LIQUIDITY TRADING ON STOCK MARKETS: DETERMINANTS OF THE HUMPED SHAPE OF THE ORDER BOOK

Dániel Havran, István Margitai and Balázs Árpád Szűcs
Department of Finance*
Corvinus University of Budapest
1093, Budapest, Fővám tér 8, Hungary
E-mail: balazsarpad.szucs@uni-corvinus.hu

April 17, 2012

KEYWORDS

Market liquidity, market microstructure, order book simulation, genetic algorithm

ABSTRACT

In this paper we examine pure limit order markets in order to find out how different trading strategies affect the shape of the order book. We concentrate on liquidity trading, which implies that any new information that would influence the fundamental value of the traded asset is excluded. The humped shape of the order book is a stylized fact studied in the literature. We build a simulation model to explain the position of this hump in a setting where liquidity takers and providers submit buy and sell orders to complete their exogenous trading needs. We show that under the above assumptions the exogenous impatience factor of liquidity traders determines the position of the hump.

1 INTRODUCTION

On electronic stock markets, market actors submit buy or sell orders into the order book. The limit order volume structures of the offered/asked prices form a humped shape. An example is shown in Figure 1.1. We model market microstructure in order to explore the determinants of stock exchange order book shapes.

The order book itself is the collection of specific buy and sell *limit* orders, submitted by patient market players, defined as Liquidity Providers (LP). The other type of market players are the impatient Liquidity Takers (LT), who are submitting buy and sell *market* orders. Transactions always consist of a market and a limit order matched to one another. The two types of players are making money on the expense of each other: transactions close to the mid-price (the average of the best bid and ask prices) favour LTs, whereas transactions far from the mid-price favour LPs. It follows that there is a trade-off between trading immediately (i.e. being impatient) and reaching a better price, which trade-off is often referred to as price impact or market impact (for a precise definition of market impact, where order volume is also included, see Bouchaud et al. (2002)).

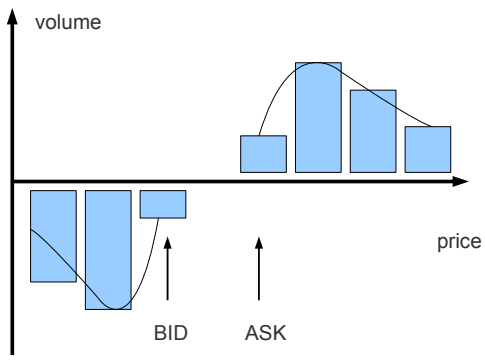

Figure 1.1: Humped shape of the order book

The limit order market literature is fairly recent, it has less than two decades of history. The most comprehensive work that introduces the limit order markets and connects liquidity to the economic literature is Parlour and Seppi (2007). The most commonly known statistical analysis is provided by Bouchaud et al. (2002), where the shape of the average order book is described. This shape has already been tested on different markets. For example, Gu et al. (2008) investigates the shape of the order book on Chinese stock markets in their empirical study. With a limit order market model, Rosu (2009) explains how the humped shape is formulated by liquidity takers and providers. He also makes suggestions to price formation. Our contribution is related to his results, we also suppose symmetric information relations among the market players. In contrast to Rosu (2009), we focus on the heterogeneity of impatience.

*This conference material was prepared with the support of the TÁMOP-4.2.2/B-10/1-2010-0023 project.

Proceedings 26th European Conference on Modelling and Simulation ©ECMS Klaus G. Troitzsch, Michael Möhring, Ulf Lotzmann (Editors)
ISBN: 978-0-9564944-4-3 / ISBN: 978-0-9564944-5-0 (CD)

We are reasoning that the humped shape of the order book can be derived from the liquidity takers' impatience, together with the heterogeneity of the liquidity providers' market expectations. In order to exclude the speculation on fundamental value, so that we could concentrate only on liquidity motives, we assume that there is no new information on asset values during the observed period. We suppose exogenously defined LTs, while the LPs' order strategies are continuously evolving over time according to a genetic algorithm.

So building on the stylized fact and the common behavioural patterns reported in the literature, we construct a model of market microstructure aiming to explain the dynamic evolution of the above mentioned hump shape in the order book. The following hypothesis will be tested in this paper.

Hypothesis: Assuming that LPs are following a genetic algorithm to adopt their strategies to the exogenous and fixed strategies of the LTs, then the position of the hump in the order book is determined by the aggregate strategy (through the impatience) of LTs.

The rest of the paper is organized as follows. The next chapter introduces the market players in detail. The third section investigates the simulation process. Section four describes our main results, while providing some further analyses based on our model. The last section concludes.

2 MARKET PLAYERS

We only model liquidity trading, therefore we must exclude all other motivations to trade, especially the arrival of new fundamental information on the asset. This results in a setting where all players stick to their initial plans concerning the direction of trades (buy or sell) as well as the overall amount they wish to trade during the simulation. As a result, there are four different types of players in our model, as shown in Table 1. Types 1 and 2 are LPs, who only submit limit orders, while types 3 and 4 are LTs, who only use market orders.

	Liquidity Provider	Liquidity Taker
Buy	1	3
Sell	2	4

Table 1: Player types

The core motivation of any player in this model (LPs and LTs alike) is the same: they wish to trade their full predefined volumes and reach the best possible volume weighted average prices while doing so. However the way they reach their goals is different. The following two subsections elaborate on the behaviour of LPs and LTs.

Liquidity Providers

Every LP has an idea (a forecast) on what is going to happen in terms of trades in the subsequent trading period. Namely they all have a guess on the price level that is furthest from the mid-price but still will be used. They base their limit order submission strategies on this forecast. On the one hand, if they expect this level to be close to the mid-price, they will also put their orders at closer price levels, because otherwise a considerable amount of their orders would remain unmatched. On the other hand, if they expect this price level to be further from the mid-price, they will then submit their orders at further levels, because they have a reason to believe these orders will still be matched. Thus they will have improved the volume weighted average price of their own trades.

The technical interpretation of this behaviour is the following. Initially every LP receives a random binary string that codes their unique expectations. This string is then converted into a scalar through equation 2.1 (similarly to Lettau (1997)) as follows:

$$\lambda = l\frac{\sum_{j=1}^{k} s_j 2^{j-1}}{2^k - 1}, \tag{2.1}$$

where λ is the above mentioned price level they are expecting, while l denotes the number of available price levels on one side, k is the length of the string, and s_j is the jth character of the string.

LPs then determine the volumes they submit on each price level using the following function form:

$$v_i(\lambda) = \frac{\frac{\lambda^i e^{-\lambda}}{i!}}{\sum_{i=1}^{l} \frac{\lambda^i e^{-\lambda}}{i!}} V, \tag{2.2}$$

where v_i denotes the volume submitted on the ith price level, and V is the latent need for transaction (inventory) of the LP. The λ is a scalar defined by equation 2.1.

Using this setup implies that each LP will submit some orders to all available price levels. The aggregate submissions of LPs shape up the order book.

Liquidity Takers

LTs differ in their levels of impatience/aggression, but they all base their market order submission strategies on the actual state of the order book. Their primary goal is also to trade, and depending on their level of impatience they are willing to clear more or less price levels, that is willing to bare more or less liquidity costs in exchange for quicker execution (although market orders are executed immediately, more aggressive players who clear more price levels trade larger volumes, hence fulfill their own trading need more quickly). Those who are very patient may sometimes choose not to trade at all in that cycle given the actual limit orders.

The technical interpretation of this behaviour is as follows. All LTs have an initial γ that is constant through-

out the simulation. This is their unique factor of impatience, that determines the worst volume weighted average price (VWAP) that they are willing to achieve while trading. Generally, if there are l price levels, for the ith LT $1 \leq VWAP_i \leq l$ holds. Each LT faces a limit of $VWAP_i \leq \gamma_i$, but if they have enough inventory left to trade, they will hit the γ limit. The smaller this γ value, the more patient the LT is, not willing to clear many price levels, being ready to wait instead. Larger γ means the LT urges the trade by letting the VWAP of their own trade go further away from the mid-price, that is clearing more price levels and thus trading the full amount in fewer rounds.

This strategy is obviously conditional on the actual state of the order book. The same γ may result in very different trading volumes depending on the depth of the book on each price level.

Note that LTs start trading at the best price levels and go towards worse ones as the better levels are cleared.

3 SIMULATION PROCESS

We used the Matlab/Octave environment to run our simulation. The simulation setup consists of cycles and phases as shown in Figure 3.1. The remaining of this section elaborates on this setup.

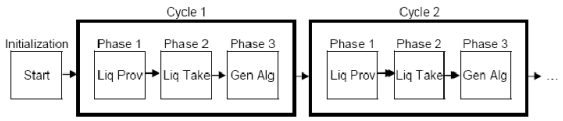

Figure 3.1: Structure of simulation

Initialization

This is the starting phase of each run that sets the initial parameters of the simulation. It is important to separate this phase from the cycles, because thus we are able to set the same initial parameters across different runs if necessary. The parameters that must be set are (1) the random strings for each LP (e.g. 1001101), (2) the random γ for each LT, (3) the number of each player type listed in Table 1, (4) the inventory of each player, which equals the latent volume they try to trade, and (5) the number of available price levels on each side of the mid-price.

Cycles

The series of cycles is the model of the trading period. Trading is continuous as the cycles follow each other. Any inventory that runs out is refilled at the beginning of the next cycle.

Phase 1 - Liquidity providers

This is the phase of the liquidity providers. When Phase 1 begins, there is an existing order book that remained from the previous cycle (except in Cycle 1). LPs are allowed to revise their previous quotes, that is they can cancel and replace those that were not matched in the previous cycle. It is achieved in model by starting Phase 1 with the cancellation of all previous orders, thus LPs face a blank order book in Phase 1.

Given that our goal is to model the evolution of the shape of the order book, it follows that the price formation falls beyond the scope of our interest in this model. This is why we continue Phase 1 by setting the mid-price to a constant. The bid-ask spread is also fixed, which means that the best prices are a constant distance away from the mid-price in both directions. All price levels beyond the best prices follow each other by one tick.

All LPs act simultaneously, which implies that their strategies exclude the response to the behaviour of each other. The result of their aggregate behaviour is the order book itself. We model this by making them all decide their submission strategies at the beginning of the phase, and then follow each other one by one in a random order (uniform distribution) ignoring the submissions of the other LPs in the same phase.

It is important to mention that the matching mechanism of the order book is First-In-First-Out (FIFO). This means that limit orders closer to the mid-price have priority when a market order arrives, and because limit orders have a timestamp, on the same price level those with an earlier timestamp have priority. This is why the random order is important in order to model the diversity in submission times.

By the end of Phase 1, the order book is built.

Phase 2 - Liquidity takers

This is the phase of the liquidity takers. They condition their order submission strategies on the actual state of the order book, in consideration of their own γ. They implicitly also consider the submissions of other LTs, given that the order book changes as a result of their actions. This is captured in our model in a way where within the same phase, they follow each other in a uniformly distributed random order. This implies that they do not face the full order book, only the remainders of it after the quicker LTs' trades.

Let us suppose for example an LT with $\gamma = 1$. This means that $VWAP = 1$ is the only acceptable scenario, hence no trade beyond the first price level is possible. In a situation where this LT is not within the first few ones to trade, it is highly probable that the first price level will be long gone by the time this LT takes its turn, it will therefore decide not to trade.

By the end of Phase 2, all transactions took place.

Phase 3 - Genetic algorithm

This is the phase of the genetic algorithm, and it is the LPs that evolve genetically. The strings of those LPs who had a better guess on the λ will survive with higher probabilities into the next cycle, because a better guess will clearly result in statistically more favorable volume weighted average prices. We use a fitness function to rank the LPs based on the success of their preliminary guesses as shown in equation 3.1.

$$f(\lambda, l, \Omega) = l - |\lambda - \sum_{i=1}^{l} \frac{\sum_j v_{i,j}^1 - \sum_j v_{i,j}^2}{\sum_j v_{i,j}^1}|, \qquad (3.1)$$

where f is the fitness value of the LP (the higher, the better) and l is the number of price levels on one side of the book. i stands for price levels, j for LPs, v^1 and v^2 are the volumes in the book at the end of Phases 1 and 2. The fitness is the highest if they got the λ correctly, and gets smaller as their guess deteriorates.

The genetic algorithm itself has two stages: crossover and mutation. The first one is the crossover stage, where a random h number is selected, and the strings are cut after the hth character. Then the strings are mixed up in pairs along this h cutoff with probability p_c. The second stage is mutation, where random characters of random strings are changed with probability p_m. For illustration of the two stages of the genetic algorithm see figure 3.2. The LPs thus evolve, and those that are more successful survive with higher probabilities.

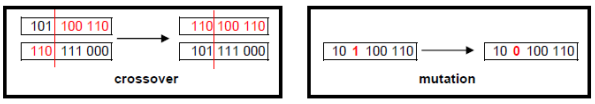

Figure 3.2: Genetic algorithm

For a better understanding of the mechanism behind the genetic algorithm driven LP behaviours, we tested the effects of LP mutations. The two factors of the evolution are the mutation and the crossover probabilities. When the crossover probability is high, there are big jumps in the time series of λ, until such a level is reached, where all strings are identical, hence mutation has no effect.

The mutation causes noisy changes. The mechanism of mutation is illustrated in Figure 3.3, where we plotted the average λ values of LPs against the simulation cycle count. The two graphs represent the zero and the positive (but small) probability cases of mutation.

4 RESULTS

First we set up three alternative simulation environments with different parameter sets regarding the LTs' impatience and the LPs' evolution. Second, we analyze the behaviour of the two kinds of players with random initial variables.

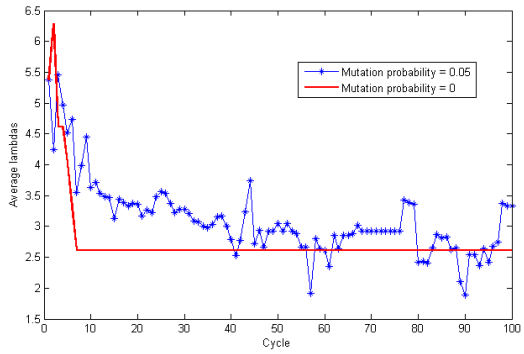

Figure 3.3: Mutation effects

We calibrated the model as follows. At the initial state, the λ of LPs is allocated from 0.5 to 9.5 in steps of one. We chose $l = 10$ as the number of price levels on one side. LPs are therefore inhomogeneous, each of them having different initial values. The setup so far corresponds to a very controlled environment with special assumptions on the initial λ parameters. The probability of the crossover in a given cycle is twenty percent, and the probability of mutation is five percent in each run.

The γ value of LTs, therefore their level of impatience is homogeneous in all the three runs, with values of 1.5 in the first run, 3.5 in the second and finally 5.5 in the third. For an overview of the different calibration setups, see Table 2.

Simulation	Liquidity Takers	Liquidity Providers		
	γ	λ_0	p_c	p_m
Run a)	1.5	0.5 ... 9.5	0.2	0.05
Run b)	3.5	0.5 ... 9.5	0.2	0.05
Run c)	5.5	0.5 ... 9.5	0.2	0.05

Table 2: Model calibrations

According to Run a)-c) the diversity of liquidity providers disappears with time, as the genetic algorithm in each running cycle improves their λ expectation. In case of Run a), after the cycle number 20, the peak of the curve is around the 1.5th price level, while in case of Run c) the hump in the order book locates around the 5th price level. In Run c) the LTs are much more aggressive as they are willing to trade at less favourable price. Hence, they clear approx. 5 price levels from the book. Figure 4.1 shows that the heterogeneous λ values are converging to the pre-defined and fix γ in each simulation.

We showed in special cases that the location of the hump in the order book is a result of a learning process of liquidity providers if liquidity takers are homogeneous. In Run a)-c) the humps after 20 cycles moved from the middle of the book to a price level of the LTs' γ.

Given that the above described results hold for arbitrary γ values, our hypothesis can be confirmed in the special setup of homogeneous LTs. Our next question was whether our hypothesis can also be confirmed for a

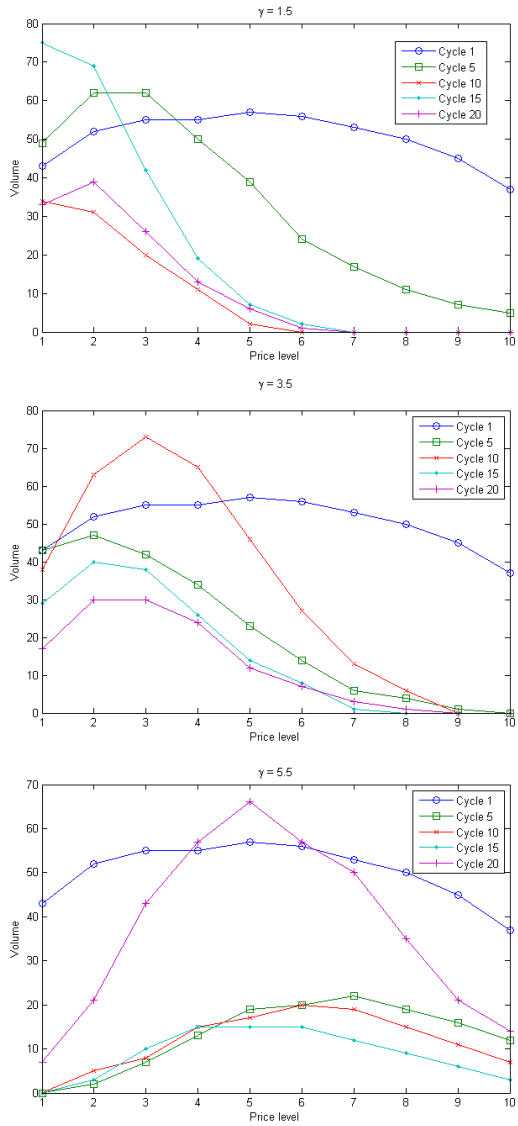

Figure 4.1: Runs a), b) and c)

more general case, where both LPs and LTs are heterogeneous.

In order to examine such a case, let each LT have a random γ impatience coefficient. After repeating the simulations 150 times, we plotted the result in Figure 4.2, which depicts how the LPs' expectations adjust to the LTs' impatience in random cases. Even after 30 running cycles in each simulation the average λ is close to the average impatience parameters of LTs. We note that the slope of the regression line on Figure 4.2 is close to one. The deviation of λ and γ points from the regression line can be reasoned by the mutation in the genetic algorithm. It can randomly improve or deteriorate the LPs' λ from the LTs' average γ in each cycle.

The deviation of the slope of regression line from one can be explained as follows. As the inventories of agents (the volume that the LTs and LPs intend to trade in a certain cycle) dynamically change, and both LTs and LPs are heterogeneous, it is not ensured that the number of

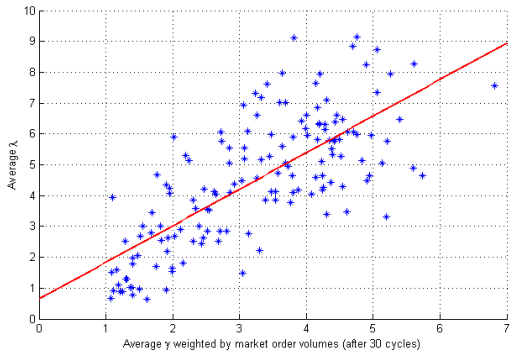

Figure 4.2: Average γ and λ

aggregately cleared price levels will be equal to the unweighted average of γ values. Hence, the LPs evolve in a way that they place the hump on a price level, which might be different from the average γ values. This is the reason why we plotted the evolved λ values against the *inventory weighted* average γ.

5 CONCLUSION

In this paper we model a pure limit order market with a heterogeneous set of players including liquidity takers and providers who buy and sell a single asset. The strategies of the liquidity takers are exogenous, whereas the strategies of the liquidity providers evolve according to a genetic algorithm that is conditioned on the behaviour of the liquidity takers. Knowing that such market microstructures tend to provide a humped shape order book, we run our simulations to show that within the above conditions, the position of the hump in the order book can be explained by the exogenous strategies (the impatience) of the liquidity takers.

The next step of further research could examine the determinants of the humped shape in a similar model, where the strategies of the liquidity takers are endogenous. Changing this single aspect may add interesting interactions to the model of the current paper.

REFERENCES

Bouchaud, J.-P., Mezard, M., and Potters, M. (2002). Statistical properties of stock order books: empirical results and models. Science & Finance (CFM) working paper archive 0203511, Science & Finance, Capital Fund Management.

Gu, G.-F., Chen, W., and Zhou, W.-X. (2008). Empirical shape function of limit-order books in the chinese stock market. *Physica A*, 387(21):5182–5188.

Lettau, M. (1997). Explaining the facts with adaptive agents: The case of mutual fund flows. *Journal of Economic Dynamics and Control*, 21:1117–1147.

Parlour, C. A. and Seppi, D. J. (2007). *Limit Order Markets: A Survey*. Handbook of Financial Intermediation & Banking.

Rosu, I. (2009). A dynamic model of the limit order book. *Review of Financial Studies*, 22(11):4601–4641.

AUTHOR BIOGRAPHIES

DÁNIEL HAVRAN holds an Assistant Professor position at the Department of Finance, Corvinus University of Budapest. He actively teaches Financial Markets, Corporate Finance and Credit Risk Management courses. His research interest is related to corporate financial flexibility, liquidity and credit risk. Earlier, he developed methodical tools and forecasting models for corporate liquidity and cash management systems. He obtained his PhD in Economics in 2011. His email is `daniel.havran@uni-corvinus.hu`.

ISTVÁN MARGITAI majored in Investments and Risk Management at the Corvinus University of Budapest. He currently works as an analyst at the Asset Liability Management Department of a Hungarian retail bank, where his main responsibility is strategic risk management. He joined the Liquidity Research Group of the Department of Finance, Corvinus University of Budapest in February 2011. His email is `margitai.i@gmail.com`

BALÁZS ÁRPÁD SZŰCS is a PhD candidate in Economics. He currently works as a Research Assistant at the Department of Finance, Corvinus University of Budapest. His research interest covers market microstructure, modeling and experimental games of financial microstructures. His email is `balazsarpad.szucs@uni-corvinus.hu`