Essays on markets over random networks and learning in Continuous Double Auctions

This dissertation has been written within the European Doctorate in Economics-Erasmus Mundus (EDE-EM) programme in order to obtain a joint doctorate degree at the Faculty of Economics and Business at the University of Amsterdam and the Department of Economics at Università Ca' Foscari Venezia.

Layout: Michiel Chr.W. van de Leur

Cover design: Coördesign, Leiden

PhD committee

Advisors	Prof. dr. C.H. Hommes (University of Amsterdam)
	Prof. dr. M. LiCalzi (Università Ca' Foscari Venezia)
Co-advisor	Dr. M. Anufriev (University of Technology Sydney)
Other members	Prof. dr. J. Arifovic (Simon Fraser University)
	Prof. dr. H. Dawid (Universität Bielefeld)
	Prof. dr. C.G.H. Diks (University of Amsterdam)
	Dr. P. Pellizzari (Università Ca' Foscari Venezia)
	Prof. dr. J. Tuinstra (University of Amsterdam)

Acknowledgements

This dissertation would not have been possible without the many people I have had the pleasure of meeting in the last three years. It has been very inspiring to work with so many excellent researchers in the Netherlands, Italy and Germany.

I am grateful for the support of the European Doctorate in Economics - Erasmus Mundus (EDE-EM) programme and the Erasmus Mundus Association for allowing me this opportunity.

My greatest gratitude goes to my supervisors at the different universities; Mikhail Anufriev for the advice and inspiring talks and the effort to meet whenever it was possible; Marco LiCalzi for the intense supervision and all the inspiration during my stay in Venice; Herbert Dawid for all the constructive discussions in Bielefeld; Cars Hommes for looking after the bigger picture of my research and the overall supervision.

Valentyn Panchenko and Jasmina Arifovic have been very kind to provide me with their input on my dissertation.

Furthermore I want to thank my entire PhD committee, Cars Hommes, Marco LiCalzi, Mikhail Anufriev, Jasmina Arifovic, Herbert Dawid, Cees Diks, Paolo Pellizzari and Jan Tuinstra, for their careful reading of this manuscript and providing me with feedback.

During the last three years I have met and worked with some very inspiring researchers at the universities that welcomed me: University of Amsterdam, Università Ca' Foscari Venezia and Universität Bielefeld. In these institutions I have had the pleasure meeting Daan, Tomasz, David, Juanxi, Thom, Marcin, Lorenzo, Nadia, Peter and Bertrand. Numerous colleagues provided me with ideas and remarks during seminars.

Whenever necessary my family and my friends helped me relax my mind. My friends within the national frisbee team deserve a huge thanks for the amazing contrast they provide in my life.

Finally I owe a lot to my close family and friends who have always supported me. Marjolein, Barry, Thomas, Sinead and Maarten, during the most difficult year in my life you were always there; I could not have coped without you.

Michiel van de Leur

September 2014

”Why are numbers beautiful?

It’s like asking why is Beethoven’s Ninth Symphony beautiful.

If you don’t see why, someone can’t tell you.

I *know* numbers are beautiful.

If they aren’t beautiful, nothing is.”

(Paul Erdős)

Contents

Chapter 1

Introduction and Thesis Outline

The design of a market and the information that is available before traders make their decision largely influence traders' behaviour and the efficiency of the market. For example the OpenBook system as introduced in 2002 by the New York Stock Exchange, opened the content of the limit order book to the public. This allows for a change in behaviour of traders, who can now condition their strategy on the full history of orders. We study whether a market design with more information, such as the OpenBook system, is preferable in terms of efficiency. More information benefits traders with a high market power and hurts others, but it is unclear whether the total profit in the market and thus efficiency will increase. We consider boundedly rational behaviour of traders and the resulting efficiency depending on the available information in the market design, to study what information should be made available to traders. In the markets examined in this dissertation traders are truthful, or behave boundedly rational. In the first case, traders offer their valuation for the asset or ask their cost, which is in general not rational. In the latter case traders are boundedly rational by only considering linear strategies or by using a learning algorithm that is based on the hypothetical payoff of strategies in the previous period. Boundedly rational behaviour is commonly modelled by putting mild restrictions on the strategy of traders or by learning algorithms. Such algorithms are used in agent-based models of financial markets since they do not impose strict assumptions on the behaviour of traders or their strategy space, and are considered in the second part of this dissertation. An underexposed type of market is a market in which trade occurs over a network, where the network structure is not entirely known to traders. An example is the spot foreign exchange market which is modelled in the first part of this dissertation by imposing mild restrictions on the strategy function of traders.

1.1 Network theory

Network theory is applicable to many research fields besides pure mathematics. In neuroscience, biological networks of the neural system are considered. In sociology networks are applied for instance to social media and relational connections. A common example in computer science is the use of networks in Google's PageRank and in operational research directed networks are used for transportation problems. In economic theory the banking crisis has led to a large literature on banking networks.

The seminal papers of Erdős and Rényi (1960, 1961) have introduced a mathematical theory on random graphs, often referred to as Erdős-Rényi graphs. We consider vertices in the network as traders and edges as links between traders. In these graphs traders are linked with an equal probability, independently of other connections. Erdős and Rényi derive phase transitions for infinitely many traders. During these phase transitions the structure of the network changes abruptly. The most surprising result of Erdős and Rényi occurs when the expected number of links per trader crosses the threshold value of one half. During this phase transition the structure of the graph changes from a collection of mainly isolated spanning trees to a network that contains a giant component of positive measure. Such a spanning tree connects a subset of traders of the graph but does not contain any cycle. Alon and Spencer (2008), Bollobás (1982) and Janson et al. (2000) summarise the work in the field of random graphs.

Markets over networks have been studied in various settings and trading mechanisms. In these markets trade may only occur between linked traders. The literature has in common that there is full knowledge of the network structure when traders determine their strategy. Spulber (2006) and Kranton and Minehart (2001) consider a market in which sellers jointly raise their ask until supply equals demand and trade occurs, known as simultaneously ascending-bid auctions. In Corominas-Bosch (2004) and Chatterjee and Dutta (1998) traders submit an offer side by side, which can be accepted or rejected by traders on the other side of the market. It is shown in Corominas-Bosch (2004) that the network can be split into different subgraphs in which the short side of the market extracts all the surplus, when all buyers have the same valuation and sellers the same cost. Intermediaries that act strategically and extract surplus are added in Easley and Kleinberg (2010) and Blume et al. (2009). In a market over a network, the power of a trader is measured in Calvó-Armengol (2001) on the basis of the number of linked traders and their links. The market power of a trader is higher when linked to more traders and when the linked traders have fewer links themselves. Moreover, a branch of network theory in economics and sociology studies the formation of links in a network, starting from Jackson and Wolinksy (1996). However, entirely random graphs are very seldomly studied in economic theory. These random graphs are important since they allow for studies on the effect of information about the network structure that is available to traders.

1.2 Learning algorithms

Learning algorithms are used in economic theory to model boundedly rational behaviour of traders. These algorithms are attractive because they do not make strict assumptions on the behaviour. For instance in reinforcement learning traders may learn to select the optimal strategy without having knowledge of the equilibrium. Genetic algorithms are developed in game theory for cobweb and overlapping generations models. In genetic algorithms every period a new generation of individuals is generated, depending on the fitness or profit of individuals in the previous period. Many agent-based models use learning to avoid making extreme assumptions about the rationality or strategies of traders. For example, the Individual Evolutionary Learning (IEL) algorithm is introduced in Arifovic and Ledyard (2003, 2007) to model the boundedly rational learning behaviour of agents in a Call Market model. In this learning algorithm traders learn to select from a pool of strategies, based on the hypothetical payoffs in the previous period. Moreover, this learning algorithm is used in a Continuous Double

Auction in Anufriev et al. (2013) to compare efficiency under full and no information about the history of others' strategies. Anufriev et al. (2013) also study the GS-environment from Gode and Sunder (1993, 1997) under the assumption that traders have zero intelligence and submit every possible offer with equal probability.

The introduction of the OpenBook system in 2002 by the New York Stock Exchange allows for studies on the effect of the information that is available to traders. This OpenBook system opened the content of the limit order book to the public, which allows experienced traders to use a full history of orders submission, instead of solely knowledge of global market statistics as under the former ClosedBook system. Boehmer et al. (2005) empirically show that this led to a decrease in price volatility and an increase in liquidity. The opening and closing of stock exchanges can be modelled with a Call Market. For such Call Markets Arifovic and Ledyard (2007) analyse experiments and simulations under the IEL algorithm, in which traders select strategies on the basis of their hypothetical performance in the previous period. Under the OpenBook system traders can directly determine the hypothetical performance of a strategy, assuming that other traders would have behaved the same. Under the ClosedBook system however, traders have to make additional assumptions to estimate the hypothetical foregone payoff of selecting another strategy. Arifovic and Ledyard (2007) show that in the OpenBook system agents try to influence the market clearing price. Agents behave as price makers and offers converge towards an equilibrium price. However, in the ClosedBook system traders learn to become pricetakers and offers diverge away from the equilibrium price range.

Anufriev et al. (2013) analyse the effect of the OpenBook system in a Continuous Double Auction. Agents enter the market and trade with an existing agent if possible. Otherwise their offers are stored in the order book until trade occurs with newly arriving traders or the book is emptied. In the IEL-algorithm the same hypothetical payoff functions as in Arifovic and Ledyard (2007) are used to value strategies. Anufriev et al. (2013) find the same bidding behaviour in a Continuous Double Auction as in the latter paper. They conclude that in the long-run, efficiency is similar in both designs and the price volatility is lower in the OpenBook system. Under this hypothetical payoff function in the formerly used ClosedBook system, where only information about past average prices is available, Anufriev et al. (2013) proved divergence of bids and asks away from the equilibrium price range. This results from the chosen ClosedBook hypothetical foregone payoff function, which only distinguishes between orders below and above the average price of the previous period. As a consequence investors trade with a high probability but may generate a very small profit. Anufriev et al. (2013) state however that ”the specification (of the ClosedBook hypothetical foregone payoff function) is a strong assumption ... which may affect (their) results of IEL”. Contrary to the latter paper, Fano et al. (2013) use a genetic algorithm in a setting closely related to the ClosedBook system, and show that traders behave as pricemakers and thus offers converge towards the equilibrium price. In this genetic algorithm, traders with the same valuation are compared on the basis of their average profit over some evaluation window, after which individuals with a low average profit take on strategies of better performing agents.

Starting from early contributions it is common in many agent-based models of order-driven financial markets that traders submit their order at a random moment during a trading session. Moreover, they are often required to make a one-dimensional decision, namely to choose a bid or ask price as in LiCalzi and Pellizzari (2006) or to forecast a future price as in Brock and Hommes (1997, 1998). For example, LiCalzi and Pellizzari (2006, 2007) compare efficiency in a Continuous Double Auction with other market protocols such as the Call Market, under boundedly rational respectively zero intelligent agents that arrive in a random sequence. Chiarella and Iori (2002) as well as Yamamoto and LeBaron (2010) use traders that submit their order at a random moment and use simple rules to make predictions about future prices, similar to Brock and Hommes (1998). In the classical financial literature many studies focus on limit and market orders. The surveys Gould et al. (2013a) and Hachmeister (2007) discuss the main theoretical, experimental and empirical papers on limit orders of informed and uninformed traders. Bae et al. (2003) and Biais et al. (1995) empirically find that the number of orders during a day follows a U-shaped distribution. Their reasoning behind this distribution is that at the beginning of the day traders desire to perform price discovery and react to events that occurred during the closing of the exchange and at the end of the day traders desire to unwind their positions. With the use of learning algorithms such as Individual Evolutionary Learning, agent-based models of financial markets can be extended to allow traders to submit their order at a chosen moment during a period. This requires an extension of the learning algorithm, in which traders are required to make a two-dimensional decision.

1.3 Dissertation outline

This dissertation consists of 4 chapters after this introduction. These chapters consider the effect of the available information in the market design on expected efficiency, in markets over networks when we assume that traders use linear markup strategies and in Continuous Double Auctions when traders use the Individual Evolutionary Learning algorithm to select their strategy. The first two chapters study efficiency in markets over random networks; in infinitely large markets when we assume that traders behave truthfully and in thin markets under boundedly rational behaviour of traders. The last two chapters consider the Individual Evolutionary Learning algorithm in Continuous Double Auctions. We introduce a new hypothetical foregone payoff function under no information about the history of others' actions and moreover extend the model by requiring traders to make the additional decision of choosing the timing of order submission.

Efficiency in Large Markets over Random Erdős-Rényi Networks

Chapter 2 follows Erdős and Rényi and derives phase transitions of bipartite graphs, depending on the probability of a link. Links are realised with the same probability, independently of each other. We find a similar transition of the bipartite graph, when the expected number of links per trader crosses the value one: the graph consists of many small isolated spanning trees below the threshold value and contains a giant component after the threshold. A market over random bipartite graphs with infinitely many traders is considered in the second part of this chapter. Agents desire to trade one unit and we assume that every trade yields the same surplus.

We study the restrictions of the network on the maximal efficiency, which can be calculated as the maximal expected number of trades divided by the number of traders on the thin side of the market, under identical valuations and costs of traders. The problem of finding the maximal number of trades is known as the Maximum Matching problem, studied for instance in Mucha and Sankowski (2004) and in West (1999). We derive bounds on expected efficiency as a function of the probability of a link, and improve these bounds for the range where the graph contains mainly spanning trees. An algorithm is introduced to construct all spanning trees and we determine the distribution of the degree of the vertices in a spanning tree.

Information and Efficiency in Thin Markets over Random Networks

A thin Erdős-Rényi market with two buyers and two sellers is considered in Chapter 3. Similar to the model of the spot foreign exchange market studied in Gould et al. (2013a), trades occur over links of the network. In contrast to their model we assume that links are realised with the same probability and independently of each other. Traders receive information about the network structure and behave strategically. We compare the equilibrium configurations for three nested information sets about the network structure; no, partial and full information. Under no information traders do not receive information about the realisation of links, but only the probability that a link is realised. The existence of one's links is given under partial information, as well as the probability of links of other traders. Under full information the entire network structure is revealed. We consider the effect of the amount of information on the allocative efficiency. This work shows that this effect is not only *non-monotonic*, but that a reversal of this non-monotonicity occurs when we switch from complete to incomplete information about traders' valuations. Contrary to Corominas-Bosch (2004), we show that under partial information about the network structure, or under incomplete information about valuations and costs, not all the surplus is necessarily extracted. Under complete information about valuations and costs, partial information about the network structure is weakly dominated. Under incomplete information about valuations and costs, we restrict attention to linear markup and markdown strategies. This type of strategies is introduced in Zhan and Friedman (2007) and a symmetric version is derived in Cervone et al. (2009). Myerson and Satterthwaite (1983) and

Chatterjee and Samuelson (1983) show for bilateral trading that Nash equilibrium strategies are monotone and piecewise linear transformations of valuations into offers. For the subset of linear markup strategies, partial information about the network structure strongly dominates full and no information.

On the role of Information under Individual Evolutionary Learning in a Continuous Double Auction

In Chapter 4 we demonstrate through simulations that the specification of the hypothetical foregone payoff functions indeed plays a crucial role in a Continuous Double Auction model under the IEL learning algorithm, as suggested by Anufriev et al. (2013). Traders use the payoff function to estimate how other strategies would have performed in the previous period. Under their hypothetical foregone payoff function bids and asks diverge away from the equilibrium price range in de ClosedBook system. This work, jointly with Mikhail Anufriev, Jasmina Arifovic and Valentyn Panchenko, introduces a new foregone payoff function, that uses more information to estimate the hypothetical foregone payoff of each possible offer, which results in bids and offers drifting towards an equilibrium price similar to Fano et al. (2013). Under this payoff function investors learn to increase their expected profit by submitting an order that has a higher possible profit. This results in a lower probability of trading, but this effect is outweighed by an increase in possible profit from trade. First we perform simulations during the learning phase of a Continuous Double Auction, to study the effect of the OpenBook system. We compare with the results of the simulations in the Call Market performed by Arifovic and Ledyard (2007), by comparing efficiency between both markets. Second, we examine the effect of the OpenBook system during long-run simulations. This allows for a comparison of the new ClosedBook hypothetical foregone payoff function with the function used in Anufriev et al. (2013). Thirdly we show robustness of our results with respect to the size of the market and the number of units a trader desires to buy or sell. As indicated in Anufriev et al. (2013) the specification of the hypothetical foregone payoff function indeed plays a crucial role and largely affects their main results.

Timing under Individual Evolutionary Learning in a Continuous Double Auction

In Chapter 5 we extend the IEL-algorithm used in Arifovic and Ledyard (2003, 2007) and in Anufriev et al. (2013) by introducing learning about the timing of order submission. In this joint work with Mikhail Anufriev, traders submit a multidimensional strategy which allows for contemporaneous learning about the submitted order and the moment of submission. In a benchmark environment with complete information about the trading history in the previous period, we study the distribution of submission moments under the extended IEL algorithm and the interrelation between the submission moments and the orders. This chapter is a step forward to a more complete model of learning in markets and is distinguished from previous research by the decision traders are required to make. Instead of a one-dimensional decision traders are required to make a two-dimensional decision; which bid or ask to submit and when to submit this offer during the trading session. We show that traders in medium size markets learn to submit around the middle of the trading session to avoid a lower profit or trading probability. Moreover, we consider the impact of competition and the size of the market on the timing of the submission. We conclude that the size of the market highly influences the preferred arrival moment. We show that the effect of the extra decision that traders are required to make is negative, by comparing general market statistics with Anufriev et al. (2013), where traders submit at a random moment during the trading period.

Chapter 2

Efficiency in Large Markets over Random Erdős-Rényi Networks

2.1 Introduction

Random graphs have been of interest since the seminal papers of Erdős and Rényi (1960, 1961). In these papers the random graph is introduced and phase transitions are derived as the number of vertices converges to infinity. The main result is that a phase transition occurs as the expected number of edges per vertex crosses the threshold value $\frac{1}{2}$. During such a phase transition the structure of the graph changes dramatically; up to the threshold the graph consists mainly of isolated trees whereas after the phase transition a giant component of positive measure arises. The work in the field of random graphs has been summarised in Alon and Spencer (2008), Bollobás (1982) and Janson et al. (2000).

The work of Erdős and Rényi on phase transitions in random graphs has not been thoroughly extended to bipartite graphs, which are graphs whose vertices can be divided in two disjoint sets in such a way that edges only occur between the sets. In this chapter we derive the phase transitions of a bipartite graph depending on the probability of an edge. We find a similar transition of the graph at the value 1; below the threshold the graph is a collection of mainly isolated spanning trees and after the transition a giant component emerges.

We consider a market over such a random bipartite graph, in which buyers and sellers are randomly linked with a certain probability. There is an equal number of buyers and sellers, who all desire to trade one unit of the good and we consider the case where the number of traders converges to infinity. For simplicity buyers assign a value of one to the good and sellers have a cost of zero. We assume that traders behave truthfully and bid one or ask zero. We study the maximal set of trades in the random bipartite graph, which depends on the characteristics of the different phases. For this so-called Maximum Matching problem many algorithms have been found, f.i. in Mucha and Sankowski (2004) and West (1999). We derive bounds on the expected efficiency, which under these simplifications can be calculated by dividing the expected number of trades in the maximum matching, by the number of traders on one side of the market. We derive an algorithm to construct all spanning trees and the distribution of the degree of the vertices. This allows for a development of tighter bounds on expected efficiency in the range consisting of mainly spanning trees.

The organisation of this chapter is as follows. The model and graph theory are considered in Section 2.2, followed by the phases of random bipartite graphs in Section 2.3. Section 2.4 derives bounds on expected efficiency in an infinitely large market over such networks. Finally, Section 2.5 concludes. The proofs are given in an appendix.

2.2 Model

We consider a market with n buyers and n sellers and we let n converge to infinity. Buyer i and seller j are linked with each other with a fixed probability p, independent of other links. Trade occurs only between linked traders. A related example of a market over networks is the spot exchange market studied in Gould et al. (2013a). In this market, traders provide a blocklist that excludes some traders on the opposite side of the market as possible trading partners. Trades are possible when both traders are not included in the blocklist of the other. The blocklist is used to protect against adverse selection and to control counterparty risk, and is thus considered exogenous. However, in contrast to the spot exchange market we assume that links are realised with equal probability and independently of each other.

Traders desire to obtain or sell one unit of a good. The valuation of a buyer equals one and the cost of a seller is zero, and there is complete information about valuations and costs. A buyer receives a profit equal to his valuation minus the transaction price after a trade, and zero otherwise. The profit of a seller that trades equals the transaction price minus his cost, otherwise the profit equals zero. We consider the maximal expected efficiency given the restrictions of the network structure and thus assume that traders are truthful and bid one or ask zero. Expected efficiency is defined as the maximal expected surplus under the network structure divided by the maximal surplus in a complete network. Because every trade results in the same surplus, it is sufficient to determine the fraction of transactions.

2.2.1 Graph theory

The market can be considered as a random bipartite graph, which is an extended Erdős-Rényi network. Two sets of labelled *vertices* V^1 and V^2 denote the sets of buyers and sellers and the set of *edges* E represents the links between traders. A graph is called *bipartite* when every edge connects a vertex in V^1 with a vertex in V^2. We consider the *number of edges* $N(n)$ as a function of the number of traders n on one side of the market; the *probability of an edge* equals $p = \mathbb{E}(\frac{N(n)}{n^2})$.

A graph G_2 is called a *subgraph* of G_1 if the vertices V_2^1 and V_2^2 of G_2 are subsets of the vertices V_1^1 and V_1^2 of G_1, and the edges E_2 of G_2 are a subset of the edges E_1 of G_1. A subgraph is called of *size k,l* if it is constructed from k and l labelled vertices. A subgraph is an *isolated* subgraph when either both or neither of the endpoints of an edge in E_1 belong to the subgraph, i.e. a vertex in the subgraph cannot be linked with a vertex outside the subgraph.

We define different types of subgraphs. A sequence of m attached edges is called a *path of size* m. A graph is *connected* if there is a path between every pair of vertices. A connected isolated subgraph of G_1 is denoted as a *component* of G_1. A *spanning tree* of size k,l occurs when k and

l vertices are connected by exactly $k+l-1$ edges. A *cycle* of size k,k occurs when k and k vertices are connected by at least $2k$ edges and a path of size $2k$ exists. In a *complete* bipartite graph an edge exists between any point in V^1 and any point in V^2.

Two graphs are *isomorphic* if there is a one-to-one correspondence between the vertices and the edges of both graphs. The *degree* of a graph G is the average number of edges of the vertices. A graph G is *balanced* if it contains no subgraph that has a larger degree than G itself.

As we consider asymptotic behaviour of the graph we often use the order of variables. The *little o* notation $a(n) = o\left(b(n)\right)$ denotes that $\lim_{n\to\infty} \frac{|a(n)|}{b(n)} = 0$ which indicates that $b(n)$ grows much faster than $a(n)$. Functions have the same growth rate when $\frac{|a(n)|}{b(n)}$ is bounded, which is indicated with the *big O* notation $a(n) = O\left(b(n)\right)$. Two functions are *similar*, denoted as $a(n) \sim b(n)$, when they are asymptotically equal and thus $\lim_{n\to\infty} \frac{a(n)}{b(n)} = 1$.

For a given property D^*, the function $D(n)$ is called a *threshold function* with respect to $N(n)$ if D^* almost surely (a.s.) is not satisfied when the ratio $\frac{N(n)}{D(n)}$ converges to zero, and a.s. is satisfied if the ratio converges to infinity:

$$\lim_{n\to\infty} \mathbb{P}_{n,N(n)}(D^*) = \begin{cases} 0 & \text{if } \lim_{n\to\infty} \frac{N(n)}{D(n)} = 0 \\ 1 & \text{if } \lim_{n\to\infty} \frac{N(n)}{D(n)} = \infty. \end{cases}$$

2.3 Phase transitions bipartite graphs

We consider the phase transitions for a bipartite graph, similar to Erdős and Rényi (1960, 1961), as the probability of an edge increases. From phase to phase the network structure of the market changes abruptly. As the probability increases the market evolves from a collection of larger and larger spanning trees to a market that contains cycles; and eventually a giant central market emerges that contains a positive fraction of all traders. In the next section we derive tighter bounds when the market consists almost surely (a.s.) solely of spanning trees. We prove most theorems, shown in the appendix, by considering the number of edges N. The Law of Large Numbers implies that if p is of some order, the number of realised links N is of order pn^2 almost surely. Hence these results also hold for the generalised random bipartite graph.

Phase 1: $p = o(\frac{1}{n}) \iff N = o(n)$

In this phase the random graph consists a.s. solely of connected subgraphs that are spanning trees (Th. 2.3.5). Hence, a.s. there are no cycles (Cor. 2.3.2). Spanning trees of size k,l only exist from the threshold $n^{2-\frac{k+l}{k+l-1}}$ on (Cor. 2.3.1). For $p \sim \rho n^{2-\frac{k+l}{k+l-1}}$ the number of spanning trees of size k,l follows a Poisson distribution with $\lambda = \frac{\rho^{k+l-1}k^{l-1}l^{k-1}}{k!l!}$ (Th. 2.3.2).

Hence during this phase the expected number of links per trader converges to zero and the market consists of infinitely many isolated submarkets up to a certain size.

Phase 2: $p \sim \frac{c}{n} \iff N \sim cn,$ for $c \leq 1$

Besides spanning trees, also cycles occur in the graph. For $c < 1$, the probability that the bipartite graph contains at least one cycle equals $1 - \sqrt{1-c^2}\mathrm{e}^{\frac{c^2}{2}}$, which is strictly smaller than 1 (Th. 2.3.8). The number of cycles of size k,k follows a Poisson distribution with $\lambda = \frac{1}{2k}c^{2k}$ (Th. 2.3.3), whereas the number of isolated cycles of size k,k follows a Poisson distribution with $\lambda = \frac{1}{2k}(c\mathrm{e}^{c})^{2k}$ (Th. 2.3.4). The total expected number of cycles is given by $\frac{1}{2}\log(\frac{1}{1-c^2}) - \frac{c^2}{2}$ (Th. 2.3.7) and the expected number of vertices that belong to a cycle equals $\frac{c^4}{1-c^2}$ (Th. 2.3.9).

Even though cycles emerge, almost every vertex belongs to a spanning tree (Th. 2.3.6). Moreover, the total number of components is given by $n - N + O(1)$ (Th. 2.3.10) and hence almost every component is a spanning tree. The possible cycles in the bipartite graph are thus negligible. The maximum number of spanning trees of size k,l, $\frac{k^{l-1}l^{k-1}}{k!l!} \cdot (\frac{k+l-1}{k+l})^{k+l-1}\mathrm{e}^{-(k+l-1)}$, is attained for $p \sim \frac{1}{n} \cdot \frac{k+l-1}{k+l}$ (Th. 2.3.2). In this phase spanning trees of all sizes exist.

For $c = 1$ the graph almost surely contains a cycle (Th. 2.3.8) and the total number of cycles is of order $\frac{1}{2}\log(n)$ (Th. 2.3.7). The expected number of components is given by $n - N + O(\log(n))$ (Th. 2.3.10).

The expected number of links per trader in this phase is given by the value c. Almost every of $n - N + O(1)$ submarkets is a spanning tree and almost every trader is part of a spanning tree.

Phase 3: $p \sim \frac{c}{n} \iff N \sim cn, \text{ for } c > 1$

As the expected number of edges exceeds 1 the structure of the bipartite graph undergoes a sudden change. The probability that a vertex belongs to a spanning tree is smaller than 1 and equals $\frac{x(c)}{c}$, where $x(c) = \sum_{v=1}^{\infty} \frac{v^{v-1}(ce^{-c})^v}{v!}$ and $v = k + l$ (Th. 2.3.6). The number of components is given by $\frac{2n}{c}\left(x(c) - \frac{x(c)^2}{2}\right)$ (Th. 2.3.10). The greatest component covers a set of vertices of positive measure, which follows directly from Blasiak and Durrett (2005).

The expected number of trading partners exceeds 1 and now a giant central market arises that covers a positive fraction of the total market. Around the central market smaller and smaller submarkets exist.

Phase 4: $pn \to \infty$

As the expected number of links converges to infinity, almost surely every trader is part of the central market. With probability zero small submarkets exist and thus the number of components is of order $O(1)$ (Th. 2.3.10).

2.4 Bounds on expected efficiency

The different phases determined in the previous section allow us to consider the restrictions of the network structure on the number of trades. Assuming truthful traders with equal valuations and costs, the number of trades corresponds directly to the extracted surplus. The expected maximal efficiency under the random network structure equals the expected maximum number of trades divided by n. The problem of calculating this expected maximal efficiency reduces to the Maximum Matching problem; this is a matching at which the number of trades is maximised. A matching in which all traders can trade is called a perfect matching. Many algorithms have been established to determine the maximum matching of a given bipartite graph.

Translating the network into a matrix allows for some necessary and sufficient conditions for a perfect matching. We can represent the network by a matrix, where rows correspond to the buyers and columns to the sellers. A value of 1 at place i,j denotes a link between buyer i and seller j; the value 0 denotes the absence of a link. A perfect matching is available iff either:

- All diagonal elements equal 1, possibly after permuting rows and/or columns.
- Every subset of sellers is linked to a subset of buyers with at least the same cardinality, often referred to as the Marriage theorem of Hall (1935).
- There does not exist a block of zeros of size $k \cdot l$ with $k + l > n$.

2.4.1 Example

The expected maximal efficiency is calculated exactly for $n = 1, ..., 4$ by determining for every number of existing links N the number of possibilities of having a certain number of maximal trades t. For example for $n = 2$ the $2^4 = 16$ possible realisations of the network are given in Fig. 2.1, where the two buyers are shown on top and the two sellers on the bottom.

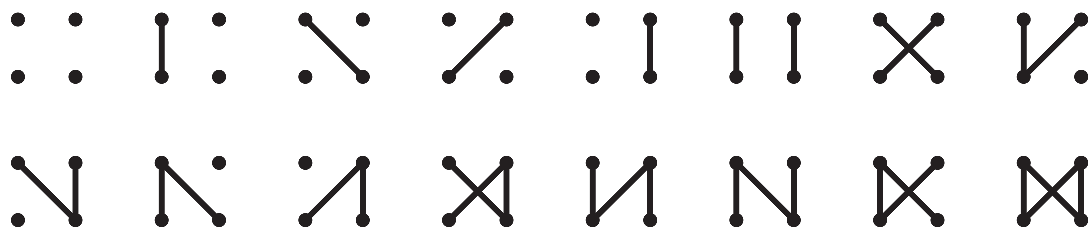


Figure 2.1: Possible network realisations for 2 buyers and 2 sellers.

We show the distribution of the maximal number of trades per number of links for $n = 3$ buyers and sellers; thus there are $2^9 = 512$ possible realisations of the network. For every possible realisation we determined the maximal number of trades and the number of links. The number of realisations with N links and a maximal possible number of trades t is shown in Table 2.1.

t \ N	0	1	2	3	4	5	6	7	8	9
0	1	0	0	0	0	0	0	0	0	0
1	0	9	18	6	0	0	0	0	0	0
2	0	0	18	72	90	45	6	0	0	0
3	0	0	0	6	36	81	78	36	9	1

Table 2.1: Example with 3 buyers and 3 sellers which shows the number of realisation of the network with N links and a maximal possible number of trades t.

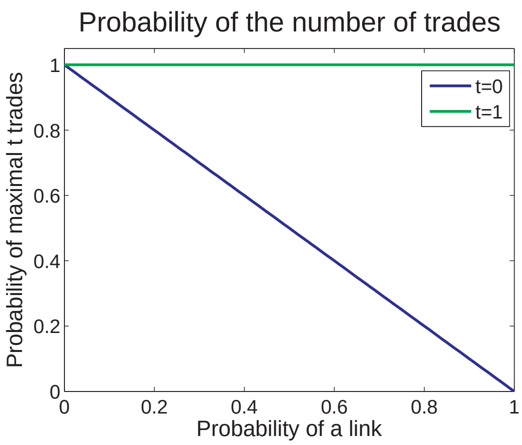
These calculations allow us to determine the distribution of the number of trades t for $n = 1, ..., 4$ buyers and n sellers, shown in Fig. 2.2.

For $n = 1, ..., 4$ buyers and sellers we show respectively the probability of full efficiency and the expected efficiency in Fig. 2.3. We observe that the probability of full efficiency increases for large p and decreases for small p. The expected efficiency is increasing in n because the expected number of links per trader increases.

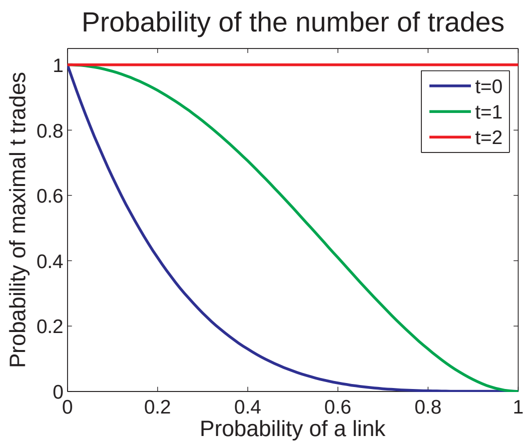
2.4.2 Infinitely many traders

The expected maximal efficiency due to restrictions of the network structure is of interest in this section in a market with infinitely many traders. As mentioned before, this market is related to the spot exchange market studied in Gould et al. (2013a). For the different phases of the random bipartite graph expected efficiency for the entire market is calculated. As the expected number of links converges to zero, we find that the expected efficiency converges to zero. When the expected number of links however converges to a constant c we derive a lower bound of $1 - \frac{1-\mathrm{e}^{-c}}{c}$ and an upper bound of $1 - \mathrm{e}^{-c}$. Finally as the market becomes almost surely connected the probability that full efficiency is attained converges to one (Th. 2.4.1).

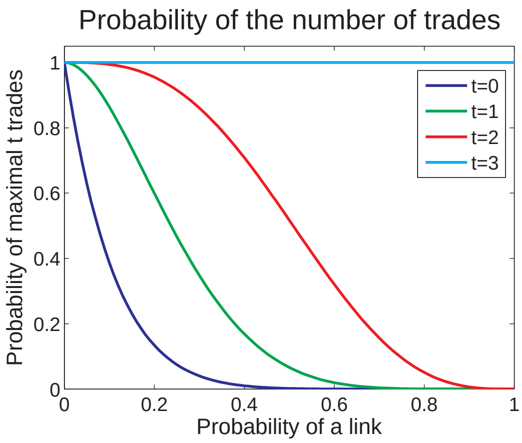
In the range $p = \frac{c}{n},\ c \leq 1$ cycles are negligible and almost every vertex of the bipartite graph belongs to a spanning tree. Hence bounds for expected efficiency can be derived for spanning trees individually and added up. We formally show that the expected efficiency is continuous



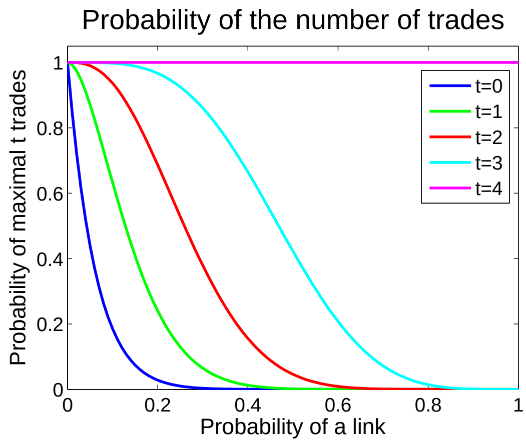
(a) 1 buyer and 1 seller.



(b) 2 buyers and 2 sellers.



(c) 3 buyers and 3 sellers.

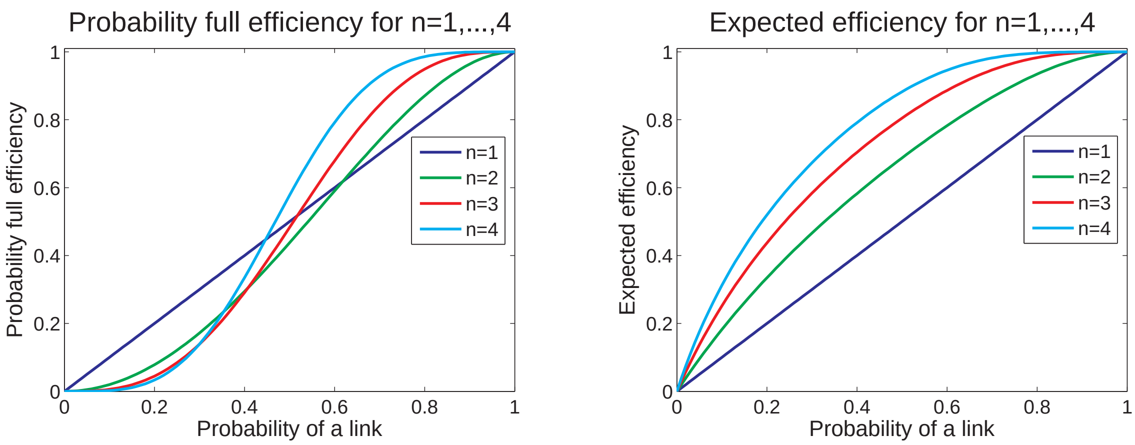


(d) 4 buyers and 4 sellers.

Figure 2.2: Distribution of the number of trades for $n = 1, ..., 4$ buyers and sellers. For every value of the probability of a link, the probability of maximal $t = 0, ..., n$ trades is given.

and especially at the point $c = 1$ of the phase transition (Th. 2.4.2).

In order to derive tighter bounds on expected efficiency in the range $p = \frac{c}{n},\ c \leq 1$ we construct an algorithm that produces all possible, undirected, spanning trees of a certain size by adding vertices one by one to a directed tree. We show that this algorithm produces exactly all spanning trees (Th. 2.4.3).



(a) Probability of full efficiency. (b) Expected efficiency.

Figure 2.3: Efficiency as a function of the size of the market.

Algorithm 2.4.1: Constructing all possible spanning trees of a bipartite graph

All possible, undirected, spanning trees of size k,l can be constructed by forming a directed tree step by step. We denote the vertices by $V^1 = \{v_1^1, ..., v_k^1\}$ and $V^2 = \{v_1^2, ..., v_l^2\}$ respectively. The set of spanning trees is equivalent to the set of directed spanning trees with root v_1^1. This algorithm produces layer by layer all the possible spanning trees:

Step 1

The node v_1^1 is linked to a non-empty subset of V^2. This subset is removed from V^2 and v_1^1 is removed from V^1.

Step 2

All the vertices that are added to the directed tree in the last step are linked to a group of disjoint subsets of the other set of vertices that satisfy:

- The number of subsets in the group equals the number of vertices added in the last step.
- The union of the group of subsets is non-empty.
- The vertices linked to the same predecessor are ordered; to prevent counting isomorphic spanning trees multiple times.

The vertices in the group of subsets are removed from the set of vertices and this step is repeated until one set of remaining vertices is empty.

Step 3

All the vertices that are added to the directed tree in the last step are linked to a group of disjoint subsets of the other set of vertices that satisfy:

- The number of subsets in the group equals the number of vertices added in the last step.
- The union of the group is equal to the set of remaining vertices.
- The vertices linked to the same predecessor are ordered; to prevent counting isomorphic spanning trees multiple times.

This algorithm can easily be extended to multipartite graphs. In every step vertices are added that are a subset of the other sets of vertices. When all but one set of vertices is empty the algorithm moves on to Step 3. The distribution of the degrees of vertices in spanning trees can be determined using Algorithm 2.4.1. We show that every vertex in a spanning tree naturally has one edge and that the remaining edges are multinomially distributed per set of vertices (Th. 2.4.4).

This allows for tighter bounds on expected efficiency by considering the number of vertices with a degree larger than one, $\#V_{\text{degree}>1}$. This number of vertices can be calculated from the multinomial distribution of the remaining edges. We derive a lower bound of $\lceil\frac{\#V_{\text{degree}>1}+1}{2}\rceil$ on the expected efficiency in a spanning tree of size $k+l > 2$ and an upper bound of $min\,(k, l, \#V_{\text{degree}>1})$ (Th. 2.4.5). Together it can provide bounds on the expected maximal efficiency of the entire market when almost every component is a spanning tree. Considering spanning trees separately we find tighter bounds on expected maximal efficiency in the range $p \sim \frac{c}{n},\ c \leq 1$:
$\sum_{1\leq i\leq k\leq\infty}\sum_{1\leq j\leq l\leq\infty}\frac{\mathrm{e}^{-(k+l)}}{i!j!}S(l-1,k-i)S(k-1,l-j)\lceil\frac{k-i+l-j+1}{2}\rceil \leq \mathbb{E}(\text{eff})$
$\leq \sum_{1\leq i\leq k\leq\infty}\sum_{1\leq j\leq l\leq\infty}\frac{\mathrm{e}^{-(k+l)}}{i!j!}S(l-1,k-i)S(k-1,l-j)\min(k,l,k-i+l-j)$ (Th. 2.4.6).

These bounds are approximated by evaluating them for $k + l \leq 140$. Fig. 2.4 shows that these bounds are indeed tighter than the bounds found when the graph is considered as a whole.

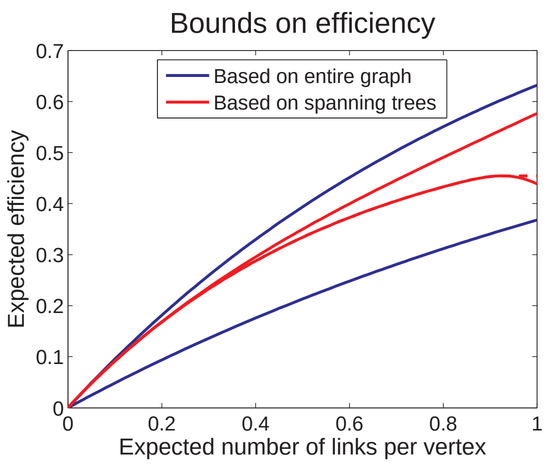


Figure 2.4: Bounds on expected efficiency on the basis of the entire graph and on the basis of spanning trees.

2.5 Concluding remarks

Following Erdős and Rényi (1960, 1961) we have constructed phase transitions for random bipartite graphs, where links are realised independently from each other with probability p. In the phase $p = o(\frac{1}{n})$ the graph consists of isolated spanning trees up to a certain size. The phase $p \sim \frac{c}{n}$, $c \leq 1$ is characterised by a graph where almost every component is a spanning tree. Such spanning trees occur of every size. The number of spanning trees follows a Poisson distribution. The greatest component is a spanning tree with zero measure. As c crosses the value 1 for $p \sim \frac{c}{n}$, the behaviour of the graph undergoes a sudden change. Besides spanning trees and small cycles, the graph contains a giant component of positive measure. As the expected number of edges per vertex converges to infinity, $p \cdot n \to \infty$, almost every vertex belongs to the giant component.

Using these phases we could find bounds for the expected efficiency in a market setting, for individual spanning trees and in general. We considered an equal number of buyers and sellers, who all desire to trade one unit of the good and consider the case where the number of traders converges to infinity. The results hold under the assumption that traders are truthful and bid or ask their valuation of 1 and cost 0 respectively. Under these settings the problem of finding the expected maximal efficiency reduces to the Maximum Matching problem. Moreover, the expected maximal efficiency can be calculated by dividing the expected number of trades in the maximum matching, by the number of traders on one side of the market. When the expected number of edges per vertex converges to zero or infinity, the expected efficiency converges to zero respectively one. In the range $p \sim \frac{c}{n}$ we have found a lower bound of $1 - \frac{1-\mathrm{e}^{-c}}{c}$ and an upper bound of $1 - \mathrm{e}^{-c}$ on expected efficiency.

These bounds can be improved in the range $p \sim \frac{c}{n},\ c \leq 1$ by considering the expected maximal efficiency of spanning trees separately. We introduced a new algorithm to construct all the spanning trees of a certain size and determined the distribution of the degrees of the vertices in spanning trees. In the phase where the bipartite graph consists mainly of spanning trees and other components can be neglected, the tighter bounds $\sum_{1 \leq i \leq k \leq \infty} \sum_{1 \leq j \leq l \leq \infty} \frac{\mathrm{e}^{-(k+l)}}{i!j!} S(l-1, k-i) S(k-1, l-j) \lceil \frac{k-i+l-j+1}{2} \rceil \leq \mathbb{E}(\text{eff})$
$\leq \sum_{1 \leq i \leq k \leq \infty} \sum_{1 \leq j \leq l \leq \infty} \frac{\mathrm{e}^{-(k+l)}}{i!j!} S(l-1, k-i) S(k-1, l-j) \min(k, l, k-i+l-j)$ are determined by considering the spanning trees separately.

As an extension of Erdős and Rényi (1960, 1961) we have found similar phase transitions for random bipartite graphs. Under an assumption about the distribution of links, random bipartite graphs describe the spot exchange market and we have derived bounds on expected maximal efficiency for every phase.

Appendix A: Theorems in Section 3

We prove most theorems by considering the number of edges N instead of the probability of a link p. The Law of Large Numbers implies that if p is of some order, the number of realised links N is of order pn^2 almost surely. Hence these results also hold for the generalised random bipartite graph. This random bipartite graph of size n,n with N edges is denoted as $\Gamma_{n,N}$.

Theorem 2.3.1

Let $k + l \geq 3$ and $k + l - 1 \leq m \leq kl$ be positive integers. $\mathbb{B}_{k,l,m}$ denotes the non-empty set of connected balanced bipartite graphs of size k,l and m edges. The threshold function for the existence of at least one subgraph isomorphic with an element in $\mathbb{B}_{k,l,m}$ is $N = O(n^{2-\frac{k+l}{m}})$.

Proof

Let $\mathbf{B}_{k,l,m} \geq 1$ be the number of graphs in $\mathbb{B}_{k,l,m}$ that can be constructed from k and l labelled vertices. $\mathbb{P}_{n,N}(\mathbb{B}_{k,l,m})$ is the probability that the random graph $\Gamma_{n,N}$ contains at least one subgraph that is isomorphic to one of the elements in $\mathbb{B}_{k,l,m}$ and can be bounded by $\mathbb{P}_{n,N}(\mathbb{B}_{k,l,m})$ $\leq \binom{n}{k}\binom{n}{l}\mathbf{B}_{k,l,m}\frac{\binom{n^2-m}{N-m}}{\binom{n^2}{N}} = O(n^k n^l \frac{(n^2-m)^{N-m}}{(n^2)^N} \cdot \frac{N!}{(N-m)!}) = O(\frac{N^m}{n^{2m-k-l}})$. This holds since as $n \to \infty$, $\binom{n}{k} = \frac{n!}{(n-k)!k!} = O(n^k)$ for $k \geq 1$ and $\binom{n^2-m}{N-m} = \frac{(n^2-m)!}{(n^2-m-(N-m))!(N-m)!}$ $=O(\frac{(n^2-m)^{N-m}}{(N-m)!})$ for arbitrary N.

The k and l labelled vertices can be selected in $\binom{n}{k}\binom{n}{l}$ different ways and the m edges can form an element of $\mathbb{B}_{k,l,m}$ in $\mathbf{B}_{k,l,m}$ ways. The remaining $N - m$ edges can be selected from the remaining $n^2 - m$ possible edges. The above expression is only an upper bound, since graphs that contain multiple subgraphs isomorphic with an element of $\mathbb{B}_{k,l,m}$ are counted multiple times. Hence it remains to show that the graph contains a subgraph isomorphic with an element of $\mathbb{B}_{k,l,m}$ if N is at least of the order $n^{2-\frac{k+l}{m}}$.

We denote the set of all subgraphs S of $\Gamma_{n,N}$ that are isomorphic with an element of $\mathbb{B}_{k,l,m}$ by $\mathbb{B}^{(n)}_{k,l,m}$. Then $\mathbb{E}(\sum_{S\in\mathbb{B}^{(n)}_{k,l,m}} \mathbb{1}_{\{S\in\Gamma_{n,N}\}}) = \sum_{S\in\mathbb{B}^{(n)}_{k,l,m}} \mathbb{E}(\mathbb{1}_{\{S\in\Gamma_{n,N}\}}) = \binom{n}{k}\binom{n}{l}\mathbf{B}_{k,l,m}\frac{\binom{n^2-m}{N-m}}{\binom{n^2}{N}}$

$\sim \frac{\mathbf{B}_{k,l,m}}{k!l!} \cdot \frac{N^m}{n^{2m-k-l}}$.

For two elements $S_i, S_j \in \mathbb{B}^{(n)}_{k,l,m}$ that do not share an edge we find that $\mathbb{E}(\sum_{S_i,S_j\in\mathbb{B}^{(n)}_{k,l,m}} \mathbb{1}_{\{S_i,S_j\in\Gamma_{n,N}\}}) = \sum_{S_i,S_j\in\mathbb{B}^{(n)}_{k,l,m}} \mathbb{E}(\mathbb{1}_{\{S_i,S_j\in\Gamma_{n,N}\}}) \leq \binom{n}{2k}\binom{n}{2l}\mathbf{B}^2_{k,l,m}\frac{\binom{n^2-2m}{N-2m}}{\binom{n^2}{N}}$
$\leq \left(\binom{n}{k}\binom{n}{l}\mathbf{B}_{k,l,m}\frac{\binom{n^2-m}{N-m}}{\binom{n^2}{N}}\right)^2 \sim \left(\mathbb{E}(\sum_{S\in\mathbb{B}^{(n)}_{k,l,m}} \mathbb{1}_{\{S\in\Gamma_{n,N}\}})\right)^2$.

For two elements $S_i, S_j \in \mathbb{B}^{(n)}_{k,l,m}$ that share s,t vertices and $1 \leq r \leq m-1$ edges we find that $\mathbb{E}(\mathbb{1}_{\{S_i,S_j\in\Gamma_{n,N}\}}) = \frac{\binom{n^2-2m+r}{N-2m+r}}{\binom{n^2}{N}} = O(\frac{N^{2m-r}}{n^{4m-2r}})$.

Since all S_i are balanced the degree of the intersection of S_1 and S_2 should be less than the degree of the subgraph S_1 (and also S_2): $\frac{r}{s+t} \leq \frac{m}{k+l}$. Hence $s+t \geq \frac{r(k+l)}{m}$, and thus the number of such pairs of subgraphs S_i, S_j is bounded by $\mathbf{B}^2_{k,l,m}\sum_{s=1}^{k}\sum_{t=\frac{r(k+l)}{m}-s}^{l}\binom{n}{k}\binom{n}{l}\binom{k}{s}\binom{l}{t}\binom{n-k}{k-s}\binom{n-l}{l-t}$
$= O(\mathbf{B}^2_{k,l,m}\sum_{s=1}^{k}\sum_{t=\frac{r(k+l)}{m}-s}^{l}\frac{n^k n^l k^s l^t (n-k)^{k-s}(n-l)^{l-t}}{k!l!s!t!(k-s)!(l-t)!})$
$= O\left(\sum_{s=1}^{k}\sum_{t=\frac{r(k+l)}{m}-s}^{l} n^k n^l (n-k)^{k-s}(n-l)^{l-t}\right) = O(\sum_{s=1}^{k}\sum_{t=\frac{r(k+l)}{m}-s}^{l} n^{2(k+l)-s-t})$
$= O(n^{2(k+l)-\frac{r(k+l)}{m}})$, since $s+t \geq \frac{r(k+l)}{m}$. So $\mathbb{E}(\sum_{S_i,S_j\in\mathbb{B}^{(n)}_{k,l,m}} \mathbb{1}_{\{S_i,S_j\in\Gamma_{n,N}\}})$
$= O\left((\frac{N^m}{n^{2m-(k+l)}})^2\sum_{r=1}^{m-1}(\frac{n^{2-\frac{k+l}{m}}}{N})^r\right)$.

We combine the above results and find that
$\mathbb{E}\left((\sum_{S\in\mathbb{B}^{(n)}_{k,l,m}} \mathbb{1}_{\{S\in\Gamma_{n,N}\}})^2\right) = \mathbb{E}(\sum_{S_i,S_j\in\mathbb{B}^{(n)}_{k,l,m}} \mathbb{1}_{\{S_i,S_j\in\Gamma_{n,N}\}})$
$\leq \mathbb{E}(\sum_{S\in\mathbb{B}^{(n)}_{k,l,m}} \mathbb{1}_{\{S\in\Gamma_{n,N}\}}) + \left(\mathbb{E}(\sum_{S\in\mathbb{B}^{(n)}_{k,l,m}} \mathbb{1}_{\{S\in\Gamma_{n,N}\}})\right)^2 + O\left((\frac{N^m}{n^{2m-(k+l)}})^2\sum_{r=1}^{m-1}(\frac{n^{2-\frac{k+l}{m}}}{N})^r\right)$.

For $\frac{N^m}{n^{2m-k-l}} = \omega \to \infty$ it holds that
$\text{Var}(\sum_{S\in\mathbb{B}^{(n)}_{k,l,m}} \mathbb{1}_{\{S\in\Gamma_{n,N}\}}) = \mathbb{E}\left((\sum_{S\in\mathbb{B}^{(n)}_{k,l,m}} \mathbb{1}_{\{S\in\Gamma_{n,N}\}})^2\right) - \left(\mathbb{E}(\sum_{S\in\mathbb{B}^{(n)}_{k,l,m}} \mathbb{1}_{\{S\in\Gamma_{n,N}\}})\right)^2$
$= \mathbb{E}(\sum_{S\in\mathbb{B}^{(n)}_{k,l,m}} \mathbb{1}_{\{S\in\Gamma_{n,N}\}}) + \left(\mathbb{E}(\sum_{S\in\mathbb{B}^{(n)}_{k,l,m}} \mathbb{1}_{\{S\in\Gamma_{n,N}\}})\right)^2 + O\left((\frac{N^m}{n^{2m-k-l}})^2\sum_{r=1}^{m-1}(\frac{n^{2-\frac{k+l}{m}}}{N})^r\right)$
$- \left(\mathbb{E}(\sum_{S\in\mathbb{B}^{(n)}_{k,l,m}} \mathbb{1}_{\{S\in\Gamma_{n,N}\}})\right)^2 = \frac{(\mathbb{E}(\sum_{S\in\mathbb{B}^{(n)}_{k,l,m}} \mathbb{1}_{\{S\in\Gamma_{n,N}\}}))^2}{\mathbb{E}(\sum_{S\in\mathbb{B}^{(n)}_{k,l,m}} \mathbb{1}_{\{S\in\Gamma_{n,N}\}})} + O\left((\frac{N^m}{n^{2m-k-l}})^2\sum_{r=1}^{m-1}(\frac{n^{2-\frac{k+l}{m}}}{N})^r\right)$
$= O\left(\frac{(\mathbb{E}(\sum_{S\in\mathbb{B}^{(n)}_{k,l,m}} \mathbb{1}_{\{S\in\Gamma_{n,N}\}}))^2}{\frac{N^m}{n^{2m-k-l}}}\right) = O\left(\frac{(\mathbb{E}(\sum_{S\in\mathbb{B}^{(n)}_{k,l,m}} \mathbb{1}_{\{S\in\Gamma_{n,N}\}}))^2}{\omega}\right)$.

Now we are able to use Chebysheff's inequality, which states that $\mathbb{P}(|X-\mu| \geq h\sigma) \leq \frac{1}{h^2}, \forall h > 0$. For $h = \frac{1}{2}\sqrt{\omega}$ we find that

$\mathbb{P}\left(|\sum_{S\in\mathbb{B}_{k,l,m}^{(n)}} \mathbb{1}_{\{S\in\Gamma_{n,N}\}} - \mathbb{E}(\sum_{S\in\mathbb{B}_{k,l,m}^{(n)}} \mathbb{1}_{\{S\in\Gamma_{n,N}\}})| \geq \frac{1}{2}\mathbb{E}(\sum_{S\in\mathbb{B}_{k,l,m}^{(n)}} \mathbb{1}_{\{S\in\Gamma_{n,N}\}})\right) = O(\frac{1}{\omega})$

$\Rightarrow \mathbb{P}\left(\sum_{S\in\mathbb{B}_{k,l,m}^{(n)}} \mathbb{1}_{\{S\in\Gamma_{n,N}\}} \leq \frac{1}{2}\mathbb{E}(\sum_{S\in\mathbb{B}_{k,l,m}^{(n)}} \mathbb{1}_{\{S\in\Gamma_{n,N}\}})\right) = O(\frac{1}{\omega})$.

As $\omega \to \infty$ we have that $\mathbb{E}(\sum_{S\in\mathbb{B}_{k,l,m}^{(n)}} \mathbb{1}_{\{S\in\Gamma_{n,N}\}}) \to \infty$ and thus a.s. $\Gamma_{n,N}$ contains a subgraph isomorphic to an element in $\mathbb{B}_{k,l,m}$ and the number of these subgraphs a.s. converges to ∞ with order of magnitude ω^m. □

Corollary 2.3.1

The threshold function for the existence of a spanning tree of size k,l with $m = k + l - 1$ edges is $N = O(n^{\frac{k+l-2}{k+l-1}})$.

Corollary 2.3.2

The threshold function for the existence of a connected subgraph of size k,l with $m = k+l \geq 3$ edges is $N = O(n)$. This connected subgraph contains precisely one cycle.

Corollary 2.3.3

The threshold function for the existence of a cycle of length $2k$ over k,k vertices with $m = 2k$ edges is $N = O(n)$, $k \geq 2$.

Corollary 2.3.4

The threshold function for the existence of a complete subgraph of size k,l with $m = k \cdot l$ edges is $p = O(n^{2-\frac{k+l}{k\cdot l}})$.

Lemma 2.3.1 (Erdős and Rényi, 1960)

Let $\epsilon_{n1}, \epsilon_{n2}, ..., \epsilon_{nl}$ be sets of l random variables on some probability space; suppose that ϵ_{ni} takes on only the values 1 or 0. If $\lim_{n\to\infty} \sum_{1\leq i_1<i_2<...<i_r\leq l} \mathbb{E}(\epsilon_{ni_1}, \epsilon_{ni_2}, ..., \epsilon_{ni_r}) = \frac{\lambda^r}{r!}$ uniformly in r for $r = 1, 2, ...$, where $\lambda > 0$ and the summation is extended over all combinations $(i_1, i_2, ..., i_r)$ of order r of the integers $1, 2, ..., l$, then $\lim_{n\to\infty} \mathbb{P}(\sum_{i=1}^{l} \epsilon_{n_i} = j) = \frac{\lambda^j e^{-\lambda}}{j!}$, $(j = 0, 1, ...)$. I.e. the distribution of the sum $\sum_{i=1}^{l} \epsilon_{n_i}$ tends for $n \to \infty$ to the Poisson-distribution with mean value λ.

Theorem 2.3.2

For $\tau_{k,l}$, the number of isolated spanning trees of size k,l in $\Gamma_{n,N}$, and $\lim_{n\to\infty} \frac{N(n)}{n^{-\frac{k+l}{k+l-1}}} = \rho > 0$ it holds that: $\lim_{n\to\infty} \mathbb{P}_{n,N}(\tau_{k,l} = j) = \frac{\lambda^j e^{-\lambda}}{j!}$, $(j = 0, 1, ...)$, with $\lambda = \frac{\rho^{k+l-1} k^{l-1} l^{k-1}}{k!l!}$. Moreover, the maximum number of spanning trees of size k,l of $n\frac{k^{l-1}l^{k-1}}{k!l!}(\frac{k+l-1}{k+l})^{k+l-1} e^{-(k+l-1)}$ is attained for $N \sim n^{\frac{k+l-1}{k+l}}$.

Proof

We denote the set of all spanning trees of size k,l that are subgraphs of $\Gamma_{n,N}$ by $T_{k,l}^{(n)}$. The variable $\epsilon(S)$ takes on the value 1 if it is an isolated subgraph and 0 otherwise. We prove the theorem by applying Lemma 2.3.1 to $\sum_{S\in T_{k,l}^{(n)}} \epsilon(S)$, which requires us to show that all its conditions are fulfilled.

We find $\mathbb{E}(\epsilon(S)) = \frac{\binom{(n-k)(n-l)}{N-k-l+1}}{\binom{n^2}{N}} \sim (\frac{N}{n^2})^{k+l-1} e^{-(k+l)\frac{N}{n}}$ by induction. For $k = 0$ and $l = 0$, $\frac{\binom{(n-k)(n-l)}{N-k-l+1}}{\binom{n^2}{N}}$ equals $\frac{(n^2)!N!(n^2-N)!}{(N+1)!(n^2-N-1)!(n^2)!} = \frac{n^2-N}{N+1} \sim \frac{n^2}{N}$, and thus the equality holds for $k = 0$ and $l = 0$. This is the basis of the induction. Dividing $\frac{\binom{(n-k)(n-l)}{N-k-l+1}}{\binom{n^2}{N}}$ by its limit $(\frac{N}{n^2})^{k+l-1} e^{-(k+l)\frac{N}{n}}$ gives $\frac{\frac{\binom{(n-k)(n-l)}{N-k-l+1}}{\binom{n^2}{N}}}{(\frac{N}{n^2})^{k+l-1} e^{-(k+l)\frac{N}{n}}} = \frac{(n^2-(k+l)n+kl)!N!(n^2-N)!}{(N-k-l+1)!(n^2-(k+l)n+kl-N+k+l-1)!(n^2)!}(\frac{n^2}{N})^{k+l-1} e^{(k+l)\frac{N}{n}}$. We disregard the negligible order term. Now we use this to construct the following step of induction, by dividing this term by the subsequent term with $k + 1$ and l (k and $l + 1$ works symmetrical):

$$\frac{\frac{(n^2-(k+l)n+kl)!N!(n^2-N)!}{(N-k-l+1)!(n^2-(k+l)n+kl-N+k+l-1)!(n^2)!}(\frac{n^2}{N})^{k+l-1} e^{(k+l)\frac{N}{n}}}{\frac{(n^2-((k+1)+l)n+(k+1)l)!N!(n^2-N)!}{(N-(k+1)-l+1)!(n^2-((k+1)+l)n+(k+1)l-N+(k+1)+l-1)!(n^2)!}(\frac{n^2}{N})^{(k+1)+l-1} e^{((k+1)+l)\frac{N}{n}}}$$

$$= \frac{\frac{(n^2-(k+l)n+kl)!}{(N-k-l+1)!(n^2-(k+l)n+kl-N+k+l-1)!}(\frac{n^2}{N})^{k+l-1} e^{(k+l)\frac{N}{n}}}{\frac{(n^2-((k+1)+l)n+(k+1)l)!}{(N-(k+1)-l+1)!(n^2-((k+1)+l)n+(k+1)l-N+(k+1)+l-1)!}(\frac{n^2}{N})^{(k+1)+l-1} e^{((k+1)+l)\frac{N}{n}}}$$

$$\sim \frac{(n^2-(k+l)n+kl)^{N-k-l+1}}{(n^2-(k+l+1)n+(k+1)l)^{N-k-l}} \frac{1}{n^2} e^{-\frac{N}{n}} \sim n^2 \frac{1}{n^2} e^{-\frac{N}{n}} \sim e^{-\frac{N}{n}} \to 1.$$

And thus the limit is shown for $k+1$ and l which concludes the induction. Hence the equation is proved for $\lim_{n\to\infty} \frac{N(n)}{n^{-\frac{k+l}{k+l-1}}} = \rho > 0$.

Moreover, for disjoint $S_1, ...S_r \in T_{k,l}^{(n)},\ k,l,r \geq 1$, it holds that $\mathbb{E}\big(\epsilon(S_1), ..., \epsilon(S_r)\big) = \frac{\binom{(n-rk)(n-rl)}{N-r(k+l-1)}}{\binom{n^2}{N}} \sim (\frac{N}{n^2})^{r(k+l-1)} \mathrm{e}^{-(k+l)r\frac{N}{n}}$ when all S_i are disjoint and zero otherwise.

An extended version of Cayley's formula states that from k and l labelled points, $k^{l-1}l^{k-1}$ different spanning trees can be formed. Hence summing over all possible r-tuples of spanning trees in $T_{k,l}^{(n)}$ gives $\sum \mathbb{E}\big(\epsilon(S_1), ..., \epsilon(S_r)\big) \sim k^{l-1}l^{k-1}\frac{\binom{n}{k}^r\binom{n}{l}^r}{r!}(\frac{N}{n^2})^{r(k+l-1)}\mathrm{e}^{-(k+l)r\frac{N}{n}}$
$\sim (\frac{k^{l-1}l^{k-1}}{k!l!})^r \frac{n^{(k+l)r}}{r!}(\frac{N}{n^2})^{r(k+l-1)}\mathrm{e}^{-(k+l)r\frac{N}{n}}$.

For $\lim_{n\to\infty} \frac{N(n)}{n^{-\frac{k+l}{k+l-1}}} = \rho > 0$ we can conclude that $\lim_{n\to\infty} \sum \mathbb{E}\big(\epsilon(S_1), ..., \epsilon(S_r)\big) = \frac{\lambda^r}{r!}$, $r = 1, 2, ...$ with λ defined as before. Hence we showed that Lemma 2.3.1 can be applied to $\tau_{k,l} = \sum_{S\in T_{k,l}^{(n)}} \epsilon(S)$.

Rewriting the above formula gives $\mathbb{E}(\tau_{k,l}) = \frac{n^2}{N} \cdot \frac{(\frac{N}{n}\mathrm{e}^{-\frac{N}{n}})^{k+l}k^{l-1}l^{k-1}}{k!l!} = n \cdot m_{k,l}(\frac{N}{n})$ with $m_{k,l}(t) = \frac{k^{l-1}l^{k-1}t^{k+l-1}\mathrm{e}^{-(k+l)t}}{k!l!}$. For k,l fixed we solve $\frac{\partial}{\partial t}m_{k,l}(t)$
$= \frac{k^{l-1}l^{k-1}}{k!l!}\mathrm{e}^{-(k+l)t}t^{k+l-2}\ (k+l-1-(k+l)t) = 0$ and hence the maximum is attained at
$t = \frac{k+l-1}{k+l}$, or $N \sim n\frac{k+l-1}{k+l}$. This maximum equals $n\frac{k^{l-1}l^{k-1}}{k!l!}(\frac{k+l-1}{k+l})^{k+l-1}\mathrm{e}^{-(k+l-1)}$. □

Theorem 2.3.3

Let $\gamma_{k,k}$ be the number of cycles of size k,k as a subgraph of $\Gamma_{n,N}$. For $N(n) \sim cn,\ c > 0$ we find that $\lim_{n\to\infty} \mathbb{P}(\gamma_{k,k} = j) = \frac{\lambda^j \mathrm{e}^{-\lambda}}{j!},\ \lambda = \frac{1}{2k}(\frac{N}{n})^{2k}$.

Proof

There are $\frac{1}{2}k!(k-1)!$ possible cycles of size k,k. Thus $\mathbb{E}(\gamma_{k,k}) = \binom{n}{k}\binom{n}{k}\frac{1}{2}k!(k-1)!\frac{\binom{n^2-2k}{N-2k}}{\binom{n^2}{N}}$
$\sim \frac{1}{2} \cdot \frac{n!n!k!(k-1)!}{k!k!(n-k)!(n-k)!} \cdot \frac{(n^2-2k)^{N-2k}}{(n^2)^N} \cdot \frac{N!}{(N-2k)!} \sim \frac{1}{2k}n^k n^k \frac{(N)^{2k}}{(n^2)^{2k}} \sim \frac{1}{2k} \cdot (\frac{N}{n})^{2k}$.

For $C_{k,k}^{(n)}$ all cycles of size k,k, let $\epsilon(S)$, $S \in C_{k,k}^{(n)}$, be equal to 1 if S is a subgraph of $\Gamma_{n,N}$ and 0 otherwise. Similar to above we find that $\mathbb{E}\left(\epsilon(S)\right) = \frac{\binom{n^2-2k}{N-2k}}{\binom{n^2}{N}} \sim (\frac{N}{n^2})^{2k}$ and $\mathbb{E}\left(\epsilon(S_1), ..., \epsilon(S_r)\right) = \frac{\binom{n^2-2kr}{N-2kr}}{\binom{n^2}{N}} \sim \frac{(n^2-2kr)^{N-2kr}}{(n^2)^N} \cdot \frac{N!}{(N-2kr)!} \sim (\frac{N}{n^2})^{2kr}$. Since there are $\frac{1}{2}k!(k-1)!$ of these cycles, if we sum over all possible r-tuples of these cycles we have: $\sum \mathbb{E}\big(\epsilon(S_1), ..., \epsilon(S_r)\big)$ $\sim \frac{\left(\binom{n}{k}\binom{n}{k}\frac{1}{2}k!(k-1)!\right)^r}{r!}(\frac{N}{n^2})^{2kr} \sim \frac{\left(\frac{1}{2}\cdot\frac{n!n!k!(k-1)!}{k!k!(n-k)!(n-k)!}\right)^r}{r!}\cdot(\frac{N}{n^2})^{2kr} \sim \frac{(\frac{1}{2k}n^k n^k)^r}{r!}\cdot(\frac{N}{n^2})^{2kr} = \frac{\left(\frac{1}{2k}(\frac{N}{n})^{2k}\right)^r}{r!} = \frac{\lambda^r}{r!}$ with $\lambda = \frac{1}{2k}(\frac{N}{n})^{2k}$. Since $\gamma_{k,k} = \sum \epsilon(S)$ we can apply Lemma 2.3.1 and hence for $N(n) \sim cn$ the number of cycles of size k,k follows a Poisson distribution with $\lambda = \frac{1}{2k}(\frac{N}{n})^{2k}$. □

Theorem 2.3.4

Let $\gamma_{k,k}^*$ be the number of isolated cycles or size k,k as a subgraph of $\Gamma_{n,N}$. For $N(n) \sim cn$, $c > 0$ we find that $\lim_{n\to\infty} \mathbb{P}(\gamma_{k,k}^* = j) = \frac{\lambda^j \mathrm{e}^{-\lambda}}{j!},\ \lambda = \frac{1}{2k}(c\mathrm{e}^c)^{2k}$.

Proof

There are $\frac{1}{2}k!(k-1)!$ possible cycles of size k,k. $\mathbb{E}(\gamma_{k,k}^*) = \frac{1}{2}k!(k-1)!\binom{n}{k}\binom{n}{k}\frac{\binom{(n-k)^2}{N-2k}}{\binom{n^2}{N}}$ $\sim \frac{1}{2}\cdot\frac{n!n!k!(k-1)!}{k!k!(n-k)!(n-k)!}(\frac{N}{n^2})^{2k}\mathrm{e}^{-2k\frac{N}{n}} \sim \frac{1}{2k}n^k n^k(\frac{N}{n})^{2k}\mathrm{e}^{-2k\frac{N}{n}} \sim \frac{1}{2k}(\frac{N}{n}\mathrm{e}^{-\frac{N}{n}})^{2k}$. We can apply Lemma 2.3.1 again and thus the number of isolated cycles of size k,k follows a Poisson distribution with $\lambda = \frac{1}{2k}(c\mathrm{e}^c)^{2k}$ for $N(n) \sim cn$. □

Theorem 2.3.5

For $N = o(n)$ and $n \to \infty$ the graph $\Gamma_{n,N}$ is a.s. the union of disjoint spanning trees.

Proof

Let T be the property that a graph is the union of disjoint spanning trees and thus $\bar{T}$ the property that the graph contains at least one cycle. Then $\mathbb{P}_{n,N}(\bar{T}) \leq \sum_{k=2}^n \binom{n}{k}\binom{n}{k}\frac{1}{2}k!(k-1)!\frac{\binom{n^2-2k}{N-2k}}{\binom{n^2}{N}}$ $= O\left(\sum_{k=2}^n \frac{n!n!\frac{1}{2}k!(k-1)!}{k!k!(n-k)!(n-k)!}\cdot\frac{(n^2-2k)^{N-2k}}{(n^2)^N}\cdot\frac{N!}{(N-2k)!}\right) = O(\sum_{k=2}^n \frac{n^k n^k}{(n^2)^{2k}}N^{2k}) = O\left(\sum_{k=2}^n (\frac{N}{n})^{2k}\right)$ $= O(\frac{N}{n})$. Thus for $N = o(n)$ we have that $\lim_{n\to\infty} \mathbb{P}_{n,N}(T) = 1$. □

Theorem 2.3.6

Let $V_{n,N}$ be the number of vertices in $\Gamma_{n,N}$ which belong to an isolated spanning tree contained in $\Gamma_{n,N}$. For $N(n) \sim cn$ we have $\lim_{n\to\infty} \frac{\mathbb{E}(V_{n,N})}{2n} = 1$ when $c \leq 1$ and $\lim_{n\to\infty} \frac{\mathbb{E}(V_{n,N})}{2n} = \frac{x(c)}{c}$ when $c > 1$, where $x(c) = \sum_{v=1}^{\infty} \frac{v^{v-1}(\mathrm{ce}^{-c})^v}{v!}$ and $v = k + l$.

Proof

For $\tau_{k,l}$ the number of isolated spanning trees of size k,l as a subgraph of $\Gamma_{k,l}$, then we have: $V_{n,N} = \sum_{k=0}^{n} \sum_{l=0}^{n} (k+l)\tau_{k,l}$ and hence $\mathbb{E}(V_{n,N}) = \sum_{k=0}^{n} \sum_{l=0}^{n} (k+l)\mathbb{E}(\tau_{k,l})$.

Using $\lim_{n\to\infty} \frac{\mathbb{E}(\tau_{k,l})}{2n} = \frac{1}{2c} \cdot \frac{(\mathrm{ce}^{-c})^{k+l} k^{l-1} l^{k-1}}{k!l!}$ we obtain for $c \leq 1$ that $\lim_{n\to\infty} \frac{\mathbb{E}(V_{n,N})}{2n}$
$= \frac{1}{2c} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{(k+l)(\mathrm{ce}^{-c})^{k+l} k^{l-1} l^{k-1}}{k!l!}$.

Remains to show that the latter term equals 1. We make use of $\sum_{k=0}^{v} \binom{v}{k} k^{v-k-1}(v-k)^{k-1}$
$= 2v^{v-2} \Rightarrow \sum_{k=0}^{v} \frac{k^{v-k-1}(v-k)^{k-1}}{k!(v-k)!} = \frac{2v^{v-2}}{v!}$ and $\sum_{v=1}^{\infty} \frac{v^{v-1}(\mathrm{ce}^{-c})^v}{v!} = c$ for $c \leq 1$, see Erdős and Rényi (1960).

Thus $\lim_{n\to\infty} \frac{\mathbb{E}(V_{n,N})}{2n} = \frac{1}{2c} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{(k+l)(\mathrm{ce}^{-c})^{k+l} k^{l-1} l^{k-1}}{k!l!}$
$= \frac{1}{2c} \sum_{v=1}^{\infty} \sum_{k+l=v} \frac{(k+l)(\mathrm{ce}^{-c})^{k+l} k^{l-1} l^{k-1}}{k!l!} = \frac{1}{2c} \sum_{v=1}^{\infty} \sum_{k=0}^{v} \frac{v(\mathrm{ce}^{-c})^v k^{v-k-1}(v-k)^{k-1}}{k!(v-k)!}$
$= \frac{1}{2c} \sum_{v=1}^{\infty} v(ce^{-c})^v \sum_{k=0}^{v} \frac{k^{v-k-1}(v-k)^{k-1}}{k!(v-k)!} = \frac{1}{2c} \sum_{v=1}^{\infty} v(ce^{-c})^v \frac{2v^{v-2}}{v!} = \frac{1}{c} \sum_{v=1}^{\infty} \frac{v^{v-1}(\mathrm{ce}^{-c})^v}{v!}$
$= 1$. So a.e. vertex belongs to a spanning tree for $c \leq 1$.

Similarly to Erdős and Rényi (1960), for $c > 1$ it holds that $\lim_{n\to\infty} \frac{\mathbb{E}(V_{n,N})}{2n}$
$= \frac{1}{c} \sum_{v=1}^{\infty} \frac{v^{v-1}(\mathrm{ce}^{-c})^v}{v!} < 1$. □

Theorem 2.3.7

The total number of cycles in $\Gamma_{n,N}$ is denoted by $C_{n,N}$. For $N(n) \sim cn,\ c < 1$ it holds that $\lim_{n\to\infty} \mathbb{E}(C_{n,N}) = \frac{1}{2}\log(\frac{1}{1-c^2}) - \frac{c^2}{2}$ and for $c = 1$ that $\lim_{n\to\infty} \mathbb{E}(C_{n,N}) \sim \frac{1}{2}\log(n)$.

Proof

For $c < 1$ we find that $\mathbb{E}(\gamma_{k,k}) = \binom{n}{k}\binom{n}{k}\frac{1}{2}k!(k-1)!\frac{\binom{n^2-2k}{N-2k}}{\binom{n^2}{N}} \sim \frac{n!n!\frac{1}{2}k!(k-1)!}{k!k!(n-k)!(n-k)!} \cdot \frac{(n^2-2k)^{N-2k}}{(n^2)^N} \cdot \frac{N!}{(N-2k)!}$ $\sim \frac{1}{2k}(\frac{N}{n})^{2k} \sim \frac{1}{2k}c^{2k}$. Since $C_{n,N} = \sum_{k=2}^{n}\gamma_{k,k}$ and $\sum_{k=1}^{\infty}\frac{z^k}{k} = -\log(1-z)$ for $z = c^2$ we find that $\lim_{n\to\infty}\mathbb{E}(C_{n,N}) = \frac{1}{2}\log(\frac{1}{1-c^2}) - \frac{c^2}{2}$, which proves the first part.

For $c = 1$ it holds that $\mathbb{E}(\gamma_{k,k})$ is similar to $\frac{1}{2k}$ and hence $C_{n,N} = \sum_{k=2}^{n}\gamma_{k,k} = \frac{1}{2}\sum_{k=2}^{n}\frac{1}{k}$. Since $\sum_{k=1}^{n}\frac{1}{k} \sim \log(n+1)$ we can conclude that $\lim_{n\to\infty}\mathbb{E}(C_{n,N}) \sim \frac{1}{2}\log(n)$. □

Theorem 2.3.8

Let C be the property that a bipartite graph contains at least one cycle. When $N(n) \sim cn,\ c \leq 1$, it holds that $\lim_{n\to\infty}\mathbb{P}_{n,N}(C) = 1 - \sqrt{1-c^2}\mathrm{e}^{\frac{c^2}{2}}$. For $c < 1$ the probability of at least one cycle is less than 1, but for $c = 1$ the bipartite graph a.s. contains a cycle.

Proof

Given that the probability that two cycles are not disjoint is negligibly small and that the number of cycles follows a Poisson distribution with mean $\lambda = \lim_{n\to\infty}\mathbb{E}(C_{n,N})$, we find for $c < 1$ that $\lim_{n\to\infty}\mathbb{P}_{n,N}(\bar{C}) = \mathrm{e}^{-\lim_{n\to\infty}\mathbb{E}(C_{n,N})} = \mathrm{e}^{-\left(\frac{1}{2}\log(\frac{1}{1-c^2})-\frac{c^2}{2}\right)} = \sqrt{1-c^2}\mathrm{e}^{\frac{c^2}{2}}$. As this converges to 0 as $c \uparrow 1$, the theorem is proved for $c \leq 1$. □

Theorem 2.3.9

The number of points of $\Gamma_{n,N}$ that are part of a cycle is denoted as $C^*_{n,N}$. For $N(n) \sim cn,\ c < 1$, it holds that $\lim_{n\to\infty}\mathbb{E}(C^*_{n,N}) = \frac{c^4}{2(1-c^2)}$.

Proof

Given that the probability that two cycles are not disjoint is negligibly small we find that $\lim_{n\to\infty}\mathbb{E}(C^*_{n,N}) \sim \lim_{n\to\infty}\sum_{k=2}^{n}2k\gamma_{k,k} = c^4\sum_{k=0}^{\infty}(c^2)^k = c^4\frac{1}{1-c^2} = \frac{c^4}{1-c^2}$, since $\sum_{k=0}^{n}z^k = \frac{1-z^{n+1}}{1-z}$. □

Theorem 2.3.10

We denote the number of components of $\Gamma_{n,N}$ by $\zeta_{n,N}$. For $N(n) \sim cn,\ c < 1$ it holds that $\mathbb{E}(\zeta_{n,N}) = n - N + O(1)$. When $c = 1$ the expected number of components is given by $\mathbb{E}(\zeta_{n,N}) = n - N + O(\log(n))$. For $c > 1$ we find $\lim_{n\to\infty} \frac{\mathbb{E}(\zeta_{n,N})}{2n} = \frac{1}{c}\left(x(c) - \frac{x(c)^2}{2}\right)$, where $x(c) = \sum_{v=1}^{\infty} \frac{v^{v-1}(\mathrm{c}\mathrm{e}^{-c})^v}{v!}$ and $v = k + l$.

Proof

In order to prove the first two parts we will make use of Theorem 2.3.7, which states the number of cycles. A new link can either connect two components or create at least one cycle. Hence a new link either increases $N - \zeta_{n,N}$ by one or increases $C_{n,N}$ by at least one. Thus it holds that $N \leq n - \zeta_{n,N} + C_{n,N}$.

For $c < 1$ the expected number of cycles is a constant and hence $\mathbb{E}(\zeta_{n,N}) = n - N + O(1)$.

Similarly when $c = 1$ the expected number of cycles is given by $\frac{1}{2}\log(n)$. Therefore it holds that $\mathbb{E}(\zeta_{n,N}) = n - N + O(\log(n))$.

Theorem 3.1 implies that the expected number of components of size k,l with $m \geq k + l$ edges is of the order $O(\frac{N^m}{n^{2m-k-l}}) = O\left((\frac{N}{n})^{k+l}\right)$, which is bounded $\forall k$,l. The number of components of size K,L or greater is trivially of order $O(\frac{2n}{K+L})$, where K,$L < n$ can be chosen arbitrarily large. Hence the number of components is similar to the number of spanning trees. As a result it holds that $\mathbb{E}(\zeta_{n,N}) \sim \sum_{k=0}^{n}\sum_{l=0}^{n} \mathbb{E}(\tau_{k,l}) \sim \frac{n^2}{N}\sum_{k=0}^{n}\sum_{l=0}^{n} \frac{k^{l-1}l^{k-1}}{k!l!}(\frac{N}{n}\mathrm{e}^{-\frac{N}{n}})^{k+l}$.

Using $\sum_{k=0}^{v} \frac{k^{v-k-1}(v-k)^{k-1}}{k!(v-k)!} = \frac{2v^{v-2}}{v!}$, it follows that $\lim_{n\to\infty} \frac{\mathbb{E}(\zeta_{n,N})}{2n}$
$= \frac{1}{2c}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \frac{k^{l-1}l^{k-1}}{k!l!}(\mathrm{c}\mathrm{e}^{-c})^{k+l} = \frac{1}{2c}\sum_{v=1}^{\infty}\sum_{k+l=v} \frac{k^{l-1}l^{k-1}}{k!l!}(\mathrm{c}\mathrm{e}^{-c})^{k+l}$
$= \frac{1}{2c}\sum_{v=1}^{\infty}\sum_{k=0}^{v} \frac{k^{v-k-1}(v-k)^{k-1}}{k!(v-k)!}(\mathrm{c}\mathrm{e}^{-c})^{v} = \frac{1}{2c}\sum_{v=1}^{\infty}(\mathrm{c}\mathrm{e}^{-c})^{v}\sum_{k=0}^{v} \frac{k^{v-k-1}(v-k)^{k-1}}{k!(v-k)!}$
$= \frac{1}{2c}\sum_{v=1}^{\infty}(\mathrm{c}\mathrm{e}^{-c})^{v}\frac{2v^{v-2}}{v!} = \frac{1}{c}\sum_{v=1}^{\infty}(\mathrm{c}\mathrm{e}^{-c})^{v}\frac{v^{v-2}}{v!} = \frac{1}{c}\left(x(c) - \frac{x(c)^2}{2}\right)$ as in Erdős and Rényi (1960).

□

Appendix B: Theorems in Section 4

Theorem 2.4.1

a) If $p = O(\frac{1}{n^\alpha})$, $\alpha > 1$ it follows that $\mathbb{E}(\text{eff}) \to 0$.

b) If $p = \frac{c}{n}$ it follows that $\mathbb{E}(\text{eff}) \to d \in [1 - \frac{1-\mathrm{e}^{-c}}{c}, 1 - \mathrm{e}^{-c}]$.

c) If $\frac{1}{p} = O(n^\alpha)$, $\alpha < 1$ it follows that $\mathbb{P}(\text{eff} = 1) \to 1$ and therefore $\mathbb{E}(\text{eff}) \to 1$.

Proof

a) We can bound the probability that buyer i trades by the probability that he has at least one link: $\mathbb{P}(\text{buyer}_i \text{ trades}) \leq 1 - (1-p)^n = 1 - (1 - \frac{1}{n^\alpha})^n = 1 - \left((1 - \frac{1}{n^\alpha})^{n^\alpha}\right)^{n^{1-\alpha}} \to 1 - \mathrm{e}^{-n^{1-\alpha}}$ $\to 1 - \mathrm{e}^0 = 0$. Since $\mathbb{P}(\text{buyer}_i \text{ trades}) \to 0$ we have that $\mathbb{E}(\text{eff}) \to 0$.

b) We select buyers one by one and if possible let them trade with a linked and available seller. The i^{th} selected buyer has at least $n - i + 1$ available sellers. Hence a lowerbound is found by considering the probability that this buyer is not linked to at least one of the available $n - i + 1$ sellers. It holds that $1 - (1 - \frac{c}{n})^{n-i+1} \leq \mathbb{P}(i^{th}$ selected buyer trades$) \leq 1 - (1 - \frac{c}{n})^n$. Similarly, this leads to $\frac{n+1-\frac{n^\alpha}{c}+\frac{(1-\frac{c}{n^\alpha})^{n+1}}{c}}{n} \leq \mathbb{E}(\text{eff}) \leq \frac{n(1-(1-\frac{c}{n})^n)}{n}$. In the limit it follows that $1 - \frac{1-\mathrm{e}^{-c}}{c} \leq \mathbb{E}(\text{eff}) \leq 1 - \mathrm{e}^{-c}$.

c) Here we make use of the matrix representation. A perfect matching and therefore a maximal number of trades is obtained if there does not exist a block of zeros, where the number of columns plus the number of rows exceeds n. We show that the probability that the maximal number of trades t is lower than n converges to zero.

$\mathbb{P}(t < n) = \mathbb{P}(\exists k, l \text{ block}, k + l > n) \leq \sum_{i=1}^{n} \mathbb{P}(\exists i(n-i+1) \text{ block})$
$\leq \sum_{i=1}^{n} \binom{n}{i}\binom{n}{n-i+1}(1-p)^{i(n-i+1)} \leq 2\sum_{i=1}^{\lfloor\frac{n+1}{2}\rfloor} \binom{n}{i}\binom{n}{n-i+1}(1-p)^{i(n-i+1)}$. As a result of Claim 2.4.1 it holds that $\mathbb{P}(\text{eff} = 1) \to 1$ and thus $\mathbb{E}(\text{eff}) \to 1$. □

Claim 2.4.1

$\lim_{n\to\infty} \sum_{i=1}^{\lfloor\frac{n+1}{2}\rfloor} \binom{n}{i}\binom{n}{n-i+1}(1-p)^{i(n-i+1)} \to 0$ for $p = O(\frac{1}{n^\alpha})$ and $\alpha < 1$.

Proof

Let us denote $\frac{\mathbb{P}(\exists(i+1)(n-i)\text{ block})}{\mathbb{P}(\exists i(n-i+1)\text{ block})}$ as f_i. We use induction to prove the claim. $f_1 = \frac{1}{1} \cdot \frac{1}{2} n(n-1)(1-p)^{n-2} < 1$ for n large enough. Furthermore $\frac{f_{i+1}}{f_i} = \frac{i}{i+2} \frac{n-i+2}{n-i}(1-p)^2 < 1,\ i \leq \lfloor \frac{n-3}{2} \rfloor$. So $f_i < 1,\ i \leq \lfloor \frac{n-1}{2} \rfloor$ and therefore it holds that $\mathbb{P}(\exists 1 \cdot n \text{ block}) \geq \mathbb{P}(\exists i(n-i+1) \text{ block}),\ i \leq \lfloor \frac{n+1}{2} \rfloor$.

Now we can conclude that
$\lim_{n\to\infty} \mathbb{P}(t < n) \leq 2 \sum_{i=1}^{\lfloor \frac{n+1}{2} \rfloor} \binom{n}{i} \binom{n}{n-i+1} (1-p)^{i(n-i+1)} \leq 2 \sum_{i=1}^{\lfloor \frac{n+1}{2} \rfloor} \binom{n}{1} \binom{n}{n} (1-p)^n$
$\leq 2 \lim_{n\to\infty} n \lfloor \frac{n+1}{2} \rfloor (1-p)^n \leq 2 \lim_{n\to\infty} n^2 (1-p)^n \leq 2 \lim_{n\to\infty} n^2 \mathrm{e}^{-n^{1-\alpha}} \to 0$ for $p = O(\frac{1}{n^\alpha})$ and $\alpha < 1$. □

Theorem 2.4.2

Although the behaviour of the graph changes abruptly when c passes through 1, the expected efficiency is continuous in c.

Proof

For any given graph G we add a link e. Obviously we have for t the maximal number of trades that $\mathbb{E}\left(t(G+e)\right) - \mathbb{E}\left(t(G)\right) \leq 1$; a removal of the link e can only lead to a decrease of at maximum one trade. So for any number of added links δ we have that $\mathbb{E}\left(t(G + \sum_{i=1}^{\delta} e_i)\right) - \mathbb{E}\left(t(G)\right) \leq \delta$. The expected efficiency as a function of p is simply the weighted sum over all possible graphs. So now we can find a δ for every $\epsilon > 0$ such that $\mathbb{E}\left(\text{eff}(p+\delta)\right) - \mathbb{E}\left(\text{eff}(p)\right) \leq \epsilon$; namely a δ such that $\lim_{n\to\infty} \frac{\delta}{n} \leq \epsilon$. This proves right continuity and similarly the expected efficiency is also left continuous. □

Theorem 2.4.3

Algorithm 2.4.1 produces the set of possible spanning trees of size k,l.

Proof

The maximal in-degree of the vertices is one and all vertices are connected, thus the algorithm only constructs spanning trees.

Next we will prove by reductio ad absurdum that this algorithm produces all the possible spanning trees. Let us suppose a spanning tree T of size k,l exists that is not produced by the algorithm. The set of points in T that are linked to v_1^1 is non-empty and a subset of V^2 and is thus produced by the algorithm. Let D be the maximal distance from node v_1^1. The sets of vertices with distance $0 < d \leq D$ from node v_1^1 must be non-empty and disjoint. If this set is empty there cannot be a point with distance D. If these sets are not disjoint then there is a point from which there are two possible paths to reach v_1^1, which is not possible in a spanning tree. Trivially since T is a spanning tree, all the vertices that do not have distance $0 < d < D$ must have distance D. Thus T must be produced by the algorithm. □

Theorem 2.4.4

A spanning tree with k buyers and l sellers has $k + l - 1$ edges. Every vertex has at least one edge. The remaining $l - 1$ edges are multinomially distributed over the buyers, and the remaining $k - 1$ edges multinomially over the sellers.

Proof

The number of spanning trees with degrees $i_1, ..., i_k$ respectively $j_1, ..., j_l$ can be computed recursively by building such a spanning tree step by step. We select a vertex or component without unfulfilled outgoing links and calculate the number of possible incoming links. The set V^1 has l outgoing links and V^2 has $k - 1$ outgoing links. A selected vertex or component with a root in V^1 can be linked to all $k - 1$ outgoing links of set V^2. A selected vertex or component with a root in V^2 can be linked to $l - 1$ outgoing links; the outgoing link of the root v_1^1 excluded. There have to be at least two vertices with degree one so we can assume without loss of generality that v_1^1 has degree one. After every step there has to be at least one component remaining without any unfulfilled outgoing links. The outgoing link of root v_1^1 has to be fulfilled last, otherwise the remaining vertices or components cannot be added to the same spanning tree. Including symmetric spanning trees for the moment, this results in $(k - 1)!(l - 1)!$

possible spanning trees.

The number of symmetries in a vertex is given by the factorial of the outdegree. We can conclude that the number of different spanning trees of size k,l and degrees $i_1, ..., i_k$ respectively $j_1, ..., j_l$ is given by $\frac{(k-1)!(l-1)!}{(i_1-1)!\cdot...\cdot(i_k-1)!\cdot(j_1-1)!\cdot...\cdot(j_l-1)!}$ and the probability of having those degrees equals $\frac{(k-1)!(l-1)!}{(i_1-1)!\cdot...\cdot(i_k-1)!\cdot(j_1-1)!\cdot...\cdot(j_l-1)!} \cdot \frac{1}{k^{l-1}l^{k-1}}$.

If we only consider the degrees of set V^1, we can calculate the number of spanning trees with degrees $i_1, ..., i_k$ by $\sum_{j_1+...+j_l=k+l-1} \frac{(k-1)!(l-1)!}{(i_1-1)!\cdot...\cdot(i_k-1)!\cdot(j_1-1)!\cdot...\cdot(j_l-1)!} \cdot \frac{1}{k^{l-1}l^{k-1}}$ $= \frac{(l-1)!}{(i_1-1)!\cdot...\cdot(i_k-1)!} \cdot \frac{1}{k^{l-1}} \sum_{j_1+...+j_l=k+l-1} \frac{(k-1)!}{(j_1-1)!\cdot...\cdot(j_l-1)!} \cdot \frac{1}{l^{k-1}} = \frac{(l-1)!}{(i_1-1)!\cdot...\cdot(i_k-1)!} \cdot \frac{1}{k^{l-1}}$ This shows that the remaining edges per set of vertices follow a multinomial distribution, independently of each other. □

Theorem 2.4.5

A lower- and upper bound for the maximum number of trades in a spanning tree with $k+l > 2$ is determined by the number of vertices with a degree of at least 2, $\#V_{\text{degree}>1}$: $LB = \lceil \frac{\#V_{\text{degree}>1}+1}{2} \rceil$ and $UB = \min(k, l, \#V_{\text{degree}>1})$.

Proof

To show the lower bound, we consider the sidebranches until a single path remains. Sidebranches start in vertices with degree larger than two and consist of an isolated path. We consider sidebranches with the root excluded that have vertices with a degree that is maximal two. Vertices in sidebranches are considered only once, and hence a sequence of sidebranches with vertices of a degree maximal two exists. In such a sidebranch the maximal number of trades is equal to $t = \lfloor \frac{\#V^b_{\text{degree}>1}+1}{2} \rfloor \geq \frac{\#V^b_{\text{degree}>1}}{2}$, where V^b denotes the vertices in this branch. By removing such branches one by one only a single path remains. Trivially, the number of possible trades in such a remaining single path is given by $\lceil \frac{\#V^p_{\text{degree}>1}+1}{2} \rceil$, where V^p are the vertices in this path. Hence the total number of possible trades is bounded from below by $\frac{\#V^b_{\text{degree}>1}}{2} + \lceil \frac{\#V^p_{\text{degree}>1}+1}{2} \rceil \geq \lceil \frac{\#V^p_{\text{degree}>1}+1}{2} \rceil$.

Trivially k and l are both an upper bound for the maximum number of trades. Since there are no vertices with degree one linked to each other in a spanning tree with $k + l > 2$, every trade includes a vertex with a degree of at least two. Hence $UB = \min(k, l, \#V_{\text{degree}>1})$ is an upper bound for the maximum number of trades. □

Theorem 2.4.6

For $p \sim \frac{c}{n} \iff N \sim cn,\ c \leq 1$ it holds that:
$\sum_{1\leq i\leq k\leq\infty} \sum_{1\leq j\leq l\leq\infty} \frac{\mathrm{e}^{-(k+l)}}{i!j!} S(l-1, k-i) S(k-1, l-j) \lceil \frac{k-i+l-j+1}{2} \rceil \leq \mathbb{E}(\text{eff})$
$\leq \sum_{1\leq i\leq k\leq\infty} \sum_{1\leq j\leq l\leq\infty} \frac{\mathrm{e}^{-(k+l)}}{i!j!} S(l-1, k-i) S(k-1, l-j) \min(k, l, k-i+l-j)$.

Proof

In a spanning tree every vertex has at least one edge and the remaining edges are multinomially distributed over the vertices. The probability that i out of k vertices do not have any of the remaining $l-1$ edges is given by $\frac{k!}{i!} \frac{S(l-1,k-i)}{k^{l-1}}$, where $S(l-1, k-i)$ is the Stirling number of the second kind. Hence the probability of a spanning tree with k,l vertices where $k-i$ and $l-j$ vertices have a degree larger than 1 equals $\frac{k^{l-1} l^{k-1}}{k!l!} \mathrm{e}^{-(k+l)} \frac{k!}{i!} \frac{S(l-1,k-i)}{k^{l-1}} \frac{l!}{j!} \frac{S(k-1,l-j)}{l^{k-1}}$
$= \frac{\mathrm{e}^{-(k+l)}}{i!j!} S(l-1, k-i) S(k-1, l-j)$.

The result is obtained by multiplying by the bounds on the expected maximal efficiency per spanning tree and summing over all sizes of the spanning tree and the number of vertices with a degree equal to 1: $\sum_{1\leq i\leq k\leq\infty} \sum_{1\leq j\leq l\leq\infty} \frac{\mathrm{e}^{-(k+l)}}{i!j!} S(l-1, k-i) S(k-1, l-j) \lceil \frac{k-i+l-j+1}{2} \rceil$
$\leq \mathbb{E}(\text{eff}) \leq \sum_{1\leq i\leq k\leq\infty} \sum_{1\leq j\leq l\leq\infty} \frac{\mathrm{e}^{-(k+l)}}{i!j!} S(l-1, k-i) S(k-1, l-j) \min(k, l, k-i+l-j)$.

□

Chapter 3

Information and Efficiency in Thin Markets over Random Networks

3.1 Introduction

In this chapter we consider a market in which transactions only occur between linked traders. These links occur as in a bipartite random network where every link is realised with the same probability, independently of each other. Regular random graphs have been introduced by Erdős and Rényi (1960, 1961). The spot foreign exchange market is studied by Gould et al. (2013a) and is an example of a market in which trade occurs through Bilateral Trading Agreements. Traders provide a block list containing trading partners with whom they prefer not to trade, to protect themselves against adverse selection and to control counterparty risk. In such a market a transaction between two traders only takes place if both are not part of the other's block list. We use the model of Gould et al. and additionally assume that links are realised with the same probability and independently of each other.

Markets over networks have been studied in various settings. Corominas-Bosch (2004) and Chatterjee and Dutta (1998) consider a market in which side by side traders submit an offer which the other traders accept or reject. In Corominas-Bosch (2004) all buyers have the same valuation and sellers the same cost; this allows the network to be split into different subgraphs.

In every subgraph the short side extracts all the possible surplus. We show that under partial information about the network structure, or under incomplete information about valuations and costs, not all the surplus is necessarily extracted. Spulber (2006) and Kranton and Minehart (2001) study simultaneously ascending-bid auctions in which sellers jointly raise their ask until supply equals demand, and then trade occurs. Easley and Kleinberg (2010) and Blume et al. (2009) introduce intermediaries who act strategically and profit from trade. The power of a trader in a network is formalised in Calvó-Armengol (2001) by considering the number of linked traders and their links. A higher market power is achieved when a trader is linked to more traders and when linked traders have fewer links themselves.

For bilateral trading Myerson and Satterthwaite (1983) and Chatterjee and Samuelson (1983) study Nash equilibrium strategies that monotonely transform valuations and costs into offers and exhibit an equilibrium in which they are piecewise linear. We restrict attention to linear markup and markdown strategies where the intensity of the markup or markdown depends on the information set that is available to the trader. These strategies have been introduced by Zhan and Friedman (2007); Cervone et al. (2009) discuss a version that is symmetric between buyers and sellers.

We consider thin markets with few traders, who trade only over existing links in a bipartite graph. These links are formed independently with the same probability p in $(0, 1)$, forming a bipartite random graph à la Erdős-Rényi. Traders behave strategically, and we derive equilibrium configurations depending on the information about the network structure that is available to traders.

Three nested information sets about the realisation of the network are compared. Under no information, traders place orders without knowing which links materialise, but simply the probability p that each link may exist. With partial information, traders know their own links and the probability p that links may exist between other market participants. Under full information, the entire structure of the network is common knowledge.

We study the effect of the quantity of information available to traders on allocative efficiency. We show that this effect is non-monotonic. Furthermore, switching from complete to incomplete information about traders' valuations flips the shape of this non-monotonicity. Under complete information about traders' valuations, we show that for any value of p both no information and full information lead to full allocative efficiency, while the partial information regime is weakly dominated. However, under a more realistic assumption of incomplete information about traders' valuations, this ranking is reversed. If traders use linear markup strategies, partial information strongly dominates full and no information for any value of p.

The organisation of this chapter is as follows. The model and the trading mechanism are described in Section 3.2, together with the markup and markdown strategies and the information sets. Efficiency under complete information about traders' valuations is studied in Section 3.3, followed by incomplete information in Section 3.4. Finally, Section 3.5 concludes.

3.2 The model

Let us consider a market over a bipartite Erdős-Rényi network. In such a network every buyer b_i and every seller s_j are connected with probability p in $(0, 1)$ independently of other links. Trade is possible only if a link exists. An example of such a market is the spot exchange market studied in Gould et al. (2013a). In comparison with this market we add the assumption that every pair of traders is linked with the same probability and independently of other links. Furthermore, in the spot exchange market trade is only possible when both traders do not include the other in their blocklist. E.g. this may occur when the trading partner does not exceed some risk requirement and therefore the network structure is considered exogenous. Nevertheless, a bijective transformation exists from the probability of a link in the spot exchange market to the probability of a link in this chapter. The probability of a link in the spot exchange market is the square of the latter probability.

A buyer desires to obtain one unit of a good and a seller seeks to sell one unit. Under complete information valuations equal one and costs equal zero, whereas under incomplete information the valuations v_i of buyers and costs c_j of sellers are uniformly distributed on the interval $[0, 1]$. This distribution is public information but the realisations are private information. The profit of a buyer is equal to his valuation minus the transaction price if he trades and zero otherwise. The profit of a seller equals the transaction price minus his cost after a trade and zero otherwise.

The probability of a link influences the expected allocative efficiency as absence of links makes some trades impossible. Furthermore, expected efficiency is reduced by strategic behaviour of traders, that could prevent feasible trades. Expected allocative efficiency is defined as the expected total realised surplus from trade divided by the expected maximal total surplus, i.e. the expected total profit of all traders divided by the expected total maximal profit.

We show our results for a market with two buyers and two sellers. In this market the maximal expected surplus equals 2 under complete information and $\frac{2}{5}$ under incomplete information of valuations and costs. The maximal expected surplus under incomplete information is derived for a full network. From the point of view of buyer b_1, he has the highest valuation with probability v_1 since the valuation of the other buyer is uniformly distributed. This results in a trade with the seller with the lowest cost if this trade is feasible. The density function of the lowest cost is given by $2 - 2c_{\text{min}}$ and this trade results in a surplus of $v_1 - c_{\text{min}}$. Similarly buyer b_1 has the lowest valuation with probability $1 - v_1$ and he trades with the seller with the highest cost, with density function $2c_{\text{max}}$, if this trade is feasible. Hence the maximal expected surplus of $\frac{2}{5}$ for a full network is obtained from

$$2\left[\int_0^1 \int_{c_{\text{min}}}^1 (v_1 - c_{\text{min}})(2 - 2c_{\text{min}})v_1 \mathrm{d}v_1 \mathrm{d}c_{\text{min}} + \int_0^1 \int_{c_{\text{max}}}^1 (v_1 - c_{\text{max}})2c_{\text{max}}(1 - v_1)\mathrm{d}v_1 \mathrm{d}c_{\text{max}}\right].$$

However, due to absence of links the maximal expected surplus is reduced, depending on the value of p. For complete and incomplete information about valuations and costs, the ratios between maximal expected surplus given the random network structure and maximal expected surplus of the full network are given by

$$\mathbb{E}(\mathrm{AE}_{\mathrm{p}}^{\mathrm{C}}) = \frac{1{\cdot}4p(1-p)^3+1{\cdot}4p^2(1-p)^2+2{\cdot}2p^2(1-p)^2+2{\cdot}4p^3(1-p)+2{\cdot}p^4}{2},$$

$$\mathbb{E}(\mathrm{AE}_{\mathrm{p}}^{\mathrm{I}}) = \frac{\frac{1}{6}{\cdot}4p(1-p)^3+\frac{1}{4}{\cdot}4p^2(1-p)^2+\frac{1}{3}{\cdot}2p^2(1-p)^2+\frac{43}{120}{\cdot}4p^3(1-p)+\frac{2}{5}{\cdot}p^4}{\frac{2}{5}}.$$

In Fig. 3.1 it is shown that a difference in reduction of efficiency, due to restrictions of the network structure, exists between complete and incomplete information. This is due to the difference in distribution of valuations and costs. Hence, under complete information every trade results in the same surplus while under incomplete information extra links not only increase the expected number of links, but also decrease the expected surplus per trade. Hence the difference between both ratios of efficiency increases.

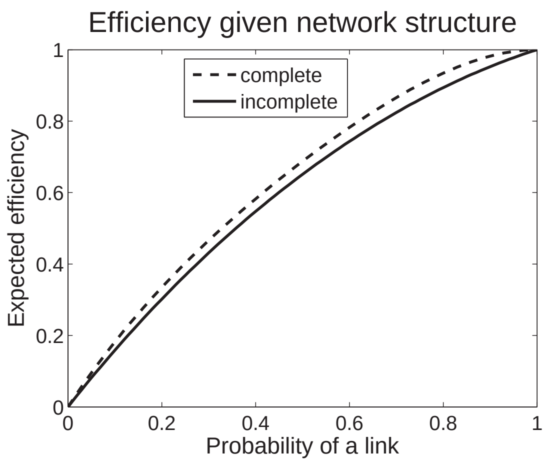


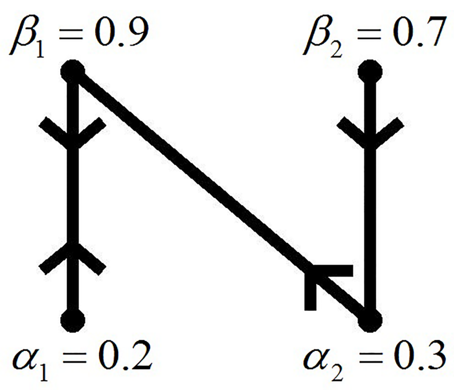
Figure 3.1: Ratios of efficiency given the probability of a link for complete and incomplete information about valuations and costs.

3.2.1 Trading mechanism

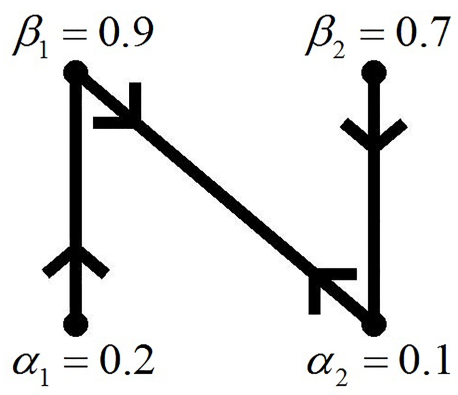
The symmetric trading mechanism consists of simultaneous submission of bids and asks by all traders after which the offers are made public. A buyer ranks his connected sellers by their asks, and a seller his linked buyers by their bids. Trades respect such preferences: preferred buyer-seller pairs are matched with each other. As long as further trades are possible, such a preferred pair naturally exists. Every seller desires to trade with the buyer with the highest bid which ensures that this buyer can trade with his preferred connection. The trade is executed at a price that is equal to the average of bid and ask and this is repeated until no further trades are possible. In contrast to some related literature, this trading mechanism gives equal power to both sides of the market.

If this trading mechanism does not lead to a unique outcome, as a result of traders that do not have a unique preferred trading partner, the trading mechanism selects the outcome that maximises total surplus. Under complete information about valuations and costs this is conservative towards our result, under incomplete information this occurs only in a nullset.

A possible realisation of the bipartite Erdős-Rényi network with bids β_i and asks α_j is given in Fig. 3.2. In the first example buyer b_1 and seller s_1 trade after which b_2 and s_2 trade; this coincides with the social optimum. In the second example however b_1 and s_2 trade. Hence the most profitable trade occurs first and therefore a social optimum is not necessarily reached.



(a) The social optimum is reached.



(b) The social optimum is not reached.

Figure 3.2: Example of the trading mechanism.

3.2.2 Markup and markdown strategies

There is an incentive for traders to act strategically and bid below their valuation and ask above their cost to obtain a higher profit. Under complete information about valuations and costs traders choose a unique strategy given the available information about the network. Under incomplete information traders choose a strategy that is depending on the realisation of their valuation or cost. We assume that traders use linear markup and markdown strategies symmetric on $[0, 1]$ from Cervone et al. (2009). These strategies transform the valuations and costs as follows:

A buyer with valuation v_i bids $\beta_i = v_i(1 - m_i^d)$.
A seller with cost c_j asks $\alpha_j = c_j + m_j^u(1 - c_j)$.

The values m_i^d and m_j^u denote the intensity of the markdown of buyer i and the markup of seller j. The higher these values, the further away bids and asks are from the valuations and costs. The Nash equilibrium markdown and markup strategies are determined on the basis of the distribution of valuation and cost of others, not on the realisation of it. Moreover, we need to take into account the information set of a trader. Hence the markdown and markup strategies will not be a simple transformation of the valuation or cost, but will also depend on the information that is available to traders about the network structure.

3.2.3 The information sets

We study the Nash equilibrium markdown and markup strategies depending on the information set available to traders. Under complete information the valuations and costs are known; under incomplete information only their distribution. Moreover, the number of traders on both sides of the market is known. We consider the following nested sets of information about the network structure, which are all common knowledge:

- No information: The probability of a link is known.
- Partial information: The probability of a link is known as well as the realisation of the own links.
- Full information: The realisation of the entire network is known.

Under no information only the minimal amount of information is available to traders. The probabilities of all networks can be calculated and hence the equilibrium strategy depends only on the probability of a link. Partial information allows a trader to base the strategy on the number of own links and hence the equilibrium strategy depends on the number of a player's own links and the probability that other links are realised. With full information the entire network is known and the equilibrium strategy is based on the realisation of all links.

We show the partitions of the possible networks of the different information sets, for two buyers and two sellers, in Figs. 3.3-3.5. Networks that are not distinguishable are shown in the same partition.

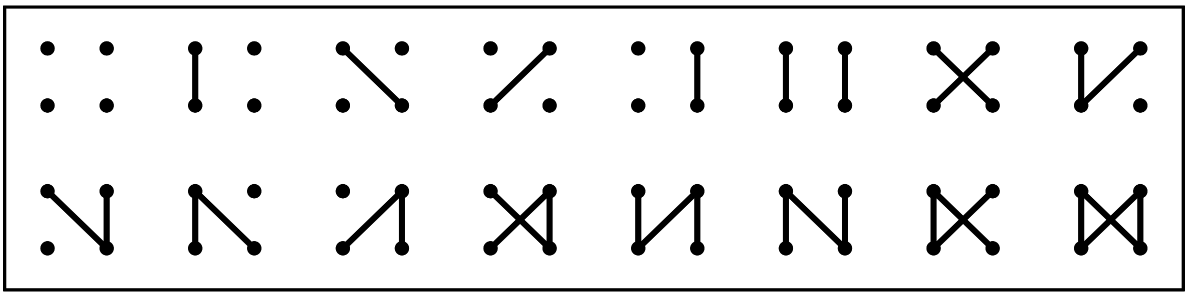


Figure 3.3: Partition under no information.

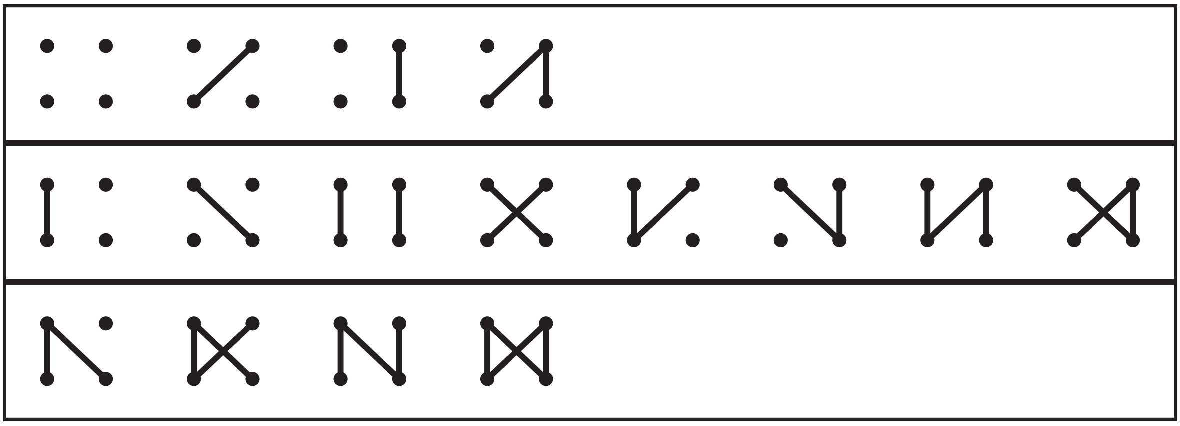


Figure 3.4: Partitions with partial information from the perspective of the top left node.

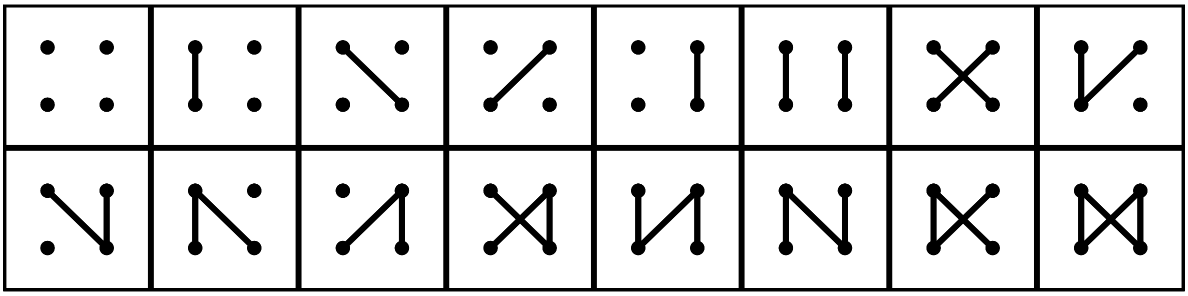


Figure 3.5: Partitions under full information from the perspective of the top left node.

3.3 Complete information about valuations and costs

To compare the expected efficiency given different information sets we consider a market with two buyers and two sellers. Under complete information we assume that valuations and costs are equal to one respectively zero, which is common knowledge. For each information set we calculate the symmetric Nash equilibrium strategies from the profit functions given in Appendix 1. These profit functions are a multiplication of the profit of trade and an indicator function that attains the value one if the trade is feasible.

No Information

Under no information traders have no knowledge about the realisation of links; but they know the probability that they occur. All the traders have the same information and thus they use the same deterministic strategy. Naturally, bids can be decreased and asks increased until they are equal. Hence in the unique symmetric Nash equilibrium buyers bid $\frac{1}{2}$ and sellers ask $\frac{1}{2}$. Given the limitations of p full efficiency is reached; i.e. strategic behaviour does not reduce efficiency.

Partial Information

For computational reasons we restrict offers of traders to the grid $[0, \frac{1}{2k}, ..., 1]$. Traders with one link may prefer to be less aggressive to outcompete other traders. This requires them to increase their bid or decrease their ask by $\frac{1}{2k}$. With a rougher grid this becomes less attractive. For a rough grid with $k < 5$ buyers bid and sellers ask $\frac{1}{2}$ in equilibrium and hence full efficiency is reached.

Below we show the equilibrium strategies as a function of p for $k = 5$, the roughest grid that does not always lead to full efficiency. First given is the markup of a trader with one link, second the markup of a trader with two links. For mixed strategies the probabilities are given by ρ_i. In the range $\frac{1}{\sqrt{11}} < p < \frac{1}{3}$ the latter equilibrium is unstable with respect to the strategy of traders with two links. If one trader deviates to the stable equilibrium it is optimal for other traders to deviate also, because this allows for trades between agents with two links.

$0 < p < \frac{5-\sqrt{5}}{10}$: $[\frac{1}{2}], [\frac{1}{2}]$.

$\frac{5-\sqrt{5}}{10} < p < \frac{\sqrt{17}-3}{4}$: $[\rho_1 \frac{2}{5}, (1-\rho_1)\frac{3}{5}], [\frac{3}{5}]$, where $\rho_1 = \frac{-4}{2p^2+3p-5}$.

$\frac{\sqrt{17}-3}{4} < p < \frac{1}{\sqrt{11}}$: $[\frac{2}{5}], [\frac{3}{5}]$.

$\frac{1}{\sqrt{11}} < p < \frac{1}{3}$: $[\frac{2}{5}], [\frac{1}{2}]$ stable and $[\frac{2}{5}], [\frac{3}{5}]$ unstable.

$\frac{1}{3} < p < \frac{5+\sqrt{5}}{10}$: $[\rho_1 \frac{3}{10}, \rho_2 \frac{2}{5}, (1-\rho_1-\rho_2)\frac{1}{2}], [\frac{1}{2}]$,
where $\rho_1 = -\frac{-8+45p-45p^2}{41(p-1)p}$ and $\rho_2 = -2\frac{9-25p+25p^2}{41(p-1)p}$.

$\frac{5-\sqrt{5}}{10} < p < 1$: $[\frac{1}{2}], [\frac{1}{2}]$.

For $k = 5$ we find that full efficiency is not attained for $\frac{5-\sqrt{5}}{10} < p < \frac{1}{\sqrt{11}}$. When the grid is sufficiently dense full efficiency is not reached for every value of p. For denser grids this area increases and hence the result may also hold without the assumption of a grid of strategies. The subset of p for which traders with one link become less aggressive increases. As k goes to infinity these traders use mixed strategies over an infinite number of strategies. When the probability p is sufficiently small, traders with two links will become more aggressive when traders with one link are less aggressive. This does not hold when the probability p is relatively large, because this will cause a profit of zero in a fully connected network. For a subset of p, strategic behaviour reduces efficiency.

Full Information

Under full information traders have full knowledge about the realisation of the network. When both sides of the market have the same size, bids and asks equal one half in the symmetric equilibrium. For traders with one link it is not profitable to be less aggressive. When one side of the market is thinner, it extracts all the possible surplus. Agents with one link are less aggressive to outcompete the other trader on the same side of the market. Given the limitations of p full efficiency is reached, i.e. efficiency is not reduced by strategic behaviour.

No and full information lead to full efficiency, given the limitations of the network structure. For a grid of possible strategies we show that under partial information the strategic behaviour of traders decreases efficiency for a subset of p. Moreover, we argue that with a denser grid efficiency is decreased for a larger range of values for p. Under no and full information, restricting strategies of traders to a grid has no effect. Hence a non-monotonicity occurs and partial information is weakly dominated. Under complete information about traders' valuations and costs it is optimal when traders either receive all or no information about the network structure.

3.4 Incomplete information about valuations and costs

Under incomplete information, valuations and costs are uniformly distributed on $[0, 1]$, where the distribution is common knowledge but the realisations are private information. To compare the expected efficiency given different information sets we again consider a market with two buyers and two sellers. The necessary calculations for every information set are shown in Appendix 2. As an example, the best response functions under full information are given below for a network where buyer b_1 is connected with both sellers. We solve these to find the Nash equilibrium strategies and calculate expected allocative efficiency, volume and profit.

Example

Network $b_1 \leftrightarrow s_1$ & $b_1 \leftrightarrow s_2$

In equilibrium it holds that $m_1^u = m_2^u = m^u$ and thus buyer b_1 trades with the seller with the lowest cost $c_{\text{min}} = \min(c_1, c_2)$, which has pdf $2 - 2c_{\text{min}}$. We denote the profit of a buyer with bid β_i trading with a seller with ask α_j as $\pi(\beta_i, \alpha_j)$ and similar for sellers. For simplicity we disregard in this notation that the offers are a function of both the strategy and the valuation or cost. The integration limits are set to indicate the region of valuations and costs in which trade occurs:

$$\frac{\partial}{\partial m_1^d} \int_0^{\frac{1-m_1^d-m^u}{1-m^u}} \int_{\frac{c_{\text{min}}+m^u(1-c_{\text{min}})}{1-m_1^d}}^{1} \pi(\beta_1, \alpha_{\text{min}})(2-2c_{\text{min}})\mathrm{d}v_1\mathrm{d}c_{\text{min}} = 0.$$

Seller s_1 only trades when its ask is lower than the ask from seller s_2: $c_1 + m_1^u(1-c_1) < c_2 + m_2^u(1-c_2)$. For a given cost c_1 this happens with probability $\mathbb{P}(\text{trade}) = 1 - \frac{(1-m_1^u)c_1+m_1^u-m_2^u}{1-m_2^u}$:

$$\left[\frac{\partial}{\partial m_1^u} \int_0^{\frac{1-m_1^d-m_1^u}{1-m_1^u}} \int_{\frac{c_1+m_1^u(1-c_1)}{1-m_1^d}}^{1} \pi(\alpha_1, \beta_1)\mathbb{P}(\text{trade})\mathrm{d}v_1\mathrm{d}c_1\right]_{\{m_2^u=m_1^u\}} = 0.$$

Solving these best response functions gives the Nash equilibrium strategies
$m^u = m_1^u = m_2^u \approx 0.110$ and $m_1^d \approx 0.341$.

The expected efficiency given the reductions invoked by absence of links is given by

$$\mathbb{E}(\text{AE}) = \frac{\int_0^{\frac{1-m_1^d-m^u}{1-m^u}} \int_{\frac{c_{\text{min}}+m^u(1-c_{\text{min}})}{1-m_1^d}}^{1} (v_1-c_{\text{min}})(2-2c_{\text{min}})\mathrm{d}v_1\mathrm{d}c_{\text{min}}}{\int_0^1 \int_{c_{\text{min}}}^1 (v_1-c_{\text{min}})(2-2c_{\text{min}})\mathrm{d}v_1\mathrm{d}c_{\text{min}}} \approx 0.858.$$

The ratio between the expected number of trades and the maximal number of trades gives the expected volume:

$$\mathbb{E}(\text{Volume}) = \frac{1}{2}\int_0^{\frac{1-m_1^d-m^u}{1-m^u}} \int_{\frac{c_{\min}+m^u(1-c_{\min})}{1-m_1^d}}^1 1\cdot(2-2c_{\min})\mathrm{d}v_1\mathrm{d}c_{\min} \approx 0.204.$$

Similarly to the best response functions above we calculate the expected profit for a trader having one link, respectively two links:

$$\mathbb{E}(\Pi^1) = \int_0^{\frac{1-m_1^d-m_1^u}{1-m_1^u}} \int_{\frac{c_1+m_1^u(1-c_1)}{1-m_1^d}}^1 \pi(\alpha_1,\beta_1)(1-c_1)\mathrm{d}v_1\mathrm{d}c_1 \approx 0.038,$$

$$\mathbb{E}(\Pi^2) = \int_0^{\frac{1-m_1^d-m^u}{1-m^u}} \int_{\frac{c_{\min}+m^u(1-c_{\min})}{1-m_1^d}}^1 \pi(\beta_1,\alpha_1)(2-2c_{\min})\mathrm{d}v_1\mathrm{d}c_{\min} \approx 0.138.$$

Comparisons

No information outperforms full information in terms of expected efficiency for values of p smaller than the benchmark $c \approx 0.106$, but for large values of p the opposite holds, as shown in Fig. 3.6. As the available information has no effect on the efficiency reduction due to absence of links, we emphasise solely the effect of strategic behaviour. Hence the ratio is shown between the realised efficiency and the maximal efficiency given the network structure. The maximum differences are reached at $p \approx 0.070$ and $p \approx 0.729$.

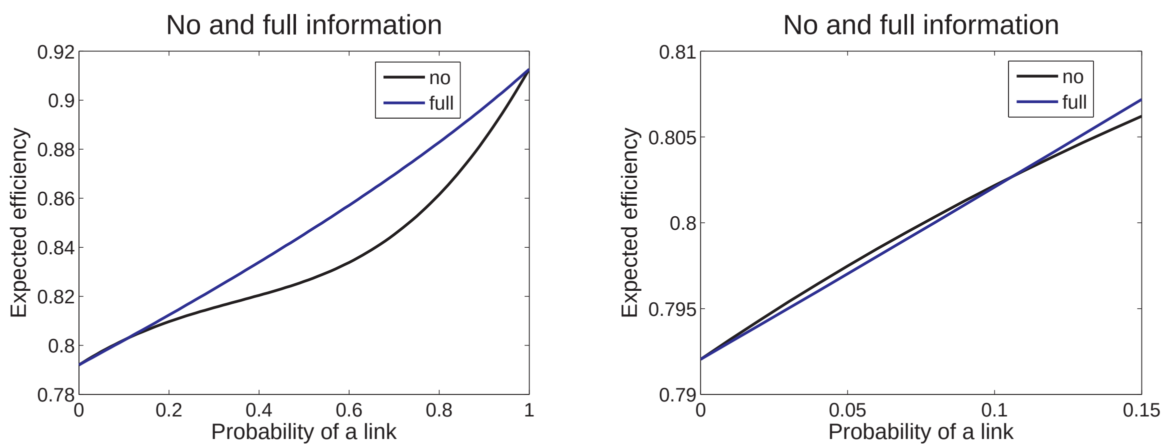


Figure 3.6: Efficiency no and full information.

We find that the amount of information available to traders has a non-monotonic effect on efficiency; irrespective of the probability of a link, partial information leads to the highest expected efficiency. Moreover, we observe that switching from complete to incomplete information reverses the shape of the non-monotonicity. We conclude that in terms of efficiency the following order of information sets holds:

$$0 < p < c : \mathbb{E}(\text{AE}_{\text{partial}}) > \mathbb{E}(\text{AE}_{\text{no}}) > \mathbb{E}(\text{AE}_{\text{full}}),$$
$$c < p < 1 : \mathbb{E}(\text{AE}_{\text{partial}}) > \mathbb{E}(\text{AE}_{\text{full}}) > \mathbb{E}(\text{AE}_{\text{no}}).$$

The maximum difference between information sets is reached near $p = \frac{2}{3}$ where the probability of having one respectively two links is equal. At this point uncertainty about the network structure is the most reduced by additional information. Fig. 3.7 shows the efficiency under strategic behaviour given the restrictions of the network structure, between the different information sets:

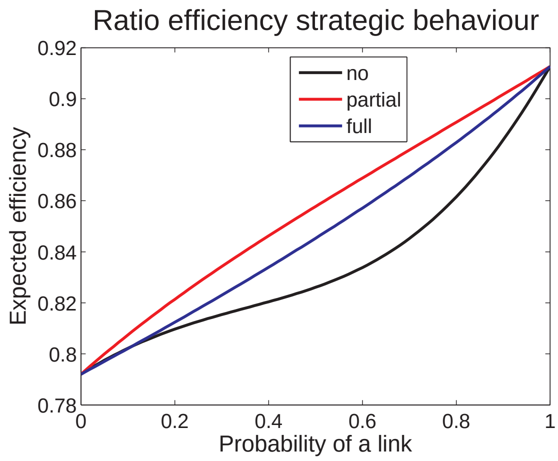
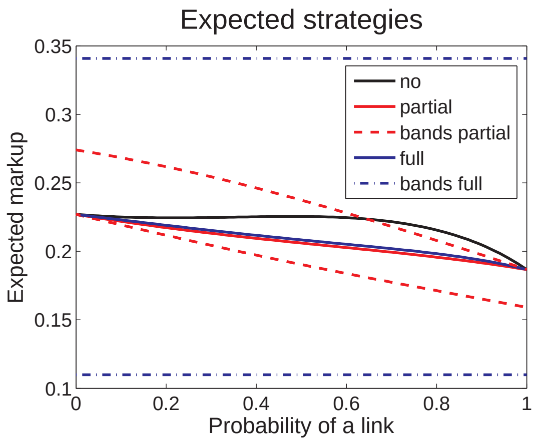
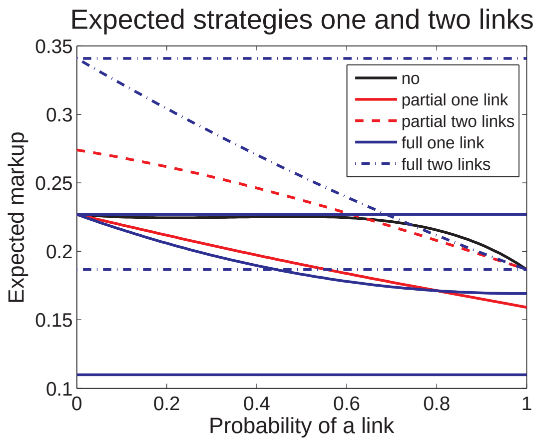


Figure 3.7: Comparison between ratios between efficiency under strategic behaviour and the maximal efficiency given the limitations of the network structure.

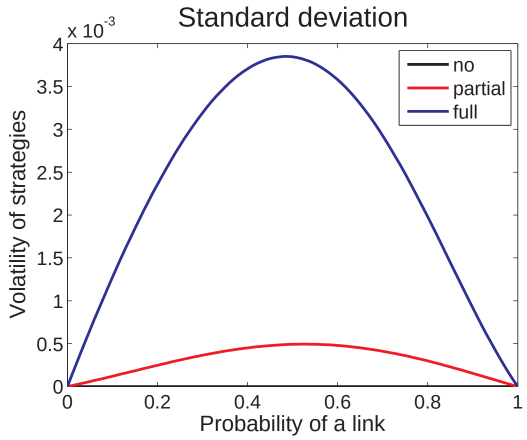
These results can be explained by the equilibrium strategies for which the average value for having one link, respectively two links, with their bands and the volatility for every value of p are displayed in Fig. 3.8. The no information strategies are the highest, but are not subject to volatility. The average partial and full information strategies are similar albeit the volatility is significantly larger in the latter case. A higher volatility in observable market power results in a



(a) Expected strategies with bands.



(b) Strategies for one and two links with bands.



(c) Volatility of strategies.

Figure 3.8: Distribution of strategies.

higher volatility of strategies, since the higher the observable market power the more aggressive offers the trader submits.

The expected value of the bargaining power measure, as in Calvó-Armengol (2001), can be calculated on the basis of the available information about the market. This measure takes on values in the interval $[0, 1]$ and is increasing in market power. For example with no information the expected bargaining power is always equal to one half. Under partial information, having two links results in an expected bargaining power larger than one half, in which case the trader will aim for a higher profit. For this trader it results in possible large profits but reduces the probability of trading. Having one link results in an expected bargaining power less than one half. Under full information certain networks lead to an even higher dispersion between traders' expected market powers.

Volatility of strategies has a negative effect; lower markups cause a slightly higher efficiency whereas higher markups may result in absence of trade. Partial information leads to the highest expected efficiency and the negative effects of higher markups for no information and high volatility for full information are similar.

For example for $p = \frac{1}{2}$, under no information a trader will always use the markup strategy 0.224. Under partial information a trader with one link uses 0.190 and a trader with two links 0.237. Under full information the markup does not only depend on the own number of links but also on the links of others. If a trader has one link his strategy ranges from 0.110 to 0.227, with two links from 0.187 to 0.341.

Volume

In Fig. 3.9 the expected number of trades, i.e. the volume, shows a similar comparison as the expected efficiency. For $p > 0.030$ we find that full information leads to a higher volume than no information, for small p the opposite holds. For every value of the probability of a link, partial information leads to the highest expected volume.

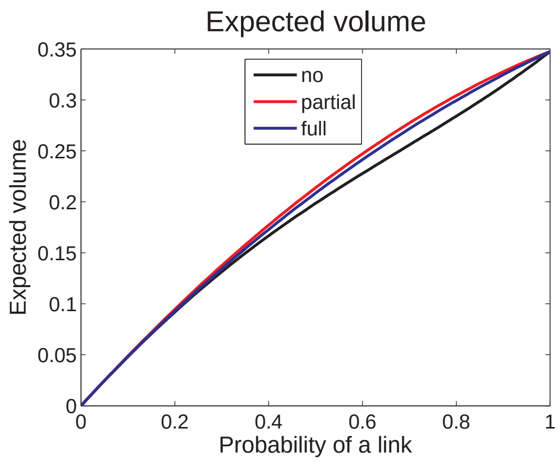


Figure 3.9: Expected volume for all information sets.

Expected profit

The expected profit for a trader that has one link, respectively two links, is shown in Fig. 3.10. A trader with one link has the highest expected profit under partial information; for $p < 0.408$ the lowest under full information, and otherwise the lowest under no information. A trader with two links has the highest expected profit under full information, the lowest under no information. Comparing partial and full information, a trader with one link has a higher expected profit under partial information and a trader with two links under full information. For any value of p the latter is dominated and hence partial information leads to the highest expected efficiency.

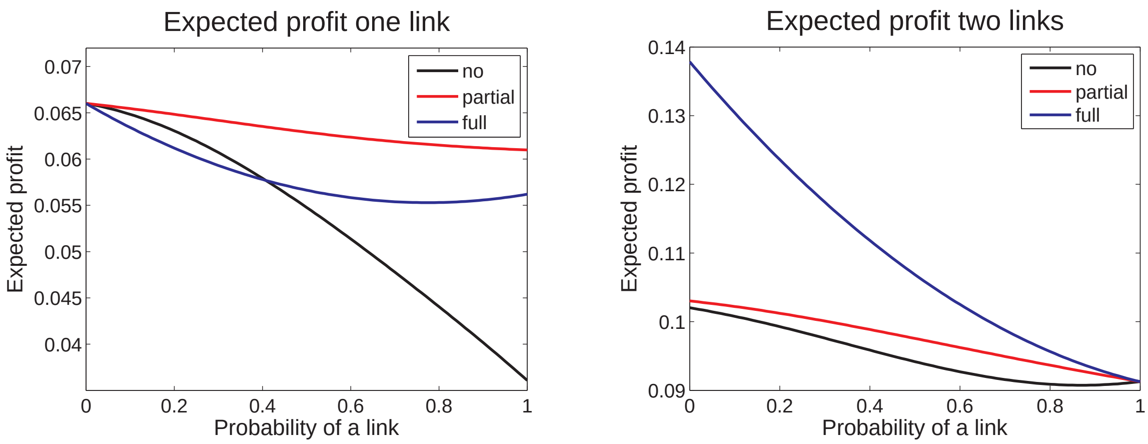


Figure 3.10: Expected profit for having one (left) and two links (right).

3.5 Concluding remarks

In a bipartite Erdős-Rényi market agents only trade in case they are linked to each other. The trading mechanism allows preferred trades to occur, not necessarily the socially optimal allocation of trades. In such a market three ordered sets of information about the network structure are considered; no, partial and full information. These information sets are compared under complete and incomplete information about valuations and costs.

With no information only the probabilities of all networks can be calculated and hence the equilibrium strategy depends on the probability of a link. Partial information allows a trader to base the strategy on the number of his own links and hence the equilibrium strategy depends on the number of a player's own links and the probability that other links are realised. With full information the entire network is known and the equilibrium strategies are based on the realisation of all links.

Under complete information about traders' valuations, in a market with two buyers and two sellers, no and full information lead to attain full efficiency for every probability of a link. Due to strategic behaviour of traders, under partial information allocative efficiency might be reduced. Hence we found that partial information is weakly dominated by no and full information and it is optimal if either everything or nothing of the realisation of the network structure is revealed to traders.

Under incomplete information about valuations and costs, expected efficiency given no and full information are comparable, when we assume that traders use markup and markdown strategies. For a small probability of a link no information outperforms and the opposite holds for a large probability. Partial information leads to the highest expected efficiency, since markups in no information and volatility of strategies in full information are higher, and thus strongly dominates no and full information. Higher markups and a larger volatility increase the probability of absence of trades and hence decrease the expected efficiency. Knowledge of the own links rather than only of the probability distribution improves efficiency, but adding knowledge of the links of others decreases efficiency. It is optimal, when only the realisation of own links is known. Therefore, more information does not necessarily lead to a higher expected allocative efficiency. Furthermore, the expected volume and the expected profit for traders when they have one link, respectively two links, are compared.

We demonstrated that the effect of the quantity of information available to traders on the allocative efficiency is non-monotonic. Moreover, the shape of this non-monotonicity flips over when we switch from complete to incomplete information about traders' valuations.

Appendix A: Profit functions complete information about valuations and costs

The profit functions of buyer b_1 under complete information about valuations and costs are shown below. These are used to calculate the Nash equilibrium markup and markdown strategies. Some best response functions are symmetric and hence we find symmetric markups, and simplifying assumptions about the strategies of others can be made.

Network 1: $b_1 \leftrightarrow s_1$
Buyer b_1 trades if $1 - m_1^d \geq m_1^u$, which results in a profit of $\pi(m_1^d, m_1^u) = 1 - \frac{1-m_1^d+m_1^u}{2}$:
$\mathbb{E}(\Pi_{b_1}) = \pi(m_1^d, m_1^u)\mathbb{1}_{\{1-m_1^d \geq m_1^u\}}$.

Network 2: $b_1 \leftrightarrow s_1$ & $b_1 \leftrightarrow s_2$
Buyer b_1 trades with the seller with the lowest ask $m^u = m_1^u = m_2^u$, if $1 - m_1^d \geq m^u$:
$\mathbb{E}(\Pi_{b_1}) = \pi(m_1^d, m^u)\mathbb{1}_{\{1-m_1^d \geq m^u\}}$.

Network 3: $b_1 \leftrightarrow s_1$ & $b_2 \leftrightarrow s_1$
b_1 only trades when its bid is higher than the bid from b_2, $1 - m_1^d > 1 - m_2^d$, or with probability one half if they are equal, if $1 - m_1^d \geq m_1^u$:
$\mathbb{E}(\Pi_{b_1}) = \pi(m_1^d, m_1^u)\mathbb{1}_{\{1-m_1^d \geq m_1^u\}}(\mathbb{1}_{\{1-m_1^d > 1-m_2^d\}} + \frac{1}{2}\mathbb{1}_{\{m_1^d = m_2^d\}})$.

Network 4: $b_1 \leftrightarrow s_1$ & $b_2 \leftrightarrow s_2$
Buyer b_1 trades if $1 - m_1^d \geq m_1^u$:
$\mathbb{E}(\Pi_{b_1}) = \pi(m_1^d, m_1^u)\mathbb{1}_{\{1-m_1^d \geq m_1^u\}}$.

Network 5: $b_1 \leftrightarrow s_2$, $b_2 \leftrightarrow s_1$ & $b_2 \leftrightarrow s_2$
Buyer b_1 is only connected to s_2. Unless b_2 and s_2 prefer to trade with each other he trades with s_2, if $1 - m_1^d \geq m_2^u$. This happens unless $m_2^d < 1 - m_1^d$ and $m_2^u < m_1^u$:
$\mathbb{E}(\Pi_{b_1}) = \pi(m_1^d, m_2^u)\mathbb{1}_{\{1-m_1^d \geq m_2^u\}}(1 - \mathbb{1}_{\{m_2^d < m_1^d\}}\mathbb{1}_{\{m_2^u < m_1^u\}})$.

Network 6: $b_1 \leftrightarrow s_1$, $b_1 \leftrightarrow s_2$ & $b_2 \leftrightarrow s_2$

If $1 - m_1^d \geq m_1^u$ and $1 - m_1^d \geq m_2^u$, buyer b_1 trades with s_1 except when he and s_2 both prefer to trade with each other. The latter happens when $1 - m_1^d > 1 - m_2^d$ and $m_2^u < m_1^u$:

$$\mathbb{E}(\Pi_{b_1}) = \pi(m_1^d, m_2^u)\mathbb{1}_{\{1-m_1^d \geq m_2^u\}}\mathbb{1}_{\{1-m_1^d > 1-m_2^d\}}\mathbb{1}_{\{m_2^u < m_1^u\}}$$
$$+ \pi(m_1^d, m_1^u)\mathbb{1}_{\{1-m_1^d \geq m_1^u\}}(1 - \mathbb{1}_{\{1-m_1^d > 1-m_2^d\}}\mathbb{1}_{\{m_2^u < m_1^u\}}).$$

Network 7: $b_1 \leftrightarrow s_1$, $b_1 \leftrightarrow s_2$, $b_2 \leftrightarrow s_1$ & $b_2 \leftrightarrow s_2$

In the symmetric equilibrium both sellers use the strategy $m^u = m_1^u = m_2^u$. Buyer b_1 can trade as long as his bid exceeds the askprice of the sellers:

$$\mathbb{E}(\Pi_{b_1}) = \pi(m_1^d, m^u)\mathbb{1}_{\{1-m_1^d \geq m^u\}}.$$

Appendix B: Efficiency under incomplete information about valuations and costs

The possible realisations of the network, without permutations, are given below. For each network we show the best response function of buyer b_1 under incomplete information about valuations and costs. Some best response functions are symmetric and hence we find symmetric markups, and simplifying assumptions about the strategies of others can be made. We denote the profit of a buyer with bid β_i trading with a seller with ask α_j as $\pi(\beta_i, \alpha_j)$ and similar for sellers. For simplicity we disregard in the notation that the offers are a function of both the strategy and the valuation or cost. The integration limits are set to indicate the region of valuations and costs in which trade occurs.

Network 1: $b_1 \leftrightarrow s_1$

$$\left[\frac{\partial}{\partial m_1^d} \int_0^{\frac{1-m_1^d-m_1^u}{1-m_1^u}} \int_{\frac{c_1+m_1^u(1-c_1)}{1-m_1^d}}^{1} \pi(\beta_1, \alpha_1) \mathrm{d}v_1 \mathrm{d}c_1\right]_{\{m_1^u=m_1^d\}} = 0$$

Network 2: $b_1 \leftrightarrow s_1$ & $b_1 \leftrightarrow s_2$

In equilibrium, $m_1^u = m_2^u = m^u$ and thus b_1 trades with the seller with the lowest cost $c_{\text{min}} = \min(c_1, c_2)$ which has pdf $2 - 2c_{\text{min}}$:

$$\frac{\partial}{\partial m_1^d} \int_0^{\frac{1-m_1^d-m^u}{1-m^u}} \int_{\frac{c_{\text{min}}+m^u(1-c_{\text{min}})}{1-m_1^d}}^{1} \pi(\beta_1, \alpha_{\text{min}})(2 - 2c_{\text{min}}) \mathrm{d}v_1 \mathrm{d}c_{\text{min}} = 0.$$

Network 3: $b_1 \leftrightarrow s_1$ & $b_2 \leftrightarrow s_1$

b_1 only trades when its bid is higher than the bid from b_2, $v_1(1 - m_1^d) > v_2(1 - m_2^d)$:

$$\left[\frac{\partial}{\partial m_1^d} \int_0^{\frac{1-m_1^d-m_1^u}{1-m_1^u}} \int_{\frac{c_1+m_1^u(1-c_1)}{1-m_1^d}}^{1} \pi(\beta_1, \alpha_1) \frac{v_1(1-m_1^d)}{(1-m_2^d)} \mathrm{d}v_1 \mathrm{d}c_1\right]_{\{m_2^d=m_1^d\}} = 0.$$

Network 4: $b_1 \leftrightarrow s_1$ & $b_2 \leftrightarrow s_2$

The network is split into two separate markets; the best response function of b_1 is the same as in network 1.

Network 5: $b_1 \leftrightarrow s_2$, $b_2 \leftrightarrow s_1$ & $b_2 \leftrightarrow s_2$

b_1 is only connected to s_2 and unless b_2 and s_2 prefer to trade with each other he can trade with s_2. This happens with probability $\mathbb{P}(\text{trade}) = 1-(1-\min\{1, \frac{v_1(1-m_1^d)}{1-m_2^d}\})(1-\frac{(1-m_2^u)c_2+m_2^u-m_1^u}{1-m_1^u})$. We disregard the possibility that the latter term is negative, since in equilibrium the markup of a trader with two links is higher than the markup of a trader with one link:

$$\left[\frac{\partial}{\partial m_1^d}\int_0^{\frac{1-m_1^d-m_2^u}{1-m_2^u}}\int_{\frac{c_2+m_2^u(1-c_2)}{1-m_1^d}}^1 \pi(\beta_1,\alpha_2)\mathbb{P}(\text{trade})\mathrm{d}v_1\mathrm{d}c_2\right]_{\{m_1^u=m_1^d, m_2^u=m_2^d\}} = 0.$$

Network 6: $b_1 \leftrightarrow s_1$, $b_1 \leftrightarrow s_2$ & $b_2 \leftrightarrow s_2$

b_1 trades with s_1 except when he and s_2 both prefer to trade with each other. This happens with probability $\frac{v_1(1-m_1^d)}{1-m_2^d}(1-\frac{(1-m_2^u)c_2+m_2^u-m_1^u}{1-m_1^u}) = \frac{v_1(1-m_1^d)}{1-m_2^d}\cdot\max\{0, \frac{(1-m_1^u)c_1+m_1^u-m_2^u}{1-m_2^u}\}$, where we disregard the possibility that the first term is negative:

$$\left[\frac{\partial}{\partial m_1^d}\int_0^{\frac{1-m_1^d-m_1^u}{1-m_1^u}}\int_{\frac{c_1+m_1^u(1-c_1)}{1-m_1^d}}^1 \pi(\beta_1,\alpha_1)(1-\frac{v_1(1-m_1^d)}{1-m_2^d}\cdot\max\{0,\frac{(1-m_1^u)c_1+m_1^u-m_2^u}{1-m_2^u}\})\mathrm{d}v_1\mathrm{d}c_1\right.$$
$$\left.+\int_0^{\frac{1-m_1^d-m_2^u}{1-m_2^u}}\int_{\frac{c_2+m_2^u(1-c_2)}{1-m_1^d}}^1 \pi(\beta_1,\alpha_2)\frac{v_1(1-m_1^d)}{1-m_2^d}(1-\frac{(1-m_2^u)c_2+m_2^u-m_1^u}{1-m_1^u})\mathrm{d}v_1\mathrm{d}c_2\right]_{\{m_1^u=m_2^d, m_2^u=m_1^d\}}$$
$$= 0.$$

Network 7: $b_1 \leftrightarrow s_1$, $b_1 \leftrightarrow s_2$, $b_2 \leftrightarrow s_1$ & $b_2 \leftrightarrow s_2$

In equilibrium, $m_1^u = m_2^u = m^u$ and thus b_1 trades with the seller with the lowest cost $c_{\text{min}} = \min(c_1, c_2)$ which has pdf $2-2c_{\text{min}}$, if his bid is higher than the bid of b_2. For a given value of $\beta_1 = v_1(1-m_1^d)$ this happens with probability $\frac{v_1(1-m_1^d)}{1-m_2^d}$. b_1 trades with the seller with the highest cost $c_{\text{max}} = \max(c_1, c_2)$ which has pdf $2c_{\text{max}}$, if his bid is lower than the bid of b_2. For a given value of $\beta_1 = v_1(1-m_1^d)$ this happens with probability $1-\frac{v_1(1-m_1^d)}{1-m_2^d}$:

$$\left[\frac{\partial}{\partial m_1^d}\int_0^{\frac{1-m_1^d-m^u}{1-m^u}}\int_{\frac{c_{\text{min}}+m^u(1-c_{\text{min}})}{1-m_1^d}}^1 \pi(\beta_1,\alpha_{\text{min}})(2-2c_{\text{min}})\frac{v_1(1-m_1^d)}{1-m_2^d}\mathrm{d}v_1\mathrm{d}c_{\text{min}}\right.$$
$$\left.+\int_0^{\frac{1-m_1^d-m^u}{1-m^u}}\int_{\frac{c_{\text{max}}+m^u(1-c_{\text{max}})}{1-m_1^d}}^1 \pi(\beta_1,\alpha_{\text{max}})2c_{\text{max}}(1-\frac{v_1(1-m_1^d)}{1-m_2^d})\mathrm{d}v_1\mathrm{d}c_{\text{max}}\right]_{\{m^u=m_2^d=m_1^d\}} = 0.$$

Full Information

In the full information setting, traders have knowledge of the entire realisation of the network. Hence Nash equilibrium strategies are calculated per possible network.

Network: $b_1 \leftrightarrow s_1$

Solving the best response function of network 1 gives the Nash equilibrium strategies $m_1^d = m_1^u \approx 0.227$. This allows us to calculate expected allocative efficiency given the limitations of the network structure, which is the ratio between the expected surplus from trade, divided by the total expected surplus:

$$\mathbb{E}(\text{AE}) = \frac{\int_0^{\frac{1-m_1^d-m_1^u}{1-m_1^u}} \int_{\frac{c_1+m_1^u(1-c_1)}{1-m_1^d}}^1 (v_1-c_1)\mathrm{d}v_1\mathrm{d}c_1}{\int_0^1\int_{c_1}^1 (v_1-c_1)\mathrm{d}v_1\mathrm{d}c_1} \approx 0.792.$$

Moreover, the expected volume, the ratio between the expected number of trades and the maximal number of trades, equals:

$$\mathbb{E}(\text{Volume}) = \frac{1}{2}\int_0^{\frac{1-m_1^d-m_1^u}{1-m_1^u}} \int_{\frac{c_1+m_1^u(1-c_1)}{1-m_1^d}}^1 1\mathrm{d}v_1\mathrm{d}c_1 \approx 0.125.$$

Similarly to the best response function above we calculate the expected profit for a trader that has one link:

$$\mathbb{E}(\Pi^1) = \int_0^{\frac{1-m_1^d-m_1^u}{1-m_1^u}} \int_{\frac{c_1+m_1^u(1-c_1)}{1-m_1^d}}^1 \pi(\beta_1,\alpha_1)\mathrm{d}v_1\mathrm{d}c_1 \approx 0.066.$$

Network: $b_1 \leftrightarrow s_1$ & $b_1 \leftrightarrow s_2$

Solving the best response functions of network 2 and a symmetric version of 3 gives the Nash equilibrium strategies $m^u = m_1^u = m_2^u \approx 0.110$ & $m_1^d \approx 0.341$:

$$\mathbb{E}(\text{AE}) = \frac{\int_0^{\frac{1-m_1^d-m^u}{1-m^u}} \int_{\frac{c_{\text{min}}+m^u(1-c_{\text{min}})}{1-m_1^d}}^1 (v_1-c_{\text{min}})(2-2c_{\text{min}})\mathrm{d}v_1\mathrm{d}c_{\text{min}}}{\int_0^1\int_{c_{\text{min}}}^1 (v_1-c_{\text{min}})(2-2c_{\text{min}})\mathrm{d}v_1\mathrm{d}c_{\text{min}}} \approx 0.858,$$

$$\mathbb{E}(\text{Volume}) = \frac{1}{2}\int_0^{\frac{1-m_1^d-m^u}{1-m^u}} \int_{\frac{c_{\text{min}}+m^u(1-c_{\text{min}})}{1-m_1^d}}^1 1\cdot(2-2c_{\text{min}})\mathrm{d}v_1\mathrm{d}c_{\text{min}} \approx 0.204,$$

$$\mathbb{E}(\Pi^1) = \int_0^{\frac{1-m_1^d-m_1^u}{1-m_1^u}} \int_{\frac{c_1+m_1^u(1-c_1)}{1-m_1^d}}^1 \pi(\alpha_1,\beta_1)(1-c_1)\mathrm{d}v_1\mathrm{d}c_1 \approx 0.038,$$

$$\mathbb{E}(\Pi^2) = \int_0^{\frac{1-m_1^d-m^u}{1-m^u}} \int_{\frac{c_{\text{min}}+m^u(1-c_{\text{min}})}{1-m_1^d}}^1 \pi(\beta_1,\alpha_1)(2-2c_{\text{min}})\mathrm{d}v_1\mathrm{d}c_{\text{min}} \approx 0.138.$$

Network: $b_1 \leftrightarrow s_1$ & $b_2 \leftrightarrow s_2$

This network is similar to the first network, except that expected volume is doubled.

Network: $b_1 \leftrightarrow s_2, b_2 \leftrightarrow s_1$ & $b_2 \leftrightarrow s_2$

Solving the best response functions of network 5 and a symmetric version of 6 gives the Nash equilibrium strategies $m_1^u = m_1^d \approx 0.169$ & $m_2^u = m_2^d \approx 0.246$:

$$\mathbb{E}(\text{AE}) = \Big[2\int_0^{\frac{1-m_2^d-m_1^u}{1-m_1^u}} \int_{\frac{c_1+m_1^u(1-c_1)}{1-m_2^d}}^1 (v_2 - c_1) \times (1-(1-\tfrac{v_2(1-m_2^d)}{1-m_1^d})(1-\max\{0, \tfrac{(1-m_1^u)c_1+m_1^u-m_2^u}{1-m_2^u}\}))\mathrm{d}v_2\mathrm{d}c_1 + \int_0^{\frac{1-m_2^d-m_2^u}{1-m_2^u}} \int_{\frac{c_2+m_2^u(1-c_2)}{1-m_2^d}}^1 (v_2-c_2)\tfrac{v_2(1-m_2^d)}{1-m_1^d}(1-\tfrac{(1-m_2^u)c_2+m_2^u-m_1^u}{1-m_1^u})\mathrm{d}v_2\mathrm{d}c_2] / [2\int_0^1\int_{c_1}^1 (v_2-c_1)(1-(1-v_2)(1-c_1))\mathrm{d}v_2\mathrm{d}c_1 + \int_0^1\int_{c_2}^1 (v_2-c_2)v_2(1-c_2)\mathrm{d}v_2\mathrm{d}c_2\Big] \approx 0.867,$$

$$\mathbb{E}(\text{Volume}) = \tfrac{1}{2}\Big[2\int_0^{\frac{1-m_2^d-m_1^u}{1-m_1^u}} \int_{\frac{c_1+m_1^u(1-c_1)}{1-m_2^d}}^1 1 \times (1-(1-\tfrac{v_2(1-m_2^d)}{1-m_1^d})(1-\max\{0, \tfrac{(1-m_1^u)c_1+m_1^u-m_2^u}{1-m_2^u}\}))\mathrm{d}v_2\mathrm{d}c_1 + \int_0^{\frac{1-m_2^d-m_2^u}{1-m_2^u}} \int_{\frac{c_2+m_2^u(1-c_2)}{1-m_2^d}}^1 1\cdot\tfrac{v_2(1-m_2^d)}{1-m_1^d}(1-\tfrac{(1-m_2^u)c_2+m_2^u-m_1^u}{1-m_1^u})\mathrm{d}v_2\mathrm{d}c_2\Big] \approx 0.293,$$

$$\mathbb{E}(\Pi^1) = \int_0^{\frac{1-m_1^d-m_2^u}{1-m_2^u}} \int_{\frac{c_2+m_2^u(1-c_2)}{1-m_1^d}}^1 \pi(\beta_1,\alpha_2) \times (1-(1-\min\{1, \tfrac{v_1(1-m_1^d)}{1-m_2^d}\})(1-\tfrac{(1-m_2^u)c_2+m_2^u-m_1^u}{1-m_1^u}))\mathrm{d}v_1\mathrm{d}c_2 \approx 0.056,$$

$$\mathbb{E}(\Pi^2) = \int_0^{\frac{1-m_2^d-m_1^u}{1-m_1^u}} \int_{\frac{c_1+m_1^u(1-c_1)}{1-m_2^d}}^1 \pi(\beta_2,\alpha_1)(1-\tfrac{v_2(1-m_2^d)}{1-m_1^d}\cdot\max\{0, \tfrac{(1-m_1^u)c_1+m_1^u-m_2^u}{1-m_2^u}\})\mathrm{d}v_2\mathrm{d}c_1 + \int_0^{\frac{1-m_2^d-m_2^u}{1-m_2^u}} \int_{\frac{c_2+m_2^u(1-c_2)}{1-m_2^d}}^1 \pi(\beta_2,\alpha_1)\tfrac{v_2(1-m_2^d)}{1-m_1^d}(1-\tfrac{(1-m_2^u)c_2+m_2^u-m_1^u}{1-m_1^u})\mathrm{d}v_2\mathrm{d}c_2 \approx 0.099.$$

Network: $b_1 \leftrightarrow s_1, b_1 \leftrightarrow s_2, b_2 \leftrightarrow s_1$ & $b_2 \leftrightarrow s_2$

Solving the best response function of network 7 gives the Nash equilibrium strategies $m^u = m_1^u = m_2^u = m_1^d = m_2^d \approx 0.187$:

$$\mathbb{E}(\text{AE}) = \Big[\int_0^{\frac{1-m_1^d-m^u}{1-m^u}} \int_{\frac{c_{\text{min}}+m^u(1-c_{\text{min}})}{1-m_1^d}}^1 (v_1 - c_{\text{min}})(2-2c_{\text{min}}) \frac{v_1(1-m_1^d)}{1-m_2^d} \mathrm{d}v_1 \mathrm{d}c_{\text{min}}$$
$$+ \int_0^{\frac{1-m_1^d-m^u}{1-m^u}} \int_{\frac{c_{\text{max}}+m^u(1-c_{\text{max}})}{1-m_1^d}}^1 (v_1 - c_{\text{max}}) 2c_{\text{max}} (1 - \frac{v_1(1-m_1^d)}{1-m_2^d}) \mathrm{d}v_1 \mathrm{d}c_{\text{max}} \Big]$$
$$/ \Big[\int_0^1 \int_{c_{\text{min}}}^1 (v_1 - c_{\text{min}})(2-2c_{\text{min}}) v_1 \mathrm{d}v_1 \mathrm{d}c_{\text{min}}$$
$$+ \int_0^1 \int_{c_{\text{max}}}^1 (v_1 - c_{\text{max}}) 2c_{\text{max}} (1-v_1) \mathrm{d}v_1 \mathrm{d}c_{\text{max}} \Big] \approx 0.913,$$

$$\mathbb{E}(\text{Volume}) = \tfrac{1}{2} \Big[2 \int_0^{\frac{1-m_1^d-m^u}{1-m^u}} \int_{\frac{c_{\text{min}}+m^u(1-c_{\text{min}})}{1-m_1^d}}^1 1 \cdot (2-2c_{\text{min}}) \frac{v_1(1-m_1^d)}{1-m_2^d} \mathrm{d}v_1 \mathrm{d}c_{\text{min}}$$
$$+2 \int_0^{\frac{1-m_1^d-m^u}{1-m^u}} \int_{\frac{c_{\text{max}}+m^u(1-c_{\text{max}})}{1-m_1^d}}^1 1 \cdot 2c_{\text{max}} (1 - \frac{v_1(1-m_1^d)}{1-m_2^d}) \mathrm{d}v_1 \mathrm{d}c_{\text{max}} \Big] \approx 0.347,$$

$$\mathbb{E}(\Pi^2) = \int_0^{\frac{1-m_1^d-m^u}{1-m^u}} \int_{\frac{c_{\text{min}}+m^u(1-c_{\text{min}})}{1-m_1^d}}^1 \pi(\beta_1, \alpha_{\text{min}})(2-2c_{\text{min}}) v_1 \mathrm{d}v_1 \mathrm{d}c_{\text{min}}$$
$$+ \int_0^{\frac{1-m_1^d-m^u}{1-m^u}} \int_{\frac{c_{\text{max}}+m^u(1-c_{\text{max}})}{1-m_1^d}}^1 \pi(\beta_1, \alpha_{\text{max}}) 2c_{\text{max}} (1-v_1) \mathrm{d}v_1 \mathrm{d}c_{\text{max}} \approx 0.091.$$

Combining the 4 possibilities of having only one link in the network, the 4 possibilities of having one trader that has two links, the 2 possibilities of having two linked pairs, the 4 possibilities of having 3 links in total, with the possibility of having a fully connected network, gives a function of the expected efficiency in terms of the probability of a link. We show the expected efficiency reduction due to strategic behaviour and the total expected efficiency. The latter is the ratio of efficiency reductions due to strategic behaviour and the limitations of the network:

$$\mathbb{E}(\text{AE}_\text{s}) = \frac{0.132 \cdot 4p(1-p)^3 + 0.215 \cdot 4p^2(1-p)^2 + 2 \cdot 0.132 \cdot 2p^2(1-p)^2 + 0.311 \cdot 4p^3(1-p) + 0.366 \cdot p^4}{\frac{1}{6} \cdot 4p(1-p)^3 + \frac{1}{4} \cdot 4p^2(1-p)^2 + \frac{1}{3} \cdot 2p^2(1-p)^2 + \frac{43}{120} \cdot 4p^3(1-p) + \frac{2}{5} \cdot p^4},$$

$$\mathbb{E}(\text{AE}_{\text{p,s}}) = \frac{0.132 \cdot 4p(1-p)^3 + 0.215 \cdot 4p^2(1-p)^2 + 2 \cdot 0.132 \cdot 2p^2(1-p)^2 + 0.311 \cdot 4p^3(1-p) + 0.366 \cdot p^4}{\frac{2}{5}}.$$

The expected volume as a function of p is given by

$$\begin{aligned}\mathbb{E}(\text{Volume}) &= 0.125 \cdot 4(1-p)^3 p + 0.204 \cdot 4(1-p)^2 p^2 + 2 \cdot 0.125 \cdot 2(1-p)^2 p^2 \\ &\quad + 0.293 \cdot 4(1-p)p^3 + 0.347 \cdot p^4 \\ &\approx 0.499 \cdot p - 0.181 \cdot p^2 + 0.039 \cdot p^3 - 0.009 \cdot p^4.\end{aligned}$$

Furthermore we calculate the expected strategy as a function of the probability of a link:

$$\begin{aligned}\mathbb{E}(m) &= 0.227 \cdot 4(1-p)^3 p + \frac{2 \cdot 0.110 + 0.341}{3} \cdot 4(1-p)^2 p^2 + 0.227 \cdot 2(1-p)^2 p^2 \\ &\quad + \frac{0.169 + 0.246}{2} \cdot 4(1-p)p^3 + 0.187 \cdot p^4 \\ &\approx \frac{0.908 \cdot p - 1.522 \cdot p^2 + 1.150 \cdot p^3 - 0.349 \cdot p^4}{1-(1-p)^4}.\end{aligned}$$

Conditioning on having one link, respectively two links, we calculate the expected profit:

$$\begin{aligned}\mathbb{E}(m^1) &= 0.227 \cdot (p(1-p) + (1-p)^2) + 0.110 \cdot p(1-p) + 0.169 \cdot p^2 \\ &\approx 0.227 - 0.117 \cdot p + 0.059 \cdot p^2, \\ \mathbb{E}(m^2) &= 0.341 \cdot (1-p)^2 + 0.246 \cdot 2p(1-p) + 0.187 \cdot p^2 \approx 0.341 - 0.191 \cdot p + 0.036 \cdot p^2.\end{aligned}$$

Moreover, the expected strategy of a trader with one link, respectively two links, is given by

$$\begin{aligned}\mathbb{E}(\Pi^1) &= 0.066 \cdot (p(1-p) + (1-p)^2) + 0.038 \cdot p(1-p) + 0.056 \cdot p^2 \\ &\approx 0.066 - 0.028 \cdot p + 0.018 \cdot p^2, \\ \mathbb{E}(\Pi^2) &= 0.138 \cdot (1-p)^2 + 0.099 \cdot 2p(1-p) + 0.091 \cdot p^2 \approx 0.138 - 0.177 \cdot p + 0.130 \cdot p^2.\end{aligned}$$

Partial Information

In this setting, traders know only the realisation of own links and the probability of the other links. Hence Nash equilibrium strategies for having one link m^1 and for having two links m^2 are found based on the best response functions given below.

If a trader has one link the possible networks are 1, 3, 4 or 5 and hence the best response function is given by

$$\Big[\frac{\partial}{\partial m_1^d}(1-p)p\int_0^{\frac{1-m_1^d-m^1}{1-m^1}}\int_{\frac{c_1+m^1(1-c_1)}{1-m_1^d}}^1 \pi(\beta_1,\alpha_1)\mathrm{d}v_1\mathrm{d}c_1$$
$$+(1-p)p\int_0^{\frac{1-m_1^d-m^2}{1-m^2}}\int_{\frac{c_1+m^2(1-c_1)}{1-m_1^d}}^1 \pi(\beta_1,\alpha_1)\frac{v_1(1-m_1^d)}{(1-m^1)}\mathrm{d}v_1\mathrm{d}c_1$$
$$+(1-p)^2\int_0^{\frac{1-m_1^d-m^1}{1-m^1}}\int_{\frac{c_1+m^1(1-c_1)}{1-m_1^d}}^1 \pi(\beta_1,\alpha_1)\mathrm{d}v_1\mathrm{d}c_1$$
$$+p^2\int_0^{\frac{1-m_1^d-m^2}{1-m^2}}\int_{\frac{c_2+m^2(1-c_2)}{1-m_1^d}}^1 \pi(\beta_1,\alpha_2)$$
$$\times\min\{1,1-(1-\tfrac{v_1(1-m_1^d)}{1-m^2})(1-\tfrac{(1-m^2)c_2+m^2-m^1}{1-m^1})\}\mathrm{d}v_1\mathrm{d}c_2\Big]_{\{m^1=m_1^d\}}=0.$$

If a trader has two links networks 2, 6 or 7 are possible and which results in the following best response function:

$$\Big[\frac{\partial}{\partial m_1^d}(1-p)^2\int_0^{\frac{1-m_1^d-m^1}{1-m^1}}\int_{\frac{c_{\min}+m^1(1-c_{\min})}{1-m_1^d}}^1 \pi(\beta_1,\alpha_{\min})(2-2c_{\min})\mathrm{d}v_1\mathrm{d}c_{\min}$$
$$+2(1-p)p[\int_0^{\frac{1-m_1^d-m^1}{1-m^1}}\int_{\frac{c_1+m^1(1-c_1)}{1-m_1^d}}^1 \pi(\beta_1,\alpha_1)\min\{1,1-\tfrac{v_1(1-m_1^d)}{1-m^1}\tfrac{(1-m^1)c_1+m^1-m^2}{1-m^2}\}\mathrm{d}v_1\mathrm{d}c_1$$
$$+\int_0^{\frac{1-m_1^d-m^2}{1-m^2}}\int_{\frac{c_2+m^2(1-c_2)}{1-m_1^d}}^1 \pi(\beta_1,\alpha_2)\frac{v_1(1-m_1^d)}{1-m^1}(1-\tfrac{(1-m^2)c_2+m^2-m^1}{1-m^1})\mathrm{d}v_1\mathrm{d}c_2]$$
$$+p^2\big[\int_0^{\frac{1-m_1^d-m^2}{1-m^2}}\int_{\frac{c_{\min}+m^2(1-c_{\min})}{1-m_1^d}}^1 \pi(\beta_1,\alpha_{\min})(2-2c_{\min})\frac{v_1(1-m_1^d)}{1-m^2}\mathrm{d}v_1\mathrm{d}c_{\min}$$
$$+\int_0^{\frac{1-m_1^d-m^2}{1-m^2}}\int_{\frac{c_{\max}+m^2(1-c_{\max})}{1-m_1^d}}^1 \pi(\beta_1,\alpha_{\max})2c_{\max}(1-\tfrac{v_1(1-m_1^d)}{1-m^2})\mathrm{d}v_1\mathrm{d}c_{\max}\big]\Big]_{\{m^2=m_1^d\}}=0.$$

The Nash equilibrium strategies are calculated with double precision from these two best response functions. The second derivatives in these points are smaller than zero; ensuring a maximum.

No Information

With no information the only knowledge about the network is the probability that a link occurs. Hence the best response function of b_1 is a weighted average over all seven best response functions given above. Because of symmetry we may assume that all other traders use the same strategy, i.e. $m=m_2^d=m_1^u=m_2^u$:

$$\Big[\frac{\partial}{\partial m_1^d} 2(1-p)^3 p \int_0^{\frac{1-m_1^d-m}{1-m}} \int_{\frac{c_1+m(1-c_1)}{1-m_1^d}}^1 \pi(\beta_1,\alpha_1)\mathrm{d}v_1\mathrm{d}c_1$$
$$+(1-p)^2p^2 \int_0^{\frac{1-m_1^d-m}{1-m}} \int_{\frac{c_{\min}+m(1-c_{\min})}{1-m_1^d}}^1 \pi(\beta_1,\alpha_1)(2-2c_{\min})\mathrm{d}v_1\mathrm{d}c_{\min}$$
$$+2(1-p)^2p^2 \int_0^{\frac{1-m_1^d-m}{1-m}} \int_{\frac{c_1+m(1-c_1)}{1-m_1^d}}^1 \pi(\beta_1,\alpha_1)\frac{v_1(1-m_1^d)}{(1-m)}\mathrm{d}v_1\mathrm{d}c_1$$
$$+2(1-p)^2p^2 \int_0^{\frac{1-m_1^d-m}{1-m}} \int_{\frac{c_1+m(1-c_1)}{1-m_1^d}}^1 \pi(\beta_1,\alpha_1)\mathrm{d}v_1\mathrm{d}c_1$$
$$+2(1-p)p^3 \int_0^{\frac{1-m_1^d-m}{1-m}} \int_{\frac{c_2+m(1-c_2)}{1-m_1^d}}^1 \pi(\beta_1,\alpha_2)(1-(1-\min\{1,\frac{v_1(1-m_1^d)}{1-m}\})(1-c_2))\mathrm{d}v_1\mathrm{d}c_2$$
$$+2(1-p)p^3[\int_0^{\frac{1-m_1^d-m}{1-m}} \int_{\frac{c_1+m(1-c_1)}{1-m_1^d}}^1 \pi(\beta_1,\alpha_1)(1-\min\{1,\frac{v_1(1-m_1^d)}{1-m}\}c_1)\mathrm{d}v_1\mathrm{d}c_1$$
$$+\int_0^{\frac{1-m_1^d-m}{1-m}} \int_{\frac{c_2+m(1-c_2)}{1-m_1^d}}^1 \pi(\beta_1,\alpha_2)\frac{v_1(1-m_1^d)}{1-m}(1-c_2)\mathrm{d}v_1\mathrm{d}c_2]$$
$$+p^4\Big[\int_0^{\frac{1-m_1^d-m}{1-m}} \int_{\frac{c_{\min}+m(1-c_{\min})}{1-m_1^d}}^1 \pi(\beta_1,\alpha_{\min})(2-2c_{\min})\frac{v_1(1-m_1^d)}{1-m}\mathrm{d}v_1\mathrm{d}c_{\min}$$
$$+\int_0^{\frac{1-m_1^d-m}{1-m}} \int_{\frac{c_{\max}+m(1-c_{\max})}{1-m_1^d}}^1 \pi(\beta_1,\alpha_{\max})2c_{\max}(1-\frac{v_1(1-m_1^d)}{1-m})\mathrm{d}v_1\mathrm{d}c_{\max}\Big]\Big]_{\{m=m_1^d\}} = 0.$$

The Nash equilibrium strategy m_1^d is solved to be the solution in $[0,1]$ of

$$\begin{aligned}\text{Root}\big[&-40+45p-41p^2+22p^3\\
&+(300-340p+203p^2-61p^3)m_1^d\\
&+(-700+875p-254p^2-77p^3){m_1^d}^2\\
&+(780-1070p+32p^2+306p^3){m_1^d}^3\\
&+(-460+670p+124p^2-258p^3){m_1^d}^4\\
&+(120-180p-40p^2+60p^3){m_1^d}^5\big]=0.\end{aligned}$$

Chapter 4

On the role of Information under Individual Evolutionary Learning in a Continuous Double Auction

4.1 Introduction

The effect of the setup of markets and the information available to traders has been studied intensively on allocative efficiency. Common examples of studied markets are the Call Market and the Continuous Double Auction (CDA). In these markets, experiments as in Cason and Friedman (1996) showed that subjects do not learn to optimise in a Bayesian sense, and hence are not fully rational. However, this paper showed that still a fast converge towards the equilibrium occurs, and hence the efficiency moves towards one quickly. This led to a large literature on simulations to model the boundedly rational behaviour of traders. In these models traders select their next strategy on the basis of the past trading history (Brock and Hommes, 1997, 1998), or by imitating other past strategies (Dawid, 1999). This branch of research distinguishes itself from the standard economic research in which traders are forward looking and hence select a rational equilibrium. One approach to model boundedly rational behaviour is to use learning algorithms, as this does not impose any strict assumptions on the behaviour of traders. The literature uses learning algorithms under full information about trading history, to describe the boundedly behaviour of traders and to show that this may still lead to full efficiency.

In this chapter we study the impact of the information about trading history that is available to traders. In 2002 the New York Stock Exchange introduced the OpenBook system, which opened the content of the limit order book to the public. This allowed experienced traders to use a full history of orders submission, instead of solely knowledge of global market statistics as under the former ClosedBook system. This decision resulted in decreased price volatility and increased liquidity in the market as shown in Boehmer et al. (2005). A number of papers describe the effect of this OpenBook system in simplified markets. In a Continuous Double Auction model, Ladley and Pellizzari (2014) show that the information of the OpenBook is useless. We use the Individual Evolutionary Learning (IEL) algorithm, introduced by Arifovic and Ledyard (2003, 2007) to model the boundedly rational learning behaviour of agents in a Call Market model, in a multi-period Continuous Double Auction that models the common stock exchanges. This individual learning algorithm allows traders to select their strategy depending on the, hypothetical, performance in the previous period. Arifovic and Ledyard (2004) show that this algorithm performs better than other learning rules, and that it suits well in environments with continuous or large strategy spaces. Under this OpenBook system traders can directly determine the hypothetical performance of a strategy, assuming that traders would have behaved the same. Under the ClosedBook system however, traders have to make additional assumptions to estimate the hypothetical foregone payoff of selecting another strategy.

The effect of the OpenBook system is analysed in a simple Call Market by Arifovic and Ledyard (2007). In a Call Market, at the end of each period a market clearing price is computed, at which price agents can trade. Their paper shows that this IEL-algorithm captures the behaviour of subjects in experiments during the learning phase. In the OpenBook system agents can influence the market clearing price, and the offers converge towards an equilibrium value. In the former ClosedBook system it becomes more difficult to influence the market clearing price and agents become pricetakers. The offers converge to the private valuations or costs of the agents. Both in experiments and simulations, Arifovic and Ledyard (2007) show that efficiency is higher in the ClosedBook system.

Anufriev et al. (2013) analyse the OpenBook system in a Continuous Double Auction. Agents place their offer when they enter the market and if possible trade with an existing agent. Otherwise the order is stored in the order book. Agents learn by using the IEL-algorithm, where the same hypothetical payoff functions as in Arifovic and Ledyard (2007) are used to value strategies based on their hypothetical performance in the previous period. Anufriev et al. (2013) find the same bidding behaviour in Open- and ClosedBook as in the latter paper. In the long-run, efficiency is similar in both designs and the price volatility is lower in the OpenBook system. In the formerly used ClosedBook system, where only information about past average prices is available, they proved divergence of bids and asks away from the equilibrium price range. This is the consequence of their choice for a payoff function that only distinguishes between offers below and above the average price of the previous period. As a result, investors trade with a high probability but may generate a very small profit. However, the latter paper states that ”the specification (of the ClosedBook hypothetical foregone payoff function) is a strong assumption ... which may affect results of IEL”.

The paper Fano et al. (2013) compares the Call Market and the Continuous Double Auction in a setting closely related to the ClosedBook system. The strategies of traders emerge over time by a genetic algorithm. Traders with the same valuation are compared on the basis of their average profit over some evaluation window, after which individuals with a low average profit take on strategies of better performing agents. Thus only information contained in the ClosedBook system is used. Similar to Arifovic and Ledyard (2007) they find that traders in a Call Market become pricetakers and offer their valuation of cost. Contrary to Anufriev et al. (2013), in a Continuous Double Auction traders become pricemakers and offers converge towards the equilibrium price.

In this chapter we demonstrate in simulations that the specification of the hypothetical payoff functions indeed plays a crucial role in a Continuous Double Auction model under the IEL learning algorithm. We show, that when agents use more information to estimate the hypothetical foregone payoff of each possible offer, bids and asks tend to drift towards each other. In this approach investors learn to increase their expected profit by submitting an order that has a higher possible profit. The probability of trading will be lower in this situation but is outweighed by the increase in possible profit. Similarly to Fano et al. (2013), in this setting bids and asks will not converge to the redemption value, but to some equilibrium price.

In the learning phase of the Continuous Double Auction, we examine the effect of the OpenBook system in simulations. Furthermore we compare these results with the simulations in the Call Market performed by Arifovic and Ledyard (2007). Hence we study whether the comparison of efficiency between Open- and ClosedBook differs between both markets.

The results of the long-run simulations allow us to compare the effect of the difference between our ClosedBook hypothetical payoff function and the function used in Anufriev et al. (2013). We study whether the differences between Open- and ClosedBook, as indicated in Anufriev et al. (2013) still hold under the new ClosedBook hypothetical foregone payoff function. Moreover, we show robustness of our results with respect to the size of the market and the number of units a trader desires to buy or sell.

This chapter is organised as follows. The Call Market and Continuous Double Auction models are described in Section 4.2. The renewed Individual Evolutionary Learning algorithm is explained in Section 4.3. The setup of the simulations as well as the parameters and methods used are described in Section 4.4. In Section 4.5 the simulations in the learning phase are compared with the Call Market results of Arifovic and Ledyard (2007). We compare the results in the long-run that are described in Section 4.6 with Anufriev et al. (2013). Robustness with respect to the number of units traded is studied in Section 4.7 and with respect to the number of traders in Section 4.8. Concluding remarks are given in Section 4.9.

4.2 Market setup

We describe the environments and the competitive equilibrium used to compare OpenBook (OP) and ClosedBook (CB). The Call Market and the Continuous Double Auction (CDA) are explained together with the benchmark environments we use.

4.2.1 The environments

Each environment is determined by a set of buyers and a set of sellers with their valuations and costs for the good, also referred to as redemption values. In each trading period $t \in \{1, \cdots, T\}$ each of the buyers $b \in \{1, \cdots, B\}$ desires to consume one unit of the good and each of the sellers $s \in \{1, \cdots, S\}$ is endowed with one unit of the good. The buyers have the same valuation of V_b per unit in every period, sellers have fixed costs of C_s that only need to be paid when a transaction occurs. Agents have knowledge of their own redemption value, but not of the values or distribution of the other agents.

The environments are shown as a vector of valuations and a vector of costs. For example the environment $\{[0.9, 0.9], [0, 0.2, 0.4]\}$ consists of two buyers with valuation 0.9 and three sellers with costs of 0, 0.2 and 0.4. We mainly use two environments from Arifovic and Ledyard (2007) and Anufriev et al. (2013) in our analysis, in order to compare with these papers.

The demand and supply functions are determined from the valuations and costs of traders. The equilibrium quantity is denoted as q^* and the interval of equilibrium prices is given by $[p_L^*, p_H^*]$. The traders that can gain a positive profit in equilibrium are called intramarginal, whereas the traders that in equilibrium cannot make a positive profit and therefore will not trade are called extramarginal. The payoff of a buyer equals $U_b(p) = V_b - p$ if he traded at price p and zero otherwise. The payoff of a seller equals $U_s(p) = p - C_s$ after a trade at price p and zero otherwise.

The allocative value of a trading period is the sum of the payoffs of all agents. The allocative efficiency is defined as the ratio between the allocative value in a trading period and the maximal allocative value. The market is fully efficient when during a period all intramarginal agents trade. The efficiency of a period can be lower when an extramarginal agent trades, or when intramarginals simply do not trade. Furthermore we study the average price, the price volatility and the number of transactions.

4.2.2 Call Market

A Call Market is often used to determine the opening- and closing prices at stock exchanges. In period t each buyer b submits a bid $b_{b,t}$ and each seller s submits an ask $a_{s,t}$. At the end of the period all offers are collected and a market clearing price is computed. This market clearing price is the midpoint of the range where supply equals demand. Each buyer that submitted a bid above this value, and each seller with an ask below this value will trade at the market clearing price.

In the ClosedBook system the market clearing price is publicly available after the period. Moreover, the own offer is known and thus traders know their own profit and own trading history. The market clearing price does not reveal the entire sequence of orders. Hence the limited information does not allow agents to determine precisely how they can influence their transaction price. In a multi-period setting all offers are converging towards the redemption values, since this increases the probability of trading and in principle does not influence the market clearing price. In this system, Arifovic and Ledyard (2007) show that agents behave as pricetakers under the Individual Evolutionary Learning algorithm.

In the OpenBook system not only the market clearing price is known, but also all offers become publicly available after the period. Full information is available and therefore agents can exactly calculate how they can influence the market clearing price, assuming that other agents submit the same offers. In a multi-period setting, offers will not converge towards the redemption values but to some equilibrium price and Arifovic and Ledyard (2007) conclude that agents behave as pricemakers.

4.2.3 Continuous Double Auction

A Continuous Double Auction model is used to describe the regular behaviour during the trading day at stock exchanges. In this model, buyers and sellers arrive in a random sequence during a trading period and submit their bid. Agents can select their order only once, before the period and hence unconditional on the state of the order book. The bid of buyer b and the ask of seller s in period t are denoted as $b_{b,t}$ and $a_{s,t}$. If an arriving order can be matched with an order from the book according to the price-time priority, the transaction takes place at the price of the order in the book and both orders are removed. If the arriving order cannot be matched, it is stored in the order book. At the end of the period the order book with unmatched orders is cleared. The latter is a strict assumption that is often made in Continuous Double Auction models as Anufriev et al. (2013). However we show that our results are robust with respect to larger markets, in which case this assumption is less important.

The prices in period t at which buyer b and seller s trade are denoted as $p_{b,t}$ and $p_{a,t}$. The payoff of a buyer equals $U_b(p) = V_b - p_{b,t}$ if he trades and zero otherwise. The payoff of a seller equals $U_s(p) = p_{s,t} - C_s$ after a trade and zero otherwise. We note that the payoff depends not only on the own offer, but also on the arrival sequence. The arrival sequence does not only determine the trading partner but also at which price the trade occurs.

In the ClosedBook system the average price of the last period is public information and each trader knows if his offer resulted in a trade and if so at which transaction price. Submitting an offer equal to the own redemption value will increase the probability of trading, but the trade may occur at the price of the own offer, yielding a profit of zero.

All offers and thus the average price are publicly available in the OpenBook system. Each agent can exactly determine what the payoff would have been for each possible offer, assuming that other agents submit the same offers and arrive in the same sequence. Agents learn to select the offer that has the highest expected payoff and the offers will converge to some equilibrium value.

4.3 Individual Evolutionary Learning algorithm

Agents learn to submit the most profitable offer by the Individual Evolutionary Learning (IEL) algorithm as introduced by Arifovic and Ledyard (2003, 2007). It is shown in the paper Arifovic and Ledyard (2004) that this algorithm performs better than other learning rules, and it suits well in environments with continuous or large strategy spaces.

At the beginning of the period every agent selects an offer from an individual pool of strategies probabilistically, depending on the hypothetical foregone payoff of these strategies. After the trading period, the new hypothetical payoffs of own strategies are determined assuming the same arrival sequence and offers of other traders. The pool of agents' strategies is also updated using these new hypothetical payoffs as follows. First, every strategy may mutate with a given probability. Second, the places in the new pool are filled by those strategies that have higher hypothetical payoffs. After every agent updated own pool, the new period starts and, again, every agent selects probabilistically a strategy.

Pool of strategies

Every trader has an individual pool of strategies, which is a subset of the continuous strategy space: the set $B_{b,t}$ of K bids $b_{b,t} \leq V_b$ for buyer b and the set $A_{s,t}$ of K asks $a_{s,t} \geq C_s$ for seller s. Hence the Individual Rationality constraint of Anufriev et al. (2013) is satisfied and traders cannot submit offers that could result in a negative profit. The offers are initially drawn from a uniform distribution on $[0, V_b]$ and $[C_s, 1]$ respectively. Even though this pool of strategies is updated every period, the number of strategies remains constant. However, it is possible that a single strategy starts to dominate the pool of strategies over time and occupies many positions in the pool.

Mutation

With a fixed small probability ρ a strategy mutates and otherwise remains the same. When a strategy mutates, a normally distributed variable with mean zero and a fixed variance is added to the old strategy.

Replication

At the end of the period and after possible mutations, the foregone payoff is calculated for each strategy while taking the strategies from others constant. Two strategies are randomly selected and compared on the basis of their hypothetical foregone payoff, after which the best performing strategy occupies a spot in the new pool of strategies. This procedure is repeated K times, in order to select a new strategy for every position in the pool. In the field of Genetic Algorithms this is denoted as a tournament selection process.

Hypothetical foregone payoff functions

After a period traders have knowledge about the profit of the strategy that they selected. However, it is vital for traders to value also the other strategies from the pool. Hence traders calculate, or estimate, the hypothetical payoff of other strategies in the previous period. This is done given the offers of traders and the sequence of order submission in the previous period. Hence traders do not take into account that others are also learning in between periods and thus behave boundedly rational. Moreover, similar to Arifovic and Ledyard (2007) and Anufriev et al. (2013) traders compare strategies solely on their performance in the previous period. It is unclear how a longer sampling period would affect the comparison between Closed- and OpenBook.

The order book and arrival sequence of the previous period are known in the OpenBook system and hence it is possible to calculate the hypothetical foregone payoffs. These can be determined exactly for every possible strategy, given the strategies of others and the arrival sequence from the previous period. We assume the same arrival sequence because of the computational problems this would yield for traders in large markets, as the number of computations will be multiplied by the factorial of the number of traders. For example, with only one buyer and one seller who in the previous period submitted ask $a_{s,t}$, the hypothetical foregone payoff of the buyers' strategy (b_i, n_i) is equal to $V_b - a_{s,t}$ when $n_i > n_{s,t}$ and $b_i \geq a_{s,t}$, equal to $V_b - b_i$ when $n_i < n_{s,t}$ and $b_i \geq a_{s,t}$ and zero otherwise. When $n_i = n_{s,t}$ one of the traders randomly arrives first, and the hypothetical foregone payoff of the strategy (b_i, n_i) equals $\frac{1}{2}(V_b - a_{s,t}) + \frac{1}{2}(V_b - b_i)$. The hypothetical foregone payoff functions are given by

$$U_{b,t}(b_i|\mathcal{J}^{OP}_{b,t}) = \begin{cases} V_b - p^*_{b,t}(b_i) & \text{if bid } b_i \text{ resulted in a transactionat price } p^*_{b,t}(b_i) \\ 0 & \text{otherwise,} \end{cases}$$

$$U_{s,t}(a_j|\mathcal{J}^{OP}_{s,t}) = \begin{cases} p^*_{s,t}(a_j) - C_s & \text{if ask } a_j \text{ resulted in a transactionat price } p^*_{s,t}(a_j) \\ 0 & \text{otherwise.} \end{cases}$$

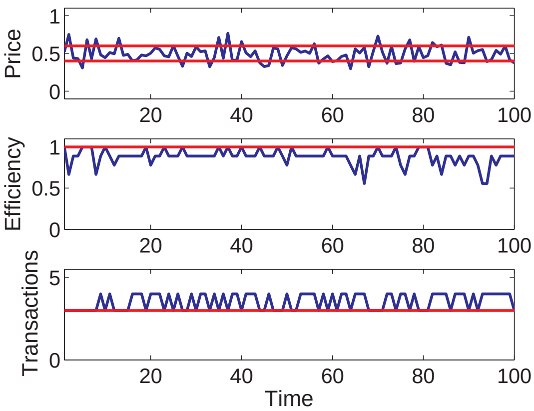
In the ClosedBook system however, only the average price of the previous period is known and it is necessary to determine some estimated payoff for every strategy. The average price, P^{av}_t, is set to the previous average price if no trade occurred. In this setup it is only possible to estimate the probability that a different strategy would have resulted in a trade. We extend the paper of Anufriev et al. (2013) by introducing a different hypothetical foregone payoff function in the ClosedBook setting, that uses more of the available information.

Anufriev et al. (2013) proved convergence of submitted bids and asks towards the redemption values of the agents in the ClosedBook system. This is the consequence of their choice for a payoff function chosen to be the closest to the Call Market payoff function of Arifovic and Ledyard (2007). Anufriev et al. (2013) state that ”this specification is a strong assumption” and that this ”may effect results of IEL”. This foregone payoff function solely distinguishes between offers below and above the average price of the previous period. The consequence is that investors trade with a high probability but may generate a very small profit. The ClosedBook hypothetical payoff function of Anufriev et al. (2013) is given by

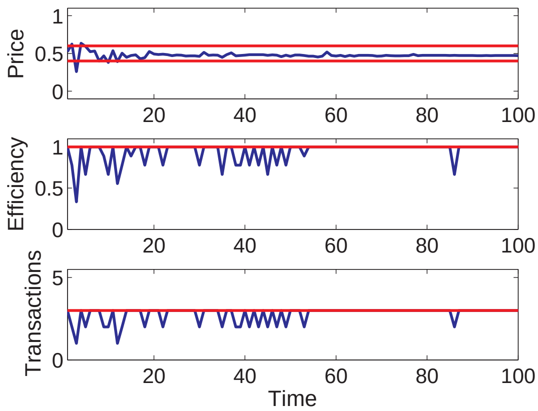
$$U_{b,t}(b_i|\mathcal{J}^{CL}_{b,t}) = \begin{cases} V_b - P^{av}_t & \text{if } b_i \geq P^{av}_t \\ 0 & \text{otherwise,} \end{cases}$$

$$U_{s,t}(a_j|\mathcal{J}^{CL}_{s,t}) = \begin{cases} P^{av}_t - C_s & \text{if } a_j \leq P^{av}_t \\ 0 & \text{otherwise.} \end{cases}$$

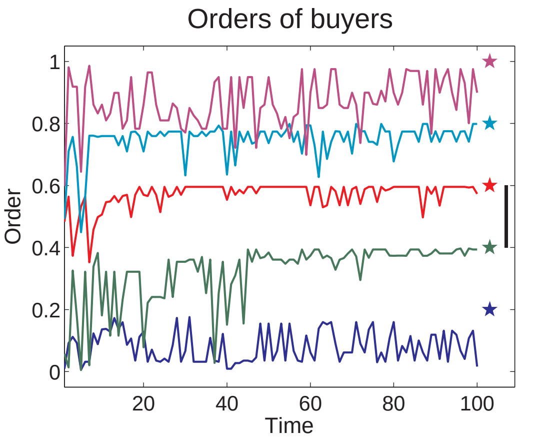
Fig. 4.1 shows a simulation in the symmetric S5-environment that has valuations and costs $\{[1, 0.8, 0.6, 0.4, 0.2], [0.8, 0.6, 0.4, 0.2, 0]\}$ and consists of 5 buyers and 5 sellers, where the payoff function of Anufriev et al. (2013) is used. The figures can be described as follows. Part (a) and (b) show the average price, efficiency and number of transactions in every period for



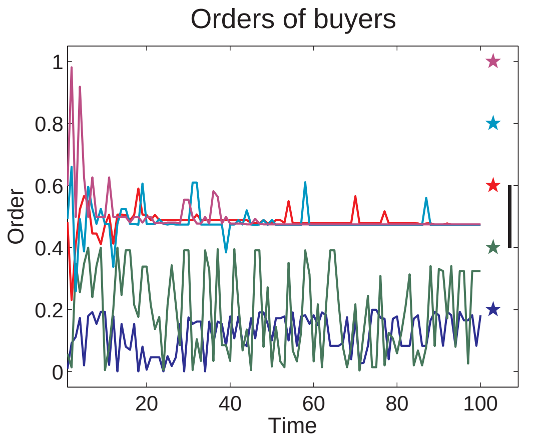
(a) Aggregate outcomes under CL.



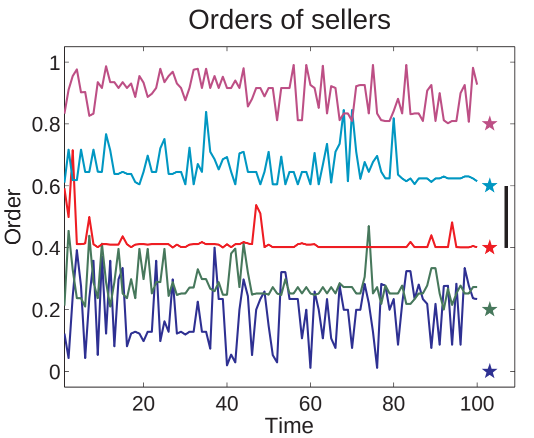
(b) Aggregate outcomes under OP.



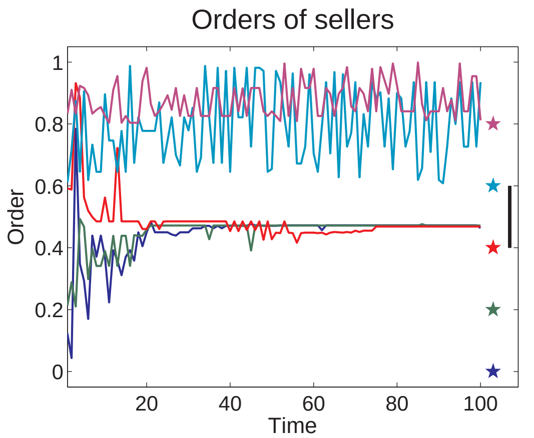
(c) Individual bids under CL.



(d) Individual bids under OP.



(e) Individual asks under CL.



(f) Individual asks under OP.

Figure 4.1: Dynamics in the S5-environment under the foregone payoff function of Anufriev et al. (2013). In the OpenBook system they observe that offers of intramarginal traders rapidly move towards the equilibrium price range. In the ClosedBook however, offers of traders move towards their valuation of cost. Extramarginal agents have more opportunities to trade, yielding a higher number of transactions and a slightly lower efficiency.

ClosedBook (CL) respectively OpenBook (OP). The horizontal lines indicate the equilibrium price range $0.4 - 0.6$, the equilibrium efficiency and the equilibrium quantity 3. The individual offers in the Open- and ClosedBook are shown in parts (c)-(f). The equilibrium price range is here indicated by the vertical line and the redemption values by the stars on the right. While in the OpenBook system offers of intramarginal traders converge towards the equilibrium price range, in ClosedBook they observe convergence of offers towards the redemption values, i.e. a divergence away from the equilibrium range of prices. As a result, extramarginal agents trade frequently and more transactions occur than the equilibrium number of 3. This leads to a lower efficiency and a higher price volatility. Under OpenBook mutation might prevent trades, hence the number of transactions and the efficiency are occasionally lower than the equilibrium value.

We now introduce a new ClosedBook hypothetical foregone payoff function, that uses more of the available information in the CDA. This payoff function uses more of the available information to derive hypothetical foregone payoffs. Both Arifovic and Ledyard (2007) and Anufriev et al. (2013) use a ClosedBook hypothetical foregone payoff function in which the main interest of agents is to trade. We will show that under the new ClosedBook payoff function intramarginal traders have a higher profit. We assume the following thought process of a buyer. A buyer who traded in the last period trades again with probability 1, if he submits a bid that is higher than the minimum of his last bid and the average price. A buyer who did not trade, trades if and only if he submits a bid above the maximum of his last bid and the average price. A bid that we assume to result in a trade, is matched with an ask price with the average price as estimated value. The trading price is equal to the bid with probability $\frac{1}{2}$ and with the same probability equal to P_t^{av}. A similar argument is used to construct the hypothetical payoff function for sellers. The new and old OpenBook hypothetical foregone payoff functions only coincide for a trader that traded at the average price. A buyer would under the old payoff function not distinguish between bids above the average price. Under the new payoff function a buyer prefers a lower bid, when he assumes that this also results in a trade. The hypothetical foregone payoff function is given by:

If the agent traded in the last period:

$$U_{b,t}(b_i|\mathcal{J}_{b,t}^{CL}) = \begin{cases} \frac{1}{2}(V_b - b_i) + \frac{1}{2}(V_b - P_t^{av}) & \text{if } b_i \geq \min(P_t^{av}, p_{b,t}) \\ 0 & \text{otherwise,} \end{cases}$$

$$U_{s,t}(a_j|\mathcal{J}_{s,t}^{CL}) = \begin{cases} \frac{1}{2}(P_t^{av} - C_s) + \frac{1}{2}(a_j - C_s) & \text{if } a_j \leq \max(P_t^{av}, p_{s,t}) \\ 0 & \text{otherwise.} \end{cases}$$

If the agent did not trade in the last period:

$$U_{b,t}(b_i|\mathcal{J}_{b,t}^{CL}) = \begin{cases} \frac{1}{2}(V_b - b_i) + \frac{1}{2}(V_b - P_t^{av}) & \text{if } b_i \geq \max(P_t^{av}, b_{b,t}) \\ 0 & \text{otherwise,} \end{cases}$$

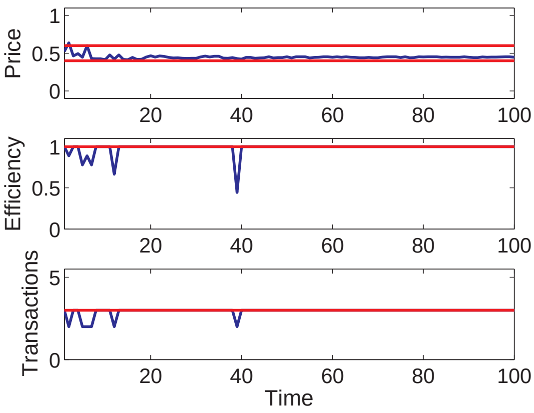
$$U_{s,t}(a_j|\mathcal{J}_{s,t}^{CL}) = \begin{cases} \frac{1}{2}(P_t^{av} - C_s) + \frac{1}{2}(a_j - C_s) & \text{if } a_j \leq \min(P_t^{av}, a_{s,t}) \\ 0 & \text{otherwise.} \end{cases}$$

We observe the impact of the new ClosedBook hypothetical foregone payoff function in Fig. 4.2. Under ClosedBook, instead of a movement of offers towards the valuations and costs of traders as seen in Fig. 4.1, the offers move towards the equilibrium price range. The OpenBook hypothetical payoff function remains unchanged. After the learning phase offers occasionally fluctuate when traders use a different strategy due to mutation. These fluctuations may reduce efficiency and occur less frequently under the ClosedBook system. Hence the efficiency and number of transactions seem higher and the price volatility lower than under OpenBook.

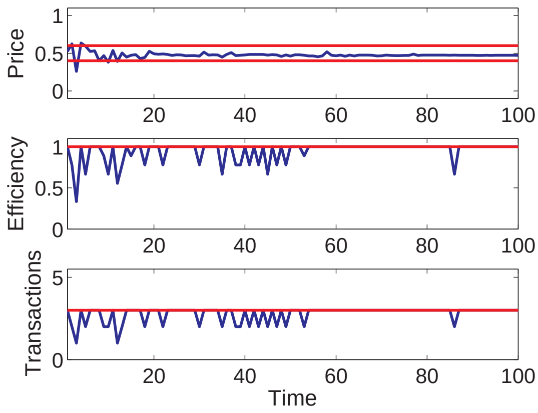
Selection of a strategy from the pool

Initially, every strategy is equally likely to be chosen. In the next periods the probability that a certain strategy is selected is proportional to its hypothetical payoff in the previous period. After mutation has taken place, the probability that a buyer b selects strategy b_i for period $t+1$ is given by

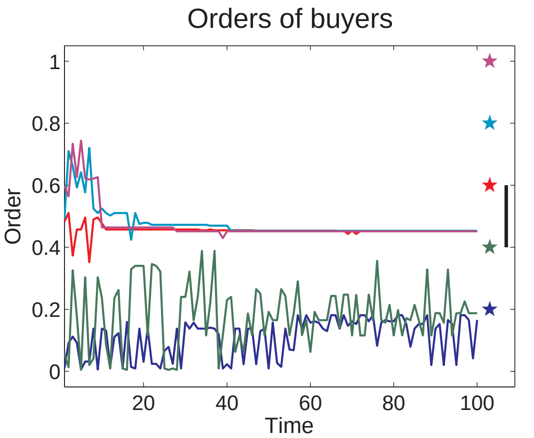
$$\pi_{b,t+1}(b_i) = \frac{U_{b,t}(b_i|\mathcal{J}_t)}{\sum_{i=1}^{K} U_{b,t}(b_i|\mathcal{J}_t)}.$$



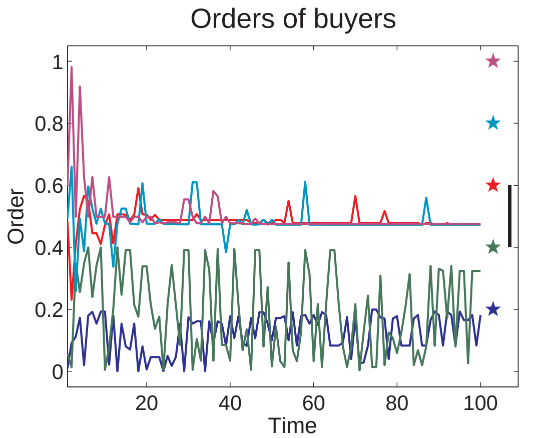
(a) Aggregate outcomes under CL.



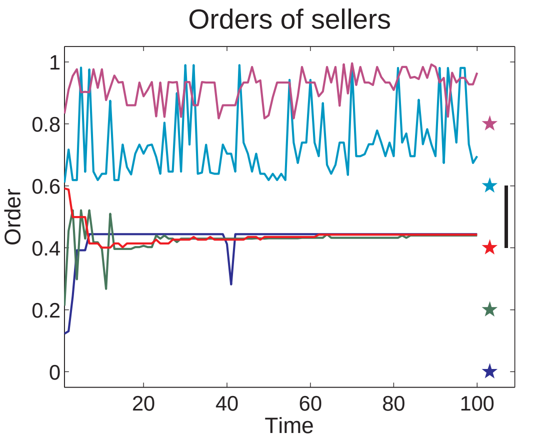
(b) Aggregate outcomes under OP.



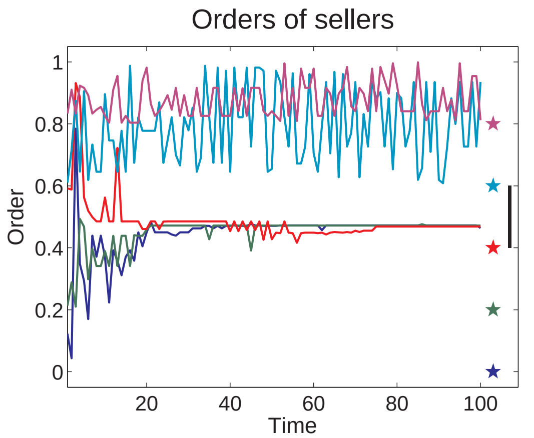
(c) Individual bids under CL.



(d) Individual bids under OP.



(e) Individual asks under CL.



(f) Individual asks under OP.

Figure 4.2: Dynamics in the S5-environment under the new ClosedBook payoff function. Where offers move towards valuations and costs in Fig. 4.1 under the ClosedBook payoff function of Anufriev et al. (2013), the new payoff function leads to a convergence of offers towards the equilibrium price range. The OpenBook payoff function is identical.

The variables used in this Individual Evolutionary Learning algorithm are the size of the individual pools, the probability and the distribution of mutation and the replication rate. In the next section the values of these variables are given, as well as the methods to compare average behaviour of Open- and ClosedBook.

4.4 Methodology

In the standard model we examine three environments used in Arifovic and Ledyard (2007) and Anufriev et al. (2013). The symmetric environment $\{[1, 0.8, 0.6, 0.4, 0.2], [0.8, 0.6, 0.4, 0.2, 0]\}$ which is denoted as S5 and $\{[1, 0.93, 0.92, 0.81, 0.5], [0.66, 0.55, 0.39, 0.39, 0.3]\}$ as the AL-environment, which both consist of 5 buyers and 5 sellers. The results of the latter environment will be shown in the appendices. The third environment is introduced by Gode and Sunder (1997) for its simplicity, which allows for obtaining intuition about the behaviour under the IEL-algorithm and the differences between Open- and ClosedBook. In this so called GS-environment we have 1 seller with cost 0, 1 buyer with valuation 1 and n buyers with valuation β. Of main concern is the GS-environment with 3 extramarginal buyers with valuation 0.5. We will find similar results for all environments and hence show a robustness with respect to the environment. The demand- and supply functions of these environments are shown in Fig. 4.3.

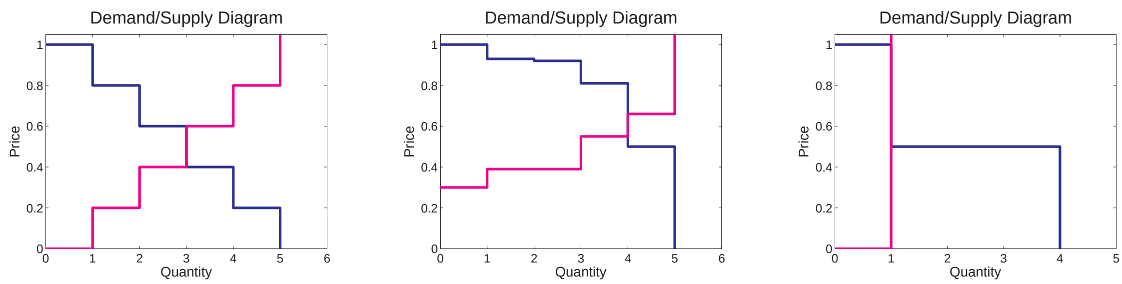


Figure 4.3: The main environments used: (a) The symmetric S5-environment $\{[1, 0.8, 0.6, 0.4, 0.2], [0.8, 0.6, 0.4, 0.2, 0]\}$ with 5 buyers and 5 sellers, (b) the AL-environment $\{[1, 0.93, 0.92, 0.81, 0.5], [0.66, 0.55, 0.39, 0.39, 0.3]\}$ with 5 buyers and 5 sellers and (c) the GS-environment $\{[1, 0.5, 0.5, 0.5], [0]\}$ with 4 buyers and 1 seller.

We study efficiency, average price, the price volatility and the number of transactions under the Individual Evolutionary Learning algorithm, where the allocative efficiency is defined as the ratio between the allocative value in a trading period and the maximal allocative value. This algorithm is studied under the parameters of Arifovic and Ledyard (2007), to be able to make a thorough comparison with their Call Market results. As a result our setup differs from Anufriev et al. (2013) where mutation is uniform. However, simulations have shown that our results are not affected by the distribution of this mutation. Every agent is given an individual pool of strategies, which is in principle of size $K = 100$. A strategy can mutate with a probability of 0.033. When a strategy mutates a normally distributed term with mean 0 and a standard deviation of 0.1 is added to the former strategy. When the mutated strategy lies outside the strategy space, a new normally distributed variable is drawn. In the replication phase K pairs are compared. All the averages are calculated over $S = 100$ random seeds.

We denote the periods $1 - 20$ as the learning phase and moreover highlight learning by considering the periods $1 - 5$ and $16 - 20$ as well. With a small probability it can happen that no transaction takes place in a period during the learning phase, since we only use few agents. In that case the average price of the previous period remains, similar to real markets.

We will see that after the learning phase the market becomes quite stable and the offers and average price only fluctuate within a certain range. We denote this behaviour as the "equilibrium" phase. During the periods $101 - 200$ we study the response of the learning algorithm to mutation.

4.5 Learning phase

The learning phase of a Call Market is studied in Arifovic and Ledyard (2007). They conclude that the IEL-algorithm with the parameters above explains their experiments quite well. In the OpenBook system agents behave as pricemakers and in the ClosedBook system as pricetakers. In their simulations and experiments the efficiency is higher in the ClosedBook system. In this section we compare simulations of the Open- and ClosedBook in a Continuous Double Auction market during periods $1 - 20$ and the subperiods $1 - 5$ and $16 - 20$. Moreover, during this learning phase we study whether the comparison of efficiency is identical to the comparison of Open- and ClosedBook in the Call Market.

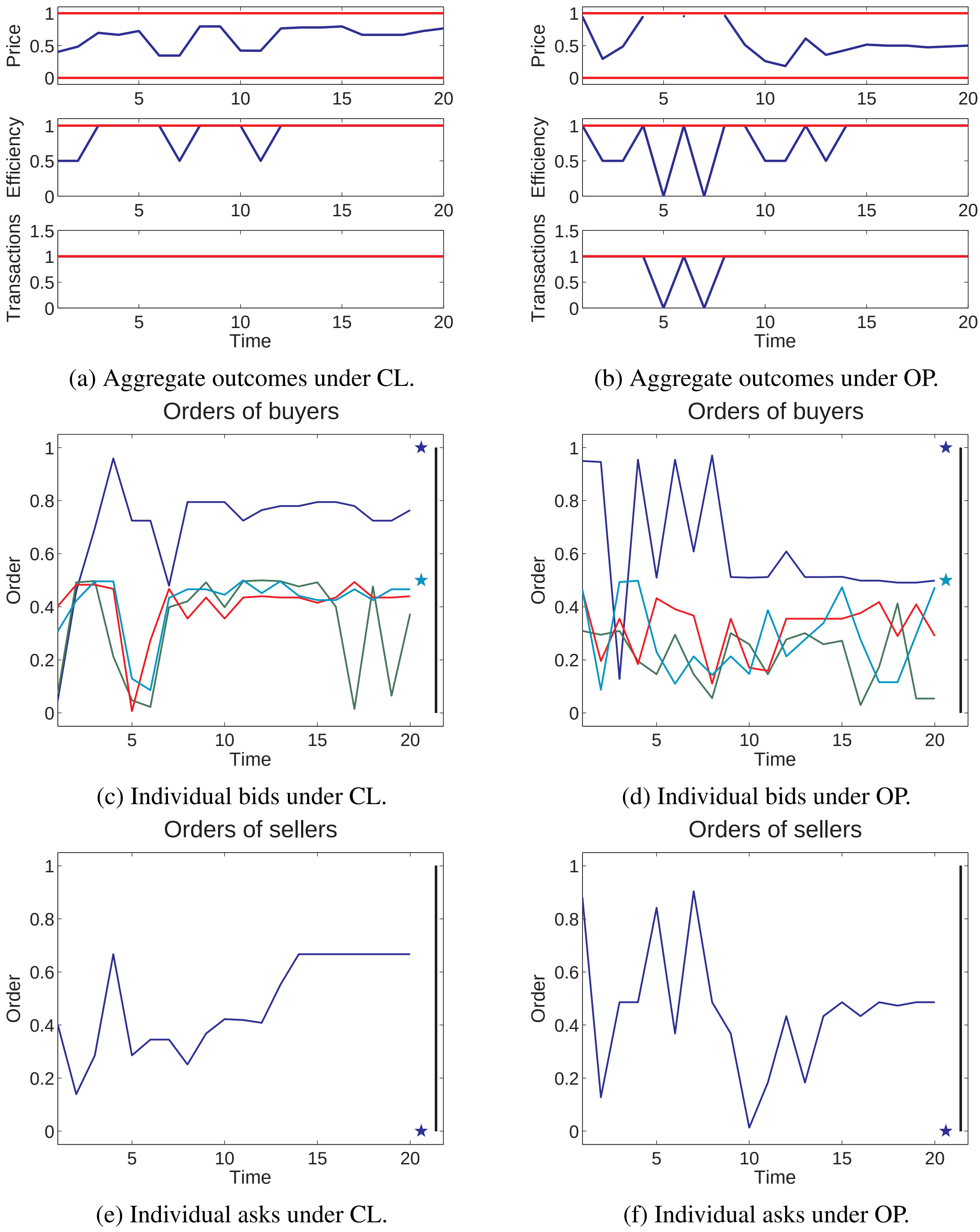
4.5.1 Gode Sunder-environment

The Gode Sunder-environment, denoted as GS-environment, is simulated in Fig. 4.4 with 3 extramarginal buyers with valuation $\beta = 0.5$. In both settings we observe that the intramarginal traders seem to coordinate on offers above the valuation of the extramarginal buyers. However, during the first periods the asks from sellers are often quite low, which may result in a trade with an extramarginal buyer.

In Anufriev et al. (2013) the expected efficiency in the ClosedBook system is close to $\frac{1}{2}+\frac{1}{2}\beta$ as n goes to infinity, but under our payoff function the expected efficiency is close to the equilibrium value of 1 even for $n = 3$. In the new setting the agents will learn to select a strategy with a higher expected profit and the buyer will in general bid and the seller ask more than β, thus the efficiency will not be lowered due to transactions with extramarginal buyers.

4.5.2 S5- and AL-environments

Fig. 4.5, and Fig. 4.12 in the appendix, show the behaviour in simulations during the learning phase over periods $1 - 20$ of the S5- en AL-environment. We notice that the orders of intramarginal traders converge fast towards the equilibrium price range. The initial pool of strategies is drawn uniformly and hence in the first periods the submitted offers are almost uniform. In these figures we observe that after five periods the orders are relatively close to each other and the standard deviation of individual offers has significantly dropped. The fast convergence originates from two effects. Strategies far away from the equilibrium price range are removed during the replication process, for example if the strategy with the lowest hypothetical foregone payoff occurs only once in the pool it cannot attain a spot in the updated pool. Second, the algorithm

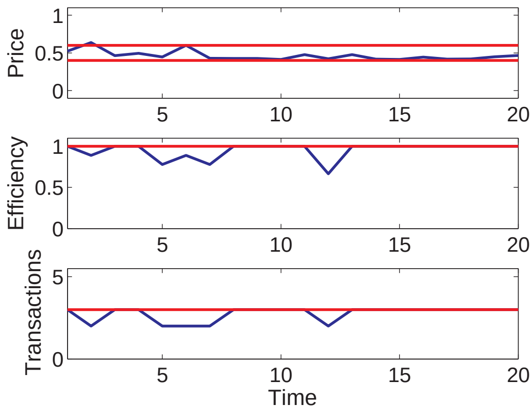


(a) Aggregate outcomes under CL. (b) Aggregate outcomes under OP.

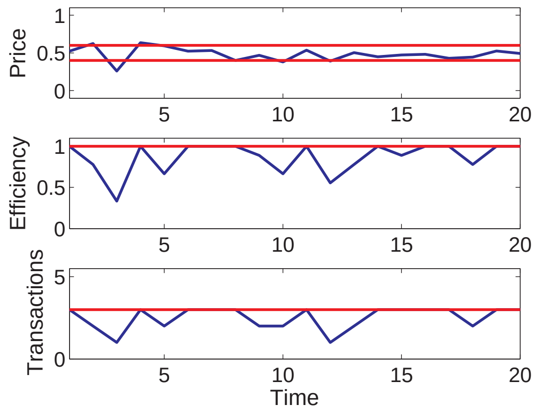
(c) Individual bids under CL. (d) Individual bids under OP.

(e) Individual asks under CL. (f) Individual asks under OP.

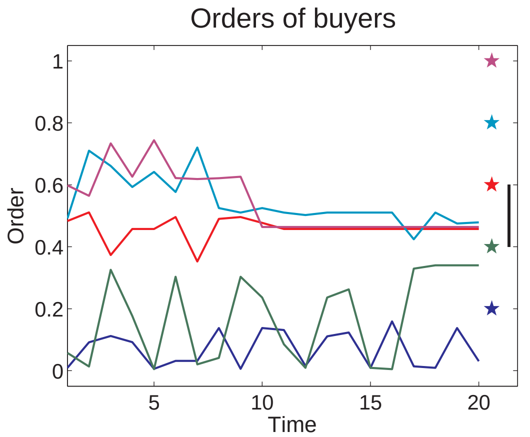
Figure 4.4: Learning phase dynamics in the GS-environment with 3 extramarginal buyers with valuation $\beta = 0.5$. Both in Open- and ClosedBook the intramarginal buyer learns to submit a bid above 0.5, such that the seller prefers to trade with this buyer. The seller increases his ask to make sure that he will trade with the intramarginal buyer and not with an extramarginal.



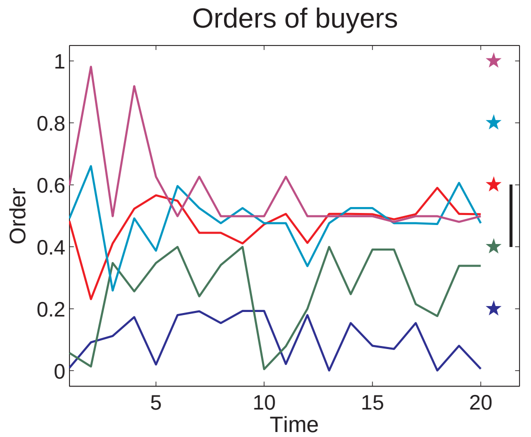
(a) Aggregate outcomes under CL.



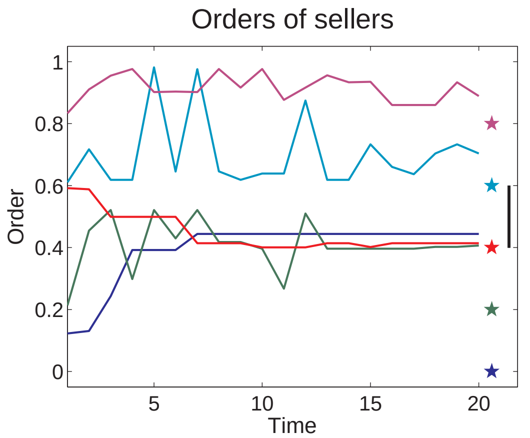
(b) Aggregate outcomes under OP.



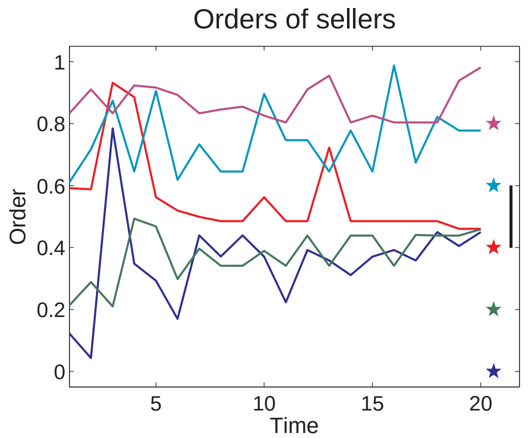
(c) Individual bids under CL.



(d) Individual bids under OP.



(e) Individual asks under CL.



(f) Individual asks under OP.

Figure 4.5: Learning phase dynamics in the symmetric S5-environment with 5 buyers and 5 sellers. In both settings a fast convergence occurs. After the first periods offers are relatively close to each other and these offers are less volatile.

selects more profitable offers, since the selection probability of a strategy is proportional to its hypothetical foregone payoff.

4.5.3 Comparison between Open- and ClosedBook

The IEL-algorithm is used to replicate the behaviour of agents. We simulated the different environments and averaged the efficiency, average price, price volatility and the number of transactions. The results of the simulations for Open- and ClosedBook are shown in Table 4.1 for the S5-environment and Table 4.5 in the appendix for the AL-environment, with the standard deviation between brackets. The t-values for comparing two means are given in Tables 4.2 and 4.6, the difference is significant at a level of 5% when the absolute t-value exceeds 1.96. These t-values are obtained by subtracting the OpenBook values from the ClosedBook values.

Of main interest is the question how the behaviour in the OpenBook system differs from the ClosedBook system. From the t-test for comparing two means, we can conclude that the efficiency is higher in the ClosedBook system, at a significance level of 5%. The average prices are similar in both systems and the price volatility is lower under ClosedBook. The number of transactions is higher in ClosedBook, but always lower than the equilibrium quantity. We show that this also holds for periods $1-5$ and $16-20$. We can now conclude that in the learning phase, the extra available information in the OpenBook setting leads to a higher price volatility and a lower efficiency and number of transactions. When under OpenBook some traders can increase their profit by submitting a more aggressive offer, they do not take into account that traders on the other side of the market do the same. This will often result in absence of trade when multiple traders are more aggressive and may lead to a higher price volatility when traders are matched with other tradingpartners. In the ClosedBook system traders will in such a case be less aggressive as the previous average price is used as a benchmark. Hence multiple traders may offer more aggressively, but this will less frequently result in absence of trade and a higher price volatility. This results in a lower expected efficiency and a higher price volatility in the OpenBook system.

	CL: closed book			OP: open book		
Period:	1-5	1-20	16-20	1-5	1-20	16-20
Efficiency	0.7771	0.9031	0.9562	0.7304	0.8433	0.8887
	(0.1109)	(0.0491)	(0.0562)	(0.0900)	(0.0477)	(0.0695)
Price	0.5001	0.5020	0.5027	0.4966	0.4930	0.4891
	(0.0881)	(0.0567)	(0.0488)	(0.0564)	(0.0414)	(0.0510)
Price Volat	0.0835	0.0546	0.0209	0.1173	0.0776	0.0320
	(0.0440)	(0.0246)	(0.0124)	(0.0503)	(0.0238)	(0.0170)
Num transact	2.3260	2.6860	2.8200	2.1680	2.5010	2.6120
	(0.3299)	(0.1536)	(0.1938)	(0.2737)	(0.1616)	(0.2324)

Table 4.1: Average outcomes during the learning phase in the symmetric S5-environment with 5 buyers and 5 sellers. We observe a clear learning effect, during the initial periods the efficiency and the number of transactions are relatively low but these increase fast during the learning phase.

	T-values		
Period:	1-5	1-20	16-20
Efficiency	3.27	8.74	7.55
Price	0.33	1.28	1.93
Price Volat	-5.06	-6.72	-5.28
Num transact	3.69	8.30	6.87

Table 4.2: T-values for testing the differences in average outcomes between ClosedBook and OpenBook during the learning phase in the symmetric S5-environment with 5 buyers and 5 sellers. The efficiency and number of transactions are significantly higher and the volatility significantly lower under ClosedBook at a 5% percent significance level. Average price is slightly higher under ClosedBook but not significantly.

4.5.4 Comparison with the Call Market

Arifovic and Ledyard (2007) find a higher efficiency in the ClosedBook system in various environments, a ClosedBook efficiency between 92% and 94% and an OpenBook efficiency between 77% and 90%. We used the same parameters in the simulations and, as in their paper, averaged over the first 20 periods of all environments. We find a ClosedBook efficiency between 86% and 90% and an OpenBook efficiency between 82% and 84%. The efficiency under ClosedBook is slightly higher in the Call Market and similar under OpenBook. We conclude that the differences in efficiency between Open- and ClosedBook are very similar for the Continuous Double Auction and the Call Market. Moreover, under ClosedBook the efficiency is higher in a Call Market than in a Continuous Double Auction.

4.6 Long-term behaviour

Arifovic and Ledyard (2007) do not consider the long-run in the Call Market, since the offers will converge to the redemption values in the ClosedBook system and mutation has little effect. The efficiency will almost always be equal to one. Arifovic and Ledyard (2007) find that under the OpenBook traders behave as pricemakers, and hence mutation does have a significant effect. Hence we argue that under OpenBook the efficiency is higher.

The long-term behaviour of agents in a Continuous Double Auction under the IEL-algorithm is studied by Anufriev et al. (2013). The foregone payoff function that they use is the same as in the Call Market model. The orders converge to the redemption values, as Anufriev et al. (2013) explain in their Result 1: ”The strategy profile under which the pool of every trader consists of messages equal to his own valuation/cost is attractive under the ClosedBook treatment in the GS-environment”.

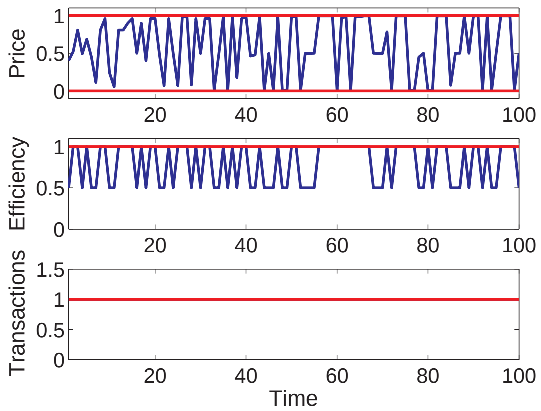
In this section we present our results under the alternative ClosedBook foregone payoff function. We use normally instead of uniformly distributed mutation, but simulations have shown that this does not affect our results. We compare Open- and Closed book simulations and the differences with the foregone payoff function from Anufriev et al. (2013). Under the foregone payoff function we introduced, also in the ClosedBook system orders will converge towards the equilibrium price range and this result does no longer hold.

4.6.1 GS-environment

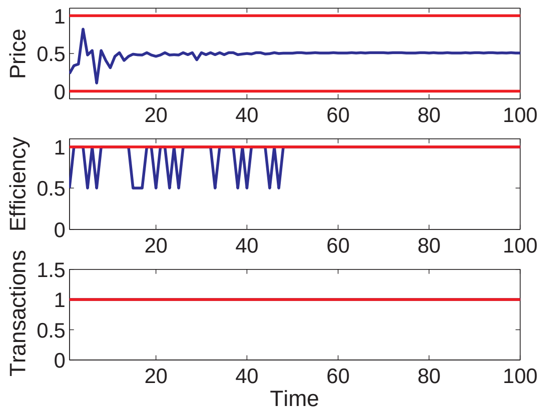
Figs. 4.6 and 4.7 show the impact of the payoff function in the ClosedBook setting, in the GS-environment with 3 extramarginal buyers with valuation $\beta = 0.5$. In Fig. 4.6 the hypothetical foregone payoff function of Anufriev et al. (2013) is used. The intramarginal buyer will submit a bid close to 1 and the seller will submit an ask close to 0. The seller will often trade with an extramarginal buyer, giving him a low payoff. The seller can increase his expected profit by asking a higher price, without a decrease in the probability of trading. In Fig. 4.7 the new hypothetical payoff functions is used, where the agents use more of the available information, resulting in both in Open- and ClosedBook in some convergence towards an equilibrium price between 0.5 and 1.

4.6.2 S5- and AL-environments

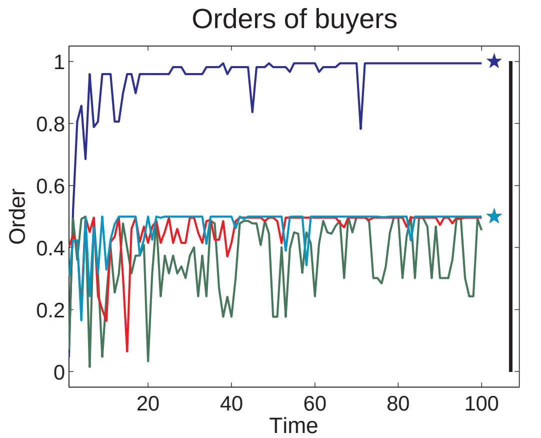
Now we have studied the long-term behaviour in the simple GS-environment, we can formally consider the other environments. Figs. 4.8 and 4.13 show the S5- and the AL-environment. In these realisations the efficiency and the number of transactions are clearly higher in the ClosedBook system. In equilibrium the offers in the ClosedBook fluctuate less, as a result of the mutual IEL of all the agents, and in particular of their evaluation of mutations and their consequences. In the OpenBook system traders observe the entire trading sequence of the previous period and base their next strategy on this. If an agent traded at the price of the trading partner, slightly higher bids or lower asks have an identical hypothetical foregone payoff. However, in the ClosedBook system, the hypothetical foregone payoff function gives a lower profit for these slightly higher bids and lower asks. Hence mutated strategies are more often selected in the OpenBook system, which may reduce efficiency in the next period when other traders condition on this mutated strategy. The arrival sequence influences the strategy of traders and



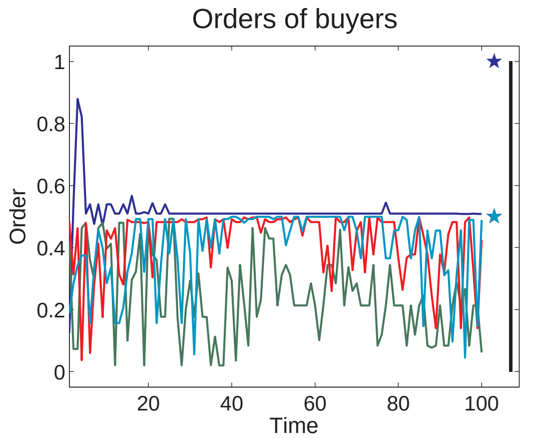
(a) Aggregate outcomes under CL.



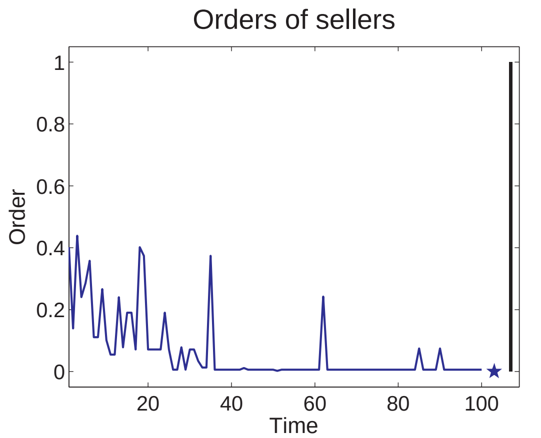
(b) Aggregate outcomes under OP.



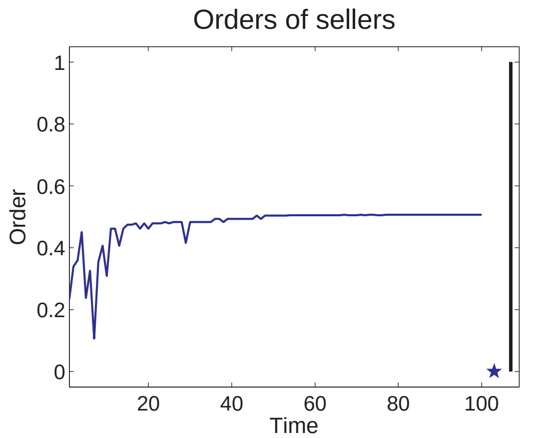
(c) Individual bids under CL.



(d) Individual bids under OP.

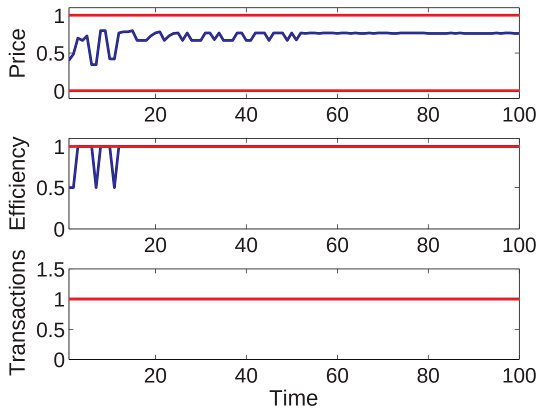


(e) Individual asks under CL.

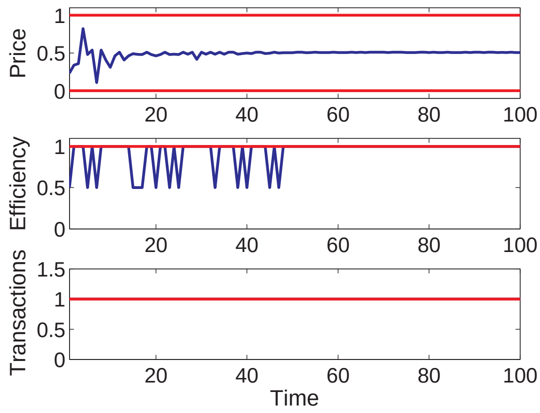


(f) Individual asks under OP.

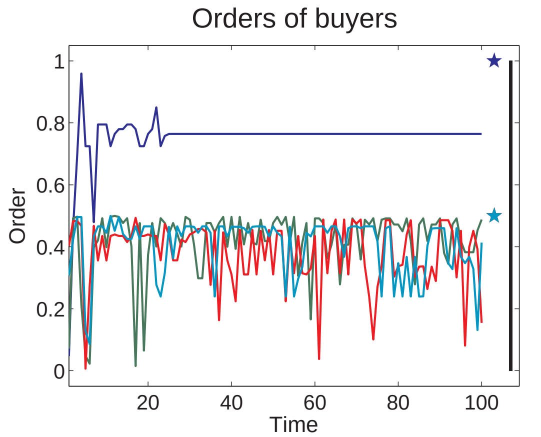
Figure 4.6: Long-term dynamics in the GS-environment with 3 extramarginal buyers with valuation $\beta = 0.5$ under the foregone payoff function of Anufriev et al. (2013). While in the OpenBook system the intramarginal buyer and seller coordinate on a price above 0.5, in the ClosedBook system the buyer bids close to 1 and the seller asks 0. Frequently an extramarginal buyer will trade which lowers the efficiency to 0.5. Hence the trading price is very volatile.



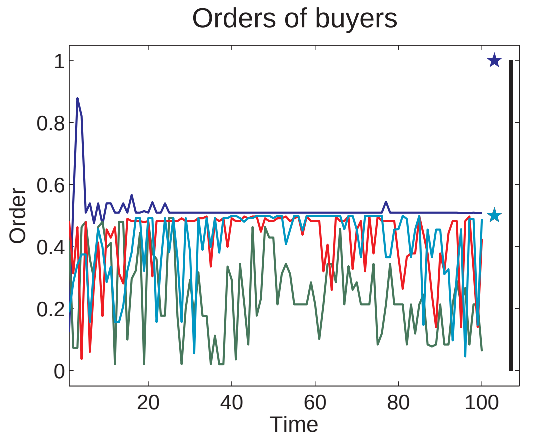
(a) Aggregate outcomes under CL.



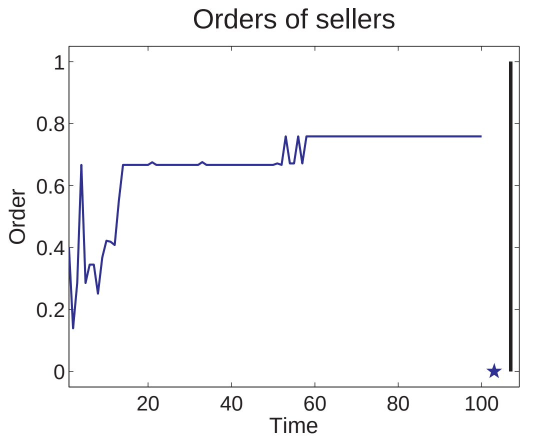
(b) Aggregate outcomes under OP.



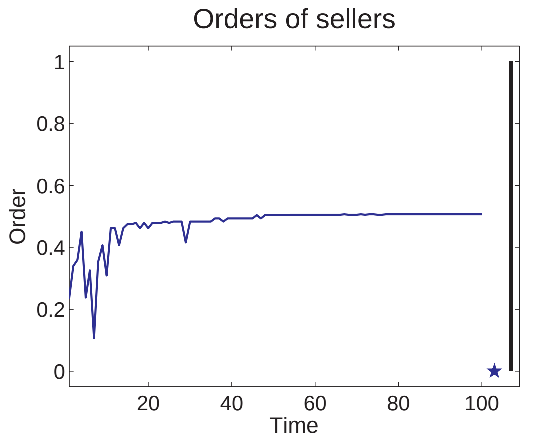
(c) Individual bids under CL.



(d) Individual bids under OP.

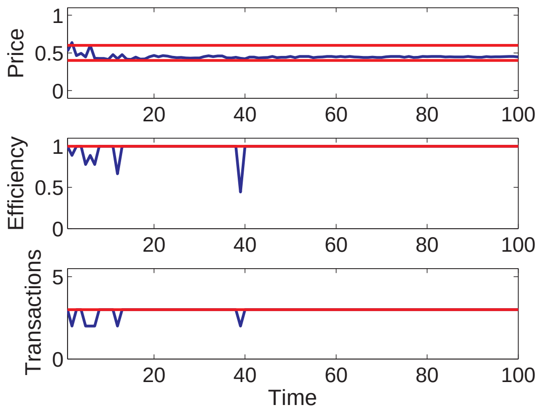


(e) Individual asks under CL.

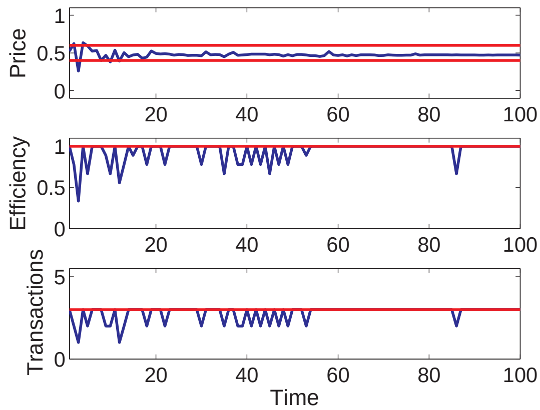


(f) Individual asks under OP.

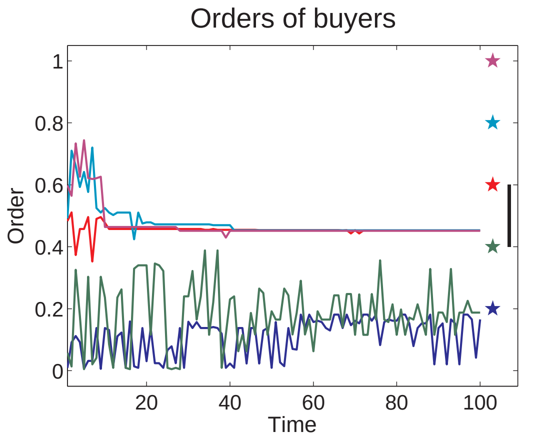
Figure 4.7: Long-term dynamics in the GS-environment with 3 extramarginal buyers with valuation $\beta = 0.5$ under the new foregone payoff function. The intramarginal buyer and seller coordinate on a price above 0.5, which ensures that they will trade with each other.



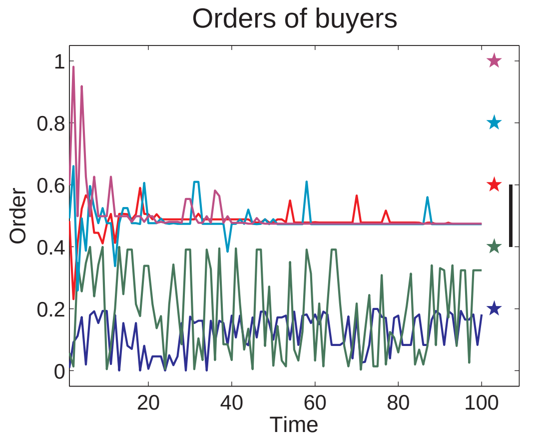
(a) Aggregate outcomes under CL.



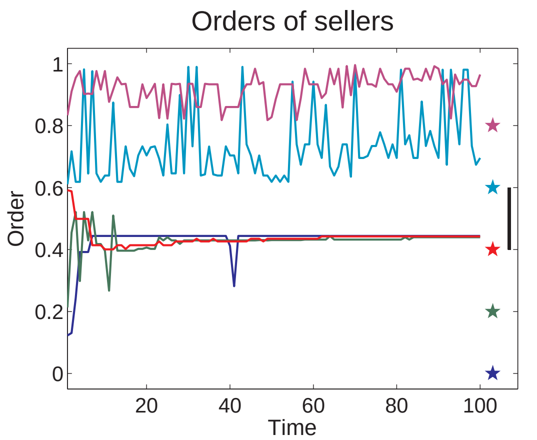
(b) Aggregate outcomes under OP.



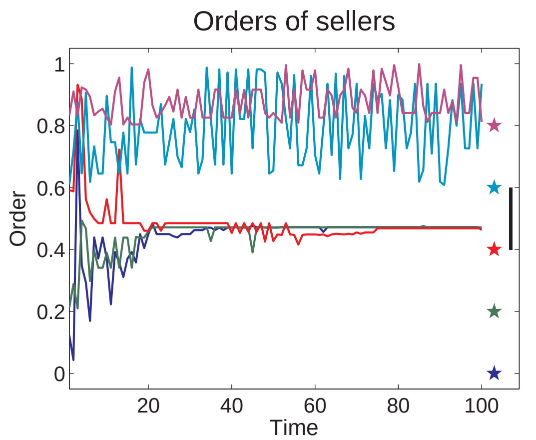
(c) Individual bids under CL.



(d) Individual bids under OP.



(e) Individual asks under CL.



(f) Individual asks under OP.

Figure 4.8: Long-term dynamics in the symmetric S5-environment with 5 buyers and 5 sellers. While under the OpenBook offers fluctuate frequently this occurs much less under ClosedBook. Hence the efficiency and the number of transactions are often lower under OpenBook.

they will occasionally offer relatively aggressive. This may lead to absence of trade if the arrival sequence in the next period is different.

4.6.3 Comparison between Closed- and OpenBook

We simulated the S5- and AL-environments and averaged the efficiency, average price, price volatility and the number of transactions, for different sizes K of the pools of strategies. In equilibrium we observe robustness with respect to the size of the pool of strategies. Anufriev et al. (2013) already showed robustness with regards to the probability of mutation; general statistics barely change as long as this probability is not too large. The results of the simulations for Open- and ClosedBook are shown in Tables 4.3 and 4.7 with the standard deviation between brackets. The t-values for comparing two means are given in Tables 4.4 and 4.8.

In all the environments the average efficiency is higher and the price volatility is lower in the ClosedBook setting. We do not observe a difference in average prices between Closed- and OpenBook. We observe a higher number of transactions for the ClosedBook setting in each environment, but this is in each case lower than the equilibrium number of transactions. All these results are significant at a level of 5%. Efficiency and number of transactions are higher when the size of the pool K increases, and the price volatility decreases. We conclude that in equilibrium, irrespective of the size of the pool of strategies, the extra available information in the OpenBook setting significantly leads to a higher price volatility and a lower efficiency and number of transactions. Hence we have shown robustness with respect to the size of the pool of strategies.

4.6.4 Comparison with the ClosedBook foregone payoff function in Anufriev et al. (2013).

The behaviour under the introduced ClosedBook hypothetical foregone payoff function is different than the behaviour in Anufriev et al. (2013). They find a divergence of offers; we find a convergence towards an equilibrium price. With the new hypothetical foregone payoff

	CL: closed book				OP: open book			
	$K = 10$	$K = 50$	$K = 100$	$K = 200$	$K = 10$	$K = 50$	$K = 100$	$K = 200$
Eff	0.9059	0.9799	0.9868	0.9918	0.8708	0.9505	0.9605	0.9607
	(0.0817)	(0.0230)	(0.0196)	(0.0144)	(0.0883)	(0.0345)	(0.0295)	(0.0362)
Price	0.5035	0.5042	0.5030	0.4986	0.4973	0.4993	0.4966	0.5031
	(0.0503)	(0.0405)	(0.0432)	(0.0442)	(0.0535)	(0.0412)	(0.0407)	(0.0426)
Vol	0.0227	0.0090	0.0066	0.0054	0.0252	0.0130	0.0122	0.0117
	(0.0114)	(0.0060)	(0.0050)	(0.0053)	(0.0120)	(0.0046)	(0.0045)	(0.0048)
Trans	2.6176	2.9362	2.9559	2.9723	2.5377	2.8461	2.8714	2.8862
	(0.3305)	(0.0724)	(0.0587)	(0.0435)	(0.3310)	(0.0989)	(0.0881)	(0.0977)

Table 4.3: Long-term average outcomes in the symmetric S5-environment with 5 buyers and 5 sellers. As the size K of the pool of strategies increases, so do the average efficiency and number of transactions, and the price volatility decreases.

	T-values			
	$K = 10$	$K = 50$	$K = 100$	$K = 200$
Eff	2.92	7.09	7.43	7.98
Price	0.84	0.85	1.08	-0.73
Vol	-1.51	-5.29	-8.32	-8.81
Trans	1.71	7.35	7.98	8.05

Table 4.4: T-values for testing the differences in long-term average outcomes between ClosedBook and OpenBook in the symmetric S5-environment with 5 buyers and 5 sellers. For every size K of the pool of strategies the average efficiency and number of transactions are significantly higher and the price volatility lower under ClosedBook. However, for a small size of the pool, $K = 10$, not all statistics are significant.

function, mutation has a smaller effect on the market under Closed Book. Hence, where they find a comparable efficiency and a higher price volatility in the ClosedBook system, under the new foregone payoff function the efficiency is higher and the price volatility lower under ClosedBook.

We argue that intramarginal traders are better off when they use the newly introduced hypothetical foregone payoff function. Efficiency and thus the total profit is significantly higher under the new payoff function. Moreover, we have shown that extramarginals trade less frequently, so that intramarginal traders receive a larger part of the total profit. Hence the intramarginal traders can increase their profit by using the new hypothetical foregone payoff function.

4.7 Multi-unit Continuous Double Auction market

So far we considered a market in which every buyer and every seller tries to trade a single unit of the good. In this section we extend this model by allowing agents to trade multiple units. Investors submit a strategy (b_i, n_i) respectively (a_j, n_j), which not only consists of a bid or an ask price, but also the number of units they desire to trade. The more units a buyer already obtained in the period the lower he values an extra unit, and similarly for sellers. In this symmetric environment the 10 valuations for a single buyer are given by $\{[1, .95, .89, .82, .74, .63, .53, .42, .3, .17]\}$, where the first value denotes the valuation for the first obtained unit and so on. The costs for a seller are symmetric to these valuations. Fig. 4.9a shows the decreasing valuation function and the increasing cost function for 5 identical buyers and sellers.

In this symmetric environment there exist only equilibria in which all investors trade 7 units. Submitting a strategy for more units can result in a loss on the additional units. Mutation of the number of units can occur, in which case one unit is added or subtracted to the strategy with equal probability. In the ClosedBook setting the payoff function needs to be adjusted to consider in how many trades a certain strategy will result. The adjusted hypothetical payoff is calculated by multiplying the payoffs of the regular payoff function by the estimated number of units that would be traded. For buyers this number of units is as follows:

If the buyer traded in the previous period and the strategy consists of fewer units than in the previous period $(n_{i,t+1} < n_{i,t})$:

$$\begin{cases} all & \text{if } b_i \geq P_t^{av} \\ \# \text{ of transaction prices in the last period} < b_i & \text{otherwise.} \end{cases}$$

If the buyer traded in the previous period and the strategy consists of more units than in the previous period $(n_{i,t+1} \geq n_{i,t})$:

$$\begin{cases} \text{all} & \text{if } b_i \geq \max(P_t^{av}, b_{b,t}) \\ \# \text{ of transactions in the previous period} & \text{if } \max(P_t^{av}, b_{b,t}) > b_i \geq \min(P_t^{av}, b_{b,t}) \\ \# \text{ of transaction prices in the last period} < b_i & \text{otherwise.} \end{cases}$$

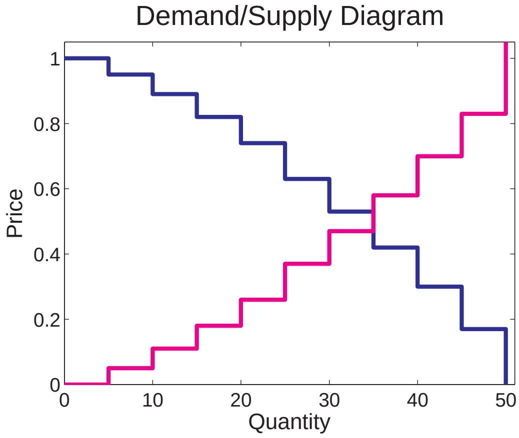
If the buyer did not trade in the previous period:

$$\begin{cases} \text{all} & \text{if } b_i > \max(P_t^{av}, b_{b,t}) \\ 0 & \text{otherwise.} \end{cases}$$

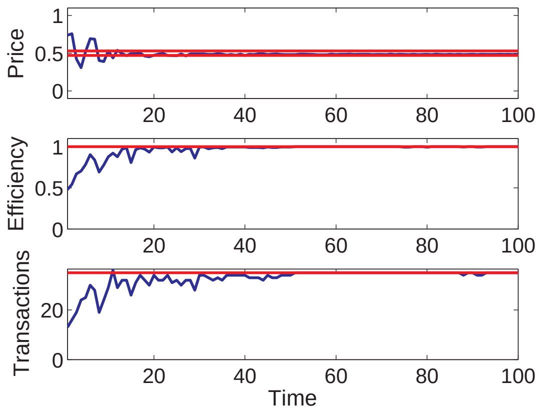
The symmetric environment in this extended market is shown in Figs. 4.9 and 4.10. Two-dimensional learning occurs as traders place an offer and a size of the offer. Both in Open- and ClosedBook these order move towards the equilibrium. We observe a robustness with respect to the number of units agents can trade. Mutation has a larger effect in the OpenBook system, resulting in a lower efficiency and number of trades and a higher price volatility.

A random environment where each investor trades 5 units in equilibrium is shown in Figs. 4.14 and 4.15 in the appendix. In this example, during early trading periods we observe a remarkable coordination on offers outside the equilibrium price, which is never observed before under IEL. Both in Open- and ClosedBook a disturbance after period 40 leads to coordination on offers closer to the equilibrium price. This may be the result of the relatively small equilibrium price range.

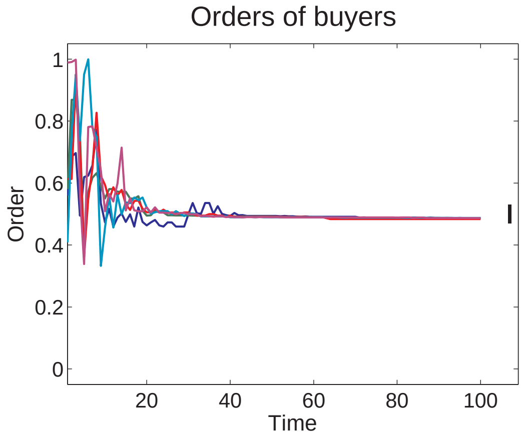
As shown in Tables 4.9-4.16 the comparisons between Open- and ClosedBook in the learning and in the equilibrium phase are slightly altered. The efficiency and number of transactions



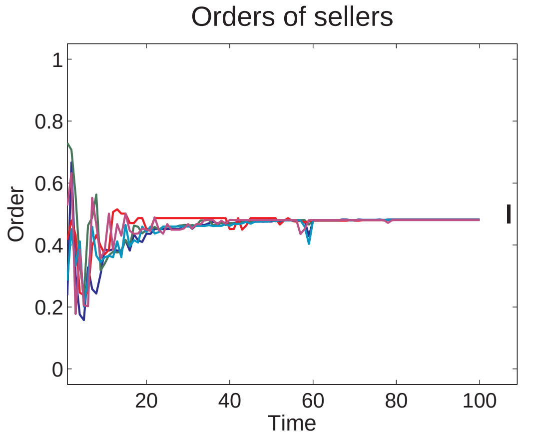
(a) Symmetric environment under CL.



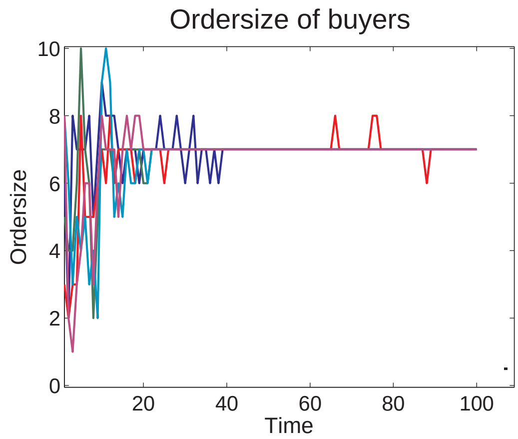
(b) Aggregate outcomes under CL.



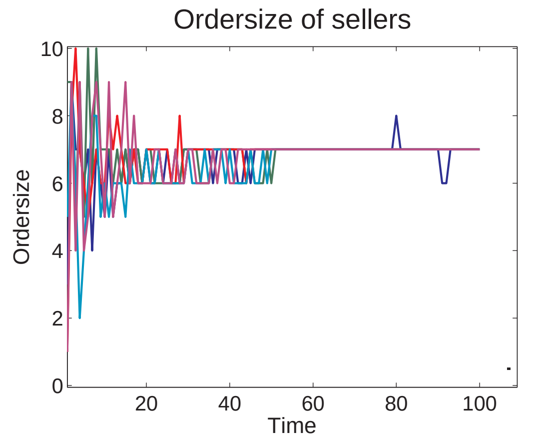
(c) Individual bids under CL.



(d) Individual asks under CL.

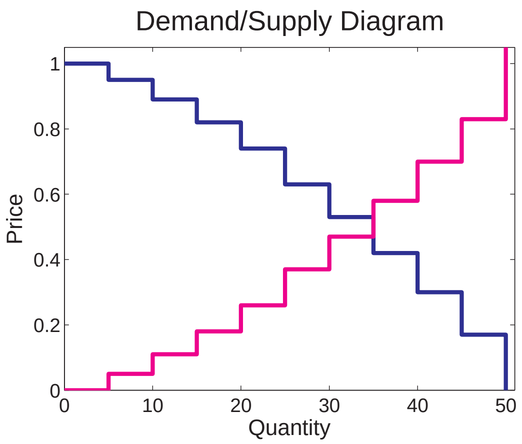


(e) Order size buyers under CL.

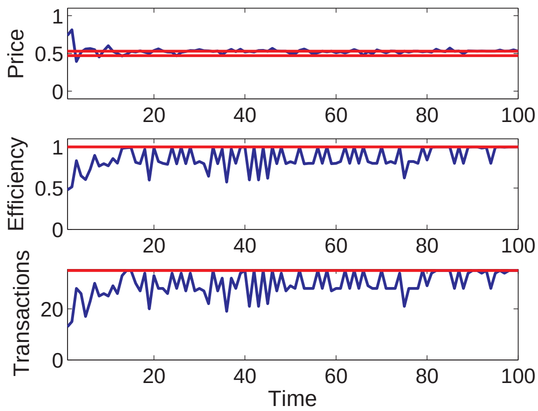


(f) Order size sellers under CL.

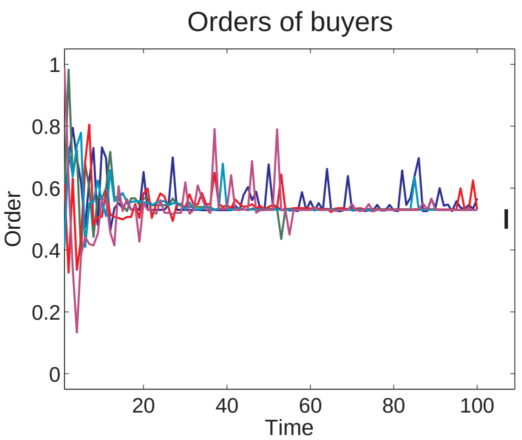
Figure 4.9: Long-term dynamics in the ClosedBook multi-unit symmetric environment with 5 buyers and 5 traders that can place an offer for a maximum of 10 units. The equilibrium offer is made for 7 units. Even though traders are required to make a two-dimensional decision, orders move towards the equilibrium. Mutation seems to have little effect after the learning phase.



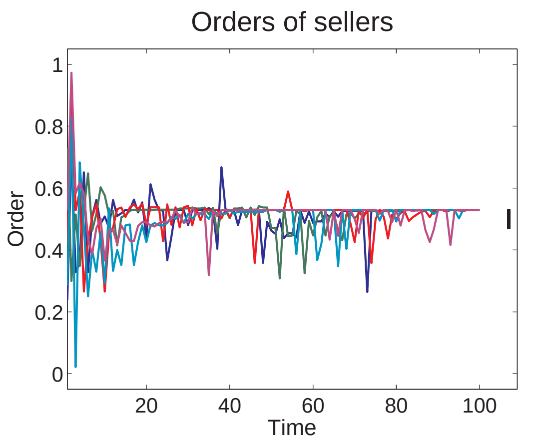
(a) Symmetric environment under OP.



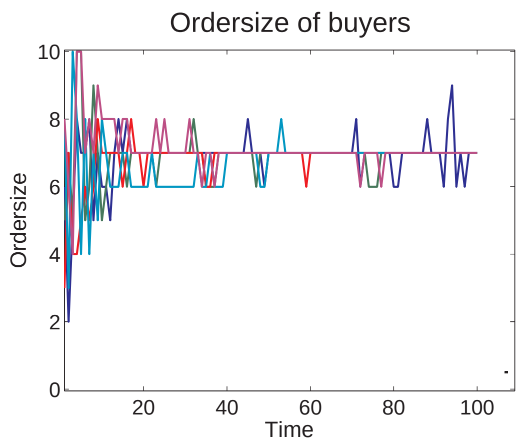
(b) Aggregate outcomes under OP.



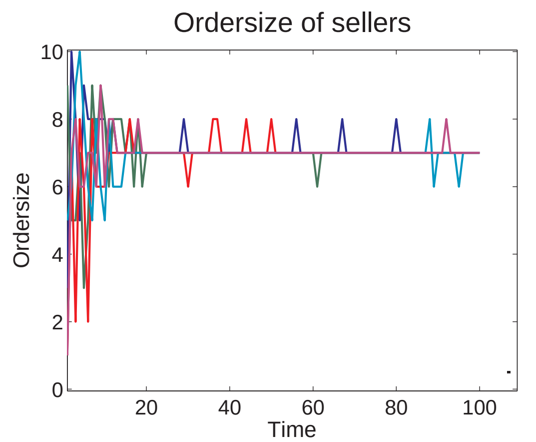
(c) Individual bids under OP.



(d) Individual asks under OP.



(e) Order size buyers under OP.



(f) Order size sellers under OP.

Figure 4.10: Long-term dynamics in the OpenBook multi-unit symmetric environment with 5 buyers and 5 sellers that can place an offer for a maximum of 10 units. The equilibrium offer is made for 7 units. Orders move towards the equilibrium quantity and price, but mutation has a larger effect than in the ClosedBook system. Efficiency and the number of transactions are often below the equilibrium value and the price volatility is higher.

remain higher under ClosedBook. However, the difference in price is now significantly larger under OpenBook. In the multi-unit symmetric environment the price volatility is significantly larger under OpenBook, but in the multi-unit random environment significantly lower. In the first environment offers converge to the equilibrium price range after the first 20 periods, but in the latter environment traders coordinate on offers outside the equilibrium price range as described above. The disturbance around period 40 increases the price volatility, which reverses the comparison between Open- and ClosedBook. Moreover, the average price is significantly larger under OpenBook, in the long run of this extended model. We find a robustness with respect to the number of units traded, but the IEL algorithm reacts slightly different when the equilibrium price range is relatively small.

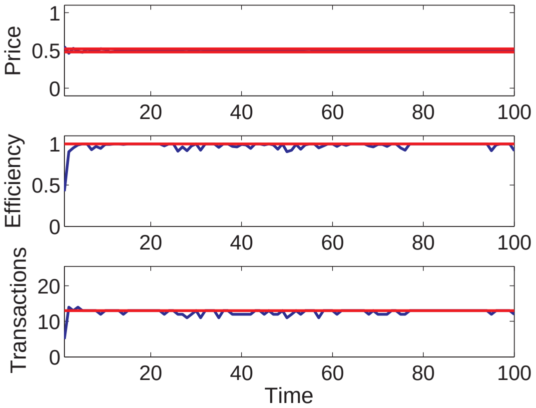
4.8 Size of the market

Robustness with respect to the size of the pool of strategies and to the number of units that each investor prefers to trade is shown in the previous sections. In this section we will consider robustness with respect to the number of investors in the market.

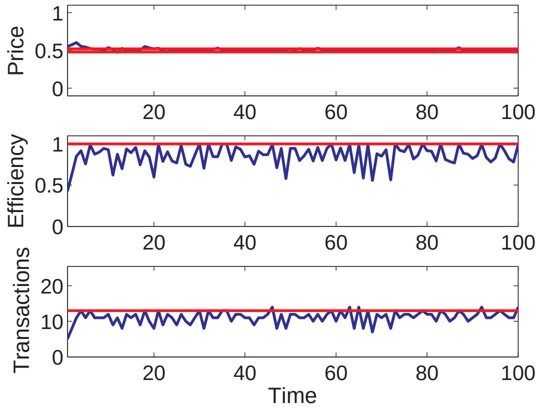
To study the robustness with respect to the size of the market we increased the number of investors in the S5- and AL-environments. This results in a setting with 15 buyers and 15 sellers shown in Fig. 4.16 of the appendix and a setting with 25 buyers and 25 sellers in Fig. 4.11. Larger markets would be more realistic, but computationally not feasible. Although the efficiency increases and volatility decreases, the comparisons between Open- and ClosedBook remain in Tables 4.17-4.24. Our results are robust to the size of the market, both during the learning and the equilibrium phase.

4.9 Concluding Remarks

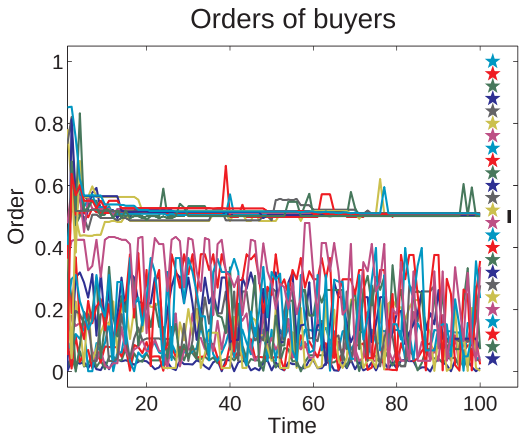
In this chapter we have studied the role of information about the trading history that is available to traders in a Continuous Double Auction market, when traders use the Individual Evolution-



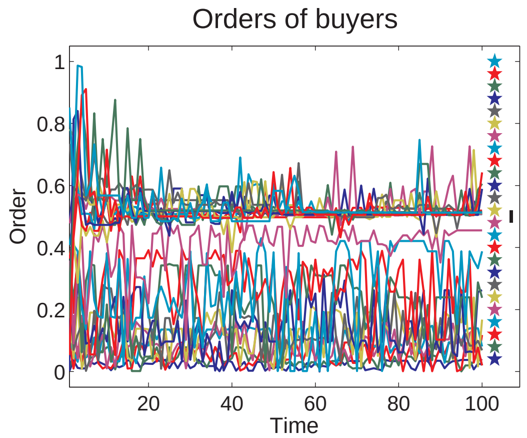
(a) Aggregate outcomes under CL.



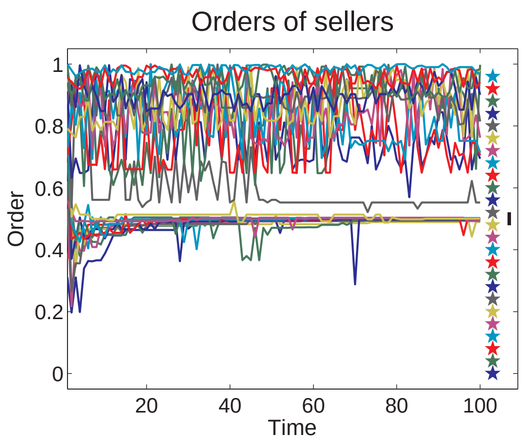
(b) Aggregate outcomes under OP.



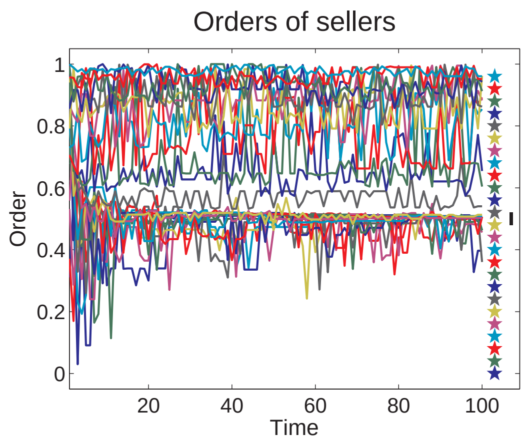
(c) Individual bids under CL.



(d) Individual bids under OP.



(e) Individual asks under CL.



(f) Individual asks under OP.

Figure 4.11: Long-term dynamics in the symmetric S25-environment with 25 buyers and 25 sellers. Both in the learning phase and long-term offers of intramarginal traders are less volatile under ClosedBook. Average efficiency and number of transactions are higher and price volatility lower under ClosedBook, showing a robustness with respect to the number of traders on both sides of the market.

ary Learning algorithm. In this learning algorithm traders select from a pool of strategies based on the, hypothetical, payoff in the previous period. In the ClosedBook system, where only information about past average prices is available, Anufriev et al. (2013) proved convergence of bids and offers towards the valuations and costs of agents. This is the consequence of their choice for a payoff function that only distinguishes between offers below and above the average price of the previous period, as in a Call Market. The consequence is that investors trade with a high probability but may generate a very small profit. We showed however, that when the payoff function uses more information to estimate the expected payoff of each possible offer, bids and offers tend to drift towards the equilibrium price range. In this approach investors learn to increase their expected profit by submitting an offer that has a higher possible profit. The probability of trading will be lower in this situation but is outweighed by the increase in possible profit. In this setting bids and asks will not diverge, but will converge towards some equilibrium price. These results are in line with Fano et al. (2013) who show that, in a setting closely related to the ClosedBook system, traders behave as pricetakers in a Call Market and as pricemakers in a Continuous Double Auction.

Both during the learning phase and in equilibrium we compared simulations of the Closed- and OpenBook treatments in different environments. In all the environments the efficiency and the number of transactions are significantly larger in the ClosedBook system. The number of transactions is in each case lower than the equilibrium number of transactions. In general, we did not observe a difference in average prices between Closed- and OpenBook. Moreover, we observed a significantly lower price volatility in the ClosedBook system. The cause is that agents in the OpenBook system are more tempted to try to influence their transaction price. We conclude that both during the learning phase and in equilibrium, the extra available information in the OpenBook treatment leads to a higher price volatility and a lower efficiency and number of transactions.

The efficiency found is comparable to efficiency in the Call Market from Arifovic and Ledyard (2007). The results differ however from the results under the payoff function of Anufriev et al. (2013), which results in a comparable efficiency and a higher price volatility in the ClosedBook system. The behaviour in the ClosedBook system is also quite different under the new payoff function; instead of a divergence of offers some convergence occurs.

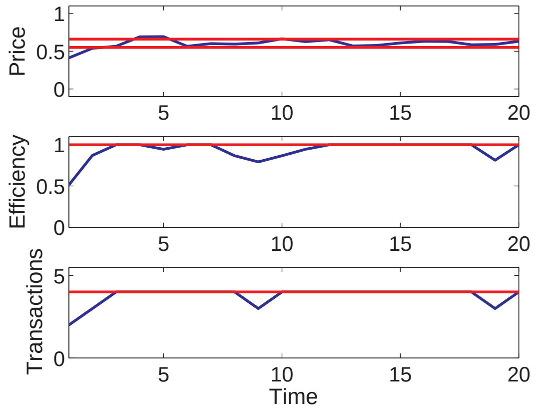
Some extensions to this Continuous Double Auction market are considered. We have shown that the above results are robust with respect to the number of units agents desire to trade and the size of the market. However, in the multi-unit random environment, price volatility is higher in the ClosedBook as a result of the relatively small equilibrium price range.

Both during the learning phase and in equilibrium, more information about the trading history leads to a higher price volatility and a lower efficiency and number of transactions. This is the result of the ClosedBook foregone hypothetical payoff function, that is introduced to estimate the expected profit in the previous period by using more of the available information in a Continuous Double Auction market. We conclude that it is optimal not to reveal the extra information of the OpenBook system.

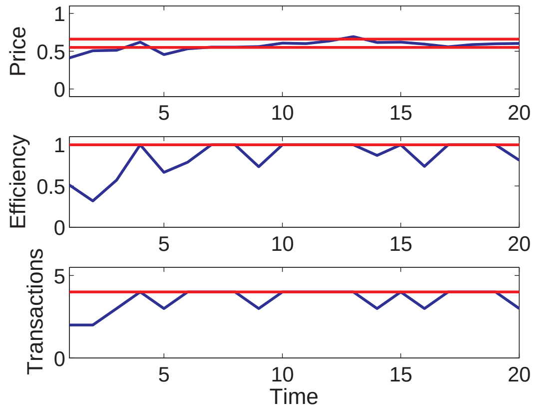
Appendix A: Learning phase

In this appendix we consider the learning phase during the periods $1-20$ for the AL-environment. This is done for the entire period, as well as the subperiods $1-5$ and $16-20$. An example is shown in Fig. 4.12 and averages in Table 4.5. Moreover, the t-values for testing the differences between ClosedBook and OpenBook are given in Table 4.6.

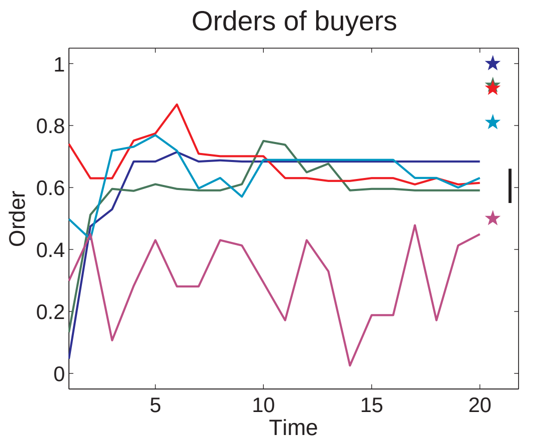
We observe a fast convergence of offers towards the equilibrium price range, both in the Open- and ClosedBook system. In the entire time span and the subperiods, the efficiency and number of transactions are significantly higher and the price volatility significantly lower under ClosedBook. The average price does not significantly differ. Traders learn over time, and hence the efficiency and number of transactions increase and the price volatility decreases during the periods $1-20$.



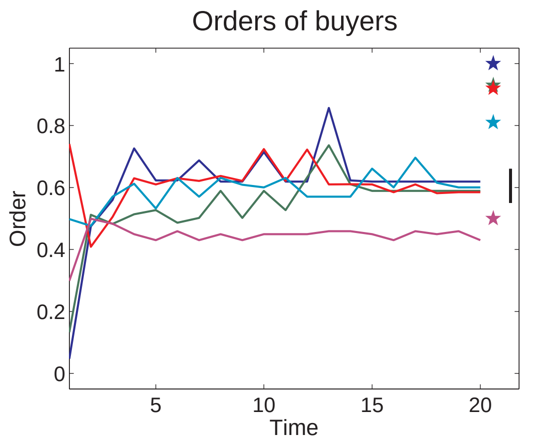
(a) Aggregate outcomes under CL.



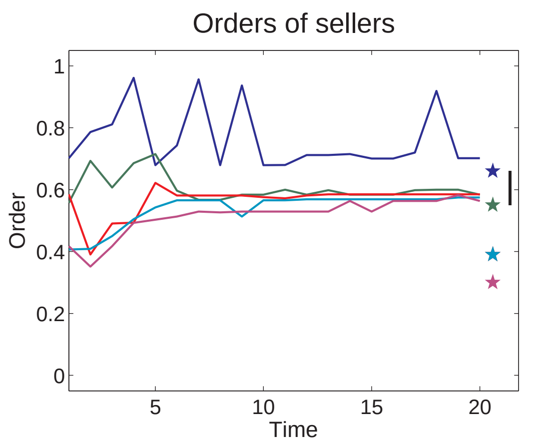
(b) Aggregate outcomes under OP.



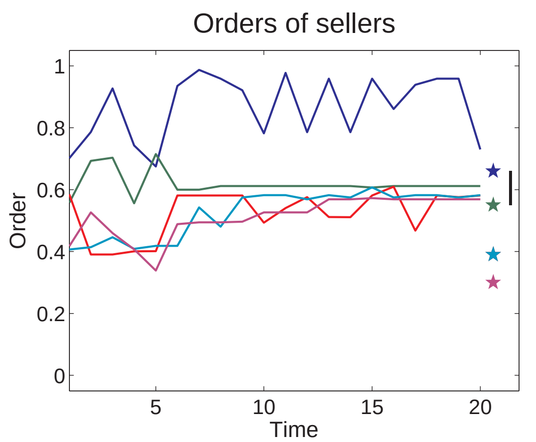
(c) Individual bids under CL.



(d) Individual bids under OP.



(e) Individual asks under CL.



(f) Individual asks under OP.

Figure 4.12: Learning phase dynamics in the AL-environment with 5 buyers and 5 sellers. Both in the Open- and ClosedBook system offers of traders move fast towards the equilibrium price range.

	CL: closed book			OP: open book		
Period:	1-5	1-20	16-20	1-5	1-20	16-20
Efficiency	0.7764	0.9059	0.9546	0.7198	0.8331	0.8805
	(0.1176)	(0.0386)	(0.0441)	(0.0719)	(0.0361)	(0.0608)
Price	0.6328	0.6316	0.6297	0.6396	0.6436	0.6459
	(0.0703)	(0.0490)	(0.0429)	(0.0488)	(0.0324)	(0.0394)
Price Volat	0.0648	0.0416	0.0163	0.0819	0.0552	0.0226
	(0.0313)	(0.0147)	(0.0078)	(0.0309)	(0.0140)	(0.0119)
Num transact	3.2000	3.6920	3.8600	2.9560	3.4330	3.6380
	(0.4680)	(0.1716)	(0.1853)	(0.2907)	(0.1460)	(0.2662)

Table 4.5: Average outcomes during the learning phase in the AL-environment with 5 buyers and 5 sellers. The efficiency and number of transactions are higher and the price volatility is lower under ClosedBook. A difference in average price is not observed. A learning effect occurs and efficiency and number of transactions increase over time and price volatility decreases.

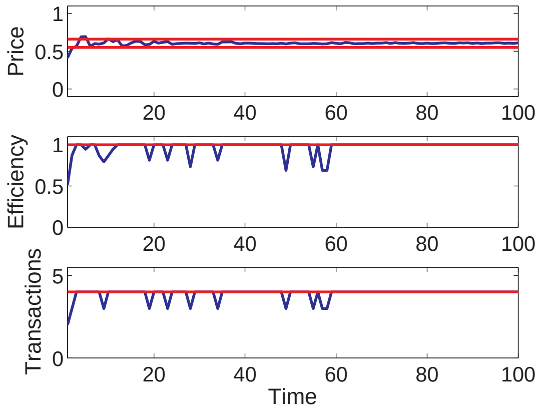
	T-values		
Period:	1-5	1-20	16-20
Efficiency	4.11	13.77	9.87
Price	-0.79	-2.04	-2.78
Price Volat	-3.89	-6.70	-4.43
Num transact	4.43	11.50	6.84

Table 4.6: T-values for testing the differences in average outcomes between ClosedBook and OpenBook during the learning phase in the AL-environment with 5 buyers and 5 sellers. The efficiency and number of transactions are significantly higher and the price volatility is significantly lower under ClosedBook. A difference in average price is not observed.

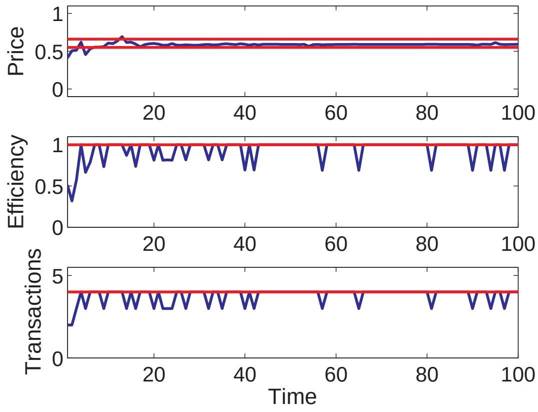
Appendix B: Equilibrium phase

This appendix considers the long term behaviour of traders during the periods $101-200$ for the AL-environment. This is done for different sizes K of the pool of strategies. Fig. 4.13 shows an example of the behaviour, with the averages given in Table 4.7. Moreover, the t-values for testing the differences between ClosedBook and OpenBook are given in Table 4.8.

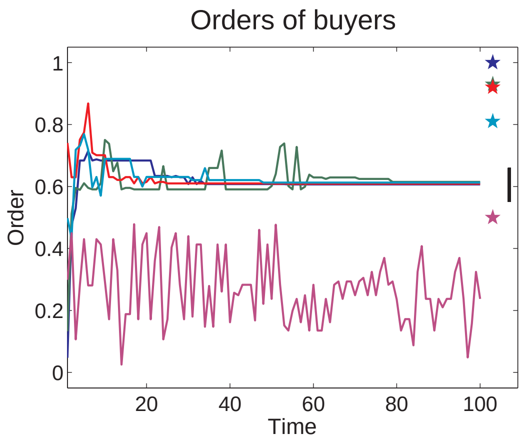
In the example we observe a smaller effect of mutation under ClosedBook. Hence the equilibrium efficiency and number of trades are more often reached. The efficiency and number of transactions are significantly higher and the price volatility significantly lower in the ClosedBook system, irrespective of the size K of the pool of strategies. We do not observe a clear significant difference in average price.



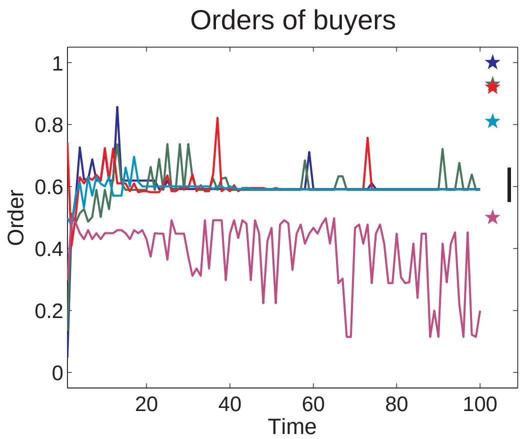
(a) Aggregate outcomes under CL.



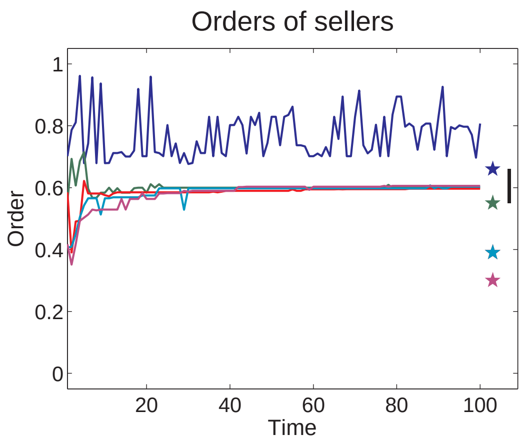
(b) Aggregate outcomes under OP.



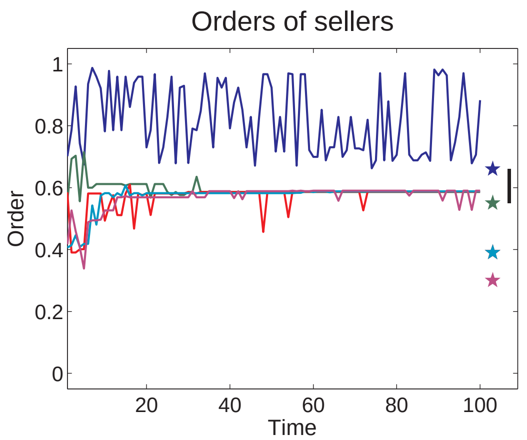
(c) Individual bids under CL.



(d) Individual bids under OP.



(e) Individual asks under CL.



(f) Individual asks under OP.

Figure 4.13: Long-term dynamics in the AL-environment with 5 buyers and 5 sellers. In both systems offers move fast towards the equilibrium price range. Under ClosedBook mutation seems to have a smaller effect and efficiency and the number of transactions more frequently attain the equilibrium value.

	CL: closed book				OP: open book			
	$K = 10$	$K = 50$	$K = 100$	$K = 200$	$K = 10$	$K = 50$	$K = 100$	$K = 200$
Eff	0.9197	0.9745	0.9806	0.9885	0.8646	0.9229	0.9259	0.9267
	(0.0584)	(0.0219)	(0.0210)	(0.0151)	(0.0647)	(0.0319)	(0.0334)	(0.0347)
Price	0.6287	0.6326	0.6239	0.6184	0.6364	0.6320	0.6359	0.6353
	(0.0383)	(0.0263)	(0.0289)	(0.0274)	(0.0423)	(0.0287)	(0.0262)	(0.0267)
Vol	0.0167	0.0089	0.0077	0.0063	0.0220	0.0144	0.0132	0.0134
	(0.0074)	(0.0050)	(0.0055)	(0.0064)	(0.0078)	(0.0044)	(0.0038)	(0.0039)
Trans	3.6898	3.9213	3.9359	3.9562	3.5087	3.7459	3.7785	3.7698
	(0.3175)	(0.0659)	(0.0649)	(0.0583)	(0.2936)	(0.1162)	(0.1120)	(0.1245)

Table 4.7: Long-term average outcomes in the AL-environment with 5 buyers and 5 sellers. The efficiency and number of transactions are higher and the price volatility is lower under ClosedBook. A clear difference in average price is not observed. These results are robust with respect to the size K of the pool of strategies.

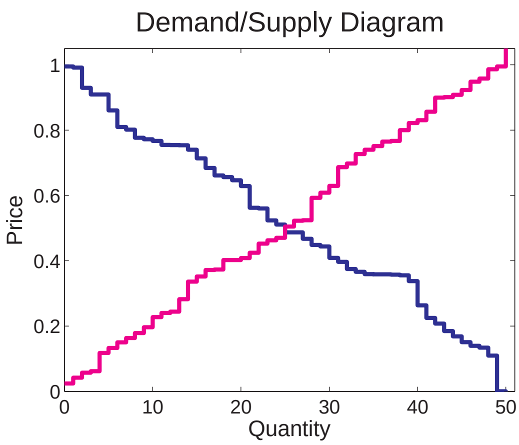
	T-values			
	$K = 10$	$K = 50$	$K = 100$	$K = 200$
Eff	6.32	13.34	13.86	16.33
Price	-1.35	0.15	-3.08	-4.42
Vol	-4.93	-8.26	-8.23	-9.47
Trans	4.19	13.13	12.16	13.56

Table 4.8: T-values for testing the differences in long-term average outcomes between ClosedBook and OpenBook in the AL-environment with 5 buyers and 5 sellers. The efficiency and number of transactions are significantly higher and the price volatility is significantly lower under ClosedBook. A clear significant difference in average price is not observed.

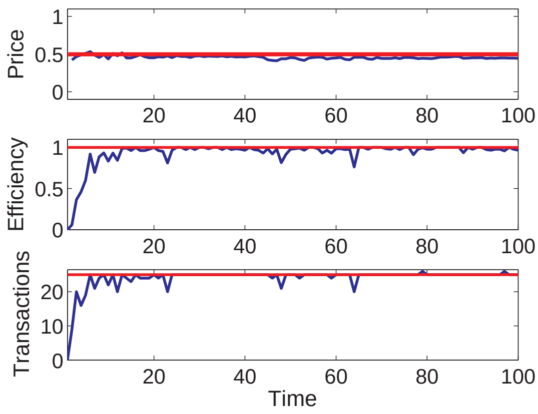
Appendix C: Multi-unit market

In this appendix we consider an extension of the model by allowing agents to trade multiple units. In the equilibrium of the random environment traders place an order for 5 units, shown in Figs. 4.14 and 4.15. The learning phase of both the symmetric and random environments is studied over the periods $1 - 20$ and the subperiods $1 - 5$ and $16 - 20$. We present the average outcomes and the t-values for testing the differences between Closed- and OpenBook in Tables 4.9-4.12 for both the symmetric and random environment. Average outcomes and t-values for the long-term, during periods $101 - 200$ and for different sizes K of the pool of strategies, are shown in Tables 4.13-4.16.

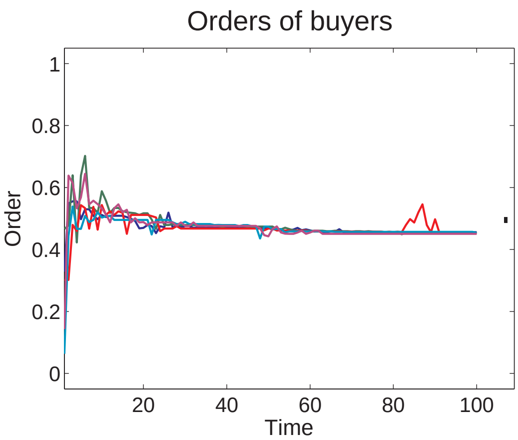
In the examples of the multi-unit random environment we observe that mutation plays a larger role under OpenBook, as full efficiency is often not obtained. During the learning phase the efficiency and number of transactions are higher under ClosedBook. The average price and price volatility do not show a significant difference. In the long-rum efficiency and number of transactions remain higher under ClosedBook and the average price becomes significantly lower. The comparison between price volatility is different between both environments. In the symmetric environment the price volatility is significantly lower under ClosedBook, and in the random environment significantly higher. The latter is the effect of the coordination of offers outside the small equilibrium price range in the ClosedBook system.



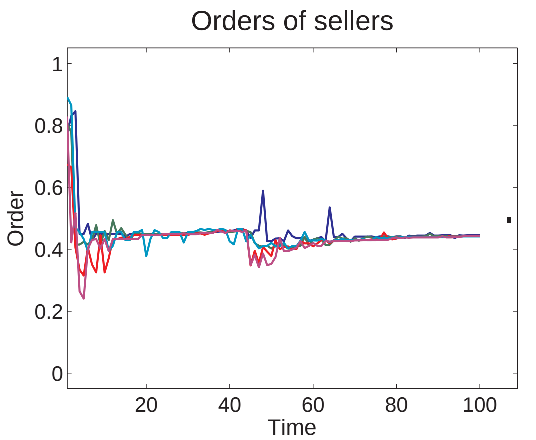
(a) Symmetric environment under CL.



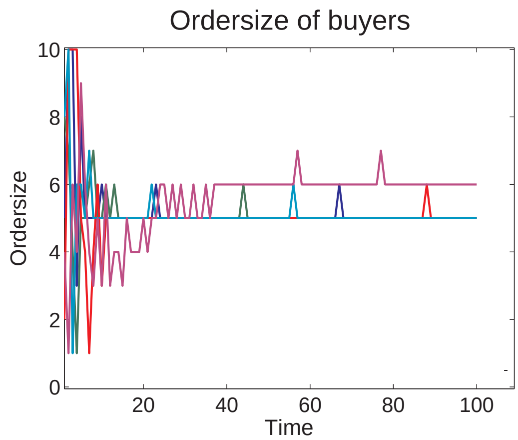
(b) Aggregate outcomes under CL.



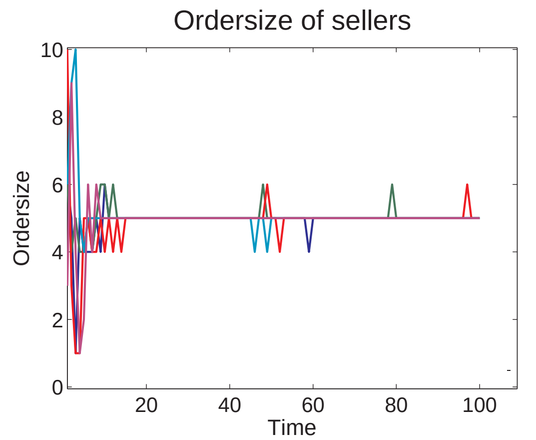
(c) Individual bids under CL.



(d) Individual asks under CL.

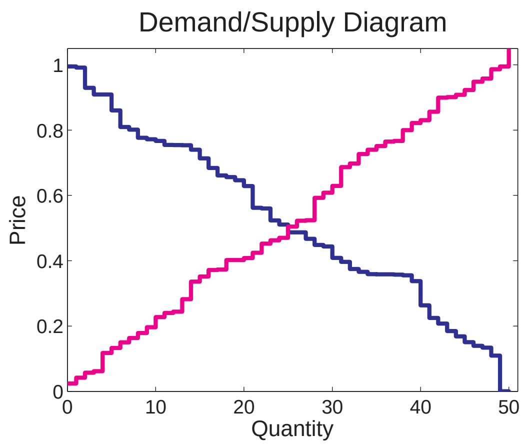


(e) Order size buyers under CL.

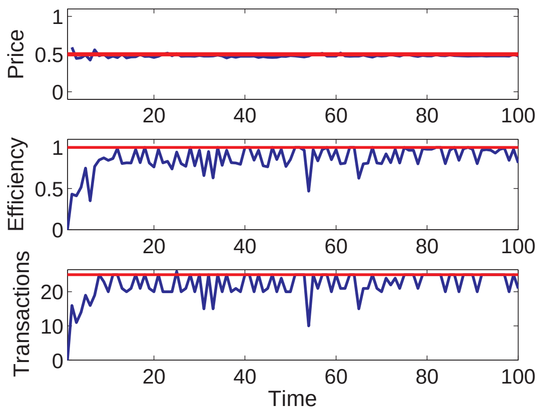


(f) Order size sellers under CL.

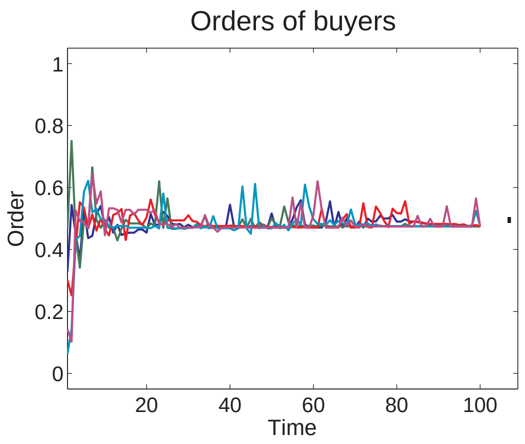
Figure 4.14: Long-term dynamics in the ClosedBook multi-unit random environment with 5 buyers and 5 sellers that can place an offer for a maximum of 10 units. The equilibrium offer is made for 5 units. Both size and offer converge. Traders coordinate their offers outside the relatively small equilibrium price range. This coordination is disturbed around period 40.



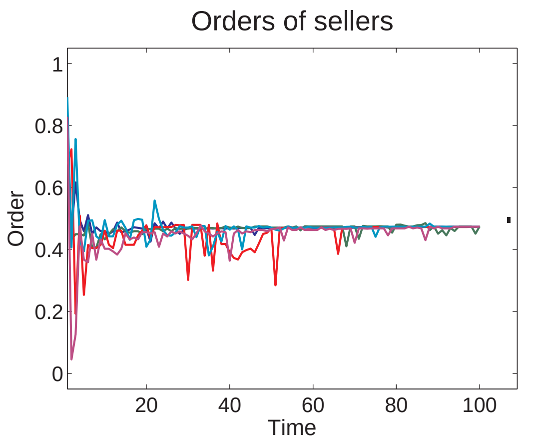
(a) Random environment under OP.



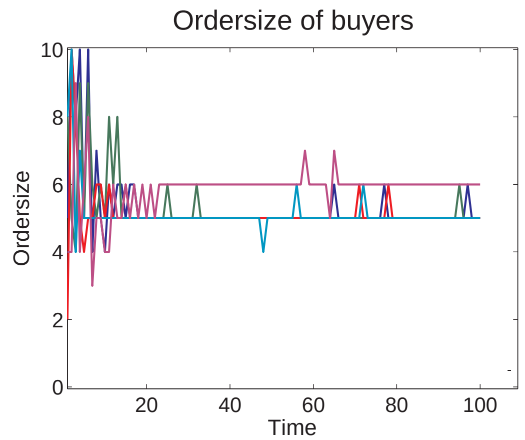
(b) Aggregate outcomes under OP.



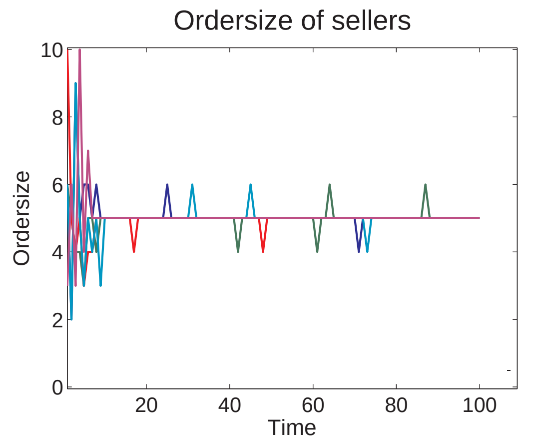
(c) Individual bids under OP.



(d) Individual asks under OP.



(e) Order size buyers under OP.



(f) Order size sellers under OP.

Figure 4.15: Long-term dynamics in the OpenBook multi-unit random environment with 5 buyers and 5 sellers that can place an offer for a maximum of 10 units. The equilibrium offer is made for 5 units. Both size and offer converge to the equilibrium value. Mutation seems to play a larger role than under ClosedBook, as full efficiency is less frequently obtained.

	CL: closed book			OP: open book		
Period:	1-5	1-20	16-20	1-5	1-20	16-20
Efficiency	0.6895	0.8657	0.9516	0.6423	0.8109	0.8878
	(0.0654)	(0.0322)	(0.0336)	(0.0604)	(0.0311)	(0.0525)
Price	0.4895	0.4863	0.4852	0.5015	0.5054	0.5078
	(0.0804)	(0.0435)	(0.0380)	(0.0623)	(0.0403)	(0.0417)
Price Volat	0.1278	0.0808	0.0273	0.1355	0.0786	0.0210
	(0.0511)	(0.0249)	(0.0121)	(0.0535)	(0.0223)	(0.0107)
Num transact	21.8240	27.8625	30.8560	21.4420	27.3770	30.1460
	(2.6538)	(1.3104)	(1.4896)	(2.5747)	(1.2164)	(2.0175)

Table 4.9: Average outcomes during the learning phase in the multi-unit symmetric environment with 5 buyers and 5 sellers that can place an offer for a maximum of 10 units. The efficiency and number of transactions are higher under ClosedBook. A clear difference in average price and price volatility is not observed. A learning effect occurs and efficiency and number of transactions increase over time and price volatility decreases.

	T-values		
Period:	1-5	1-20	16-20
Efficiency	5.30	12.24	10.24
Price	-1.18	-3.22	-4.01
Price Volat	-1.04	0.66	3.90
Num transact	1.03	2.72	2.83

Table 4.10: T-values for testing the differences in average outcomes between ClosedBook and OpenBook during the learning phase in the multi-unit symmetric environment with 5 buyers and 5 sellers that can place an offer for a maximum of 10 units. The efficiency and number of transactions are significantly higher under ClosedBook. A clear significant difference in average price and price volatility is not observed.

	CL: closed book			OP: open book		
Period:	1-5	1-20	16-20	1-5	1-20	16-20
Efficiency	0.5395	0.8201	0.9614	0.4850	0.7593	0.8889
	(0.0883)	(0.0444)	(0.0413)	(0.0675)	(0.0352)	(0.0544)
Price	0.4951	0.4862	0.4822	0.5094	0.5078	0.5101
	(0.0559)	(0.0368)	(0.0361)	(0.0512)	(0.0332)	(0.0336)
Price Volat	0.0849	0.0559	0.0236	0.0839	0.0543	0.0205
	(0.0391)	(0.0174)	(0.0089)	(0.0360)	(0.0157)	(0.0098)
Num transact	16.5240	21.5830	24.0320	16.1740	20.4600	22.3260
	(2.6932)	(1.1873)	(1.0307)	(2.3771)	(0.9495)	(1.3516)

Table 4.11: Average outcomes during the learning phase in the multi-unit random environment with 5 buyers and 5 sellers that can place an offer for a maximum of 10 units. The efficiency and number of transactions are higher under ClosedBook. A clear difference in average price and price volatility is not observed. A learning effect occurs and efficiency and number of transactions increase over time and price volatility decreases.

	T-values		
Period:	1-5	1-20	16-20
Efficiency	4.90	10.73	10.61
Price	-1.89	-4.36	-5.66
Price Volat	0.19	0.68	2.34
Num transact	0.97	7.39	10.04

Table 4.12: T-values for testing the differences in average outcomes between ClosedBook and OpenBook during the learning phase in the multi-unit random environment with 5 buyers and 5 sellers that can place an offer for a maximum of 10 units. The efficiency and number of transactions are significantly higher and average price significantly lower under ClosedBook. A clear difference in price volatility is not observed.

	CL: closed book				OP: open book			
	$K = 10$	$K = 50$	$K = 100$	$K = 200$	$K = 10$	$K = 50$	$K = 100$	$K = 200$
Eff	0.9749	0.9781	0.9783	0.9783	0.9388	0.9372	0.9442	0.9399
	(0.0244)	(0.0254)	(0.0263)	(0.0254)	(0.0219)	(0.0260)	(0.0219)	(0.0199)
Price	0.4721	0.4764	0.4795	0.4766	0.5038	0.5040	0.5040	0.4991
	(0.0171)	(0.0195)	(0.0232)	(0.0189)	(0.0307)	(0.0376)	(0.0324)	(0.0312)
Vol	0.0077	0.0055	0.0054	0.0055	0.0120	0.0110	0.0102	0.0106
	(0.0042)	(0.0044)	(0.0055)	(0.0049)	(0.0037)	(0.0036)	(0.0039)	(0.0038)
Trans	32.7328	33.0609	33.1029	33.0880	31.9787	31.7267	32.0048	31.8543
	(2.0657)	(2.2211)	(2.2171)	(2.2109)	(1.5151)	(1.6777)	(1.7313)	(1.6213)

Table 4.13: Long-term average outcomes in the multi-unit symmetric environment with 5 buyers and 5 sellers that can place an offer for a maximum of 10 units. The efficiency and number of transactions are higher and the price volatility lower under ClosedBook. However, also the average price is lower under ClosedBook. These results are robust with respect to the size K of the pool of strategies.

	T-values			
	$K = 10$	$K = 50$	$K = 100$	$K = 200$
Eff	11.01	11.25	9.96	11.90
Price	-9.02	-6.52	-6.15	-6.17
Vol	-7.68	-9.67	-7.12	-8.22
Trans	2.94	4.79	3.90	4.50

Table 4.14: T-values for testing the differences in long-term average outcomes between ClosedBook and OpenBook in the multi-unit symmetric environment with 5 buyers and 5 sellers that can place an offer for a maximum of 10 units. The efficiency and number of transactions are significantly higher and the price volatility significantly lower under ClosedBook. However, also the average price is significantly lower under ClosedBook.

	CL: closed book				OP: open book			
	$K = 10$	$K = 50$	$K = 100$	$K = 200$	$K = 10$	$K = 50$	$K = 100$	$K = 200$
Eff	0.9524	0.9470	0.9517	0.9517	0.9306	0.9391	0.9410	0.9426
	(0.0267)	(0.0310)	(0.0294)	(0.0250)	(0.0251)	(0.0286)	(0.0267)	(0.0271)
Price	0.4502	0.4463	0.4516	0.4517	0.4989	0.5014	0.4978	0.5028
	(0.0354)	(0.0321)	(0.0348)	(0.0359)	(0.0302)	(0.0279)	(0.0270)	(0.0268)
Vol	0.0200	0.0182	0.0181	0.0175	0.0103	0.0089	0.0092	0.0091
	(0.0076)	(0.0082)	(0.0074)	(0.0082)	(0.0027)	(0.0028)	(0.0031)	(0.0028)
Trans	23.8183	23.6129	23.7963	23.6853	23.2905	23.4898	23.5340	23.5920
	(0.8207)	(0.9692)	(0.8932)	(0.9102)	(0.7143)	(0.7415)	(0.7692)	(0.7354)

Table 4.15: Long-term average outcomes in the multi-unit random environment with 5 buyers and 5 sellers that can place an offer for a maximum of 10 units. The efficiency, number of transactions and price volatility are higher and the average price lower under ClosedBook. Price volatility is higher under ClosedBook due to the disturbances between subsequent coordinations of offers outside the equilibrium price. These results are robust with respect to the size K of the pool of strategies.

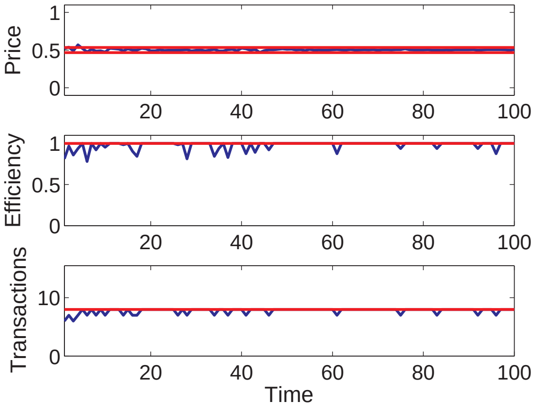
	T-values			
	$K = 10$	$K = 50$	$K = 100$	$K = 200$
Eff	6.44	1.87	2.69	2.47
Price	-10.47	-12.96	-10.49	-11.41
Vol	12.03	10.73	11.09	9.69
Trans	4.85	1.01	2.23	0.80

Table 4.16: T-values for testing the differences in long-term average outcomes between ClosedBook and OpenBook in the multi-unit random environment with 5 buyers and 5 sellers that can place an offer for a maximum of 10 units. The efficiency, number of transactions and price volatility are significantly higher and the average price significantly lower under ClosedBook.

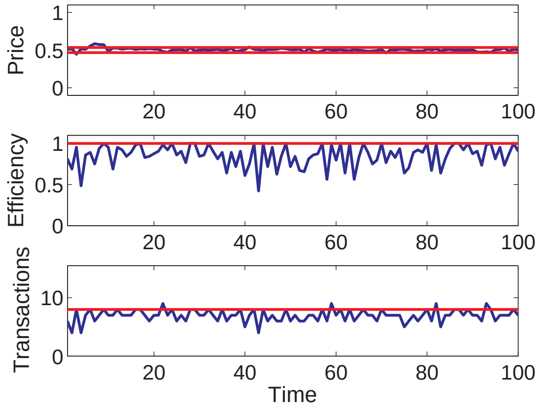
Appendix D: Size of the market

This appendix considers larger markets that consist of 15 buyers and 15 sellers or 25 buyers and 25 sellers. An example of the S15-environment is shown in Fig. 4.16. Both for the learning phase and the long-term we present average outcomes and the t-values for testing the differences between ClosedBook and OpenBook. These are given in Tables 4.17-4.24 for both the S15- and the S25-environment. The learning phase is studied over the periods $1-20$ and the subperiods $1-5$ and $16-20$ and the long-term during periods $101-200$ for different sizes K of the pool of strategies.

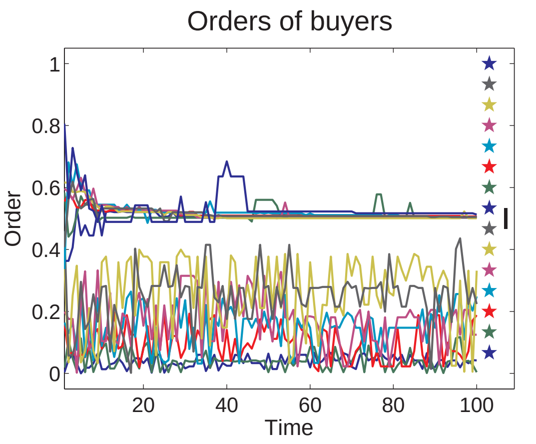
In the example under ClosedBook full efficiency is often attained, but only rarely in the OpenBook system. Hence we can conclude that the results shown in the S5-environment are robust with respect to the size of the market. The efficiency and the number of transactions are significantly higher under ClosedBook, and the price volatility significantly lower. The average price does not differ significantly between Open- and ClosedBook.



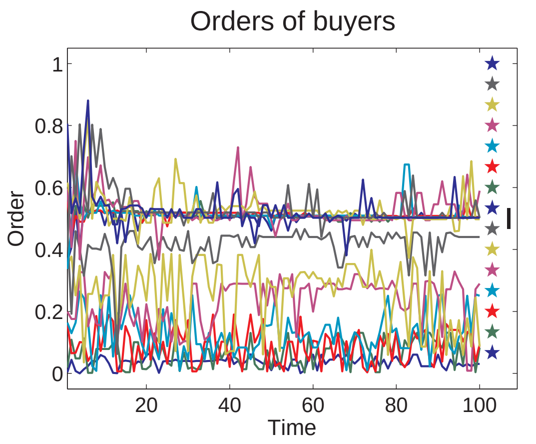
(a) Aggregate outcomes under CL.



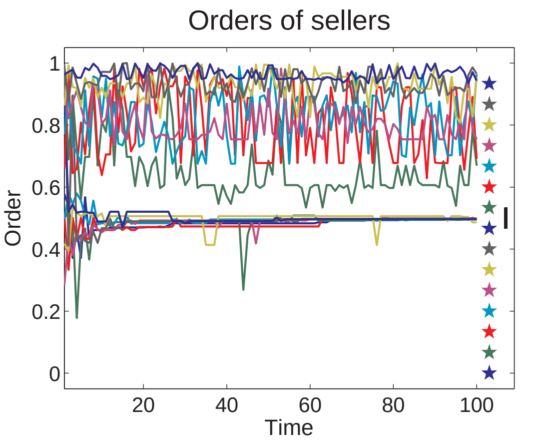
(b) Aggregate outcomes under OP.



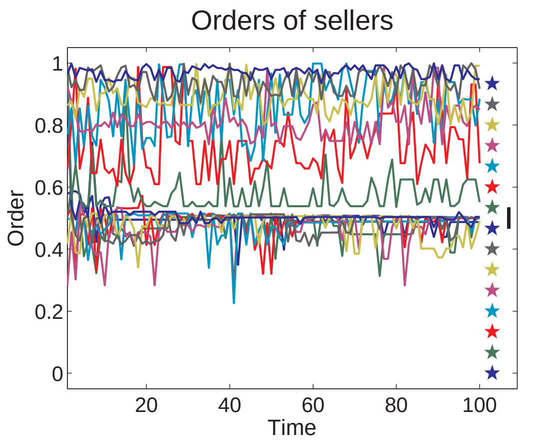
(c) Individual bids under CL.



(d) Individual bids under OP.



(e) Individual asks under CL.



(f) Individual asks under OP.

Figure 4.16: Long-term dynamics in the symmetric S15-environment with 15 buyers and 15 sellers. With more traders in the market offers again move towards the equilibrium price range. Where under ClosedBook mutation incidentally leads to a decreased efficiency, this occurs under OpenBook in most periods.

	CL: closed book			OP: open book		
Period:	1-5	1-20	16-20	1-5	1-20	16-20
Efficiency	0.8246	0.9258	0.9664	0.7514	0.8401	0.8744
	(0.0415)	(0.0207)	(0.0288)	(0.0490)	(0.0231)	(0.0464)
Price	0.5048	0.5026	0.5008	0.5015	0.5003	0.4995
	(0.0379)	(0.0219)	(0.0176)	(0.0314)	(0.0193)	(0.0230)
Price Volat	0.0549	0.0336	0.0125	0.0574	0.0406	0.0199
	(0.0197)	(0.0094)	(0.0056)	(0.0242)	(0.0106)	(0.0093)
Num transact	6.7700	7.3650	7.5920	6.1060	6.7280	6.9000
	(0.4552)	(0.2222)	(0.2953)	(0.4278)	(0.2015)	(0.4045)

Table 4.17: Average outcomes during the learning phase in the symmetric S15-environment with 15 buyers and 15 sellers. The efficiency and number of transactions are higher and price volatility lower under ClosedBook. A difference in average price is not observed. A learning effect occurs and efficiency and number of transactions increase over time and price volatility decreases.

	T-values		
Period:	1-5	1-20	16-20
Efficiency	11.40	27.63	16.85
Price	0.67	0.79	0.45
Price Volat	-0.80	-4.94	-6.82
Num transact	10.63	21.24	13.82

Table 4.18: T-values for testing the differences in average outcomes between ClosedBook and OpenBook during the learning phase in the S15-environment with 15 buyers and 15 sellers. The efficiency and number of transactions are significantly higher and price volatility significantly lower under ClosedBook. A significant difference in average price is not observed.

	CL: closed book			OP: open book		
Period:	1-5	1-20	16-20	1-5	1-20	16-20
Efficiency	0.8417	0.9356	0.9725	0.7640	0.8387	0.8652
	(0.0362)	(0.0172)	(0.0192)	(0.0398)	(0.0211)	(0.0406)
Price	0.5032	0.5028	0.5031	0.5017	0.5007	0.4992
	(0.0290)	(0.0158)	(0.0116)	(0.0274)	(0.0148)	(0.0155)
Price Volat	0.0433	0.0263	0.0098	0.0395	0.0293	0.0142
	(0.0169)	(0.0080)	(0.0040)	(0.0185)	(0.0075)	(0.0061)
Num transact	11.3120	12.1510	12.4600	10.0800	10.9110	11.1380
	(0.5638)	(0.2832)	(0.3510)	(0.5895)	(0.2804)	(0.5208)

Table 4.19: Average outcomes during the learning phase in the symmetric S25-environment with 25 buyers and 25 sellers. The efficiency and number of transactions are higher and price volatility lower under ClosedBook. A difference in average price is not observed. A learning effect occurs and efficiency and number of transactions increase over time and price volatility decreases.

	T-values		
Period:	1-5	1-20	16-20
Efficiency	14.44	35.60	23.89
Price	0.38	0.97	2.01
Price Volat	1.52	-2.74	-6.03
Num transact	15.10	31.11	21.05

Table 4.20: T-values for testing the differences in average outcomes between ClosedBook and OpenBook during the learning phase in the S25-environment with 25 buyers and 25 sellers. The efficiency and number of transactions are significantly higher and price volatility significantly lower under ClosedBook. A significant difference in average price is not observed.

	CL: closed book				OP: open book			
	$K = 10$	$K = 50$	$K = 100$	$K = 200$	$K = 10$	$K = 50$	$K = 100$	$K = 200$
Eff	0.9340	0.9791	0.9827	0.9844	0.8636	0.9103	0.9146	0.9196
	(0.0316)	(0.0111)	(0.0117)	(0.0125)	(0.0337)	(0.0215)	(0.0195)	(0.0245)
Price	0.4996	0.5010	0.5007	0.5010	0.5015	0.5038	0.5014	0.5012
	(0.0240)	(0.0129)	(0.0143)	(0.0154)	(0.0225)	(0.0148)	(0.0137)	(0.0164)
Vol	0.0112	0.0052	0.0052	0.0047	0.0181	0.0107	0.0104	0.0100
	(0.0038)	(0.0018)	(0.0022)	(0.0024)	(0.0043)	(0.0025)	(0.0019)	(0.0018)
Trans	7.1764	7.7946	7.8386	7.8599	6.6505	7.2267	7.3048	7.3640
	(0.3919)	(0.1082)	(0.0989)	(0.1054)	(0.3692)	(0.1677)	(0.1502)	(0.1624)

Table 4.21: Long-term average outcomes in the S15-environment with 15 buyers and 15 sellers. The efficiency and number of transactions are higher and price volatility lower under ClosedBook. A difference in average price is not observed. These results are robust with respect to the size K of the pool of strategies.

	T-values			
	$K = 10$	$K = 50$	$K = 100$	$K = 200$
Eff	15.24	28.69	29.95	23.56
Price	-0.58	-1.43	-0.35	-0.09
Vol	-12.02	-17.85	-17.89	-17.67
Trans	9.77	28.46	29.68	25.61

Table 4.22: T-values for testing the differences in long-term average outcomes between ClosedBook and OpenBook in the S15-environment with 15 buyers and 15 sellers. The efficiency and number of transactions are significantly higher and price volatility significantly lower under ClosedBook. A significant difference in average price is not observed.

	CL: closed book				OP: open book			
	$K = 10$	$K = 50$	$K = 100$	$K = 200$	$K = 10$	$K = 50$	$K = 100$	$K = 200$
Eff	0.9456	0.9800	0.9859	0.9864	0.8647	0.9046	0.9057	0.9052
	(0.0214)	(0.0086)	(0.0079)	(0.0080)	(0.0305)	(0.0170)	(0.0148)	(0.0174)
Price	0.5021	0.5030	0.5022	0.5011	0.5010	0.5016	0.5002	0.5002
	(0.0186)	(0.0096)	(0.0085)	(0.0094)	(0.0174)	(0.0083)	(0.0080)	(0.0088)
Vol	0.0083	0.0043	0.0037	0.0035	0.0145	0.0085	0.0085	0.0085
	(0.0020)	(0.0011)	(0.0014)	(0.0013)	(0.0033)	(0.0015)	(0.0013)	(0.0013)
Trans	11.7892	12.6831	12.7884	12.7947	10.9008	11.7102	11.7794	11.8101
	(0.4615)	(0.1521)	(0.1274)	(0.1190)	(0.4903)	(0.2301)	(0.1869)	(0.2145)

Table 4.23: Long-term average outcomes in the S25-environment with 25 buyers and 25 sellers. The efficiency and number of transactions are higher and price volatility lower under Closed-Book. A difference in average price is not observed. These results are robust with respect to the size K of the pool of strategies.

	T-values			
	$K = 10$	$K = 50$	$K = 100$	$K = 200$
Eff	21.71	39.58	47.81	42.40
Price	0.43	1.10	1.71	0.70
Vol	-16.07	-22.58	-25.12	-27.20
Trans	13.19	35.27	44.61	40.14

Table 4.24: T-values for testing the differences in long-term average outcomes between Closed-Book and OpenBook in the S25-environment with 25 buyers and 25 sellers. The efficiency and number of transactions are significantly higher and price volatility significantly lower under ClosedBook. A significant difference in average price is not observed.

Chapter 5

Timing under Individual Evolutionary Learning in a Continuous Double Auction

5.1 Introduction

In many agent-based models of order-driven financial markets traders submit their order at a random moment during a trading period and are required to make a one-dimensional decision; to choose a bid or ask price as in LiCalzi and Pellizzari (2006) or to forecast a future price as in Chiarella and Iori (2002). However, in a Continuous Double Auction (CDA) the moment of order submission plays a crucial role; submitting at the end of the period will yield a lower probability of trading, submitting at the beginning of the period will most likely result in a trade at the own submitted price which yields a lower profit. Allowing traders to submit their order at their preferred moment may influence these effects as traders may decide to condition their offer on the moment of submission. In agent-based models learning is often used to avoid making extreme assumptions about the rationality of traders and to select between multiple equilibria. With non-random timing learning of agents becomes multidimensional; not only learning about the offer but also learning about the timing of submission is of great importance.

In this chapter we introduce learning about the timing of order submission in an agent-based model. Traders also learn about the offer that they submit, and hence we extend the Individual Evolutionary Learning (IEL) algorithm used in Arifovic and Ledyard (2003, 2007) and in Anufriev et al. (2013) to a multidimensional version and allow for contemporaneous learning about the moment of submission and about the submitted orders. In the IEL algorithm traders select from a pool of possible strategies. After a trading period the hypothetical payoff is calculated for every possible strategy and some strategies are replaced with randomly modified strategies. Adopting the IEL algorithm to incorporate the decision about timing, we study the distribution of preferred submission moments, the interrelation between these moments and the submitted orders, and also the impact of the size of the market on the timing of submission and the offers of traders. In simulations we find that the distribution of the submission moments highly depends on the size of the market.

Starting from early contributions it is common that investors in agent-based models make a decision about the price of the order, but not about its timing. It is typically assumed that they submit their one unit orders at a random moment in the trading period and that between periods their learning is only one-dimensional: buyers learn which bid to submit and sellers learn which price to ask. Sometimes agents directly learn bids and asks, sometimes their bids and asks depend on the expectations and learning is over the space of prediction rules. For examples of the former approach, see LiCalzi and Pellizzari (2006, 2007) who compare efficiency in the CDA with other market protocols such as the call market, under boundedly rational resp. zero intelligent agents; and Bottazzi et al. (2005) who focus on the properties of price time series under different trading protocols. In the market protocols with sequential trade these papers assume that agents arrive in a random sequence. For examples of the latter approach see Chiarella and Iori (2002) who study properties of asset pricing under Continuous Double Auctions and other mechanisms in a model with heterogeneous expectations, Yamamoto and LeBaron (2010) who study the number of order splits and Anufriev and Panchenko (2009) who study the switching between forecasting rules. Again, in all simulations under Continuous Double Auctions, agents submit orders in a random sequence.

Arifovic and Ledyard (2003, 2007) introduced the Individual Evolutionary Learning algorithm to model the boundedly rational learning behaviour of agents in a Call Market model. Anufriev et al. (2013) use IEL in a Continuous Double Auction and compare efficiency under full and no information about the history of orders. Furthermore the latter paper studies the GS-environment from Gode and Sunder (1993, 1997) in the case where traders have zero intelligence and submit every possible offer with equal probability. In Chapter 4 we have shown that the results of Anufriev et al. (2013) depend on the hypothetical foregone payoff function that is chosen, under no information about the order history. This chapter extends the model in Anufriev et al. (2013) by considering multidimensional learning in which traders also learn about the moment of order submission.

An important feature of IEL is that it is essentially a backward-looking learning process. This approach contrasts with the standard economic approach where optimising agents make their decision and use all information rationally. In Friedman (1991) traders can submit orders at any moment of time and can also improve their outstanding orders. Traders regard other's orders as random, update their beliefs about the order distribution using Bayes' formula and submit orders on the basis of their updated distribution. The classical financial literature contains many studies on limit and market orders. Rosu (2009) and Parlour and Seppi (2008) provide a survey about the theoretical research on limit and market orders under random arrival of traders. The surveys Gould et al. (2013b) and Hachmeister (2007) discuss the main theoretical, experimental and empirical papers on limit orders of informed and uninformed traders. Bloomfield et al. (2005) performed an experiment on the choice between limit and market orders by informed and uninformed traders over time. As the period advances, uninformed traders use more market orders and informed traders more limit orders. Bae et al. (2003) and Chung et al. (1999) empirically consider the number of limit and market orders during a trading day and their relation with spread, order size and price volatility. Biais et al. (1995) determine the empirical distribution of large and small trades, orders and cancelations during a trading day. These papers find a U-shaped distribution of orders during a day. They explain this finding as motivated by the desire of traders to perform price discovery in the beginning of the day and react to events during the closing of the exchange. At the end of the day traders desire to unwind their positions. In this chapter there is no modelling of the news process; information is equal for all traders and does not evolve during periods. Rather we are interested in how traders with given valuation and cost are able to find their trading moment and strategy during the period.

The distribution of submission moments is studied in a benchmark environment under full information about trading history. We find that under the IEL-algorithm investors in a medium size market learn to submit their order around the middle of the trading period to avoid a lower trading probability or lower profit. Moreover, we observe an increasing bid function and decreasing ask function over time, similar to Fano and Pellizzari (2011). We show that the size of the market and competition between traders influence the distribution or submission moments. Furthermore of interest are the placed offer, efficiency, profit and the probability of trading as a function of the submission moment. General market statistics are compared with the setup in Anufriev et al. (2013) where traders submit orders at random moments.

The organisation of this chapter is as follows. The model and the trading mechanism are described in Section 5.2, followed by the extended Individual Evolutionary Learning algorithm in Section 5.3 and the methodology used. The distribution of the preferred moment of submitting and its relation to the bid and ask are described in a benchmark environment in Section 5.4. The impact of the size of the market is considered in Section 5.5. In Section 5.6 we study submission moments and their relation with offers as the amount of competition changes. The Gode-Sunder environment is attractive for its simplicity and studied in Section 5.7. Finally, Section 5.8 concludes.

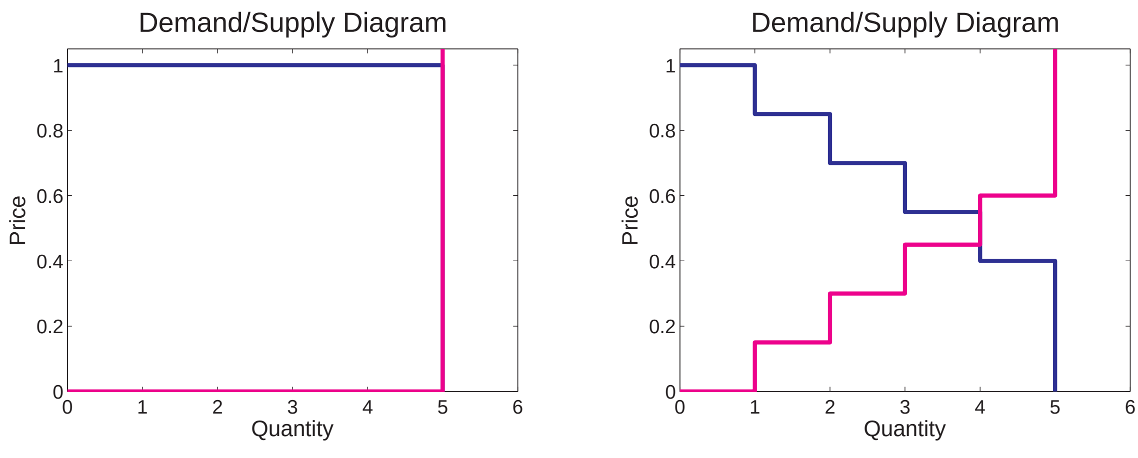
5.2 Market setup

We describe the environments and the trading mechanism in which we study the simultaneous decision about the time of order submission and the submitted offer. Each trader buys or sells in a Continuous Double Auction market one unit of the good and has to decide the moment of submission and the offered price.

5.2.1 The environments

Each environment is determined by a set of buyers and a set of sellers with their redemption values for the good. In each trading period $t \in \{1, \cdots, T\}$ each of the buyers $b \in \{1, \cdots, B\}$ likes to consume one unit of the good and each of the sellers $s \in \{1, \cdots, S\}$ is endowed with one unit of the good. Such a trading period consists of the time moments $\{0, 1, ..., 100\}$. The buyers have a fixed valuation of V_b per unit, sellers have fixed costs of C_s that only needs to be paid in case of a transaction. Agents know their own redemption value, but not the values of the other agents.

We will denote the environments by vectors of valuations and costs. For instance, $\{[1, 1], [0, 0.1, 0.2]\}$ denotes an environment with two buyers having identical valuations 1 and 1 and three sellers with costs 0, 0.1 and 0.2. The supply and demand functions of the benchmark environment $\{[1, 1, 1, 1, 1], [0, 0, 0, 0, 0]\}$ and the main symmetric environment $\{[1, 0.85, 0.7, 0.55, 0.4], [0.6, 0.45, 0.3, 0.15, 0]\}$ are shown in Fig. 5.1.



(a) The benchmark environment.

(b) The main symmetric environment.

Figure 5.1: The demand and supply functions of the main environments used, the benchmark environment $\{[1, 1, 1, 1, 1], [0, 0, 0, 0, 0]\}$ and the main symmetric environment $\{[1, 0.85, 0.7, 0.55, 0.4], [0.6, 0.45, 0.3, 0.15, 0]\}$.

Based on the valuations and costs the demand and supply functions can be determined. We denote the equilibrium quantity by q^* and the interval of equilibrium prices by $[p_L^*, p_H^*]$. The traders that can gain a positive profit in equilibrium are denoted as intramarginals, whereas the traders that cannot make a positive profit and therefore will not trade in equilibrium are called extramarginals. The payoff of a buyer equals $U_b(p) = V_b - p$ if he traded at price p and zero otherwise. The payoff of a seller equals $U_s(p) = p - C_s$ after a trade at price p and zero otherwise.

5.2.2 Continuous Double Auction

A Continuous Double Auction model is used to describe the regular behaviour at stock exchanges. During a trading period buyers and sellers arrive at their preferred moment and immediately submit their order. The bid of buyer b and the ask of seller s in trading period t are denoted as $b_{b,t}$ and $a_{s,t}$ and their arrival times as $n_{b,t}$ and $n_{s,t}$. If an arriving order can be matched with the best order from the book, the transaction takes place at the price of the order in the book. If the arriving order cannot be matched, it is stored in the order book. At the end of the period the order book with unmatched orders is cleared.

If in period t buyer b traded, the transaction price of this buyer is denoted as $p_{b,t}$. Similarly the transaction price of seller s is denoted as $p_{s,t}$. Hence the payoff of a buyer equals $U_{b,t}(p) = V_b - p_{b,t}$ if he traded and zero otherwise. The payoff of a seller equals $U_{s,t}(p) = p_{s,t} - C_s$ after a trade and zero otherwise. We note that the payoff depends not only on the own offer, but also on the trading sequence. For example, if there are only one buyer and one seller, given their offers $b_{b,t} > a_{s,t}$ a buyer will get higher payoff if he will submit his order after the seller, as this will yield a transaction price equal to the ask of the seller. In this chapter we will focus on learning of traders about their timing of submitting the order and about the price of submission. That is why we assume that no order can be cancelled and restrict the traders to buy or sell only one unit of the good. The effects of learning about cancellation and size of the order are left for further research.

5.2.3 Nash equilibria

In the extended model where traders are required to make a two-dimensional decision, a multiplicity of possible long run outcomes may exist. Let us consider a one-period version of our model, where valuations and costs are common knowledge. In this one-period model traders are required to select a strategy consisting of a submission moment and an order only once. Then irrespective of the environment, one set of Nash equilibria exists in which every intramarginal trader submits the same offer in the equilibrium price range, at any possible arrival moment. Hence in this equilibrium the timing of order submission is of no importance. Trivially this constitutes an equilibrium. If a trader adjust its offer price this will result in a lower profit if the offer is more conservative or in absence of trade if more aggressive. However, after a deviation of any offer the arrival moment does play an important role. It is optimal for traders on the other side of the market to arrive at moment $n = 100$ if the deviation leads to a more conservative offer. A more aggressive offer leads to absence of trade for one of the traders on the other side of the market. Hence it is optimal for traders on the other side to arrive at moment $n = 0$.

Furthermore other Nash equilibria may exist in the one-period model, depending on the environment. A trivial example consists of one buyer and two sellers, such that the sellers attempt to outcompete each other. An example of a Nash equilibrium in which not all traders submit the same offer consists of the two sellers submitting an ask price of 0 and the buyer submitting a bid price b at a later moment than both sellers.

The Nash equilibria of this one-period model are possible long run outcomes of the multi-period model used in this chapter. Under the Individual Evolutionary Learning algorithm we find that the offers of traders converge towards the equilibrium price range, such as in the first Nash equilibrium of the one-period model. However, the timing of order submission is of importance; traders learn the optimal submission moment when offer prices of intramarginal traders are not all identical.

5.3 Individual Evolutionary Learning algorithm

Agents learn which strategy to select by a multidimensional version of the Individual Evolutionary Learning algorithm as introduced by Arifovic and Ledyard (2003). Every agent can choose from a set of strategies, which consists of an offer and a submission moment. These strategies might mutate from time to time to allow for some sort of experimentation. Based on how these strategies would have performed in the last trading period, some strategies are replaced by better performing ones. At the beginning of the next period one strategy is selected with a probability proportional to the hypothetical foregone payoff.

Past offers and arrival moments and thus the average price are publicly available. This setup is comparable with the OpenBook setting of Chapter 4. Therefore, after the trading period each agent can determine exactly what his payoff would have been for each possible strategy, assuming no changes in the behaviour of other agents. Agents learn to select the strategy that has the highest hypothetical payoff in the previous period.

Pool of strategies

Every trader has an individual set of strategies: the set $B_{b,t}$ of *K* randomly drawn pairs of bids and arrival moments (b_i, n_i) for buyer b and the set $A_{s,t}$ of *K* randomly drawn pairs of asks and arrival moments (a_j, n_j) for seller s. Offers are initially drawn from a uniform distribution on $[0, V_b]$ and $[C_s, 1]$ respectively and the arrival moments are uniformly drawn from the set $\{0, 1, 2, ..., 100\}$. The periods correspond to days in reality and the arrival moments to the time period of a trading day. Note that the set of arrival moments is larger than the number of traders, which in most environments equals 10. This is done to prevent a large random component in the sequence of submission. When two or more traders decide to submit their order at the same moment in time the orders are handled randomly. Expanding the set of arrival moments reduces this random component.

Traders observe not only the sequence in which traders arrived in the last period, but also their actual moment of submitting. Under the OpenBook system introduced in Chapter 4, the full order book is shown to traders in the New York Stock Exchange. Hence the actual submission moments are known and not only the sequence of submission. Under this assumption agents evaluate the hypothetical payoff in the previous period on the basis of the moment of submission of other traders. This assumption is important as the following example illustrates. If there are 10 traders and trader i desires to submit in place 9 and the submitted moments of others are for example $\{1, 1, 1, 1, 1, 1, 1, 1, 3\}$ (trader i excluded), he prefers to submit at moment 2. This results in submitting the order at place 9, ceteris paribus. We compare with the setup where only the sequence of submissions is known and traders select at which position in the sequence they prefer to arrive, from the set $[1, 2, ..., 10]$. Calculating foregone payoffs is more difficult as it is uncertain in which place in the sequence a certain submission moment results. If the trader in this example prefers to maximise the probability of submitting in place 9, he would submit moment 9 and ceteris paribus arrive at place 10.

Mutation

A part of a strategy mutates with a fixed small probability ρ. A normally distributed variable with mean zero is added to the part of the strategy that mutates while the other part may remain unchanged. The mutated arrival moment is rounded to the nearest integer $n \in \{0, 1, 2..., 100\}$. The variance of the normally distributed variable depends on which part of the strategy mutates. This distribution is truncated; when the mutated strategy lies outside the strategy space, a new normally distributed variable is drawn.

Replication

After the trading period has ended and some strategies have possibly mutated, the foregone payoffs are calculated for each strategy while taking the chosen strategies of others from this period constant. Replication consists of a comparison of two strategies, randomly selected from the pool of strategies. The strategy with the highest foregone hypothetical payoff will obtain a place in the updated pool of strategies. This is repeated K times to fill the entire updated pool.

Hypothetical foregone payoff functions

Calculating the foregone payoffs is a straightforward task under full information about the trading history. Since the entire order book and the arrival moments of others in the last period are known, the foregone payoff can precisely be determined for every possible strategy, given that others remain to use the same strategies.

For example, with only one buyer and one seller who in the previous period submitted ask $a_{s,t}$, the hypothetical foregone payoff of the buyers' strategy (b_i, n_i) is equal to $V_b - a_{s,t}$ when $n_i > n_{s,t}$ and $b_i \geq a_{s,t}$, and is equal to $V_b - b_i$ when $n_i < n_{s,t}$ and $b_i \geq a_{s,t}$ and zero otherwise. When $n_i = n_{s,t}$ one of the traders randomly arrives first, and the hypothetical foregone payoff of the strategy (b_i, n_i) equals $\frac{1}{2}(V_b - a_{s,t}) + \frac{1}{2}(V_b - b_i)$. The hypothetical foregone payoff functions are in general given by

$$U_{b,t}(b_i, n_i) = \begin{cases} V_b - p^*_{b,t}(b_i, n_i) & \text{if strategy } (b_i, n_i) \text{ resulted in a tradeat price } p^*_{b,t}(b_i, n_i) \\ 0 & \text{otherwise,} \end{cases}$$

$$U_{s,t}(a_j, n_j) = \begin{cases} p^*_{s,t}(a_j, n_j) - C_s & \text{if strategy } (a_j, n_j) \text{ resulted in a tradeat price } p^*_{s,t}(a_j, n_j) \\ 0 & \text{otherwise.} \end{cases}$$

Selection of a strategy from the pool

After the hypothetical foregone payoffs are determined each trader has to select a strategy for the next trading period. The probability that a certain strategy is selected is proportional to its hypothetical foregone payoff. In the first period every strategy is equally likely to be chosen. For a buyer b_i the probability of selecting strategy (b_i, n_i) for period $t + 1$ is given by

$$\pi_{b,t+1}(b_i, n_i) = \frac{U_{b,t}(b_i,n_i)}{\sum_{i=1}^{K} U_{b,t}(b_i,n_i)}.$$

This Individual Evolutionary Learning algorithm depends on some variables, such as the size of the individual pools, the probability and the distribution of mutation and the replication rate. The next section shows the values of these variables that are used in the simulations, as well as the characteristics used to describe the overall outcome in a trading period.

5.3.1 Methodology

The benchmark environment $\{[1,1,1,1,1],[0,0,0,0,0]\}$ is considered to show the basic results of IEL-learning regarding the moment of arrival. We simulate this environment with different numbers of traders to study the effect of the size of the market. Furthermore we will use symmetric environments which mainly consist of five buyers and five sellers to study the impact of changes in the amount of competition and the size of the equilibrium price range on the distribution of arrival moments. Some of these environments as the main symmetric environment $\{[1,0.85,0.7,0.55,0.4],[0.6,0.45,0.3,0.15,0]\}$ are introduced in Arifovic and Ledyard (2007) and Anufriev et al. (2013). We will show that the results are not robust with respect to the environment; when one side of the market is much larger the other side will extract their power and submit their order as late as possible.

We study the long-run distribution of arrival moments and the expected offer per submission moment. Also of interest are the *allocative efficiency*, which is the ratio between the allocative value in a trading period and the maximal possible allocative value. The allocative value of a trading period is the sum of the payoffs of all agents. It is fully efficient when all intramarginal investors trade during a period. Efficiency can be lower when an extramarginal investor trades, or when intramarginals do not trade at all. Furthermore we study the *average transaction price*, the *price volatility* and the *number of transactions*. All these characteristics are considered per trading period as well as per possible arrival moment and are compared with the one dimensional model.

The Individual Evolutionary Learning algorithm is used with most of the parameters of Arifovic and Ledyard (2007). Every agent is given an individual pool of strategies of size $K = 300$. A part of a strategy mutates with a probability of 0.033; in the case that the offer mutates a normally distributed term with mean 0 and a standard deviation of 0.1 is added to the offer, in the case that the arrival moment mutates a normally distributed term with mean 0 and a standard deviation of 10 is added to the arrival moment. The mutated offer is truncated on the bounds of the interval $[0,1]$ and the mutated arrival moment is rounded to the nearest integer

$n \in \{0, 1, 2..., 100\}$. When the mutated strategy lies outside the strategy space, a new normally distributed variable is drawn. This mutation differs from Anufriev et al. (2013), where a uniform distribution is used to form the new strategy. In the replication phase K pairs are compared.

All the *averages* are calculated over $S = 3000$ *random seeds*. The benchmark environment $\{[1, 1, 1, 1, 1], [0, 0, 0, 0, 0]\}$ shows that 3000 seeds is indeed sufficient. After a number of periods the market becomes more or less stable and the offers and average price only fluctuate within a certain range, mainly due to mutation. We denote this behaviour as an *"equilibrium"* in which the offers of intramarginals are close to the equilibrium price range and the agents choose the time to submit that showed to perform the best given these offers. All the results are averaged over periods $41 - 50$ to avoid the random impact of the first learning periods. We show that the distribution of submission moments is stable after 40 periods and thus the impact of the first learning periods is negligible. This is done by conducting a two-sample Kolmogorov-Smirnov test on periods 39 and 40. The test statistic $D = 0.0022$ lies outside the critical region $D > 0.0159$ for $\alpha = 0.001$. Chapter 4 studies this impact by considering both the learning and the equilibrium phase.

5.4 Benchmark environment

In this section we focus on the distribution of arrival moments and its correlation with the chosen offer, in a benchmark environment. Important characteristics as efficiency, variance of transaction prices and volume are considered both per trading period as per possible arrival moment. We consider simulations of the basic environment with the following redemption values: $\{[1, 1, 1, 1, 1], [0, 0, 0, 0, 0]\}$. In this environment all buyers and all sellers are identical and buyers and sellers are symmetric. In Fig. 5.2 we show the distribution of arrival times and the expected offer per arrival moment for every trader individually, where buyers are represented by solid lines and sellers by dotted lines. Averages are shown in solid black lines. Bids of buyers are positively correlated with the moment of arrival, asks of sellers are negatively correlated. Buyers' valuations are shown in the offer function plotted by stars and sellers' costs by circles.

When multiple traders have the same valuation or cost, colours do not correspond to the colours of the offer functions, but are shown in black.

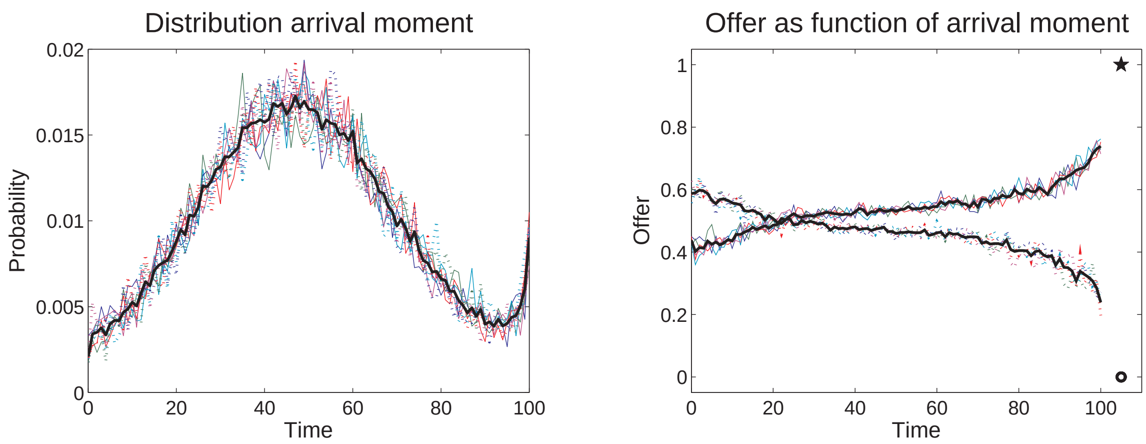


Figure 5.2: Distribution of learned submission moments (left) and offers as function of submission moment (right) in environment $\{[1, 1, 1, 1, 1], [0, 0, 0, 0, 0]\}$. Traders learn to submit their order more often during the middle of the period. The increasing bid function and decreasing ask function show that early submitting traders are more aggressive.

We derive more intuition about the distribution of submission moments by further investigating the environment with 5 buyers and 5 sellers. In Section 5.5 we study robustness with respect to the number of traders. This environment is attractive since it is an intermediate case. In the environment with one trader on either side of the market the probability of trading does not play a role, and as the number of traders converges to infinity the effect of timing on the expected profit from trade fades away. However, in this intermediate environment both effects play a significant role.

Moment of order submission

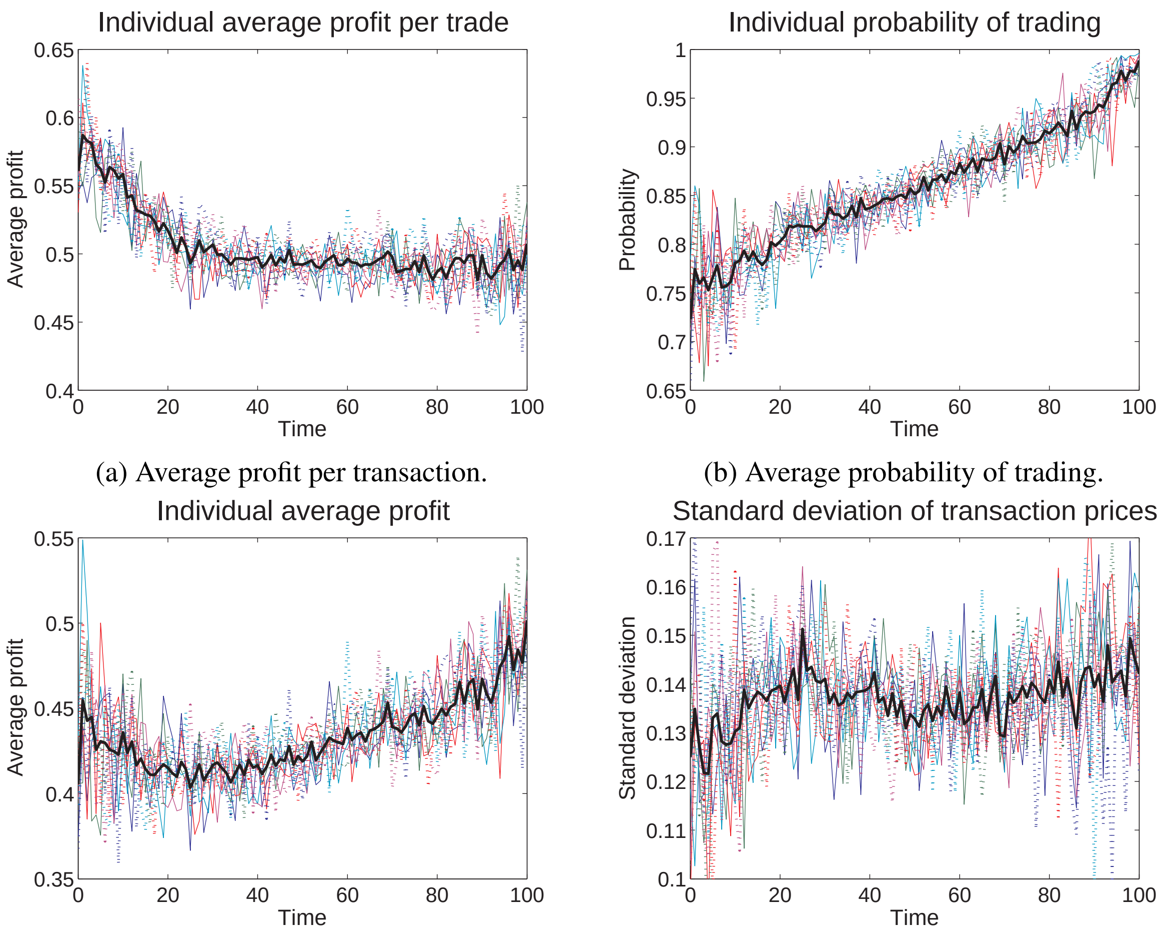
With respect to the preferred moment of submitting we observe in Fig. 5.2 that agents desire to submit their order in the middle part of the trading period. This illustrates the trade-off any individual trader faces: submitting the same offer earlier increases the probability that a trade will occur at the price of their own offer which results in a lower expected profit, submitting later decreases the probability of trading. A peak at the end of the period exists and we observe an increasing bid function and decreasing ask function. The latter conclusion is also drawn in the paper Fano and Pellizzari (2011). Traders who submit their offer late, bid close to the valuation or ask close to the cost, which yields a high probability of trading. It makes sense for such a late offer to be submitted as late as possible. This ensures that the trade occurs at the preferable price of the other trader, which outweighs the minimal decrease in the probability of trading.

Traders learn which arrival moment performs best when the offers of intramarginal traders are not identical. Under the IEL algorithm, mutation is the main cause that strategies do not entirely converge. Due to mutation the trade-off between the probability of trading and the expected profit is of importance and the moment of submission plays a crucial role. This results in the distribution shown in Fig. 5.2. The micro-motives of this distribution are further investigated in Section 5.5.

Offer

Our main observation, with regards to the offer that agents submit in relation with the preferred moment of submitting, is that the earlier they intend to submit their order, the further bids are from the valuations and asks from the costs. If an agent prefers to submit at a late moment, he intends for a lower profit to increase his probability of trading. Thus we find a positive correlation between bids and time and negative between asks and time.

Characteristics per possible arrival moment are shown in Fig. 5.3. The average profit per transaction is shown in panel (a), conditional on the arrival moment. This excludes the instances where a trader arrives at that moment but does not trade. The next panel shows the probability of trading; defined as the number of trades divided by the number of arrivals at a given moment. Panel (c) shows the average profit over all the instances in which a trader arrives at that moment. Included in this average are the instances at which no trade occurs and a trader receives a profit of 0. This panel is the product of the first two panels; the probability of trading and the average profit of a transaction. Finally the standard deviation of transaction prices is shown for every arrival moment in panel (d).



(a) Average profit per transaction.

(b) Average probability of trading.

(c) Average profit.

(d) Standard deviation of transaction prices.

Figure 5.3: Characteristics environment $\{[1,1,1,1,1],[0,0,0,0,0]\}$ per possible arrival moment. The average profit per transaction is decreasing and the average probability of trading increasing. The average profit, which is the product of the two, is U-shaped. The standard deviation of transaction prices is increasing at the end of the period and is μ-shaped.

The average profit per transaction is decreasing and the probability of trading is increasing during the period. Early submitting traders submit a very aggressive offer, a bid below and a ask above one half. This will often not result in a trade, but if a trade occurs this will yield a very high profit. An early or late arrival leads to a higher average profit, but also a higher variance of transaction prices. A higher variance of the price at which the trade occurs (if late), or a higher variance because of more occasions where the agent does not trade (if early). In the replication process early and late strategies are therefore often removed from the pool and that is why agents more often submit in the middle of the trading period. It is remarkable that in the IEL-algorithm traders learn to arrive more frequent at moments that yields a lower average profit. This may be the result of the learning algorithm, which only considers the profit in the previous period. It is likely that when the IEL-algorithm is modified in such a way that hypothetical profit is averaged over multiple periods, a reversal will occur and traders more frequently submit at the beginning or the end of the period.

Let us suppose that traders do not condition their offer on the moments of submitting, thus always submit the same bid or ask. In this case the probability of trading would be decreasing over time because fewer possible trading partners remain, whereas the expected profit per trade is increasing over time because it is more likely to occur at the preferable price of the other trader. However, the buyers' bid function is increasing and the sellers' ask function is decreasing. This results in a reversal of these two effects.

5.4.1 Knowledge of the submission moments

We compare the results in the benchmark environment with 5 buyers and 5 sellers to the setting where traders only observe the sequence of arrivals and the offers of traders, but not the actual submission moments. In the latter setting traders choose from the set $[1, 2, ..., 10]$ in which position in the sequence they want to submit their order. We observe in Fig. 5.4 that knowledge of the actual arrival moment in addition to the sequence impacts the arrival moment distribution. This is confirmed with a two-sample Kolmogorov-Smirnov test. The test statistic $D = 0.0530$ lies in the critical region $D > 0.0051$ for $\alpha = 0.001$. The additional knowledge reduces the kurtosis and increases the peak at the end of the period.

5.4.2 Allowing the choice of submission moment

For the two-dimensional model with timing and the one-dimensional model without timing, we measure the overall results for efficiency, number of trades and volatility in the benchmark environment with 5 buyers and 5 sellers over trading periods $41 - 50$. These characteristics and

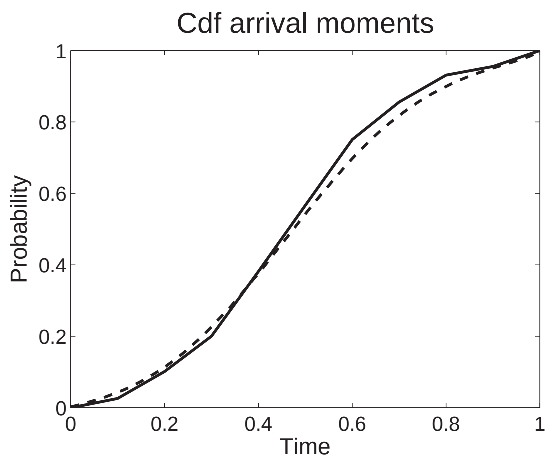


Figure 5.4: CDF of arrival times for knowledge of sequence (solid line) and actual arrival moments (dotted). The distribution of arrival moments is significantly altered when knowledge of the actual arrival moments of others in the previous period is added. This additional information reduces the kurtosis and increases the peak in the distribution of arrival moments at the end of the period.

	With timing	Without timing
Efficiency	0.8567	0.8952
	(0.1718)	(0.0447)
Price Volatility	0.0208	0.0177
	(0.0101)	(0.0077)
Number of transactions	4.2837	4.3520
	(0.8591)	(0.1290)

Table 5.1: Average outcomes with and without timing in environment $\{[1,1,1,1,1],[0,0,0,0,0]\}$. The average efficiency and the average number of trades significantly decrease, and the average price volatility significantly increases when traders are allowed to submit orders at their preferred moment.

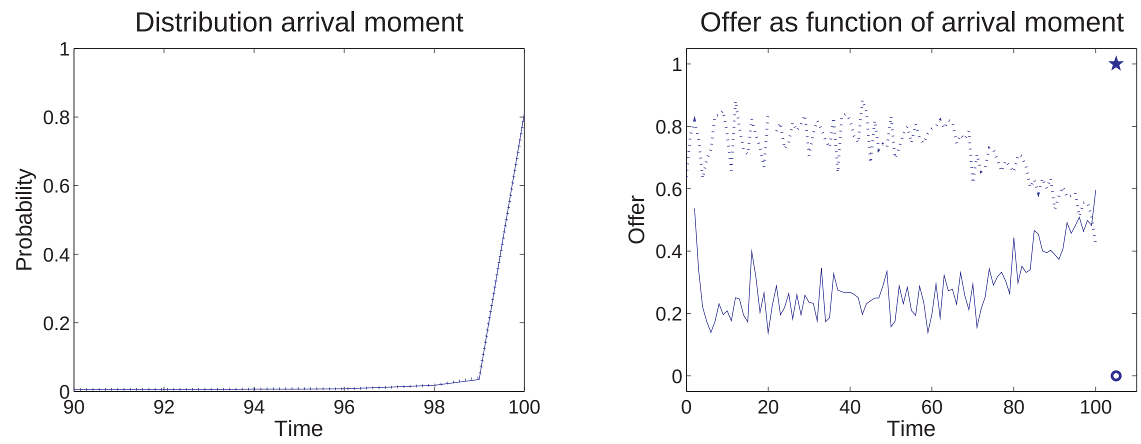
their standard deviations are shown in Table 5.1. Allowing traders to submit at their preferred moment has a negative effect; the average efficiency and the average number of transactions decrease, and the average price volatility increases. These comparisons are all significant at a significance level of 1%. It is optimal not to allow traders this extra decision, since it results in a lower efficiency and thus a lower profit.

5.5 Size of the market

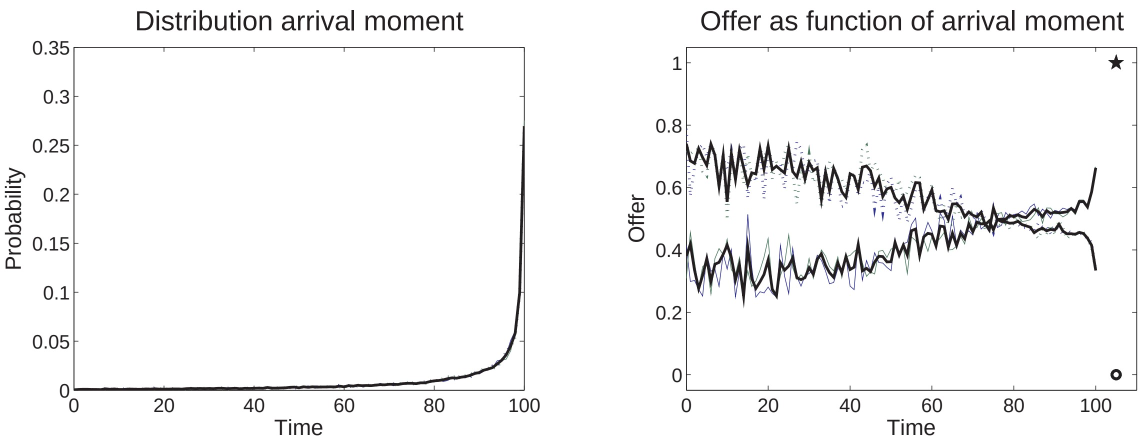
The size of the market plays a crucial role on the distribution of arrival moments. Two forces are important for the expected profit: the expected profit from a transaction and the probability of trading. In thinner markets the first has a larger impact than in thicker markets. In Fig. 5.5 we show the distribution of arrival moments and the average offer per trading moment depending on the number of traders. The results are given for 1, 2, 5, 10 and 15 traders on either side of the market.

Moment of order submission

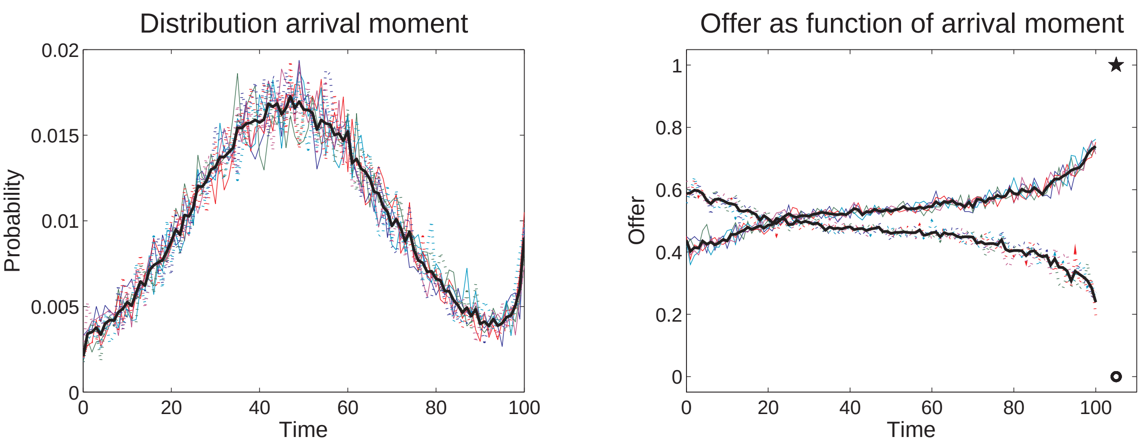
In a market with only one buyer and one seller the moment of arrival does not affect the probability of trading. It is optimal to submit the order at moment 100, since submitting after the other trader results in a trade at the price of the other trader. In the simulations we find indeed that traders submit as late as possible. Submitting later strongly dominates submitting earlier in the IEL algorithm and hence in a market with one buyer and one seller the IEL algorithm selects submission moment 100. A similar distribution of submission moments is shown for two buyers and two sellers. When the size of the market increases, the probability of trading does play a role. Also the effect of the moment of arrival on the expected profit from trade decreases, since the probability that the transaction price equals the own offer tends towards one half for every arrival moment. With five traders on either side of the market traders submit around the middle of the period. The larger the size of the market, the earlier traders arrive. The simulations suggest that the moment of arrival will converge to zero as the size of the market converges to infinity, which would in the benchmark environment with infinitely many traders be optimal as the effect of the expected profit from trade disappears.



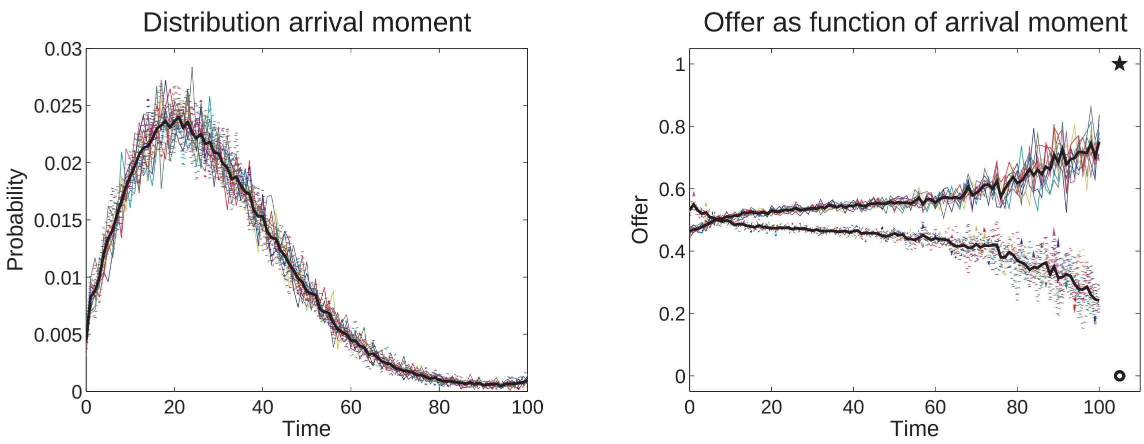
Dynamics in environment with 1 buyer with valuation 1 and 1 seller with cost 0.



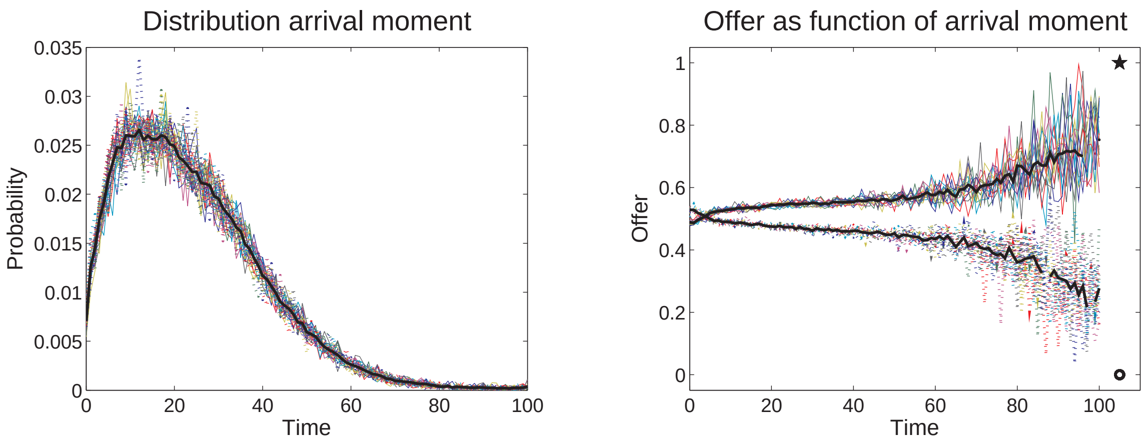
Dynamics in environment with 2 buyers with valuation 1 and 2 sellers with cost 0.



Dynamics in environment with 5 buyers with valuation 1 and 5 sellers with cost 0.



Dynamics in environment with 10 buyers with valuation 1 and 10 sellers with cost 0.



Dynamics in environment with 15 buyers with valuation 1 and 15 sellers with cost 0.

Figure 5.5: Distribution of learned submission moments (left) and offers as function of submission moment (right), for $1, 2, 5, 10$ and 15 traders on either side of the market. In the environment with one buyer and one seller both traders submit as late as possible, since the probability of trading does not depend on the submission moment. As the size of the market increases, this probability does play a role and traders submit earlier and earlier, and moreover bids increase and asks increase for every possible arrival moment.

Offer

Irrespective of the arrival moment traders on average submit a higher bid and a lower ask as the size of the market increases. Average bids and asks of early arriving traders are below respectively above one half, but the intersection points of the bid and ask function with one half approaches $n = 0$. For traders arriving late the average bids and asks are further away from one half and thus closer to their valuation and cost.

5.6 Competition

In this section we study how the distribution of arrival moments and the function of average offers is affected by different aspects of competition. Sets of environments are considered in which the size of one side of the market increases, competition to extramarginal traders increases, additional extramarginal traders enter or the range of equilibrium prices decreases.

5.6.1 Decreasing competition between buyers, increasing competition between sellers

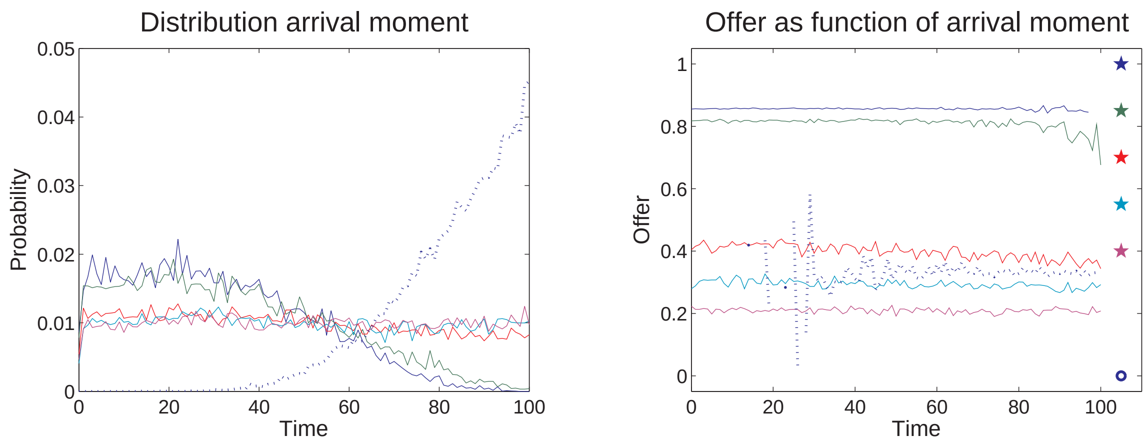
Competition between buyers decreases and competition between sellers increases as more and more sellers enter, as illustrated in Fig. 5.6. These environments range from a large difference in size between both sides of the market, towards a symmetric environment.

Moment of order submission

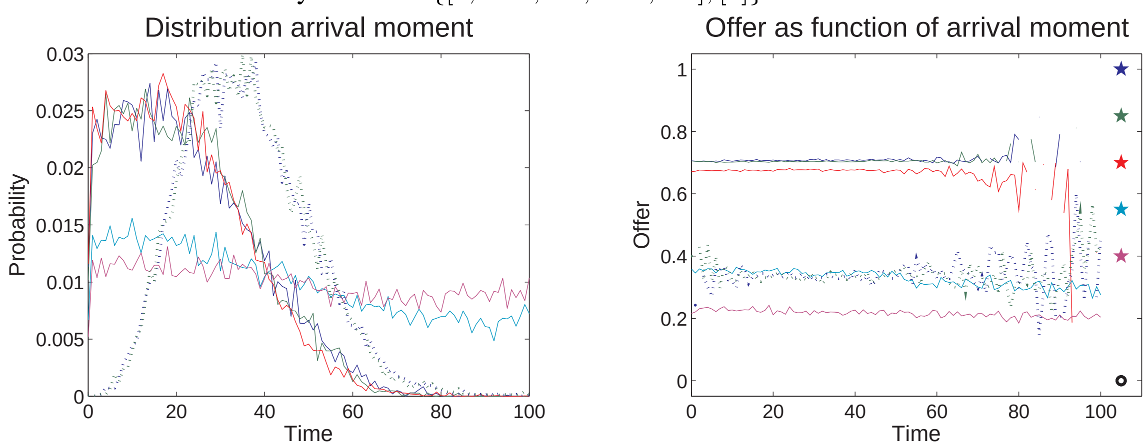
When there is little competition between sellers they tend to trade late. After more and more sellers enter, the intramarginal sellers trade earlier and earlier. Intramarginal buyers trade early to outcompete extramarginals. In the final environment in Fig. 5.6 the distribution is symmetric and traders prefer not to trade too early. The extramarginal buyers submit every moment with the same probability in the first environment, and act more like the intramarginals as more sellers enter; in which case they have more opportunities to trade.

Offer

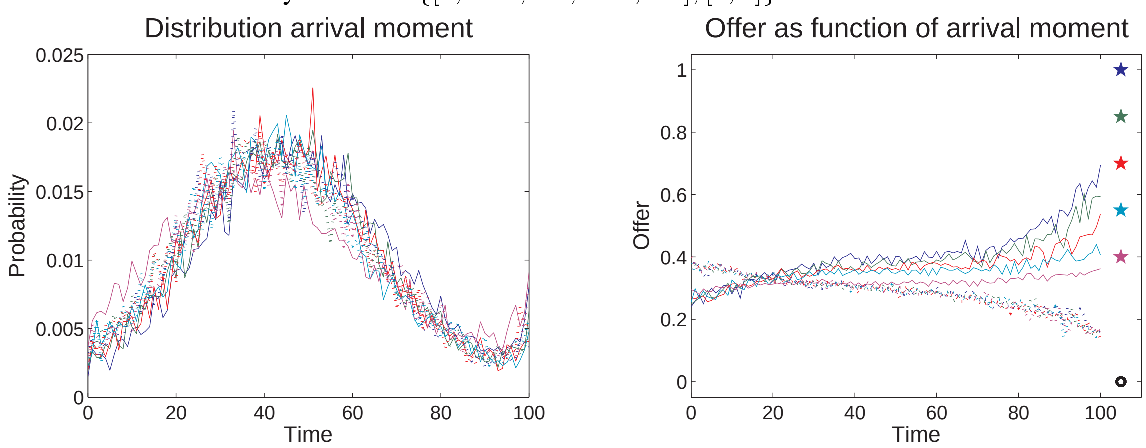
Sellers submit lower asks as more sellers enter; since their market power decreases they can be less aggressive. The intramarginal buyers submit a bid higher than the extramarginal buyers. When a seller enters, the buyer that becomes the most competitive extramarginal buyer significantly increases his bid. The other intramarginal buyers lower their bid, so that they slightly overbid the most competitive extramarginal buyer. The average offer of extramarginals is not affected by their moment of submitting their order.



Dynamics in $\{[1, 0.85, 0.7, 0.55, 0.4], [0]\}$ environment.



Dynamics in $\{[1, 0.85, 0.7, 0.55, 0.4], [0, 0]\}$ environment.



Dynamics in $\{[1, 0.85, 0.7, 0.55, 0.4], [0, 0, 0, 0, 0]\}$ environment.

Figure 5.6: Distribution of learned submission moments (left) and offers as function of submission moment (right) with decreasing competition between buyers and increasing competition between sellers. With little competition between sellers, they tend to trade late in order to trade with the buyer who submitted the highest bid, and intramarginal buyers trade early. When more sellers are added to the environment the arrival moments move towards the middle of the period. Intramarginal buyers slightly overbid the most competitive extramarginal buyer.

5.6.2 Increasing competition to extramarginal traders

Competition between intramarginal traders and the two extramarginals increases in Fig. 5.7 as their valuation and cost become closer to the equilibrium price range.

Moment of order submission

When extramarginals in Fig. 5.7 have little opportunity to trade they have no clear preferred arrival moment. As they can compete more with intramarginal traders they prefer to trade earlier to increase their probability of trading. The intramarginals that set the range of equilibrium prices, and thus are the traders that face the most competition from extramarginals, trade earlier than the other intramarginals when competition to extramarginals is less. They behave more similar to the other intramarginals when competition increases.

Offer

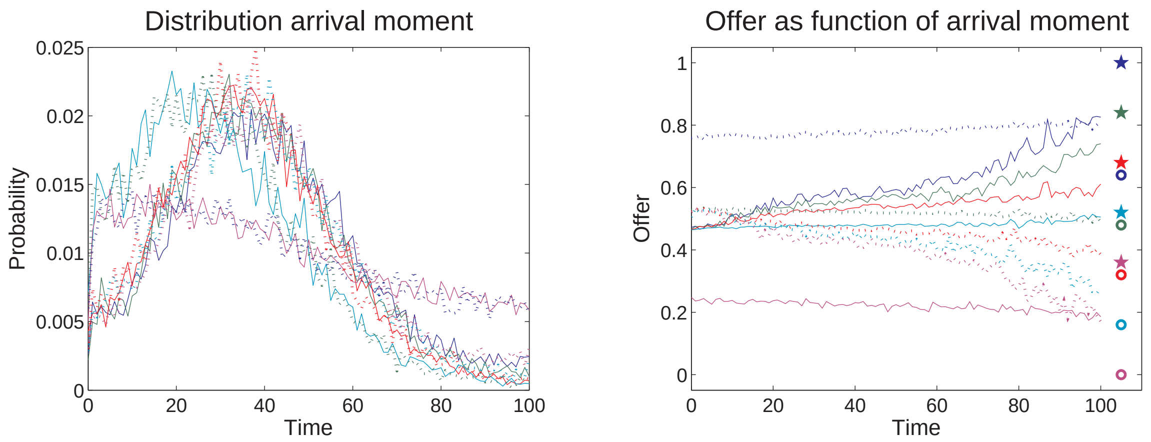
When competition increases and the valuation and the cost of the extramarginals get closer to the equilibrium price range, they post less aggressive offers. As a result, the intramarginals that set the range of equilibrium prices also submit less aggressive offers to ensure that their offers are better than the offers from the extramarginals.

5.6.3 Extramarginal traders entering

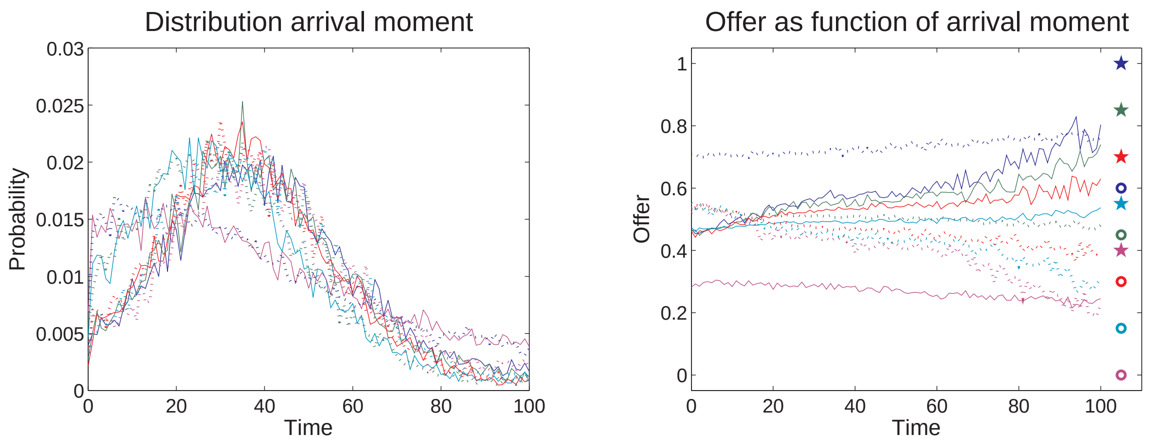
In Fig. 5.8 we show environments with zero, one and two extramarginals on either side of the market. Over these environments the same intramarginals face an increasing probability of absence of trade due to competition to extramarginals.

Moment of order submission

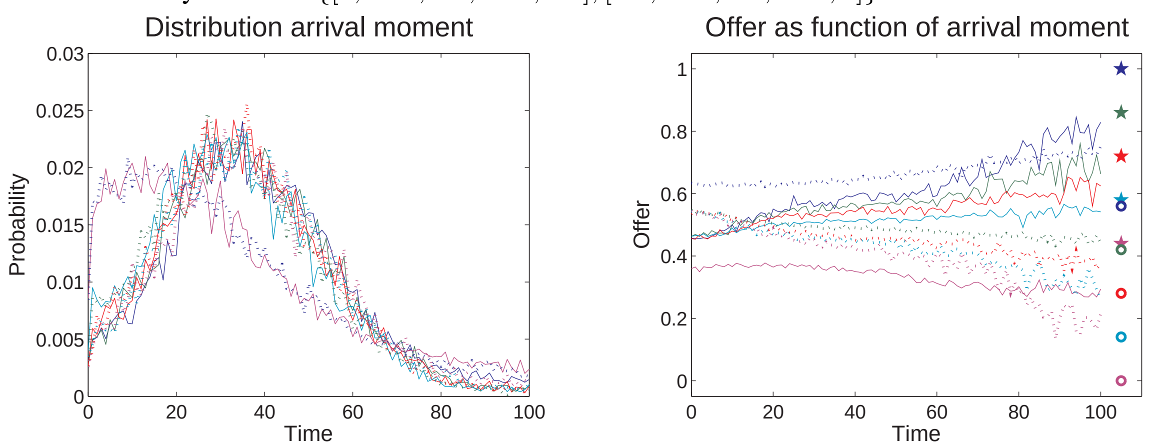
Intramarginals tend to trade earlier when extramarginals enter. The intramarginals that set the range of equilibrium prices tend to trade even earlier than the rest of the intramarginals. The most competitive extramarginals in the last environment prefer to trade earlier to increase their probability of outcompeting an intramarginal.



Dynamics in $\{[1, 0.84, 0.68, 0.52, 0.36], [0.64, 0.48, 0.32, 0.16, 0]\}$ environment.

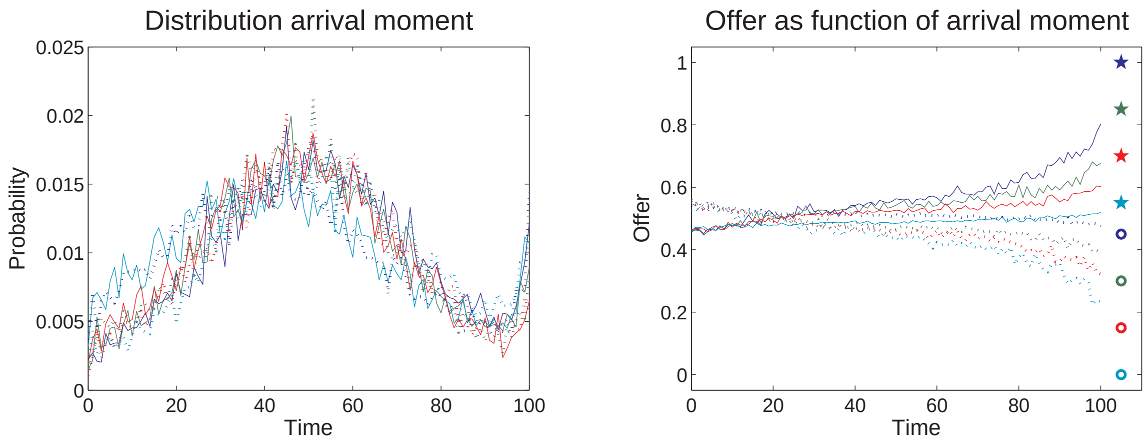


Dynamics in $\{[1, 0.85, 0.7, 0.55, 0.4], [0.6, 0.45, 0.3, 0.15, 0]\}$ environment.

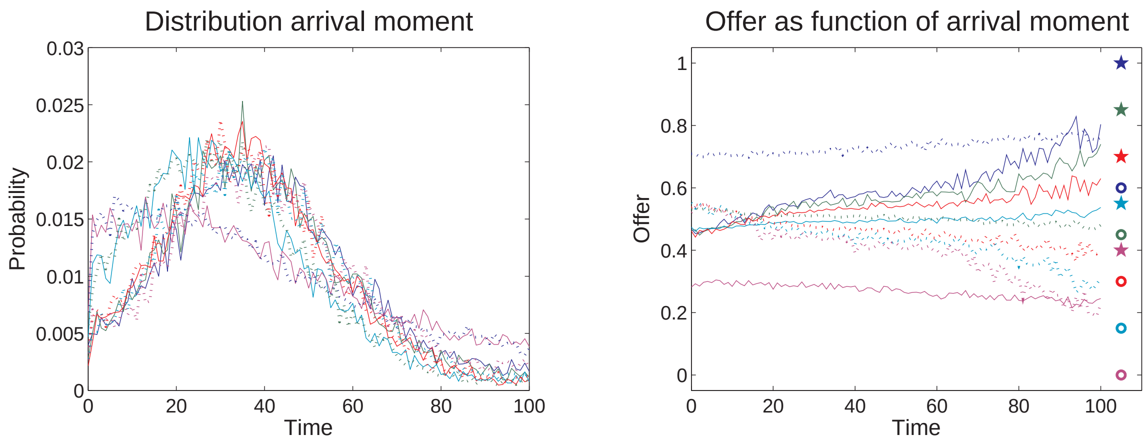


Dynamics in $\{[1, 0.86, 0.72, 0.58, 0.44], [0.56, 0.42, 0.28, 0.14, 0]\}$ environment.

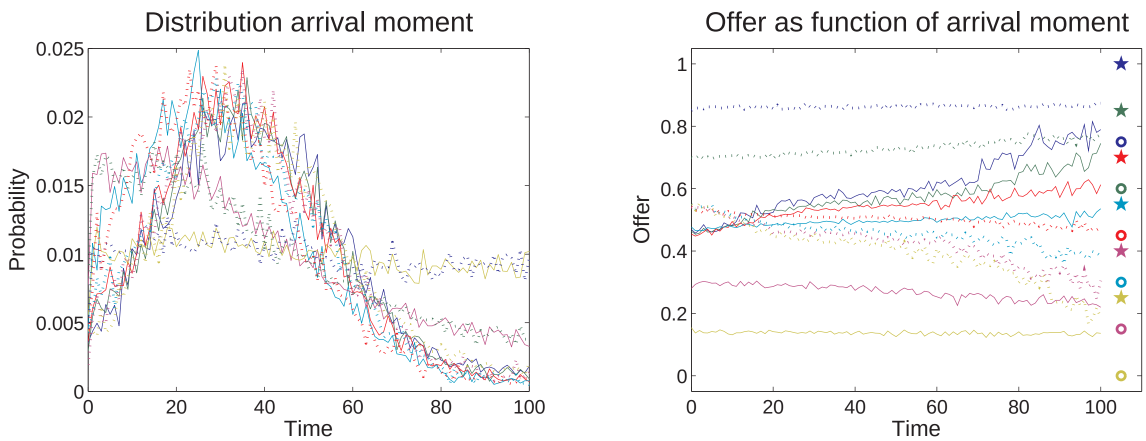
Figure 5.7: Distribution of learned submission moments (left) and offers as function of submission moment (right) with increasing competition to extramarginal traders. As competition to extramarginal traders increases, these traders learn to submit early and post less aggressive offers to increase their probability of trading. The weakest intramarginal traders submit later and submit less aggressive offers.



Dynamics in $\{[1, 0.85, 0.7, 0.55], [0.45, 0.3, 0.15, 0]\}$ environment.



Dynamics in $\{[1, 0.85, 0.7, 0.55, 0.4], [0.6, 0.45, 0.3, 0.15, 0]\}$ environment.



Dynamics in $\{[1, 0.85, 0.7, 0.55, 0.4, 0.25], [0.75, 0.6, 0.45, 0.3, 0.15, 0]\}$ environment.

Figure 5.8: Distribution of learned submission moments (left) and offers as function of submission moment (right) as more intramarginal traders enter the market. This increasing competition forces intramarginal traders to submit their order earlier.

Offer

There is not a clear effect of extramarginal traders entering in Fig. 5.8 on the average offer of intramarginals. Hence the intramarginals solely respond to extra competition to extramarginals by trading earlier, not by altering their offer function. The extramarginal traders submit a similar offer for every arrival moment.

5.6.4 Decreasing range of equilibrium prices

Over the environments of Fig. 5.9, all valuations decrease by 0.02 and all costs increase, which reduces the equilibrium price range. Thus competition between intramarginal traders is increased and competition to extramarginals decreased.

Moment of order submission

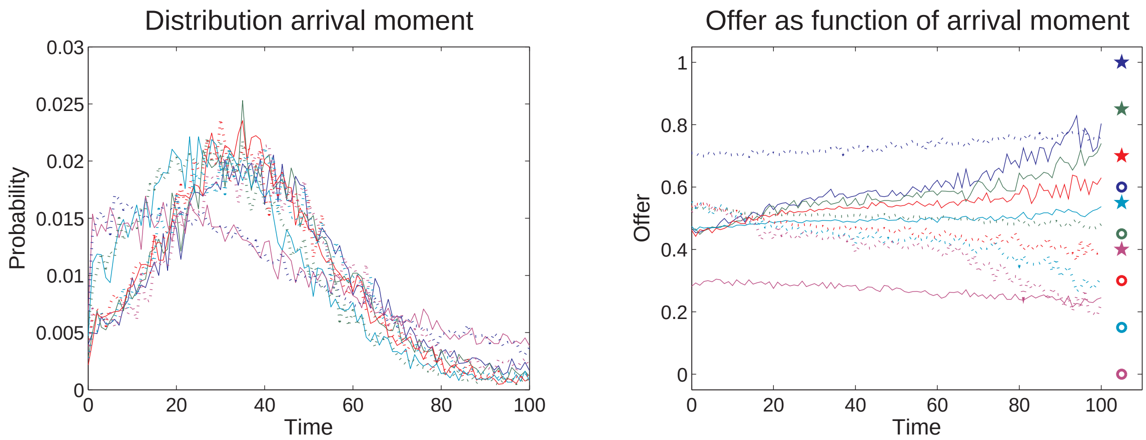
As the range of equilibrium prices decreases in Fig. 5.9, the intramarginals that set the range of equilibrium prices tend to trade earlier. The remaining intramarginals alter their moment of arrival very limited. Extramarginals prefer to trade early, but this effect decreases as the equilibrium price range decreases and their valuation and cost are relatively further away from this range.

Offer

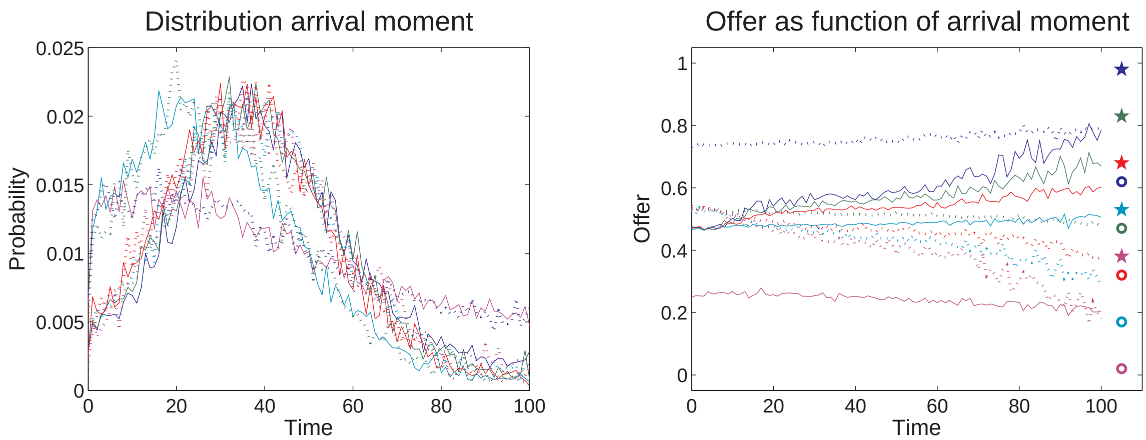
The offer functions of traders that set the equilibrium price range become more constant as the equilibrium price range decreases. The average offer function lies within the equilibrium price range. There is no clear impact of the increasing competition on the other intramarginals and the decreasing competition on extramarginals. However, the change in valuations and costs naturally causes the average bid to decrease and the average ask to increase.

5.7 Gode-Sunder environments

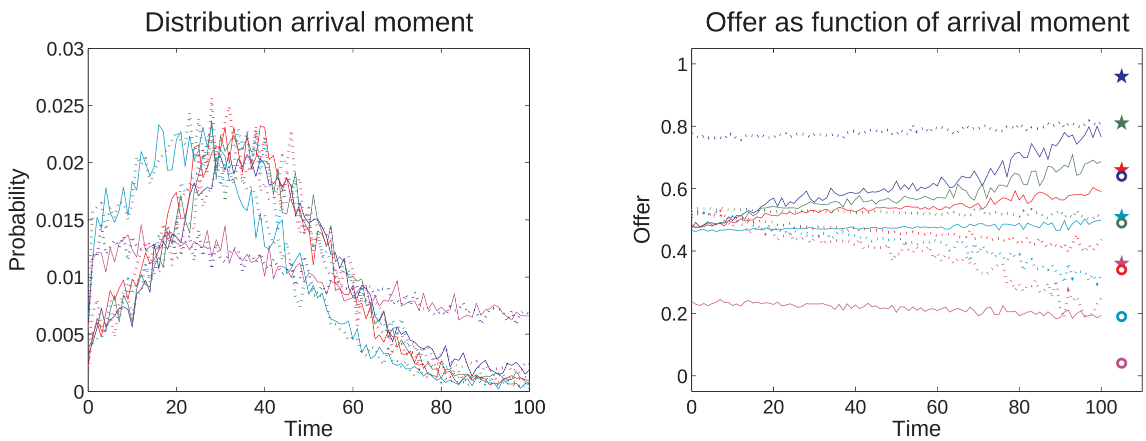
Gode and Sunder (1997) study the impact of extramarginals in the so-called GS-environments. These environments consist of one seller with cost 0, one buyer with valuation 1 and a set of ex-



Dynamics in $\{[1, 0.85, 0.7, 0.55, 0.4], [0.6, 0.45, 0.3, 0.15, 0]\}$ environment.



Dynamics in $\{[0.98, 0.83, 0.68, 0.53, 0.38], [0.62, 0.47, 0.32, 0.17, 0.02]\}$ environment.



Dynamics in $\{[0.96, 0.81, 0.66, 0.51, 0.36], [0.64, 0.49, 0.34, 0.19, 0.04]\}$ environment.

Figure 5.9: Distribution of learned submission moments (left) and offers as function of submission moment (right) when the range of equilibrium prices decreases. The least competitive intramarginals submit earlier and the extramarginal traders later.

tramarginal buyers with valuation β. In this analysis we consider three extramarginal buyers under different values of β. The demand and supply function of the GS-environment with $\beta = 0.5$ is shown in Fig. 5.10. Anufriev et al. (2013) determine efficiency in these GS-environments in a simple CDA. They show that the efficiency under the IEL-algorithm is very close to one and significantly larger than under Zero Intelligence. We investigate the impact of timing in the GS-environments in Fig. 5.11 and compare efficiency with Anufriev et al. (2013) in Fig. 5.12.

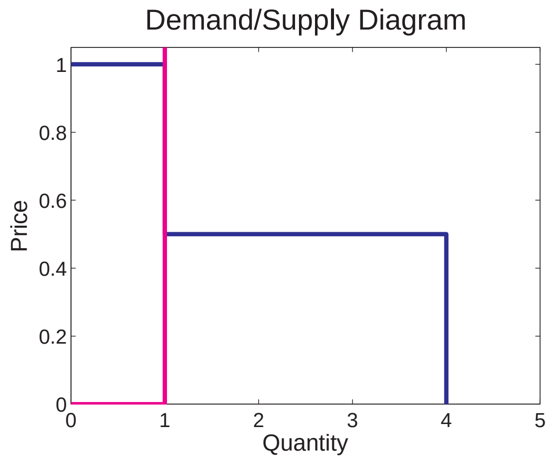


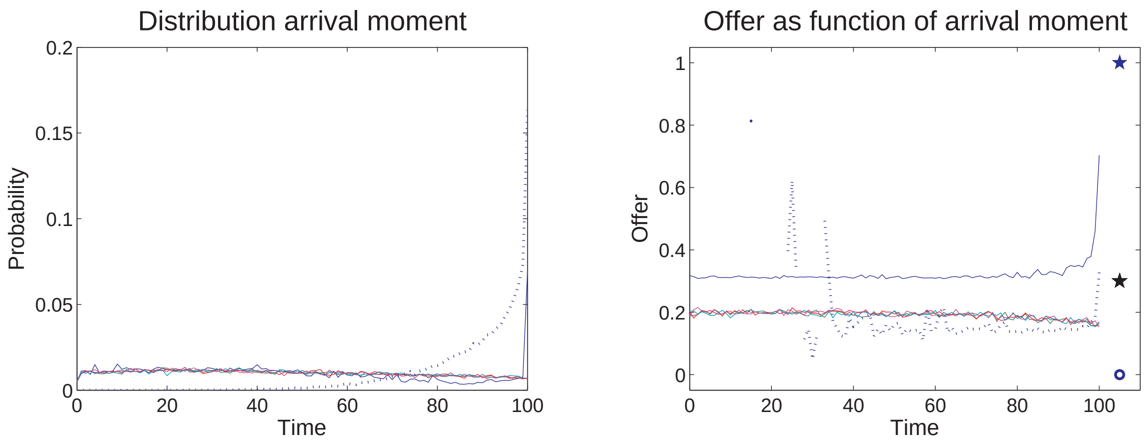
Figure 5.10: The demand and supply function of the GS-environment with 1 buyer with valuation 1, 1 seller with cost 0 and 3 extramarginal buyers with valuation $\beta = 0.5$.

Moment of order submission

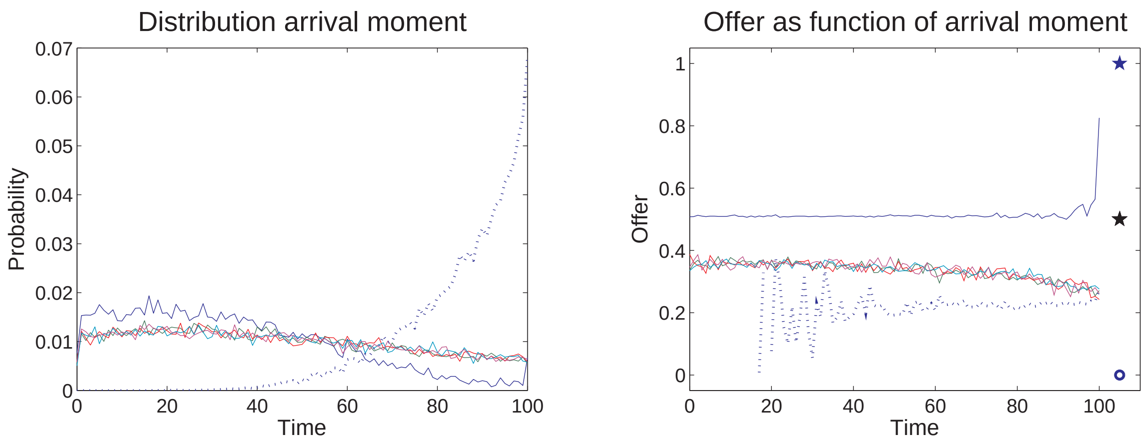
The seller submits late and makes use of his market power to face the best possible bid and trade at the price of the buyer. As β increases the seller has a weaker incentive to enter late. The intramarginal buyer faces more competition and is forced to trade earlier. As extramarginals are competing more with the intramarginal buyer they tend to trade earlier. Arriving early can be used to outperform other buyers if the seller submits early.

Offer

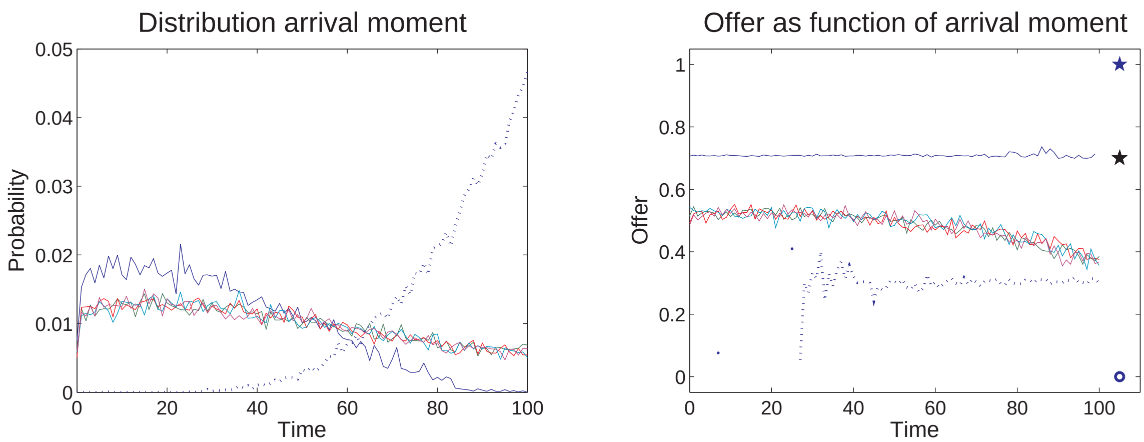
The intramarginal buyer increases his bid as β increases to outbid other buyers. The seller increases his ask to trade more often with the intramarginal buyer; which yields a higher profit. Early arriving extramarginals relatively bid higher to outbid other early arriving buyers.



Dynamics in GS-environment with $\beta = 0.3$.



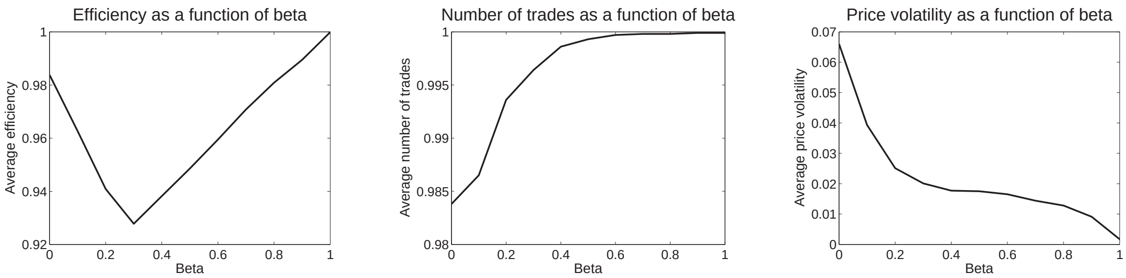
Dynamics in GS-environment with $\beta = 0.5$.



Dynamics in GS-environment with $\beta = 0.7$.

Figure 5.11: Distribution of learned submission moments (left) and offers as function of submission moment (right) in the GS-environment with 3 extramarginal buyers with valuation β. The seller submits his offer late in order to make use of his market power to trade against the best possible bid. As β increases the intramarginal buyer faces more competition, requiring him to increase his bid.

The efficiency function shows the same non-monotonicity as in Gode and Sunder (1997) and Anufriev et al. (2013). However, as stated earlier, with a more complex decision problem for traders that yields more freedom we find a lower efficiency than in Anufriev et al. (2013). The number of trades is increasing and the volatility decreasing in β.



(a) Average efficiency. (b) Average number of trades. (c) Average volatility.

Figure 5.12: Characteristics of the GS-environments with 3 extramarginal buyers with valuation β. The efficiency as a function of β follows a U-shape similar to Gode and Sunder (1997) and Anufriev et al. (2013). The number of transactions is increasing and the price volatility decreasing in β.

5.8 Concluding Remarks

In the Continuous Double Auction the moment of submitting the order plays a crucial role; submitting at the end of the trading period may yield a lower probability of trading, submitting at the beginning of the period will most likely result in a trade at the own submitted price which yields a lower profit. This chapter is a step forward to a more complete model of learning in markets. Moreover, it is distinguished from other papers by the decision of traders. Instead of a one-dimensional decision traders are required to make a two-dimensional decision; which bid or ask to submit and when to submit the offer during the trading period. We showed that the size of the market and competition between traders influence this distribution.

The distribution of arrival moments is studied in a benchmark environment under full information about trading history. We found in simulations that under the Individual Evolutionary

Learning algorithm intramarginal traders learn to submit their order around the middle of the trading period. This result holds for a medium size market with a comparable number of traders on each side. If one side of the market is thinner it can extract more profit by submitting later. Our main observation with regards to the offer that agents submit, in relation with the preferred moment of submitting, is that the earlier they submit their order, the higher profit they aim for. If an agent submits his order at a late moment, he submits a conservative offer to increase his probability of trading. As a result, an early or late arrival results in a higher expected profit. However, an early arrival increases the risk of not trading and a late arrival results in a higher price volatility. Therefore agents tend to trade more often in the middle of the period. This shows that in the IEL-algorithm traders learn not to submit risky strategies, resulting from the algorithm that considers only the performance of strategies in the previous period. A possible future research subject would be to adjust the IEL-algorithm in such a way that the average profit over multiple periods is considered. This may result in a more realistic outcome and traders may prefer to enter at the beginning or the end of the period.

Allowing traders to submit at their preferred moment has a negative effect; the expected efficiency and the expected number of trades decrease significantly and the expected price volatility significantly increases. Hence, allowing traders to make this extra decision results in a lower expected profit. It is optimal not to allow traders this extra decision.

When the size of the market increases, the probability of trading and the probability that trade occurs at the price of the own offer change. The larger the size of the market, the earlier traders submit their order. It appears that the moment of submission will converge to zero as the size of the market converges to infinity. Irrespective of the submission moment traders on average submit a higher bid and a lower ask as the size of the market increases. We conclude that the size of the market is of great importance to the distribution of submission moments.

Ceteris paribus as competition increases in some sense, the probability of trading decreases. Intramarginal traders are hence forced to trade earlier and to submit less aggressive offers to cope with the increased competition, which increases the probability of trading. Extramarginal traders have a clearer preference for their submission moment when they have more opportunity to trade, in which case they submit early to outcompete others.

We found that under the Individual Evolutionary Learning algorithm investors in a medium size Continuous Double Auction market learn to submit their order around the middle of the trading period to avoid a lower trading probability or profit. The earlier traders submit the more aggressive offer they submit and thus aim for a higher profit. In a large market the latter effect reduces and traders submit earlier and earlier. Moreover, we have shown how the distribution of submission moments and the expected offer as a function of the submission moment change with the amount of competition and the size of the market.

Bibliography

N. Alon, and J.H. Spencer (2008), *The Probabilistic Method*, John Wiley & Sons.

M. Anufriev, J. Arifovic, J. Ledyard and V. Panchenko (2013). "Efficiency of continuous double auctions under individual evolutionary learning with full or limited information", *Journal of Evolutionary Economics* **23**, 539-573.

M. Anufriev and V. Panchenko (2009). "Asset prices, traders' behavior and market design", *Journal of Economic Dynamics and Control* **33**, 1073-1090.

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

J. Arifovic and J. Ledyard (2004), "Scaling up learning models in Public Good games", *Journal of Public Economic Theory* **6**, 203-238.

J. Arifovic and J. Ledyard (2007), "Call market book information and efficiency", *Journal of Economic Dynamics and Control* **31**, 1971-2000.

K. Bae, H. Jang and K.S. Park (2003), "Traders' choice between limit and market orders: evidence from NYSE stocks", *Journal of Financial Markets* **6**, 517-538.

B. Biais, P. Hillion and C. Spatt (1995), "An empirical analysis of the limit order book and the order flow in the Paris Bourse", *Journal of Finance* **50**, 1655-1689.

J. Blasiak and R. Durrett (2005), "Random Oxford Graphs", *Stochastic Processes and their Applications* **115-8**, 1257-1278.

R. Bloomfield, M. O'Hara and G. Saar (2005), "The "make or take" decision in an electronic market: Evidence on the evolution of liquidity", *Journal of Financial Economics* **75**, 165-199.

L. Blume, D. Easley, J. Kleinberg and É. Tardos (2009), "Trading networks with price-setting sgents", *Games and Economic Behavior* **67**, 36-50.

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

B. Bollobás (1982), *Annals of Discrete Mathematics: Graph Theory*, North-Holland Publishing Company.

G. Bottazzi, G. Dosi and I. Rebesco (2005), "Institutional architectures and behavioral ecologies in the dynamics of financial markets", *Journal of Mathematical Economics* **41**, 197-228.

W.A. Brock and C.H. Hommes (1997), "A rational route to randomness", *Econometrica* **65**, 1059-1096.

W.A. Brock and C.H. Hommes (1998), "Heterogeneous beliefs and routes to chaos in a simple asset pricing model", *Journal of Economic Dynamics and Control* **22**, 1235-1274.

A. Calvó-Armengol (2001), "Bargaining power in communication networks", *Mathematical Social Sciences* **41**, 69-87.

T.N. Cason and D. Friedman (1996), "Price formation in double auction markets", *Journal of Economic Dynamics and Control* **20**, 1307-1337.

R. Cervone, S. Galavotti and M. LiCalzi (2009), "Symmetric equilibria in double auctions with markdown buyers and markup sellers", in: C. Hernandez, M. Posada and A. Lopez-Paredes, Artificial Economics, Springer, 81-92.

K. Chatterjee and B. Dutta (1998), "Rubinstein auctions: On competition for bargaining partners", *Games and Economic Behavior* **23**, 119-145.

K. Chatterjee and W. Samuelson (1983), "Bargaining under incomplete information", *Operation Research* **31-5**, 835-851.

C. Chiarella and G. Iori (2002), "A simulation analysis of the microstructure of double auction markets", *Quantitative Finance* **2**, 346-353.

K.H. Chung, B.F. van Ness and R.A. van Ness (1999), "Limit orders and the bid-ask spread", *Journal of Financial Economics* **53**, 255-287.

M. Corominas-Bosch (2004), “Bargaining in a network of buyers and sellers”, *Journal of Economic Theory* **115**, 35-77.

H. Dawid (1999), “On the convergence of genetic learning in a double auction market”, *Journal of Economic Dynamics and Control* **23**, 1545-1567.

D. Easley and J. Kleinberg (2010), *Networks, crowds and markets*, Cambridge University Press.

P. Erdős and A. Rényi (1960), “On the evolution of random graphs”, *Mathematical Institute of the Hungarian Academy of Sciences* **5**, 17-61.

P. Erdős and A. Rényi (1961), “On the strength of connectedness of a random graph”, *Acta Mathematica Academiae Scientiarum Hungarica* **12**, 261-267.

E. Even-Dar, M. Kearns and S. Suri (2007), “A network formation game for bipartite exchange economies”, *SODA ’07 Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, 697-706.

S. Fano and P. Pellizzari (2011), “Time-dependent trading strategies in a continuous double auction”, in: Lecture Notes in Economics and Mathematical Systems **652**, Springer Berlin-Heidelberg, 165-176.

S. Fano, M. LiCalzi and P. Pellizzari (2013), “Convergence of outcomes and evolution of strategic behavior in double auctions”, *Journal of Evolutionary Economics* **23-3**, 513-538.

D. Friedman (1991), “A testable model of double auction markets”, *Journal of Economic Behavior and Organization* **15**, 47-70.

D. Gode and S. Sunder (1993), “Allocative efficiency of markets with zero-intelligence traders: market as a partial substitute for individual rationality”, *Journal of Political Economy* **101**, 119-137.

D. Gode and S. Sunder (1997), “What makes markets allocationally efficient?”, *The Quarterly Journal of Economics* **112**, 603-630.

M.D. Gould, M.A. Porter, N. Hautsch, S. Williams, M. McDonald, D.J. Fenn and S.D. Howison (2013), “Modelling limit order books with bilateral trading agreements”, manuscript.

M.D. Gould, M.A. Porter, S. Williams, M. McDonald, D.J. Fenn and S.D. Howison (2013), "Limit order books", arXiv:1012.0349.

A. Hachmeister (2007), *Informed traders as liquidity providers: evidence from the German equity market*, Der Deutsche Universitäts-Verlag.

P. Hall (1935), "On Representatives of Subsets", *Journal of the London Mathematical Society* **10-1**, 26-30.

M.O. Jackson and A. Wolinksy (1996), "A strategic model of social and economic networks", *Journal of Economic Theory* **71**, 44-74.

S. Janson, T. Luczak and A. Rucinski (2000), *Random Graphs*, John Wiley & Sons.

E. Kranton and D. Minehart (2001), "A theory of buyer-seller networks", *The American Economic Review* **91-3**, 485-508.

D. Ladley and P. Pellizzari (2014), "The simplicity of optimal trading in order book markets", in: R. Dieci, X.-Z. He and C. Hommes (eds.), Economic Dynamics and Financial Modelling, Springer, 183-199.

M. LiCalzi and P. Pellizzari (2006), "The allocative effectiveness of market protocols under intelligent trading", in: C. Bruun (ed.), Advances in Artificial Economics, Springer, 17-29.

M. LiCalzi and P. Pellizzari (2007), "Simple market protocols for efficient risk-sharing", *Journal of Economic Dynamics and Control* **31**, 3568-3590.

M. Mucha and P. Sankowski (2004), "Maximum Matchings via Gaussian Elimination", *Proc. 45th IEEE Symp. Foundations of Computer Science*, 248-255.

R. Myerson and M. Satterthwaite (1983), "Efficient mechanisms for bilateral trading", *Journal of Economic Theory* **29**, 265-281.

C.A. Parlour and D.J. Seppi (2008), "Limit order markets: a survey", in: Handbook of financial intermediation and banking, Elsevier North-Holland, 63-96.

A. Polanski (2007), "Bilateral bargaining in networks", *Journal of Economic Theory* **134**, 557-565.

I. Rosu (2009), “A dynamic model of the limit order book”, *Review of Financial Studies* **22**, 4601-4641.

D. Spulber (2006), “Firms and networks in two-sided markets”, in: Handbooks in Information Systems, Elsevier, 137-200.

D.B. West (1999), *Introduction to Graph Theory*, Prentice Hall.

R. Yamamoto and B. LeBaron (2010), “Order-splitting and long-memory in an order-driven market”, *The European Physical Journal B* **73**, 51-57.

W. Zhan and D. Friedman (2007), “Markups in double auction markets”, *Journal of Economic Dynamics and Control* **31**, 2984-3005.

Summary

The behaviour of traders has previously been studied extensively in different market designs. The setup of a market contains the information available to traders, the decisions traders have to make and the trading mechanism. Markets over networks, where transactions may occur between connected traders, have been studied mainly under full information about the network structure. In many agent-based models, for instance on Continuous Double Auctions, traders submit orders at random moments during a period under full or limited information about trading history. In both situations, traders simply have to select the optimal deterministic offer.

To study the effect of the design of the market, we have extended these models in this dissertation. In markets over networks we have introduced randomness and in equilibrium we have derived bounds on the maximal efficiency given the network structure. Moreover, under strategic behaviour of traders, we derived the effect of the available information about the network structure on the expected allocative efficiency. This effect depends also on the information about traders' valuations. We studied an alternative payoff function used in the Evolutionary Individual Learning algorithm under a Continuous Double Auction. Furthermore we extended this model by allowing traders to submit a two dimensional decision; their order and their preferred submission moment during the period, and studied the distribution of these moments. We compared with the original model to conclude whether it is optimal to allow traders this extra decision.

In Chapter 2 random bipartite networks are considered, similar to Erdős and Rényi (1960, 1961), where links between buyers and sellers are realised independently from each other with an equal

probability. We considered a market over such a network, which models the foreign spot exchange market. For infinitely large networks we derived phase transitions, where the structure of the network changes abruptly. When the expected number of links per trader converges to zero, the network almost surely consists of isolated spanning trees, which connect a subset of traders of the graph but do not contain any cycle. We show that a remarkable change in structure occurs when the expected number of links per trader crosses the threshold value one. The structure of the network changes from a collection of relatively small spanning trees, to a network that contains a giant central market. As the expected number of links per trader converges to infinity, almost every trader is contained in the giant market. We derive bounds on maximal efficiency given the network structure, and improve these bounds in the phases where almost every trader is part of a spanning tree, by studying the number of traders that have more than one connected trader.

Chapter 3 extends this setup by considering the efficiency reduction in equilibrium due to strategic behaviour, under different information sets about the network structure. In a thin Erdős-Rényi market with two buyers and two sellers a trading mechanism is used that allows preferred trades to occur, not necessarily the trades that construct a globally optimal allocation. We have compared three ordered information sets about the network structure; no, partial and full information. Under no information traders only know the probability of a link, under partial information the existence of own links is revealed and under full information the entire network structure is known to all traders. Under complete information about traders' valuations and costs, partial information is weakly dominated and hence it is optimal if either everything or nothing of the network structure is revealed to traders. Under incomplete information about valuations and costs we have found that no and full information lead to a comparable efficiency, assuming that traders use markup strategies. Partial information dominates strongly, since volatility of strategies under full information is higher and under no information traders offer more aggressively. Thus under incomplete information about valuations, it is optimal if traders know the existence of the own links, but not of the links of other traders. We can conclude that the quantity of information about the network structure that is available to traders, has a non-monotonic effect on allocative efficiency. Switching from complete to incomplete information about valuations and costs reverses the shape of this non-monotonicity.

In Chapter 4 the role of the information about the trading history that is available to traders is studied in a Continuous Double Auction market. Traders use the Individual Evolutionary Learning algorithm to determine their offer for the next period, based on the hypothetical payoff in the previous period. We introduced a new hypothetical payoff function when only information about past average prices is available, that uses more of the available information. We have shown that during the learning phase and in equilibrium, the efficiency and the number of transactions are higher than under full information about the trading history. Moreover, the price volatility is lower. This comparison of efficiency is in line with the results of Arifovic and Ledyard (2007) in a Call Market. However, when only past average prices are known the behaviour found is quite different than in Anufriev et al. (2013); instead of a divergence of offers, some convergence occurs. This behaviour is in line with Fano et al. (2013), who show that traders behave as pricemakers when only past profits and average prices are available. Moreover, we have found that these results are robust with respect to the size of the market and the number of units that agents desire to trade. Under the introduced hypothetical payoff function we have found that more information about the trading history leads to a higher price volatility and a lower efficiency and number of transactions.

Chapter 5 studies the timing of order submission. The Individual Evolutionary Learning algorithm is extended by requiring traders to make a two-dimensional decision: to choose the offer and the moment of submitting this offer. We have found that traders in a medium size market learn to submit their order around the middle of the period to balance the probability of trading and the expected profit from trade. Moreover, early submitted offers are more aggressive to gain a higher profit if trade occurs. Offers that are submitted late are less aggressive in order to increase the probability of trading. As a result, submitting early or late results in a higher expected profit, but respectively also in a higher risk of not trading or a higher price volatility. Traders learn to trade in the middle of the period, showing that in the IEL-algorithm traders learn not to select risky strategies. We showed that it is optimal not to allow traders to submit their offer at their preferred moment, since this results in a lower expected efficiency and a higher expected price volatility. As the size of the market or competition between traders increases, traders learn to submit their offer earlier and to submit a more conservative offer.

A general conclusion of this dissertation is that market design has a large impact on allocative efficiency. In random Erdős-Rényi markets the information about the network structure that is available to traders has a non-monotone effect on efficiency. This non-monotonicity is opposite under complete and incomplete information about traders' valuations and costs. In a Continuous Double Auction, information about the trading history reduces expected efficiency when traders use the Individual Evolutionary Learning algorithm. Allowing traders to choose their submission moment has a negative effect on allocative efficiency.

Samenvatting (Summary in Dutch)

Het gedrag van handelaren is in de literatuur uitgebreid onderzocht in verschillende marktdesigns. De opzet van een markt bevat de beschikbare informatie voor handelaren, de beslissingen die handelaren moeten nemen en het handelsmechanisme. Markten over netwerken, waarin transacties kunnen optreden tussen verbonden handelaren, zijn voornamelijk bestudeerd bij volledige informatie over de structuur van het netwerk. In veel agent gebaseerde modellen, bijvoorbeeld voor Continuous Double Auctions, plaatsen handelaren biedingen op willekeurig momenten gedurende een periode, met volledige of met beperkte informatie over de handelshistorie. In beide situaties hoeven handelaren dan alleen het optimale deterministische bod te selecteren.

Om te onderzoeken wat het effect is van het design van de markt, hebben we deze modellen in dit proefschrift uitgebreid. In markten over netwerken hebben we onzekerheid geïntroduceerd en in het evenwicht grenzen voor de maximale efficiëntie gegeven de structuur van het netwerk afgeleid. Bovendien bekeken we, bij strategisch gedrag van handelaren, het effect van de informatie die beschikbaar is over de structuur van het netwerk op de verwachte efficiëntie. Dit effect is ook afhankelijk van de informatie over de waarderingen van handelaren. We bestudeerden een alternatieve winstfunctie die gebruikt wordt in het Evolutionaire Individuele Leer algoritme, in een Continuous Double Auction. Dit model hebben we verder uitgebreid door handelaren een tweedelige beslissing voor te leggen; hun bod en het door hen geprefereerde moment tijdens de periode om dit bod te plaatsen, en bestudeerden de verdeling van het moment van plaatsen. We vergeleken onze resultaten met het oorspronkelijke model om te onderzoeken of het optimaal is om handelaren deze extra beslissing te laten nemen.

In Hoofdstuk 2 worden stochastische bipartiete netwerken beschouwd, vergelijkbaar met Erdős and Rényi (1960, 1961), waarbij connecties tussen kopers en verkopers onafhankelijk van elkaar worden gerealiseerd met dezelfde kans. We hebben een markt over een dergelijk netwerk onderzocht, hetgeen de spotmarkt voor buitenlandse valuta modelleert. Voor oneindig grote netwerken hebben we faseovergangen afgeleid, waarbij de structuur van het netwerk abrupt verandert. Als het verwachte aantal connecties per handelaar naar nul convergeert, bestaat het netwerk vrijwel zeker uit geïsoleerde opspannende bomen, die deelverzamelingen van handelaren verbinden maar geen cycli bevatten. We hebben aangetoond dat er een opmerkelijke verandering in de structuur optreedt wanneer het verwachte aantal connecties per handelaar de waarde één overschrijdt. De structuur van het netwerk verandert van een verzameling van relatief kleine opspannende bomen, naar een netwerk dat één grote centrale markt bevat. Wanneer het verwachte aantal connecties per handelaar naar oneindig convergeert, maakt bijna elke handelaar onderdeel uit van de grote markt. We leidden grenzen voor de maximale efficiëntie gegeven deze structuur van het netwerk af en verbeterden deze grenzen in de fasen waar bijna elke handelaar onderdeel uitmaakt van een opspannende boom, door het aantal handelaren dat meer dan één aangesloten handelaar heeft te bestuderen.

In Hoofdstuk 3 wordt dit model uitgebreid door in het evenwicht te kijken naar de efficiëntievermindering ten gevolge van strategisch gedrag, bij verschillende aannames met betrekking tot de informatie over de structuur van het netwerk. In een Erdős-Rényi markt met twee kopers en twee verkopers wordt een handelsmechanisme gebruikt waarbij de geprefereerde transacties plaatsvinden, niet per se de transacties die tot een sociaal optimale allocatie leiden. We hebben drie geneste informatieverzamelingen over de structuur van het netwerk vergeleken; geen, partiële en volledige informatie. Bij geen informatie kennen handelaren alleen de kans op een connectie, bij partiële informatie zijn alleen de eigen connecties bekend en bij volledige informatie is de hele netwerkstructuur bekend bij alle handelaren. Bij complete informatie over waarderingen en kosten van handelaren, wordt partiële informatie zwak gedomineerd en dus is het optimaal indien ofwel alles ofwel niets van de netwerkstructuur bekend wordt gemaakt aan handelaren. Onder incomplete informatie over waarderingen en kosten leiden geen en volledige informatie tot een vergelijkbare efficiëntie, onder de veronderstelling dat handelaren gebruik maken van zogenaamde opslagstrategieën. Partiële informatie domineert sterk, omdat enerzijds de volatiliteit van de strategieën bij volledige informatie hoger is en anderzijds handelaren agressiever bieden bij geen informatie. Onder incomplete informatie over waarderingen, is het optimaal als handelaren de eigen connecties kennen, maar niet de connecties van andere handelaren. We concluderen dat de hoeveelheid informatie over de structuur van het netwerk die beschikbaar is voor handelaren, een niet-monotoon effect heeft op de efficiëntie. Veranderen van complete naar incomplete informatie over waarderingen en kosten, leidt tot een omkering van deze niet-monotoniciteit.

In Hoofdstuk 4 is de rol van de informatie over de handelshistorie die beschikbaar is voor handelaren onderzocht in een Continuous Double Auction markt. Handelaren gebruiken het Individuele Evolutionaire Leer algoritme om hun bod voor de volgende periode te bepalen, aan de hand van de hypothetische winst in de voorgaande periode. We introduceerden een nieuwe hypothetische winst functie die meer informatie in een Continuous Double Auction markt gebruikt, als uit het verleden alleen informatie over gemiddelde aandelenprijzen beschikbaar is. We hebben aangetoond dat tijdens de leerfase en in het evenwicht, de efficiëntie en het aantal transacties significant hoger zijn dan onder complete informatie over de handelshistorie. Bovendien is gebleken dat de prijsvolatiliteit lager is. Deze vergelijking van efficiëntie komt overeen met de resultaten van Arifovic and Ledyard (2007) in een Call Market. Wanneer uit het verleden louter gemiddelde prijzen bekend zijn, leidt deze nieuwe winstfunctie tot ander gedrag dan onder de oude winstfunctie, die bestudeerd is in Anufriev et al. (2013); in plaats van een divergentie van biedingen, vindt enige convergentie plaats. Dit gedrag komt overeen met Fano et al. (2013), die laten zien dat handelaren proberen de transactieprijzen te beïnvloeden wanneer alleen historische winsten en gemiddelde aandelenprijzen bekend zijn. Onze resultaten bleken robuust met betrekking tot de omvang van de markt en het aantal eenheden dat agenten willen verhandelen. Gegeven de geïntroduceerde hypothetische winstfunctie hebben wij geconstateerd dat meer informatie over de handelshistorie leidt tot een hogere prijsvolatiliteit en een lagere efficiëntie en aantal transacties.

Hoofdstuk 5 beschrijft de keuze van het moment om een bod in te doen. Het Individuele Evolutionaire Leer algoritme wordt uitgebreid door van handelaren te vragen om een tweedelige beslissing te nemen: het bod zelf en het moment om dit bod te plaatsen. Handelaren in een middel grote markt leren om hun bod rond het midden van de periode te plaatsen, om de kans op een transactie en de verwachte winst uit een transactie tegen elkaar af te wegen. Vroeg geplaatste biedingen zijn agressiever om een hogere winst te behalen. Aanbiedingen die laat ingediend worden zijn minder agressief om de kans op een transactie te verhogen. Als gevolg hiervan, leidt het vroeg dan wel laat plaatsen van een bod tot een hogere verwachte winst, maar respectievelijk ook tot een hoger risico op het uitblijven van een transactie of een hogere prijsvolatiliteit. Handelaren leren om hun bod in het midden van de periode te plaatsen, hetgeen laat zien dat het leeralgoritme er toe leidt dat handelaren leren om de risicovolle strategieën niet te selecteren. Het is dus optimaal om handelaren niet toe te staan om te kiezen wanneer zij hun bod plaatsen, aangezien dit leidt tot een lagere verwachte efficiëntie en een hogere prijsvolatiliteit. Wanneer de omvang van de markt of concurrentie tussen handelaren toeneemt, leren handelaren om conservatiever te bieden en om hun bod eerder in te dienen.

Een algemene conclusie van dit proefschrift is dat het marktdesign een grote impact op de efficiëntie heeft. In willekeurige Erdős-Rényi markten heeft de informatie over de structuur van het netwerk die beschikbaar is voor de handelaren een niet-monotoon effect op de efficiëntie. Deze niet-monotoniciteit is precies omgekeerd als we de gevallen met complete en incomplete informatie over waarderingen en kosten van handelaren vergelijken. In een Continuous Double Auction vermindert informatie over de handelshistorie de verwachte efficiëntie als handelaren gebruik maken van het Individuele Evolutionaire Leer algoritme. Toestaan om handelaren te laten kiezen wanneer zij hun bod plaatsen heeft een negatief effect op de efficiëntie.