UNIVERSITY OF LEEDS

Article:

Ladley, D, Lensberg, T, Palczewski, J et al. (1 more author) (2015) Fragmentation and stability of markets. Journal of Economic Behavior and Organization, 119. 466 - 481. ISSN 0167-2681

https://doi.org/10.1016/j.jebo.2015.09.013

Fragmentation and stability of markets

Daniel Ladley[a,1], Terje Lensberg[b], Jan Palczewski[c], Klaus Reiner Schenk-Hoppé[d]

[a]*Department of Economics, University of Leicester, United Kingdom. Email address: d.ladley@leicester.ac.uk*
[b]*Department of Finance, NHH–School of Economics, Norway. Email address: terje.lensberg@nhh.no*
[c]*School of Mathematics, University of Leeds, United Kingdom and Faculty of Mathematics, University of Warsaw, Poland. Email address: j.palczewski@leeds.ac.uk*
[d]*Economics, School of Social Sciences, University of Manchester, United Kingdom and Department of Finance, NHH–School of Economics, Norway. Email address: klaus.schenk-hoppe@manchester.ac.uk*

Abstract

Trading skills are highly rewarded in practice but largely ignored in theoretical models of financial markets. This paper demonstrates the importance of skills by examining their interaction with market fragmentation and market stability. We consider a computational model where traders' abilities to accurately price assets are endogenous. In contrast to models that do not consider skills, we find that centralising markets can lead to higher price volatility and less resilience to shocks because it increases the equilibrium proportion of unskilled traders.

Keywords: Skills; market fragmentation; volatility; market resilience.
JEL: D47; D83; G11.

1. Introduction

Skills command huge premia in the job market for professionals, particularly in the financial sector. The ability to identify mispriced assets and the

[1]Financial support by the Swiss National Center of Competence in Research 'Financial Valuation and Risk Management' (NCCR FINRISK) and the European Commission under the Marie Curie Intra-European Fellowship Programme (grant agreement PIEF-GA-2010-274454) as well as computing time from the Norwegian Metacenter for Computational Science (NOTUR), the Leicester Grid Computer ALICE and the Lagrange computing cluster at NHH are gratefully acknowledged. This paper was completed during a visit of the last three authors to the Hausdorff Research Institute for Mathematics at the University of Bonn in the framework of the Trimester Program Stochastic Dynamics in Economics and Finance. We are grateful to the reviewers and an associate editor for helpful comments.

most profitable trades, even without informational advantages, can result in big returns both for the individual and the employer (Coval et al., 2005). Despite this, skills are mostly ignored in the academic literature on asset markets. Bayesian rationality, the mainstay of decision theory, simply disregards the possibility that some investors may not have the ability to make good decisions: In the standard model all agents are able to solve the appropriate optimisation problem. In that world, information matters, but skills do not.

In this paper we consider the opposite scenario: a world with complete and symmetric information - but one in which individuals differ in their trading skills. We present a computational equilibrium model of a market for financial assets in which traders develop skills endogenously. Skill is defined in this paper as a trader's ability, under full symmetric information, to price financial contracts accurately. For instance, a skilled trader could follow an asset pricing theory that enables her to make good decisions on the purchase or sale of incorrectly valued assets. Although skill is irrelevant in perfectly efficient markets, it matters when a trader has to post a price at which others can trade, e.g., in order-driven markets, or when asset prices in a competitive market deviate from their fair values.

Empirical studies have shown that markets are usually populated by skilled and unskilled traders. The first group is able to consistently make above average risk-adjusted returns and the latter is not, e.g., Oliven and Rietz (2004); Barras et al. (2010); Fama and French (2010); Barber et al. (2014). Traders develop these skills through learning (e.g. Coates and Page, 2009), potentially participating in unprofitable trades early in their career in the hope that they will acquire the ability to submit profitable orders with experience (Coval et al., 2005). Mimicking the trading strategies of superior investors can be profitable: For instance Grinblatt et al. (2012) find that investors with a high IQ exhibit superior stock picking skill and that copying their trades leads to abnormal annual returns of 11% over low IQ investors.

In this paper we apply an evolutionary learning process where natural selection ensures that profitable trading strategies survive and spread in the market.[2] This approach is motivated by the above evidence and by human decision-making behaviour in ‘large worlds’, i.e., situations in which it is impossible to foresee all potential consequences of one’s actions. Binmore (2007) argues that in large worlds Bayesian decision theory has to be replaced by other forms of learning and decision-making. Some progress towards that end has been made in game theory, notably by Ellison and Fudenberg (1993), where the players apply decision rules learned in one set of games when playing a new game in which they have no prior experience. This learning mechanism is a key feature in the model considered here. It describes an adaptive market in the sense of Lo (2004): The agents must either learn and become skilled or free-ride on others’ skills, and their incentives to do the one or the other depends upon the institutional setting and the behaviour of other traders. This differs from the approach taken by Grossman and Stiglitz (1980) in which individuals decide whether or not to purchase information.

An important feature of this modelling approach is that risk preferences and other trader characteristics are endogenous (Lensberg et al., 2015). In particular, there are no utility functions and no preassigned roles as skilled/unskilled trader. Instead, the model implements Alchian’s (1950) ‘as if’ view of rational behaviour as the outcome of a competitive evolutionary process. We find that market prices are rational but, at the same time, there is substantial heterogeneity with respect to skill and trading behaviour.

[2]Yeh (2008) employs a similar method to show that more intelligent traders may contribute to increased market efficiency.

We consider market structures with multiple trading venues, in which the traders quote prices at which other traders can buy or sell. The market is centralised if there are few venues and many traders at each venue, and fragmented if there are many venues and few traders at each venue. We find that a move from a fragmented to a centralised market structure can harm market stability and increase volatility by adversely affecting the proportion of skilled traders. Centralised markets protect unskilled traders from the consequences of bad decisions by allowing them to free-ride on the prices discovered by skilled traders. In the real world, traders free-ride by placing market orders, i.e. orders to buy (or sell) at any price. In the model, traders can place buy or sell market orders by quoting very high or very low prices.

In centralised markets with many traders, transaction prices tend to be efficient, and small market orders tend to have little price impact. Therefore, the incentives to acquire skills are weak, and, in equilibrium, most traders are unskilled. In fragmented markets, market orders have more price impact. Consequently, skilled traders, who quote prices close to fundamental values, make money by trading with unskilled traders who do not, and therefore most traders are skilled in equilibrium. As a result, fragmented markets are more resilient. Inter-market price variation, defined as the variation in prices between trading venues is, however, increasing in market fragmentation.

Regulators are concerned with the potential lack of competition among exchanges; economies of scale and network externalities make these institutions natural monopolies (Mendelson, 1987). Regulations such as MiFID in Europe and Regulation ATS and RegNMS in the US aimed to increase competition for order flow (Fink et al., 2006). As a consequence there has been a rise in the number of alternative trading venues such as crossing networks and dark pools which attract traders away from the main exchanges (Stoll, 2008). Despite the increase in the number of places to trade, the US market, for instance, has become, as described by O'Hara and Ye (2011), a 'single virtual market with multiple points of entry'.

Market centralisation during the last two decades has been accompanied by a dramatic increase in the stock market participation of non-professional investors (Bogan, 2008) who adopt investment strategies without fully understanding the risks (Bikhchandani et al., 1998). Our model shows that the increase in the proportion of unskilled traders following market centralisation is an equilibrium phenomenon.

Section 2 sets out a model of a market with multiple trading venues and endogenous skills. Section 3 discusses data collection and performance measurement. Section 4 presents results on the relationship between market fragmentation and pricing errors, the prevalence of skills, and the consequences for price volatility and market resilience. Section 5 concludes. Tests of model convergence and robustness are provided in Appendix AppendixA.

2. Model

We consider markets populated by agents who trade option contracts at one or more trading venues during a sequence of trading rounds. A *market* for an option contract is a list (I, N, A, C), where I is the number of traders; N is the number of traders per venue; A is an assignment of venues to traders, and C is the option contract. The set $\{1, 2, ..., I\}$ is a *population* of traders, and the pair $M = (I, N)$ is a *market structure.*

To run the model, we first select a market structure M, and a seed for the random number generator. Then for each trading round $t = 1, 2, ..., T$, we randomly choose an assignment A_t of venues to traders, and a random option contract C_t to obtain the market (M, A_t, C_t). Next, all agents trade in their respective venues, and then (imperfectly) hedge their positions until the contract expires at $t + 1$. Then a new market opens; the traders are again randomly assigned to venues to trade a new randomly selected option contract. A trader therefore competes with a random selection of traders in each round and, over the sequence of rounds, with the whole population of traders.

We consider a set $\mathcal{M}$ of 33 different market structures, with 3 possible population sizes I and 11 venue sizes N. It is defined as

$$\mathcal{M} := \{2000, 4000, 8000\} \times \{2, 4, 8, 16, 32, 64, 128, 250, 500, 1000, 2000\} \quad (1)$$

Trade is bilateral if $N = 2$, fragmented for small N, centralised for large N, and fully centralised for $N = I$. The number of venues in a market structure (I, N) is the integer part of I/N. If N is a divisor of I, every trader will participate in every market. Otherwise there will be some degree of non-participation. For the market structures in $\mathcal{M}$, the non-participation rate is 4% for $(I, N) = (2000, 128)$, and less than 1% for all other market structures. Repeated random assignment of venues to traders ensures that all traders are treated equally with respect to market participation.

Market. In each venue of every market, the buyers and sellers of the contract are determined as follows. All traders simultaneously quote prices at which they are indifferent between buying and selling one option contract. The transaction price is then obtained as the median of these quotes.[3] At this price supply is equal to demand. Traders with quotes above (below) the market price buy (sell) one unit of the contract. Ties are resolved by randomly assigning traders to be buyers and sellers.

All contracts are European call options with three months time to expira-

[3]Since all venues have an even number of traders, the median quote is the mean of the two nearest quotes.

tion.[4] The price dynamic of the underlying asset follows a geometric Brownian motion

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t)$$

with drift μ and volatility σ. Traders also have access to a money market with interest rate r. Time is measured in years and the parameter values for drift, volatility and interest rate are given per annum. The initial spot price $S(0)$ is fixed at 100 without loss of generality due to the scalability of the price process. The set of all option contracts is a subset of $\mathbb{R}^3$ defined as

$$\mathcal{C} := [80, 120] \times [0.1, 0.3] \times [0.01, 0.06], \quad (2)$$

with typical element $C = (K, \sigma, r)$.

Gains and losses. At each issue date, a random option contract is drawn from $\mathcal{C}$; its price is established at each venue and traders exchange premiums. Traders then hedge their exposure to the option contract until expiry by trading in the underlying asset and the money market. At the beginning of each trading day, their portfolio is adjusted to match the Black-Scholes delta-hedging strategy. When the contract expires, sellers pay the option payoff (if any) to buyers.

The profit of a call option writer who delta-hedges the short position in the 3-month option at times $t_n = n\,\Delta t$ with $n = 0, 1, ..., 65$ and $\Delta t = 1/264$, from

[4] The model uses European option contracts as these assets have a well known, closed form pricing rule. The conclusions we draw, however, are applicable to any asset market. For example, in equity markets skilled traders use information about companies to identify mispriced shares. The form of the mapping between a company's descriptive statistics and its asset price is, however, less clear than the case considered in this paper.

the issue time $t_0 = 0$ to the time of expiry $T = t_{66}$ evolves as:[5]

$$v(t_{n+1}) = e^{r\Delta t}[v(t_n) - \phi(t_n)S(t_n)] + \phi(t_n)[S(t_{n+1}) - S(t_n)]$$

where

$$\phi(t_n) = \Phi\left(\frac{\log(S(t_n)/K) + (r + \sigma^2/2)(T - t_n)}{\sigma\sqrt{T - t_n}}\right)$$

and Φ is the standard normal cumulative distribution function, $v(0) = 0$. With option price P, the seller's payoff at expiry is

$$\Pi = Pe^{rT} + v(T) - [S(T) - K]^+ \tag{3}$$

and that of the buyer is $-\Pi$ because premium payments, hedge positions and payments net to zero. Since the hedging strategy is not continuous, it provides only an imperfect hedge. There is, therefore, a stochastic element to the returns from all options.

Learning. Each trader has a pricing function which determines her quote. A pricing function associates a real number to every option contract $C \in \mathcal{C}$. In contrast to standard asset pricing models, these pricing functions are not derived by solving explicitly formulated optimisation problems. Instead, traders replace their pricing functions when they underperform relative to other traders. A new pricing function is obtained by randomly combining elements from the pricing functions of more successful traders in the same population. Between replacements, the trader's pricing function stays constant. In particular, there is no updating of parameters in response to new information. In this sense, individual traders don't learn. All learning takes place at the population level through competition and natural selection.

[5]For simplicity we assume that the 66 trading days from issue to expiry, a quarter of the annual 264 trading days, are equally distributed across time.

Success is measured relative to the market rather than some artificial benchmark. A trader's pricing function is more successful than that of another trader if, on average, it produces more wealth.

We represent the learning process by a genetic programming algorithm with tournament selection as in Lensberg and Schenk-Hoppé (2007).[6] Pricing functions are implemented as computer programmes. A programme is a list of up to 256 instructions. Each instruction consists of an operator and one or two operands. The set of operators is $\{+\ , -, /, \times$, max, min, change sign, exp, log, $\Phi\}$, where Φ is the standard normal CDF. Operands consist of the three option parameters (strike, volatility and interest rate); 2^{13} numerical constants; and three temporary variables. Traders price options by executing their programmes. Their quotes are given by the value of the first temporary variable on completion of the programme.

To model replacement of pricing functions, we use tournaments that are run after each option is traded and the payoffs calculated. Initially each trader is equipped with zero wealth and a randomly generated pricing function. From then on, each trader continues to use her current pricing function until it is replaced by a new one according to the following algorithm:

1. *Tournament:* Randomly (and uniformly) select four traders in the population and rank them by accumulated wealth.
2. *Reproduction:* Replace the programmes of the two with the lowest rank by copies of those of the two with the highest rank.
3. *Crossover:* With probability χ_1, swap two randomly selected sublists of instructions of equal length between the two programmes.

[6]Genetic Programming was introduced by Koza (1992) as an extension of genetic algorithms (Holland, 1975). This technique has been used in economics and finance in multiple applications, for instance, Noe et al. (2003, 2006) investigate optimal security design, whilst Chen and Yeh (2002); Yeh and Yang (2010) both use it to model trading strategies in an artificial stock market.

4. *Mutation:* For each programme, with probability χ_2, an instruction, an operation or operand in a programme is randomly selected, and replaced by a new random instruction, operation or operand.

In every tournament, the two parent programmes are left unchanged, while those of the two offspring are created as described in the algorithm. All four traders keep their wealth unchanged. Alternatively, we could have reset the (mostly negative) wealth of the offspring programmes to the zero population mean. However, that would improve their initial wealth; increase their rate of reproduction; and, since most offspring are unskilled, slow down the learning process. To support learning, the offspring should display some evidence of skill before they get a chance to reproduce. To that end, we give them an initial handicap by preserving their wealth, and use discounting to ensure that the handicap is only temporary.

At every option market date, the accumulated wealth (positive or negative) of each trader is discounted by the multiplicative factor $(1 - \tau)$, where τ is the 3 month discount rate.[7] Discounting provides poor traders with good pricing functions with a chance to overtake wealthy traders with inferior ones, while preserving wealth ratios and tournament rankings. Without discounting, the genetic algorithm would not be able to produce increasingly better pricing functions.

The parameter values for the model are presented in Table 1. It takes about 10 million trading rounds to evolve a model to a converged state. We provide some margin by adding another 7.5 million rounds. For every trading round we carry out $I/2000$ tournaments. Each tournament involves four pricing rules, of which two are replaced, hence pricing functions are replaced at a rate of 1/1000

[7] In the model, the option markets are open for trade every 3 months, and 3 months is the time to expiry of the traded options.

Table 1: Parameter values.

Population size	$I \in \{2000, 4000, 8000\}$
Trading venue size	$N \in \{2, 4, 8, 16, 32, 64, 128, 250, 500, 1000, 2000\}$
Trading rounds T	17,5 million
Tournaments per trading round	$I/2000$
Tournament size	4 traders
Crossover probability χ_1	0.50
Mutation probability χ_2	0.05
Discount rate (3 month) τ	1%
Initial stock spot price $S(0)$	100
Stock price drift μ	6%
Option strike price K	uniformly drawn from $[80, 120]$
Stock price volatility σ	uniformly drawn from $[10\%, 30\%]$
Interest rate r	uniformly drawn from $[1\%, 6\%]$
Penalty if quote < 0 or > 40	5

per trading round.

3. Data and performance measurement

In this section we describe the data and the performance measures that will be used to assess the quality of different market structures. Our main interest is in trader skill, but we will also consider various measures of price variation.

Data for the analyses are generated as follows. For each market structure $(I, N) \in \mathcal{M}$, we carry out 50 independent model runs by choosing 50 different seeds for the random number generator. During each run, which consists of $T = 17.5$ million trading rounds, the market varies randomly across rounds with respect to the assignment of venues to traders and the traded option contract. At the end of each run, the pricing functions of the converged models are recorded and used to generate data for the statistical analysis.

All performance measures are based on a fixed, representative set $\hat{\mathcal{C}}$ of 27 option contracts. This set is given by all combinations of the following parameter values:

$$K = 95, 100, 105;\ \sigma = 0.15, 0.2, 0.25;\ \text{and}\ r = 0.02, 0.035, 0.05. \tag{4}$$

The traders' pricing decisions on this set of options are evaluated only to collect data; there is no corresponding change in their wealth.

The performance measures will be based on various aggregates of quotes and transaction prices, and we will distinguish carefully between the two. For that purpose, the following definitions will be useful:

Definition 1. A *trader quote* is a price quote made by an individual trader.

Definition 2. A *transaction price* is the median quote among all traders in a venue. (Since all trading venues have an even number of traders, the transaction price is the mean of the two quotes closest to the median.)

Definition 3. A *market mid price* is the median transaction price among all venues in a market.

Definition 4. A *market mid quote* is the median quote among all traders in a population. (This is as a proxy for the sort of market-wide consolidated price quotes provided by the SIPs (Security Information Processors) in the US stock market.)

In measuring skill, choosing a proper benchmark is a key issue. Figure 1 depicts the distributions of trader quotes for the option contract $(K, \sigma, r) = (95, 0.2, 0.035)$ in two market structures; one bilateral and one fully centralised. The middle spike of the left graph is shown in detail on the right.

An obvious measure of skill is the deviation between a trader's quote and the Black-Scholes price, which is marked by the vertical lines in Figure 1. The left graph illustrates that the quotes tend to cluster in three distinct areas: very low prices, very high prices, and at some value in between, fairly close to the Black-Scholes price. Any measure of skill based on the distance to some benchmark will therefore be dominated by the extreme quotes, with low sensitivity to variations in quotes within the middle cluster.

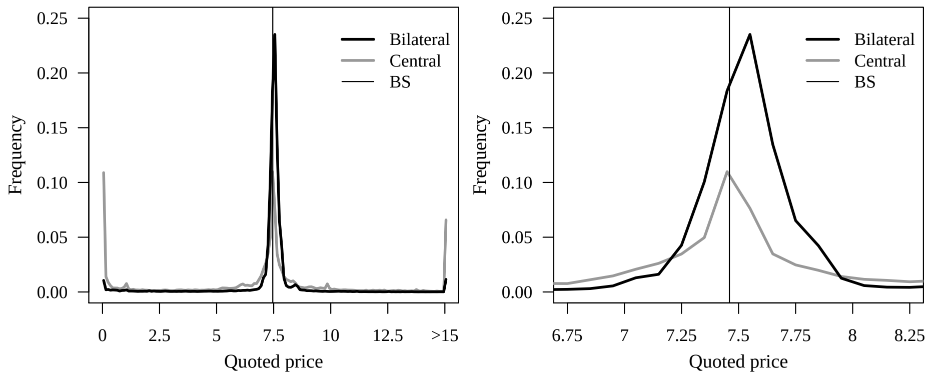

Figure 1: Histogram of traders' price quotes for the option contract $(K, \sigma, r) = (95, 0.2, 0.035)$ under centralised trade (grey line) and bilateral trade (black line) in a population of 4,000 traders. The left and right graphs show respectively the whole distribution of quotes, and the details of the middle clusters. The Black-Scholes (BS) price is 7.46, represented by the vertical lines in the histograms. The histogram is calculated over 50 independent model simulations for each venue size.

Traders who quote either very high or very low prices for the option are the ones we would like to call unskilled. But such behaviour does not necessarily make them stupid. In real markets, many traders use market orders[8] to execute their trades. In centralised markets with many traders, transaction prices tend to be efficient, and small market orders tend to have little price impact. Unskilled traders can then use market orders to free ride on the prices discovered by skilled traders. In fragmented markets, the price impact of market orders is larger, and the cost of being unskilled higher.

In the model, the traders can easily submit market orders by quoting a very high price to implement a buy order, or a very low price for a sell order. But the cost of doing so is high in fragmented markets. To see this, consider a bilateral market for a contract with fundamental value p_f, where the two traders, denoted s and b, quote prices p_s and p_b such that $p_s < p_b$. Then trader s is the seller, b is the buyer, and the transaction price is $p_t = (p_s + p_b)/2$. The traders' expected

[8]A market order is an order to buy (or sell) at any price.

profits, π_s and π_b, satisfy

$$\pi_b = -\pi_s = p_f - p_t = p_f - (p_s + p_b)/2. \tag{5}$$

Equation (5) shows that both traders have an incentive to quote a price close to that of the other trader, and in Nash equilibrium, they both quote the fundamental price.[9] Extreme quotes are expensive because the price impact of a quote change amounts to 50% of the change. They are less expensive in larger venues, where a quote change has no impact on the transaction price as long as the quote stays on one and the same side of the spread between the two quotes closest to the median.

The other prominent feature of Figure 1 are the spikes in the distributions of quotes close to the Black-Scholes price. These are the quotes of the traders we would like to call skilled. The figure suggests to measure the number of skilled traders by counting those who quote prices within a neighbourhood of the Black-Scholes price, as illustrated in the right-hand graph of the figure. The graph shows that both quote distributions, in particular the bilateral one, have more mass to the right of the Black-Scholes price. The main reason for this apparent anomaly is not that our traders quote biased prices, but that the Black-Scholes price is a biased measure of the fair option price.

The bias arises because the traders can only adjust their hedge positions once per day, which exposes them to some unhedged risk. The traders' quotes will then depend on their risk aversion. Figure 2 depicts the Black-Scholes price for the option contract $(K, \sigma, r) = (95, 0.2, 0.035)$, along with the maximal buy price and the minimal sell price for an option trader with constant risk aversion 2. The figure shows that the Black-Scholes price is biased downwards relative

[9] Consider two traders, i and j. If $p_j = p_f$ then $\pi_i = 0$ for $p_i = p_f$, otherwise $\pi_i < 0$.

to the mid price between the buy and sell prices.

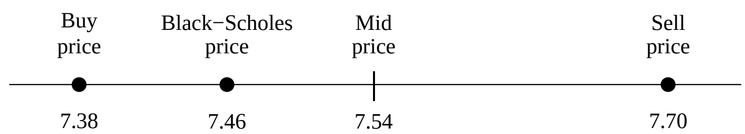

Figure 2: Buy price, sell price, mid price, and Black-Scholes price for the option contract $(K, \sigma, r) = (95, 0.2, 0.035)$. The buy and sell prices are indifference prices, obtained by Monte Carlo simulation, for an agent with constant absolute risk aversion 2.

This suggests to measure skill relative to the mid price between the buy and sell prices. But that too depends on the traders' risk aversion. For example, increasing the risk aversion from 2 to 4 increases the mid price from 7.54 to 7.71. Since risk aversion is an endogenous (and only implicit) variable of the model, the mid price in Figure 2 is not a feasible benchmark for measuring skill.

An unbiased measure of skill can be obtained if we are willing to assume that the markets mostly get it right. We can then measure a trader's skill by counting the number of contracts in $\hat{\mathcal{C}}$ for which her quote lies within a neighbourhood of the market mid quote.

In view of these considerations, we introduce the following performance measures and associated data sets. Unless otherwise stated, the data are collected by means of a single evaluation of each individual pricing function in the converged models for each contract in the set $\hat{\mathcal{C}}$ of test option contracts.

Skill is a measure how well a trader can price option contracts. We use the market mid quotes for all the 27 contracts in $\hat{\mathcal{C}}$ as a benchmark. We then consider the quotes made by a trader i for these contracts and count the number of quotes that lie within a $\pm 10\%$ band around the corresponding market mid quote[10]. The resulting number $k_i \in \{0, ..., 27\}$ is trader i's skill, and $(1/I)\sum_{i=1}^{I} k_i$ is the population skill. Our data set consists of mean population skill conditional on

[10]The particular choice of the percentage by which a quote can differ from the market mid quote is not critical. One obtains qualitatively similar results with band widths of 5% and 20%.

market structure for each one of the 50 model runs with each market structure. This yields a total of $33 \times 50 = 1,650$ observations.

Trader pricing error measures the extent to which the price quotes of individual traders deviate from Black-Scholes prices. For each contract in $\hat{\mathcal{C}}$ and each trader in the population, we compute the absolute difference between the trader's quote and the Black-Scholes price. The trader pricing error is the mean of these differences across all traders in the population. Our data set consists of mean trader pricing errors conditional on 33 market structures and 27 option contracts, a total of 891 observations.

Market pricing error measures the extent to which market mid quotes deviate from Black-Scholes prices. The objective is to have a measure of pricing error which is less sensitive to extreme quotes than the trader pricing errors. For each contract in $\hat{\mathcal{C}}$, we compute the absolute difference between the market mid quote and the Black-Scholes price. Our data set consists of mean market pricing errors conditional on 33 market structures and 27 option contracts, a total of 891 observations.

Trader wealth is used to measure the association between skill and wealth. For each market structure $(I, N) \in \mathcal{M}$, we have a set of 50 converged models. We restart each model with zero wealth for each trader and run the model for 100,000 trading rounds with the tournament process turned off to keep the traders' pricing functions intact. We then make a record of each trader's wealth, his skill, and the market structure parameters I and N. Finally, we construct a data set by computing mean wealth conditional on market structure and trader skill. This yields a total of 924 observations (3 population sizes, 11 venue sizes and 28 levels of skill ranging from 0 to 27).

Price volatility measures the time-series variability of market mid prices. We do 10,000 rounds with the converged models with the tournament process turned on. For every 10 rounds, we compute the market mid price of each option contract in $\hat{\mathcal{C}}$ to get a sample X of 1,000 observations for each run and contract. The price volatility of contract C for that run is the volatility of sample X. We compute mean volatility conditional on 33 market structures and 27 option contracts to obtain a data set with 891 observations.

Price dispersion is a measure of the variability in transaction prices across trading venues. Market structures (I, N) with $N = I$ are excluded since they have only one venue and no price dispersion. This is the case for $(I, N) = (2000, 2000)$ only, but in order to preserve orthogonality among the treatment variables, we also exclude the two other market structures with $N = 2000$.

To collect data, we run 10,000 rounds with the converged models with the tournaments process turned on. For every 10 rounds, we compute the transaction prices for each venue and each option contract in $\hat{\mathcal{C}}$ to get a sample of 1,000 observations for each run, venue and contract. For each run and contract, we compute the standard deviation of transaction prices across market venues and take the average of these standard deviations across the set of 1,000 observations to get the price dispersion for that run and contract. We compute mean volatility conditional on 30 market structures and 27 option contracts to obtain a data set with 810 observations.

Price sensitivity is measured by exposing the market to shocks in terms of entries into the market by traders with extreme option valuations. Let $P(0)$ be market mid quote in the population without additional traders, and let $P(J)$ resp. $P(-J)$ denote the market mid quote when $J > 0$ additional individuals who post the highest resp. lowest feasible quote are added to the market. The shock size J can take on 7 values; 1%, 2%, 5%, 10%, 20%, 40%, and 80% of the population size I. For each value of J, we calculate the corresponding price sensitivity

$$\frac{P(J) - P(-J)}{P(0)} \Big/ \frac{J}{I} . \tag{6}$$

We compute mean price sensitivity conditional on 33 market structures, 27 options and 7 shock sizes to obtain a data set with 6,237 observations.

Table 2: Summary statistics for dependent variables.

Statistic	Obs	Mean	St.Dev	Skew	Min	Median	Max
			Panel A: Raw Variables				
Market pricing error	891	0.082	0.064	3.017	0.029	0.061	0.453
Trader pricing error	891	1.168	0.594	0.065	0.207	1.213	2.595
Skill	1,650	16.350	5.967	0.151	2.604	15.521	26.267
Wealth (1000)	924	-0.147	1.182	-5.008	-8.732	0.164	0.423
Price volatility	891	0.047	0.026	0.415	0.004	0.047	0.125
Price dispersion	810	0.141	0.210	2.509	0.012	0.073	0.818
Price sensitivity	6,237	0.447	0.612	3.082	0.005	0.231	5.630
			Panel B: Transformed Variables				
log(Market pricing error)	891	-2.676	0.527	1.252	-3.543	-2.799	-0.792
log(Trader pricing error)	891	-0.016	0.637	-0.657	-1.575	0.193	0.954
logit(Skill/27)	1,650	0.631	1.245	0.658	-2.237	0.302	3.580
exp(Wealth)	924	1.072	0.316	-2.185	0.000	1.179	1.527
log(Price volatility)	891	-3.234	0.666	-0.830	-5.464	-3.058	-2.076
log(Price dispersion)	810	-2.555	0.971	0.867	-4.405	-2.614	-0.200
log(Price sensitivity)	6,237	-1.559	1.339	-0.395	-5.297	-1.465	1.728

Note: Prices are option prices

Table 2 provides summary statistics for raw and transformed versions of the seven performance measures. The raw variables in Panel A display some positive and negative skewness, partly because most of them have a bounded range. The two pricing errors and the three measures of price variation have a lower bound at zero, and varying degrees of positive skewness. Some of these variables are natural pairs. We therefore log transform all of them to obtain less skewness in general while preserving comparability between the pairs. Skill is bounded between 0 and 27, and we use a logit transformation of Skill/27 to obtain an unbounded range. Wealth displays strong negative skewness, which is caused by large numbers of unlucky unskilled traders. We apply an exponential transformation to obtain some reduction in skewness at the expense of a bounded range.

4. Results

The discussion of the results follows the list of performance measures: Pricing errors, skill, wealth distribution, and the three measures of price variation.

We explore the properties of the model by regressing each performance measure on the market structure parameters I (population size) and N (venue size). In addition, we use a dummy variable 'Fragmented' to distinguish between fragmented and highly centralised markets. The dummy variable is set to 1 for markets with venues of 250 traders or less. At about this venue size most of the dependent variables exhibit a kink when plotted against venue size. We include the Black-Scholes price as an additional regressor, wherever that is relevant. This improves precision and provides some information on whether the traders distinguish systematically between cheap and expensive options. Alternatively, we could have included each separate option parameter, but their main effect appears to go through the Black-Scholes price. All regressions will be run with log-transformed versions of the explanatory variables.

Our model runs are structured in such a way that the regressors are orthogonal, hence there is no multicollinearity in the data sets, except between venue size and its cross product with the dummy variable. We use OLS to estimate the parameters and robust (HC3) standard errors to deal with heteroskedasticity. Since our data sets are quite large, we use a 1% significance level for hypothesis testing.

4.1. Pricing errors

We first investigate the relationship between market structure and pricing errors. The pricing error of a quote is the absolute difference between the quote and the Black-Scholes price. We estimate one model for the quotes made by individual traders and one for market mid quotes. Table 3 contains the results.

Table 3: Pricing errors measured as the absolute difference between price quotes and Black-Scholes option prices. Two models are estimated. One for quotes by individual traders and one for market mid quotes. Robust standard errors in parentheses.

	Dependent variable: log(Pricing error)			
	Individual traders		Market	
	(1)	(2)	(3)	(4)
Constant	0.174	1.301^{***}	0.460^{*}	-0.598^{**}
	(0.125)	(0.099)	(0.197)	(0.208)
log(Venue size)	0.264^{***}	0.039^{***}	-0.140^{***}	0.034
	(0.004)	(0.008)	(0.007)	(0.020)
log(Population size)	-0.155^{***}	-0.155^{***}	-0.308^{***}	-0.308^{***}
	(0.015)	(0.009)	(0.023)	(0.018)
Fragmented		-1.645^{***}		1.309^{***}
		(0.061)		(0.142)
log(Venue size) × Fragmented		0.320^{***}		-0.305^{***}
		(0.010)		(0.022)
log(Black-Scholes price)		0.195^{***}		0.054^{**}
		(0.011)		(0.020)
Observations	891	891	891	891
Adjusted R^2	0.837	0.937	0.444	0.634
Note:		$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$		

As discussed in Section 3 and illustrated in Figure 1, unskilled traders effectively submit market orders by quoting very high or very low prices. Moreover, the price impact of market orders is lower in more centralised markets. We therefore expect to find an increasing relationship between venue size and individual pricing errors.

Model (1) of Table 3 supports this hypothesis. The model fits the data well, and the coefficient on venue size is positive and highly significant. Increasing the population size reduces the pricing error, but the effect is smaller and the coefficient estimate is less precise. In Model (2) we add the 'Fragmented' dummy variable and its interaction with log(Venue size), along with the Black-Scholes price. The table shows that the interaction term picks up almost all the effect of venue size on individual pricing errors. The positive relationship between venue size and individual pricing errors is therefore limited to market structures that are fragmented at the outset. The positive coefficient on the Black-Scholes price reveals that pricing errors are larger for more expensive options.

Models (3) and (4) contain analogous results for market pricing errors. Again we find that the effect of venue size on pricing error is limited to markets that are fragmented at the outset. Compared to the model of individual pricing errors (Model 2), this feature is more pronounced in the market model (4), as can be seen by the insignificant coefficient on venue size and the large increase in R^2 from Model (3) to (4). However, the sign of log(Venue size) $\times$ Fragmented has switched from positive to negative between Model (2) and (4). Thus in fragmented markets, an increase in the venue size reduces the market pricing error.

To see why, recall first that due to imperfect hedging, fair market prices are higher than Black-Scholes prices, and more so for more risk averse traders, cf. Figure 2 and the related discussion. A closer look at the data reveals that in general, market mid prices are higher than Black-Scholes prices, and more so in more fragmented markets. This indicates that the observed negative relationship between venue size and market pricing error is due to more risk aversion among traders in fragmented markets.

One may wonder why we do not observe the same negative relationship for individual pricing errors. Note first that individual pricing errors exceed market pricing errors by a factor of 20, cf. Panel A of Table 2. The reason, as pointed out in Section 3, is that the individual pricing errors are dominated by extreme quotes made by unskilled traders (the left and right clusters in Figure 2), while the market pricing errors are dominated by the quotes made by skilled traders (the middle cluster in Figure 2). Therefore, the large pricing errors of the unskilled traders dominate the small errors of the skilled ones in the model for individual pricing errors.

4.2. Skill

We investigate how skill varies with the market structure. The key variable is venue size (market centralisation), which we expect to have a negative impact on skill. Three model specifications are used: Model (1) measures the overall impact of venue size on skill, Model (2) looks at the effect of population size, and Model (3) considers whether the effect of venue size differs between markets that are more or less fragmented at the outset. Table 4 provides the results.

Table 4: Trader skill. We use the 27 options from the set $\mathcal{C}$ of test option contracts to measure skill. A trader's skill is the number of options from this set she prices within ±10% of the market mid price. Robust standard errors in parentheses.

	Dependent variable:		
	logit(Skill/27)		
	(1)	(2)	(3)
Constant	2.499***	0.066	−2.411***
	(0.043)	(0.278)	(0.429)
log(Venue size)	−0.450***	−0.450***	−0.046
	(0.009)	(0.009)	(0.050)
log(Population size)		0.293***	0.293***
		(0.033)	(0.029)
Fragmented			2.973***
			(0.349)
log(Venue size) × Fragmented			−0.600***
			(0.051)
Observations	1,650	1,650	1,650
Adjusted R^2	0.623	0.640	0.718
Note:			*p<0.05; **p<0.01; ***p<0.001

Model (1) shows that centralisation has a negative effect on skill. As expected, the coefficient of log(Venue size) is negative and highly significant, and the model produces a good fit with $R^2 = 62.3\%$. In fragmented markets, individual quotes have more price impact, and this puts pressure on the traders to develop skills. Model (2) shows that population size has a significant positive effect on skill, but this effect has low economic importance, as measured by its impact on R^2.

In Model (3), we add the 'Fragmented' dummy and its interaction with venue size. As with pricing errors, we find that the direct effect of venue size is picked up by its interaction with the Fragmentation dummy, and hence, that the effect of increased venue size on skill is limited to markets that are relatively fragmented at the outset. In highly centralised markets, further centralisation does not decrease skill.

In supporting large numbers of unskilled traders along with a few skilled ones, the centralised markets in our model resemble those examined by Oliven and Rietz (2004). Unskilled traders only have a small effect on the price and receive very little feedback on their strategy; instead they free-ride on the skills of others. Barber and Odean (2001) observe that with easier market access through internet trading, there has been a growth in free on-line advice. Unskilled traders can then get a free ride by basing their decisions on this information.

4.3. Skill and wealth

If skill is valuable, one may wonder why some traders are unskilled. Clearly, there must be some cost associated with skill, and this cost arises because skill entails model risk for the traders who attempt to acquire it. Although our traders do not obtain their pricing functions through conscious deliberation, the evolutionary learning process works as if they do. Suppose a trader who is about to change his pricing function could choose between a robust rule of thumb (e.g. always quote a very high price) or a combination of two complex pricing functions, both of which are sensitive to small variations in the data. In a centralised market, the first alternative is safe because the trader will do no worse than other buyers. The second alternative is risky because a combination of two sensitive pricing functions might result in one that loses money systematically. The equilibrium number of skilled traders will therefore depend on the trade off between the expected cost of model risk and the expected benefits of skill.

The marginal reward to skill should be highest in those market structures where skill matters most, i.e., in bilateral markets where the transaction price is halfway between the two traders' quotes. It should be lowest in fully centralised markets where the expected price impact is small. As we have seen in the previous analysis, skill increases with market fragmentation. We therefore expect skill to increase wealth and more so in more fragmented markets. Traders who free-ride on the prices discovered by skilled traders face little model risk, and in equilibrium they should not be rewarded.

Table 5 contains results on the relationship between skill and wealth conditional on the market structure. The table shows that wealth increases with increasing skill (Model 1), with no significant difference between fragmented and centralised markets, (Model 2). In Model (3), we include the cross product of skill and venue size among the the set of regressors. By considering the combined effect of the three variables involving skill and venue size, the following picture emerges: (i) Skill has a positive effect on wealth across all venue sizes, but this effect is small in highly centralised markets. (ii) Venue size has a strong positive effect on wealth for low levels of skill, but the effect is close to zero for skill levels beyond 15. To paraphrase: Traders in fragmented markets would like to be skilled, but skill is a minor concern for traders in highly centralised markets. Unskilled traders would prefer to trade in centralised markets, but highly skilled traders do not care where they trade because high skill is always rewarded.

Table 5: Skill and wealth. We run the converged models for 100,000 trading rounds with the tournament process turned off, and with zero initial wealth for all traders. After the last round of each run, we collect wealth data for each individual trader, which yields a total of 8.7 million observations. We aggregate the data by computing the mean wealth for each population size, venue size, and each level of skill between 0 and 27. This gives 924 observations (3 population sizes, 11 venue sizes ranging between 2 and 2,000, and 28 levels of skill). Robust standard errors in parentheses.

	Dependent variable:			
	exp(Wealth)			
	(1)	(2)	(3)	(4)
Constant	0.609***	0.533***	0.142*	0.383*
	(0.024)	(0.050)	(0.056)	(0.159)
Log(1+Skill)	0.201***	0.209***	0.350***	0.351***
	(0.009)	(0.009)	(0.022)	(0.023)
log(Venue size)		0.014	0.091***	0.091***
		(0.007)	(0.012)	(0.012)
log(1+Skill) × log(Venue size)			−0.030***	−0.030***
			(0.005)	(0.005)
log(Population size)				−0.029
				(0.017)
Observations	924	924	924	924
Adjusted R^2	0.730	0.739	0.800	0.803
Note:		*p<0.05; **p<0.01; ***p<0.001		

4.4. Option price volatility and dispersion

In this section, we examine the time-series and cross-sectional variation of option prices conditional on the market structure. The main question of interest is whether market quality, in terms of low price variation across time and space, is better in centralised or fragmented markets. The raw data consist of transaction prices for individual trading venues. For the volatility analysis, we use time series of returns based on market mid prices (median transaction prices across trading venues). For the analysis of cross-sectional price dispersion, we use time series of standard deviations across market venues. Table 6 presents the results for time-series volatility.

Price volatility increases with centralisation, (Model 1) and the effect is limited to markets that are fragmented at the outset, (Model 3). The mechanism

Table 6: Price volatility. We run the the converged models for 10,000 rounds with the tournaments process turned on. For every 10 rounds, we compute the market mid price of each option contract in $\hat{\mathcal{C}}$ to get a sample X of 1,000 observations for each run and contract. The price volatility of contract C for that run is the volatility of sample X. We compute mean volatility conditional on 33 market structures and 27 option contracts to obtain a data set with 891 observations. Robust standard errors in parentheses.

	Dependent variable:			
	log(Volatility)			
	(1)	(2)	(3)	(4)
Constant	−4.344***	−1.948***	−0.642***	−0.925***
	(0.030)	(0.143)	(0.172)	(0.139)
log(Venue size)	0.268***	0.268***	0.060***	0.060***
	(0.006)	(0.005)	(0.018)	(0.013)
log(Population size)		−0.289***	−0.289***	−0.289***
		(0.018)	(0.014)	(0.013)
Fragmented			−1.516***	−1.516***
			(0.130)	(0.095)
log(Venue size) × Fragmented			0.291***	0.291***
			(0.020)	(0.015)
log(Black-Scholes price)				0.199***
				(0.013)
Observations	891	891	891	891
Adjusted R^2	0.769	0.829	0.880	0.906
Note:		*p<0.05; **p<0.01; ***p<0.001		

which drives the effect of centralisation on skill in fragmented markets appears to drive its impact on time-series volatility as well. As illustrated in Figure 1, a fragmented market displays a higher density of quotes around the median price. Transaction prices are therefore less prone to change in response to noise from unskilled traders' decisions.

Model (2) shows that price volatility is lower in larger populations of traders. The estimated coefficient is large and significant with a noticeable effect on the model fit. Still, from an economic perspective, increased venue size in ex ante fragmented markets, which leads to a loss of skill, is the main cause of increased time-series volatility.

In Model (4) we add the Black-Scholes price to the list of regressors. This yields more precise estimates for the market structure variables, but the coefficients do not change due to orthogonality among the regressors. The model shows that price volatility is higher for more expensive options.

Our results on price volatility agree with those of Smith et al. (1988), who find that markets with a large proportion of unskilled or inexperienced traders have larger price fluctuations.

We next consider the relationship between market structure and cross-sectional variation of prices. Table 7 shows the results.

Table 7: Price dispersion. We run the converged models for 10,000 rounds with the tournament process turned on. Market structures with venue size 2,000 are excluded. For every 10 rounds, we compute the transaction prices for each venue and each option contract in $\hat{\mathcal{C}}$ to get a sample of 1,000 observations for each run, venue and contract. For each run and contract, we compute the standard deviation of transaction prices across market venues and take the average of these standard deviations across the set of 1,000 observations to get the price dispersion for that run and contract. We compute mean volatility conditional on 30 market structures and 27 option contracts to obtain a data set with 810 observations. Robust standard errors in parentheses.

	Dependent variable:			
	log(Price dispersion)			
	(1)	(2)	(3)	(4)
Constant	−0.923***	1.396***	2.790***	1.095***
	(0.050)	(0.225)	(0.390)	(0.222)
log(Venue size)	−0.429***	−0.429***	−0.637***	−0.429***
	(0.010)	(0.010)	(0.047)	(0.010)
log(Population size)		−0.280***	−0.280***	−0.280***
		(0.027)	(0.027)	(0.026)
Fragmented			−1.376***	
			(0.320)	
log(Venue size) × Fragmented			0.200***	
			(0.050)	
Black-Scholes price				0.211***
				(0.028)
Observations	810	810	810	810
Adjusted R^2	0.767	0.793	0.794	0.807
Note:				*p<0.05; **p<0.01; ***p<0.001

Model 1 shows that cross-sectional price variation decreases with market centralisation. Price dispersion is lower in larger populations of traders (Model

2), and the effect of population size on price dispersion is virtually identical to its effect on volatility. In Model (3), we include the 'Fragmented' dummy and its cross-product with venue size. The effect of these variables appears to be significant, but including them adds nothing to the model fit. The estimated coefficients and standard errors of Model (3) indicate that this only picks up the multicollinearity between venue size and its cross-product with the dummy variable. Since Model (3) appears to be misspecified, we exclude the dummy terms from model (4) and conclude that venue size has a uniform negative marginal effect on price dispersion.

The uniform impact of centralisation on price dispersion results from two opposite effects. In fragmented markets, there are many venues and many different transaction prices. In centralised markets, there are few venues with many traders in each. Other things being equal, it follows by the law of large numbers that price dispersion is lower in more centralised markets. On the other hand, in fragmented markets, the traders are more skilled, and therefore, other things being equal, price dispersion is lower in fragmented markets. In our preferred model (4) it appears that the law of large numbers beats loss of skill with respect to the impact of centralisation on price dispersion.

Our results on time-series and cross-sectional volatility mirror those in Madhavan (1995). In our model, the skilled traders effectively act as market makers. They set the market price and stand ready to take either side of the contract at this price. With endogenous skills, prices are more volatile in fragmented markets, but their post-trade information on prices is more precise in comparison with centralised markets.

4.5. Resilience

We use our measure of price sensitivity (6) to measure the impact of centralisation on market resilience. This measure captures the response of the market mid quote to exogenous shocks. The shocks are modelled as an injection of traders with extreme valuations of the option contract into an existing population. The market is resilient if it has a low price sensitivity to such shocks. Table 8 collects the results.

Table 8: Price sensitivity. We expose the models to shocks by injecting traders with extreme option valuations into the populations. Let $P(0)$ be market mid quote in the population without additional traders, and let $P(J)$ resp. $P(-J)$ denote the market mid quote when $J > 0$ additional individuals who post the highest resp. lowest feasible quote are added to the market. The shock size J can take on 7 values; 1%, 2%, 5%, 10%, 20%, 40%, and 80% of the population size I. For each value of J, we calculate the price sensitivity $((P(J) - P(-J))/P(0))/(J/I)$. We compute mean price sensitivity conditional on 33 market structures, 27 options and 7 shock sizes to obtain a data set with 6,237 observations. Robust standard errors in parentheses.

	Dependent variable:			
	log(Price sensitivity)			
	(1)	(2)	(3)	(4)
Constant	-3.398^{***}	-0.606^{***}	1.782^{***}	3.040^{***}
	(0.027)	(0.177)	(0.294)	(0.236)
log(Venue size)	0.443^{***}	0.443^{***}	0.055	0.055^{*}
	(0.006)	(0.006)	(0.035)	(0.027)
log(Population size)		-0.337^{***}	-0.337^{***}	-0.337^{***}
		(0.021)	(0.019)	(0.016)
Fragmented			-2.852^{***}	-2.852^{***}
			(0.244)	(0.191)
log(Venue size) × Fragmented			0.572^{***}	0.572^{***}
			(0.036)	(0.028)
log(Black-Scholes price)				-0.882^{***}
				(0.016)
Observations	6,237	6,237	6,237	6,237
Adjusted R^2	0.522	0.542	0.601	0.731
Note:				$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Model (1) shows that price sensitivity increases with market centralisation. This is a direct effect of the loss of skill. Fragmented markets absorb shocks better because they have more skilled traders who quote prices close to the market mid quote. Population size reduces price sensitivity (Model 2), but this effect is not important, as measured by its impact on R^2.

In Model (3), we include the "Fragmentation" dummy and its cross-product with venue size. As usual, the effect of centralisation on price sensitivity is limited to markets that are relatively fragmented at the outset. Model (4) shows that market resilience is better for more expensive options, and this effect is large and economically important.

The difference in resilience between fragmented and centralised markets is illustrated in an example. We consider a scenario in which a market attracts (unskilled) bullish traders over a prolonged number of rounds. We model such a trader as one who quotes the highest feasible price for every option. These traders therefore increase the market mid quote. The effect will be stronger, the sparser the distribution of quotes close to the median quote.

We analyse two cases of different degrees of market centralisation. In the first case, trade is bilateral, and in the second case, trade is centralised at one venue. The initial trader population at round zero in both cases is taken from the equilibrium of the bilateral market with 4,000 traders. For each round in a sequence of 5,000 trading rounds, another bullish trader enters the population with 50% probability. After round 5,000 the market evolves without further exogenous shocks.

Figure 3 shows time series of the market mid quote of one representative option contract, and the proportions of skilled traders, unskilled buyers and unskilled sellers. *Skilled traders* denote traders who price all 27 test options within $\pm 10\%$ of the market mid quote. *Unskilled buyers (sellers)* are traders whose quotes are at least 10% above (below) the market mid quote for all 27 test options. Panel (a) shows that in the bilateral market an exogenous inflow of bullish traders increases the market mid quote to levels above the initial value. However, as soon as the inflow of bullish traders stops, both the mid quote and the number of unskilled buyers fall back to their original levels and the market returns to its equilibrium.

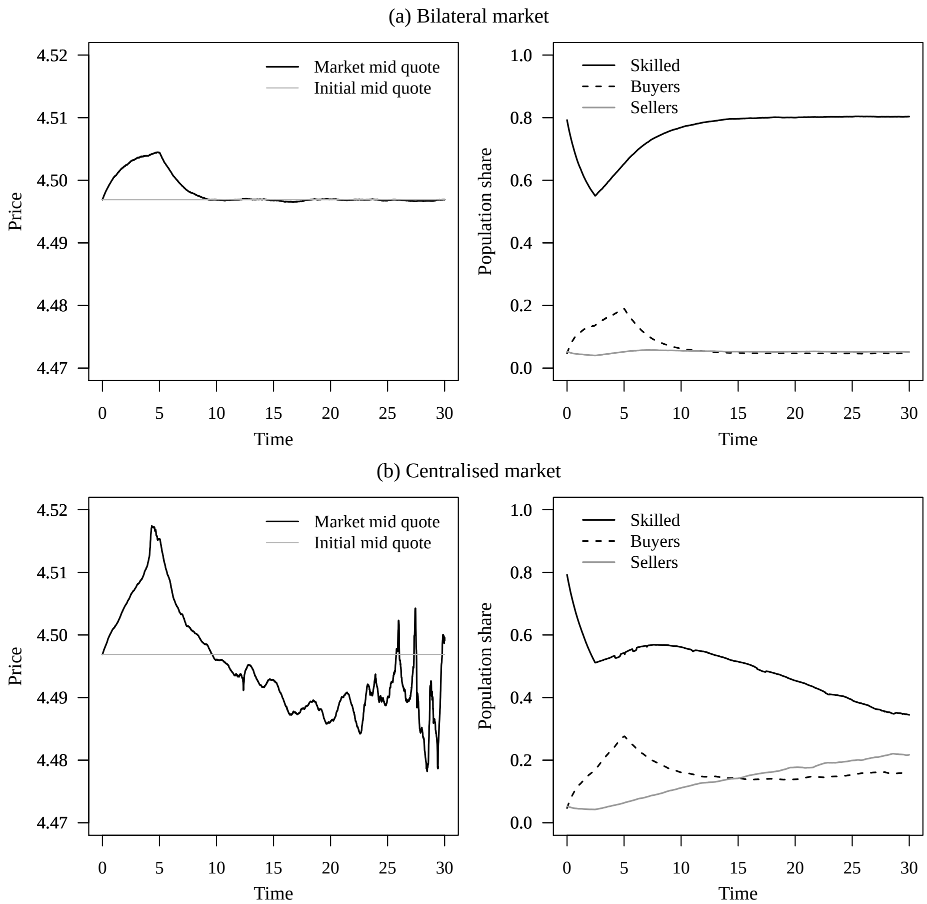

Figure 3: Market price and frequency of trader types in a bilateral market (Panel (a)) and in a centralised market (Panel (b)). Initial population of 4,000 traders taken from a market with bilateral trade. Arrival of bullish traders with constant quote equal to 40 during the first 5,000 rounds (in each round, with probability 0.5, one such trader is added). Time in units of 1,000. Left-hand panels: Time series of the market mid quote and its initial level in the bilateral market. Right-hand panels: Percentage of skilled traders, unskilled buyers (including bullish traders) and unskilled sellers.

Panel (b) in Figure 3 depicts the scenario in which the inflow of bullish traders coincides with a change in market structure from bilateral to centralised trade. The inflow leads to a period during which the option is overvalued. The market mid quote increases more than in the bilateral case because the decline in the proportion of skilled traders is more pronounced. Unskilled sellers outperform unskilled buyers and do at least as well as skilled traders (who are both on the same side of the market as unskilled sellers). Therefore the proportion of unskilled sellers grows at the expense of skilled traders. As soon as the inflow of bullish traders stops, the market mid quote sharply declines and undershoots the initial mid quote. Price volatility increases as skilled traders are replaced by unskilled individuals. This is caused by the diminishing benefit of skill in more centralised markets. To put the negative effect into perspective, we note that initially 76% of the traders are skilled, but by round 30,000 (not shown in figure), their proportion is reduced to 28% and asymptotically becomes 5.6%. In contrast, the proportion of unskilled buyers and sellers increases from 12.6% to 62.4%.

5. Conclusion

The paper demonstrates that the ability of traders to make good decisions, rather than taken for granted, can be considered as an equilibrium phenomenon. To this end, we propose a model where skills are endogenous and study the effect of market centralisation on trader skill. We find that the level of skill decreases as markets become more centralised. This has implications for option price volatility and market resilience. Volatility is lowest, and resilience highest in bilateral markets, but these benefits come at the cost of higher price dispersion across trading venues. Our results suggest that the transparency and cost efficiency of centralised markets must be weighed against higher price volatility and lower resilience in those markets.

By viewing skills as endogenous to markets, our findings highlight a hidden cost of moving the trade of complex assets towards centralised exchanges. While it may be socially desirable to have an asset traded at one price, centralising trade to accommodate unskilled investors can be counterproductive for market stability and price efficiency.

AppendixA. Model convergence and sensitivity

The appendix contains two tests of the numerical model. First, we check whether the model is in a converged state when we extract the data for the analyses. If the model has converged, we can be reasonably confident that our results represent equilibrium phenomena. Second, we check the robustness of our results by measuring the sensitivity of the model to changes in some of the key parameters of the genetic programming algorithm.

AppendixA.1. Convergence

To test convergence, we use data on market pricing errors across two different sample periods. Each sample period consists of 10,000 observations sampled at 10 round intervals. The two samples are 5 million rounds apart. We estimate our model of market pricing errors (model (4) in Table 3) with the addition of a dummy T2 for the second sample period, and the cross-products of that dummy with all the explanatory variables.

Alternatively, we could have used skill or trader pricing errors for the test. However, skill is measured against an endogenous variable, which could then exhibit a trend that would not be picked up by the test. A test based on pricing errors will detect such trends because pricing errors are measured against a fixed benchmark; the Black-Scholes price. We use market pricing errors, rather than trader pricing errors as our test variable, because trader pricing errors are dominated by extreme quotes by unskilled traders, as discussed in Section 4.1.

Table A.9 contains the results of the convergence test.[11] None of the terms

[11]Observe that the coefficients in Table A.9 differ somewhat from their counterparts in model (4) of Table 3, despite the fact that the two models are identical except for the dummy terms in Table A.9, and that the R^2 is higher. The reason is that the data for Table 3 are

Table A.9: Convergence. We run the evolved models for 5.1 million trading rounds and sample data from the first and last 100,000 rounds for every 10th round. This yields two samples of 10,000 observations for each run, 5 million rounds apart. For each run and sample, we compute the mean market pricing error, which is the dependent variable in this regression. The dummy variable T2 represents the second sample period, and the regression includes the cross-products of T2 with all other regressors. An F-test on the terms involving T2 is used to test the null hypothesis that market pricing errors, conditional on market structure and Black-Scholes option values, are the same in both sample periods. Robust standard errors in parenthesis.

	Dependent variable:	
	log(Market pricing error)	
Constant	-0.862^{***}	(0.222)
log(Venue size)	-0.064^{**}	(0.024)
log(Population size)	-0.274^{***}	(0.017)
Fragmented	1.234^{***}	(0.171)
log(Venue size) × Fragmented	-0.291^{***}	(0.025)
log(Black-Scholes price)	0.182^{***}	(0.019)
T2	−0.263	(0.300)
T2 × log(Venue size)	0.008	(0.031)
T2 × log(Population size)	0.018	(0.024)
T2 × Fragmented	−0.022	(0.221)
T2 × log(Venue size) × Fragmented	0.011	(0.033)
T2 × log(Black-Scholes price)	0.018	(0.027)
Observations	1,782	
Adjusted R^2	0.787	
Test of structural break between sample periods H_0: T2 × X = 0 for all ordinary regressors X F-statistic: 2.163 on 6 and 1770 DF, P-value: 0.044		
Note:	$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$	

involving the dummy are significant, hence none of the coefficients seem to exhibit any trend. We next use an F-test to test whether market pricing errors, conditional on market structure and Black-Scholes option values, are the same in both sample periods. This amounts to testing whether the terms involving the dummy are jointly zero. The F-test yields a P-value of 0.044, too large to be significant at the 1% level. We can therefore assume that the model has converged, and that the data for our analyses are extracted from an equilibrium state.

collected at a single point in time, while the data for Table A.9 are averages across 10,000 trading rounds.

AppendixA.2. Sensitivity analysis

We next consider the model's sensitivity to the following key parameters of the genetic programming algorithm: Discount rate, mutation rate and the complexity of trader pricing functions, as measured by the maximal number of programme instructions. We vary the parameters one by one, and quantify the resulting change in skill relative to the base model. The results are obtained by doing 50 model runs for each parameter configuration and market structure. To aid clarity we focus on two market structures, one bilateral, and one fully centralised, both with a population size of 2,000. Table A.10 contains the results. We use raw (untransformed) skill data for the sake of exposition, and non-parametric tests to avoid issues with non-normality.

Table A.10: Sensitivity analysis. Mean skill of traders in venue sizes 2 and 2000 in models with population size 2000. Base case and 6 model variants (higher discount rate, higher mutation rates and less complexity of trading strategies). Mann-Whitney U tests of differences in distributions between the base model and the variants. P-values in parentheses.

Model variant	Base model	Discounting	Mutation rate		Complexity	
		Higher	Higher	Highest	Lower	Lowest
Discount rate (%)	1	2	1	1	1	1
Mutation rate (%)	5	5	50	95	5	5
Program length	256	256	256	256	128	64
			Venue size 2			
Mean skill	25.74	25.77	22.93	20.84	25.19	24.48
Median skill	25.82	25.81	23.33	20.71	25.29	24.74
P-value		(0.637)	(0.000)	(0.000)	(0.000)	(0.000)
			Venue size 2000			
Mean skill	9.69	6.55	9.45	7.82	8.33	7.75
Median skill	9.55	6.37	8.90	6.98	7.53	6.93
P-value		(0.000)	(0.450)	(0.004)	(0.006)	(0.001)
Observations	50	50	50	50	50	50

In the base model, the average level of skill is about 25.5 for the bilateral market structure, and about 9.5 for the centralised structure. We first consider an increase in the discount rate from 1 to 2%. Higher discounting has no effect on skill in the bilateral model, but it reduces skill in the centralised model from 9.5 to 6.5. A higher discount rate reduces the importance of distant events with respect to the ranking of agents in tournaments. Traders in the bilateral market receive feedback on their decisions in every trading round, but traders in centralised markets can vary their decisions widely without getting any feedback. A higher discount rate is therefore likely to have a stronger negative impact on learning in centralised markets.

We next look at the effect of an increase in the mutation rate[12] from 5% to 50% and 95%. In general, higher mutation rates reduce skill in both markets. An increase from 5% to 50% has no significant in the centralised market, but the percentage reduction in skill from an increase in the mutation rate from 5% to 95% is about the same for both market structures. More mutations creates more noise in the evolutionary reproduction process, which inhibits the role of the copy and crossover operations.

Finally, we consider a reduction in the complexity of the traders' pricing functions by reducing their programme size from 256 to 128 and 64. In the bilateral market, smaller programs have a significant but small effect on skill. In the centralised market, the effect is much larger, even in absolute terms. As programme size can be said to limit the reasoning abilities of the traders, this seems to indicate that traders in centralised markets face a more difficult task than traders in bilateral markets, although it is hard to say why this might be the case.

[12]The mutation rate as well as the specific way in which mutations are carried out is known to potentially have a strong impact on convergence, e.g., Maschek (forth), although this is very model-dependent. In our setting, as confirmed by Table A.10, the mutation probability is not a critical parameter. Two aspects of our approach help to avoid the solution getting trapped in a sub-optimal state. (1) We use GP rather than GA, and (2) mutation works as a two-step process. Whether or not a program will be mutated is governed by probability χ_2. If it is, then a second random draw decides whether a (randomly chosen) instruction, operand or operator is changed. Therefore both minor and major random changes can occur.

References

Alchian, A. A., 1950. Uncertainty, Evolution, and Economic Theory. Journal of Political Economy 58, 211–221.

Barber, B. M., Lee, Y.-T., Liu, Y.-J., Odean, T., 2014. The cross-section of speculator skill: Evidence from day trading. Journal of Financial Markets 18, 1–24.

Barber, B. M., Odean, T., 2001. The internet and the investor. Journal of Economic Perspectives 15, 41–54.

Barras, L., Scaillet, O., Wermers, R., 2010. False discoveries in mutual fund performance: Measuring luck in estimated alphas. Journal of Finance 65, 179–216.

Bikhchandani, S., Hirshleifer, D., Welch, I., 1998. Learning from the behavior of others: Conformity, fads, and informational cascades. Journal of Economic Perspectives 12, 151–170.

Binmore, K., 2007. Rational decisions in large worlds. Annals of Economics and Statistics 86, 25–41.

Bogan, V., 2008. Stock market participation and the internet. Journal of Financial and Quantitative Analysis 43, 191–211.

Chen, S.-H., Yeh, C.-H., 2002. On the emergent properties of artificial stock markets: The efficient market hypothesis and the rational expectations hypothesis. Journal of Economic Behavior & Organization 49, 217–239.

Coates, J. M., Page, L., 2009. A note on trader Sharpe ratios. PLoS ONE 4, 1–5.

Coval, J. D., Hirshleifer, D., Shumway, T. G., 2005. Can individual investors beat the market? Finance 0412005, EconWPA.

Ellison, G., Fudenberg, D., 1993. Rules of thumb for social learning. Journal of Political Economy 101, 612–643.

Fama, E. F., French, K. R., 2010. Luck versus skill in the cross-section of mutual fund returns. Journal of Finance 65, 1915–1947.

Fink, J., Fink, K. E., Weston, J. P., 2006. Competition on the NASDAQ and the growth of electronic communication networks. Journal of Banking and Finance 30, 2537–2559.

Grinblatt, M., Keloharju, M., Linnainmaa, J. T., 2012. IQ, trading behavior, and performance. Journal of Financial Economics 104, 339–362.

Grossman, S. J., Stiglitz, J. E., 1980. On the impossibility of informationally efficient markets. American Economic Review 70, 393–408.

Holland, J. H., 1975. Adaptation in Natural and Artificial Systems. MIT Press.

Koza, J. R., 1992. Genetic Programming: On the Programming of Computers by Means of Natural Selection. MIT Press.

Lensberg, T., Schenk-Hoppé, K. R., 2007. On the evolution of investment strategies and the Kelly rule—A Darwinian approach. Review of Finance 11, 25–50.

Lensberg, T., Schenk-Hoppé, K. R., Ladley, D., 2015. Costs and benefits of financial regulation: Short-selling bans and transaction taxes. Journal of Banking and Finance 51, 103–118.

Lo, A. W., 2004. The adaptive markets hypothesis: Market efficiency from an evolutionary perspective. Journal of Portfolio Management 30, 15–29.

Madhavan, A., 1995. Consolidation, fragmentation, and the disclosure of trading information. Review of Financial Studies 8, 579–603.

Maschek, M. K., forth., Economic modeling using evolutionary algorithms: The influence of mutation on the premature convergence effect. Computational Economics. DOI 10.1007/s10614-015-9485-8

Mendelson, H., 1987. Consolidation, fragmentation, and market performance. Journal of Financial and Quantitative Analysis 22, 189–207.

Noe, T. H., Rebello, M. J., Wang, J., 2003. Corporate financing: An artificial agent-based analysis. Journal of Finance 58, 943–973.

Noe, T. H., Rebello, M. J., Wang, J., 2006. The evolution of security designs. Journal of Finance 61, 2103–2135.

O'Hara, M., Ye, M., 2011. Is market fragmentation harming market quality? Journal of Financial Economics 100, 459–474.

Oliven, K., Rietz, T. A., 2004. Suckers are born but markets are made: Individual rationality, arbitrage, and market efficiency on an electronic futures market. Management Science 50, 336–351.

Smith, V. L., Suchanek, G. L., Williams, A. W., 1988. Bubbles, crashes, and endogenous expectations in experimental spot asset markets. Econometrica 56, 1119–1151.

Stoll, H. R., 2008. Future of securities markets: Competition or consolidation? Financial Analysts Journal 64, 15–26.

Yeh, C.-H., 2008. The effects of intelligence on price discovery and market efficiency. Journal of Economic Behavior & Organization 68, 613–625.

Yeh, C.-H., Yang, C.-Y., 2010. Examining the effectiveness of price limits in an artificial stock market. Journal of Economic Dynamics and Control 34, 2089–2108.