THREE ESSAYS ON BARGAINING AND CHEAP TALK

Anna Ilyina

A DISSERTATION

In

Economics

Presented to the Faculties of the University of Pennsylvania in Partial Fulfillment of the Requirement for the Degree of Doctor of Philosophy

2001

Supervisor of Dissertation

Graduate Group Chairperson

ABSTRACT

THREE ESSAYS ON BARGAINING AND CHEAP TALK

Anna Ilyina

George Mailath

It is generally impossible to design an ex post efficient mechanism for bilateral trading when the traders' valuations for an object of trade are private, there are no outside subsidies and the traders can refuse to carry out a transaction if they expect to lose money. The first two essays focus on the problem of bilateral bargaining with multiple objects and private valuations, where efficiency loss can occur not only due to the failure to reach an agreement to trade when positive gains from trade exist, but also due to the failure to fully realize potential gains from trade because of selecting an ex post inefficient object (not the one with the largest spread between the traders' valuations). The problem is studied for general, as well as for the special class of trading mechanisms (k-double auctions), and also for simple and rich trading environments and under different assumptions about the information structure of the game. It is proposed to distinguish between fully ex post efficient mechanisms and the mechanisms, which are ex post efficient with respect to the object selection (given that trade occurs, the traders always select the best object, but trade does not always occur when gains from trade exist). In simple trading environments, the second best mechanism is always ex post efficient with respect to the object selection. For rich trading environments, the necessary conditions for the second-best mechanism to be ex post efficient with respect to the object selection are formulated. When the information structure is such that the optimal choice of the object depends on the private information of only one of the traders, it is always possible to design a mechanism that is ex post efficient with respect to the object selection.

The third essay develops an epistemic framework for the Sender-Receiver Cheap Talk Games that is used to formulate sufficient epistemic conditions for credible neologism (Farrell, 1985) and credible announcement (Matthews et al., 1991). It is shown that the players' knowledge of their opponents' rationality must be bounded in order for them to accept a message as credible neologism or as credible announcement.

ACKNOWLEDGEMENTS

I am grateful to my advisors George Mailath and Stephen Morris for their valuable advice, support and patience.

CONTENTS

LIST OF ILLUSTRATIONS

Chapter 1

Efficient Mechanisms for Bilateral Trading with Multiple Objects

1.1 Introduction

Most market transactions involve complex negotiations between a buyer and a seller over various attributes of the object of trade, besides its price, such as quality, technical specifications and delivery date. This paper studies a buyer and a seller negotiating the terms of trade. The buyer demands at most one good, which is selected from a set of trading alternatives containing several indivisible objects. As an example, one can think of any objects that belong to the same product category and have a variety of special features (like, different models of cars or computers), in which case the seller's reservation prices (or the buyer's willingness to pay) for different goods may not be the same.

In the first-best, *ex post* efficient, outcome of the bargaining process, the buyer and seller always trade an object, which generates the highest total gains from trade and they do not trade only if there are no gains from trade. Thus, the first question to be addressed in this paper is whether an *ex post* efficient allocation is generally feasible in sufficiently realistic trading environments, in which the traders have several alternatives to choose from.

The analysis of the bilateral trading problem presented in this chapter captures the key features of a typical market environment, where the traders can refuse to carry out a transaction if they expect to lose money and there are no outside subsidies. In such environment, the *ex post* efficient allocation is always feasible if the trading outcome is contingent on perfect information about the traders' preferences. However, when reservation values are private, the problem of when to trade and which object should be traded becomes more complicated and the *ex post* efficient outcome is not always feasible. Intuitively, because of information asymmetries the traders have incentives to lie about their reservation values in order to extract informational rents, which often creates distortions leading to the failure to realize positive gains from trade.

The trading problem, where a buyer and a seller are bargaining over the price of a single object for which their reservation values are private (independent random variables distributed with positive continuous probability densities), was studied by Myerson and Satterthwaite (1983), who showed that *ex post* efficiency is generally not feasible in a typical market environment (when the intersection of the supports of the density functions is an interval). The trading problem analyzed in this paper, where a buyer and a seller trade only one good and their reservation values for all trading alternatives are independently and identically distributed, is a natural generalization of the single-object bargaining problem studied by Myerson and Satterthwaite. The issues addressed in this paper seem to be important in a variety of contexts in the area of institutional and organizational design including such problems as the assignment of (technology) licenses and optimal procurement of goods and services.

Another extension of the single-object bilateral trading problem was analyzed by McAfee (1991), who studied the possibility of implementing an *ex post* efficient allocation in a bilateral asymmetric information environment with continuous quantities and each trader's private information represented by a scalar.[1] His main finding was that the implementation of *ex post* efficient quantity requires that 'even the worst type of one agent will trade with the best type of the other' (p.52, 1991). In other words, the implementation of *ex post* efficient quantity is impossible without outside subsidies when the supports of the

[1]McAfee assumes that the seller's possible cost functions are convex nondecreasing and parametrized by s, the seller's private information, and that the buyer's possible value functions are strictly concave and increasing and parametrized by t, the buyer's private information.

buyer's value functions and the seller's cost functions coincide (this result is essentially parallel to the result of Myerson and Satterthwaite). Other papers on bargaining with multiple units and dimensions include Spulber (1988), which analyses the interim-efficient mechanisms for a special case of McAfee's model and Linhart and Radner (1988), which studies non- Bayesian mechanisms of bilateral trading with several objects. A good overview of the literature on multidimensional screening can be found in the paper by Rochet and Choné (1998), which analyses an optimal pricing problem of a multiproduct monopolist facing a diverse population of consumers.

Besides being more realistic than a single-object bargaining model, the problem of bargaining with several potential objects of trade presents an opportunity to study the trade-offs between the two types of ex post efficiency loss. When the traders are involved in bargaining over several objects, they have to decide not only on whether to trade or not, but also on which object should be exchanged. Therefore, the efficiency loss can occur not only due to the failure to reach an agreement to trade when positive gains from trade exist, but also due to the failure to fully realize potential gains from trade because of not choosing the best possible object (i.e. the one, which maximizes the difference between the traders' valuations, provided that the latter is positive).

When would it make sense to distinguish between these two types of *ex post* efficiency loss? Here is an example of a game, where such distinction may be relevant. Consider a buyer, who demands only one good and who sequentially faces n sellers, with each seller offering several models of the same good. The buyer's reservation values for all possible trading alternatives are drawn from the same probability distribution prior to each new interaction. The game ends when trade takes place. Although the analysis of the problem described above is beyond the scope of the paper, intuitively it seems that in such a game the buyer may prefer to employ a strategy that increases the probability of selecting the best object in every interaction with a seller, conditional on the event that trade takes place.

The main distinction between the trading environments studied in this chapter and the one in McAfee's paper is that this model assumes that each players' reservation values for all potential objects of trade are independently and identically distributed (as, for example, the buyer's tastes and therefore his valuations of different models of cars may be uncorrelated), while in McAfee's paper, the traders' valuations of different quantities are correlated. Thus, in contrast to the McAfee's paper, the uncertainty in this model is multidimensional and so are the players' types.

Throughout this paper we will study the bilateral trading problem with *two* potential objects of trade, i.e. the simplest framework that allows us to see how the presence of several trading alternatives affects the players' incentives to reveal their private information.

The first part of the analysis focuses on a relatively simple trading environment, where each trader's reservation value for every object is either high or low. It turns out that in such environment, any incentive feasible mechanism which maximizes total *ex ante* expected gains from trade always guarantees an *ex post* efficient object selection (although the *ex post* efficiency is not always incentive feasible). It remains unclear, however, whether the same property obtains in more general environments, for example, where the players' valuations are drawn from the probability distributions defined on the sets containing an arbitrary finite number of elements. For such trading environments, the necessary conditions for a second-best mechanism to be *ex post* efficient with respect to the object selection are presented in the last section of the chapter.

The plan of the chapter is as follows. Section 1.2 introduces notation and definitions. Section 1.3 presents the characterization of the first- and second-best mechanisms for bilateral trading with two potential objects of trade in a simple trading environment. Section 1.4 analyses the possibility of designing a second-best mechanism that guarantees the *ex post* efficient object selection in more general (rich) environments.

1.2 Notation and Definitions

Let $Q = \{q_0, q_1, q_2\}$ denote the set of all trading alternatives, where q_0 is the no-trade option and q_1, q_2 are two distinctly different indivisible objects. The traders' reservation values for all potential objects of trade are independent random variables. Each trader's private information is represented by a vector, where $V \equiv (V_0, V_1, V_2)$ and $C \equiv (C_0, C_1, C_2)$ denote the buyer's type and the seller's type, respectively. The traders' reservation values for the no-trade option are type independent and normalized to zero and for any $i \in \{1, 2\}$, V_i, C_i are distributed on the interval $[0, 1]$.

The realizations of the traders' reservation values for both potential objects of trade q_1 and q_2 are their private information. The joint probability distribution $F(V, C)$, where $F(V, C) = \left(\prod_{i=1}^{2} F_i^B(V_i)\right)\left(\prod_{i=1}^{2} F_i^S(C_i)\right)$, is commonly known.[2]

We will assume that the traders are involved in some bargaining process, which can be modeled as an incomplete information game, where players choose their actions conditional on their private information. The outcome of this Bayesian game is a vector of probabilities of selecting q_0, q_1 or q_2 and a vector of the corresponding transfer payments. As it follows from the Revelation Principle, in order to characterize an optimal trading procedure we can focus (without loss of generality) on the class of direct trading mechanisms where an outcome is directly assigned to every possible list of types and the traders find it optimal to reveal their types truthfully.[3]

Let P_i be the price of q_i and let π_i be the probability that q_i is selected. Then, for any $i \in \{0, 1, 2\}$, P_i is defined as a mapping from the players' type space into the set of all positive real numbers, $P_i : [0,1]^2 \times [0,1]^2 \to R_+$, and π_i is defined as a mapping from the type space into the interval $[0, 1]$, $\pi_i : [0,1]^2 \times [0,1]^2 \to [0,1]$. We will assume that for any $(V, C) \in [0,1]^2 \times [0,1]^2$, $\sum_{i=0}^{2} \pi_i(V, C) = 1$, which implies that only one object can be traded.

Suppose that both traders are risk neutral. The type V buyer's expected utility defined as a function of his report $\tilde{V}$, given the mechanism $(\bar{P}, \bar{\pi})$, is as follows[4]

$$EU^B\left(\tilde{V} \mid V, (\bar{P}, \bar{\pi})\right) \equiv E_C\left\{\sum_{i=0}^{2}\left(V_i \cdot \pi_i\left(C, \tilde{V}\right) - P_i\left(C, \tilde{V}\right)\right)\right\}$$

[2] F, f denote the c.d.f and p.d.f , respectively.

[3] For a comprehensive discussion of Revelation Principle and direct trading mechanisms, see Fudenberg and Tirole (1991).

[4] E_V and E_C denote the expectation operators, where expectation is taken over the set of all possible types of the buyer and seller, respectively.

The type C seller's expected utility defined as a function of his report $\tilde{C}$, given the mechanism $(\bar{P}, \bar{\pi})$, is

$$EU^S\left(\tilde{C} \mid C, (\bar{P}, \bar{\pi})\right) \equiv E_V\left\{ \sum_{i=0}^{2} \left(P_i\left(\tilde{C}, V\right) - C_i \cdot \pi_i\left(\tilde{C}, V\right)\right)\right\}$$

The traders' reports can take any values in the interval $[0, 1]$.

We will also assume that traders can refuse to participate in the direct revelation game when their interim expected gains from trade are negative (*interim individual rationality*).

Definition 1 *A direct trading mechanism $(\bar{P}, \bar{\pi})$ is* ***Bayesian incentive compatible*** *if for any $C \in [0,1]^2, C' \in [0,1]^2$,*

$$EU^S\left(C \mid C, (\bar{P}, \bar{\pi})\right) \geq EU^S\left(C' \mid C, (\bar{P}, \bar{\pi})\right) \qquad (IC^S)$$

and for any $V \in [0,1]^2, V' \in [0,1]^2$,

$$EU^B\left(V \mid V, (\bar{P}, \bar{\pi})\right) \geq EU^B\left(V' \mid V, (\bar{P}, \bar{\pi})\right) \qquad (IC^B)$$

The trading mechanism is ***individually rational*** *if for any $C \in [0,1]^2$, $V \in [0,1]^2$,*

$$EU^S\left(C \mid C, (\bar{P}, \bar{\pi})\right) \geq 0 \qquad (IR^S)$$

$$EU^B\left(V \mid V, (\bar{P}, \bar{\pi})\right) \geq 0 \qquad (IR^B)$$

Whenever the mechanism is both incentive compatible and individually rational we will refer to it as *incentive feasible.*

When the players are bargaining over several trading alternatives, one can distinguish between two types of *ex post* efficiency loss: (1) the traders can fail to reach an agreement to trade when positive gains from trade exist; and (2) when trade occurs, the traders can fail to realize *maximal* potential gains from trade by not choosing an object which maximizes the difference between the buyer's valuation and the seller's cost.

Thus, we can distinguish between

- fully *ex post* efficient (the first-best) trading mechanisms;
- trading mechanisms, which are *ex post* efficient with respect to the object selection (only), i.e. when trade takes place, the object selection is always *ex post* efficient, but the players do not always trade when positive gains from trade exist; and
- trading mechanisms which allow for both types of *ex post* efficiency loss.

The mechanisms described above are formally introduced in Definitions 2 and 3 below.

Definition 2 *A trading mechanism is* ***ex post efficient with respect to the object selection*** *if for any* $i, j \in \{0, 1, 2\}$, $\pi_i(V, C) \geq 0$ *implies that* $V_i - C_i \geq V_j - C_j$.

Full *ex post* efficiency requires that the players always trade when there are positive gains from trade and that they always choose an *ex post* efficient object.

Definition 3 *A direct trading mechanism* $\left(\overline{P}, \overline{\pi}\right)$ *is* ***ex post efficient (the first-best),*** *whenever it is ex post efficient with respect to the object selection and*

$$\pi_0(V, C) = \begin{cases} 1, & \text{if } V_i < C_i \text{ for all } i \in \{1, 2\} \\ 0, & \text{if } V_i \geq C_i \text{ for some } i \in \{1, 2\} \end{cases}$$

When it is not possible to ensure that every agent's type always makes the best trade, the next best option is to design a trading mechanism, which maximizes the total expected gains from trade.

Let $G(\bar{P},\overline{\pi})$ denote the total *ex ante* expected gains from trade (sum of the buyer's and seller's *ex ante* expected utilities) that could be achieved through a direct trading mechanism $(\bar{P},\overline{\pi})$. From our definition of the state-contingent prices and utilities, it follows that the balanced budget condition is automatically satisfied and G does not depend on the transfer payments.

Definition 4 *$(\bar{P},\overline{\pi})$ is a second-best mechanism whenever $(\bar{P},\overline{\pi}) = \underset{\overline{\pi},\overline{P}}{\operatorname{argmax}} \{G(\bar{P},\overline{\pi})\}$, subject to all incentive compatibility and individual rationality constraints.*

In the next section, we will address the question of whether the second-best mechanism can be *ex post* efficient with respect to the object selection in a simple trading environment, where buyer's (seller's) valuation for each object is either high or low.

1.3 Simple Trading Environments

Consider a simple trading environment, where the seller's cost as well as the buyer's valuation for an object i, where $i \in \{1,2\}$, can take one of the two values: **high** or **low**.

We will start with a brief overview of the characterization of incentive feasible mechanisms for bargaining with single object (due to Matsuo, 1989). We will use the single object case as a base of reference for the analysis of the bargaining problem with two objects.

1.3.1 Bilateral Trading with Single Object

Suppose there is only one potential object of trade owned by the seller. And suppose that the seller's cost for this object, $\mathbf{C}$, and the buyer's valuation, $\mathbf{V}$, are independent random variables such that $\mathbf{V}$ is distributed on the set $\left\{V^H, V^L\right\}$ and $\mathbf{C}$ is distributed on the set $\left\{C^H, C^L\right\}$. The players' beliefs are given by $Pr\left\{\mathbf{C} = C^H\right\} = \varepsilon$ and $Pr\left\{\mathbf{V} = V^L\right\} = \mu$, $0 < \varepsilon < 1$, $0 < \mu < 1$.

The problem of designing an incentive feasible mechanism for negotiating a trade in this environment was analyzed by Matsuo (1989). There are four possible configurations of the traders' reservation values that have to be considered:

(i) $V^H > V^L > C^H > C^L$, where mutually beneficial trade is always possible;

(ii) $C^H > V^H > V^L > C^L$, where mutually beneficial trade is not possible for the seller's high type;

(iii) $V^H > C^H > C^L > V^L$, where mutually beneficial trade is not possible for the buyer's low type;

(iv) $V^H > C^H > V^L > C^L$, where mutually beneficial trade is not always possible for both players.

Note, that only in *(iv)* the problem of *when* to trade is non-trivial, in the sense that the optimal decision depends on both players' private information. So, it is not surprising that *(iv)* is the only configuration of parameter values for which the first-best mechanism is not always incentive feasible. In *(i)*, *(ii)* or *(iii)*, the first-best mechanism is always (for any parameter values) incentive feasible. One can verify that in *(i)* – *(iii)* the first-best mechanism is to trade at every state where the valuation reported by the buyer is greater than or equal to the cost reported by the seller and not to trade, otherwise. An optimal transfer payment at all states where trade occurs, p, must be such that in *(i)* $C^H \leq p \leq V^L$, in *(ii)* $C^L \leq p \leq V^L$ and in *(iii)* $C^H \leq p \leq V^H$ and at the state where trade does not occur, the transfer payment is equal to zero. In *(iv)*, the mechanism design problem is more complicated and will be discussed below in some detail.

Suppose that $V^H > C^H > V^L > C^L$. The mechanism for bilateral trading presented in Table 1 specifies the trading decision and the transfer payment for every pair of the players' reports. One can verify that such mechanism is *ex post* efficient whenever it satisfies all incentive compatibility and individual rationality conditions.

Table 1

		Seller's reports	
		C^L	C^H
Buyer's reports	V^H	p^2, trade occurs	p^1; trade occurs
	V^L	p^3, trade occurs	p^4; (no trade)

Matsuo's main result, restated in Theorem 5, is that when trade is not always possible, there is a region of parameter values (beliefs and reservation values), where the first-best mechanism is not incentive feasible.

Theorem 5 *Suppose that $V^H > C^H > V^L > C^L$. Then, Bayesian incentive compatible, individually rational and ex post efficient mechanism exists if and only if*

$$(1-\mu)\varepsilon V^H + (1-\varepsilon) V^L \geq (1-\mu) C^H + \mu(1-\varepsilon) C^L \tag{1.1}$$

The interpretation of (1.1) becomes more transparent if we rewrite it as follows:

$$\begin{aligned}&(1-\mu)\varepsilon\left(V^H - C^H\right) + (1-\varepsilon)\mu\left(V^L - C^L\right) + (1-\varepsilon)(1-\mu)\left(V^H - C^L\right) \geq \\ &(1-\mu)(1-\varepsilon)\left(C^H - C^L\right) + (1-\mu)(1-\varepsilon)\left(V^H - V^L\right)\end{aligned} \tag{1.2}$$

Inequality (1.2) tells us that the first-best mechanism is incentive feasible if and only if the total *ex ante* expected gains from trade (the l.h.s. of (1.2)) exceed the total expected informational rents accruing to the traders (the r.h.s. of (1.2)). Intuitively, the expression in the r.h.s. of (1.2) can be referred to as informational rents because, whenever the true state is $\left(V^H, C^L\right)$, the difference $\left(C^H - C^L\right)$ is the least the type C^L can gain (under the worst possible price scheme from the seller's point of view)[5] by pretending that his type is C^H and the difference $\left(V^H - V^L\right)$ is the least the type V^H can gain by pretending that his type is V^L. Clearly, the higher the probability that the buyer's valuation is V^H, the stronger the seller's incentives to lie to the designer. And, similarly, the higher the probability that the seller's cost is C^L, the stronger the buyer's incentives to misrepresent his private information.

Let's focus on the region of parameter values where *ex post* efficiency is not incentive feasible. Consider the problem of designing a trading mechanism, which maximizes total *ex ante* expected gains from trade within the incentive feasible region.

Assuming that $0 \leq \pi(V,C) \leq 1$ for any realization (V,C), such that $V \geq C$, the inequality (1.1) can be rewritten as follows:

$$\begin{aligned}
&(1-\mu)\left(\varepsilon\left(V^H - C^L\right) - \left(C^H - C^L\right)\right)\cdot\pi\left(V^H, C^H\right) + \\
&(1-\varepsilon)\left(\mu\left(V^H - C^L\right) - \left(V^H - V^L\right)\right)\cdot\pi\left(V^L, C^L\right) + \\
&(1-\varepsilon)(1-\mu)\left(V^H - C^L\right)\cdot\pi\left(V^H, C^L\right) \geq 0
\end{aligned} \tag{1.3}$$

Because (1.3) does not depend on the transfer payments, the characterization of the second-best trading mechanism reduces to solving a simple linear programming (LP) problem, where we need to find the state contingent probabilities of trade that maximize the total *ex ante* expected gains from trade within the incentive feasible region defined by (1.3) and the non-negativity constraints.

Thus, for any second-best mechanism $(\bar{P}^*, \bar{\pi}^*)$, $\bar{\pi}^*$ is a solution of the following LP problem:

$$\max_{\bar{\pi}}\left\{\left(V^H - C^H\right)\cdot\pi\left(V^H, C^H\right) + \left(V^L - C^L\right)\cdot\pi\left(V^L, C^L\right) + \left(V^H - C^L\right)\cdot\pi\left(V^H, C^L\right)\right\}$$

$$\text{subject to}\begin{cases}(1.3)\\ 0 \leq \pi(V,C) \leq 1 \text{ for any } (V,C)\end{cases}$$

[5]The exact gain from such deviation would be equal to the price differential $\left(p^1 - p^2\right)$.

For any second best mechanism, the probability of trade at state $\left(V^H, C^L\right)$ is always equal to one, the probabilities of trade at states $\left(V^H, C^H\right)$, $\left(V^L, C^L\right)$ may be lower than one depending on parameter values.[6]

1.3.2 Bilateral Trading with Two Objects: The First-Best Mechanisms

Suppose that there are two potential objects of trade, q_1 and q_2, which are initially owned by the seller. The traders' reservation values for every potential object of trade are drawn from the same joint probability distribution defined on the set $\left\{V^H, V^L, C^H, C^L\right\}$. For any $i \in \{1, 2\}$, $\Pr\left\{C_i = C^H\right\} = \varepsilon$, $\Pr\left\{C_i = C^L\right\} = (1-\varepsilon)$, $0 < \varepsilon < 1$ and $\Pr\left\{V_i = V^L\right\} = \mu$, $\Pr\left\{V_i = V^H\right\} = (1-\mu)$, $0 < \mu < 1$.

In what follows we will focus on two configurations of the players' reservation values, where mutually beneficial trade is always possible and where the seller's high cost exceeds the buyer's low value. [7]

First, consider the most simple case, where mutually beneficial trade is possible for any realization of the players' types:

$$\boxed{V^H > V^L > C^H > C^L}$$

When $\underline{V^H - C^H < V^L - C^L}$, the low-cost object is always the best trade. The relevant price and object-selection schedule is shown in Table 2.

[6]For a complete characterization of the solution of the LP program, see the Appendix.

[7]The analysis of other configurations of the traders' reservation values is not presented here, as it does not offer any new insights into the problem.

Table 2

Seller's reports

	(C^H,C^L)	(C^L,C^H)	(C^L,C^L)	(C^H,C^H)
Buyer's reports (V^H,V^H)	$p^5;[q_2]$	$p^5;[q_1]$	$p^2;[q_1]\vee[q_2]$	$p^1;[q_1]\vee[q_2]$
(V^H,V^L)	$\boxed{p^{10};[q_2]}$	$p^9;[q_1]$	$p^6;[q_1]$	$p^7;[q_1]$
(V^L,V^H)	$p^9;[q_2]$	$\boxed{p^{10};[q_1]}$	$p^6;[q_2]$	$p^7;[q_2]$
(V^L,V^L)	$p^8;[q_2]$	$p^8;[q_1]$	$p^3;[q_1]\vee[q_2]$	$p^4;[q_1]\vee[q_2]$

In Table 2, each player's report is a vector, with the first element being his reservation value for q_1 and the second element being his reservation value for q_2. For every pair of reports (V,C), the designer specifies an outcome $(p;[q])$, where $[q]$ is the object of trade and p is its price. When the object selection is denoted $[q_i]$, it means that $\pi_i(V,C)=1$ and when the object selection is denoted $[q_1]\vee[q_2]$, it means that for any $i\in\{1,2\}$, $\pi_i(V,C)\in[0,1]$, where $\sum_{i=1}^{2}\pi_i(V,C)=1$.[8]

When $\underline{V^H-C^H>V^L-C^L}$, the high-value object is always the best trade. The relevant price and object-selection schedule is shown in Table 3. [9]

Table 3

Seller's reports

	(C^H,C^L)	(C^L,C^H)	(C^L,C^L)	(C^H,C^H)
Buyer's reports (V^H,V^H)	$p^5;[q_2]$	$p^5;[q_1]$	$p^2;[q_1]\vee[q_2]$	$p^1;[q_1]\vee[q_2]$
(V^H,V^L)	$\boxed{p^{11};[q_1]}$	$p^9;[q_1]$	$p^6;[q_1]$	$p^7;[q_1]$
(V^L,V^H)	$p^9;[q_2]$	$\boxed{p^{11};[q_2]}$	$p^6;[q_2]$	$p^7;[q_2]$
(V^L,V^L)	$p^8;[q_2]$	$p^8;[q_1]$	$p^3;[q_1]\vee[q_2]$	$p^4;[q_1]\vee[q_2]$

[8] $[q_1]\vee[q_2]$ will be interpreted as $\pi_1=\pi_2=\frac{1}{2}$ in the derivation of the incentive feasibilility condition.

[9] Note that it is possible to focus on the price schedule with only ten (not sixteen) distinctly different prices (without loss of generality) because both players' reservation values for the two objects are independently and identically distributed. Since the conditional probability of the seller's type being either (C^H,C^L) or (C^L,C^H), given that his reservation values for the two objects are different, is $\frac{1}{2}$, the two cases are symmetric. Same is true for the buyer's types (V^H,V^L) or (V^L,V^H).

Note that at states $((V^H, V^L), (C^H, C^L))$ and $((C^L, C^H), (V^L, V^H))$, where the seller's high-cost realization coincides with the buyer's high-value realization, it is optimal to trade the low-value-low-cost object whenever $V^H - C^H < V^L - C^L$ (as in Table (2)) and it is optimal to trade the high-value-high-cost object whenever $V^H - C^H > V^L - C^L$ (as in Table (3)). Note, that at these two states the optimal object selection is the same as the one preferred by the trader with the largest difference between the reservation values for q_1 and q_2.

When $V^H > V^L > C^H > C^L$, the Bayesian incentive compatible, individually rational and ex post efficient mechanisms exist for any beliefs and reservation values. The argument goes as follows. Clearly, the first-best mechanism has to be such that $p^1 = p^2 = p^3 = p^4$, $p^6 = p^7$, $p^5 = p^8$, since otherwise either the buyer's types $(V^H, V^H), (V^L, V^L)$ or the seller's types $(C^H, C^H), (C^L, C^L)$ would prefer to imitate one another. Also, note that although p^1, p^5, p^6 do not have to be equal, uniform pricing mechanism is always optimal. Notice, that whenever the price is the same at all states, the only way any of the traders can gain anything by misrepresenting his type is if he can induce the designer to choose a different object of trade. However, none of the traders can unilaterally affect the object selection at any state where any of them would prefer to do so. For instance, when $V^H - C^H < V^L - C^L$, the buyer would prefer to trade the high-value-high-cost object at the state $((C^H, C^L), (V^H, V^L))$, but he cannot induce his preferred choice by misrepresenting his type because the mechanism designer only needs to know the seller's type to be able to make an ex post efficient object selection.

Next, suppose that the trade is not always possible, i.e.

$$\boxed{V^H > C^H > V^L > C^L}$$

Here, the mechanism design problem is more complicated. For $V^H - C^H < V^L - C^L$, the relevant price and object selection schedule is shown in Table 4.

Table 4

		Seller's reports			
		(C^H, C^L)	(C^L, C^H)	(C^L, C^L)	(C^H, C^H)
Buyer's reports	(V^H, V^H)	$p^5; [q_2]$	$p^5; [q_1]$	$p^2; [q_1] \vee [q_2]$	$p^1; [q_1] \vee [q_2]$
	(V^H, V^L)	$\boxed{p^{10}; [q_2]}$	$p^9; [q_1]$	$p^6; [q_1]$	$p^7; [q_1]$
	(V^L, V^H)	$p^9; [q_2]$	$\boxed{p^{10}; [q_1]}$	$p^6; [q_2]$	$p^7; [q_2]$
	(V^L, V^L)	$p^8; [q_2]$	$p^8; [q_1]$	$p^3; [q_1] \vee [q_2]$	p^4; no trade

The incentives of the buyers' types (V^H, V^H), (V^H, V^L), (V^L, V^H) (and also those of the seller's types (C^L, C^L), (C^H, C^L), (C^L, C^H)) to imitate each other can always be eliminated by selecting the 'right' prices (for some parameter values, there exists a particularly simple first-best mechanism, such that $p^2 = p^5 = p^6 = p^9 = p^{10}$, $p^1 = p^7$ and $p^3 = p^8$). However, the buyer's types (V^H, V^H), (V^H, V^L), (V^L, V^H) may still want to imitate (V^L, V^L) and the seller's types (C^L, C^L), (C^H, C^L), (C^L, C^H) may still want to imitate (C^H, C^H) for certain configurations of the traders' reservation values.

Proposition 6 states the incentive feasibility conditions for all possible configurations of the traders' reservation values when $V^H > C^H > V^L > C^L$.

Proposition 6 *Suppose that $V^H > C^H > V^L > C^L$ and $V^H - C^H < V^L - C^L$. Then, Bayesian incentive compatible, individually rational and ex post efficient trading mechanisms exist if and only if*

$$\begin{aligned}&\left[(1-\mu^2)\varepsilon^2 + \mu(1-\mu)(1-\varepsilon)^2\right]V^H + \left[(1-\varepsilon^2) - \mu(1-\mu)(1-\varepsilon)^2\right]V^L \geq \\ &\left[(1-\mu^2) - \varepsilon(1-\varepsilon)(1-\mu^2)\right]C^H + \left[\mu^2(1-\varepsilon^2) + \varepsilon(1-\varepsilon)(1-\mu^2)\right]C^L\end{aligned} \quad (1.4)$$

Suppose that $V^H > C^H > V^L > C^L$ and $V^H - C^H > V^L - C^L$. Then, Bayesian incentive compatible, individually rational and ex post efficient trading mechanisms exist if and only *if*

$$\begin{aligned}&\left[\left(1-\mu^2\right)\varepsilon^2+\mu\left(1-\mu\right)\left(1-\varepsilon^2\right)\right]V^H+\left[\left(1-\varepsilon^2\right)-\mu\left(1-\mu\right)\left(1-\varepsilon^2\right)\right]V^L\geq\\&\left[\left(1-\mu^2\right)-\varepsilon\left(1-\varepsilon\right)\left(1-\mu\right)^2\right]C^H+\left[\mu^2\left(1-\varepsilon^2\right)+\varepsilon\left(1-\varepsilon\right)\left(1-\mu\right)^2\right]C^L\end{aligned} \tag{1.5}$$

Suppose that $V^H > C^H > V^L > C^L$ and $V^H - C^H = V^L - C^L$. Then, Bayesian incentive compatible, individually rational and ex post efficient trading mechanisms exist if and only if

$$\begin{aligned}&\left[\left(1-\mu^2\right)\varepsilon^2+\mu\left(1-\mu\right)\left(1-\varepsilon\right)\right]V^H+\left[\left(1-\varepsilon^2\right)-\mu\left(1-\mu\right)\left(1-\varepsilon\right)\right]V^L\geq\\&\left[\left(1-\mu^2\right)-\varepsilon\left(1-\varepsilon\right)\left(1-\mu\right)\right]C^H+\left[\mu^2\left(1-\varepsilon^2\right)+\varepsilon\left(1-\varepsilon\right)\left(1-\mu\right)\right]C^L\end{aligned} \tag{1.6}$$

Proof: see Appendix.

In order to understand how the presence of the second potential object of trade affects the players' incentives to reveal their private information, let's take a closer look at the case where the low-cost object is always the best trade. Let's rewrite (1.4) as follows:

$$\begin{aligned}&\boxed{\left(1-\mu^2\right)\varepsilon^2V^H+\left(1-\varepsilon^2\right)V^L}+\mu\left(1-\mu\right)\left(1-\varepsilon\right)^2\left(V^H-V^L\right)\geq\\&\boxed{\left(1-\mu^2\right)C^H+\mu^2\left(1-\varepsilon^2\right)C^L}-\varepsilon\left(1-\varepsilon\right)\left(1-\mu^2\right)\left(C^H-C^L\right)\end{aligned} \tag{1.7}$$

Recall that in the single-object case, the first-best mechanism is incentive feasible if and only if

$$\begin{aligned}&\left(1-\mu\right)\varepsilon V^H+\left(1-\varepsilon\right)V^L\geq\\&\left(1-\mu\right)C^H+\mu\left(1-\varepsilon\right)C^L\end{aligned} \tag{1.8}$$

The first two terms on both sides of the inequalities (1.7) and (1.8) look very similar, except that in (1.7) we have the probability of having *at least one* high-value object out of the two potential objects of trade, $(1-\mu^2)$, instead of the probability of a high value realization for one potential object of trade, $(1-\mu)$ (as in (1.8)).

The second terms in each side of (1.7) have no analogues in the incentive feasibility condition for the single object bargaining problem and, therefore, require some explanation. Consider the buyer's type $\left(V^H, V^L\right)$ reporting problem. If type $\left(V^H, V^L\right)$ reports $\left(V^L, V^L\right)$, then, given that the seller's type is $\left(C^L, C^L\right)$, there is a fifty percent chance that he will end up trading the low-value object. So, by misrepresenting his type, the buyer can (a) gain by paying a lower price for the selected object of trade, and (b) lose by not trading his preferred (high-value) object. In the latter case, the buyer's type $\left(V^H, V^L\right)$ expected utility loss would be equal to $\frac{(1-\varepsilon)^2}{2}\left(V^H - V^L\right)$. Therefore, $\mu(1-\mu)(1-\varepsilon)^2\left(V^H - V^L\right)$ is a measure of the buyer's expected loss associated with not trading his preferred object.

What is interesting here is that because the mechanism designer can discourage types $\left(V^H, V^H\right)$ and $\left(C^L, C^L\right)$ from reporting $\left(V^H, V^L\right)$ and $\left(C^H, C^L\right)$, respectively, by setting the 'right' prices, the expected informational rents do not depend on whether $\left(V^H - C^H\right)$ is higher, lower or equal to $\left(V^L - C^L\right)$.

Let G_1 denote the total *ex ante* expected gains from trade in the first-best mechanism for the case when $V^H - C^H > V^L - C^L$, let G_2 denote the total *ex ante* expected gains from trade in the first-best mechanism for the case when $V^H - C^H < V^L - C^L$, and let G_3 denote the total *ex ante* expected gains from trade in the first-best mechanism for the case when $V^H - C^H = V^L - C^L$.[10]

One can verify that the inequalities (1.4), (1.5), (1.6) can be rewritten as follows:

$$G_j \geq \left(1-\mu^2\right)(1-\varepsilon)\left(C^H - C^L\right) + \left(1-\varepsilon^2\right)(1-\mu)\left(V^H - V^L\right) \tag{1.9}$$

where $j \in \{1, 2, 3\}$. The l.h.s. of (1.9) represents the total *ex ante* expected gains from trade and the r.h.s. of (1.9) represents the total informational rents (same expression for

[10]For instance,

$$\begin{aligned} G_2 \equiv\ & \left(V^H - C^L\right)\left[(1-\mu)^2\left(1-\varepsilon^2\right) + 2\mu(1-\mu)(1-\varepsilon)\right] \\ & + \left(V^H - C^H\right)\left(1-\mu^2\right)\varepsilon^2 + \left(V^L - C^L\right)\left[\mu^2\left(1-\varepsilon^2\right) + 2\mu(1-\mu)\varepsilon(1-\varepsilon)\right] \end{aligned}$$

all three configurations of parameter values). Thus, the interpretation of the incentive feasibility condition in the two-object case is essentially the same as in the single object case i.e. the total ex ante expected gains from trade should exceed the total informational rents.

Note that when we add the second potential object of trade, the probability of trade, the total *ex ante* expected gains from trade and the expected informational rents become larger, compared to the single-object bargaining problem.

When trade is not always possible in bilateral bargaining with one good, a more accurate information towards V^H or C^L may be detrimental to *ex post* efficiency. Same is true here, e.g. higher probability that the seller's type is $\left(C^L, C^L\right)$ gives stronger incentives to the buyer's types $\left(V^H, V^H\right)$ and $\left(V^H, V^L\right)$ to report $\left(V^L, V^L\right)$.

1.3.3 Bilateral Trading with Two Objects: The Second-Best Mechanisms

Consider the configuration of reservation values ($V^H > C^H > V^L > C^L$), such that trade is not always possible, and suppose that the traders' beliefs and reservation values are such that the first-best mechanism is not incentive feasible.

Without loss of generality, we can assume that the probability of trade is equal to one at any state, where the buyer has high valuation for at least one object or where the seller has low cost for at least one object (because there is no uncertainty about the existence of gains from trade at any such state).

Let η_1 denote the probability of trade at states $\left(\left(C^H, C^H\right), \left(V^H, V^H\right)\right)$, $(\left(C^H, C^H\right)$, $\left(V^H, V^L\right))$ and $(\left(C^H, C^H\right), \left(V^L, V^H\right))$ (we can assume that it is the same at all three states without loss of generality). At any of these states, the players' reservation values for the *ex post* efficient object of trade are V^H, C^H.

Let η_3 denote the probability of trade at states $(\left(V^L, V^L\right), \left(C^L, C^L\right))$, $(\left(V^L, V^L\right)$, $\left(C^H, C^L\right))$ and $(\left(V^L, V^L\right), \left(C^L, C^H\right))$ (we can assume that it is the same at all three states without loss of generality). At any of these states, the players' reservation values for the *ex post* efficient object of trade are V^L, C^L.

Let ξ denote the probability of trading the *ex post* efficient object at the states $((C^H,C^L),(V^H,V^L)),((V^L,V^H),(C^L,C^H))$.[11]

The relevant price and object-selection schedule is shown in Table 5[12].

Table 5

		Seller reports (C^H,C^L)
Buyer's reports	(V^H,V^H)	p^2; $[q_2]$
	(V^H,V^L)	p^2; trade q_1 with prob. $(1-\xi)$ trade q_2 with prob. ξ
	(V^L,V^H)	p^2; $[q_2]$
	(V^L,V^L)	p^3; trade q_2 with prob. η_4 trade q_1 with prob. 0

		Seller reports (C^L,C^H)
Buyer's reports	(V^H,V^H)	p^2; $[q_1]$
	(V^H,V^L)	p^2; $[q_1]$
	(V^L,V^H)	p^2; trade q_1 with prob. ξ trade q_2 with prob. $(1-\xi)$
	(V^L,V^L)	p^3; trade q_1 with prob. η_4 trade q_2 with prob. 0

[11]Here, symmetry is assumed without loss of generality.

[12]Note that we don't need to specify a transfer payment for the *ex post* inefficient object that can be traded at states $((C^H,C^L),(V^H,V^L)),((V^L,V^H),(C^L,C^H))$ with probability $(1-\xi)$, because it does not affect the derivation of the incentive feasibility condition. Same is true for other states where ex post inefficient object can be traded with positive probability.

		Seller reports (C^L, C^L)
Buyer's reports	(V^H, V^H)	p^2; $[q_1] \vee [q_2]$
	(V^H, V^L)	p^2; $[q_1]$
	(V^L, V^H)	p^2; $[q_2]$
	(V^L, V^L)	p^3; trade q_1 with prob. $\frac{1}{2}\eta_3$ trade q_2 with prob. $\frac{1}{2}\eta_3$

		Seller reports (C^H, C^H)
Buyer's reports	(V^H, V^H)	p^1; trade q_1 with prob. $\frac{1}{2}\eta_1$ trade q_2 with prob. $\frac{1}{2}\eta_1$
	(V^H, V^L)	p^1; trade q_1 with prob. η_1 trade q_2 with prob. 0
	(V^L, V^H)	p^1; trade q_1 with prob. 0 trade q_2 with prob. η_1
	(V^L, V^L)	p^4; (no trade)

The conditional probability of trading the *ex post* efficient object, given that trade occurs, at all states where the buyer's type is (V^L, V^L) and at all states where the seller's type is (C^H, C^H), is equal to one, because the buyer's type (V^H, V^H) and the seller's type (C^L, C^L) incentive problems can not be mitigated by allowing the *ex post* inefficient object to be sometimes traded at these states.

Lowering ξ below 1 would make sense only if the mechanism designer wanted to make honest reporting more rewarding for the type (V^H, V^L). Notice, that the buyer's type (V^H, V^L) would actually prefer to trade the *ex post* inefficient object at state $((V^H, V^L), (C^H, C^L))$. However, lower ξ weakens the seller's type (C^H, C^L) incentives to report truthfully. Thus, in order to determine the net impact of ξ on both traders' incentives, we have to take a closer look at the incentive feasibility condition.

When $V^H > C^H > V^L > C^L$ and $V^H - C^H < V^L - C^L$, the second-best mechanism

$(\eta_1^*, \eta_3^*, \xi^*)$ is a solution of the following LP problem:

$$\max_{\eta_1,\eta_3,\xi} \{G(\eta_1,\eta_3,\xi)\}$$
subject to
$$G(\eta_1,\eta_3,\xi) \geq \eta_1 (1-\mu^2)(1-\varepsilon)\left(C^H - C^L\right) + \eta_3 (1-\varepsilon^2)(1-\mu)\left(V^H - V^L\right)$$
$$0 \leq \eta_1 \leq 1$$
$$0 \leq \eta_3 \leq 1$$
$$0 \leq \xi \leq 1$$

where

$$\begin{aligned} G(\eta_1,\eta_3,\xi) \equiv\ & \left(V^H - C^L\right)\left((1-\mu)^2\left(1-\varepsilon^2\right) + 2\mu(1-\mu)(1-\varepsilon)\right) \\ & + \left(V^H - C^H\right)\left(\eta_1\left(1-\mu^2\right)\varepsilon^2 + 2\mu(1-\mu)\varepsilon(1-\varepsilon)(1-\xi)\right) \\ & + \left(V^L - C^L\right)\left(\eta_3\mu^2\left(1-\varepsilon^2\right) + 2\mu(1-\mu)\varepsilon(1-\varepsilon)\xi\right) \end{aligned}$$

Because $V^H - C^H < V^L - C^L$, the objective function is strictly increasing in ξ and lowering ξ shifts the boundary of the incentive feasibility region inward. So, in order to make honest reporting incentive feasible, the mechanism designer has to lower either η_1 or η_3.

Proposition 7 summarizes the argument presented above.

Proposition 7 *Suppose that $V^H > C^H > V^L > C^L$. The second-best trading mechanism is always ex post efficient with respect to the object selection.*

Note that when $V^H - C^H \neq V^L - C^L$, the player with smaller difference between reservation values always prefers to trade the *ex post* inefficient object at states $\left(\left(C^H, C^L\right), \left(V^H, V^L\right)\right)$, $\left(\left(V^L, V^H\right), \left(C^L, C^H\right)\right)$, while the player with larger difference always prefers to trade the *ex post* efficient object. Therefore, decreasing the probability of trading *ex post* efficient object only reduces total expected gains from trade and, at the same time, weakens the players' incentives to report truthfully.

Thus, in simple trading environments, where each player's reservation values for both potential objects of trade are independently and identically distributed, any second-best mechanism is *ex post* efficient with respect to the object selection.[13] The next logical question is under what conditions (if any) this property is generalizable.

1.4 Rich Trading Environments with Finite Types

Consider a trading environment, where the traders' reservation values are drawn from the discrete probability distributions defined on the set $\{0, ..., 1\}$. First, we will look at the single object case.

1.4.1 Bilateral Trading with Single Object: Incentive Feasible Mechanisms

Suppose that there is one potential object of trade. And suppose that the buyer's valuation for this object is drawn from the discrete probability distribution $F^B(V)$ defined on the set $\{0, ..., 1\}$ and the seller's cost is drawn from the discrete probability distribution $F^S(C)$ defined on the set $\{0, ..., 1\}$, where $0 < f^S(C) < 1$, $0 < f^B(V) < 1$.

Consider the following configuration of the players' reservation values. Suppose that for every pair of types $V^{(k)}, C^{(k)}$ of the same rank k, where $k \in N$, $C^{(k+1)} > V^{(k)} \geq C^{(k)}$ (which means that gains from trade are positive whenever the buyer's type is at least of the same rank as the seller's type).[14] Let $\Delta_{(V^k)} \equiv V^{(k+1)} - V^{(k)}$ and $\Delta_{(C^k)} \equiv C^{(k)} - C^{(k-1)}$. To keep things simple, let's assume that $\Delta_{(V^k)} = \Delta_{(C^k)} = \Delta$, for any $k \in N$. The generalized incentive feasibility condition for the discrete type sets is presented in Proposition 8.

[13] This result may not hold if we change our assumptions about the distributions of the players' reservation values.

[14] This assumption is needed to make sure that the players' valuation sets overlap properly.

Proposition 8 *For any incentive feasible mechanism* $(\bar{P}, \bar{\pi})$, *such that for any* $(V,C) \in \{0,...,1\} \times \{0,...,1\}$,

$$\begin{cases} 0 \leq \pi(V,C) \leq 1, \text{ if } V \geq C \\ \pi(V,C) = 0, \text{ otherwise} \end{cases}$$

the following has to be true:

$$\sum_{V=0}^{1}\sum_{C=0}^{1}\left(\left(V-\Delta\frac{\left(1-F^{B}(V)\right)}{f^{B}(V)}\right)-\left(C+\Delta\frac{F^{S}(C-\Delta)}{f^{S}(C)}\right)\right)\pi(V,C)\,f^{S}(C)\,f^{B}(V)\geq 0$$

Proof: see Appendix.

We will refer to the expression $\left(V-\Delta\frac{(1-F^{B}(V))}{f^{B}(V)}\right)$ as the buyer's 'virtual valuation' and we will refer to the expression $\left(C+\Delta\frac{F^{S}(C-\Delta)}{f^{S}(C)}\right)$ as the seller's 'virtual cost'.[15] Proposition 8 states that the direct mechanism for bilateral trading is incentive feasible if and only if the total expected gains from trade for the players' 'virtual types' are non-negative. The incentive feasibility condition can be rewritten as follows:

$$\begin{aligned} &\sum_{V=0}^{1}\sum_{C=0}^{1}(V-C)\,\pi(V,C)\,f^{S}(C)\,f^{B}(V) \qquad (1.10)\\ \geq\ &\Delta\sum_{V=0}^{1}\sum_{C=0}^{1}\left(1-F^{B}(V)\right)f^{S}(C)\,\pi(V,C)+\Delta\sum_{V=0}^{1}\sum_{C=0}^{1}F^{S}(C-\Delta)\,f^{B}(V)\,\pi(V,C)\end{aligned}$$

The l.h.s of (1.10) is the total *ex ante* expected gains from trade and the r.h.s. of (1.10) is the total informational rents. Note, that the structure of informational rents is essentially the same as in the simple trading environment. The first term in the r.h.s of (1.10) is a

[15]Indeed, the expressions described above are the discrete approximations of the 'virtual' reservation values introduced by *Myerson* and *Satterthwaite* in their model of bilateral bargaining over the price of a single object with the continuum of types, (*Myerson* and *Satterthwaite*,1983). Assuming that the buyer's valuation and the seller's cost are distributed with strictly positive densities $g^{B}(V)$ and $g^{S}(C)$, respectively, the buyer's 'virtual valuation' is equal to $\left(V-\frac{1-G^{B}(V)}{g^{B}(V)}\right)$ and the seller's 'virtual cost' is equal to $\left(C+\frac{G^{S}(C)}{g^{S}(C)}\right)$.

measure of expected gains from all possible downward deviations from the true type for all possible types of the buyer, and the second term in the r.h.s. of (1.10) is a measure of expected gains from all possible upward deviations from the true type for all types of the seller.

1.4.2 Bilateral Trading Model with Two Objects: Incentive Feasible Mechanisms

Suppose that there are two potential objects of trade and that the buyer's valuation for every object q_i is drawn from the discrete probability distribution $F^B(V_i)$ defined on the set $\{0, ..., 1\}$, where $0 < f^B(V_i) < 1, i \in \{1,2\}$. The seller's cost for every q_i is drawn from the discrete probability distribution $F^S(C_i)$ defined on $\{0, ..., 1\}$, where $0 < f^S(C_i) < 1, i \in \{1,2\}$. We will assume that for every pair of $V^{(k)}, C^{(k)}$ of the same rank k, where $k \in N$, $C^{(k+1)} > V^{(k)} \geq C^{(k)}$ and for any $k \in N$, $V^{(k+1)} - V^{(k)} = C^{(k)} - C^{(k-1)} = \Delta$.

What are the necessary conditions for a second-best mechanism to be *ex post* efficient with respect to the object selection?

Define r_i^B to be a mapping from the buyer's type space into the real line, $r_i^B : \{0, ..., 1\} \times \{0, ..., 1\} \to R$, for $i \in \{1,2\}$, and, similarly, let r_i^S be a mapping from the seller's type space into the real line, $r_i^S : \{0, ..., 1\} \times \{0, ..., 1\} \to R$, for $i \in \{1,2\}$.

And suppose that given an incentive feasible mechanism $(\bar{P}, \bar{\pi})$, there exist r_i^B, r_i^S, for $i \in \{1,2\}$, such that the incentive feasibility condition can be represented as follows (which is essentially parallel to the incentive feasibility condition for the single object trading problem described in Proposition 8):

$$\sum_C \sum_V \sum_{i=1}^{2} \left[\left(r_i^B(V) - r_i^S(C) \right) \pi_i(V, C) \right] \mathrm{f}^B(V) \mathrm{f}^S(C) \geq 0 \qquad (1.11)$$

where $\mathrm{f}^S(C) = f^S(C_1) f^S(C_2)$, $\mathrm{f}^B(V) = f^B(V_1) f^B(V_2)$ and $\pi_i(V, C)$ is the probability of trading object q_i at state (V, C).

We can think of $r_i^B(V)$ as the type V buyer's 'virtual valuation' for object q_i and, similarly, we can think of $r_i^S(C)$ as the type C seller's 'virtual cost' of object q_i. Here, we are assuming that it is possible to eliminate prices from the optimization problem by collapsing all IC and IR constraints into one inequality which contains only state-contingent probabilities of trading different objects.

Then, any incentive feasible trading mechanism $(\bar{P}^*, \bar{\pi}^*)$, which maximizes total *ex ante* expected gains from trade, must be such that $\bar{\pi}^*$ is a solution of the following LP problem:

$$\max_{\bar{\pi}} \left[\sum_C \sum_V [(V_1 - C_1)\pi_1(V,C) + (V_2 - C_2)\pi_2(V,C)] f^B(V) f^S(C) \right]$$

subject to

$$\sum_C \sum_V \sum_{i=1}^{2} \left[\left(r_i^B(V) - r_i^S(C) \right) \pi_i(V,C) \right] f^B(V) f^S(C) \geq 0 \tag{1.12}$$

$$\begin{aligned} \pi_1(V,C) + \pi_2(V,C) &\leq 1 \\ \pi_1(V,C) &\geq 0 \\ \pi_2(V,C) &\geq 0 \end{aligned} \tag{1.13}$$

for any $(V,C) \in \{0,...,1\}^2 \times \{0,...,1\}^2$

Let λ_0 denote the dual variable associated with (1.12) and $\lambda_{(V,C)}$ denote the dual variable associated with (1.13). Then, we can write the dual LP problem as follows:

$$\min_{\lambda} \left[-\lambda_0 \times 0 + \sum_V \sum_C \lambda_{(V,C)} \right]$$

subject to

$$-\lambda_0 \left(r_1^B(V) - r_1^S(C) \right) f^B(V) f^S(C) + \lambda_{(V,C)} \geq (V_1 - C_1) f^B(V) f^S(C) \tag{1.14}$$

$$-\lambda_0 \left(r_2^B(V) - r_2^S(C) \right) f^B(V) f^S(C) + \lambda_{(V,C)} \geq (V_2 - C_2) f^B(V) f^S(C) \tag{1.15}$$

$$\lambda_{(V,C)} \geq 0, \; \lambda_0 \geq 0$$

for any $(V,C) \in \{0,...,1\}^2 \times \{0,...,1\}^2$

Let's rewrite (1.14) and (1.15) as

$$\lambda_{(V,C)} \geq \left((V_1 - C_1) + \lambda_0 \left(r_1^B(V) - r_1^S(C)\right)\right) f^B(V) f^S(C) \quad (1.16)$$
$$\lambda_{(V,C)} \geq \left((V_2 - C_2) + \lambda_0 \left(r_2^B(V) - r_2^S(C)\right)\right) f^B(V) f^S(C) \quad (1.17)$$

Clearly, $\lambda^*_{(V,C)} = \max\left\{ \max_{i\in\{1,2\}} \left\{\left((V_i - C_i) + \lambda_0 \left(r_i^B(V) - r_i^S(C)\right)\right) f^B(V) f^S(C)\right\}, 0\right\}$.

Suppose that $(V_1 - C_1) + \lambda_0 \left(r_1^B(V) - r_1^S(C)\right) > (V_2 - C_2) + \lambda_0 \left(r_2^B(V) - r_2^S(C)\right)$. Then, at the optimum, (1.17) must be satisfied as strict inequality and the corresponding dual variable $\pi_2(V,C) = 0$. Otherwise, it must be the case that (1.16) is satisfied as strict inequality and the corresponding dual variable $\pi_1(V,C) = 0$. Note that if the 'virtual' reservation values are equal to the actual reservation values, *ex post* inefficient object is never traded.

Let $v_i(V)$ and $c_i(C)$ denote the buyer's type V and seller's type C expected informational rents, respectively, which arise because of the traders' incentives to misrepresent their true valuations for object q_i, i.e.

$$v_i(V) = V_i - r_i^B(V)$$

and

$$c_i(C) = r_i^S(C) - C_i$$

Proposition 9 summarizes the analysis presented above.

Proposition 9 *Consider a state* (V, C)*, where* $V_j \geq C_j, V_i \geq C_i$ *and* $V_j - C_j \geq V_i - C_i$ *for* $i, j \in \{1, 2\}$ *(*$\Leftrightarrow q_j$ *is the ex post efficient object) and a second-best mechanism* $(\bar{P}^*, \bar{\pi}^*)$, *such that* $\bar{\pi}^*$ *solves the LP problem described above. Then,* $\pi_i^*(V, C) = 0$, *whenever*

$$(v_j(V) - v_i(V)) + (c_j(C) - c_i(C)) < \left(\frac{1 + \lambda_0^*}{\lambda_0^*}\right)((V_j - C_j) - (V_i - C_i)) \tag{1.18}$$

Thus, the ex post inefficient object q_i is never traded in a second-best mechanism $(\bar{P}^*, \bar{\pi}^*)$ only if the difference between the gains from trading q_j and the gains from trading q_i is large enough to offset the sum of the differences between buyer's and seller's informational rents associated with trading q_j and q_i. An immediate and intuitively obvious implication of Proposition 9 is that **if each trader's 'virtual valuations' for both objects differ from his true reservation values by equal amounts, the outcome of the bargaining process is always ex post efficient with respect to the object selection.**

Corollary 10 *A second-best mechanism* $(\bar{P}^*, \bar{\pi}^*)$ *is ex post efficient with respect to the object selection if and only if* $V_j - C_j \geq V_i - C_i$ *implies (1.18).*

Note, that since the r.h.s. of (1.18) is positive (by construction), in order to show that the mechanism $(\bar{P}^*, \bar{\pi}^*)$ is *ex post* efficient with respect to the object selection, it would be sufficient to show that the l.h.s. of (1.18) is non-positive for any (V, C), such that $V_j - C_j \geq V_i - C_i$. However, the latter cannot be easily demonstrated without explicitly deriving the incentive feasibility condition. The analysis of a special class of trading mechanisms, k-double auctions, presented in the next chapter shows that the difference between the buyer's valuations for the high and low value objects always exceeds the difference between his bids for the high and low value object for all types of the buyer and the same is true for the seller. The latter implies that k-double auction mechanisms are unlikely to be *ex post* efficient with respect to the object selection in sufficiently rich trading environments.

APPENDIX

Proposition A1

Suppose that $V^H > C^H > V^L > C^L$ and the first-best mechanism is not incentive feasible. Let $C^H = \lambda^H V^H + \left(1-\lambda^H\right) C^L$ and let $V^L = \left(1-\lambda^L\right) V^H + \lambda^L C^L$, where $0 \le \lambda^H \le 1$, $0 \le \lambda^L \le 1$. Then, for any second-best mechanism $(\bar{P}^*, \bar{\pi}^*)$ the following has to be true:

(i) $\pi^*\left(V^H, C^L\right) = 1$, $\pi^*\left(V^H, C^H\right) = 1$, $\pi^*\left(V^L, C^L\right) = \frac{(1-\mu)\left(V^H - C^H\right)}{(1-\varepsilon)\left(\left(V^H - V^L\right) - \mu\left(V^H - C^L\right)\right)}$, *whenever* $(\mu - \varepsilon)\lambda^H \lambda^L > \mu\lambda^H - \varepsilon\lambda^L$;

(ii) $\pi^*\left(V^H, C^L\right) = 1$, $\pi^*\left(V^H, C^H\right) = \frac{(1-\varepsilon)\left(V^L - C^L\right)}{(1-\mu)\left(\left(C^H - C^L\right) - \varepsilon\left(V^H - C^L\right)\right)}$, $\pi^*\left(V^L, C^L\right) = 1$, *whenever* $(\mu - \varepsilon)\lambda^H \lambda^L < \mu\lambda^H - \varepsilon\lambda^L$;

(iii) whenever $(\mu - \varepsilon)\lambda^H \lambda^L = \mu\lambda^H - \varepsilon\lambda^L$, *the optimum can be any point on the interval connecting the points in (i) and (ii).*

Proof.

For any second-best mechanism $(\bar{P}^*, \bar{\pi}^*)$, $\bar{\pi}^*$ is a solution of the following LP problem:

$$\max_{\bar{\pi}} \left\{ \begin{array}{l} \pi\left(V^H, C^H\right) \times \left(V^H - C^H\right) + \\ \pi\left(V^L, C^L\right) \times \left(V^L - C^L\right) + \\ \pi\left(V^H, C^L\right) \times \left(V^H - C^L\right) \end{array} \right\}$$

subject to

$$\begin{array}{l} \pi\left(V^H, C^H\right) \times (1-\mu)\left(\varepsilon\left(V^H - C^L\right) - \left(C^H - C^L\right)\right) + \\ \pi\left(V^L, C^L\right) \times (1-\varepsilon)\left(\mu\left(V^H - C^L\right) - \left(V^H - V^L\right)\right) + \\ \pi\left(V^H, C^L\right) \times (1-\varepsilon)(1-\mu)\left(V^H - C^L\right) \ge 0 \end{array}$$

and

$0 \le \pi(V, C) \le 1$, for any realization (V, C)

Consider the case when $\varepsilon \le \frac{\left(C^H - C^L\right)}{\left(V^H - C^L\right)}$ and $\mu \le \frac{\left(V^H - V^L\right)}{\left(V^H - C^L\right)}$.

Let's fix the value of the total gains from trade as some level z, and rewrite the objective function in terms of $\pi\left(V^L, C^L\right)$ as follows:

$$\pi\left(V^L,C^L\right)=\frac{z-(1-\varepsilon)(1-\mu)\left(V^H-C^L\right)}{\mu(1-\varepsilon)(V^L-C^L)}-\pi\left(V^H,C^H\right)\times\frac{\varepsilon(1-\mu)\left(V^H-C^H\right)}{\mu(1-\varepsilon)(V^L-C^L)}$$

Similarly, rewrite the incentive feasibility constraint in terms of $\pi\left(V^L,C^L\right)$ as well (since $\varepsilon<\frac{(C^H-C^L)}{(V^H-C^L)}$ and $\mu<\frac{(V^H-V^L)}{(V^H-C^L)}$, all expressions in the parentheses are now positive)

$$\frac{(1-\mu)\left(V^H-C^L\right)}{((V^H-V^L)-\mu(V^H-C^L))}-\pi\left(V^H,C^H\right)\times\frac{(1-\mu)\left(\left(C^H-C^L\right)-\varepsilon\left(V^H-C^L\right)\right)}{(1-\varepsilon)((V^H-V^L)-\mu(V^H-C^L))}\geq\pi\left(V^L,C^L\right)$$

Now, the optimal solution of the LP problem can be described on the ($\pi\left(V^H,C^H\right)$, $\pi\left(V^L,C^L\right)$) plane.

Let a denote the slope of the objective function, where $a=\frac{\varepsilon(1-\mu)\left(V^H-C^H\right)}{(1-\varepsilon)\mu(V^L-C^L)}$ and

let b denote the slope of the constraint, where $b=\frac{(1-\mu)\left(\left(C^H-C^L\right)-\varepsilon\left(V^H-C^L\right)\right)}{(1-\varepsilon)((V^H-V^L)-\mu(V^H-C^L))}$. Then,

(i) whenever $a>b$, $\pi^*\left(V^H,C^H\right)=1$, $\pi^*\left(V^L,C^L\right)=\frac{(1-\mu)\left(V^H-C^H\right)}{(1-\varepsilon)((V^H-V^L)-\mu(V^H-C^L))}$;

(ii) whenever $a<b$, $\pi^*\left(V^H,C^H\right)=\frac{(1-\varepsilon)\left(V^L-C^L\right)}{(1-\mu)((C^H-C^L)-\varepsilon(V^H-C^L))}$, $\pi^*\left(V^L,C^L\right)=1$

(iii) whenever $a=b$, the optimum can be any point on the interval connecting the points in (i) and (ii).

The remaining two cases are trivial.

If $\varepsilon\geq\frac{(C^H-C^L)}{(V^H-C^L)}$ and $\mu\leq\frac{(V^H-V^L)}{(V^H-C^L)}$, the second-best mechanism is such that $\pi^*\left(V^H,C^H\right)=1$ and $\pi^*\left(V^L,C^L\right)=\frac{(1-\mu)\left(V^H-C^H\right)}{(1-\varepsilon)((V^H-V^L)-\mu(V^H-C^L))}$.

If $\mu\geq\frac{(V^H-V^L)}{(V^H-C^L)}$ and $\varepsilon\leq\frac{(C^H-C^L)}{(V^H-C^L)}$, the second-best mechanism must be such that $\pi^*\left(V^H,C^H\right)=\frac{(1-\varepsilon)\left(V^L-C^L\right)}{(1-\mu)((C^H-C^L)-\varepsilon(V^H-C^L))}$ and $\pi^*\left(V^L,C^L\right)=1$. ■

As it follows from Proposition A1, it is optimal to lower the probability of trade at the state where potential gains from trade are smaller. And whenever the gains from trade at states $\left(V^H, C^H\right)$, $\left(V^L, C^L\right)$ are equal, it is optimal to lower the probability of trade at state $\left(V^H, C^H\right)$, if the probability of the buyer's worst type is higher than the probability of the seller's worst type, and it is optimal to lower the probability of trade at state $\left(V^L, C^L\right)$, otherwise.

Proof of Proposition 6

We will show that (1.4) is both necessary and sufficient condition for the first-best mechanism to be incentive feasible when $V^H > C^H > V^L > C^L$ and $V^H - C^H < V^L - C^L$. The proofs for other configurations are very similar.

(Necessity)

Suppose that (1.4) fails and for some parameter values there exists an *ex post* efficient mechanism which is incentive feasible.

Let's rewrite the buyer's types (V^H, V^L) and (V^H, V^H) IC constraints versus (V^L, V^L) as follows:

$$V^H - \frac{2\varepsilon(1-\varepsilon)(p^5 - p^8) + (1-\varepsilon)^2(p^2 - p^3)}{\varepsilon^2} \geq p^1 \quad (1.19)$$

$$V^H - \frac{2\varepsilon(1-\varepsilon)\left(\frac{p^9+p^{10}}{2} - p^8\right) + (1-\varepsilon)^2(p^6 - p^3) - \frac{(1-\varepsilon)^2}{2}\left(V^H - V^L\right)}{\varepsilon^2} \geq p^7 \quad (1.20)$$

After multiplying both sides of (1.19) by $(1-\mu)^2$ and both sides of (1.20) by $2\mu(1-\mu)$, and then adding them up, we obtain the following inequality:

$$\begin{aligned}&(1-\mu^2)\varepsilon^2 V^H + \mu(1-\mu)(1-\varepsilon)^2\left(V^H - V^L\right) + (1-\mu^2)\left(2\varepsilon(1-\varepsilon)p^8 + (1-\varepsilon)^2 p^3\right) \geq \\ &(1-\mu)^2\left(2\varepsilon(1-\varepsilon)p^5 + (1-\varepsilon)^2 p^2\right) + 2\mu(1-\mu)\left(2\varepsilon(1-\varepsilon)\frac{p^9+p^{10}}{2} + (1-\varepsilon)^2 p^6\right) + \\ &(1-\mu)^2\varepsilon^2 p^1 + 2\mu(1-\mu)\varepsilon^2 p^7\end{aligned} \quad (1.21)$$

Also, let's rewrite the seller's types (C^H, C^L) and (C^L, C^L) IC constraint versus (C^H, C^H) as follows:

$$p^3 \geq \frac{2\mu(1-\mu)(p^7 - p^6) + (1-\mu)^2(p^1 - p^2)}{\mu^2} + C^L \quad (1.22)$$

$$p^8 \geq \frac{2\mu(1-\mu)\left(p^7 - \frac{p^9+p^{10}}{2}\right) + (1-\mu)^2(p^1 - p^5) - \frac{(1-\mu^2)}{2}\left(C^H - C^L\right)}{\mu^2} + C^L \quad (1.23)$$

After multiplying both sides of (1.22) by $(1-\varepsilon)^2$ and both sides of (1.23) by $2\varepsilon(1-\varepsilon)$, and then adding them up, we obtain the following inequality:

$$\begin{aligned}&\mu^2\left((1-\varepsilon)^2 p^3 + 2\varepsilon(1-\varepsilon)p^8\right) + \varepsilon(1-\varepsilon)\left(1-\mu^2\right)\left(C^H - C^L\right) \geq \\ &\left(1-\varepsilon^2\right)\left(2\mu(1-\mu)p^7 + (1-\mu)^2 p^1\right) + \mu^2\left(1-\varepsilon^2\right)C^L \\ &-(1-\varepsilon)^2\left(2\mu(1-\mu)p^6 + (1-\mu)^2 p^2\right) - 2\varepsilon(1-\varepsilon)\left(2\mu(1-\mu)\tfrac{p^9+p^{10}}{2} + (1-\mu)^2 p^5\right)\end{aligned} \tag{1.24}$$

Adding (1.21) and (1.24) we have the following inequality:

$$\begin{aligned}&\varepsilon(1-\varepsilon)\left(1-\mu^2\right)\left(C^H - C^L\right) + \varepsilon^2\left(1-\mu^2\right)V^H + \mu(1-\mu)(1-\varepsilon)^2\left(V^H - V^L\right) \geq \\ &\mu^2\left(1-\varepsilon^2\right)C^L + \left(2\mu(1-\mu)p^7 + (1-\mu)^2 p^1\right) - \left(2\varepsilon(1-\varepsilon)p^8 + (1-\varepsilon)^2 p^3\right)\end{aligned} \tag{1.25}$$

Adding the buyer's type $\left(V^L, V^L\right)$ and the seller's type $\left(C^H, C^H\right)$ IR conditions we have the following inequality

$$\begin{gathered}\left(2\mu(1-\mu)p^7 + (1-\mu)^2 p^1\right) - \left(2\varepsilon(1-\varepsilon)p^8 + (1-\varepsilon)^2 p^3\right) \geq \\ \left(1-\mu^2\right)C^H - \left(1-\varepsilon^2\right)V^L\end{gathered} \tag{1.26}$$

(1.24) together with (1.26) are inconsistent with the converse of (1.4) ⊗ ■

(Sufficiency)

Suppose that (1.4) holds. Then, we should be able to specify a price vector that satisfies all incentive compatibility and individual rationality conditions.

Let $p^3 = p^8 = V^L$, $p^1 = p^7 = C^H$ and let $p^6 = C^H$, then p^5 can be defined as follows:

$$\min\left\{C^H, \frac{\varepsilon^2(1-\mu)^2\left(V^H-C^H\right)+\left(1-\varepsilon^2\right)(1-\mu)^2V^L-(1-\varepsilon)^2\left(1-\mu^2\right)C^H+\mu^2(1-\varepsilon)^2\left(V^L-C^L\right)+2\mu(1-\mu)(1-\varepsilon)^2p^6}{2\varepsilon(1-\varepsilon)(1-\mu)^2}\right\} \geq p^5 \tag{1.27}$$

$$p^5 \geq \max \left\{V^L, P\right\},$$

where

$$P = \frac{2\mu(1-\mu)(1-\varepsilon)^2 p^6 + 2\varepsilon(1-\varepsilon)\left(1-\mu^2\right)C^H}{2\varepsilon(1-\varepsilon)(1-\mu)^2} - \frac{2\varepsilon(1-\varepsilon)\mu^2\left(V^L-C^L\right)+\varepsilon(1-\varepsilon)\left(1-\mu^2\right)\left(C^H-C^L\right)+2\mu(1-\mu)\varepsilon^2\left(V^H-C^H\right)+\mu(1-\mu)(1-\varepsilon)^2\left(V^H-V^L\right)+2\mu(1-\mu)\left(1-\varepsilon^2\right)V^L}{2\varepsilon(1-\varepsilon)(1-\mu)^2} \tag{1.28}$$

Because (1.4) implies that the l.h.s of (1.27) is greater than or equal to the r.h.s. of (1.28), p^5 is well defined.

Then given p^5 and p^6, (1.27) must hold and can be rewritten as follows

$$(1-\mu)^2 \varepsilon^2 \left(V^H - C^H\right) + (1-\mu)^2 \left(1-\varepsilon^2\right) V^L - (1-\mu)^2 2\varepsilon (1-\varepsilon) p^5 \geq$$
$$(1-\varepsilon)^2 \left(1-\mu^2\right) C^H - \mu^2 (1-\varepsilon)^2 \left(V^L - C^L\right) - 2\mu (1-\mu)(1-\varepsilon)^2 p^6$$

which implies that p^2 is well defined as well

$$\min\left\{C^H, \frac{\varepsilon^2\left(V^H-C^H\right)+\left(1-\varepsilon^2\right)V^L-2\varepsilon(1-\varepsilon)p^5}{(1-\varepsilon)^2}\right\} \geq p^2 \geq \max\left\{V^L, \frac{\left(1-\mu^2\right)C^H-\mu^2\left(V^L-C^L\right)-2\mu(1-\mu)p^6}{(1-\mu)^2}\right\} \tag{1.29}$$

Also, given p^5 and p^6, (1.28) must hold and can be rewritten as follows

$$2\mu(1-\mu)\varepsilon^2\left(V^H - C^H\right) + 2\mu(1-\mu)\left(1-\varepsilon^2\right)V^L +$$
$$\mu(1-\mu)(1-\varepsilon)^2\left(V^H - V^L\right) - 2\mu(1-\mu)(1-\varepsilon)^2 p^6$$
$$\geq$$
$$2\varepsilon(1-\varepsilon)\left(1-\mu^2\right)C^H - 2\varepsilon(1-\varepsilon)\mu^2\left(V^L - C^L\right) -$$
$$\varepsilon(1-\varepsilon)\left(1-\mu^2\right)\left(C^H - C^L\right) - 2\varepsilon(1-\varepsilon)(1-\mu)^2\left(p^5\right)$$

which implies that $\frac{p^9+p^{10}}{2}$ is well defined too and

$$
\begin{aligned}
\min\left\{C^H, \frac{\varepsilon^2(V^H-C^H)+(1-\varepsilon^2)V^L+\frac{(1-\varepsilon)^2}{2}(V^H-V^L)-(1-\varepsilon)^2p^6}{2\varepsilon(1-\varepsilon)}\right\} &\geq \frac{p^9+p^{10}}{2} \\
\frac{p^9+p^{10}}{2} \geq \max\left\{V^L, \frac{(1-\mu^2)C^H-\mu^2(V^L-C^L)-\frac{(1-\mu^2)}{2}(C^H-C^L)-(1-\mu)^2(p^5)}{2\mu(1-\mu)}\right\} &
\end{aligned}
\tag{1.30}
$$

It is easy to verify that for the price vector described above all incentive compatibility and individual rationality conditions are satisfied. (Example: Suppose that the traders reservation values are $V^H = 8, V^L = 3, C^H = 6, C^L = 0$; and the prior beliefs are uniform $\mu = \varepsilon = \frac{1}{2}$. Let $p^3 = p^8 = 3$, $p^1 = p^7 = 6$, $p^6 = 6$; then $\min\{6, 4\} \geq p^5 \geq \max\{3, -1.5\}$, $\min\{6, 11 - 2p^5\} \geq p^2 \geq 3$ and $\min\{6, 3.75\} \geq \frac{p^9+p^{10}}{2} \geq \max\{3, 3 - .5p^5\}$ and we can set $p^{10} = 3$). ■

Proof of Proposition 8

(Necessity)

Consider the 'upward' incentive compatibility constraints for the seller's type C.

Type C, such that $C \leq 1 - \Delta$, weakly prefers not to report $(C + \Delta)$ whenever the following 'one-step' incentive compatibility constraint holds

$$EU^S(C\,|C) \geq EU^S(C + \Delta|C) \tag{1.31}$$

Or, equivalently, whenever

$$EU^S(C|C) \geq EU^S(C + \Delta|C + \Delta\ ) + (\Delta) \sum_{V=0}^{1} \pi(V, C + \Delta) f^B(V)$$

(In what follows, we will use $EU^S(C)$ in place of $EU^S(C|C)$)

Generally, for any $k \in \{0, 1, .., K\}$, type $(C + k\Delta)$ weakly prefers not to report $(C + (k+1)\Delta)$, whenever

$$EU^S(C + k\Delta) \geq EU^S(C + (k+1)\Delta) + (\Delta) \sum_{V=0}^{1} \pi(V, C + (k+1)\Delta) f^B(V) \tag{1.32}$$

One can verify that the seller's type C would not want to report $(C + k'\Delta)$ for any $k' \in \{0, 1, .., K\}$, as long as for every $k < k'$ the corresponding 'one-step' incentive compatibility constraint (1.32) holds.

Adding all one-step upward incentive compatibility constraints for the seller's type C, we obtain the following inequality

$$EU^S(C) \geq EU^S(1) + \sum_{t'=C+\Delta}^{1} (\Delta) \sum_{V=0}^{1} \pi(V, t') f^B(V) \tag{1.33}$$

The r.h.s of (1.33) is the expected gains from all possible upward deviations for the seller's type C.

Multiplying (1.33) by $f^S(C)$ and adding across all possible types we obtain the aggregate incentive compatibility constraint for the seller.

$$\sum_{C=0}^{1-\Delta} EU^S(C) f^S(C) \geq EU^S(1) + \sum_{V=0}^{1} \sum_{C=0}^{1-\Delta} \left(\sum_{t'=C+\Delta}^{1} (\Delta) \pi(V,t') \right) f^S(C) f^B(V) \quad (1.34)$$

Consider the 'downward' incentive compatibility constraints for the buyer's type V.

Type V, such that $V \geq 0 + \Delta$, prefers not to imitate type $V - \Delta$ whenever

$$EU^B(V|V) \geq EU^B(V'|V) \quad (1.35)$$

Or, equivalently, whenever

$$EU^B(V|V) \geq EU^B(V-\Delta|V-\Delta) + (\Delta) \sum_{C=0}^{1} \pi(V-\Delta, C) f^S(C)$$

(In what follows, we will use $EU^B(V)$ in place of $EU^B(V|V)$)

Generally, for any $k \in \{0,1,..,K\}$, type $(V-k\Delta)$ weakly prefers not to report $(V-(k+1)\Delta)$, whenever

$$EU^B(V-k\Delta) \geq EU^B(V-(k+1)\Delta) + (\Delta) \sum_{C=0}^{1} \pi(V-(k+1)\Delta, C) f^S(C) \quad (1.36)$$

Type V would weakly prefer not to report $(V-k'\Delta)$ for any $k' \in \{0,1,..,K\}$, as long as for every $k < k'$, the corresponding one-step downward incentive compatibility constraint (1.36) holds.

Adding all 'one-step' 'downward' incentive compatibility constraints for the buyer's type V, we have the following inequality

$$EU^B(V) \geq EU^B(0) + \sum_{t=0}^{V-\Delta} (\Delta) \sum_{C=0}^{1} \pi(t, C) f^S(C) \quad (1.37)$$

The r.h.s. of (1.37) is the expected gains from all possible downward deviations for the buyer's type V.

Multiplying (1.37) by $f^B(V)$ and adding across all possible types of the buyer we obtain the following aggregate incentive compatibility constraint for the buyer:

$$\sum_{V=\Delta}^{1} EU^B(V)\, f^B(V) \geq EU^B(0) + \sum_{C=0}^{1}\sum_{V=\Delta}^{1}\left(\sum_{t=0}^{V-\Delta}(\Delta)\,\pi(t,C)\right) f^B(V)\, f^S(C) \quad (1.38)$$

At the optimum the individual rationality constraints for the agents' worst types are binding ($EU^B(0) + EU^S(1) = 0$). Adding up the aggregate incentive compatibility constraints for the buyer and the seller we obtain the following condition:

$$\begin{aligned} &\sum_{V=0}^{1}\sum_{C=0}^{1}(V-C)\,\pi(V,C)\, f^S(C)\, f^B(V) \quad (1.39)\\ \geq\ &\sum_{C=0}^{1}\sum_{V=\Delta}^{1}\left(\sum_{t=0}^{V-\Delta}(\Delta)\,\pi(t,C)\right) f^B(V)\, f^S(C) + \\ &\sum_{V=0}^{1}\sum_{C=0}^{1-\Delta}\left(\sum_{t'=C+\Delta}^{1}(\Delta)\,\pi(V,t')\right) f^S(C)\, f^B(V) \end{aligned}$$

Then, the r.h.s of (1.39) can be rewritten as follows:

$$\begin{aligned} &(\Delta)\sum_{C=0}^{1}\sum_{V=0}^{1-\Delta}\pi(V,C)\left(\sum_{\tau=V+\Delta}^{1} f^B(\tau)\right) f^S(C)\\ &+(\Delta)\sum_{V=0}^{1}\sum_{C=\Delta}^{1}\pi(V,C)\left(\sum_{\tau'=0}^{C-\Delta} f^S(\tau')\right) f^B(V)\\ =\ &(\Delta)\sum_{C=0}^{1}\sum_{V=0}^{1-\Delta}\pi(V,C)\left(1-F^B(V)\right) f^S(C)\\ &+(\Delta)\sum_{V=0}^{1}\sum_{C=\Delta}^{1}\pi(V,C)\left(F^S(C-\Delta)\right) f^B(V) \end{aligned}$$

Finally, we can rewrite (1.39) as follows:

$$\sum_{V=0}^{1}\sum_{C=0}^{1}\left(\left[V-(\Delta)\frac{(1-F^{B}(V))}{f^{B}(V)}\right]-\left[C+(\Delta)\frac{F^{S}(C-\Delta)}{f^{S}(C)}\right]\right)\pi(V,C)\,f^{S}(C)\,f^{B}(V)\geq 0. \blacksquare$$

(Sufficiency)

In order to complete the proof we need to construct a transfer payment function.

Let $\bar{\pi}^{S}(C)=\sum_{V=0}^{1}\pi(V,C)\,f^{B}(V)$ and $\bar{\pi}^{B}(V)=\sum_{C=0}^{1}\pi(V,C)\,f^{S}(C)$. Consider the following discrete type adaptation of the transfer payment function due to Myerson and Satterthwaite (1983):

Let

$$P(V,C)=\begin{cases}\frac{p(V,C)}{\pi(V,C)}, \text{when } \pi(V,C)>0\\ p(V,C), \text{ when } \pi(V,C)=0\end{cases}$$

and

$$\begin{aligned}p(V,C) &= V\bar{\pi}^{B}(V)-(\Delta)\sum_{t=0}^{V-\Delta}\bar{\pi}^{B}(t)+C\bar{\pi}^{S}(C)-(\Delta)\sum_{t'=0}^{C}\bar{\pi}^{S}(t')+\\ &\quad 0*\bar{\pi}^{B}(0)+\sum_{t'=0}^{C}t'\left[1-F^{S}(t')\right]\left[-\left(\bar{\pi}^{S}(t')-\bar{\pi}^{S}(t'-\Delta)\right)\right]\end{aligned}$$

where the last two terms represent a constant such that

$$EU^{B}(0)=0*\bar{\pi}^{B}(0)-\sum_{t'=0}^{\bar{C}}p(\underline{V},t')\,f^{S}(t')=0$$

One can verify that given this transfer payment function, all individual rationality and incentive compatibility constraints are satisfied and all seller's one-step upward incentive compatibility constraints as well as all buyer's one-step downward incentive compatibility constraints are binding. ■

Chapter 2

K-double Auction Mechanisms for Bargaining with Multiple Objects

2.1 Introduction

Consider a bargaining problem where a buyer and a seller negotiate the transfer of a (single) object to be chosen from several alternatives. The traders' valuations of different objects are private. The object selection depends on the prices quoted by the buyer and seller for all available alternatives. An object, for which the spread between the buyer's bid and the seller's offer is the largest, is traded only if this spread is non-negative. The traders bargain according to certain rules, which are commonly known.

The analysis of the bargaining process is restricted to a particular class of trading procedures - the sealed-bid $k-$double auction mechanisms. In the context of bilateral bargaining over the price of a single object, a sealed-bid $k-$double auction is a simple price selection rule that requires both parties to submit their bids for an object in 'sealed envelopes', the envelopes are opened simultaneously and trade occurs if and only if the buyer's bid is above the seller's offer. The price at which the transfer takes place is a convex combination of the prices quoted by the traders, $p = kv + (1 - k)\,c$, where v, c are the buyer's bid, seller's offer, respectively (parameter $k \in [0, 1]$ is commonly known). Despite their simplicity, the sealed-bid $k-$double auctions capture important features of the bargaining process and, therefore, can provide valuable insights into more realistic and complex bargaining procedures.

Because in a typical market environment[1], the outcome of any k−double auction is not *ex post* efficient[2], the k− double auction literature mostly focused on the analysis of the properties of the 'second-best' outcomes, i.e. the **ex-ante efficiency ratios**[3] (to see what proportion of potential gains from trade could be realized in a particular k-double auction) and the **allocative efficiency** (to see how gains from trade could be divided among the traders in k-double auctions).

The research on k−double auctions with a single object showed that for generic priors the *ex ante* efficient equilibria in $k \in (0,1)$ double auctions did not exist and that only the seller's offer and buyer's bid auctions ($k = 0$ and $k = 1$) were *ex ante* efficient (see Satterthwaite and Williams (1989)). For the special case of uniformly distributed reservation values, Chatterjee and Samuelson (1983), characterized a linear equilibrium of the 'split-the-difference' ($k = 0.5$) double auction, which allowed the traders to achieve maximal total *ex ante* expected gains from trade, given incentive compatibility, individual rationality and balanced budget constraints (the latter was proven later by Myerson and Satterthwaite (1983)). Thus, equal bargaining power ($k = 0.5$) was shown to lead to a smaller efficiency loss.

The literature on bargaining with 'multiple units and dimensions' under two-sided uncertainty is more scarce (see, for example, Linhart and Radner (1989)), because sufficiently realistic trading mechanisms for bargaining with multiple objects tend to be quite complicated.

In the problem of bargaining with multiple objects, the first best outcome is achieved when the object of trade is always such that the spread between the buyer's valuation and

[1]The term 'typical', here, refers to a market environment, where (i) the existence of positive gains from trade is not known to the traders *ex ante*; (ii) the parties can refuse to participate in the bargaining process after they receive their private information.

[2]See Myerson and Satterthwaite (1983).

[3]Here, the *ex ante* efficiency ratio refers to the ratio of the maximal *ex ante* expected gains from trade that could be achieved in a Bayesian equilibrium of a k− double auction to the maximal potential *ex ante* expected gains from trade.

the seller's cost is maximal and non-negative. Depending on the information structure and the choice of the trading mechanism, the failure to achieve the first best outcome may be either due to trading an *ex post* inefficient object or due to not trading at all at some states where positive gains from trade exist. In Chapter 1, I proposed to distinguish between **fully ex post efficient** trading mechanisms and the mechanisms that are **ex post efficient with respect to the object selection** only. The latter guarantee the *ex post* efficient object selection at any state where trade occurs, but do not guarantee that the players always trade when gains from trade exist. In this Chapter, we will address the question of **whether/when the $k-$double auction mechanisms can be ex post efficient with respect to the object selection given different assumptions about the information structure of the bargaining problem.**

It is a well known fact that when uncertainty about the traders' valuations is two-sided, the *ex post* efficient outcome is generally infeasible and that inefficiency arises because the traders attempt to use their private information to increase their personal shares in total gains from trade. The main advantage of working with a $k-$ double auction framework is that it allows to see how the players bids are distorted away from their true valuations and determine **the type of distortions** that lead to *ex post* inefficient object selection.

In what follows, the analysis will focus on the choice between **two** potential objects of trade, denoted q_1 and q_2.[4]

The bargaining process is modeled as a game of incomplete information, Γ, which proceeds as follows:

Stage 1: both traders simultaneously submit their bids for all potential objects of trade. Given the players' bids, the price for each potential object of trade is set according to a k-double auction price selection rule.

Stage 2: given the players' bids for each object q_i, where $i \in \{1,2\}$, an object with the largest spread between the buyer's bid and the seller's offer is selected, provided that the spread is non-negative, and there is no trade, otherwise.

[4]The generalization to an arbitrary number of objects will be presented whenever possible.

The games $\Gamma 1, \Gamma 2, \Gamma 3$ and $\Gamma 4$ differ with respect to the information structure of the basic bargaining game Γ described above. The rules of the game as well as the value of parameter k are common knowledge.

Game	Valuations	Information structure
$\Gamma 1$	$V_1, V_2 \overset{i.i.d.}{\sim} U(0,1)$	All valuations are private and the probability
	$C_1, C_2 \overset{i.i.d}{\sim} U(0,1)$	distributions are commonly known.
$\Gamma 2$	$C_1 = 0, V_2 = 1$	The seller's cost of q_1 and the buyer's value of q_2
	$C_2, V_1 \overset{i.i.d.}{\sim} U(0,1)$	are commonly known; other valuations are private and
		the probability distributions are commonly known.
$\Gamma 3$	$V_1 = V_2 = V \sim U(0,1)$	All valuations are private and the probability
	$C_1, C_2 \overset{i.i.d.}{\sim} U(0,1)$	distributions are commonly known
$\Gamma 4$	$C_1, V_1 \in [0,1], V_1 > C_1$	Valuations for q_1 are commonly known;
	$C_2, V_2 \overset{i.i.d}{\sim} U(0,1)$	valuations for q_2 are private and the distributions
		of the traders' valuations for q_2 are commonly known

If both traders were to quote their true valuations, then the price of any potential object of trade in a k−double auction would be equal to $kV_i + (1-k)C_i$ and the outcome of the bargaining process would be fully *ex post* efficient. However, in most cases, the traders prefer to misreport their private information. The following is an overview of the main results:

In $\Gamma 1$, none of the traders knows *ex ante* whether gains from trade exist and which object is *ex post* efficient. The analysis of $\Gamma 1$ shows that when the optimal choice of the object of trade depends on both players' private information, the most common unmediated trading mechanism (where one party proposes a price schedule ($k = 1$ or $k = 0$) and the other either picks an object of trade or chooses not to trade) does not guarantee an *ex post* efficient object selection.

In Γ2, the existence of gains from trade is commonly known, but the *ex post* efficient object is not known *ex ante*. Not surprisingly, the outcome of the bargaining process is not fully *ex post* efficient: although the players always trade, they do not always select the best object. Interestingly, compared to $k = 1$ or $k = 0$ auction mechanisms, equal distribution of the bargaining power ($k = 0.5$) reduces the range of transactions where the *ex post* inefficient object is traded, and thus, results in a relatively smaller efficiency loss.

In Γ3, it is commonly known that the buyer is indifferent between two trading alternatives, which implies that the low-cost object is *ex ante* commonly known to be the *ex post* efficient choice. However, none of the players knows *ex ante* whether positive gains from trade exist. The analysis of Γ3 suggests that for any $k \in \{0, 0.5, 1\}$, the outcome is *ex post* efficient with respect to the object selection, although it is not fully *ex post* efficient. This result is easily generalizable to an arbitrary number of potential objects of trade.

In Γ4, there is no uncertainty about the existence of mutually beneficial trade, since the players' valuations for q_1 are common knowledge. In Γ4, trade always occurs, but the *ex post* efficient object is not always selected for any $k \in \{0, 0.5, 1\}$.

The *ex ante* efficiency ratio tends to increase when we add a second potential object of trade.

2.2 The Model

Suppose that there are two potential objects of trade, $\{q_1, q_2\}$. Each trader's type is given by a pair of reservation values. We will assume that each player's valuation for any potential object of trade is a random variable distributed on the interval $[0, 1]$. Denote by $\mathbf{V} \in [0, 1]^2$, the buyer's type, where $\mathbf{V} = (V_1, V_2)$, and by $\mathbf{C} \in [0, 1]^2$, the seller's type, where $\mathbf{C} = (C_1, C_2)$.

The bargaining process is modelled as a two-stage game:

Stage 1: the traders simultaneously submit their bids for all potential objects of trade (v_i denotes the buyer's bid for q_i and c_i denotes the seller's bid for q_i). Given the players' bids, the price schedule is determined by a k-double auction rule, where for any $i \in \{1,2\}$, such that $c_i \leq v_i$, $p_i = kv_i + (1-k)\, c_i$.

Stage 2: given the players' bids v_i, c_i, for $i \in \{1,2\}$, the object selection procedure γ is as follows:

$$\begin{aligned}
\gamma(c_1, c_2, v_1, v_2) &= q_1, \text{ iff } v_1 - c_1 > v_2 - c_2 \text{ and } v_1 \geq c_1 \\
\gamma(c_1, c_2, v_1, v_2) &= q_2, \text{ iff } v_1 - c_1 < v_2 - c_2 \text{ and } v_2 \geq c_2 \\
\gamma(c_1, c_2, v_1, v_2) &= \emptyset, \text{ iff } v_i - c_i < 0, \text{ for } i \in \{1,2\}
\end{aligned}$$

where $\emptyset$ denotes the 'no-trade' option.[5] We will assume that whenever $v_1 - c_1 = v_2 - c_2$ and $v_i \geq c_i$, for $i \in \{1,2\}$, the object of trade is selected at random. The object selection procedure γ is a mapping from the players' bids into $Q \cup \{\emptyset\}$.

Whenever trade occurs and q_i is selected, the buyer's profit is $(V_i - p_i)$ and the seller's profit is $(p_i - C_i)$.

The player's bidding strategy maps his type into a pair of bids. Denote by $\overline{\beta} : [0,1]^2 \rightarrow [0,1]^2$ the buyer's bidding strategy, and by $\overline{\alpha} : [0,1]^2 \rightarrow [0,1]^2$, the seller's bidding strategy. For $i \in \{1,2\}$, define the buyer's bidding function for an object q_i, $\beta_i : [0,1]^2 \rightarrow [0,1]$, where $v_i = \beta_i(\mathrm{V})$ is the buyer's type V bid for q_i, and the seller's bidding function for an object q_i, $\alpha_i : [0,1]^2 \rightarrow [0,1]$, where $c_i = \alpha_i(\mathrm{C})$ is the seller's type C bid for q_i.

In what follows, we will focus on bidding strategies with the following properties: for at least one object q_i, $i \in \{1,2\}$,

(1) β_i, α_i are twice continuously differentiable on $[0,1]^2$;

[5] Note, that in Chapter 1 the 'no-trade' options was denoted by q_0.

(2) β_i is strictly increasing in V_i; α_i is strictly increasing in C_i;

(3) $0 \le \beta_i(\mathbf{V}) \le V_i$ for any $\mathbf{V} \in [0,1]^2$; $C_i \le \alpha_i(\mathbf{C}) \le 1$ for any $\mathbf{C} \in [0,1]^2$.

(1) and (2) are technical assumptions that we need to ensure the existence of a well-behaved solution of each player's optimization problem; (3) is equivalent to the requirement that the players' optimal bids must satisfy the interim individual rationality condition. (Note that assumptions (1)-(3) do not preclude the possibility that the buyer can optimally bid $-\infty$ for some q_i for all realizations of his type and, similarly, that the seller can optimally bid $+\infty$ for some q_i for all realizations of his type).

2.3 The Game Γ1

2.3.1 The Buyer's Bid Double Auction $\Gamma1\backslash k = 1$

Suppose that $V_1, V_2 \overset{i.i.d.}{\sim} U(0,1)$ and $C_1, C_2 \overset{i.i.d.}{\sim} U(0,1)$. Each player's type is his private information and only probability distributions of the players' valuations are commonly known. Also, suppose that the buyer proposes the prices for all potential objects of trade.

The bargaining game proceeds as follows:

Stage 1: the buyer announces v_1, v_2 and the seller announces c_1, c_2. The price schedule is determined according to the $k = 1$ double auction price selection rule, i.e. $p_i = v_i$ for any $i \in \{1,2\}$.

Stage 2: given the players' bids, the object of trade is chosen according to the object selection procedure γ described in Section 2.2.

Without loss of generality, we can think of a game, where the buyer proposes the price schedule and the seller either picks an object of trade or chooses not to trade. Note that, when both players' valuations for all potential objects of trade are private, the trading decision (whether to trade or not) and the choice of the object of trade depends on both players' private information about q_1 and q_2 in a non-trivial way.

We will begin by giving a general description of *Bayesian-Nash Equilibrium* of $\Gamma 1\backslash k=1$.

The seller's type C **optimal trading strategy**, given the buyer's bids v_1, v_2, is as follows:

$$\begin{aligned}
\gamma^*(\mathrm{C}, v_1, v_2) &= q_1, \text{ iff } C_2 - C_1 > v_2 - v_1 \text{ and } v_1 \geq C_1 \\
\gamma^*(\mathrm{C}, v_1, v_2) &= q_2, \text{ iff } C_2 - C_1 < v_2 - v_1 \text{ and } v_2 \geq C_2 \\
\gamma^*(\mathrm{C}, v_1, v_2) &= \emptyset, \text{ iff } v_i - C_i < 0, \text{ for } i \in \{1,2\}
\end{aligned}$$

Whenever $v_1 - c_1 = v_2 - c_2$ and $v_i \geq c_i$, for $i \in \{1,2\}$, the object of trade is selected at random.

Note that the buyer's bid v_i is relevant only when q_i is selected by the seller. So, the buyer's bidding strategy $\overline{\beta}$ is a **best response** to the seller's trading strategy γ^* if for all $\mathbf{V} \in [0,1]^2$, the buyer's bid for an object q_i, $v_i = \beta_i(\mathbf{V})$, maximizes the buyer's type $\mathbf{V}$ expected gains from trade, given that object q_i is selected by the seller according to the trading strategy γ^*. Whenever this is the case, the strategy pair $\left(\overline{\beta}, \gamma^*\right)$ constitutes a *Bayesian-Nash Equilibrium* of $\Gamma 1\backslash k=1$.

The derivation of the buyer's best response bidding strategy.

The buyer's type V gains from trading q_i, given that trade occurs and q_i is selected by the seller, is $\Pi_i^B = (V_i - v_i)$.

Then, the buyer's type V expected gains from trade are

$$E\Pi^B(v_2, v_1|\mathbf{V}) = \Pr\{q_1 \text{ is selected } |v_2, v_1\}\,\Pi_1^B + \Pr\{q_2 \text{ is selected } |v_2, v_1\}\,\Pi_2^B$$

In order to derive the buyer's bidding strategy, we partition the space of all action configurations of the buyer, such that $v_1 \in (0,1)$ and $v_2 \in (0,1)$, into two regions:

REGION I: where $v_2 \geq v_1$ and, therefore, the probabilities of selecting q_1 and q_2 are[6]

$$\begin{aligned}
\Pr\{q_1 \text{ is selected } |v_2, v_1\} &= v_1 - v_2 v_1 + \frac{1}{2} v_1^2 \\
\Pr\{q_2 \text{ is selected } |v_2, v_1\} &= v_2 - \frac{(v_1)^2}{2}
\end{aligned}$$

REGION II: where $v_2 < v_1$ and, therefore, the probabilities of selecting q_1 and q_2 are

$$\begin{aligned}
\Pr\{q_1 \text{ is selected } |v_2, v_1\} &= v_1 - \frac{(v_2)^2}{2} \\
\Pr\{q_2 \text{ is selected } |v_2, v_1\} &= v_2 - v_2 v_1 + \frac{1}{2} v_2^2
\end{aligned}$$

Then, the buyer's type **V** expected utility is

$$E\Pi^B(v_1, v_2|\mathbf{V}) = \begin{cases} (V_1 - v_1)\left(v_1 - v_2 v_1 + \frac{1}{2} v_1^2\right) + (V_2 - v_2)\left(v_2 - \frac{(v_1)^2}{2}\right), & \text{when } v_2 \geq v_1 \\ (V_1 - v_1)\left(v_1 - \frac{(v_2)^2}{2}\right) + (V_2 - v_2)\left(v_2 - v_2 v_1 + \frac{1}{2} v_2^2\right), & \text{when } v_2 < v_1 \end{cases}$$

In REGION I, where $v_2 \geq v_1$, the f.o.c.'s of the buyer's maximization problem are as follows:

$$\frac{\partial E\Pi^B}{\partial v_1} = 0 \Leftrightarrow -v_1 + v_2 v_1 - \frac{1}{2} v_1^2 + (V_1 - v_1)(1 - v_2 + v_1) - (V_2 - v_2) v_1 = 0$$

$$\frac{\partial E\Pi^B}{\partial v_2} = 0 \Leftrightarrow -(V_1 - v_1) v_1 + \frac{1}{2} v_1^2 - 2v_2 + V_2 = 0$$

The buyer's optimal bidding functions are given by the solution of (2.1)

[6] For details, see Appendix.

$$\begin{cases} -v_1 + v_2 v_1 - \frac{1}{2}v_1^2 + (V_1 - v_1)(1 - v_2 + v_1) - (V_2 - v_2)v_1 = 0 \\ -(V_1 - v_1)v_1 + \frac{1}{2}v_1^2 - 2v_2 + V_2 = 0 \end{cases} \tag{2.1}$$

and are as follows:

$$v_2 = \frac{1}{2}V_2 - \frac{1}{2}V_1 v_1 + \frac{3}{4}v_1^2$$

and

$$v_1 = \rho$$

where ρ is a root of the polynomial (2.2)

$$9Z^3 + (-9V_1 - 6)Z^2 + \left(-8 + 2V_2 + 4V_1 + 2V_1^2\right)Z - 2V_2V_1 + 4V_1 \tag{2.2}$$

In REGION II, the solution is symmetric:
$v_1 = -\frac{1}{2}V_2 v_2 + \frac{3}{4}v_2^2 + \frac{1}{2}V_1$ and $v_2 = \rho$,where ρ is a root of the following polynomial

$$9Z^3 + (-9V_2 - 6)Z^2 + \left(-8 + 2V_1 + 4V_2 + 2V_2^2\right)Z - 2V_2V_1 + 4V_2$$

Lemma 11 states that the buyer's best response bid for the high-value object is always greater than or equal to his best response bid for the low-value object.[7]

Lemma 11 *The buyer's best response bids are such that $v_2^* \geq v_1^*$ if and only if $V_2 \geq V_1$.*

The plot of the buyer's best response bidding function for q_1 (see Figure 1) corresponds to the numerical solution of the following system equation, assuming that $v_2 \geq v_1$ and (1) – (3) described in Section 2.2 hold,

[7]For the proof, see Appendix.

$$\begin{cases} 9v_1^3 + (-9V_1 - 6)\,v_1^2 + (-8 + 2V_2 + 4V_1 + 2V_1^2)\,v_1 - 2V_2V_1 + 4V_1 = 0 \\ V_1 \in (0,1) \\ V_2 \in (0,1) \\ v_1 \in (0,1) \end{cases}$$

Note that the bidding function in Figure 1 is defined only for the region where $V_2 \geq V_1$.

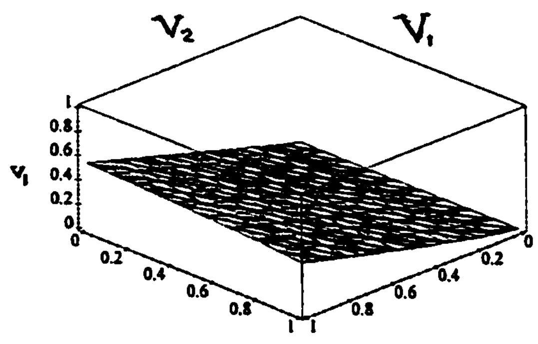

Figure 1: The best response bidding function of the buyer for the low-value object.

In order to characterize the *Bayesian-Nash Equilibrium* of $\Gamma 1 \backslash k = 1$, define $V_H \equiv \max\{V_1, V_2\}$ and $V_L \equiv \min\{V_1, V_2\}$. The buyer's bids for the high-value object, $v_H = \beta_H(\mathbf{V})$, and the low-value object, $v_L = \beta_L(\mathbf{V})$, are given by the solution of (2.1) (where v_H corresponds to v_2 and v_L corresponds to v_1), subject to the restrictions that $v_H \in (0,1)$ and $v_L \in (0,1)$. The bidding functions $\beta_H(\mathbf{V}), \beta_L(\mathbf{V})$ together with the trading strategy γ^* form a *Bayesian-Nash Equilibrium* of the game $\Gamma 1 \backslash k = 1$ only if there is no alternative bidding strategy that yields a strictly better payoff for the buyer. Suppose that the buyer deliberately precludes the possibility of trading the low-value object by bidding $\tilde{\beta}_L(\mathbf{V}) \leq 0$ for any realization of his type $\mathbf{V}$. Then, since there is only one feasible trade left (the high-value object), the buyer's optimal bid for the high-value object is $v_H = \frac{1}{2}V_H$. One can verify that every type $\mathbf{V}$ of the buyer can do strictly better if he submits the bids $v_H = \frac{1}{2}V_H$ and $v_L \in (0,1)$ instead. Thus, the buyer's optimal bidding strategy configuration must be such that $v_H \in (0,1)$ and $v_L \in (0,1)$. The best response bidding functions $\beta_H(\mathbf{V}), \beta_L(\mathbf{V})$, restricted to take the values in the $(0,1)$ interval can be calculated numerically.[8]

Table 1 presents the numerical values of the best response bids for some types of the buyer.

Table 1

V_1	V_2	v_1	v_2	$V_2 - V_1$	$v_2 - v_1$	$\frac{v_2 - v_1}{V_2 - V_1}$
1	1	.4227	.4227	0	0	...
.8	1	.3279	.4495	.2	.1216	.6081
.5	1	.1927	.4797	.5	.2870	.5739
.2	1	.0711	.4967	.8	.4256	.5320
0	1	0	.5	1	.5	.5

It is worth noting that the difference between the buyer's bids is always smaller than the difference between his valuations and the ratio $\frac{v_2 - v_1}{V_2 - V_1}$ tends to increase when $(V_2 - V_1)$ becomes smaller, which implies that the players are more likely to end up trading the *ex post* inefficient object when the difference between the buyer's valuations is relatively small.

Proposition 12 summarizes the analysis presented above.

Proposition 12 *The Bayesian-Nash Equilibrium outcome of the game* $\Gamma 1 \backslash k = 1$ *is not always ex post efficient with respect to the object selection.*

---

[8] For details, see Appendix.

In order to show that the auction mechanism in $\Gamma 1 \backslash k=1$ is not *ex post* efficient with respect to the object selection, it is sufficient to present a numerical example. Consider the players' types $V_1 = 0.5, V_2 = 1$ and $C_1 = 0.1, C_2 = 0.4$. Then, the buyer's optimal bids are $v_2 = .47968$, $v_1 = .19271$. Since, $V_1 - C_1 = 0.4$ and $V_2 - C_2 = 0.6$, q_2 is the *ex post* efficient choice, but, because $v_1 - C_1 = .09271$ and $v_2 - C_2 = .07968$, the seller chooses q_1.

Thus, the most common unmediated trading mechanism (where one party proposes the price schedule and the other party chooses the object of trade) does not guarantee the *ex post* efficient object selection when the trading decision and the object selection depend on both players' private information in a non-trivial way.

2.3.2 The Split-the-Difference Double Auction $\Gamma 1 \backslash k = 0.5$

Suppose that the traders have equal bargaining power. The sequence of events is the same as in $\Gamma \backslash k = 1$.

Stage 1: the buyer announces v_1, v_2 and the seller announces c_1, c_2. The price schedule is determined according to the $k = 0.5$ double auction price selection rule, i.e. for any $i \in \{1,2\}$, such that $v_i \geq c_i$, $p_i = (v_i + c_i)/2$.

Stage 2: given the players' bids v_1, v_2, c_1, c_2, the object selection rule is as follows:

$$\begin{aligned}
\gamma(v_1, v_2, c_1, c_2) &= q_1, \text{ iff } v_1 - c_1 > v_2 - c_2 \text{ and } v_1 \geq c_1 \\
\gamma(v_1, v_2, c_1, c_2) &= q_2, \text{ iff } v_1 - c_1 < v_2 - c_2 \text{ and } v_2 \geq c_2 \\
\gamma(v_1, v_2, c_1, c_2) &= \emptyset, \text{ iff } v_i < c_i, \text{ for } i \in \{1,2\}
\end{aligned}$$

Whenever $v_1 - c_1 = v_2 - c_2$ and $v_i \geq c_i$, for $i \in \{1,2\}$, the object of trade is selected at random.

The general description of the *Bayesian-Nash Equilibrium* of the Split-the-Difference Double Auction game is as follows.

The players choose their bids for an object q_i, conditional on the event that q_i is selected in the final stage of the game. Therefore, the buyer's bidding strategy $\overline{\beta}$ is a best response to the seller's bidding strategy $\overline{\alpha}$, if for all $\mathbf{V} \in [0,1]^2$, the buyer's bid for the object q_i, $v_i = \beta_i(\mathbf{V})$, maximizes the buyer's type $\mathbf{V}$ expected gains from trade given that (*i*) the seller's bidding strategy is $\overline{\alpha}$; (*ii*) the seller's valuation of q_i has conditional distribution $\tilde{F}_i(C_i)$, given that q_i is selected according to γ. And, the seller's bidding strategy $\overline{\alpha}$ is a best response to the buyer's bidding strategy $\bar{\beta}$, if for all $\mathbf{C} \in [0,1]^2$, the seller's bid for the object q_i, $c_i = \alpha_i(\mathbf{C})$, maximizes type $\mathbf{C}$ seller's expected gains from trade given that (*i*) the buyer's bidding strategy is $\overline{\beta}$; (*ii*) the buyer's valuation of q_i has conditional distribution $\tilde{G}_i(V_i)$, given that q_i is selected according to γ. Whenever this is the case, a pair $\left(\overline{\alpha},\overline{\beta}\right)$ constitutes a *Bayesian-Nash Equilibrium* of $\Gamma\backslash k = 0.5$.

Unfortunately, it is not possible to obtain a closed form solution of the game $\Gamma 1\backslash k = 0.5$. The characterization of the *Bayesian-Nash Equilibrium* of a simpler split-the-difference double auction $\Gamma 2\backslash k = 0.5$ (where uncertainty is one-dimensional, given that $V_2 = 1$ and $C_1 = 0$ are commonly known) will be presented in the next section.

2.4 The Game $\Gamma 2$

2.4.1 The Buyer's Bid Double Auction $\Gamma 2\backslash k = 1$

Suppose that $V_2 = 1$ and $C_1 = 0$ are commonly known, while V_1 and C_2 are distributed uniformly on the $[0,1]$ interval independently of each other and, consider the game in which the buyer proposes the price schedule. Although there is no uncertainty about the existence of positive gains from trade, the optimal object selection is not known *ex ante*.

Consider the game $\Gamma 2\backslash k = 1$, where $V_2 = 1$ and $C_1 = 0$ are commonly known, V_1 and C_2 are distributed uniformly on $[0,1]$ independently of each other. There exists a *Bayesian-Nash Equilibrium* of $\Gamma 2\backslash k = 1$, such that for any $C_2 \in [0,1], V_1 \in [0,1]$

$$\begin{cases} \alpha_1^*(\mathbf{C}) = 0 \\ \alpha_2^*(\mathbf{C}) = C_2 \end{cases} \tag{2.3}$$

$$\begin{cases} \beta_1^*(V) = 0 \\ \beta_2^*(V) = \frac{1}{2} - \frac{1}{2}V_1 \end{cases} \tag{2.4}$$

In the equilibrium described above, the players always trade, but the equilibrium outcome is not fully *ex post* efficient, because there is a range of transactions where the *ex post* inefficient object is traded. See Figure 2 below.

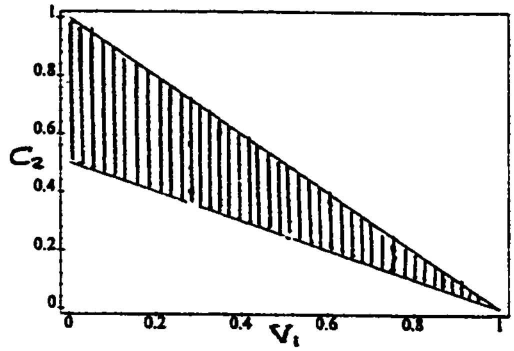

Figure 2: The range of ex post inefficient transactions in the linear equilibrium of the game $\Gamma 2 \backslash k = 1$.

Thus, whenever the optimal object selection depends on both players' private information in a non-trivial way, the equilibrium in which trade always occurs may not be fully *ex post* efficient.

2.4.2 The Split-the-Difference Double Auction $\Gamma 2 \backslash k = 0.5$

Consider the game $\Gamma 2 \backslash k = 0.5$, where the players have equal bargaining power. Proposition 13 describes the linear equilibria of the game $\Gamma 2 \backslash k = 0.5$, in which each object is traded with positive probability.

Proposition 13 *Consider the game* $\Gamma 2 \backslash k = 0.5$, *where* $V_2 = 1$ *and* $C_1 = 0$ *are commonly known,* V_1 *and* C_2 *are distributed uniformly on* $[0, 1]$ *independently of each other.*

(1) The equilibria, in which the seller's bidding functions $\alpha_1(\mathrm{C})$ and $\alpha_2(\mathrm{C})$ depend on C_2 in a non-trivial way and/or the buyer's bidding functions $\beta_1(\mathrm{V})$ and $\beta_2(\mathrm{V})$ depend on V_1 in a non-trivial way, do not exist.

(2) There exists a linear Bayesian-Nash Equilibrium such that for any $C_2 \in [0,1], V_1 \in [0,1]$ $\alpha_1^*(\mathrm{C}) = \beta_1^*(\mathrm{V}) = 0$ and

$$\begin{cases} \alpha_1^*(\mathrm{C}) = 0 \\ \alpha_2^*(\mathrm{C}) = \frac{2}{3}C_2 + \frac{1}{4} \end{cases} \tag{2.5}$$

$$\begin{cases} \beta_1^*(\mathrm{V}) = 0 \\ \beta_2^*(\mathrm{V}) = \frac{3}{4} - \frac{2}{3}V_1 \end{cases} \tag{2.6}$$

(Similarly, one can show that there is a linear equilibrium, where $\alpha_2^*(\mathrm{C}) = \beta_2^*(\mathrm{V}) = 1)$[9]

In order to see that the players' bidding strategies (2.5), (2.6) constitute an equilibrium, it is important to understand the trade-offs that are involved in determining the optimal bids. Suppose that the seller follows the strategy described in Proposition 13 and consider the buyer's problem. Since $\alpha_1^*(\mathrm{C}) = 0$, the buyer can set his best response bid v_1 as close to 0 as possible. When $v_1 = 0$, the buyer captures maximal gains from trading q_1, provided that q_1 is selected, which would then be the case if and only if $v_2 < c_2$. In order to guarantee that q_1 is always selected, the buyer could bid $v_2 = -\infty$ for any realization of his type, but the lower V_1 the higher the buyer's relative (expected) gains from trading q_2 compared to the gains from trading q_1.

Note that in equilibrium described in Proposition 13, q_2 is never traded when $V_1 > \frac{3}{4}$. Consider a buyer, whose valuation for q_1 is $V_1 \leq \frac{3}{4}$. The l.h.s. of (2.7) is the buyer's expected gains from trade in the linear equilibrium described in Proposition 12 and the r.h.s. of (2.7) is the buyer's expected gains from always trading q_1 for all realizations of the seller's type (i.e from bidding $v_2 = -\infty$ for any value of V_2):

[9] The proof is in Appendix.

$$\int_{0}^{\frac{3}{4}-V_1}\left(1-\frac{1-\frac{2}{3}V_1+\frac{2}{3}t_2}{2}\right)dt_2+\int_{\frac{3}{4}-V_1}^{1}(V_1)\,dt_2 \geq V_1 \tag{2.7}$$

Inequality (2.7) holds if and only if

$$\frac{9}{32}+\frac{1}{2}V_1^2+\frac{1}{4}V_1 \geq V_1 \tag{2.8}$$

which is true for all $V_1 \in \left[0, \frac{3}{4}\right]$. Thus, the buyer's strategy (2.6) is the best response to the seller's strategy (2.5).

Since q_2 is the *ex post* efficient choice if and only if $1 - C_2 \geq V_1$, it implies that the *ex post* efficient object is not always traded in equilibrium described in Proposition 13. Thus, although the players always trade, the trading mechanism $\Gamma 2 \backslash k = 0.5$ is not fully *ex post* efficient .

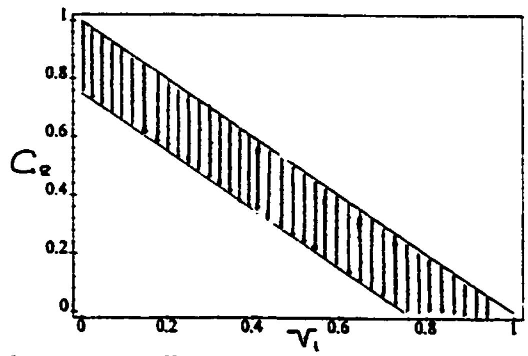

Figure 3: The range of ex post inefficient transactions in the linear equilibrium of the game $\Gamma 1 \backslash k = 0.5$

The comparison of the total *ex ante* expected gains from trade for the linear equilibria of the games $\Gamma 1 \backslash k = 1$ and $\Gamma 1 \backslash k = 0.5$ (see Section 2.7) suggests that the *ex post* efficiency loss in $k = 0.5$ double auction is smaller than in $k = 1$ double auction. Thus, a more equal distribution of the bargaining power reduces the range of *ex post* inefficient transactions, i.e. the ones where object selection is not *ex post* efficient.

2.5 The Game Γ3

2.5.1 The Buyer's Bid Auction $\Gamma 3 \backslash k = 1$

In Γ3, the trading decision (whether to trade or not) depends on both players' private information about q_1 and q_2. However, the *ex post* efficient object selection, conditional on the event that trade occurs, depends only on the seller's private information and the low cost object is known to be the best choice *ex ante*.

Suppose that $V_1 = V_2 = V \sim U(0,1)$ is common knowledge. And let's assume that the buyer submits the same bid for both objects, i.e. $v_1 = v_2 = v$. Then, the seller's type C **optimal trading strategy**, given the buyer's bid(s), is as follows:

$$\begin{aligned}
\gamma^*(C,v) &= q_1, \text{ iff } C_2 < C_1 \text{ and } v \geq C_1 \\
\gamma^*(C,v) &= q_2, \text{ iff } C_2 > C_1 \text{ and } v \geq C_2 \\
\gamma^*(C,v) &= \emptyset, \text{ iff } v < C_L, \text{ where } C_L = \min\{C_1, C_2\}
\end{aligned}$$

Whenever $v_1 - C_1 = v_2 - C_2$ and $v_i \geq C_i$, for $i \in \{1,2\}$, the object of trade is selected at random.

The seller always chooses the low-cost object, and given that $V_1 = V_2 = V$, it implies that he always chooses an object, which maximizes the difference between the players valuations.

Denote by F_L the distribution function of C_L, where $F_L(t) = 2t - t^2$.

The buyer's type **V** expected profit is

$$E\Pi^B(v|\mathbf{V}) = (V - v)\left[\Pr\{C_1 \leq v, C_1 \leq C_2\} + \Pr\{C_2 \leq v, C_2 \leq C_1\}\right]$$

$$
\begin{aligned}
&= (V-v)\left[\frac{1}{2}\Pr\{C_1 \le v|\, C_1 = C_L\} + \frac{1}{2}\Pr\{C_2 \le v|\, C_2 = C_L\}\right] \\
&= (V-v)\Pr\{C_L \le v\} \\
&= (V-v)\int_0^v (2(1-x))\,dx \\
&= (V-v)\left(2v - v^2\right)
\end{aligned}
$$

Solving the buyer's optimization problem we obtain the buyer's best response bidding strategy, given the seller's trading strategy γ^* and our assumption that $v_1 = v_2 = \beta^*(V)$:

$$
\forall i \in \{1,2\}, \forall V \in [0,1], \beta_i^*(V) = \frac{1}{3}V + \frac{2}{3} - \frac{1}{3}\sqrt{(V^2 - 2V + 4)} \tag{2.9}
$$

Proposition 14 *The game* $\Gamma 3\backslash k = 1$ *has a unique Bayesian-Nash Equilibrium, where the seller always chooses the low-cost object, provided that the buyer's bid is above the seller's cost, and the buyer uses the following optimal bidding strategy:* $\forall i \in \{1,2\}, \forall V \in [0,1], \beta_i^*(V) = \frac{1}{3}V + \frac{2}{3} - \frac{1}{3}\sqrt{(V^2 - 2V + 4)}$.[10]

As the number of trading alternatives increases, it is optimal for the buyer to bid uniformly more aggressively for all objects. Consider the general case when the number of potential objects of trade is N. The buyer's best response bid v^* is implicitly given by the following expression:

$$
V - v^* = \frac{1 - (1 - v^*)^N}{N(1 - v^*)^{N-1}}
$$

where $F_L(x) = 1 - (1 - x)^N$ is the distribution function of $C_L \equiv \min\{C_1, C_2, ..., C_N\}$.

Figure 4 shows the buyer's best response bidding functions for $N = 1, N = 2, N = 3$. (Recall that for $N = 1$, $v^* = \frac{1}{2}V$)

[10] *The proof is in Appendix.*

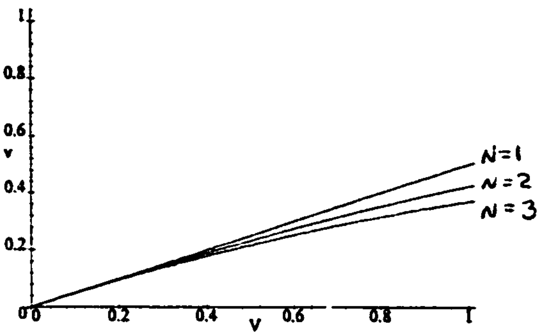

Figure 4: The best response bidding function of the buyer for N=1, N=2, N=3 in the game $\Gamma3\backslash k = 1$.

If Q is a continuum and the seller's costs for all objects are i.i.d $U(0,1)$, then the buyer knows for sure that there is an object for which the seller's cost is equal to zero and it is, therefore, optimal for the buyer to bid $v = 0$ for all potential objects of trade. So, when the number of potential objects of trade increases, the uncertainty about whether trade occurs or not decreases and so does the seller's share of the total gains from trade.

Whenever the buyer is indifferent between q_1 and q_2, the seller always chooses an *ex-post* efficient object for any number of potential objects of trade.

In fact, it is easy to show that, given the same information structure ($V_1 = V_2 = V \sim U(0,1)$ and $C_1, C_2 \overset{i.i.d}{\sim} U(0,1)$), the seller's bid double auction $\Gamma3\backslash k = 0$ is also *ex post* efficient with the respect to the object selection.

2.5.2 The Split-the-Difference Double Auction $\Gamma3\backslash k = 0.5$

Suppose that the traders have equal bargaining power and $V_1 = V_2 = V \sim U(0,1)$ is common knowledge. Let's assume that the buyer submits the same bid for both objects, i.e. $v_1 = v_2 = v$.

If the gains from trade are positive, the object selection rule is to trade an object for which the seller submits the lowest bid, i.e.

$$\begin{aligned}
\gamma(c_1, c_2, v) &= q_1, \text{ iff } c_2 < c_1 \text{ and } v \geq c_1 \\
\gamma(c_1, c_2, v) &= q_2, \text{ iff } c_2 > c_1 \text{ and } v \geq c_2 \\
\gamma(c_1, c_2, v) &= \emptyset, \text{ iff } v < c_L, \text{ where } c_L = \min\{c_1, c_2\}
\end{aligned}$$

Whenever $v_1 - c_1 = v_2 - c_2$ and $v_i \geq c_i$, for $i \in \{1, 2\}$, the object of trade is selected at random.

Let $\alpha_L(\mathrm{C})$ denote the seller's bidding strategy for the low-cost object and let $\alpha_H(\mathrm{C})$ denote the seller's bidding strategy for the high-cost object. If for any $\mathrm{C} \in [0,1]^2$, $\alpha_L(\mathrm{C}) \leq \alpha_H(\mathrm{C})$, then the high-cost object is never traded and the best response bidding functions can be obtained by solving the following optimization problems for the buyer and seller, respectively:

$$\max_v E\Pi^B(v) = \max_v \left\{ \int_0^{\alpha_L^{-1}(v)} \left(V - \frac{v + \alpha_L(t)}{2} \right) 2(1-t)\, dt \right\}$$

$$\max_{c_L} E\Pi^S(c_L) = \max_{c_L} \left\{ \int_{\beta^{-1}(c_L)}^{1} \left(\frac{\beta(z) + c_L}{2} - C_L \right) dz \right\}$$

A pair of best response strategies $(\alpha_L^*(C_L), \beta^*(V))$ is a solution of the following system of differential equations:

$$\begin{cases}
\frac{\partial \beta^{-1}(x)}{\partial x} = \frac{(1-\beta^{-1}(x))}{2(c_L - x)} \\
\frac{\partial \alpha_L^{-1}(x)}{\partial x} = \frac{\alpha_L^{-1}(x)(2-\alpha_L^{-1}(x))}{4(1-\alpha_L^{-1}(x))(V-x)}
\end{cases} \tag{2.10}$$

Although the system (2.10) does not admit neither linear nor any other higher order polynomial solutions, there exist a family of solutions of (2.10), where both functions $\beta(.)$ and $\alpha_L(.)$ are invertible and strictly increasing on the $[0, 1]$ interval (which can be calculated numerically).

Finally, we need to verify that $\alpha_L^*(\mathrm{C}) \leq \alpha_H^*(\mathrm{C})$. The argument goes as follows. Suppose not. Let $\alpha_L^*(\mathrm{C}) = \alpha_H^*(\mathrm{C}) = c^*$ for some $\mathrm{C} \in [0,1]^2$, such that $C_1 \neq C_2$, so that the likelihood that an object is selected is the same for both objects. Then, given that trade occurs ($c^* \leq \beta^*(V)$), the seller's profit is $\frac{1}{2}\left(\frac{v+c^*}{2} - C_L\right) + \frac{1}{2}\left(\frac{v+c^*}{2} - C_H\right)$. Clearly, the seller can lower his bid for the low-cost object by ε, which would guarantee the selection of the low-cost object, and increase his gains from trade by the positive amount ($\frac{1}{2}(C_H - C_L) - \frac{1}{2}\varepsilon$). So, in equilibrium, $\alpha_L^*(\mathrm{C}) \neq \alpha_H^*(\mathrm{C})$. (Similar argument can be made to show that $\alpha_L^*(\mathrm{C})$ cannot be higher than $\alpha_H^*(\mathrm{C})$). An argument for $v_1^* = v_2^* = v^*$ is essentially the same as the one presented in the previous section.

Thus, whenever one trader is indifferent between all potential objects of trade, the outcome of the bargaining problem is *ex post* efficient with respect to the object selection irrespective of the distribution of the bargaining power.

2.6 The Game Γ4

2.6.1 The Buyer's Bid Double Auction $\Gamma 4\backslash k = 1$

Suppose that the players actual valuations of the object q_1, $V_1 \in [0,1]$, $C_1 \in [0,1]$, are commonly known and $V_1 > C_1$, but their valuations of the object q_2 are private. Suppose that it is commonly known that V_2 is distributed uniformly on the $[0,1]$ interval, independently of the value of V_1 and the seller's type; and C_2 is distributed uniformly on the $[0,1]$ interval, independently of C_1 and the buyer's type.

It is easy to show that since the value of C_1 is commonly known, the buyer's optimal bid for q_1 is $v_1^* = C_1$.

The seller's optimal trading strategy is trivial: he always chooses q_2 whenever it is feasible

$$\gamma^*(\mathrm{C}, v_2) = q_1, \text{ iff } v_2 < C_2$$

$$\gamma^*(\mathbf{C}, v_2) \;=\; q_2, \text{ iff } v_2 \geq C_2$$

The buyer's type $\mathbf{V}$ expected gains from trade, given the seller's trading strategy γ^* are as follows:

$$E\Pi^B(v_2|\ \gamma^*, \mathbf{V}) = (V_2 - v_2)\Pr\{q_2 \text{ is selected}|\ v_2\} + (V_1 - C_1)\Pr\{q_1 \text{ is selected}|\ v_2\}$$

The buyer's optimization problem

$$\max_{v_2}[(V_2 - v_2)v_2 + (V_1 - C_1)(1 - v_2)]$$

has a unique solution

$$v_2^* = \frac{1}{2}V_2 - \frac{1}{2}(V_1 - C_1) \tag{2.11}$$

It is straighforward to verify that the strategy (v_1^*, v_2^*) described above is not dominated by a strategy $\left(-\infty, \frac{1}{2}V_2\right)$ (when the latter is implemented, q_1 is never traded). Let $\Pi_1 = V_1 - C_1$, then the buyer's type $\mathbf{V}$ would prefer (v_1^*, v_2^*) if and only if

$$\begin{aligned} \frac{1}{4}V_2^2 + \frac{1}{4}\Pi_1^2 + \Pi_1 - \frac{1}{2}V_2\Pi_1 &\geq \frac{1}{4}V_2^2 \\ \frac{1}{2}\Pi_1 + 2 &\geq V_2 \end{aligned}$$

which holds for all $V_2 \in [0, 1]$.

The following Proposition summarizes the analysis presented above.

Proposition 15 *The buyer's bidding strategy*

$$v_1^* = C_1$$
$$v_2^* = \begin{cases} \frac{1}{2}V_2 - \frac{1}{2}(V_1 - C_1), \textit{ if } V_2 \geq (V_1 - C_1) \\ 0, \textit{ otherwise} \end{cases}$$

and the seller's trading strategy

$$\begin{cases} \gamma^*(\mathbf{C}, v_2) = q_1, & \textit{iff } v_2^* < C_2 \\ \gamma^*(\mathbf{C}, v_2) = q_2, & \textit{iff } v_2^* \geq C_2 \end{cases}$$

constitute a unique Bayesian-Nash Equilibrium of the game $\Gamma 4 \backslash k = 1$.

Note that when $V_2 < (V_1 - C_1)$, q_1 is selected by default.

However, the object q_2 is not always traded when it is *ex post* efficient. The boundary separating the region where q_2 is traded from the region where q_1 is traded is given by $V_2 = (V_1 - C_1) + 2C_2$, while the boundary separating the region where the *ex post* efficient object is selected from the region where the *ex post in*efficient object is selected is given by $V_2 = (V_1 - C_1) + C_2$.

2.6.2 The Split-the-Difference Double Auction $\Gamma 4 \backslash k = 0.5$

Now, suppose that the players have equal bargaining power. An argument similar to the one presented in Appendix for the game $\Gamma 2$ can be made here to show that $\beta_1^*(V_1) = p_1$, $\alpha_1^*(C_1) = p_1$, where $C_1 \leq p_1 \leq V_1$.

The object selection rule is as follows:

$$\begin{aligned} \gamma(v_2, c_2, C_1, V_1, p_1) &= q_2, \text{ whenever } v_2 \geq c_2 \\ \gamma(v_2, c_2, C_1, V_1, p_1) &= q_1, \text{ whenever } v_2 < c_2 \end{aligned}$$

Then, the buyer's type V expected profit is

$$E\Pi^B(v_2|\mathbf{V}, p_1, \gamma, \overline{\alpha}) = \int_{\alpha_2^{-1}(v_2)}^{1} (V_1 - p_1)\, dt_2 + \int_{0}^{\alpha_2^{-1}(v_2)} \left(V_2 - \frac{v_2 + \alpha_2(t_2)}{2}\right) dt_2$$

and the seller's type C expected profit is

$$E\Pi^S\left(c_2|C,p_1,\gamma,\beta\right)=\int\limits_{\beta_2^{-1}(c_2)}^{1}\left(\frac{\beta_2\left(z_2\right)+c_2}{2}-C_2\right)dz_2+\int\limits_{0}^{\beta_2^{-1}(c_2)}\left(p_1-C_1\right)dz_2$$

Proposition 16 *In the game $\Gamma 4\backslash k=0.5$, where V_1 and C_1 are commonly known, V_2 and C_2 are distributed uniformly on the interval $[0,1]$ independently of V_1 and C_1 and each other, there exists a class of linear equilibria such that for any $C_2\in[0,1]$, $V_2\in[0,1]$,*

$$\begin{cases}\alpha_1^*(C)=p_1\\ \alpha_2^*(C)=\frac{2}{3}C_2+\frac{1}{4}+p_1-\frac{3}{4}C_1-\frac{1}{4}V_1\end{cases}\tag{2.12}$$

$$\begin{cases}\beta_1^*(V)=p_1\\ \beta_2^*(V)=\frac{2}{3}V_2+\frac{1}{12}+p_1-\frac{3}{4}V_1-\frac{1}{4}C_1\end{cases}\tag{2.13}$$

where $C_1\leq p_1\leq V_1$.[11]

The object q_2 is selected if and only if $V_2\geq C_2+\frac{1}{4}+\frac{3}{4}(V_1-C_1)$. The boundary separating the region where the *ex post* efficient object is selected from the region where the *ex post* inefficient object is selected is given by $V_2=C_2+(V_1-C_1)$. Therefore, the equilibria characterized in Proposition 16 are not *fully ex post* efficient.

Thus, whenever the gains from trading q_1 are commonly known, both players adjust their bidding strategies for q_2 so that the latter could not be selected unless the gains from trading q_2 are at least as large as the gains from trading q_1. However, since both traders demand a premium on all possible trades of an object for which their valuations are not commonly known (the object q_2), they end up not always trading q_2 whenever it is an *ex post* efficient choice.

The efficiency properties of the buyer's bid and split-the-difference double auctions for the games Γ1, Γ2, Γ3 and Γ4 will be presented in the next section.

---

[11] *The proof is in Appendix.*

2.7 Efficiency

2.7.1 Bargaining with One Object

We will start with the description of the efficiency properties of the linear equilibria in k-double auctions with single object and uniform priors. In the split-the-difference double auction, the famous Chatterjee-Samuelson (1983) linear equilibrium generates the highest total *ex ante* expected gains from trade.

Trading mechanism (k−double auction)	$k = 1$	$k = 0.5$
Potential gains from trade	.16667	.16667
Maximal equilibrium gains from trade	.125	.14063
'Second-best' gains from trade	.14063	.14063
Expected gains from trade for the buyer	.083333	.070315
Expected gains from trade for the seller	.041667	.070315
Ex ante efficiency Ratio	.74999	.84376
The range of trade	$V \geq 2C$	$V \geq C + \frac{1}{4}$

2.7.2 Bargaining with Two Objects

The Game Γ2

The total potential gains from trade in the game Γ2, where $V_2 = 1$ and $C_1 = 0$ are commonly known, are as follows:

$$\int_0^1 \left(\int_0^{1-C_2} (1 - C_2)\, dV_1 + \int_{1-C_2}^{1} (V_1)\, dV_1 \right) dC_2 = \frac{2}{3} = .66667$$

The following is a summary of the efficiency properties of the linear equilibria of the games $\Gamma 2 \backslash k = 1$ and $\Gamma 2 \backslash k = 0.5$ (for details, see Appendix)[12]:

Trading mechanism (k–double auction)	$\Gamma 2 \backslash k = 1$	$\Gamma 2 \backslash k = 0.5$ ($V_2 = 1$ and $C_1 = 0$)
Potential gains from trade	.66667	.66667
Maximal equilibrium gains from trade	.625	.640625
Expected gains from trade for the buyer	.58333	.5703125
Expected gains from trade for the seller	.041667	.0703125
Ex ante Efficiency Ratio	.9375	.96094

Not surprisingly, the split-the-difference double auction has a higher *ex ante* efficiency ratio than the buyer's bid auction.

The Game $\Gamma 3$

The total *ex ante* expected gains from trade obtained in a unique Bayesian-Nash Equilibrium of the game $\Gamma 3 \backslash k = 1$ are

$$\int_0^1 \left(\int_0^{\frac{1}{3}V + \frac{2}{3} - \frac{1}{3}\sqrt{V^2 - 2V + 4}} 2(V - C)(1 - C)\, dC \right) dV = .19601$$

potential (truth-telling) gains from trade are

$$\int_0^1 \left(\int_0^V 2(V - C)(1 - C)\, dC \right) dV = .25$$

[12]Note that the numbers in the table are correspond to the equilibrium of the game $\Gamma 1 \backslash k = 0.5$, where $\alpha_1^*(C) = \beta_1^*(V) = 0$, which explains why the gains from trade are not equally divided between the traders. There is also another linear equilbrium in the game $\Gamma 1 \backslash k = 0.5$, where $\alpha_2^*(C) = \beta_2^*(V) = 1$.

and the *ex ante* efficiency ratio is

$$E = \frac{.19601}{.25} = .78404$$

Trading mechanism (k–double auction)	$\Gamma 3/k = 1$
Potential gains from trade	.25
Maximal equilibrium gains from trade	.19601
'Second-best' gains from trade	.21039
Expected gains from trade for the buyer	.13733
Expected gains from trade for the seller	.05868
Ex ante Efficiency Ratio	.78404
The range of trade	$C \leq \frac{1}{3}V + \frac{2}{3} - \frac{1}{3}\sqrt{V^2 - 2V + 4}$

Note that when we add a second trading alternative, the total expected gains from trade increase, the buyer's gains from trade increase, while the seller's gains from trade decrease. The efficiency of $k = 1$ double auction rises as well, implying that the effect of having two alternatives to choose from outweighs the effect of having a tighter trading boundary (the latter is due to the fact that the buyer is always expecting the seller to choose the low-cost alternative and therefore, is bidding more aggressively for both objects the higher his reservation value).

The trading mechanism in $\Gamma 3 \backslash k = 1$ is not fully *ex post* efficient, because there is a positive probability that trade does not occur even if the seller's low cost is below the buyer's valuation.

The Game Γ4

For the special case of the game Γ4, when $V_1 = \frac{1}{2}$ and $C_1 = 0$, the total potential gains from trade are as follows:

$$\int_0^1 \left(\int_0^{\frac{1}{2}-C_2} (V_2 - C_2)\, dV_2 + \int_{\frac{1}{2}-C_2}^{1} \left(\frac{1}{2}\right) dV_2 \right) dC_2 = \frac{5}{8} = .625$$

The following is the summary of the efficiency results for the class of the linear equilibria of the game $\Gamma 4$, given that $V_1 = \frac{1}{2}$ and $C_1 = 0$ are commonly known (for details, see Appendix):

Trading mechanism (k–double auction)	$\Gamma 3 \backslash k = 1$	$\Gamma 3 \backslash k = 0.5$
Potential gains from trade	.625	.625
Maximal equilibrium gains from trade	.51563	.51758
Buyer's expected gains from trading q_1	.46875	$.46484 - .92969 p_1$
Buyer's expected gains from trading q_2	.041667	$.043945 - .070313 p_1$
Seller's expected gains from trading q_1	0	$.92969 p_1$
Seller's expected gains from trading q_2	.0052083	$.0087891 + .070313 p_1$
Ex ante Efficiency Ratio	.82501	.82813

Note that since in the game $\Gamma 4 \backslash k = 1$, the price of q_1 is equal to C_1, the seller's gains from trading q_1 are equal to zero.

APPENDIX

The calculation of the probabilities of trading q_1 and q_2 in the game $\Gamma 1 \backslash k = 1$.

Consider REGION I, where $v_2 \geq v_1$, then

$$\begin{aligned}
& \Pr\{q_1 \text{ is selected } | v_2, v_1\} \\
= & \Pr\{q_1 \text{ is selected, both objects are feasible } | v_2, v_1\} \\
& + \Pr\{q_1 \text{ is selected, only } q_1 \text{ is feasible } | v_2, v_1\} \\
= & \Pr\{C_1 \leq v_1, C_2 \leq v_2, v_2 - v_1 \leq C_2 - C_1 \leq 1\} + \Pr\{C_1 \leq v_1, C_2 > v_2\} \\
= & \int_0^{v_1} \Pr\{C_1 = t_1, C_2 \leq v_2, v_2 - v_1 \leq C_2 - C_1\} \, dt_1 + (1 - v_2) v_1 \\
= & \int_0^{v_1} \Pr\{(v_2 - v_1) + t_1 \leq C_2 \leq v_2\} \, dt_1 + (1 - v_2) v_1 \\
= & \int_0^{v_1} (v_1 - t_1) \, dt_1 + v_1 - v_1 v_2 \\
= & v_1 - v_2 v_1 + \frac{1}{2} v_1^2
\end{aligned}$$

$$\begin{aligned}
& \Pr\{q_2 \text{ is selected } | v_2, v_1\} \\
= & \Pr\{q_2 \text{ is selected, both objects are feasible } | v_2, v_1\} \\
& + \Pr\{q_2 \text{ is selected, only } q_2 \text{ is feasible } | v_2, v_1\} \\
= & \Pr\{C_2 \leq v_2, C_1 \leq v_1, C_2 - C_1 \leq v_2 - v_1\} + \Pr\{C_1 > v_1, C_2 \leq v_2\} \\
= & \int_0^{v_2} \Pr\{C_2 = t_2, C_1 \leq v_1, C_2 - C_1 \leq v_2 - v_1\} \, dt_2 + (1 - v_1) v_2 \\
= & \int_0^{v_2} \Pr\{t_2 - (v_2 - v_1) \leq C_1 \leq v_1\} \, dt_2 + (1 - v_1) v_2 \\
= & \int_{v_2 - v_1}^{v_2} (v_2 - t_2) \, dt_2 + \int_0^{v_2 - v_1} (v_1) \, dt_2 + v_2 - v_2 v_1
\end{aligned}$$

$$= v_2 - \frac{1}{2} v_1^2$$

Proof of Lemma 11

Let v_2^*, v_1^* denote the optimal bids.

($\Leftarrow$) Suppose that $V_2 \geq V_1$, then $v_2^* \geq v_1^*$, if and only if

$$\begin{aligned} & (V_1 - v_1^*)\left(v_1^* - v_2^* v_1^* + \frac{1}{2} v_1^{*2}\right) + (V_2 - v_2^*)\left(v_2^* - \frac{(v_1^*)^2}{2}\right) \\ \geq \; & (V_1 - v_1^*)\left(v_1^* - \frac{(v_2^*)^2}{2}\right) + (V_2 - v_2^*)\left(v_2^* - v_2^* v_1^* + \frac{1}{2} v_2^{*2}\right) \end{aligned}$$

$$\begin{aligned} (V_1 - v_1^*)\left(\frac{1}{2}(v_2^*)^2 - v_2^* v_1^* + \frac{1}{2} v_1^{*2}\right) & \geq (V_2 - v_2^*)\left(\frac{1}{2} v_1^{*2} - v_2^* v_1^* + \frac{1}{2} v_2^{*2}\right) \Rightarrow \\ (V_1 - v_1^*)(v_2^* - v_1^*)^2 & \geq (V_2 - v_2^*)(v_2^* - v_1^*)^2, \\ \text{since for any } v_2^*, v_1^*, (v_2^* - v_1^*)^2 & \geq 0 \Rightarrow (v_2^* - v_1^*) \geq (V_2 - V_1) \\ \text{which, given our assumption, implies that } v_2^* & \geq v_1^* \end{aligned}$$

($\Rightarrow$) Solving (1) for V_1, V_2 we obtain:

$$\begin{aligned} V_1(v_1, v_2) &= \frac{1}{2} v_1 \frac{-4 + 2v_2 - 3v_1 + 3v_1^2}{-1 + v_2 - v_1 + v_1^2} \\ V_2(v_1, v_2) &= \frac{1}{2} \frac{-v_1^2 + 3v_1^2 v_2 - 4v_2 + 4v_2^2 - 4v_2 v_1}{-1 + v_2 - v_1 + v_1^2} \end{aligned}$$

Define $D \equiv V_2(v_1, v_2) - V_1(v_1, v_2) = \frac{1}{2} \frac{-v_1^2 + 3v_1^2 v_2 - 4v_2 + 4v_2^2 - 4v_2 v_1}{-1 + v_2 - v_1 + v_1^2} - \frac{1}{2} v_1 \frac{-4 + 2v_2 - 3v_1 + 3v_1^2}{-1 + v_2 - v_1 + v_1^2}$, then whenever $v_2 \geq v_1$, $D \geq 0$, and whenever $v_2 < v_1, D < 0$. See the illustration below. ■

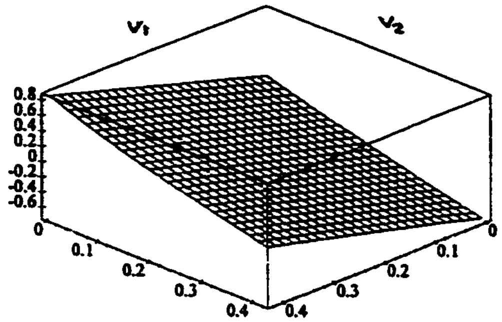

Figure 7. The spread between the buyer's valuations in the game $\Gamma 1 \backslash k = 1$.

The derivation of the buyer's best response bidding functions in the game $\Gamma 1 \backslash k = 1$.

For the buyer's type $\mathbf{V}' = (1,\ 1 - x)$, the solution of the maximization problem is such that

$$v_1^* = \rho$$

and

$$v_2^* = \frac{1}{2} - \frac{1}{2}\rho + \frac{1}{2}\rho x + \frac{3}{4}\rho^2$$

where ρ is a root of polynomial (2.2)

$$9Z^3 + (-15 + 9x) Z^2 + \left(2x^2 - 8x\right) Z + 2 - 2x \tag{2.14}$$

In general, the polynomial (2.2) has three roots

$$\begin{aligned}
\rho_1 &= \tfrac{5}{9} - \tfrac{1}{3}x + [A(x) - B(x)] \\
\rho_2 &= \tfrac{5}{9} - \tfrac{1}{3}x - \tfrac{1}{2}\left[\left(1 + i\sqrt{3}\right) A(x) - \left(1 - i\sqrt{3}\right) B(x)\right] \\
\rho_3 &= \tfrac{5}{9} - \tfrac{1}{3}x - \tfrac{1}{2}\left[\left(1 - i\sqrt{3}\right) A(x) - \left(1 + i\sqrt{3}\right) B(x)\right]
\end{aligned}$$

where

$$A(x) = \sqrt[3]{\left(-\frac{2}{81}x^2 + \frac{4}{81}x + \frac{44}{729} + \frac{1}{243}\sqrt{(-3x^6 + 18x^5 - 75x^4 + 180x^3 - 957x^2 + 1602x - 1521)}\right)}$$
$$B(x) = \frac{-\frac{1}{27}x^2 + \frac{2}{27}x - \frac{25}{81}}{A(x)}$$

The numerical solution of

$$\begin{cases} 9Z^3 + (-15 + 9x) Z^2 + (2x^2 - 8x) Z + 2 - 2x = 0 \\ x \in (0,1) \\ Z \in (0,1) \end{cases}$$

is shown in Figure 8.

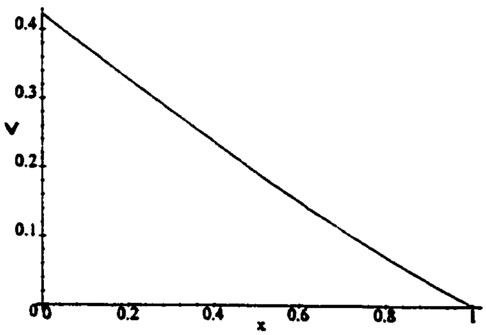

Figure 8: The best response bidding function for the buyer's type $V = (1, 1 - x)$.

Proof of Proposition 13

Given that $C_1 = 0$ and $V_2 = 1$ are commonly known, the buyer's and seller's bidding functions can be expressed in terms of V_1 and C_2, respectively, i.e. the buyer's bids can be written as $v_1 = \beta_1(V_1)$ and $v_2 = \beta_2(V_1)$ and the seller's bids can be written as $c_1 = \alpha_1(C_2)$ and $c_2 = \alpha_2(C_2)$.

Consider the buyer's problem.

Suppose that $\alpha_1(C_2)$ is non-increasing, $\alpha_2(C_2)$ is strictly increasing and $\alpha_1(C_2) \leq \alpha_2(C_2)$ for any $C_2 \in [0,1]$. Define $\bar{v} \equiv \alpha_2\left(\alpha_1^{-1}(v_1)\right)$ and let's partition the subspace of all configurations of the buyer's bids, such that $v_2 \geq v_1$, into two regions:

REGION I: where $v_2 \geq \bar{v}$, which implies that $\alpha_2^{-1}(v_2) \geq \alpha_1^{-1}(v_1)$ and

REGION II: where $v_2 < \bar{v}$, which implies that $\alpha_2^{-1}(v_2) < \alpha_1^{-1}(v_1)$.

Note, that given our assumptions about the seller's bidding functions, q_1 is feasible, i.e. $v_1 \geq \alpha_1(C_2)$, whenever $\alpha_1^{-1}(v_1) \leq C_2 \leq 1$ and q_2 is feasible, i.e. $v_2 \geq \alpha_2(C_2)$, whenever $0 \leq C_2 \leq \alpha_2^{-1}(v_2)$.

Consider REGION I. First of all, note that, in REGION I, there is a non-empty interval $[\alpha_1^{-1}(v_1), \alpha_2^{-1}(v_2)]$, where both objects q_1 and q_2 are feasible. Let C_2^* be implicitly defined by $v_2 - \alpha_2(C_2^*) = v_1 - \alpha_1(C_2^*)$. Also, let's define a new function $\tilde{\alpha}(C_2) \equiv \alpha_2(C_2) - \alpha_1(C_2)$,which is also invertible and strictly increasing. Then, we can express C_2^* in terms of the difference between the buyer's bids, i.e. $C_2^* = \tilde{\alpha}^{-1}(v_2 - v_1)$. Clearly, C_2^* must lie in the interval $\left[\alpha_1^{-1}(v_1), \alpha_2^{-1}(v_2)\right]$, because it is the only subinterval of $[0,1]$ where both objects are feasible.

Then, q_2 should be selected whenever $0 \leq C_2 \leq \tilde{\alpha}^{-1}(v_2 - v_1)$ and q_1 should be selected whenever $\tilde{\alpha}^{-1}(v_2 - v_1) \leq C_2 \leq 1$, and the buyer's optimization problem can be written as follows:

$$\max_{v_1, v_2} E\Pi^B(v_1, v_2) = \max_{v_1, v_2} \left(\int_0^{\tilde{\alpha}^{-1}(v_2 - v_1)} (V_2 - \frac{v_2 + \alpha_2(t_2)}{2}) dt_2 + \int_{\tilde{\alpha}^{-1}(v_2 - v_1)}^{1} (V_1 - \frac{v_1 + \alpha_1(t_2)}{2}) dt_2 \right) \tag{2.15}$$

Differentiating the buyer's objective function with respect to v_1, v_2, we obtain the following f.o.c's

$$\frac{\partial \tilde{\alpha}^{-1}(x)}{\partial x} \left((V_2 - \frac{v_2 + \alpha_2(\tilde{\alpha}^{-1}(x))}{2}) - (V_1 - \frac{v_1 + \alpha_1(\tilde{\alpha}^{-1}(x))}{2}) \right) = \frac{\tilde{\alpha}^{-1}(x) - 1}{2} \tag{2.16}$$

$$\frac{\partial \tilde{\alpha}^{-1}(x)}{\partial x} \left((V_2 - \frac{v_2 + \alpha_2(\tilde{\alpha}^{-1}(x))}{2}) - (V_1 - \frac{v_1 + \alpha_1(\tilde{\alpha}^{-1}(x))}{2}) \right) = \frac{\tilde{\alpha}^{-1}(x)}{2} \tag{2.17}$$

which imply that $\tilde{\alpha}^{-1}(x) - 1 = \tilde{\alpha}^{-1}(x)$ and, therefore, the solution of the system of equations (2.16), (2.17) does not exist.

Suppose that the seller's bidding function for q_1 is constant and such that $\alpha_1(C_2) = 0$, for all $C_2 \in [0,1]$. Then, the buyer's best response bid for the object q_1 is $v_1^* = 0$, and the buyer's best response bid for the object q_2 is a solution of the following optimization problem:

$$\max_{v_2} E\Pi^B(v_2) = \max_{v_2} \left\{ \int_0^{\alpha_2^{-1}(v_2)} \left(V_2 - \frac{v_2 + \alpha_2(t_2)}{2} \right) dt_2 + \int_{\alpha_2^{-1}(v_2)}^{1} (V_1 - 0)\, dt_2 \right\} \tag{2.18}$$

At the points where derivatives exist

$$\frac{\partial E\Pi^B(v_2)}{\partial v_2} = ((V_2 - v_2) - (V_1 - 0)) \frac{\partial \alpha_2^{-1}(v_2)}{\partial v_2} - \left(\frac{1}{2}\right) \alpha_2^{-1}(v_2) \tag{2.19}$$

$\frac{\partial E\Pi^B(v_2)}{\partial v_2} = 0$ characterizes the local maximum of $E\Pi^B(v_2)$ if and only if

$$3\left(\frac{\partial \alpha_2^{-1}(v_2)}{\partial v_2}\right)^2 - \alpha_2^{-1}(v_2)\left(\frac{\partial \alpha_2^{2-1}(v_2)}{\partial v_2}\right) > 0 \tag{2.20}$$

(For the configuration of the buyer's bids in REGION II, there is a non-empty subset of the values of C_2, for which at least one of the objects cannot be traded, therefore, such configurations are not considered here).

Consider the seller's problem.

Similarly, after we pin down $c_1 = 0$, we can re-write the seller's optimization problem as follows:

$$\max_{c_2} E\Pi^S(c_2) = \max_{c_2} \left\{ \int_0^{\beta_2^{-1}(c_2)} \left(\frac{c_2 + \beta_2(t_1)}{2} - C_2 \right) dt_1 + \int_{\beta_2^{-1}(c_2)}^{1} (0 - C_1)\, dt_1 \right\} \tag{2.21}$$

At the points where derivatives exist

$$\frac{\partial E\Pi^S(c_2)}{\partial c_2} = \frac{\partial \beta_2^{-1}(c_2)}{\partial c_2}((c_2 - C_2) - (0 - C_1)) + \frac{1}{2}\beta_2^{-1}(c_2) \tag{2.22}$$

and $\frac{\partial E\Pi^S(c_2)}{\partial c_2} = 0$ characterizes the local maximum of $E\Pi^S(c_2)$ if and only if

$$3\left(\frac{\partial \beta_2^{-1}(c_2)}{\partial c_2}\right) - \frac{\beta_2^{-1}(c_2)}{\left(\frac{\partial \beta_2^{-1}(c_2)}{\partial c_2}\right)}\left(\frac{\partial \beta_2^{2-1}(c_2)}{\partial c_2}\right) < 0 \tag{2.23}$$

A linear equilibrium of the bargaining problem can be derived using the method of undetermined coefficients.

Suppose that $\alpha_2(C_2) = a + bC_2$, then $\alpha_2^{-1}(v_2) = \frac{v_2 - a}{b}$ and suppose that $\beta_2(V_1) = d + fV_1$, then $\beta_2^{-1}(c_2) = \frac{c_2 - d}{f}$. Then, the system of differential equations obtained from the players' f.o.c's (2.19) and (2.22)

$$\begin{aligned} \frac{\partial \beta_2^{-1}(c_2)}{\partial c_2} &= \frac{-\beta_2^{-1}(c_2)}{2((c_2 - C_2) - (0 - C_1))} \\ \frac{\partial \alpha_2^{-1}(v_2)}{\partial v_2} &= \frac{\alpha_2^{-1}(v_2)}{2((V_2 - v_2) - (V_1 - 0))} \end{aligned}$$

can be rewritten as follows

$$\begin{cases} \frac{1}{b}((V_2 - v_2) - (V_1 - 0)) = \frac{1}{2}\left(\frac{v_2 - a}{b}\right) \\ \frac{1}{f}((c_2 - C_2) - (0 - C_1)) = -\frac{1}{2}\left(\frac{c_2 - d}{f}\right) \end{cases} \tag{2.24}$$

The solution of (2.24) is

$$\begin{aligned} v_2 &= \tfrac{2}{3}(V_2 - V_1 + 0) + \tfrac{1}{3}a \\ c_2 &= \tfrac{1}{3}d + \tfrac{2}{3}(C_2 + 0 - C_1) \end{aligned}$$

which implies that

$$\begin{cases} a = \frac{1}{3}d + \frac{2}{3}(0 - C_1) \\ d = \frac{2}{3}(V_2 + 0) + \frac{1}{3}a \\ f = -\frac{2}{3} \\ b = \frac{2}{3} \end{cases} \tag{2.25}$$

Solving (2.25), we obtain

$$
\begin{aligned}
f &= -\tfrac{2}{3} \\
b &= \tfrac{2}{3} \\
d &= \tfrac{3}{4}V_2 + 0 - \tfrac{1}{4}C_1 \\
a &= \tfrac{1}{4}V_2 + 0 - \tfrac{3}{4}C_1
\end{aligned}
$$

which gives us

$$
\begin{aligned}
c_2 &= \tfrac{2}{3}C_2 + \left(\tfrac{1}{4} + 0\right) \\
v_2 &= \left(\tfrac{3}{4} + 0\right) - \tfrac{2}{3}V_1
\end{aligned}
\tag{2.26}
$$

Then, the object q_2 is feasible, i.e. $v_2 \geq c_2$, if and only if $\frac{3}{4} - C_2 \geq V_1$. One can verify that the bidding functions in (2.26) satisfy the second-order conditions of the players' maximization problems. ■

Proof of Proposition 14

Suppose that the buyer submits different bids for q_1 and q_2, v_1 and v_2, respectively, such that $v_2 \geq v_1$.

The best response bids are given by the solution to the following system of equations (see Section 2.3.1 for details):

$$
\begin{cases}
-v_1 + v_2 v_1 - \frac{1}{2}v_1^2 + (V - v_1)(1 - v_2 + v_1) - (V - v_2)\, v_1 = 0 \\
-(V - v_1)\, v_1 + \frac{1}{2}v_1^2 - 2v_2 + V = 0
\end{cases}
$$

There are two solutions:

Solution 1: $\begin{cases} v_1 = \frac{1}{3}V - \frac{2}{3} \\ v_2 = \frac{1}{2}V + \frac{1}{3} - \frac{1}{12}V^2 \end{cases}$ and

Solution 2: $\begin{cases} v_1 = \rho \\ v_2 = \rho \end{cases}$ where ρ is a root of $3Z^2 + (-2V - 4)\, Z + 2V$,

which has the following two roots: $\frac{1}{3}V + \frac{2}{3} + \frac{1}{3}\sqrt{(V^2 - 2V + 4)}$
$\frac{1}{3}V + \frac{2}{3} - \frac{1}{3}\sqrt{(V^2 - 2V + 4)}$

However, only one solution satisfies assumptions (1),(2),(3) of Section 2.2, i.e.

$v_1 = v_2 = \frac{1}{3}V + \frac{2}{3} - \frac{1}{3}\sqrt{(V^2 - 2V + 4)}$ ■

Proof of Proposition 16

The f.o.c's of the players' optimization problems form the following system of differential equations:

$$\begin{cases} \frac{\partial \alpha_2^{-1}(v_2)}{\partial v_2} = \frac{\alpha_2^{-1}(v_2)}{2((V_2 - v_2) - (V_1 - p_1))} \\ \frac{\partial \beta_2^{-1}(c_2)}{\partial c_2} = \frac{(1 - \beta_2^{-1}(c_2))}{2((c_2 - C_2) - (p_1 - C_1))} \end{cases} \tag{2.27}$$

Assuming that $\alpha_2(C_2) = a + bC_2$, which implies that $\alpha_2^{-1}(v_2) = \frac{v_2 - a}{b}$, and that $\beta_2(V_1) = d + fV_1$, which implies that $\beta_2^{-1}(c_2) = \frac{c_2 - d}{f}$. We can rewrite (2.27) as follows:

$$\begin{cases} \frac{1}{b}((V_2 - v_2) - (V_1 - p_1)) = \frac{1}{2}\left(\frac{v_2 - a}{b}\right) \\ \frac{1}{f}((c_2 - C_2) - (p_1 - C_1)) = \frac{1}{2}\left(1 - \frac{c_2 - d}{f}\right) \end{cases} \tag{2.28}$$

Solving (2.28), we obtain

$$\begin{aligned} c_2 &= \tfrac{2}{3}C_2 + \tfrac{1}{3}f + \tfrac{2}{3}p_1 - \tfrac{2}{3}C_1 + \tfrac{1}{3}d \\ v_2 &= \tfrac{2}{3}V_2 - \tfrac{2}{3}V_1 + \tfrac{2}{3}p_1 + \tfrac{1}{3}a \end{aligned}$$

which implies that

$$\begin{cases} a = \frac{1}{3}f + \frac{2}{3}p_1 - \frac{2}{3}C_1 + \frac{1}{3}d \\ d = -\frac{2}{3}V_1 + \frac{2}{3}p_1 + \frac{1}{3}a \\ f = \frac{2}{3} \\ b = \frac{2}{3} \end{cases}$$

and that

$$\begin{aligned} f &= \tfrac{2}{3} \\ b &= \tfrac{2}{3} \\ d &= \tfrac{1}{12} - \tfrac{3}{4}V_1 + p_1 - \tfrac{1}{4}C_1 \\ a &= \tfrac{1}{4} - \tfrac{1}{4}V_1 + p_1 - \tfrac{3}{4}C_1 \end{aligned}$$

Therefore, the player's bidding functions for q_1 and q_2 are as follows:

$$\begin{aligned} c_2 &= \frac{2}{3}C_2 + \frac{1}{4} + \left(p_1 - \frac{3}{4}C_1 - \frac{1}{4}V_1\right) \\ v_2 &= \frac{2}{3}V_2 + \frac{1}{12} + \left(p_1 - \frac{3}{4}V_1 - \frac{1}{4}C_1\right) \end{aligned}$$

■

The calculation of the ex ante expected gains from trade obtained in the linear equilibrium of the game $\Gamma 2 \backslash k = 1$.

The buyer's *ex ante* expected gains from trade are

$$\int_0^1 \left(\int_0^{\frac{1}{2}-\frac{1}{2}V_1} \left(1 - \frac{1-V_1}{2}\right) dC_2 + \int_{\frac{1}{2}-\frac{1}{2}V_1}^{1} (V_1)\, dC_2 \right) dV_1 = \frac{7}{12}$$

The seller's *ex ante* expected gains from trade are

$$\int_0^{\frac{1}{2}} \left(\int_0^{1-2C_2} \left(\frac{1-V_1}{2} - C_2\right) dV_1 + \int_{1-2C_2}^{1} (0)\, dV_1 \right) dC_2 = \frac{1}{24}$$

The total *ex ante* expected gains from trade are $(\frac{7}{12} + \frac{1}{24} = \frac{5}{8})$

$$\int_0^1 \left(\int_0^{\frac{1}{2}-\frac{1}{2}V_1} (1 - C_2)\, dC_2 + \int_{\frac{1}{2}-\frac{1}{2}V_1}^{1} (V_1)\, dC_2 \right) dV_1 = \frac{5}{8}$$

The calculation of the ex ante expected gains from trade obtained in the linear equilibrium of the game $\Gamma 2 \backslash k = 0.5$.

The buyer's *ex ante* expected utility is

$$\int_{0}^{3/4}\left(\int_{0}^{\frac{3}{4}-V_1}\left(1-\frac{1-\frac{2}{3}V_1+\frac{2}{3}C_2}{2}\right)dC_2+\int_{\frac{3}{4}-V_1}^{1}(V_1)\,dC_2\right)dV_1+\int_{3/4}^{1}(V_1)\,dV_1=\frac{73}{128}$$

The seller's *ex ante* expected utility is

$$\int_{0}^{\frac{3}{4}}\left(\int_{0}^{\frac{3}{4}-C_2}\left(\frac{1-\frac{2}{3}V_1+\frac{2}{3}C_2}{2}-C_2\right)dV_1+\int_{\frac{3}{4}-C_2}^{1}(0)\,dV_1\right)dC_2=\frac{9}{128}$$

and the total ex ante expected gains from trade are $(\frac{73}{128}+\frac{9}{128}=\frac{41}{64})$, or

$$\int_{0}^{3/4}\left(\int_{0}^{3/4-C_2}(1-C_2)\,dV_1+\int_{3/4-C_2}^{1}(V_1)\,dV_1\right)dC_2+\int_{3/4}^{1}\left(\int_{0}^{1}(V_1)\,dV_1\right)dC_2=\frac{41}{64}$$

The calculation of the ex ante expected gains from trade for the special case of the game Γ3 (when $C_1=0$ and $V_1=\frac{1}{2}$)

In the linear equilibrium of the game $\Gamma3\backslash k=0.5$, given that $C_1=0$ and $V_1=\frac{1}{2}$,

the buyer's *ex ante* expected utility is

$$\begin{aligned}&\int_{\frac{1}{4}+\frac{3}{4}(\frac{1}{2})}^{1}\left(\int_{0}^{y-\frac{1}{4}-\frac{3}{4}((\frac{1}{2}))}\left(y-\left(\frac{1}{3}y+\frac{1}{6}+p_1-\frac{1}{2}\left(\frac{1}{2}\right)-\frac{1}{2}(0)+\frac{1}{3}x\right)\right)dx\right)dy\\&+\int_{\frac{1}{4}+\frac{3}{4}(\frac{1}{2})}^{1}\left(\int_{y-\frac{1}{4}-\frac{3}{4}((\frac{1}{2}))}^{1}\left(\frac{1}{2}-p_1\right)dx\right)dy\\&+\int_{0}^{\frac{1}{4}+\frac{3}{4}(\frac{1}{2})}\left(\int_{0}^{1}\left(\frac{1}{2}-p_1\right)dx\right)dy=\frac{521}{1024}-p_1\end{aligned}$$

the seller's *ex ante* expected utility is

$$\int_0^{\frac{3}{4}-\frac{3}{4}\left(\frac{1}{2}\right)}\left(\int_{x+\frac{1}{4}+\frac{3}{4}\left(\left(\frac{1}{2}\right)-(0)\right)}^{1}\left(\left(\frac{1}{3}y+\frac{1}{6}+p_1-\frac{1}{2}\left(\frac{1}{2}\right)-\frac{1}{2}(0)+\frac{1}{3}x\right)-x\right)dy\right)dx$$
$$+\int_0^{\frac{3}{4}-\frac{3}{4}\left(\frac{1}{2}\right)}\left(\int_0^{x+\frac{1}{4}+\frac{3}{4}\left(\left(\frac{1}{2}\right)-(0)\right)}(p_1-(0))\,dy\right)dx$$
$$+\int_{\frac{3}{4}-\frac{3}{4}\left(\frac{1}{2}\right)}^{1}\left(\int_0^1(p_1-(0))\,dy\right)dx=\frac{9}{1024}+p_1$$

and the total *ex ante* expected gains from trade are ($\frac{521}{1024}-p_1+\frac{9}{1024}+p_1=\frac{265}{512}$) or

$$\int_0^{\frac{3}{4}-\frac{3}{4}\left(\frac{1}{2}\right)}\left(\int_{x+\frac{1}{4}+\frac{3}{4}\left(\left(\frac{1}{2}\right)-(0)\right)}^{1}(y-x)\,dy+\int_0^{x+\frac{1}{4}+\frac{3}{4}\left(\left(\frac{1}{2}\right)-(0)\right)}\left(\frac{1}{2}\right)dy\right)dx$$
$$+\int_{\frac{3}{4}-\frac{3}{4}\left(\frac{1}{2}\right)}^{1}\left(\int_0^1\left(\frac{1}{2}\right)dy\right)dx=\frac{265}{512}$$

In the linear equilibrium of the game $\Gamma3\backslash k=1$, assuming that $C_1=0$ and $V_1=\frac{1}{2}$, the buyer's *ex ante* expected gains from trade are (note that since the buyer proposes the price $p_1=0$),

$$\int_{1/2}^{1}\left(\int_0^{\frac{1}{2}y-\frac{1}{2}\left(\frac{1}{2}\right)}\left(y-\left(\frac{1}{2}y-\frac{1}{2}\left(\frac{1}{2}\right)\right)\right)dx+\int_{\frac{1}{2}y-\frac{1}{2}\left(\frac{1}{2}\right)}^{1}\left(\frac{1}{2}-0\right)dx\right)dy$$
$$+\int_0^{1/2}\left(\int_0^1\left(\frac{1}{2}-0\right)dx\right)dy=\frac{49}{96}$$

the seller's *ex ante* expected gains from trade

$$\int_0^{\frac{1}{2}-\frac{1}{2}\left(\frac{1}{2}\right)}\left(\int_{2x+\frac{1}{2}}^{1}\left(\left(\frac{1}{2}y-\frac{1}{2}\left(\frac{1}{2}\right)\right)-x\right)dy+\int_0^{2x+\frac{1}{2}}(0-0)\,dy\right)dx$$
$$+\int_{\frac{1}{2}-\frac{1}{2}\left(\frac{1}{2}\right)}^{1}\left(\int_0^1(0-0)\,dy\right)dx=\frac{1}{192}$$

and the total *ex ante* expected gains from trade ($\frac{49}{96}-\frac{15}{16}p_1+\frac{1}{192}+\frac{15}{16}p_1=\frac{33}{64}$) or

$$\int_0^{1/4}\left(\int_{2x+\frac{1}{2}}^{1}(y-x)\,dy+\int_0^{2x+\frac{1}{2}}\left(\frac{1}{2}\right)dy\right)dx+\int_{1/4}^{1}\left(\int_0^{1}\left(\frac{1}{2}\right)dy\right)dx=\frac{33}{64}$$

Chapter 3

Credibility and Rationality in Cheap Talk Games

"In each case we are presented with alternative pictures.
The need to choose between these pictures seems very compelling;
but the non-pictorial content of the pictures is unclear..."
Michael Dummet, "The Logical Basis of Metaphysics"

3.1 Introduction

Cheap talk is an exchange of costless and non-verifiable messages. In a game with multiple equilibria it naturally comes into the picture as a plausible explanation for why a particular equilibrium is being played. It seems obvious that in a pure coordination game any meaningful communication should lead to the selection of a (unique) Pareto dominant equilibrium. But even in a game where players have conflicting preferences over different equilibrium outcomes, the pre-play communication can reveal some information and therefore, lead to some degree of coordination.

In order to examine the potential of cheap talk as an equilibrium coordination device, we can add one or more rounds of pre-play communication to the original game and analyze the process of the focal beliefs' formation assuming that the rules, which determine how the players transmit information and interpret other players' signals, are commonly known.

After the idea of equilibrium selection through cheap-talk was first introduced by Farrell (1985), a variety of new equilibrium refinements appeared in the literature. The early development of the cheap-talk refinement theories (Neologism-Proofness (Farrell, 1985) and Announcement-Proofness (Matthews, Okuno-Fujiwara and Postlewaite, 1991) was motivated by the idea that a plausible equilibrium of communication game had to be 'stable' with respect to potential deviations that could be induced by credible messages. Later on, the 'non-equilibrium' refinement theories of credible message rationalizability (Rabin, 1990 and Zapater, 1991) were constructed on the assumptions of common knowledge of rationality and credible messages and were aimed at explaining both equilibrium and off-equilibrium behavior of the players within a single framework.

This paper considers the class of games known as the Sender-Receiver Cheap Talk Games, where only the Sender has private information and makes statements in the first stage of the game and only the Receiver takes actions in the second stage of the game. Thus, the Sender can affect the outcome of the game only if he convinces the Receiver to change his beliefs. There is also a common knowledge of 'communication theory' (the latter will be referred to as 'common language') according to which the Receiver interprets the Sender's messages. Clearly, according to any reasonable theory the Sender should not be able to affect the Receiver's beliefs unless he reveals his information truthfully. Thus, a 'common language' includes the message credibility conditions that should be used by the Receiver to determine whether any given statement is true or false.

The main objective of this paper is to develop a unified framework for comparing different cheap talk refinements in terms of the relative strength of the underlying epistemic assumptions (what the players know about each others' conjectures and rationality) and the complexity of the language of communication. The latter should enable us to make better judgments about the applicability of cheap talk refinements in different contexts.

Section 3.2 of this paper introduces the Sender-Receiver Cheap Talk Game (SRCTG), as well as some definitions and general remarks on the structure of the language of communication. Section 3.3 presents a brief overview of the Theories of Neologism and Announcement Proofness. Section 3.4 develops an epistemic framework for the SRCTG and presents the formalization of the Stiglitz Critique. Section 3.5 presents the epistemic model of fixed equilibrium rationalizability. The sufficient epistemic conditions for credible neologism and credible announcement are formulated and proven in Section 3.6. Section 3.7 discusses the refinements based on the concept of credible message rationalizability and relates them to the Theories of Neologism and Announcement Proofness. Section 3.8 compares the communication-proof equilibrium outcomes obtained by applying different refinements to the SRCTG. And, Section 3.9 concludes the paper.

3.2 The Sender-Receiver Cheap Talk Game

Denote by $\Gamma^0(\pi)$ the base game, that is the game where only one of the players, who will be referred to as the Receiver in the extended cheap talk game, takes actions:[1]

$\Gamma^0(\pi) = \{T, \pi, A, u^s, u^r\}$, where

T = finite set of Sender's types;

π = the Receiver's prior beliefs about the Sender's type;

A = finite set of actions available to the Receiver;

u^s = the Sender's payoff function, $u^s : A \times T \rightarrow \Re$;

u^r = the Receiver's payoff function, $u^r : A \times T \rightarrow \Re$.

Denote by $\Gamma^1(\pi)$ the original game $\Gamma^0(\pi)$ 'extended' by adding one stage of pre-play message exchange.

[1] In what follows, I will try to follow as closely as possible the notation from Matthews, Okuno-Fujiwara and Postlewaite (1991).

$\Gamma^1(\pi) = \{T, \pi, \Im^r, \Im^s, u^s, u^s\}$, where

M = set of messages available to the Sender;

τ = the Sender's 'talking' strategy, $\tau : T \to \Delta M$; and $\Im^s$ is the Sender's strategy set;

α = the Receiver's 'action' strategy, $\alpha : M \to \Delta A$; and $\Im^r$ is the Receiver's strategy set.[2]

$\beta : M \to \Delta T$ is the belief revision function by which the Receiver updates his beliefs about the Sender's type.

The interim expected payoffs are defined as follows :

$$U^s(\tau, \alpha|t) = \sum_{m\in M} \sum_{a\in A} \tau(m|t)\alpha(a|m)u^s(a,t)$$

$$U^r(\tau, \alpha|\beta) = \sum_{t\in T} \sum_{m\in M} \sum_{a\in A} \beta(t|m)\tau(m|t)\alpha(a|m)u^r(a,t)$$

Denote by $\Sigma^1(\pi)$ the set of Perfect Bayesian Equilibria (PBE) of $\Gamma^1(\pi)$ with $\sigma \equiv (\tau, \alpha, \beta)$ being the typical element of this set and by $\Sigma^0(\pi)$ the equilibrium set of $\Gamma^0(\pi)$.

Throughout this paper, we will focus on the class of equilibrium refinements of the Sender-Receiver Cheap Talk Game (SRCT Game) which are based on Farrell's '*common language*' and '*rich language*' assumptions. Joseph Farrell was the first who proposed to analyze the pre-game communication between the players who were using natural language to communicate the pay-off relevant information. Because in a SRCT Game, the language is used strategically not all statements made by the Sender can (or should) be believed by the Receiver. According to Farrell (1985), a message should be believed only if it satisfies certain *credibility criteria* that are commonly known among the players. A '*common language*' assumption implies that everybody knows that everybody knows that everybody knows ...

[2]Note that τ and α are defined in such a way that both players can employ mixed strategies.

ad infinitum that the Receiver believes the Sender's statement only if it is credible and only then the Receiver adjusts his beliefs according to a particular belief-revision function β. The '*rich language*' assumption means that there are sufficiently many messages in a common language to enable the Sender to upset any equilibrium of $\Gamma^1(\pi)$ if the Sender prefers an out-of-equilibrium action of the Receiver to his equilibrium payoff.

Let's take a closer look at the structure of the language of communication. In any language, we can distinguish between the syntax (the set of symbols) and the semantics (the meanings of symbols).

First, consider a fairly simple language, where the collection of messages (symbols), M, available to the Sender consists of all subsets of T, excluding $\emptyset$. Using such language, the Sender can only say that his type belongs to a particular subset of the type space, like "My type is in $\{a, b\}$". The *meaning* of every message $m \in M$ is given by the beliefs it induces. Assuming that both players are Bayesian rational, consider the following class of the Receiver's belief-revision functions.

Definition 17 *Let $B(t|m)$ be a class of conditional probability distributions, where $\beta(t|m) \in B(t|m)$ is defined as follows:*

$$\beta(t|m) = \frac{\pi(t)p(t)}{\sum_{t' \in T} \pi(t')p(t')}, \quad \textit{where} \begin{cases} p(t) = 1, & \textit{if } t \in m \\ p(t) \in [0,1], & \textit{if } t \notin m \end{cases}$$

Suppose that the prior beliefs are uniform. Then, when using $\beta(t|m)$, the Receiver interprets the Sender's message "My type is in $\{a, b\}$" as implying that the Sender's type can be either a or b or, possibly, something else.

Denote by $B^L(t|m)$ an element of the class $B(t|m)$ such that $p(t) = 1$, if $t \in m$ and $p(t) = 0$, if $t \notin m$. When using $B^L(t|m)$, the Receiver interprets every message m *literally* (takes it at its face value), so that his conditional beliefs, given m, are always concentrated *only* on the types $t \in m$. For example, the message "My type is in $\{a, b\}$"

would be understood as follows: "It is equally likely that the Sender's true type is either a or b, but not c or d or anything else".

Now, consider a richer syntax, where, in addition to referring to a subset of types, say $\{a, b\}$, the Sender can say something about the relative likelihoods of him being either type a or type b. For $D \subset T$, define a probability distribution $\delta : D \to \Delta M$ and suppose that the Sender can announce (m, d), where $d \equiv (D, \delta)$.

Definition 18 *Let $\tilde{B}(t|m)$ be a class of conditional probability distributions, where $\tilde{\beta}(t|m) \in \tilde{B}(t|m)$ is defined as follows:*

$$\tilde{\beta}(t|(m,d)) = \frac{\pi(t)p(t)}{\sum_{t' \in T} \pi(t')p(t')}, \text{ where } \begin{cases} p(t) = \delta(m|t), & \text{if } t \in D \\ p(t) \in [0,1], & \text{if } t \notin D \end{cases}$$

Similarly, denote by $\tilde{B}^L(t|m)$ an element of the class $\tilde{B}(t|m)$ such that $p(t) = \delta(m|t)$, if $t \in D$ and $p(t) = 0$, if $t \notin D$. For instance, given that the prior beliefs are uniform and the Receiver is using $\tilde{B}^L(t|m)$, he would interpret the Sender's announcement "My type is in $\{a, b\}$, but it is twice more likely to be an a rather than a b" as implying that the Sender can be of type a with probability $\frac{2}{3}$ or he can be of type b with probability $\frac{1}{3}$, but not anything else". Example 1 illustrates the distinction between $B(t|m)$ and $\tilde{B}(t|m)$.

Example 1: Consider a game $\Gamma^0(\pi)$, where $T = \{t_1, t_2\}$, $A = \{A_1, A_2, A_3, A_4\}$, $\pi(t_1) = \pi(t_2) = 1/2$ and the pay-off matrix is as follows:

	A_1	A_2	A_3	A_4
t_1	0, 3.5	3, 3	0, 2	0, 0
t_2	0, 0	3, 1	0, 2	0, 3.5

The Sender cannot achieve his preferred outcome unless he can induce the Receiver to take action A_2. The Receiver would choose A_2 if he believed, for example, that he was facing type t_1 with probability 0.6 and that he was facing type t_2 with probability 0.4. Suppose that the Sender could send a signal S that the Receiver would interpret as $\beta(t_1|S) = 0.6$ and $\beta(t_2|S) = 0.4$. Given that the priors are uniform, the latter would be consistent with the Sender's talking strategy $\tau(S|t_1) = 0.6$ and $\tau(S|t_2) = 0.4$. Clearly, if the announcement S was believed, the Sender would strictly prefer to make it.

For a given syntax, the relative strength of the credibility criteria can only reflect the level of the players' sophistication (the level of mutual knowledge of rationality) that would be required for the players to determine the credibility of any given message. On the other hand, for a given level of the players' sophistication, the syntactic differences between the languages of communication can lead to the selection of different equilibrium outcomes. As it will become clear in the next section, the differences between the Farrell's refinement theory and the one developed by Matthews et al. are underpinned by the syntactic differences described above.

3.3 The Theories of Neologism- and Announcement-Proofness

3.3.1 Neologism-Proof Equilibria

The Farrell's cheap talk refinement is based on the counterfactual reasoning which is frequently used in standard equilibrium refinements. We start with a putative equilibrium (an equilibrium that is expected to be played) and test its stability against various off-equilibrium claims that could be made by the Sender in the cheap talk stage of the game. A statement is considered to be a " 'neologism' relative to the putative equilibrium if it is sent with zero probability in that equilibrium" (Matthews et al (1991), p.254). Credible neologisms upset the putative equilibrium.

The language of communication in Farrell's model (1985) is as follows:

The set of messages M consists of all subsets of T, excluding $\emptyset$. For each non-empty subset m of T, a neologism is a statement, interpreted as 'My type is in m'.

The Receiver believes neologism 'm' whenever it is credible.

A neologism is **believed**, if and only if it causes the Receiver to adopt the following beliefs:

$$\beta^L(t|m) = \frac{\pi(t)p(t)}{\sum_{t' \in T} \pi(t')p(t')}, \text{ where } \begin{array}{l} p(t) = 1, \text{ if } t \in m \\ p(t) = 0, \text{ if } t \notin m. \end{array}$$

Note that the belief revision rule is the same as $B^L(t|m)$ discussed in the previous section.

Definition 19 *A neologism 'm' is* ***credible*** *relative to an equilibrium* (τ, α, β) *of* $\Gamma^1(\pi)$ *if*

(1) $U^s(m, \alpha^*|t) > U^s(\tau, \alpha|t)$, $\forall t \in m$

(2) $U^s(m, \alpha^*|t) \leq U^s(\tau, \alpha|t)$, $\forall t \notin m$, *where* $\alpha^* \in \Sigma^0(\beta^L(t|m))$ *is an element of the set of the Receiver's optimal responses given his beliefs* $\beta^L(t|m)$.

Thus, a credible neologism has the following meaning: "My type is in m, and if my type wasn't in m, I would have preferred to stick with the equilibrium instead of sending this message".

Finally, an equilibrium is **neologism-proof**, if there are no neologisms credible relative to it.

Example 2: Consider the following game (Matthews et al (1991), p.262).

	A_1	A_2	A_3
t_1	3,3	1,0	2,2
t_2	1,0	0,3	2,2

$T = \{t_1, t_2\}$ and $A = \{A_1, A_2, A_3\}$ are the type and the action sets, respectively. The priors are $\pi(t_1) = \pi(t_2) = 1/2$. Message $\{t_1\}$ is a neologism relative to the no-communication equilibrium A_3. It is credible because only the Sender's type t_1 can get strictly higher payoff by inducing the Receiver to deviate from the no-communication equilibrium.

3.3.2 Announcement-Proof Equilibria

In the Theory of Announcement-Proofness (Matthews et al (1991)) the credibility of an off-equilibrium message is determined by considering the entire set of possible statements that could be used by the Sender to upset a putative equilibrium. The corresponding language of communication is relatively more complex:

$d = (\delta, D)$ is the Sender's **announcement strategy** in $\Gamma^1(\pi)$, where $D \subseteq T$ is the non-empty subset of deviant types, and $\delta : D \to \Delta M$ is a 'talking' strategy for the deviant types. Deviant types are the ones who wish to deviate from a putative equilibrium.

An **announcement** is a pair (m, d), where $m \in \delta(D)$, and $\delta(D)$ is the set of messages sent with positive probability by the types in D according to the 'talking' strategy δ.

The Receiver believes the announcement whenever it is credible.

An announcement is **believed,** if and only if it causes the Receiver to adopt the following beliefs:

$$\tilde{\beta}^L(t|(m,d)) = \frac{\pi(t)p(t)}{\sum_{t' \in T} \pi(t')p(t')}, \text{where} \begin{array}{l} p(t) = \delta(m|t), \text{ if } t \in D \\ p(t) = 0, \text{ if } t \notin D. \end{array}$$

Note that the belief revision rule is the same as $\tilde{B}^L(t|m)$ discussed in the previous section.

Definition 20 *The announcement* (m, d) *is* ***credible*** *relative to the equilibrium* (τ, α, β) *of* $\Gamma^1(\pi)$ *if*

(1′) $U^s((m,d), \alpha^*|t) \geq U^s(\tau, \alpha|t), \forall t \in D$ *(strict for some* $t \in D$*)*, $m \in \delta(\{t\})$

(2′) $U^s((m,d), \alpha^*|t) < U^s(\tau, \alpha|t), \forall t \in T \backslash D$

(3′) $U^s((m,d), \alpha^*|t) \geq U^s((m'',d), \alpha''|t)$, $\forall t \in D$, $m \in \delta(\{t\})$, $m'' \in \delta(D) \backslash \{m\}$, *where* $\alpha^* \in \Sigma^0(\tilde{\beta}^L(t|(m,d))$ *and* $\alpha'' \in \Sigma^0(\tilde{\beta}^L(t|(m'',d))$

(4′) *if* $\exists d' = (\delta', D')$ *that satisfies* (1′) − (3′) *relative to the same equilibrium, then* $U^s((m,d), \alpha^*|t) \geq U^s((m',d'), \alpha'|t), \forall t \in D \cap D', m \in \delta(\{t\}), m' \in \delta'(\{t\}), \alpha' \in \Sigma^0(\tilde{\beta}^L(t|(m',d'))$.

Condition (1′) is that deviant types should always prefer the announcement to the putative equilibrium. Condition (2′) is that non-deviant types should prefer equilibrium to the announcement. Condition (3′) says that each deviant type should prefer the message that is optimal for that type. Condition (4′) says that if there exist two announcement strategies that both satisfy (1′), (2′), (3′), the one which is at least weakly dominated by the other should not be credible. In Matthews et al (1991), condition (1′) must hold even for pessimistic deviant types and condition (2′) must hold even for optimistic non-deviant types of the Sender, provided that there could be more than one optimal response to any announcement. The interpretation of credible announcement is this: "My type is in D, and I am sending message m according to strategy $\delta(.|t)$. If I had been type t' in D, I would have made an announcement that differed only in so far as m would have been chosen according to strategy $\delta(.|t')$. If my type had not been in D, I would not have used announcement strategy δ." (Matthews et al (1991), p.259)

An equilibrium is **announcement proof**, if there are no announcements that are credible relative to it.

Example 3: Consider the following game (Matthews et al (1991), p.256) and suppose that the prior beliefs are $\pi(t_1) = \pi(t_2) = 1/2$

	A_1	A_2	A_3
t_1	4,3	3,0	1,2
t_2	3,0	4,3	1,2

In this game, there are no credible neologisms that could upset the no-communication equilibrium. Consider the following announcement (m, d): $m \in \{\{t_1\}, \{t_2\}\}$, $D = \{t_1, t_2\}$, $\delta(t_1) = \{t_1\}, \delta(t_2) = \{t_2\}$. The announcement (m, d) is credible relative to the no-communication equilibrium A_3, since both types at least weakly prefer to deviate from it (condition (1′)) and each type chooses the statement that induces the Receiver to take the Sender's preferred action (condition (3′)). Note, that there are no credible announcements, as well as credible neologisms, that could upset either one of the communication equilibrium outcomes A_2, A_3.

In what follows, we will explicitly specify what the players need to know about the game and each other in order to be able to infer whether a particular message m is credible neologism and/or credible announcement.

3.4 Two-Player Epistemic Model for the SRCT Game

3.4.1 The Belief System

Let Ω denote the set of all possible states of the World, where a typical element of the set $\omega \in \Omega$ is a complete description of all information that the players have in the beginning of the game (*before* the Sender learns his type t). Let $P^i_\omega(.)$ denote the probability distribution on Ω induced by player i's Theory (about his opponent' actions, beliefs and rationality) at state ω. The lower case $p^i_\omega(E)$ denotes the probability that player i assigns to event E at state ω, where E is a subset of Ω.

The Receiver's conjectures φ^r_ω at state ω is a probability distribution on M conditional on T (i.e. $\varphi^r_\omega \in [\Delta M]^T$) and the Sender's conjecture φ^s_ω at state ω is a probability distribution on A conditional on M (i.e. $\varphi^s_\omega \in [\Delta A]^M$).

Let $\tau_{\omega_k} \in [\Delta M]^T$ be the Sender's talking strategy at state ω_k, then the Receiver's conjecture $\varphi^r_\omega(m|t)$ can be specified as follows:

$$\varphi^r_\omega(m|t) = \sum_{\omega_k \in \Omega} p^r_\omega(\omega_k)\, \tau_{\omega_k}(m|t)$$

Also, let $\alpha_{\omega_k} \in [\Delta A]^M$ be the Receiver's action strategy at state ω_k, then the Sender's conjecture $\varphi^s_\omega(a|m)$ can be specified as follows:

$$\varphi^s_\omega(a|m) = \sum_{\omega_k \in \Omega} p^s_\omega(\omega_k)\, \alpha_{\omega_k}(a|m)$$

Assuming that the Receiver's prior beliefs π_ω about the Sender's type $t \in T$ are commonly known among the players at every state ω, the Receiver's belief-revision function β_ω, where $\beta_\omega \in [\Delta T]^M$, can be defined as follows:

$$\beta_\omega(t|m) = \frac{\pi_\omega(t)\,\varphi^r_\omega(m|t)}{\sum\limits_{t' \in T} \pi_\omega(t')\,\varphi^r_\omega(m|t')},$$

for all $t \in T$ and all $m \in M$, such that $\varphi^r_\omega(m|t) > 0$ for some t.

Denote by $[m] \equiv \{\omega \in \Omega : \tau_\omega(m|t) > 0 \text{ for some } t \in T\}$ the event that the Sender announces message m for some realization of his type $t \in T$.

3.4.2 The Knowledge Operator

For every player i, where $i \in \{s, r\}$, define the knowledge operator $K^i : 2^\Omega \to 2^\Omega$. For every event E, K^iE is the set of all $\omega \in \Omega$ such that player i knows E (i.e. $p^i_\omega(E) = 1$). We will assume that the operator K^i has the following properties:

[$K1$] $K^iE \subset E$ ('*non-delusion*')

[$K2$] $E \subset F$ implies $K^iE \subset K^iF$ (*Modus Ponens*)

[$K3$] $K^iE \subset K^iK^iE$ ('*knowing that you know*')

[$K4$] $\neg K^iE \subset K^i\neg K^iE$ ('*knowing that you don't know*')

Assumption [$K1$] implies that player i cannot know something that is not true. For instance, at some state $\omega \in \Omega$, the Sender cannot know that the Receiver interprets message m literally unless the Receiver's beliefs are indeed consistent with the literal interpretation of message m. Assumption [$K2$] implies that if player i ever discovers that E then he must also know F. Assumption [$K3$] means that if player i knows E then he always knows that he knows E and assumption [$K4$] means that if player i does not know E then he always knows that he does not know E.

Define $MKE \equiv K^iE \cap ... \cap K^jE$, where the superscripts i, j refer to different players. If $\omega \in MKE$, then event E is mutually known at state ω. Define $CKE \equiv MKE \cap MK(MKE) \cap MK(MK(MKE)) \cap ...$; if $\omega \in CKE$, call E commonly known at ω.

In what follows, we will assume that at every state $\omega \in \Omega$ the game itself as well as the epistemic model are commonly known among the players.

3.4.3 Bayesian Rationality

The standard definition of extensive form Bayesian rationality requires that each player chooses a strategy which is a best response to his beliefs at each information set that is not precluded by the very same strategy. Also, at each decision node, the players' conjectures should be consistent with all information they have and upon arrival of new information they should use some version of Bayes rule to update their beliefs. Thus, some information sets may never be reached because there are no consistent beliefs that support a strategy profile that reaches these nodes with positive probability.

In a SRCT Game, any information set of the Receiver could be reached during the pre-play communication from any $t \in T$. Therefore, in the absence of any restrictions on how the players interpret messages sent in the cheap-talk stage of the game, any $\tau \in \Im^s$ is rationalizable.

Let $A^*(m) \equiv \{a^* : argmax_{a \in A} \sum_{t \in T} \beta(t|m)u^r(a,t)\}$ be the set of Receiver's optimal actions in response to message m and let $A^* \equiv \bigcup_{m \in M} A^*(m)$ be the set of Receiver's optimal responses to all messages in M.

Suppose that both players, Sender and Receiver, are Bayesian rational and denote by $R^r(0) \equiv \{\alpha^* \in [\Delta A^*]^M\}$ the set of the Receiver's Bayesian rational 'action' strategies and by $R^s(0) \equiv \{\tau^* \in [\Delta M]^T\}$ the set of the Sender's Bayesian rational 'talking' strategies.

Also, let $R^r(1)$ denote the set of Receiver's optimal 'action' strategies consistent with the Receiver's knowing that the Sender is Bayesian rational and $R^s(1)$ denote the set of the Sender's optimal 'talking' strategies consistent with the Sender's knowing that Receiver is Bayesian rational.

Formally, for any $n \in Z_+$, where n corresponds to the number of levels of mutual knowledge of rationality among the players, $R^i(n)$, $\varphi^i(n)$, where $i \in \{r, s\}$, can be defined inductively as follows:

[R1] $R^s(n+1) \equiv \{\tau \in R^s(n) : \exists \varphi^s(n+1) \in co(R^r(n))$, such that τ is optimal against $\varphi^s(n+1)\}$

[R2] $R^r(n+1) \equiv \{\alpha \in R^r(n) : \exists \varphi^r(n+1) \in co(R^s(n))$, such that α is optimal against $\varphi^r(n+1)\}$

where $co(R^i(n))$ is the convex hull of player i's mixed strategies and $\varphi^i(n)$ is a player i's nth order conjecture profile.

Also, let $\bar{R}^i(n) \equiv \bigcap_{k \leq n} R^i(k)$ for $i \in \{r, s\}$. Then, whenever $\bar{R}^r(1) \neq \emptyset$ and $\bar{R}^s(1) \neq \emptyset$, we can define

$$MKR \equiv \{\omega \in \Omega : \tau_\omega \in \bar{R}^s(1)\} \cap \{\omega \in \Omega : \alpha_\omega \in \bar{R}^r(1)\}$$

which represents the event that Bayesian rationality is mutually known among the players and where R denotes the event that both players are Bayesian rational. Generally, whenever $\bar{R}^r(n) \neq \emptyset$ and $\bar{R}^s(n) \neq \emptyset$, we can define

$$MK^nR \equiv \{\omega \in \Omega : \tau_\omega \in \bar{R}^s(n)\} \cap \{\omega \in \Omega : \alpha_\omega \in \bar{R}^r(n)\}$$

which represents the event that Bayesian rationality is nth order mutually known among the players.

Let $CK(\Gamma^1)$ denote the event that a particular game Γ^1 is commonly known among the players. Whenever $CK(\Gamma^1) \cap MK^nR \cap [m] \neq \emptyset$, it implies that there is a state ω, where the nth order mutual knowledge of Bayesian rationality is consistent with the Sender's announcing m in the game Γ^1. In the absence of any restrictions on the players' beliefs, the set $CK(\Gamma^1) \cap MK^nR \cap [m]$ is non-empty for any $m \in M$ and for any $n \in Z_+$.

3.4.4 Common Language

Suppose that we want to describe a conventional interpretation of messages sent in the cheap talk stage of the game. For any $\mathcal{M} \subseteq M$, define

$$L(\mathcal{M}) \equiv \left\{\omega \in \Omega : \beta_\omega(t|m) = B^L(t|m) \text{ for all } m \in \mathcal{M}\right\}$$

and

$$\tilde{L}(\mathcal{M}) \equiv \left\{\omega \in \Omega : \ \beta_\omega(t|m) \ = \ \tilde{B}^L(t|m) \text{ for all } m \in \mathcal{M} \right\}$$

where the classes of conditional probability distributions $B^L(t|m)$ and $\tilde{B}^L(t|m)$ were defined in Section 3.2. The set $L(m)$ (as well as $\tilde{L}(m)$) is a collection of the states of the World, where the Receiver's conjectures are consistent with the literal interpretation of message m.

Whenever $CK(\Gamma^1) \cap MK^n R \cap L(m) \neq \emptyset$, it implies that there is a state ω, where the literal interpretation of message m in the game Γ^1 is consistent with the nth order mutual knowledge of rationality among the players. Note that any interpretation of message m such that $\tau_{\cdot\omega}(m|t) = 0$ for all $t \in T$ would be consistent with any level of mutual knowledge of rationality at state ω.

3.4.5 The Stiglitz Critique

Example 2 considered in Section 3.3 is often used to illustrate the *Stiglitz Critique* (also known as the *Stiglitz Paradox*) in the context of the cheap talk games.

Example 2 (revisited): Suppose that the priors are uniform and the pay-off matrix is as follows:

	A_1	A_2	A_3
t_1	3,3	1,0	2,2
t_2	1,0	0,3	2,2

The question is whether there are any credible statements that the Sender can make during the pre-play communication that would induce the Receiver to deviate from the no-communication equilibrium A_3. Here, the *Stiglitz Paradox* can be stated as follows: if it is commonly known that message $\{t_1\}$ is credible, then it is not credible. In order to see how we arrive at contradiction, let message $\{t_1, t_2\}$ be the no-communication equilibrium message, which is interpreted by the Receiver as uninformative, and suppose that the credibility of message $\{t_1\}$ is commonly known. Then, it should be commonly known that the Receiver would interpret message $\{t_1\}$ literally. But then, the Receiver should also know that only type t_2 of the Sender would announce $\{t_1, t_2\}$ and, therefore, his best response to message $\{t_1, t_2\}$ would be to choose A_2. Knowing that, both types of the Sender would then prefer to announce $\{t_1\}$, which would imply that the Receiver should not believe $\{t_1\}$ and message $\{t_1\}$ is not credible.

Using the epistemic model developed in this Section, we can show that at any state of the world, where both players are rational and the game described in Example 2 together with the literal interpretation of message $\{t_1\}$ are commonly known, either the Receiver or the Sender cannot know that his opponent is rational.

Proposition 21 *Any state $\omega \in \Omega$, where the game described in Example 2 is commonly known and $\omega \in CKL(\{t_1\}) \cap R^r \cap R^s$, must be such that $\omega \notin MKR$.*

Proof.

First of all, note that common knowledge of the literal interpretation of message $\{t_1\}$ is not sufficient to imply that Receiver should interpret any message $m \neq \{t_1\}$ as a signal than could be sent only by type t_2 of the Sender.

At any $\omega \in CKL(\{t_1\})$, $\beta_\omega(t_1|\{t_1\}) = \frac{\varphi^r_\omega(\{t_1\}|t_1)}{\varphi^r_\omega(\{t_1\}|t_1)+\varphi^r_\omega(\{t_1\}|t_2)} = 1$ and $\beta_\omega(t_2|\{t_1\}) = \frac{\varphi^r_\omega(\{t_1\}|t_2)}{\varphi^r_\omega(\{t_1\}|t_1)+\varphi^r_\omega(\{t_1\}|t_2)} = 0$, which implies that $\varphi^r_\omega(\{t_1\}|t_2) = 0$ and $\varphi^r_\omega(\{t_1\}|t_1) > 0$. Then, at any $\omega \in CKL(\{t_1\})$, the Receiver believes that type t_2 of the Sender never announces message $\{t_1\}$, i.e. for any ω_k, such that $p^r_\omega(\omega_k) > 0$, $\tau_{\omega_k}(\{t_1\}|t_2) = 0$ and $\tau_{\omega_k}(m|t_2) > 0$ for some $m \in M \setminus \{t_1\}$. The latter does not imply that $\tau_{\omega_k}(m|t_1) = 0$.

<u>Step I</u>: At any state of the world, where literal interpretation of message $\{t_1\}$ is commonly known and, in addition, the Receiver is rational and the Sender knows that Receiver is rational, the type t_2 of the Sender who sends any message that could distinguish him from type $t_1$cannot be rational.

Consider two states of the world ω_1 and ω_2, as described in Table 1, such that $\omega_1, \omega_2 \in CKL(\{t_1\}) \cap MKR^r$. The Receiver' Theory is such that $p^r(\omega_1) = 1 - \varepsilon$ and $p^r(\omega_2) = \varepsilon$, i.e. at states ω_1 and ω_2, the Receiver assigns probability $(1-\varepsilon)$ to him being at state ω_1 and he assigns probability ε to him being at state ω_2. The Sender's Theories at states ω_1, ω_2

are such that $p^{s}_{\omega_1}(\omega_1) = 1$ and $p^{s}_{\omega_2}(\omega_2) = 1$, i.e. at each state he knows exactly what state he is at. Note that in Table 1, we simply picked two states where type t_2 of the Sender chooses at least one message other than $\{t_1\}$ with positive probability.

Table 1

	R1: $\begin{cases} \alpha(A_1 \mid \{t_1\}) = 1 \\ \alpha(A_2 \mid \{t_2\}) = 1 \\ \alpha(A_3 \mid \{t_1, t_2\}) = 1 \end{cases}$
S1: $\begin{cases} \tau(\{t_1\} \mid t_1) = 1 \\ \tau(\{t_2\} \mid t_2) = 1/2 \\ \tau(\{t_1, t_2\} \mid t_2) = 1/2 \end{cases}$	$1, 1-\varepsilon$ (ω_1)
S2: $\begin{cases} \tau(\{t_1\} \mid t_1) = 1/2 \\ \tau(\{t_2\} \mid t_2) = 1/2 \\ \tau(\{t_1, t_2\} \mid t_1) = 1/2 \\ \tau(\{t_1, t_2\} \mid t_2) = 1/2 \end{cases}$	$1, \varepsilon$ (ω_2)

The imputed beliefs of the Receiver conditional on M are as follows: for $k \in \{1, 2\}$

$$\begin{cases} \beta_{\omega_k}(t_1 \mid \{t_1\}) = 1 \\ \beta_{\omega_k}(t_1 \mid \{t_2\}) = 1 \\ \beta_{\omega_k}(t_1 \mid \{t_1, t_2\}) = \frac{\frac{1}{2}\varepsilon}{\frac{1}{2}\varepsilon + \frac{1}{2}} = \frac{\varepsilon}{1+\varepsilon} \\ \beta_{\omega_k}(t_2 \mid \{t_1, t_2\}) = \frac{\frac{1}{2}}{\frac{1}{2}\varepsilon + \frac{1}{2}} = \frac{1}{1+\varepsilon} \end{cases}$$

For $\varepsilon \in \left[\frac{1}{2}, 1\right]$, the Receiver's optimal response to message $\{t_1, t_2\}$ is A_3 and the optimal responses to messages $\{t_1\}$ and $\{t_2\}$ are A_1 and A_2, respectively. So, given that $\varepsilon \in \left[\frac{1}{2}, 1\right]$, the Receiver is rational at both states ω_1 and ω_2, while the Sender is irrational at both states.[3]

<u>Step II</u>: At any state of the world, where literal interpretation of message $\{t_1\}$ is commonly known, both players are rational and the Sender knows that the Receiver is rational, the Receiver cannot know that the Sender is rational.

Suppose that at states ω_1 and ω_2, described in Table 2, the Sender never announces $\{t_2\}$, conditional on his type being t_2, but he can announce either $\{t_1\}$ or $\{t_1, t_2\}$ with positive probability, conditional on his type being t_1 (see Table 2 below)

Table 2

	$R1$: $\begin{cases} \alpha(A_1 \| \{t_1\}) = 1 \\ \alpha(A_3 \| \{t_1, t_2\}) = 1 \end{cases}$
$S1$: $\begin{cases} \tau(\{t_1\} \| t_1) = 1 \\ \tau(\{t_1, t_2\} \| t_2) = 1 \end{cases}$	$1, 1 - \varepsilon$ (ω_1)
$S2$: $\begin{cases} \tau(\{t_1\} \| t_1) = 1/3 \\ \tau(\{t_1, t_2\} \| t_1) = 2/3 \\ \tau(\{t_1, t_2\} \| t_2) = 1 \end{cases}$	$1, \varepsilon$ (ω_2)

[3]Note that if $\varepsilon \in \left[0, \frac{1}{2}\right]$, the Receiver is irrational at both states ω_1 and ω_2, implying that the Sender cannot know that he is rational, which would contradict our assumption.

The imputed beliefs of the Receiver conditional on M are as follows: for $k \in \{1,2\}$

$$\begin{cases} \beta_{\omega_k}(t_1 | \{t_1\}) = 1 \\ \beta_{\omega_k}(t_1 | \{t_1, t_2\}) = \frac{\frac{2}{3}\varepsilon}{\frac{2}{3}\varepsilon + 1} \\ \beta_{\omega_k}(t_2 | \{t_1, t_2\}) = \frac{1}{\frac{2}{3}\varepsilon + 1} \end{cases}$$

For $\varepsilon \in \left[\frac{3}{4}, 1\right]$, the Receiver's optimal response to message $\{t_1, t_2\}$ is A_3 and his optimal response to message $\{t_1\}$ is A_1. So, for $\varepsilon \in \left[\frac{3}{4}, 1\right]$, the Receiver is rational at both states ω_1 and ω_2. The Sender is rational at state ω_1 but not at state ω_2. Moreover, at state ω_1, the Sender knows that the Receiver assigns probability ε to him being irrational (as state ω_1 the Sender knows that Receiver does not know that the Sender is rational).

So, we have shown that at any state of the world, where both Sender and Receiver are rational, the Sender knows that the Receiver is rational and the literal interpretation of message $\{t_1\}$ is commonly known, the Receiver cannot know that the Sender is rational. Thus, in Example 2, literal interpretation of message $\{t_1\}$ is inconsistent with the mutual knowledge of rationality among the players. ■

One important remark that has to be made here is that the inconsistency between the literal interpretation of message $\{t_1\}$ and the mutual knowledge of rationality among the players in the game in Example 2 is present irrespective of whether we look at it as a two-player or as a $(T+1)$-player extensive form game.

Consider a $(T+1)$-player representation of the SRCT game, where there is only one Receiver, but the Sender(s) at different information sets $t \in T$ of the extensive form game $\Gamma^1(\pi)$ are viewed as independent players. Define a state space $\tilde{\Omega}$ as a cross product of the players' belief space Ω and the Sender's type space T, i.e. $\tilde{\Omega} \equiv \Omega \times T$. A generic element of $\tilde{\Omega}$ is $\tilde{\omega} \equiv (\omega, t)$, where any realization of $t \in T$ can be viewed as a state of Nature observed *only* by the Sender.

Consider the game described in Example 2. Suppose that the Receiver is rational and that he believes that both t_1 and t_2 are equally likely, but he does not know exactly what state he is at, i.e. $p^r(\tilde{\omega}_1) = \frac{1-\varepsilon}{2}$, $p^r(\tilde{\omega}_2) = \frac{1}{2}$ and $p^r(\tilde{\omega}_3) = \frac{\varepsilon}{2}$ (See Table 3 below).

Table 3

	R1 $\begin{cases} \alpha(A_1 \vert \{t_1\}) = 1 \\ \alpha(A_3 \vert \{t_1, t_2\}) = 1 \end{cases}$
S1 $\tau(\{t_1\} \vert t_1) = 1$	$1, \frac{1-\varepsilon}{2}$ $(\tilde{\omega}_1)$
S2 $\tau(\{t_1, t_2\} \vert t_2) = 1$	$1, \frac{1}{2}$ $(\tilde{\omega}_2)$
S3 $\begin{cases} \tau(\{t_1\} \vert t_1) = 1/3 \\ \tau(\{t_1, t_2\} \vert t_1) = 2/3 \end{cases}$	$1, \frac{\varepsilon}{2}$ $(\tilde{\omega}_2)$

As one can see from Table 3, whenever the literal interpretation of message $\{t_1\}$ is commonly know and both players are rational, then either the Sender's or the Receiver's rationality cannot be mutually known.

In application to the game in Example 2, both theories of Neologism and Announcement Proofness stipulate that Receiver should believe that message $\{t_1\}$ has been sent by the Sender's type t_1 (i.e. that $\{t_1\}$ is credible), because it would be irrational for the Sender's type t_1 not to send such message knowing that Receiver would interpret it literally, and, on the other hand, it would be irrational for the Sender's type t_2 to send message $\{t_1\}$ knowing that Receiver would interpret it literally. Thus, the Receiver cannot know that message $\{t_1\}$ is credible without also knowing that the Sender is rational. The argument underlying the Stiglitz critique, which implicitly assumes an epistemic framework similar to the one described above, is that in the game in Example 2 it is impossible for the Receiver to know that the Sender is rational and, at the same time, continue to believe that message $\{t_1\}$ is credible. Thus, it appears that by applying the theories of Neologism and Announcement Proofness to the game in Example 2 we inevitably arrive at contradiction. The reason for this is because in the framework of our epistemic model the possibility that the non-deviant types[4] could have sent an off-equilibrium message cannot be ruled out unless the Receiver considers all implications of him believing every off-equilibrium message and the latter being commonly known among the players.

However, the common languages in both theories implicitly contain certain assumptions that eliminate this inconsistency. Indeed, the inconsistency pointed out by Stiglitz is eliminated if an explicit distinction between equilibrium and off-equilibrium messages becomes part of the common language used by the players (and is, therefore, itself commonly known).

3.5 Epistemic Model of Fixed-Equilibrium Rationalizability

First, we will introduce the concept of a *derived game*, similar to the one developed in the signalling games literature (see Sobel et al (1990)). Suppose that now the Sender can either choose some (fixed) pay-off and end the game or, alternatively, he can choose to participate in a derived game $\hat{\Gamma}(\pi)$, where the pay-off matrix is exactly the same as in the original game $\Gamma^1(\pi)$, but the set of messages available to the Sender is only a subset of messages of the original game. Consider a putative equilibrium σ^e of $\Gamma^1(\pi)$. Let $\hat{\Gamma}^1_{\sigma^e}(\pi)$ be a derived game, where the set $\hat{M}$ of messages available to the Sender consists of all messages that are not in the support of the putative equilibrium 'talking' strategy τ^e, i.e. $\hat{M} \equiv M \backslash support\,(\tau^e)$

Consider Example 2 again and suppose that the Sender has a choice of either taking the pay-off equal to 2 (which he would have received in the no-communication equilibrium of the game $\Gamma^1(\pi)$ in Example 2), or he can choose to play the derived game $\hat{\Gamma}^1_{\sigma^e}(\pi)$, where he can only send some message identifying himself either as type t_1 or as type t_2. Denote by m^e the message that automatically triggers the putative equilibrium response from the Receiver, i.e. the Receiver's action set conditional on the Sender's choosing message m^e is a singleton $\{A_3\}$. Then, the Receiver's conjecture φ^{τ}_{ω} is a probability distribution on the

[4]The types that prefer to stick with the equilibrium given that Receiver believes every off-equilibrium message.

set $\hat{M}$ conditional on T and the Sender's conjecture $\varphi_{\tilde{\omega}}^{s}$ is a probability distribution on A conditional on $\hat{M}$.

Consider three states of the World, $\tilde{\omega}_1, \tilde{\omega}_2$ and $\tilde{\omega}_3$, such that the players' beliefs are as described in Table 4 below.

Table 4

	$\begin{cases} \alpha\left(A_1 \mid \{t_1\}, \hat{\Gamma}\right) = 1 \\ \alpha\left(A_2 \mid \{t_2\}, \hat{\Gamma}\right) = 1 \end{cases}$ and $\left\{ \alpha(A_3 \mid m^e) = 1 \right.$
$S1$ $\begin{cases} \tau(\{t_1\}, \hat{\Gamma} \mid t_1) = 1 \\ \tau(m^e \mid t_1) = 0 \end{cases}$	$1, \frac{1}{2}$ $(\tilde{\omega}_1)$
$S2$ $\begin{cases} \tau\left(\{t_2\}, \hat{\Gamma} \mid t_2\right) = 0 \\ \tau(m^e \mid t_2) = 1 \end{cases}$	$1, \frac{1-\varepsilon}{2}$ $(\tilde{\omega}_2)$
$S3$ $\begin{cases} \tau(\{t_1\}, \hat{\Gamma} \mid t_2) = 1 \\ \tau(m^e \mid t_2) = 0 \end{cases}$	$1, \frac{\varepsilon}{2}$ $(\tilde{\omega}_2)$

Suppose that the Receiver knows that the Sender is rational (i.e. $\varepsilon = 0$). Then, he should expect the Sender's type t_1 to choose $\hat{\Gamma}$ and announce message $\{t_1\}$, and he should expect the Sender's type t_2 to choose the putative equilibrium pay-off. Thus, if $\varepsilon = 0$, both players are rational and there is also common knowledge of rationality, which is consistent with the common knowledge of the literal interpretation of message $\{t_1\}$.

One important point has to be made here is that neither the mutual knowledge of rationality nor the common knowledge of the literal interpretation of message $\{t_1\}$ are necessary for the rational Sender's type t_1 to announce $\{t_1\}$ and for the rational Receiver to respond to $\{t_1\}$ by playing A_1 in the game in Example 2.

In order to see why the mutual knowledge of rationality is not necessary for the rational Sender's type t_1 to announce $\{t_1\}$ and for the rational Receiver to respond by playing A_1, note that even if Receiver is not certain that the Sender is rational and $\varepsilon < \frac{1}{2}$, his best response to message $\{t_1\}$ is A_1 and, knowing that, the rational type t_1 of the Sender should announce $\{t_1\}$ (see Table 4)

In order to see why the common knowledge of the literal interpretation of message $\{t_1\}$ is not necessary either, consider the subset of the states of the world described in Table 5.

Table 5

	R1: $\begin{cases} \alpha\left(A_1 \mid \{t_1\}, \hat{\Gamma}\right) = 1 \\ \alpha\left(A_2 \mid \{t_2\}, \hat{\Gamma}\right) = 1 \\ \alpha\left(A_3 \mid m^e\right) = 1 \end{cases}$	R2: $\begin{cases} \alpha\left(A_3 \mid \{t_1\}, \hat{\Gamma}\right) = 1 \\ \alpha\left(A_2 \mid \{t_2\}, \hat{\Gamma}\right) = 1 \\ \alpha\left(A_3 \mid m^e\right) = 1 \end{cases}$
S1: $\begin{cases} \tau(\{t_1\}, \hat{\Gamma} \mid t_1) = 1 \\ \tau(m^e \mid t_1) = 0 \end{cases}$	$1-\nu, \frac{1}{2}$ ($\tilde{\omega}_1$)	$\nu, \frac{1}{2}$ ($\tilde{\omega}_4$)
S2: $\begin{cases} \tau\left(\{t_2\}, \hat{\Gamma} \mid t_2\right) = 0 \\ \tau(m^e \mid t_2) = 1 \end{cases}$	$1, \frac{1-\varepsilon}{2}$ ($\tilde{\omega}_2$)	$0,0$ ($\tilde{\omega}_5$)
S3: $\begin{cases} \tau(\{t_1\}, \hat{\Gamma} \mid t_2) = 1 \\ \tau(m^e \mid t_2) = 0 \end{cases}$	$1-\xi, \frac{\xi}{2}$ ($\tilde{\omega}_3$)	$\xi, \frac{1}{2}$ ($\tilde{\omega}_6$)

At states $\tilde{\omega}_1$ and $\tilde{\omega}_4$, the Sender is rational even if $\nu > 0$, i.e. if he believes that Receiver may not always interpret message $\{t_1\}$ literally. Then, at state $\tilde{\omega}_1$, the Receiver assigns probability $\frac{1}{2}$ to the Sender's type t_1 assigning probability $(1-\nu)$ to the event that Receiver will respond to message $\{t_1\}$ by playing A_3. The latter implies that at state $\tilde{\omega}_1$ the literal interpretation of message $\{t_1\}$ is not commonly known. (However, note that at states $\tilde{\omega}_3$ and $\tilde{\omega}_6$, the Sender can only be rational if $\xi = 0$, i.e. if he knows that he is at the state where Receiver does not interpret message $\{t_1\}$ literally.)

The sufficient epistemic conditions for a particular off-equilibrium message to be a credible neologism or a credible announcement will be formulated and proven in the next section.

3.6 Sufficient Epistemic Conditions for Credible Neologism and Credible Announcement

Consider a putative equilibrium σ^e of $\Gamma^1(\pi)$. Let $\hat{\Gamma}^1_{\mathcal{M},\sigma^e}(\pi)$ be a derived game, where the set of messages available to the Sender consists of all messages in $\mathcal{M} \subseteq \hat{M}$, where $\hat{M} \equiv M \backslash support\,(\tau^e)$. Let $\Gamma_{\mathcal{M},\sigma^e}$ denote a game where the players can either choose the putative equilibrium pay-offs by selecting (m^e, α^e), or they can choose to participate in the derived game $\hat{\Gamma}^1_{\mathcal{M},\sigma^e}(\pi)$.

Proposition 22 *Consider a putative equilibrium σ^e of $\Gamma^1(\pi)$ and an off-equilibrium message $m \in \hat{M}$. $CK(\Gamma_{m,\sigma^e} \cap L(m)) \cap MKR \cap [m] \neq \emptyset$, then message m is a **credible neologism** relative to the putative equilibrium σ^e.*

Proof. Suppose that $CK(\Gamma_{m,\sigma^e} \cap L(m)) \cap MKR \cap [m] \neq \emptyset$.

Then, at every state $\tilde{\omega} \in CK(\Gamma_{m,\sigma^e} \cap L(m)) \cap MKR \cap [m]$, $\alpha^*_{\tilde{\omega}} \in \overline{R}^r(1)$, where $\alpha^*_{\tilde{\omega}}(m) \in \Delta A^*(m)$ and $A^*(m) = \{a^* : argmax_{a \in A} \sum_{t \in T} \beta(t|m) u^r(a,t)$,where $\beta(t|m) \in B^L(t|m)\}$, and $\tau^*_{\tilde{\omega}} \in \overline{R}^s(1)$, where $\tau^*_{\tilde{\omega}}(m|t) > 0$ for every $t \in m$. Therefore, it must be the case that $U^s(m, \alpha^*(m)\,|t) \geq U^s(\tau^e, \alpha^e|t)$ for all $t \in m$. (Condition (1) of the Definition 19 holds)

Suppose that $\exists t' \notin m$, such that $U^s(m, \alpha^*(m)\,|t') > U^s(\tau^e, \alpha^e|t')$. Then, at every state $\tilde{\omega} \in CK(\Gamma_{m,\sigma^e} \cap L(m)) \cap MKR \cap [m]$, the Sender's 'talking' strategy $\tau'_{\tilde{\omega}}$ would be such that $\tau'_{\tilde{\omega}}(m|t') > 0$ and, therefore, the Receiver's best response $\alpha^*_{\tilde{\omega}} \in R^r(1)$ would be such that $\alpha^*_{\tilde{\omega}}(m) \notin \Delta A^*(m)$. $\otimes$ (Condition (2) of the Definition 19 holds). ■

Corollary 23 *Suppose that a putative equilibrium σ^e of $\Gamma^1(\pi)$ is **neologism-proof.** Then it must be the case that for every message $m \in \hat{M}$, $CK(\Gamma_{m,\sigma^e} \cap L(m)) \cap MKR \cap [m] = \emptyset$.*

The main distinction between the Theory of Neologism-Proofness and the Theory of Announcement-Proofness is that the latter allows the Sender to choose any message from the set of all off-equilibrium messages $\hat{M}$, which could potentially induce the Receiver to take the Sender's most preferred action. The syntactic differences between the languages of communication corresponding to these theories have been discussed in Section 3.2.

Proposition 24 *Consider a putative equilibrium σ^e of $\Gamma^1(\pi)$ and an off-equilibrium message $m \in \hat{M}$. Whenever $CK\left(\Gamma_{\hat{M},\sigma^e} \cap \tilde{L}\left(\hat{M}\right)\right) \cap MKR \cap [m] \neq \emptyset$, then message m is a **credible announcement** relative to the putative equilibrium σ^e.*

Proof. Suppose that $CK\left(\Gamma_{\hat{M},\sigma^e} \cap \tilde{L}\left(\hat{M}\right)\right) \cap MKR \cap [m] \neq \emptyset$.

Then at every state $\tilde{\omega} \in CK\left(\Gamma_{\hat{M},\sigma^e} \cap \tilde{L}\left(\hat{M}\right)\right) \cap MKR \cap [m]$, $\alpha^*_{\tilde{\omega}} \in \overline{R}^r(1)$, where $\alpha^*_{\tilde{\omega}}(m) \in \Delta\tilde{A}^*(m)$ and $\tilde{A}^*(m) = \{a^* : argmax_{a \in A} \sum_{t \in T} \beta(t|m) u^r(a,t)$, where $\beta(t|m) \in \bar{B}^L(t|m)\}$ and $\tau^*_{\tilde{\omega}} \in \overline{R}^s(1)$, where $\forall t \in m$, $\tau^*_{\tilde{\omega}}(m|t) > 0$. Therefore, $U^s(\tau^*, \alpha^*|t) \geq U^s(\tau^e, \alpha^e|t)$ for all $t \in m$ (Condition (1′) of the Definition 20 holds).

(By contradiction) Suppose that for some $t' \notin m$, $U^s(m, \alpha^*(m)|t') > U^s(\tau^e, \alpha^e|t')$, then it must be the case that $CK\left(\Gamma_{\hat{M},\sigma^e} \cap \tilde{L}\left(\hat{M}\right)\right) \cap MKR \cap [m] = \emptyset$ (Condition (2′) of the Definition 20 holds; the argument is the same as in Proposition 22)

(By contradiction) Suppose that for some $t' \in m \cap m'$, $U^s(m', \alpha^*(m')|t') > U^s(m, \alpha^*(m)|t')$, where $\alpha^*(m) \in \Delta\tilde{A}^*(m)$, $\alpha^*(m') \in \Delta\tilde{A}^*(m')$, but then
$CK\left(\Gamma_{\hat{M},\sigma^e} \cap \tilde{L}\left(\hat{M}\right)\right) \cap MKR \cap [m] = \emptyset$ (given that the Sender is rational and that he knows how the Receiver would respond to m and m', the Sender could not have chosen m over m'). (Condition (3′) of the Definition 20 holds)

(By contradiction) Suppose that for some $D' \subset T$, there is a 'talking' strategy τ', such that *support*$(\tau') \nsubseteq M^e$ and τ' is optimal against some α', such that $\alpha'(m) \in \Delta\tilde{A}^*(m)$

for any $m \in$ *support* (τ') and that for some $D'' \subset T$, there is a 'talking' strategy τ'', such that *support* $(\tau'') \not\subseteq M^e$ and τ'' is optimal against some α'', such that $\alpha''(m) \in \Delta\tilde{A}^*(m)$ for any $m \in$ *support* (τ''). Also, suppose that for both τ' and τ'' conditions $(1'), (2'), (3')$ of the Definition 20 hold. Then, whenever there is a Sender's type $t \in D' \cap D''$, such that $U^s(\tau', \alpha'|t) > U^s(\tau'', \alpha''|t)$, it must be the case that $\tau' \in \overline{R}^s(1)$, $\tau'' \notin \overline{R}^s(1)$ and, therefore, for any $m \in M$, such that $\tau''(m|t) > 0$, but $\tau'(m|t) = 0$, $CK\left(\Gamma_{\hat{M},\sigma^e} \cap \tilde{L}\left(\hat{M}\right)\right) \cap MKR \cap [m] = \emptyset$. (Condition $(4')$ of the Definition 20 holds). ■

Corollary 25 *Suppose that a putative equilibrium σ^e of $\Gamma^1(\pi)$ is* ***announcement-proof****. Then it must be the case that for every message $m \in \hat{M}$, $CK\left(\Gamma_{\hat{M},\sigma^e} \cap \tilde{L}\left(\hat{M}\right)\right) \cap MKR \cap [m] = \emptyset$.*

An equilibrium selection theory for the SRCT Game, where common knowledge of credible messages is fully consistent with the common knowledge of rationality among the players without asymmetric treatment of equilibrium and non-equilibrium messages, is discussed in the next section.

3.7 Credible Message Rationalizability

A 'non-equilibrium' cheap-talk refinement based on the assumption that rationality is common knowledge was introduced by Rabin (1990), who showed that existence and common knowledge of credible messages could reduce the set of rationalizable strategies and allow to predict when the meaningful communication could develop. The common language corresponding to Rabin's theory of credible message rationalizability is such that 'honesty' remains the focal policy.

Let $\mathfrak{X}$ be a collection of mutually exclusive subsets of T, $X_j \in \mathfrak{X}$.

Denote by $\tilde{A}^*(X_j)$ the set of the Receiver's optimal responses, given that his beliefs are concentrated on the types $t \in X_j$. Also, denote by $Y^*(X_j, \mathfrak{X})$ the set of the Sender's types that do not belong to the set X_j, but may want to imitate the types in X_j.

Definition 26 *$Y^*(X_j, \mathfrak{X})$ is the set of types in $T \backslash T_{\mathfrak{X}}$ (all types that are not in the type profile $\mathfrak{X}$), excluding any type t with the property that either:*

(1) $\tilde{A}^*(X_j) = \{a^* \in A : a^* \in \arg\min_{a \in A^*} u^s(a, t)\}$ *or*

(2) $\exists X_k \in \mathfrak{X} : u^s(a^*, t) < u^s(a, t) \ \forall a^* \in \tilde{A}^*(X_j), \forall a \in \tilde{A}^*(X_k)$.

Thus, $Y^*(X_j, \mathfrak{X})$ does not include the types that would obtain their worst possible payoff by imitating the types in X_j or who could do better by imitating the types that belong to some other subset X_k.

Definition 27 *Let $B(X_j, \mathfrak{X})$ be a set of probability distributions such that $\beta \in B(X_j, \mathfrak{X})$ is defined as follows:*

$$\beta(t|X_j) = \frac{\pi(t)p(t)}{\sum_{t' \in T} \pi(t')p(t')}, \ \textit{where} \begin{cases} p(t) = 1, \ \textit{if } t \in X_j \\ p(t) = [0,1], \ \textit{if } t \in Y^*(X_j, \mathfrak{X}), \\ p(t) = 0, \ \textit{otherwise} \end{cases}$$

Definition 28 *$\mathfrak{X}$ is a credible message profile if, for all $X_j \in \mathfrak{X}$,*

(1) $\forall t \in X_j$, $\tilde{A}^*(X_j) = \arg\max_{a \in A^*} u^s(a, t)$

(2) $\forall \beta \in B(X_j, \mathfrak{X})$, $\tilde{A}^*(X_j) = \arg\max_{a \in A^*} \sum_{t \in T} \beta(t|X_j) u^r(a, t)$

A message profile is considered credible if given that Receiver believes the literal meaning of the statements, the types sending these messages obtain their best possible payoffs and the messages are true enough in the sense that believing that some types other than the types in X_j might have sent the same message as the types in X_j does not affect the Receiver's optimal response to those messages.

Let $M(\mathfrak{X}^*)$ denote the maximal credible message profile. For any game $\Gamma^1(\pi)$, there exists a unique maximal credible message profile, which is the coarsest partition of the Sender's types who could send a credible message. And a Credible Message Equilibrium (CME), which is an equilibrium formed by the strategies that are consistent with the Theory of Credible Message Rationalizability, exists for all simple communication games (See Rabin (1990) for details).

Proposition 29 *For every message* $m \notin M(\mathfrak{X}^*)$, $CKR \cap CK\left(\tilde{L}(M)\right) \cap [m] = \emptyset$.

Proof. (By contradiction) Suppose that for some message $m \notin M(\mathfrak{X}^*)$, $CKR \cap CK\left(\tilde{L}(M)\right) \cap [m] \neq \emptyset$. Then, at every state $\omega \in CKR \cap CK\left(\tilde{L}(M)\right) \cap [m]$, $\alpha^*_{\tilde{\omega}} \in \overline{R}^r(\infty)$, where $\alpha^*_{\tilde{\omega}}(m) \in \triangle\tilde{A}^*(m)$, and $\tau^*_{\tilde{\omega}} \in \overline{R}^s(\infty)$, where $\forall t \in m$, $\tau^*_{\tilde{\omega}}(m|t) > 0$. Also, $CKR \cap CK\left(\tilde{L}(M)\right) \cap [m] \neq \emptyset$ implies that for any level of mutual knowledge of rationality neither of the types $t \notin m$ prefers to be the first, second or any higher order imitator of the types $t \in m$.

Then, all types $t \in m$ can send a credible message m consistent with the common knowledge of rationality and for any $m' \in M\backslash\{m\}$, such that $m' \supset t$ and $m' \subset m$, $u^s(a,t) > u^s(a',t)$, $\forall a \in \tilde{A}^*(m)$, $\forall a' \in \tilde{A}^*(m')$. But then it means that m must belong to the maximal credible message profile $M(\mathfrak{X}^*)$.$\otimes$ ■

Thus, the Theories of Neologism-Proofness and Announcement-Proofness impose weaker credibility requirements on the off-equilibrium messages than the Theory of Credible Message Rationalizability. Indeed, according to the Theories of Neologism-Proofness and Announcement Proofness, some equilibria could be 'defeated' by a message sent according to some 'talking' strategy which is not credible message rationalizable. Therefore, if we were to rank these refinements according to the relative strength of epistemic conditions reflecting the level of sophistication required for the players to be able to evaluate the credibility of an off-equilibrium message, we could refer to NPE and APE as strongly undefeated (compared to CME). So, while both approaches attempt to eliminate the implausible equilibria by restricting the set of allowable belief profiles, the Theory of Credible Message Rationalizability does that by introducing simple language of communication and imposing strong requirement on the level of the players' sophistication (common knowledge of rationality) at all states of the world, while the Theory of Neologism (or Announcement)-Proofness attempts to eliminate the implausible equilibria by introducing rather complicated common language of communication which includes explicit references to (each) putative equilibrium.[5]

3.8 Communication Equilibria

In this section, we will compare the collections of message that are credible neologisms and/or credible announcements with the collection of messages that constitute a credible message profile. Denote by $\mathcal{K}^*$ the message profile containing all messages that are credible neologisms relative to some equilibrium σ of $\Gamma^1(\pi)$ and denote by $\mathcal{A}^*$ the message profile containing all messages such that for each m there is an announcement (m, d), where $\delta(m|t) > 0$ for $\forall t \in D$, $D \subset T$, which is a credible relative to some equilibrium σ of $\Gamma^1(\pi)$.

Proposition 30 *In general, $\mathcal{K}^* \neq \mathcal{A}^* \neq M(\mathfrak{X}^*)$. For the games which have more than one PBE, $M(\mathfrak{X}^*) \subseteq \mathcal{A}^*$.*[6]

Thus, in a game $\Gamma^1(\pi)$, which has more than one equilibrium, the types that can send messages which constitute a (unique) maximal credible message profile should also be able to make credible announcements, therefore, if a putative equilibrium is not a CME, then it cannot be an APE either.

[5] It is possible to construct a credible message rationalizability theory that uses richer language and allows for more communication (See for example, Zapater (1991)).

[6] *For the proof, see Appendix.*

Corollary 31 *If an equilibrium is CME, but not APE, it must be the no-communication equilibrium.*

Corollary 32 *If $\Gamma^1(\pi)$ is a game of pure coordination (where for any $t \in T$, $argmax_{a \in A^*} u^s(t,a)$ $= argmax_{a \in A^*} u^r(t,a)$), then a separating CME is also APE and NPE.*

Corollary 31 states that whenever CME is not APE, communication is ineffective. Therefore, the Theory of Announcement-Proofness rules out exactly the same 'implausible' informative equilibrium outcomes as the Theory of Credible Message Rationalizability.

3.9 Discussion

The language of communication used by the players in a strategic interaction reflects how sophisticated they are and how much they know about each other. If the players share common background, one would expect them to coordinate successfully in any game of common interest. Thus, the plausibility of any cheap talk refinement is highly context dependent.

The Theories of Neologism and Announcement-Proofness follow the logic of forward induction, treating the deviations from the putative equilibrium as signals (not as trembles or mistakes). In the context of the cheap talk games, the common language assumption eliminates the ambiguity in the interpretation of deviations. Each player knows the meaning of each potential deviation, provided that he knows that his opponent is rational and that he also knows how each deviation would be interpreted by the Receiver given that certain (commonly known) conditions are satisfied. Because the Theory of Neologism- Proofness implicitly assumes that the Receiver is not certain whether the Sender chooses the 'best' deviation from a putative equilibrium, it seems more appropriate in the context, where the players do not have enough information about each other and therefore, the Receiver has to interpret the Sender's signals with caution.

By fixing the meaning of the putative equilibrium messages, we implicitly assume that the players' knowledge of their opponents' rationality is bounded.[7] Suppose that the Sender has not deviated from the equilibrium path although he had an opportunity to do so. Shouldn't the Receiver also interpret the absence of a deviation as a signal? As we have seen in Example 2, the mutual knowledge of rationality entails exactly this question, forcing the players to revise the meaning of the putative equilibrium message as well. In order to get around this problem and avoid the inconsistency between common knowledge of credible messages and mutual knowledge of rationality, the theory of fixed equilibrium rationalizability treats equilibrium and non-equilibrium states differently. The story often used to justify this approach is that the players "arrive at equilibrium behavior following the period of learning and experimentation", (Matthews et al. (1990), p.310) and once the equilibrium behavior becomes a 'routine', they attempt to conjecture what happens if they deviate from the equilibrium path. Although such explanation seems plausible in the context of a learning model, which presents a complete theory of play (on and off-the equilibrium), it seems less plausible when the theory of equilibrium selection hinges on the assumption of a commonly known language of communication.

The Theory of Neologism-Proofness is *coherent* (i.e. it implies the predicted behavior), if it is commonly known among the players that the equilibrium that is being played is **neologism-proof**. In other words, at all neologism-proof equilibrium states of the world, it should be commonly known that at every off-equilibrium state where some message $m \in \hat{M}$ is announced and the literal interpretation of message m in Γ_{m,σ^*} is commonly known, the rationality cannot be mutually known among the players (i.e. it should be commonly known that for any $m \in \hat{M}, CK(\Gamma_{m,\sigma^*} \cap L(m)) \cap MKR \cap [m] = \emptyset$). Similarly, the Theory of Announcement-Proofness is *coherent*, if it is commonly known among the players that the equilibrium that is being played is **announcement-proof** (i.e. it should be commonly known that for any $m \in \hat{M}, CK\left(\Gamma_{\hat{M},\sigma^*} \cap \tilde{L}\left(\hat{M}\right)\right) \cap MKR \cap [m] = \emptyset$). Thus, the coherence of both refinement theories requires that at every 'credible message proof - equilibrium' state of the world, it should be commonly known among the players that the common

[7]It has been shown (Reny (1993), Bicchieri (1993)), that the backward induction argument relies on the bounded rationality assumption as well.

knowledge of credible off-equilibrium messages is inconsistent with the mutual knowledge of rationality. Effectively, it means that the players must know virtually everything about off-equilibrium states and virtually nothing about equilibrium states in order to be able to choose between different equilibrium paths.

APPENDIX

Proof of Proposition 30:

To prove the first part of the Proposition, it is sufficient to show that $\mathcal{K}^* \neq M(\mathfrak{X}^*)$. (To see that $\mathcal{K}^* \neq \mathcal{A}^*$, see Matthews et al. (1991) for counterexamples).

Proof. First, we have to show that $\exists m$, such that $m \in \mathcal{K}^*$, but $m \notin M(\mathfrak{X}^*)$.

Consider a neologism K, which is credible relative to some σ. Suppose there exist some $t \notin \mathcal{K}^*$, such that $U^s(\sigma|t) > U^s(\sigma'|t)$ for any other σ' , including the one that might be induced by K ($t \notin K$). Suppose there exists $a^* = argmin_{a \in A}\ u^s(a,t) = argmax_{a \in A} \sum_{t \in T} \beta(t|\{t\})u^r(a,t)$, where $\{t\}$ is a message that allows the Receiver to identify the Sender as type t. If we sequentially delete all such messages that might reveal the Sender's identity (in fact announcing any such message would be a strictly dominant strategy for the Sender), then given that there is no other $K_k \in \mathcal{K} : u^s(a^*,t) < u^s(a,t)$, $\forall a^* \in A^*(K)$, $\forall a \in A^*(K_k)$ pooling with K might be the only remaining undominated strategy for the Sender , irrespective of the fact that $U^s(s|t) > U^s(s'|t)$ for any other σ' , including the one that might be induced by K. Therefore, K does not belong to a credible message profile.

We also have to show that $\exists m$, such that $m \in M(\mathfrak{X}^*)$, but $m \notin \mathcal{K}^*$ Consider neologisms K', K'', such that $U^s(\sigma_2|t) > U^s(\sigma_o|t)$ and $U^s(\sigma_2|t) > U^s(\sigma|t)$ for all t in K' and $U^s(\sigma_1|t) > U^s(\sigma_o|t)$ and $U^s(\sigma_1|t) > U^s(\sigma_2|t)$ for all t in K'' and also for all t in K' , $U^s(\sigma_1|t) > U^s(\sigma_o|t)$ and for all t in K', $U^s(\sigma_2|t) > U^s(\sigma_o|t)$, then none of these neologisms is credible relative to so , since condition (2) of the definition of credible neologism fails, but $\{\{K''\},\{K'\}\}$ is a credible message profile, which upsets σ_0. ■

Denote by D^* the type profile that contains only those types, who can make credible announcements.

For the games which have more than one PBE, $M(\mathfrak{X}^*) \subseteq \mathcal{A}^*$.

Proof. Suppose that the game $\Gamma^1(\pi)$ has more than one equilibrium (if the game has one equilibrium, then the only messages in the credible message profile are the putative equilibrium messages). We have to show that if the types in $X_j \in \mathfrak{X}^*$can send messages that belong to the credible message profile, then the corresponding types $D_j \in D^*$, where D_j contains exactly those types that are in X_j, can make credible announcements.

Consider $X_j \in \mathfrak{X}^*$ and some equilibrium σ, such that precisely the types in $D_j = X_j$ would want to deviate from it. Since for all t in X_j, $A^*(X_j) = \{a^*|a \in argmax_{a \in A^*} u^s(a,t)\}$, then there must exist some 'talking strategy' δ, where $M(X_j) \in \delta(D_j)$, such that for all $t \in D_j$, there is some message $m \in M(X_j)$, such that $U^s(m,d|t) \geq U^s(\sigma|t)$ (strict for some t).

For any $m'' \in \delta(D_j)\backslash\{m\}$, such that $m'' \notin M(X_j)$, it has to be the case that there exist some t'' in D_j ,such that $U^s(m", d|t") < U^s(m, d|t")$, which implies that conditions (1') and (3') of the definition of credible announcement hold. Since none of the types outside X_j could imitate the types in X_j (by definition of the credible message profile) it has to be the case that condition (2')$U^s(m,d|t) \leq U^s(\sigma|t),, \forall t \in T\backslash D_j$ holds as well ($\iff Y^*(X_j,T) = \emptyset$) Consider some $X_k \neq X_j$, $X_k \in \mathfrak{X}^*$, such that for $D_k = X_k$, $\exists\delta' : D_k \longrightarrow \triangle M$ and conditions (1'),(2'),(3') of the definition of credible announcement hold. Suppose that $D_k \cap D_j \neq \emptyset$, but then it has to be the case that $A^*(X_k) = A^*(X_j) = A^*(X_j \cup X_k)$, since otherwise either $(X_k - X_j)$ or $(X_j - X_k)$ should not be in $\mathfrak{X}^*$ which would contradict our assumption. So, condition (4') trivially holds, since for any $X_k, X_j \in \mathfrak{X}^*$, such that $A^*(X_k) \neq A^*(X_j)$, $X_k, X_j$and therefore D_k, D_j have to be disjoint. ■

BIBLIOGRAPHY

Aumann R., Correlated Equilibrium as an Expression of Bayesian Rationality, *Econometrica 55* (1987), 1-19.

Aumann R. and A Brandenburger, Epistemic Conditions for Nash Equilibrium, *Working Paper, 91-042*, 1991.

Bernheim B.D, Rationalizable Strategic Behavior, *Econometrica 52* (1984), 1007-1028.

Bicchieri C., "Rationality and Coordination", 1993.

Brandenburger A. and Dekel E., Rationalizability and Correlated Equilibria, *Econometrica 55* (1987), 1391-1402.

Brandenburger A. and Dekel E., Hierarchies of Beliefs and Common Knowledge, *Journal of Economic Theory, 59* (1993),189-198.

Chatterjee, K., and Samuelson, W., Bargaining under incomplete information, *Oper. Res.*, 31 (1983): 835-851.

Dummett M., "The Logical Basis of Metaphysics", Harvard University Press, 1993.

Farrell J., Credible Neologisms in Games of Communication, *MIT Working Paper N.386,* 1985.

Fudenberg D. and J. Tirole, "Game Theory", The MIT Press, Cambridge, 1991.

Gresik T., Ex Ante Efficient, *ex post* Individually Rational Trade, *Journal of Economic Theory* 53 (1991), 131-145.

Gul F., Rationality and Coherent Theories of Strategic Behavior, *Research Paper #1990,* 1991.

Hinticca J., "Knowledge and Belief", Cornell University Press, 1962.

Koons R., "Paradoxes of Belief and Strategic Rationality", 1992.

Leninger W., P.B.Linhart and R.Radner, Equilibria of the Sealed-Bid Mechanism for Bargaining with Incomplete Information, *Journal of Economic Theory 48* (1989), 63-106.

Linhart P.B. and R.Radner, Minimax-Regret Strategies for Bargaining over Several Variables, *Journal of Economic Theory 48* (1989), 152-178.

Matsui A., Cheap-Talk and Cooperation in a Society, *Journal of Economic Theory, 54* (1991), 245-258.

Matsuo T., On Incentive Compatible, Individually Rational, and *ex post* Efficient Mechanisms for Bilateral Trading, *Journal of Economic Theory 49* (1989), 189-194.

Matthews S., Okuno-Fujiwara M., Postlewaite A.., Refining Cheap-Talk Equiliria, *Journal of Economic Theory, 55* (1991), 247-273.

McAfee R.P. and McMillan J., Multidimensional Incentive Compatibility and Mechanism Design, *Journal of Economic Theory 46* (1988), 335-354.

McAfee R.P., Efficient Allocation with Continuous Quantities, *Journal of Economic Theory 53* (1991), 51-74.

Myerson, R., and Satterthwaite, M., Efficient Mechanisms for Bilateral Trading, *Journal of Economic Theory* 29 (1983), 265-281.

Nöldeke G., Samuelson L., The Evolutionary Foundations of Backward and Forward Induction , *Working Paper*, 1992.

Pearce D. G., Rationalizable Strategic Bahavior and the Problem of Perfection, *Econometrica 52* (1984), 1029-1050.

Rabin M., Communication Between Rational Agents, *Journal of Economic Theory, 51* (1990), 144-170.

Rabin M., Corrigendum, *Journal of Economic Theory, 58*, (1992),110-111.

Reny P., Common Belief and the Theory of Games with Perfect Information, *Journal of Economic Theory, 59* (1993), 257-274.

Rochet, J., and Choné, P., Ironing, Sweeping and Multidimensional Screening, *Econometrica,* 1998.

Satterthwaite M. and S.Williams, Bilateral Trade with the Sealed Bid k-Double Auction: Existence and Efficiency, *Journal of Economic Theory 48* (1989), 107-133.

Sobel J., Stole L., Zapater I., Fixed-Equilibrium Rationalizability in Signaling Games, *Journal of Economic Theory 52* (1990), 304-331

Spulber D., Bargaining and Regulation with Asymmetric Information about Supply and Demand, *Journal of Economic Theory 44* (1988), 251-268

Warneryd K., Cheap-Talk, Coordination and Evolutionary Stability, *Games and Economic Bahavior 5* (1993), 532-546.

Williams S., Efficient Performance in Two Agent Bargaining, *Journal of Economic Theory 41* (1987), 154-172

Zapater I., Generalized Communication Between Rational Agents, *Working Paper N. 91-10*, Department of Economics, Brown University, March 1991.

INDEX