Efficiency of Continuous Double Auctions under Individual Evolutionary Learning with Full or Limited Information *

Mikhail Anufriev [a] Jasmina Arifovic [b] Valentyn Panchenko [c]

[a] *CeNDEF, Department of Economics, University of Amsterdam, Roetersstraat 11, NL-1018 WB Amsterdam, Netherlands*

[b] *Department of Economics, Simon Fraser University, Burnaby, B.C., Canada V5A 1S6*

[c] *School of Economics, University of New South Wales, Sydney, NSW 2052, Australia*

First Version: July 2009
This Version: February 2010

This is PRELIMINARY version of the paper.
Please do not quote it without the authors' permission.

Abstract

In this paper we explore how specific aspects of agents behavior affect the efficiency of the market outcome. In particular, we are interested whether learning behavior with and without information about actions of other participants improves market efficiency. We consider a simple market of a homogeneous good populated by buyers and sellers. The valuations of the buyers and the costs of the sellers are given exogenously. Agents are involved in the consequent trading sessions, which are organized as a continuous double auction with electronic book. Using Individual Evolutionary Learning mechanism agents submit price bids and offers, trying to learn the most profitable strategy by looking at their realized and counterfactual or "foregone" payoffs. We compare the outcomes of the continuous double auction under the full and limited information treatments, focusing both on the informational and allocative efficiency of markets and also on the individual

*We thank the participants of the SCE-2009 conference in Sydney for useful comments on earlier draft of this paper. MA acknowledges the financial support by the EU 7^{th} framework collaborative project "*Monetary, Fiscal and Structural Policies with Heterogeneous Agents (POLHIA)*", grant no.225408. Usual caveats apply.

learning outcome. We also compare theses outcome with the expected efficiency under zero-intelligence agents.

Keywords: Allocative Efficiency, Continuous Double Auction, Individual Evolutionary Learning

1 Introduction

A question of "What makes markets allocatively efficient?" has attracted a lot of attention in recent years. Methodology focusing on Zero Intelligent (ZI) agents initiated in Gode and Sunder (1993) has led to the conclusion that the rules of the market and not individual rationality are responsible for market's allocative efficiency.[1] ZI traders do not have memory and do not behave strategically, submitting random orders subject to budget constraints. Thus any effect on efficiency is attributed solely to the change of market rules. Gode and Sunder (1993) find that market organized as a continuous double auction (CDA) is highly efficient and in some cases allows ZI traders to extract around 99% of possible surplus. This result obtained some critical attention. Gode and Sunder (1997) have found that a number of specific rules of the CDA is required to guarantee this efficiency. LiCalzi and Pellizzari (2008) have shown that the allocative efficiency of the CDA would drop substantially if every transaction did not force agents submit new orders. In their words, the high efficiency results in Gode and Sunder (1993) are driven by order book "resampling".

The results of experiments starting with Smith (1962) show quick convergence towards competitive equilibrium, also resulting in high allocative efficiency of the CDA. A natural question arises about significance of individual rationality for this outcome. Gjerstad and Shachat (2007) note that budget constraints of the ZI agent, i.e., constraints on submitting orders clearly resulting in losses, is a part of agents' individual rationality. The role of rationality is even more important in the markets where the ZI agents do not extract a maximum possible surplus. A standard economic approach suggests solving for a rational equilibrium under specified market rules and the assumption of forward-looking, strategical, optimizing agents. Examples of such approach include Easley and Ledyard (1993), Friedman (1984) and Gjerstad and Dickhaut (1998). In our opinion, the fully rational approach is not completely satisfactory because of two reasons. First, given complexity of the CDA market and the high dimension of strategy space results are obtained only under auxiliary assumptions which limit either information or strategies available to participants or both. Full solution is, perhaps, not feasible anyway. Second, and more importantly, behavioral and experimental literature shows that people fail to optimize and behave strategically even in more simple situations and that models with simple learning behavior fit observed outcomes better (Erev and Roth, 1998).

In this paper we follow an intermediate approach between ZI and full rationality. More precisely, we analyze allocative efficiency in the market with boundedly rational participants learning in a fixed environment. While the demand and supply schedules are not changing from one trading session to another, the agents' bidding behavior does. We use the Individual Evolutionary Learning (IEL) algorithm, introduced in Arifovic and Ledyard (2003). According to the algorithm agents select their strategies (limit order prices) on the basis of their not only actual, but also counterfactual performance.

[1]See Duffy (2006) and Ladley (2010) for recent reviews of the ZI-literature.

We distinguish between learning based on the information available in the open order book and learning based only on aggregate market information, when the order book is closed. Similar questions were recently analyzed in Arifovic and Ledyard (2007) for *call auction* market, while we address them here for the CDA market. Openness of the order book is related to the questions of the market design which recently draw attention in the literature.[2] For example, in January 2002 the NYSE introduced OpenBook system which effectively opened the content of the limit order book to public. Boehmer, Saar, and Yu (2005) find that this increasing transparency affected investors' trading strategies and resulted in decreased price volatility and increased liquidity.

We analyze whether and how learning affects the allocative efficiency and study what kind of observable agents' behavior emerges as an outcome of the learning process. We find that learning may result in sizeable increase in efficiency. We also find that market transparency influences trading strategies and results in different market outcomes. This is consistent with recent works of Bottazzi, Dosi, and Rebesco (2005) and Anufriev and Panchenko (2009) in the field of Heterogeneous Agent Models in finance suggest that the level of allocative efficiency is a joint outcome of market rules and individual rationality.

The rest of the paper is organized as follows. The market environment is explained in Section 2, where we also recall the definition of allocative efficiency and derive a benchmark for ZI traders. The model for learning behavior of agents is introduced in Section 3 and its effects on market outcomes are discussed in Section 4. In Section 5 we report various robustness tests which have been performed. Section 6 concludes.

2 Model

We start with describing environment and defining competitive equilibrium as a benchmark against which the outcomes under different learning rules will be compared. We then proceed by explaining the continuous double auction mechanism. Finally, we study the allocative efficiency under ZI trading.

2.1 Environment

Suppose we have a fixed number $B + S$ market participants, B buyers and S sellers. At the beginning of trading session $t \in \{1, \dots, T\}$, each seller is endowed with one unit of commodity and each buyer wishes to consume one unit of commodity. The same agents transact during T trading sessions. Throughout the paper index $b \in \{1, \dots, B\}$ denotes the buyer and index $s \in \{1, \dots, S\}$ denotes the seller.

We consider a situation in which a valuation of every buyer and cost of every seller are fixed over time.[3] Buyers' valuations of a good are given by V_b, which are received

[2] Alternative interpretation of the closed order book is that agents choose not to use full information from the order book.

[3] Such *fixed environment* setup is common to the theoretical, simulation and experimental literature, for corresponding examples see Satterthwaite and Williams (2002), Arifovic and Ledyard (2007) and Gode and

when a unit is bought. Seller's costs are given by C_s, which are paid when a unit is sold. It is assumed that each trader knows his valuation/cost, but the traders do not know the valuations and costs of others. Traders care about utility defined as their surplus obtained from trade, i.e.

$$\begin{aligned} U_b(p) &= \begin{cases} V_b - p & \text{if buyer } b \text{ traded at price } p \\ 0 & \text{if buyer } b \text{ did not trade}, \end{cases} \\ U_s(p) &= \begin{cases} p - C_s & \text{if seller } s \text{ traded at price } p \\ 0 & \text{if seller } s \text{ did not trade}. \end{cases} \end{aligned} \tag{2.1}$$

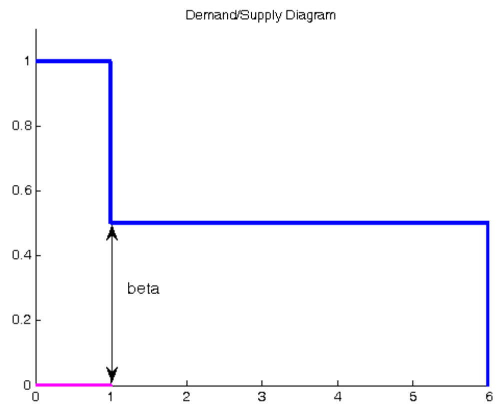
Given the set of valuations, $\{V_b\}_{b=1}^{B}$, and costs, $\{C_s\}_{s=1}^{S}$, we can build step-wise aggregate demand and supply curves, whose intersection determines the competitive equilibrium. This outcome will serve as a theoretical benchmark, as it maximizes the mutual benefits from trade. More specifically, the intersection of demand and supply determines a unique[4] equilibrium quantity $q^* \geq 0$ and, in general, an interval of the equilibrium prices $[p_L^*, p_H^*]$. This situation is illustrated in Fig. 1 for two different market environments. The units, which trade at an equilibrium price results in a nonnegative utility, are called *intramarginal* (on the figure they are to left from the the equilibrium quantity), and the agents who trade these units are called intramarginal buyers (IMBs) and intramaginal sellers (IMSs). The units, which trade at an equilibrium price would result in a negative utility, are called *extramarginal* (on the figure they are to right from the the equilibrium quantity), and the agents correspondent to these units are EMBs and EMSs. The sum of all utilities of buyers and sellers gives the *allocative value* of a trading session. When a transaction price belongs to the competitive equilibrium interval this value is maximized and is equal to the difference between the sum of the valuations of all IBMs and the sum of the costs of all IMSs. The *allocative efficiency* of a particular trading outcome is defined relatively to this benchmark.

In this paper we consider two market environments. In Fig. 1(a) we present the market introduced in Gode and Sunder (1997). There is one seller, $S = 1$, offering a unit which costs $C_1 = 0$, and $B = 1 + n$ buyers who wish to consume one unit, one of which has valuation $V_1 = 1$ and others have the same valuations equal to $\beta < 1$. The equilibrium price range is between β and 1. The n buyers with valuation β are (EMBs) and when the IMS transacts with one of them the efficiency is $\beta < 1$. A transaction between the IMS and the IMB results in a competitive outcome with efficiency equal to 1. This "GS-environment" may seem too stylized, but it is analytically tractable and provides good intuition. Moreover, varying β, we can demonstrate that the allocative efficiency of the CDA depends on the environment. The general patterns predicted by

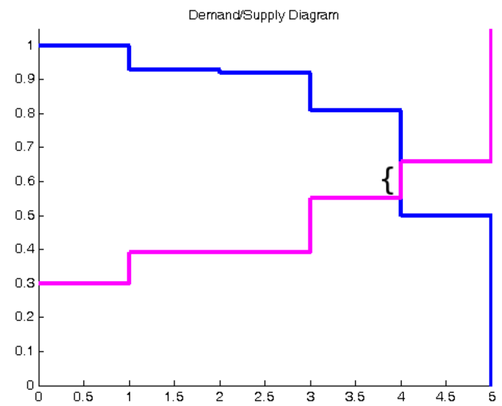
Sunder (1993).

[4]This is guaranteed by assuming that in a special case when there exists a buyer whose reservation value coincides with the cost of a seller, these sellers and buyers trade maximum possible quantity.

this simple model are confirmed by simulations with other environments.



(a) GS-market, from Gode and Sunder (1997).



(b) AL-market, from Arifovic and Ledyard (2007).

Figure 1: Competitive outcome for two market configurations considered in this paper.

The second environment we consider is depicted in Fig. 1(b). It is one of the configurations for which Arifovic and Ledyard (2007) study efficiency of the call auction. There are 5 buyers and 5 sellers in this market, with 4 IMBs and 4 IMSs.[5] The interval of equilibrium prices $[p_L^*, p_H^*]$ is shown by the figure bracket.

2.2 Continuous Double Auction

Continuous double auction (CDA) is a popular mechanism used in the electronic trade common to the stock exchanges nowadays. The market clearing is not synchronized and accommodated by an order book which stores all unsatisfied orders. If a newly submitted order finds a "matching order," it is satisfied at the price of this matching order. A matching order is defined as an order stored in the opposite side of the book at whose price the transaction with a newly arrived order is possible. If there are many orders which match the incoming order, the matching order with which the trade occurs is selected according to the price-time priority. If the submitted order does not find a matching order, it is stored in the book. All agents submit their orders (bid or ask depending on the agent's type) during a trading session.[6] There are multiple trading sessions. At the end of each trading session the order book is cleared by removing all the unsatisfied orders, so that the next session starts with an empty book.

The sequence of traders arrival to the market is randomly permuted for every trading session. We assume that during a trading session every trader can submit only one order, which, if not cleared during a transaction, will stay in the book until the end of the trading session.[7]

[5] Note that two sellers have the same costs.
[6] Each trading session can be thought as a trading "day"
[7] This assumption implies that *multiple rounds* of bidding are excluded from the analysis of this paper.

For a given set of agents' orders and their arrival sequence, the CDA mechanism described above generates a (possible empty) sequence of transactions. The prices at which buyer b and seller s traded during trading session t are denoted by $p_{b,t}$ and $p_{s,t}$, while their orders are given by $b_{(b,t)}$ and $a_{(s,t)}$, respectively. In case b traded with s, price $p_{b,t}$=$p_{s,t}$ is the price of this transaction. It is equal to $b_{(b,t)}$ if b arrived before s and is equal to $a_{(s,t)}$, otherwise. According to (2.1), buyer b who traded at price $p_{b,t}$ extracts utility $V_b - p_{b,t}$, while the buyer who did not trade over the session gets 0. Similarly, seller s who succeeded in selling the unit at price $p_{s,t}$ receives utility $p_{s,t} - C_s$, while the seller who did not trade gets 0. Note that in the CDA market the utility of the traders depend not only on their submitted orders but also on the sequence of their trades.

2.3 Market Efficiency with ZI-traders

A useful benchmark for efficiency of a market mechanism is given by its performance when the traders are Zero Intelligent (ZI). Every trading period ZI traders submit random orders, drawing them independently from a uniform distribution. Gode and Sunder (1993) distinguish between constrained and unconstrained ZI traders. Unconstrained ZI traders can draw orders from a whole interval $[0,1]$, while constrained traders are not allowed to bid higher than their valuation or ask lower than their cost. Gjerstad and Shachat (2007) attribute this restriction to the individual rationality (IR) in the order submission, rather than a market rule. We follow their terminology and distinguish between agents "with IR" and "without IR".

2.3.1 GS-environment

We derive an analytic expression for the allocative efficiency of the CDA with ZI traders for the GS-environment depicted in Fig. 1(a), when the number of extramarginal buyers $n \to \infty$. Note that in our setup a trading session may result in no transaction, whereas Gode and Sunder (1997) guarantee transaction by introducing the unlimited number of trading rounds.

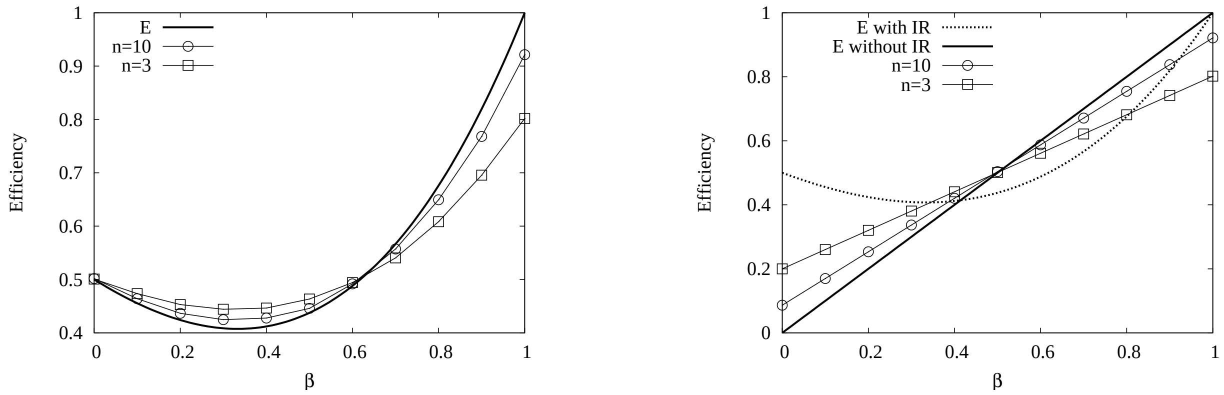
Proposition 2.1. *Consider the CDA in the GS-environment when $n \to \infty$. The expected allocative efficiency under ZI agents with IR is given by*

$$E = 0.5\big(1 + \beta^3 + \beta^2 - \beta\big), \tag{2.2}$$

the expected allocative efficiency under ZI agents without IR is given by

$$E = \beta,$$

Gode and Sunder (1997) show that multiple rounds (until all possible transactions occur) result in higher efficiency due to absence of losses caused by absence of trade. We also do not clear and "resample" the book after every transaction. Resampling would increase efficiency of the market, because orders submitted far from the equilibrium range of price would have a chance to be corrected.



(a) Agents with IR. (b) Agents without IR.

Figure 2: Allocative efficiency in the GS-environment with ZI agents. Theoretical expected efficiency E is compared with average efficiency for finite number of traders. Average is taken over 100 trading periods and 100 random seeds.

Proof. See Appendix A. □

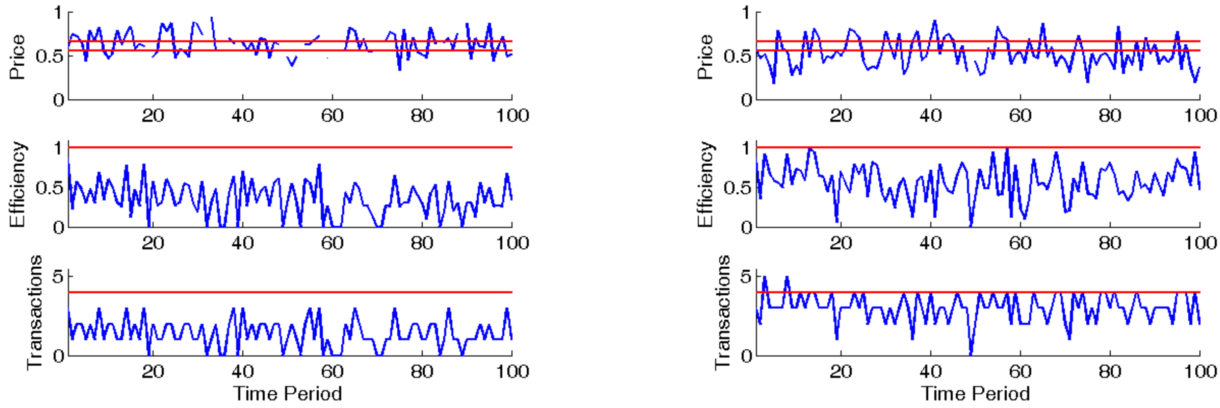
Consider first the case of ZI with IR. The solid line in Fig. 2(a) shows the theoretical efficiency (2.2) as a function of β. Its parabolic shape reflects a trade-off between the probability of inefficient transaction and the size of inefficiency. The probability of a transaction with an EMB increases in β, while the losses of allocative efficiency due to this transaction decrease in β. The probability of no trade decreases with β. Comparing (2.2) with Eq. (6) from Gode and Sunder (1997) we observe that efficiency in a market with one trading round is lower than in a market with unlimited trading rounds. In our setup efficiency can be lower than 1 not only due to a transaction with an extramarginal trader but also due to absence of trade.

Fig. 2(a) also shows the average allocative efficiency in the same market environment with a finite number n of EMBs. The average is computed over $T = 100$ trading sessions and $S = 100$ random seeds. We observe that the effect of finite number of agents is not strong. As number of agents n increases the average efficiency over the simulation runs converges to the theoretical efficiency derived in Proposition 2.1.

Analogously, Fig. 2(b) shows the efficiency under ZI without IR. Since now the probability of a transaction with an EMB is high (goes to 1 when $n \to \infty$) for any β, the trade-off between the probability of an inefficient transaction and the size of the inefficiency (equal to β) disappears. It explains a linear shape of the efficiency curve. Comparison with the IR case reveals a surprising conclusion. The absence of the IR in order submission may lead to higher efficiency for markets with high β.

2.3.2 AL-environment

Next we analyze outcomes under the ZI benchmark in the environment considered in Arifovic and Ledyard (2007) as shown in Fig. 1(b). An important difference with respect



(a) Agents with IR.

(b) Agents without IR.

Figure 3: Efficiency and price in AL-environment with ZI agents. Solid lines indicate equilibrium price range, equilibrium efficiency and equilibrium number of transactions

to the previously considered GS-environment is that now more than one transaction can occur during one trading session. Therefore, we may observe several transaction prices. In this case we report an average price of all transactions during this session.

A well known result of Gode and Sunder (1993), obtained for a similar environment, is that the allocative efficiency is close to 100%. It is obtained under assumption that the multiple rounds of bidding are allowed and the book is "resampled". We want to verify this claim relaxing this assumption and allowing only one trade per an agent in any given trading session. We simulate the trading under ZI agents with and without IR for 100 trading sessions. Fig. 3 shows dynamics of the (average) price, efficiency and number of transactions. On the price panel we show the equilibrium price range with two horizontal lines. The aggregate price is volatile, and is often outside of the equilibrium range. In the case, when the IR is imposed, the sessions when no transaction occurs are more frequent than in the case without IR.

Table 1 reports the average allocative efficiency, average price, and the average number of transactions over $T = 100$ trading sessions, as well as price volatility (standard deviation) over T periods. All these statistics are also averaged over $S = 100$ random seeds. We observe that the average allocative efficiency in the AL-environment with ZI agents is far from 100%, with the lower efficiency in the case with IR relatively to the case without IR. The low efficiency is mainly explained by the low number of transactions, which is below 4 transactions expected in competitive equilibrium. In the case with IR the number of transactions is much lower than in the case without IR. Finally, notice that the IR constraints have positive impact on the equilibrium price discovery. In the case with IR the average price is closer to the equilibrium range and price volatility is lower as opposed to the case without IR.

To summarize, our simulations with ZI agents show that the efficiency in the market does depend on the market environment (rather than only on CDA rules) and is typically

	with IR	without IR
Efficiency	0.3717	0.5752
Price	0.6211	0.4989
Price Volatility	0.1226	0.1666
Number of Transactions	1.4787	3.1176

Table 1: Aggregate outcomes in the AL-environment with ZI agents.

much lower than 100%. Further, imposing IR in agents' order submissions does not necessary improve allocative efficiency.

3 Individual Evolutionary Learning

In this paper we investigate outcomes of the market under a simple evolutionary learning mechanism with full and limited information, which reinforces successful and discourage unsuccessful strategies. We compare these outcomes to the results on the efficiency of the market populated by ZI traders. An observed action of every agent during a trading round is one submitted order.

The evolution of the orders is modeled by the Individual Evolutionary Learning (IEL) algorithm which involves the following steps:

- specification of a space of strategies (or messages);
- limiting this space to a small pool of strategies individual for every trader;
- choosing one message from the pool on the basis of its performance measure;
- evolving the pool using experimentation and replication.

The IEL can be considered as a version of the genetic algorithms adopted to the behavioral economics.

Messages

We assume that a message, $\varepsilon_{b,t}(\varepsilon_{s,t})$, represents a potential bid (or ask) order price from buyer b (or seller s) at trading session t. In our base treatment we do not allow a violation of the IR constraints , that is, we require $\varepsilon_{b,t} \leq V_b$ and $\varepsilon_{s,t} \geq C_s$. Under alternative treatments without IR constraints these restrictions will not be imposed and we will let traders to evolve and learn not to submit orders which lead to individual losses. We assume that possible orders belong to the continuous interval $[0, 1]$.

Individual Pool

Even if there is a continuum of possible messages, every agent will be restricted at every time to choose between a limited amount of them. The pool of messages (bids) available for submission at time t by buyer b is denoted by $B_{b,t}$. The pool of messages (asks) available for submission at time t by seller s is denoted by $A_{s,t}$. Every period the pool of each agent is updated, but the number of messages in the pool is fixed and equals to J. In the benchmark simulations $J = 100$. Some of the messages in the pool might be identical, so that in an agent may be choosing from J or less possible alternatives. Initially, the individual pools contains J strategies drawn, independently for each agent, from the uniform distribution on the interval $[0, 1]$.

The pool used at time t is updated before the following trading session by subsequent application of two algorithms, experimentation (or mutation) and replication. During *experimentation* stage, any message from the old pool can be replaced with a small probability by some new message. In such a way for every buyer and seller the intermediate pools are formed. More specifically, each message is removed from the pool with a small probability of experimentation, ρ, or remains in the pool with probability $1 - \rho$. In case if a message is removed, it is replaced by a new message drawn from a distribution, $\mathcal{P}$. In the benchmark simulations $\rho = 0.03$ and distribution $\mathcal{P}$ is uniform on the interval $[0, 1]$.

At the *replication* stage two randomly chosen messages from the just-formed (intermediate) pool are compared one with another, and the best of them occupies a place in new pool $B_{b,t+1}$ for a buyer or $A_{s,t+1}$ for a seller. For every agent such process is independently repeated J times (with replacement), in order to fill all the places in the new pool. The comparison is made according to a performance measure which is defined below. During replication we, therefore, increase an amount of "successful" messages in the pool at the expense of less successful messages.

Calculating the Foregone Utilities

How good is a given message? To answer this questions, every agent applies some counterfactual analysis. Indeed, only the message which has actually been used last period delivers a known payoff given by (2.1). A learning agent would also like to infer a foregone payoff from alternative strategies. Notice this is a boundedly rational reasoning, since our agent ignores the analogous learning process of all the other agents.

The calculation of foregone payoff is also made according to (2.1), but the price of transaction is notional and depends on the amount of information which is available to the agent. We distinguish between two treatments which we call *open book* (OP) and *closed book* (CL) information treatment. Under the OP treatment each agent uses full information about all bids, offers and prices from the previous period. Only the identity of bidders are not known preventing a direct access to the behavioral strategies used by others. Under the CL treatment the agents are informed only about some price aggregate, say, *average* price from the previous session, P_t^{av}. Note that the availability and use of the information from the book may be attributed either to market design (openness of the market, costs of open book access) or to individual behavior (willingness to buy information, possibility to process it), or both.

Let $\mathcal{I}_t$ denote the largest possible information set after the trading session t. It includes the orders of all buyers and sellers as well as sequence in which they arrive at the market. Under the CL treatment this whole set is not known to traders: they know only their own bids and asks as well as an average price. Thus, under the CL treatment the information sets of buyers and sellers in the end of session t are given as

$$\mathcal{I}_{b,t}^{\mathrm{CL}} = \{b_{b,t}, P_t^{\mathrm{av}}\} \cup \mathcal{I}_{b,t-1}^{\mathrm{CL}}, \quad \mathcal{I}_{s,t}^{\mathrm{CL}} = \{a_{s,t}, P_t^{\mathrm{av}}\} \cup \mathcal{I}_{s,t-1}^{\mathrm{CL}}.$$

The order book of the past period cannot be reconstructed with this information. Hence, agent can use only average price of the previous session as an indication for possible realized price given alternative message submitted.[8] Under the CL treatment, agents foregone utilities are given by

$$\begin{aligned} U_{b,t}(\varepsilon_b|\mathcal{I}_{b,t}^{\mathrm{CL}}) &= \begin{cases} V_b - P_t^{\mathrm{av}} & \text{if } \varepsilon_b \geq P_t^{\mathrm{av}} \\ 0 & \text{otherwise} \end{cases} \\ U_{s,t}(\varepsilon_s|\mathcal{I}_{s,t}^{\mathrm{CL}}) &= \begin{cases} P_t^{\mathrm{av}} - C_s & \text{if } \varepsilon_s \leq P_t^{\mathrm{av}} \\ 0 & \text{otherwise} \end{cases}. \end{aligned} \tag{3.1}$$

Under the OP treatment, agent knows the state of the order book at every moment in the previous trading session. Assuming that he arrives at the same time as he did arrive, an agent can find the (notional) price of transaction for any alternative message, $p_{b,t}^*$, and find his own payoff using (2.1). Thus, the foregone utilities under OP treatment are given by

$$U_{b,t}(\varepsilon_b|\mathcal{I}_{b,t}^{\mathrm{OP}}) = \begin{cases} V_b - p_{b,t}^*(\varepsilon_b) & \text{if order } \varepsilon_b \text{ of buyer } b \text{ is satisfied} \\ 0 & \text{otherwise} \end{cases}$$

$$U_{s,t}(\varepsilon_s|\mathcal{I}_{s,t}^{\mathrm{OP}}) = \begin{cases} p_{s,t}^*(\varepsilon_s) - C_s & \text{if order } \varepsilon_s \text{ of seller } s \text{ is satisfied} \\ 0 & \text{otherwise} \end{cases},$$

with properly specified information sets $\mathcal{I}_{b,t}^{\mathrm{OP}}, \mathcal{I}_{s,t}^{\mathrm{OP}} \subset \mathcal{I}_t$.

Selection of Message from Pool

When the new pool is formed, one of the messages is drawn randomly with a certain *selection probability* and the corresponding order is submitted for trading session $t+1$. The selection probability is also based upon foregone utilities from the previous period. For example, for buyer b the selection probability of each particular message $\varepsilon_{b,t+1}$ is computed as

$$\pi_{b,t+1}(\varepsilon_{b,t+1}) = \frac{U_{b,t+1}(\varepsilon_{b,t+1}|\mathcal{I}_t)}{\sum_{\varepsilon \in B_{b,t+1}} U_{b,t+1}(\varepsilon|\mathcal{I}_t)}, \tag{3.2}$$

[8]Other plausible possibility is to consider the closed price of the day. This modification does not influence our results.

Parameter	Symbol	Value (Range)
Number of strategies in a pool	J	100
Probability of experimentation	ρ	0.03
Distribution of experimentation	$\mathcal{P}$	$U([0,1])$
Individual Rationality constraint	IR	enforced
Transitory period	$\mathcal{T}$	100
Number of trading periods	T	100
Number of random seeds	S	100

Table 2: Parameter values used in baseline simulations.

where $\mathcal{I}_t$ is an information set, which varies depending on the type of market, and $U_{b,t+1}$ is performance of the corresponding message. Under IR all messages have non-negative performances, which guarantees that (3.2) gives a number between 0 and 1.[9]

Other specifications for selection probabilities are also possible. Popular choices in the literature are discrete choice models (probit or logit type). Our simulations showed that use of an alternative specification does not effect the results. This is mostly due to the replication stage which in several rounds replaces most of the strategies in the pool with similar relatively well performing strategies.

4 Market Efficiency under IEL

In our simulations performed with learning agents we concentrate on four different aggregate variables: allocative efficiency, session-average price, its volatility and number of transactions. As before we compute the average values of these variables over $T = 100$ consecutive trading periods after $\mathcal{T} = 100$ transitory periods. To eliminate a dependence on a realization of particular random sequence we average the above numbers over $S = 100$ random seeds.

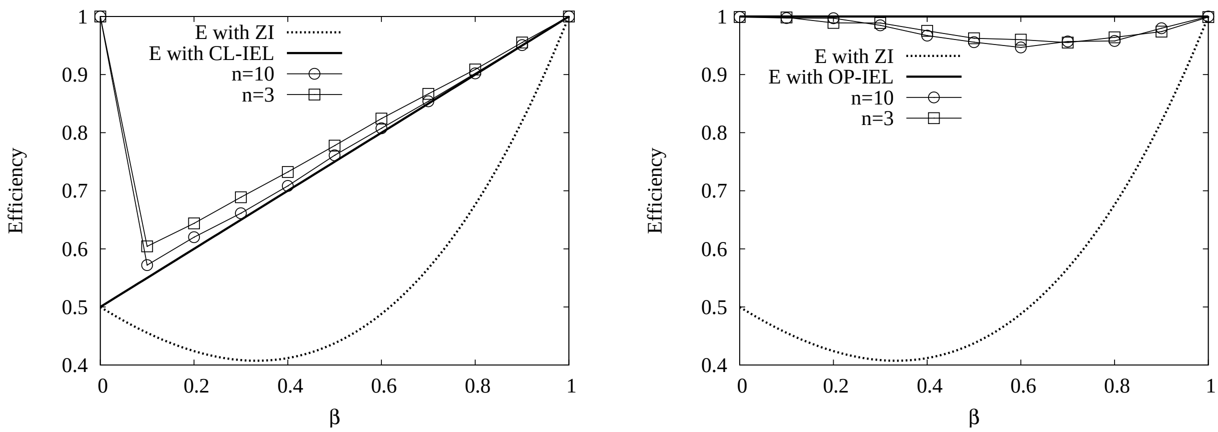
Table 2 summarizes the parameters of the IEL model which we use in the baseline simulation throughout this Section. Notice that the IR is enforced in the baseline treatment.

4.1 GS-environment

To study the effects of IEL learning and the impact of market transparency on allocative efficiency, we compare these treatments with ZI benchmark analyzed in Section 2.3.1.

Figs. 4(a) and 4(b) show an allocative efficiency under the IEL with CB and OB, respectively. For the GS-environment simulated with $n = 3$ and $n = 10$ EMBs we observe a significant increase in allocative efficiency with respect to the ZI-benchmark shown by dotted line. The solid line indicates the theoretical expected efficiency for

[9]This does not necessarily hold without IR and we add a constant equal to 1 to the performance of every message to insure that (3.2) generates a number between 0 and 1.



(a) CL treatment.

(b) OP treatment.

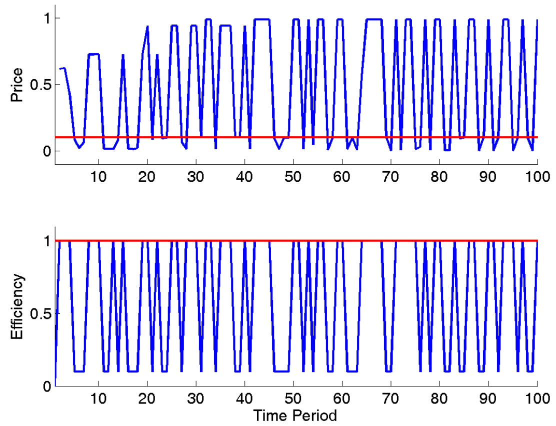
Figure 4: Efficiency in the GS-environment under IEL as compared to ZI-benchmark.

$n \to \infty$ derived below. The allocative efficiency under the IEL practically does not depend on n, the number of EMBs. Notice the difference caused by transparency of the book. The allocative efficiency is higher under the OP treatment, actually very close to 100% for any β, while under the CL treatment there is a positive linear dependence between the efficiency and β.

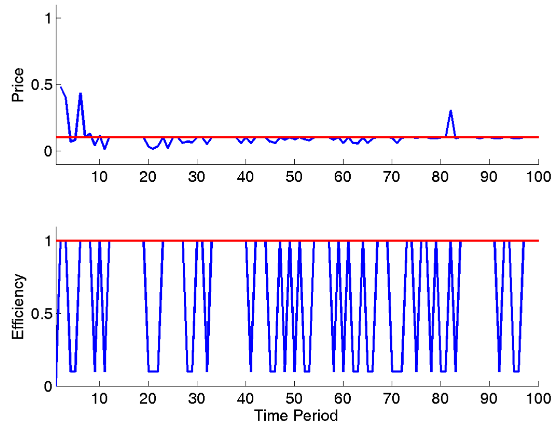
In order to explain these results for the aggregate market outcomes we look into individual strategies and their evolution. An important question is whether and where the IEL driven individual strategies converge under different treatments. Figs. 5 and 6 show the evolution of the market during first 100 trading sessions for $\beta = 0.1$ and $\beta = 0.5$, respectively. Upper panels show the evolution of market price and efficiency under CL (left panel) and OP (right panel) treatment. Horizontal lines indicate β on the price panel and 100% efficiency on the efficiency panel. We observe that the price is much more volatile under the CL and is stable and close to β under the OP. The efficiency is permanently changing between β and 1 under CB for both values of β. Such behavior characterizes also the long run as we infer from Fig. 4(a). Under OP the efficiency is also initially changing between β and 1, but then converges to 1. (In case $\beta = 0.1$ convergence occurs after 100 periods and is not shown on the plot.) An outlier in period 91 for $\beta = 0.5$ on Fig. 6(b) is the result of agents' experimentation.

On the panels (c) and (d) of Figs. 5 and 6 we show the evolution of individual bids and asks for buyers and sellers. Their valuations/costs are denoted by stars in the right part of the figure. The range of equilibrium prices is indicated by a vertical line. We observe that in the CL treatment shown in panels (c) the orders of the intramarginal traders tend to their valuations/costs and that other traders (EMBs) exhibit somewhat similar behavior. The following result shows, the strategy profile with pools consisting of such strategies is attracting in the sense that any individual deviation away from this strategy profile will not survive in the long run.

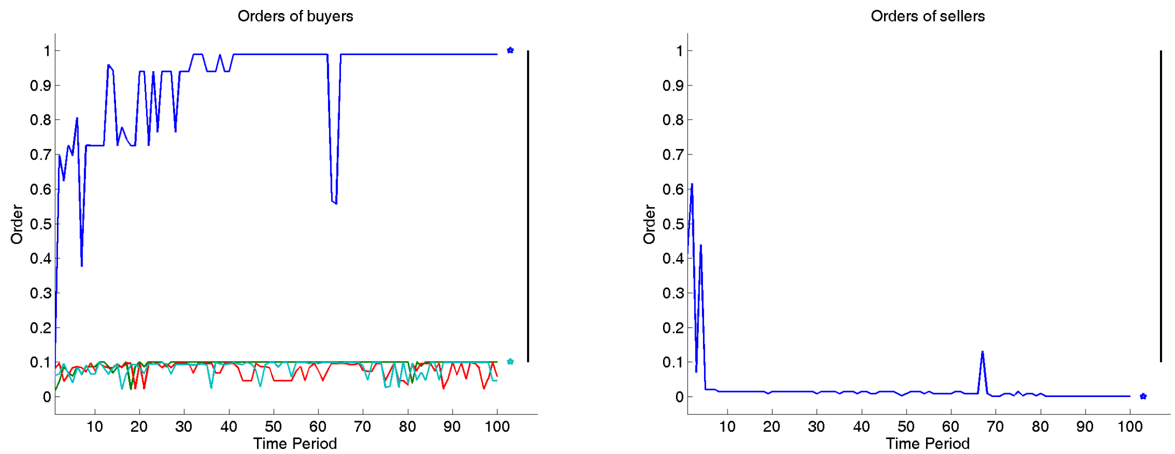
Result 1. *The strategy profile under which the pool of every trader consists of strate-*



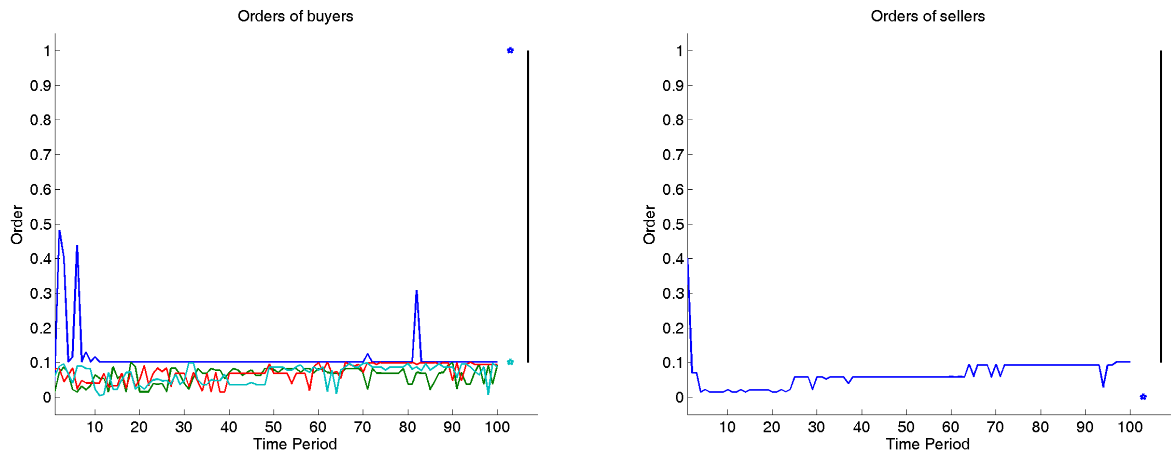
(a) Aggregate outcomes under CL.



(b) Aggregate outcomes under OP.

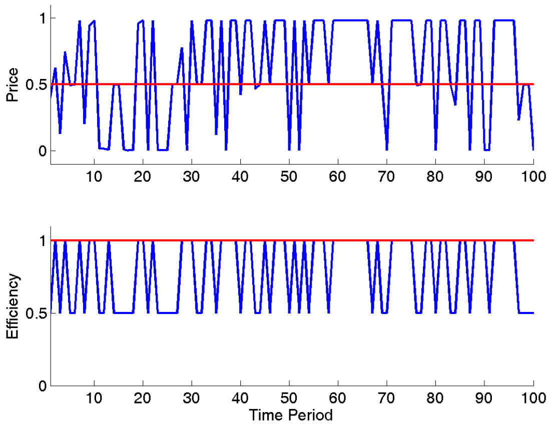


(c) Individual bids (left) and asks (right) under CL.

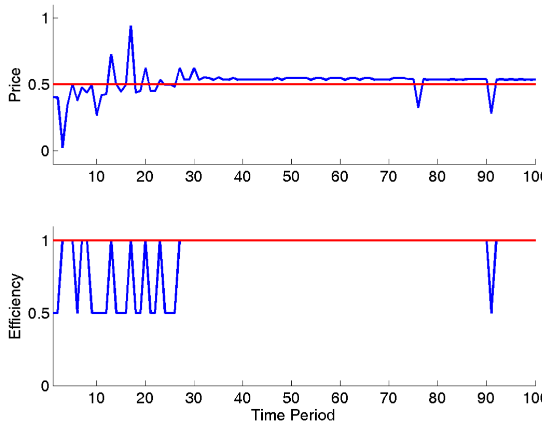


(d) Individual bids (left) and asks (right) under OP.

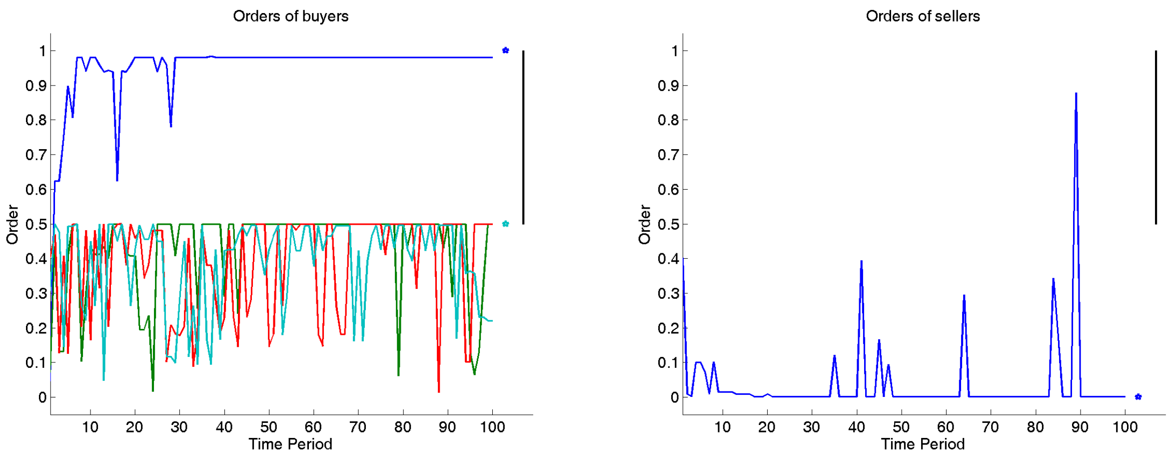
Figure 5: Dynamics in the GS-environment with 3 EMBs with $\beta = 0.1$. Horizontal lines indicate β on the panel for price and 100% efficiency on the panel for efficiency. In the right part of the panels (c) and (d) stars denote valuations/costs of agents and vertical line shows equilibrium price range.



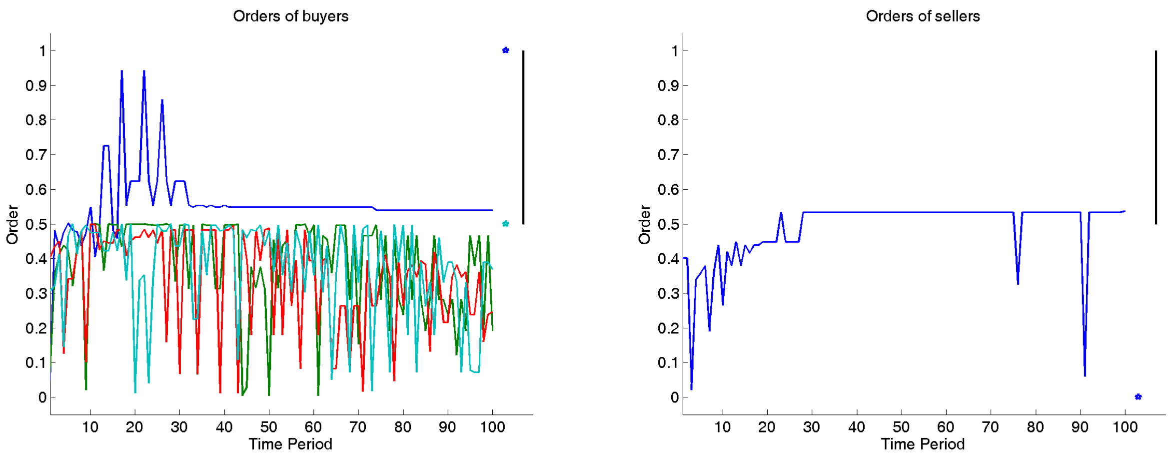
(a) Aggregate outcomes under CL.



(b) Aggregate outcomes under OP.



(c) Individual bids (left) and asks (right) under CL.



(d) Individual bids (left) and asks (right) under OP.

Figure 6: Dynamics in the GS-environment with 3 EMBs with $\beta = 0.5$. Horizontal lines indicate β on the panel for price and 100% efficiency on the panel for efficiency. In the right part of the panels (c) and (d) stars denote valuations/costs of agents and vertical line shows equilibrium price range.

gies equal to his own valuation/cost is attracting under the CL treatment in the GS-environment.

Under this equilibrium configuration the price oscillates in the range $[0, 1]$*, and the expected efficiency is given by*

$$E^{CL} = \frac{1 + \beta n}{(n+1)(n+2)} + \frac{1}{n+2}\left(1 + n\frac{\beta + 1}{2}\right). \tag{4.1}$$

Proof. See Appendix. □

When number of agents $n \to \infty$ the expression (4.1) converges to $(1 + \beta)/2$, shown by a solid line in 4(a).

In the market with the OP treatment the evolution of individual strategies is remarkably different. In the Figs. 5(d) and 6(d) we observe that intramarginal traders are able to coordinate on one price and submit the orders close to this price. We have the following result.

Result 2. *For any price p from the equilibrium price range* $[\beta, 1]$ *the strategy profile under which the pools of the IMB and the IMS consist of strategies equal to this price is stable under the OP treatment in the GS-environment.*

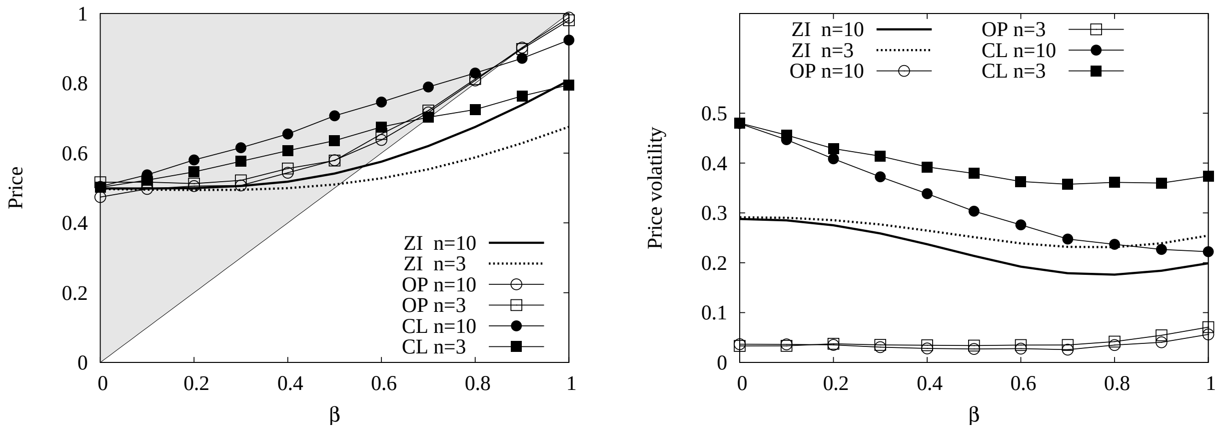
Under such configuration the price is stable and the expected efficiency is given by

$$E^{OP} = 1$$

Proof. See Appendix. □

The last result shows that in the OP treatment there are multiple equilibria with any price within the equilibrium range. For example, in Fig. 6(d) the strategies of the IMB and IMS converged to the same submitted orders approximately equal to 0.53. Of course, the EMBs never trade in such market and all their strategies in the pools have equal probabilities, see the left panel. On the other hand, in Fig. 5(d) the strategies of the IMB and IMS did converge to 0.1 only in the session $t = 95$. Before this period the EMBs had their chances to trade and could learn to submit orders close to 0.1.

Even if Result 2 implies the 100% allocative efficiency, due to experimentation the efficiency may drop in some periods. This happens around the period 91 in Fig. 6(d). After previous trading round the seller's pool was dominated by the orders equal to 0.53, which is the price at period 90. An experimentation adds a strategy 0.06 to the seller's pool, which survives replication stage. In fact, the price p_{90} was determined by the buyer's order (the seller at $t = 90$ arrived after the buyer) and so all the strategies below p_{90} have the same hypothetical payoffs. Even if the strategy 0.06 belongs to the seller's pool at time 91, a probability to use this strategy as an order is only $1/J = 1/100$. Whenever such order is submitted, the price will be lower than previously observed 0.53. In this particular case, $p_9 1 =???$ equal to the order of one of the EMB. Notice that after



(a) Price.

(b) Price volatility.

Figure 7: Average price and price volatility under IEL in the GS-environment as compared to the ZI-benchmark.

this trading round, the seller will re-evaluates his strategies, and strategy 0.53 will have higher hypothetical payoff than 0.06.

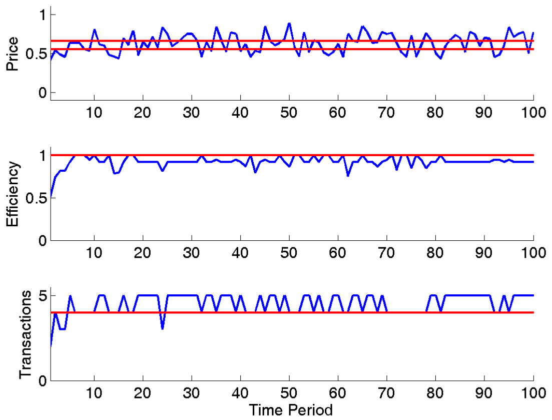
To summarize, the information used by the agents under the IEL shapes their strategy pool in the long-run. This pool affects the aggregate dynamics, which feeds back by providing a ground for selection of active strategies within the pool. When the book is closed (CL treatment), agents react on commonly available signal (price of the transaction) and learn to submit their own valuation. This leads to higher opportunity of trade, but also to larger price volatility, as we observed in Figs. 5(a) and 6(a). Instead, when the book is open (OP treatment), agents can adapt to the stable strategies. Such individual behavior results in a stable price behavior at the aggregate level.

Fig. 7 shows the average price and price volatility and confirms the above finding. In particular, we observe that under OP treatment the price is within the competitive price range denoted by shaded area for any β. This is not the case for the CL and ZI. The price volatility is at the lowest level for OP and the largest for CL.

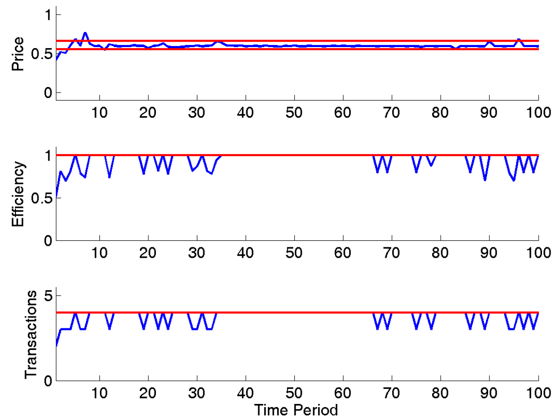
4.2 AL-environment

Does the results about aggregate dynamics and individual behavior observed in the stylized GS-environment also hold under alternative environments? Fig. 8 shows market aggregates (price, efficiency and number of transactions), as well as evolution of the individual trading orders (bids/asks) over time for the AL-environment. Two horizontal lines on the panel for price indicate equilibrium price range, while the line on the panel for transactions shows the equilibrium number of transactions equal to 4.

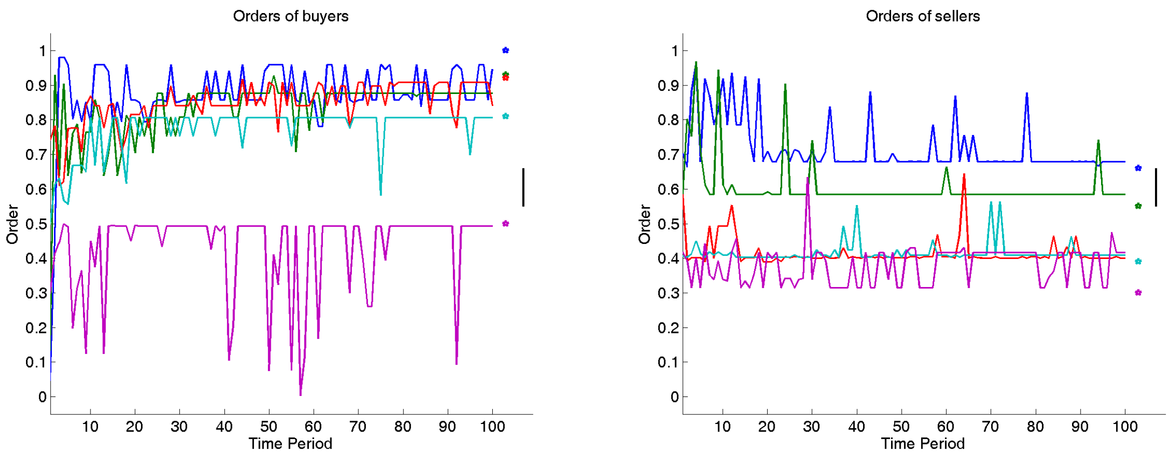
As before, the price is less volatile under the OP treatment and lie within the equilibrium range, while in case of the CL treatment the price is often outside the equilibrium range. The efficiency under OP and CL treatments is comparable. Interestingly, a loss



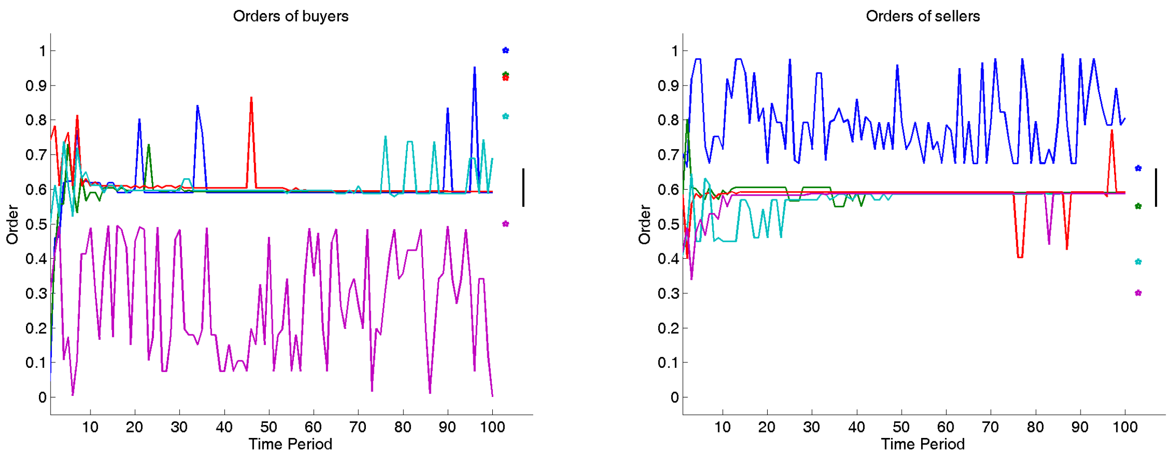
(a) Aggregate outcomes under CL.



(b) Aggregate outcomes under OP.

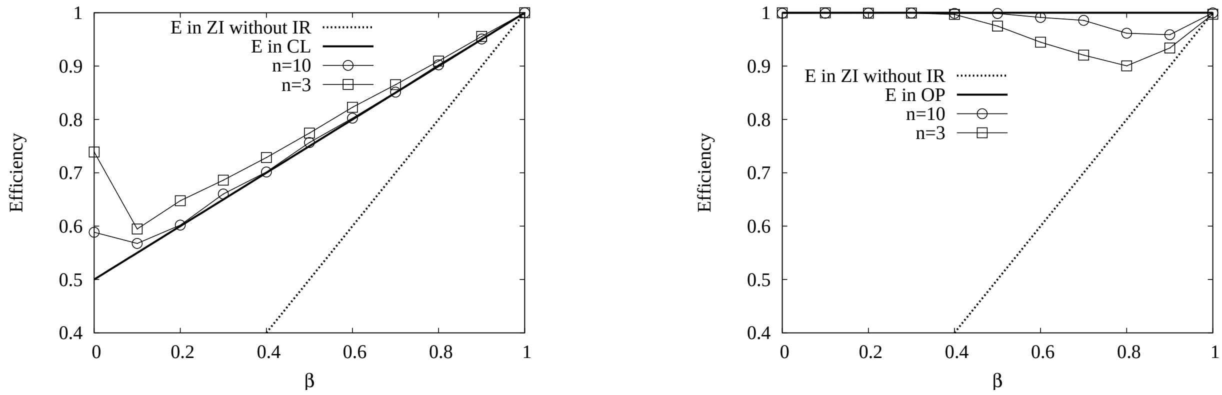


(c) Individual bids (left) and asks (right) under CL.



(d) Individual bids (left) and asks (right) under OP.

Figure 8: Dynamics in the AL-environment. Horizontal lines indicate equilibrium price range on the panel for price, equilibrium efficiency on the panel for efficiency and equilibrium number of transactions on the panel for transactions. In the right part of the plots for individual strategies stars denote valuations/costs of agents and vertical line shows equilibrium price range.



(a) CL treatment. (b) OP treatment.

Figure 9: Efficiency in the GS-environment populated by agents without IR.

of efficiency under the CL is attributed to overtrading. This is a consequence of larger than equilibrium range of price fluctuations which also contain the valuations/costs of two extramarginal traders. Under the OP treatment the loss of efficiency occurs due to smaller than equilibrium number of transactions.

As for the individual strategies, under the OP (Fig. 8(d)) the intramarginal traders coordinate on one price as we have seen in the GS-environment. The Result 2 still holds. However, under the CL (Fig. 8(c)) traders' orders converge to their valuations/costs only if the latter fall within the range of price fluctuations. It follows from (3.1) that the IEL process creates an upward pressure only on those buyers' orders which lie below average price of the last trading session, P_t^{av} (and downward pressure only on those sellers' orders which lie above P_t^{av}). Whereas in the GS-environment every order could become a transacted and, hence, average price, so that ultimately the price fluctuated within a whole range $[0, 1]$, in the AL-environment the price P_t^{av} average out the individual orders. It leads to smaller range of fluctuation and does not allow traders with relatively extreme valuations/costs learn.

5 Robustness

5.1 Role of Individual rationality

We find that generally IEL learning is robust towards a violation of the IR constraints. In the GS-environment the efficiency in the case under the IR, Fig. 4, is close to the efficiency obtained without the IR, see Fig. 9.

In the AL-environment the removal of individual rationality constraints slightly promotes learning of agents' valuations/costs under the CL. This is a natural consequence of higher volatility in submitted orders and, therefore, in price as observed in Fig. 10. At the same time, absence of IR impairs the coordination to one price and prolongs the time necessary for convergence under the OP. Fig. 10(b) shows that this results in some periods of relatively low efficiency and low number of transactions.

5.2 Role of IEL parameters

Table 3 summarizes the efficiency, price, price volatility and number of transactions averaged for different combinations of the probability of experimentation, ρ, and the size of the strategy pool, J.

Under the CL treatment an increase in the probability of experimentation results in minor decrease of efficiency. Note that this is accompanied by a considerable decrease in the number of transactions. Due to distortions introduced by the experimentation, for higher ρ individual orders lie further from the traders' valuations/costs than for smaller ρ. This reduces a probability of transaction for any trade and leads to a lower range for price and, hence, lower price volatility. Nevertheless, the extramarginal traders may still trade substituting intramarginal traders leading to some efficiency loss. Larger size of the pool J increases the frequency of experimentation and has a similar effect as ρ.

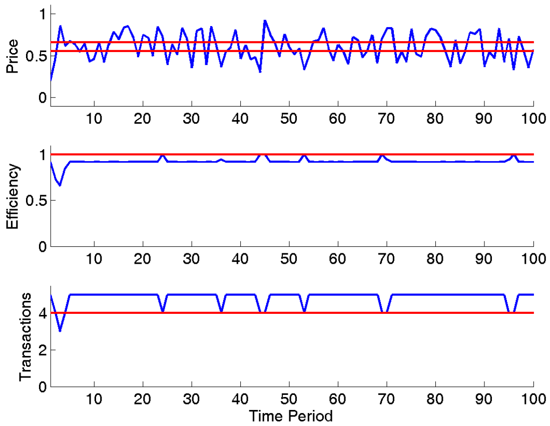
Under the OP treatment, the effect of ρ is non-monotone. For $\rho \leq 0.10$ the increase in ρ corresponds to the increase in efficiency, number of transaction and decrease in price volatility, while for $\rho = 0.30$ all these measures reverse their direction. A trade-off between the speed of convergence to the equilibrium configuration and the frequency of deviation from this equilibrium may explain this non-monotonicity. In turn, the effect of pool size seems to be monotone, i.e., larger J improves allocative efficiency, increases number of transaction and reduces price volatility.

Higher level of experimentation leads to faster convergence to one price while higher standard deviation introduces more price distortions. The observations under the uniform noise are fairly similar.

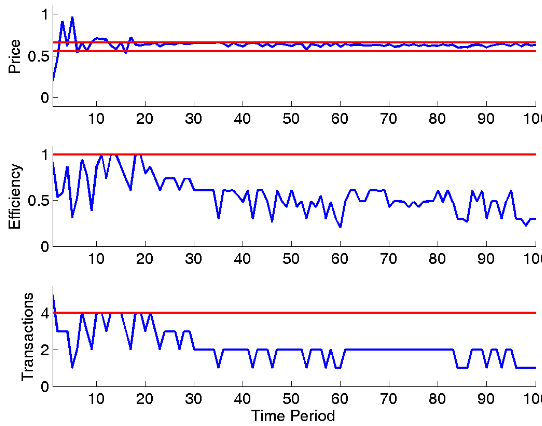
6 Conclusion

In this paper we analyzed the effects of learning with full and limited information and compared them with ZI benchmark. Participants are modeled as neither entirely rational nor fully irrational. Instead, they use the IEL mechanism to learn which orders to trade. Two main Darwinian ideas are inherent to this mechanism. First is *experimentation*, which means that agents are allowed to use, in principle, any strategy at some period of time. Second is *selection with reinforcement*, so that strategies with higher past payoffs have higher representation in the strategy pool and higher probability to be used. Furthermore, agents are comparing strategies not only on the basis of actual, but also foregone payoffs.

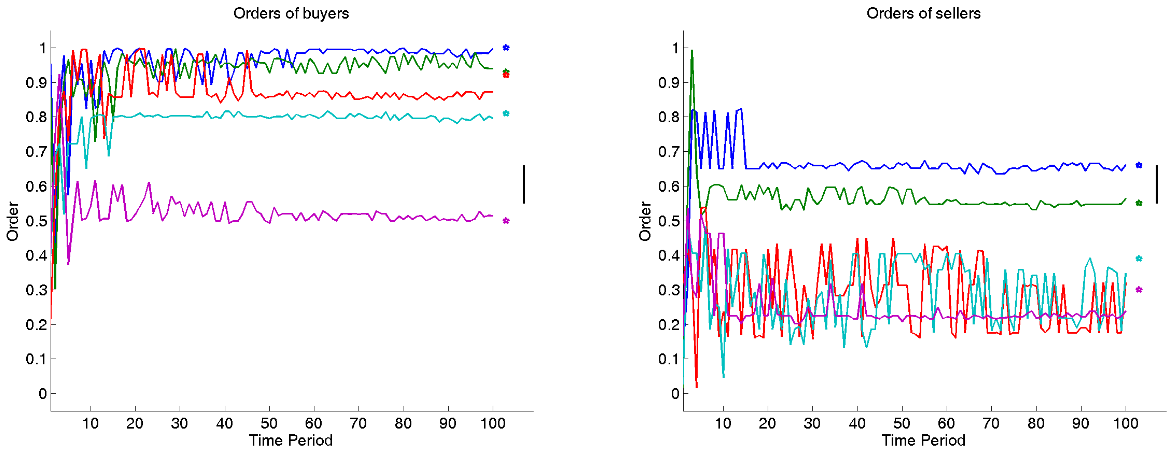
We derive allocative efficiency for ZI traders and show *via* simulations that learning leads to a higher efficiency. We also investigate the role of trade transparency and show that strategies used by trades are strikingly different in the CL and OP treatments.



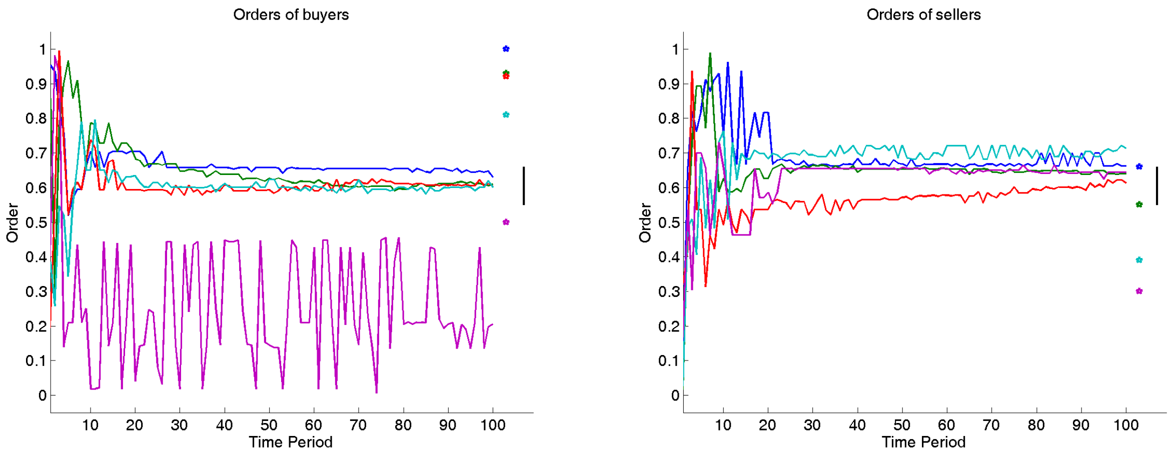
(a) Aggregate outcomes under CL.



(b) Aggregate outcomes under OP.



(c) Individual bids (left) and asks (right) under CL.



(d) Individual bids (left) and asks (right) under OP.

Figure 10: Dynamics in the AL-environment without IR. Horizontal lines indicate equilibrium price range on the panel for price, equilibrium efficiency on the panel for efficiency and equilibrium number of transactions on the panel for transactions. In the right part of the plots for individual strategies stars denote valuations/costs of agents and vertical line shows equilibrium price range.

	CL: closed book				OP: open book			
	$J = 10$	$J = 50$	$J = 100$	$J = 200$	$J = 10$	$J = 50$	$J = 100$	$J = 200$
$\rho = 0.01$								
Efficiency	0.935	0.931	0.931	0.931	0.867	0.912	0.919	0.925
Price	0.627	0.634	0.641	0.642	0.648	0.633	0.641	0.640
Price Volat	0.111	0.123	0.125	0.124	0.040	0.021	0.018	0.018
Num Transact	4.180	4.601	4.725	4.759	3.571	3.736	3.780	3.784
$\rho = 0.03$								
Efficiency	0.932	0.931	0.930	0.930	0.887	0.915	0.925	0.929
Price	0.634	0.638	0.640	0.642	0.643	0.640	0.636	0.633
Price Volat	0.120	0.124	0.126	0.127	0.035	0.023	0.022	0.021
Num Transact	4.226	4.579	4.643	4.675	3.652	3.774	3.801	3.810
$\rho = 0.10$								
Efficiency	0.932	0.930	0.929	0.929	0.896	0.926	0.935	0.936
Price	0.636	0.638	0.637	0.638	0.642	0.642	0.636	0.638
Price Volat	0.121	0.130	0.129	0.131	0.042	0.028	0.027	0.025
Num Transact	4.135	4.290	4.325	4.355	3.686	3.802	3.788	3.811
$\rho = 0.30$								
Efficiency	0.922	0.926	0.927	0.927	0.845	0.850	0.851	0.852
Price	0.643	0.642	0.640	0.641	0.648	0.645	0.642	0.642
Price Volat	0.110	0.112	0.112	0.112	0.070	0.064	0.062	0.062
Num Transact	3.963	4.048	4.058	4.074	3.498	3.520	3.514	3.532

Table 3: Aggregate outcomes of the open and close book CDA under varying ρ and J averaged over 100 random seeds, 100 trading sessions after 100 transient trading sessions.

Under the CL traders tend to learn their valuations/costs, while under the OP intra-marginal traders coordinate on one price with the efficient price range. This results in differences on the aggregate level, i.e. higher price volatility and overtrading under the CL relatively to the OP treatment. The allocative efficiency is comparable in both cases, however the sources of the inefficiencies are different.

Some of the assumptions used in the paper may be not fully realistic. Hence it would be interesting and important to relax them. First, canceling of the orders can be introduced to model intermediate situations between two extremes: no-resampling as here and resampling as in Gode and Sunder (1993). Second, submission of multiple orders can be introduced to model intermediate situations between two extremes: one-round as here and unbounded amount of multiple rounds as in Gode and Sunder (1997). Third, learning on the different scale than trading can be also considered analogous to what is done in Fano, LiCalzi, and Pellizzari (2009).

Appendix

Proof of Proposition 2.1. First, let us consider ZI agents with IR. We consider in turn different situations of the outcome of trading session, evaluating probability of each situation and its allocative efficiency.

1. IMB arrives before IMS, probability $P = 0.5$, and
   (a) IMB bids $b < \beta$, and
      i. IMS asks $a < \beta$, $P = \beta^2$ - then EMB trades and efficiency is β, or
      ii. IMS asks $a > \beta$, $P = \beta(1-\beta)$ - no transaction occurs, efficiency is 0
   (b) IMB bids $b > \beta$, and
      i. IMB bids $b >$ IMS a, $P = \frac{(1-\beta)^2}{2} + \beta(1-\beta)$ - then IMB trades and efficiency is 1, or
      ii. IMB bids $b <$ IMS a, $P = \frac{(1-\beta)^2}{2}$ - no transaction occurs, efficiency is 0
2. IMB arrives after IMS, $P = 0.5$, and
   (a) IMS $a < \beta$, $P = \beta$ - then EMB trades and efficiency is β, or
   (b) IMS $a > \beta$, and
      i. IMB $b >$ IMS a, $P = \frac{(1-\beta)^2}{2}$ - then IMB trades and efficiency is 1, or
      ii. IMB $b <$ IMS a, $P = \frac{(1-\beta)^2}{2} + \beta(1-\beta)$ - no transaction occurs, efficiency is 0

Expected efficiency is given by the following expression

$$E = \frac{1}{2}(\beta^3 + \frac{(1-\beta)^2}{2} + \beta(1-\beta)) + \frac{1}{2}(\beta^2 + \frac{(1-\beta)^2}{2}) = \frac{1}{2}(1 + \beta^3 + \beta^2 - \beta)$$

Second, we consider the case of ZI agents without IR. In this case there is no difference in bidding behavior between IMB and EMB. Thus, when a number of the EMB traders converges to infinity with probability 1 an EMB will trade. Such trade delivers efficiency β. □

Proof of Result 1. Consider the rule for the foregone utilities (3.1), which agents use in their learning procedure. Under the GS-environment there is only one price during the trading session, $p_t = P_t^{av}$. After this price is realized the IMB (IMS) receives the same nonnegative utility for any message above (below) p_t and zero utility for all other messages. The same holds for the EMBs, apart from the fact that imposed IR does not allow them submit orders above β.

Suppose now that everybody starts with configuration described, and that one of the traders, say an EMB, has a mutant strategy $\beta' < \beta$ in his pool. Depending on the order of arrival the price during the session will be 1, β, β' or 0, all with positive probabilities. In all cases all the messages in the agent's pool will receive the same utility (zero for the first two cases and β for the last two cases) and we expect to find one mutant in the pool after replication. In the next period a new mutant β'' enters the pool. If $\beta'' > \beta'$ the new mutant dominates the old mutant in the long-run because the expected utility of the new mutant is higher than the expected utility of the old mutant (If the trading price happens to be β'', strategy β'' receives utility β while strategy β' receives zero utility. For other possible prices 1, β, β' or 0, both strategies receive the same utility). By the a similar reasoning, if $\beta'' < \beta'$ the new mutant is dominated the old mutant in the long-run. Hence only mutations towards attracting configuration of own valutions/costs will survive in the long run. The same reasoning holds for other types of traders, which proves stability.

Obviously such configuration leads to the oscillating price. To prove (4.1) we consider the following situations:

1. the IMS arrives first, probability $P = 1/(n+2)$, and
   (a) the IMB arrives next, probability $P = 1/(n+1)$, and efficiency is 1, or,
   (b) the EMB arrives next, probability $P = n/(n+1)$, and efficiency is β

   In this situation the price is 0.
2. the IMB arrives first, probability $P = 1/(n+2)$, and efficiency is 1. In this situation the price is 1.
3. an EMB arrives first, probability $P = n/(n+2)$, and
   (a) the IMS arrives before the IMB, probability $P = 1/2$, and efficiency is β, in this situation the price is β, or,
   (b) the IMS arrives after the IMB, probability $P = 1/2$, and efficiency is 1, in this situation the price is 1.

Summing it up we obtain (4.1) for the efficiency. □

Proof of Result 2. Suppose that both intramarginal traders have homogeneous pools with p_t strategies inside. Consider an arbitrary mutative strategy by the IMB. If this strategy is larger than p_t then it will be dominated by incumbent strategies in the sessions when the IMB arrives before the IMS (i.e., with probability 1/2) and will have an equal chance to survive otherwise. If this strategy is smaller than p_t then the IMB does not trade at all and the mutant is eliminated from the pool in any case. The same reasoning holds for the IMS, which proves stability.

Obviously, such strategy profile leads to the price at level p with efficiency equal to 1.

□

References

Anufriev, M., and V. Panchenko (2009): "Asset Prices, Traders' Behavior and Market Design," *Journal of Economic Dynamics & Control*, 33, 1073 – 1090.

Arifovic, J., and J. Ledyard (2003): "Computer testbeds and mechanism design: application to the class of Groves-Ledyard mechanisms for the provision of public goods," manuscript.

——— (2007): "Call market book information and efficiency," *Journal of Economic Dynamics and Control*, 31(6), 1971–2000.

Boehmer, E., G. Saar, and L. Yu (2005): "Lifting the veil: An analysis of pre-trade transparency at the NYSE," *Journal of Finance*, 60(2), 783–815.

Bottazzi, G., G. Dosi, and I. Rebesco (2005): "Institutional Architectures and Behavioral Ecologies in the Dynamics of Financial Markets: a Preliminary Investigation," *Journal of Mathematical Economics*, 41, 197–228.

Duffy, J. (2006): "Agent-Based Models and Human Subject Experiments," in *Handbook of Computational Economics Vol. 2: Agent-Based Computational Economics*, ed. by K. Judd, and L. Tesfatsion. Elsevier/North-Holland (Handbooks in Economics Series).

Easley, D., and J. Ledyard (1993): "Theories of price formation and exchange in double oral auctions," in *The double auction market: Institutions, theories, and evidence*, ed. by D. Friedman, and J. Rust, pp. 63–97. Perseus Books.

Erev, I., and A. E. Roth (1998): "Prediction how people play games: Reinforcement learning in games with unique strategy equilibrium," *American Economic Review*, 88, 848–881.

Fano, S., M. LiCalzi, and P. Pellizzari (2009): "Evolution and Convergence of Strategic Behavior in Continuous Double Auctions," mimeo.

Friedman, D. (1984): "On the efficiency of experimental double auction markets," *The American Economic Review*, 74(1), 60–72.

Gjerstad, S., and J. Dickhaut (1998): "Price formation in double auctions," *Games and Economic Behavior*, 22(1), 1–29.

Gjerstad, S., and J. Shachat (2007): "Individual rationality and market efficiency," *Purdue University Economics Working Papers*, 1204.

Gode, D., and S. Sunder (1993): "Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality," *Journal of Political Economy*, pp. 119–137.

——— (1997): "What Makes Markets Allocationally Efficient?," *Quarterly Journal of Economics*, 112(2), 603–630.

Ladley, D. (2010): "Zero Intelligence in Economics and Finance," *Knowledge Engineering Review*, forthcoming.

LiCalzi, M., and P. Pellizzari (2008): “Zero-Intelligence Trading Without Resampling,” in *Complexity and Artificial Markets*, ed. by K. Schredelseker, and F. Hauser, pp. 3–14. Springer.

Satterthwaite, M., and S. Williams (2002): “The optimality of a simple market mechanism,” *Econometrica*, pp. 1841–1863.

Smith, V. L. (1962): “An Experimental Study of Competitive Market Behavior,” *The Journal of Political Economy*, 70(2), 111–137.