\underline{p} \in \{p_1, p_2\}$, and (z_k, p_k) is assigned if $p = p_k \geq \underline{p}$.

Theorem 6. Suppose that Φ is the following discrete distribution: There exist $C = \{c_1, c_2\} \subseteq \Theta_1$ and $V = \{v_1, v_2\} \subseteq \Theta_2$ with $c_1 < v_1 < c_2 < v_2$ such that $\Pr(c \notin C) = \Pr(v \notin V) = 0$. Let $\Phi_{jk} = \Pr((c, v) = (c_j, v_k))$. Then,

- if $\frac{\Phi_{11}}{\Phi_{22}} \geq \frac{(v_2 - v_1)(v_2 - c_2)}{(c_2 - c_1)(v_1 - c_1)}$, then a two-price-for-seller mechanism with $(1, v_1), (\frac{v_1 - c_1}{c_2 - c_1}, c_2)$ is optimal.

- if $\frac{\Phi_{11}}{\Phi_{22}} < \frac{(v_2 - v_1)(v_2 - c_2)}{(c_2 - c_1)(v_1 - c_1)}$, then a two-price-for-buyer mechanism with $(1, c_2), (\frac{v_2 - c_2}{v_2 - v_1}, v_1)$ is optimal.

Proof. Treat $C \times V$ as the true type space. By Corollary 3, the highest achievable guarantee of the expected welfare is upper bounded by

$$\begin{aligned} W^* = \sup_{(z_{jk}, p_{jk})_{j,k}} \quad & \sum_{j,k} (v_k - c_j) z_{jk} \Phi_{jk} \\ \text{sub.to} \quad & (p_{1k} - c_1) z_{1k} \geq (p_{2k} - c_1) z_{2k}, \ \forall k, \\ & (v_2 - p_{j2}) z_{j2} \geq (v_2 - p_{j1}) z_{j1}, \ \forall j, k, \\ & (p_{2k} - c_2) z_{2k} \geq 0, \ \forall k, \\ & (v_1 - p_{j1}) z_{j1} \geq 0, \ \forall j. \end{aligned}$$

As in the case with continuous type spaces, these LDIC conditions induce lower bounds for the agents' information rents, which then induce the surplus constraints:

$$\begin{aligned} W^* \leq \sup_{(z_{jk})_{j,k}} \quad & \sum_{j,k} (v_k - c_j) z_{jk} \Phi_{jk} \\ \text{sub.to} \quad & (v_1 - c_2) z_{21} \geq 0, \\ & (v_2 - c_2) z_{22} \geq (v_2 - v_1) z_{21}, \\ & (v_1 - c_1) z_{11} \geq (c_2 - c_1) z_{21}, \\ & (v_2 - c_1) z_{12} \geq (c_2 - c_1) z_{22} + (v_2 - v_1) z_{11}. \end{aligned}$$

For this relaxed problem, first, because $v_1 < c_2$, z_{21} must be zero. Because $v_2 - c_1 > 0$, z_{12} is not bounded from above except that $z_{12} \leq 1$, we have $z_{12} = 1$. For z_{11} and z_{22}, because the problem is linear and both $v_1 - c_1$ and $v_2 - c_2$ are positive, one of them equals one, while the other is determined so as to satisfy $v_2 - c_1 = (c_2 - c_1) z_{22} + (v_2 - v_1) z_{11}$.

If we have $z_{11} = 1$, then $z_{22} = \frac{v_1 - c_1}{c_2 - c_1}$, and the objective is

$$(v_2 - c_1)\Phi_{12} + (v_1 - c_1)\Phi_{11} + (v_2 - c_2)\frac{v_1 - c_1}{c_2 - c_1}\Phi_{22}. \tag{6}$$

If we have $z_{22} = 1$, then $z_{11} = \frac{v_2 - c_2}{v_2 - v_1}$, and the objective is

$$(v_2 - c_1)\Phi_{12} + (v_1 - c_1)\frac{v_2 - c_2}{v_2 - v_1}\Phi_{11} + (v_2 - c_2)\Phi_{22}. \tag{7}$$

Therefore, if $\frac{\phi_{11}}{\phi_{22}} \geq \frac{(v_2 - v_1)(v_2 - c_2)}{(c_2 - c_1)(v_1 - c_1)}$, then the first way is better, and vice versa.

The two-price mechanisms in the statements attain these upper bound levels of expected welfare, and therefore, they are the optimal mechanisms.

$\square$

Remark 3. We can extend the results to a more general case. For example, under certain conditions on Φ, a K-price mechanism becomes optimal. In the binary case, it turns out that it is a "probability zero event" (in terms of Φ) that an agent does not have a dominant strategy, but with more than two types, it is possible that with a positive probability, some types of an agent have multiple admissible actions.

Also, Φ can be an atomless distribution, but in that case, for the present, we could only show that a two-price mechanism can be approximately optimal (for example, when Φ converges to the two-by-two distribution in the theorem). It is an open question if there is a class of continuous distributions where a two-price (or K-price) mechanism is exactly optimal.

5.5 Without balanced budget

In this section, we consider an environment without balanced budget (e.g., auction). The mechanism designer's utility function is $w(z, t, \theta)$, which is decreasing in each t_i. As in the balanced-budget case, we assume that any mechanism has an "opt-out" message for each agent, and whenever agent i chooses that message, $(z_i, t_i) = 0$ is assigned for him. In the following, we consider the case with $\Theta_i = [0, 1]$ for each i, but the similar results hold for the case with finite type spaces as well.

As in Theorem ??, under certain conditions on the environment, the highest achievable guarantee of the designer's expected utility is upper bounded by the following IRLB bound:

$$\begin{aligned} \max_{f(\cdot)} \quad & \int_{\theta} w(f(\theta), \theta)\ d\Phi \\ \text{sub.to} \quad & U_i(\theta) \geq \int_0^{\theta_i} z_i(\tilde{\theta}_i, \theta_{-i})\ d\tilde{\theta}_i,\ \forall i, \theta. \end{aligned}$$

Without balanced budget, all IRLBs are satisfied with equality:[16]

We call each constraint with equality $ICFOC_i(\theta)$.

Suppose that the solution to the relaxed problem is a monotonic allocation rule $(z^*(\theta), t^*(\theta))_\theta$. Then, the allocation rule is dominant-strategy incentive compatible and ex post individually rational.[17]

As an example, suppose that $w(z, t, \theta) = \sum_i \theta_i z_i - \lambda \sum_i t_i$ for some constant $\lambda > 0$. Then, the designer's objective is a weighted sum of the agents' total surplus and monetary residual.[18] Then, the relaxed problem for W^* is

[16]Otherwise, we can decrease a transfer by a small amount without violating any other constraints nor decreasing the objective. With the exact balanced budget, this logic does not apply, because any decrease in the transfer to one of the agents implies an increase in the transfer to the other agent.

[17]For example, see **?**.

[18]A simple story would be that the mechanism designer can be a residual claimant for the net transfers. In this case, the expected (not exact) budget balance may be the only necessary requirement, and $\lambda > 0$ corresponds to the "shadow price" for the expected budget balance constraint. An alternative situation is that the mechanism designer is a

given as follows:

$$\max_{z(\cdot),t(\cdot)} \quad \int_\theta \sum_i \theta_i z_i - \lambda \sum_i t_i \ d\Phi$$
$$\text{sub.to} \quad U_i(\theta) = \int_0^{\theta_i} z_i(\tilde{\theta}_i, \theta_{-i}) \ d\tilde{\theta}_i \ \forall i, \theta.$$

Now, replacing $t_i(\theta)$ in the objective by the $ICFOC_i(\theta)$, and applying integration by parts, the objective function becomes the following.

$$\begin{aligned} & \int_\theta \sum_i \theta_i z_i - \lambda \sum_i [\int_{\underline{\theta}_i}^{\theta_i} z_i(\tilde{\theta}_i, \theta_{-i}) \ d\tilde{\theta}_i - \theta_i z_i(\theta)] \ d\Phi \\ = \ & \int_\theta \sum_i \Big[((1+\lambda)\theta_i - \lambda \frac{1-\Phi_i(\theta_i|\theta_{-i})}{\phi_i(\theta_i|\theta_{-i})})\Big] z_i(\theta) \ d\Phi, \end{aligned}$$

where $\Phi_i(\theta_i|\theta_{-i})$ and $\phi_i(\theta_i|\theta_{-i})$ denote the conditional cdf and pdf of θ_i given θ_{-i}. Suppose that the monotone hazard rate conditions are satisfied for the conditional distributions: $\frac{1-\Phi_i(\theta_i|\theta_{-i})}{\phi_i(\theta_i|\theta_{-i})}$ is non-increasing in θ_i for any θ_{-i}, then we obtain a monotone trading rule as the solution to this expected welfare maximization problem, which is dominant-strategy incentive compatible and ex post individually rational. Thus, there is no improvement over the optimal dominant-strategy mechanism.

This result provides a foundation to restrict attention to dominant-strategy mechanisms in this setting.[19] The result is also related to a result obtained by Chung and Ely (2007). They show that the mechanism that maximizes the worst-case expected revenue (corresponding to $\lambda \to \infty$) is dominant-strategy incentive compatible if Φ satisfies affiliation and monotone hazard rate condition. Note that the "worst-case" in their definition is based on the robust partial implementation of Bergemann and Morris (2005).

government who is concerned not only about the agents' welfare, but also the "tax payers". Then, she may desire to maximize the weighed sum of the expected welfare of the agents and the tax payers, as in Laffont and Tirole (1993). In this case, λ represents the shadow price of the transfer from the tax payers to the agents. This becomes equivalent to a revenue maximization problem, if $\frac{1}{\lambda} \to 0$.

[19]For example, see **?**.

References

Abreu, D., and H. Matsushima (1992): "Virtual Implementation in Iteratively Undominated Strategies: Complete Information," *Econometrica*, 60(5), 993–1008.

Barelli, P. (2009): "On the genericity of full surplus extraction in mechanism design," *Journal of Economic Theory*, 144(3), 1320–1332.

Bergemann, D., and S. Morris (2005): "Robust mechanism design," *Econometrica*, 73(6), 1771–1813.

——— (2010): "Robust Implementation in General Mechanisms," forthcoming, *Games and Economic Behavior*.

Bergemann, D., S. Morris, and O. Tercieux (2010): "Rationalizable Implementation," forthcoming, *Journal of Economic Theory*.

Chatterjee, K., and W. Samuelson (1983): "Bargaining under incomplete information," *Operations Research*, 31, 835–851.

Chung, K.-S., and J. Ely (2007): "Foundations of Dominant-Strategy Mechanisms," *Review of Economic Studies*, 74(2), 447–476.

Copic, J., and C. Ponsati (2008): "Robust Bilateral Trade and Mediated Bargaining," *Journal of the European Economic Association*, 6(2-3), 570–580.

Crémer, J., and R. McLean (1985): "Optimal selling strategies under uncertainty for an discriminating monopolistwhen demands are interdependent," *Econometrica*, 53, 345–361.

——— (1988): "Full extraction of the surplus in Bayesian and dominant strategy auctions," *Econometrica*, 56, 1247–1258.

D'ASPREMONT, C., J. CREMER, AND L.-A. GERARD-VARET (2004): "Balanced Bayesian mechanisms," *Journal of Economic Theory*, 115(2), 385–396.

HAGERTY, K. M., AND W. P. ROGERSON (1987): "Robust trading mechanisms," *Journal of Economic Theory*, 42(1), 94–107.

HEIFETZ, A., AND Z. NEEMAN (2006): "On the Generic (Im)Possibility of Full Surplus Extraction in Mechanism Design," *Econometrica*, 74(1), 213–233.

JACKSON, M. O. (1992): "Implementation in undominated strategies: a look at bounded mechanisms," *Review of Economic Studies*, 59, 757–775.

KOSENOK, G., AND S. SEVERINOV (2008): "Individually rational, budget-balanced mechanisms and allocation of surplus," *Journal of Economic Theory*, 140(1), 126–161.

KUNIMOTO, T., AND R. SERRANO (2010): "A New Necessary Condition for Implementation in Iteratively Undominated Strategies," Discussion paper.

LAFFONT, J.-J., AND E. MASKIN (1980): "A Differential Approach to Dominant Strategy Mechanisms," *Econometrica*, 48(6), 1507–20.

LAFFONT, J. J., AND J. TIROLE (1993): *A Theory of Incentives in Regulation and Procurement*. MIT Press, Cambridge.

MOULIN, H. (1979): "Dominance Solvable Voting Schemes," *Econometrica*, 47(6), 1137–51.

MYERSON, R., AND M. SATTERTHWAITE (1983): "Efficient mechanism for bilateral trading," *Journal of Economic Theory*, 28, 265–281.

NEEMAN, Z. (2004): "The relevance of private information in mechanism design," *Journal of Economic Theory*, 117(1), 55–77.

Smith, D. (2010): "A Prior Free Efficiency Comparison of Mechanisms for the Public Good Problem," mimeo, University of Michigan.

Srivastava, S., and M. A. Trick (1996): "Sophisticated voting rules: the case of two tournaments," *Social Choice and Welfare*, 13(3), 275–289.

Wilson, R. (1987): "Game-Theoretic Analyses of Trading Processes," in *Advances in Economic Theory: Fifth World Congress*, ed. by T. Bewley, chap. 2, pp. 33–70. Cambridge University Press, Cambridge UK.