Network motifs for microeconomic analysis

Steve Phelps
Centre for Computational Finance & Economic Agents (CCFEA)
University of Essex
sphelps@essex.ac.uk

Katarzyna Musial-Gabrys
Department of Informatics
King's College London
katarzyna.musial@kcl.ac.uk

ABSTRACT

In this paper, we study the microeconomics of the price discovery process in a double auction market for a risky asset. The traditional approach to studying market efficiency relies on analysing aggregate supply and demand functions. However, such analyses are inadequate to capture important details of interactions that occur in electronic markets at high-frequency time scales where only a small number of traders determine the price, and where market microstructure plays a prominent role. Across many other disciplines, network science has been able to explain how complex interactions between microscopic components give rise to macroscopic phenomena. Drawing on these successes, we model transactions between individuals traders as a time-varying directed graph, and we show that the topology of this dynamic network can explain the efficiency of the price discovery process.

1. INTRODUCTION

A common approach to studying the efficiency of auction mechanisms, including the double-auction mechanisms used to run financial exchanges, posits that agents with private valuations reach either a game-theoretic equilibrium, or alternatively a steady-state in their learning of bidding strategies. One then appeals to the properties of the resulting supply and demand schedules to explain the efficiency of the market [2, 5, 19].

In this paper we take an alternative approach which models the financial exchange as an interdependent-values scenario where no single agent has full and accurate information about the fair value of the asset being traded. We take an existing model of financial markets from the literature [8, 12] which allows not only trading based on the fair price, but also speculative trading based on an analysis of past returns. These models have been used to explain empirical phenomena such as fat-tailed return distributions and volatility-clustering [3] that are not easily accounted for by models based on the rational expectations assumptions.

We use this model to try and analyse the factors which contribute to deviations between the observed market-price and the underlying fundamental value of the risky asset. The root mean-squared difference between the log of these two values is our definition of market efficiency for the purposes of this paper, and is a metric for the efficiency of the price-discovery process; that is, it measures to what extent the market is able to track the "true" price of the asset being traded.

The standard approach to explaining price discovery is the microeconomics of supply and demand functions. The different expectations of each trader about the asset's valuation give rise to different supply and demand schedules at the level of individual traders. The market mechanism then attempts to aggregate these individual supply and demand functions and determine the single *equilibrium* price which maximises the volume traded. However, there are two key sources of inefficiency in this process. Firstly, traders may choose not to reveal their true supply or demand in order to attempt to "game" the market. Secondly, because fundamentals can move very quickly, the market mechanism attempts to clear the market by filling orders as soon as a potential match is found in the market — hence the *continuous* double-auction mechanism used in financial exchanges. These factors can cause a discrepency between the market's view of supply and demand, as measured by the prices and quantities of unmatched orders, and the actual supply and demand in the market as represented by individual traders' unobservable supply and demand functions. Several authors refer to this as the "sluggish" matching of supply and demand induced by market microstructure [11].

Given that supply and demand are not accurately tracked by the market mechanism, we explore whether a lower-level description of the operation of the market in terms of actual transactions between pairs of agents has any explanatory power for the efficiency of the market. We model the transactions between agents as a time-varying directed graph, and we explore whether this emergent network can be used to explain deviations from the fundamental price.

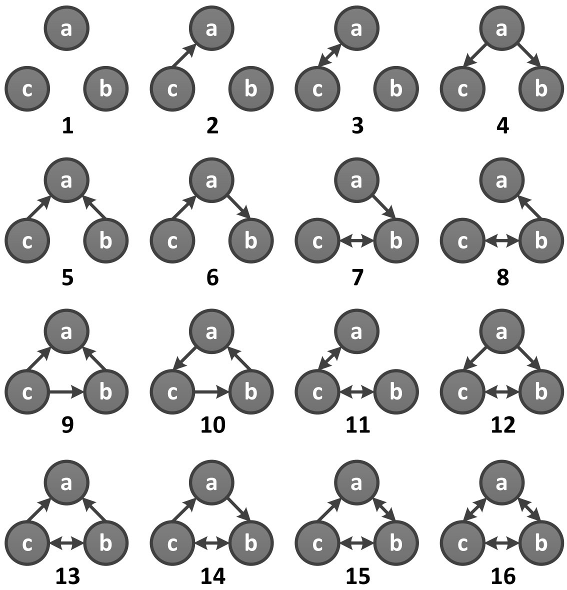
This network is a highly dynamic complex structure and can be described in terms of complex system in which its global topology and characteristics are the results of local interactions between its elements [6]. Thus, we believe that modelling of interactions at the micro level is a key step to advanced analysis of market characteristics.

In order to do this, we use *network motif analysis* — a powerful tool to discover and analyse connections between entities at the local level. Network motifs are small subgraphs (see Figure 1) that reflect local topology and can be used to discover profile and properties of the network.

Motif analysis stems from bioinformatics and theoretical biology [9, 10], where it was applied to the investigation

Appears in:

Figure 1: Motifs consisting of all possible directed triads.



of huge network structures like transcriptional regulatory networks, gene networks or food webs [16, 17]. Although the global topological organization of metabolic networks is well understood, their local structural organization is still not clear. At the smallest scale, network motifs have been suggested to be the functional building blocks of network biology. To date, several interesting properties of large *biological* network structures were reinterpreted or discovered with help of motif analysis [16, 21, 23].

Recent research has seen the development of more sophisticated approaches, among them motif analysis which aims to characterize the network by the difference between its structures and an ensemble of random networks of the same size and degree distribution. A biased distribution of local network structures (subgraphs), a.k.a. network motifs is widely observed in complex biological or technology-based networks. The basic method utilizing such subgraphs is the well–known triad census, allowing to reason about the functional connection patterns of the nodes [22].

In this work, we apply motif analysis to *economics*, and explore to what extent these motifs can explain the efficiency of the price discovery process. That is, we attempt to directly map from an analysis of the interactions between individual agents in the market at the micro-level, and the resulting efficiency of the market at the macro-level. For this study, we use motifs 4 to 16 of all possible directed triads that are depicted in Figure 1.

2. THE MODEL

We use a class of agent-based model that has been demonstrated in previous studies by other researchers to replicate many of the statistical properties observed in empirical data [8, 12, 20][1]. The model simulates not only the detailed micro-structural operation of a financial market, but also how agents form their valuations for the asset being traded based on observations of past prices and the bidding behaviour of others; that is, using the terminology of auction-theory, we model the financial market as an interdependent values scenario.

The market is modelled as an order-driven exchange, typical of that used to trade equities in electronic markets such as the London Stock Exchange (LSE). In an order-driven exchange, agents can submit limit-orders which are offers to buy, or sell, at a specified price and quantity. Orders are matched using a continuous double auction (CDA) [4]. A buy order can be matched with a sell order if the buy price is greater than the sell price, and vice versa. When new orders arrive they are executed immediately at the price of the earliest order, provided that a corresponding match can be found. The fulfilled orders are then removed from the exchange. If orders cannot be matched immediately, they are queued on the "order book" until either they are matched or their expiry time is reached. Buy orders and sell orders each have their own priority queue and are ranked in descending and ascending order respectively. The highest outstanding buy order and the lowest sell order are called the best bid and the best ask respectively, and the pair of prices corresponding to the best bid and ask is called the market quote.

As discussed above, we model agents' valuations using an interdependent-values scenario. We therefore decompose the agents' decision problem into two separate components: (i) calculation of their expectations of the future price of the asset (which determines their valuation thereof); and (ii) their execution strategy, which determines how they place bids or asks in the market as a function of their valuation. Although there has been great deal of work within the agents community examining the latter, there has been relatively little exploration of the inherent inter-dependency of valuations. The fact that financial markets are, in effect, interdependent-valuation auctions is intrinsic to an understanding of phenomena such as speculative bubbles and crashes, and therefore in this paper we concentrate our discussion on agents' expectations rather than their execution strategy.

The expectations model is based on an "adaptive expectations" framework [14] in which agents make trading decisions based on forecasts of the next period price, which are then updated through a social learning process. Agent's expectations are modelled according to a chartist, fundamentalist and noise framework [15]. Chartists believe future prices can be forecast by extrapolating from past prices, in contradiction to the efficient markets hypothesis. On the other hand, fundamentalists believe that there is a fair price for the asset, to which the future price will revert. In a market in which not every agent is rational, the best class of rule to use is not necessarily the fundamentalist forecast. For example, if a significant fraction of the market uses chartist rules, this may create trends away from the fair price, and an

9[1]Although this is not our model, we devote space to formally describing it here in order to make the paper self-contained, and also because there are some subtle technical differences between our implementation and existing models. Firstly, chartist return forecasts are calculated using logarithmic returns rather than simple returns in order to avoid a bias towards positive drift. Secondly, learning occurs probabilistically at every time step rather than being scheduled at every 5000 steps in order to avoid artifacts relating to strong temporal synchronisation of learning. Finally, in line with the learning-classifier system literature, we use an exponential moving average of forecast errors in order to determine fitness rather than a moving window.

agent using a fundamentalist rule might do better by switching to a chartist rule; i.e., chartist expectations can become self-fulfilling. Finally, noise traders form their expectations independently at random and trade on this noise believing that it is a signal.

The model is implemented as a discrete-event simulation of an entire trading day. The trading day is divided into 2×10^5 discrete time steps, each of equal duration. This gives a resolution of approximately 150 milliseconds on a real-time scale. The unequal spacing of events that are typically observed in empirical market data sampled at high-frequency is modelled using a Bernoulli process; at each time step, an randomly chosen agent arrives at the simulation with probability λ. Inter-arrival times are therefore approximately exponentially distributed, as per a continuous Poisson arrival model.

Each agent maintains a single position in the market which it revises according to its valuation policy. The market is populated with a total of n agents. When the i^{th} agent arrives at the simulation it either places a new limit order, or amends its existing order, based on its valuation $v_{(i,t)}$. The price of the order is calculated using a strategy similar to the Zero-Intelligence Constrained (ZI-C) behaviour described in [5]; every agent maintains a randomly chosen markup $\delta_i \sim U(0, \delta_{max})$ and the price of the order is set to

$$\rho_{(i,t)} = v_{(i,t)}(1 + \delta_i) \tag{1}$$

for a sell order, or

$$\rho_{(i,t)} = v_{(i,t)}(1 - \delta_i) \tag{2}$$

for a buy order. The direction of the order (buy or sell) is determined by the agent's expectation of the next period price v_i compared to the current market price p_t: if $v_{(i,t)} > p_t$ then a buy order is placed (the agent takes a long position), otherwise a sell order (short position).

Agents decide their valuations $v_{(i,t)}$ as a function of the financial returns of the asset, i.e. the time-series of changes in the logarithmic prices sampled at a particular frequency:

$$r_{j,\Delta_t} = log(p_{j\Delta t}) - log(p_{(j-1)\Delta_t}) \tag{3}$$

where p_t is the market price observed at time t, and Δ_t is the sampling interval. The market price p_t is defined as either: the price of the transaction that occurred at time t, or the middle of the quote if no transaction occurred.

2.1 Valuations

The valuation $v_{(i,t)}$ of the i^{th} agent is calculated by making a forecast of the next period logarithmic return $\hat{r}_{(t,i)}$. Agents using a chartist policy ($\hat{\text{rc}}$) use a simple moving average of past returns:

$$\hat{\text{rc}}_{(i,t)} = \sum_{j=t/\Delta t - w_i}^{t/\Delta t} r_{j,\Delta_t}/w_i \tag{4}$$

where w_i is the window size used by agent i. Agents using a fundamentalist forecasting rule (rf) have information about the fair value of the asset f_t and forecast accordingly:

$$\hat{\text{rf}}_{(i,t)} = log(p_t) - log(f_t). \tag{5}$$

Agents using a noise-trader rule (rn) make random return forecasts:

$$\hat{\text{rn}}_{(i,t)} = \epsilon_t \tag{6}$$

where ϵ_t are $i.i.d.$ random variates drawn from a standard normal distribution $N(0,1)$.

As in [8], we allow agents to make forecasts using a linear combination of all three types of rule in order to estimate the return over the next time period τ_i. Let

$$\boldsymbol{R}_{(i,t)} = (\hat{\text{rf}}_{(i,t)}, \hat{\text{rc}}_{(i,t)}, \hat{\text{rn}}_{(i,t)}) \tag{7}$$

denote the vector of return forecasts made according to each of the fundamentalist, chartist and noise rules respectively. Let $\boldsymbol{S}_{(i,t)} \in \mathbb{R}^3$ denote a vector of coefficients which define the forecasting strategy of agent i at time t. Then the return forecast $\hat{r}_{(i,t)}$ is given by:

$$\hat{r}_{(i,t)} = \boldsymbol{S}_{(i,t)} \cdot \boldsymbol{R}_{(i,t)}. \tag{8}$$

The valuation of agent i is then given by their forecast of the price at time $t + \tau$:

$$v_{(i,t)} = p_t e^{\hat{r}_{(i,t)}/\tau_i}. \tag{9}$$

This can then be substituted into equations 1 and 2 in order to determine the agent's limit price.

In order to model the arrival of market news, changes in the growth rate of the fundamental price follow a random walk:

$$df_t = \mu_f f_t + \sigma_f f_t \; dW_t \tag{10}$$

where σ_f is the volatility, μ_f is the drift and dW_t is a standard Wiener process discretised over the time interval Δt:

$$dW_t \sim \sqrt{\Delta t} N(0,1) \tag{11}$$

2.2 Initial conditions

The initial values for agents' strategy coefficients are random variables $S_{(i,0)} = (\text{SF}, \text{SC}, \text{SN})$ with distributions:

$$\begin{aligned} \text{SF} &\sim |N(0, \sigma_f)|, \\ \text{CF} &\sim N(0, \sigma_c), \\ \text{SN} &\sim |N(0, \sigma_n)| \end{aligned} \tag{12}$$

where σ_f, σ_c and σ_n are the standard deviations of the fundamentalist, chartist and noise components respectively. In a static experiment treatment these values remain constant over the course of a simulation run. Under a learning treatment, they evolve over time according to the learning model specified below.

2.3 Learning

As in [12, 20], we model adaptive expectations [14] by allowing agents to learn their forecasting strategy $\boldsymbol{S}$ using a model of social learning implemented in the form of a simple co-evolutionary genetic algorithm.

Each agent records the exponential moving average of its forecast error as the market progresses:

$$\bar{\eta}_{(i,t)} = \alpha(\hat{r}_{(i,t)} - r_{(i,t)})^2 + (1 - \alpha)\bar{\eta}_{(i,t-w_\eta)}. \tag{13}$$

The fitness of agent i is then given by

$$\phi_{(i,t)} = 1/(1 + \bar{\eta}_{(i,t)}). \tag{14}$$

Imitation occurs with probability λ_r per time step for any given agent. When imitation occurs, i imitates another agent by randomly selecting a partner j from the remainder of the population with probability proportionate to fitness:

$$p(agent_{j\neq i}) = \phi_{(j,t)} / \sum_{k\neq i} \phi_{(k,t)}, \tag{15}$$

and then agent i inherits its strategy from agent j; that is, $\boldsymbol{S}_(i,t) = \boldsymbol{S}_{(j,t-1)}$.

Mutation occurs with probability λ_m per time step for any given agent. When mutation occurs the agent re-initialises its strategy by redrawing its coefficients $\boldsymbol{S}$ from the distributions specified in Section 2.2.

2.4 Efficiency metric

The efficiency of the market is defined as the negative of the root mean squared difference between the logarithm of the market price and the underlying fundamental price:

$$\text{EA} = -\sqrt{\sum_{t=1}^{D} (\log f_t - \log p_t)^2 / D} \tag{16}$$

2.5 Trade network

We analyse the dynamic network that *emerges* from trade between the agents in this market. This is in contrast to other studies which impose an exogenous network structure on the agents within the market [1]. Rather the graph structure is an endogenous feature of the model which emerges indirectly from decisions taken by the agents. Asset and information then flow around this network, which in turn can effect agents' decisions. Thus the network structure and agents' strategies co-evolve with each other in a similar vein to adaptive network models that have been used to study cooperative behaviour [18].

We denote the total quantity traded between a pair of agents a_i and a_j at time t as $q_{(i,j,t)}$. The edges of the trade network are given by the time-weighted average of the q values. We use an exponential moving average with a time weighting of α together with a threshold parameter γ. The time-weighted average quantity traded between agents is given by:

$$\bar{q}_{(i,j,t)} = \alpha \times q_{(i,j,t)} + (1-\alpha) \times \bar{q}_{(i,j,t-1)} \tag{17}$$

The trade network Q_t is a directed, weighted graph consisting of all pairs of agents whose time-weighted average quantity is above the threshold γ:

$$Q_t = \{(a_i, a_j, \bar{q}_{(i,j,t)}) : \bar{q}_{(i,j,t)} \geq \gamma\} \tag{18}$$

For the results reported in this paper we use $\gamma = 1 - 10^{-2}$ and $\alpha = 10^{-5}$, and we sample Q_t at intervals of 2×10^3 time steps.

3. METHODOLOGY

We analyse the model numerically by simulating the market as a Monte–Carlo simulation. For each simulation we initialise free parameters by drawing random values *i.i.d.* from the distributions listed in Table 3.

In order to account for the fact that these are merely initial conditions and do not tell us how the learning process converges, we also record the mean and standard deviation of agents' strategy coefficients:

Table 1: Summary of experimental treatments

S	Static population with no learning
L	Population with imitation learning
LN	Population with nearest–neighbour learning

Symbol	Definition
$n \sim U(80, 120)$	number of traders
$\sigma_c \sim U(0.1, 3)$	Standard deviation of chartist distribution
$\sigma_n \sim U(0.1, 3)$	Standard deviation of noise-trader distribution
$\sigma_f \sim U(0.1, 3)$	Standard deviation of fundamentalist distribution
$\delta_{max} \sim U(0, 1)$	Maximum value of the markup distribution
$\lambda \sim U(0.1, 0.9)$	Probability of agent arrival per time step
$\lambda_r \sim U(0.1, 0.9)$	Imitation probability per agent per time step
$\lambda_m \sim U(0.1, 0.2)$	Mutation probability per agent per time step
$w_i \sim U(1, 100)$	Chartist window size
$w_\eta \sim U(1, 100)$	Sampling interval for forecast errors
$\tau_i \sim U(5, 10)$	Forecast time horizon
$\sigma_f \sim U(0, 1.0)$	Volatility of fundamental price process
$\mu_f \sim U(0, 0.1)$	Drift of fundamental price process
$p_0 \sim U(100, 100)$	Initial price

Table 2: Summary of model parameters and their associated distributions

$$\bar{\boldsymbol{S}}_t = \sum_{i=1}^{n} \boldsymbol{S}_{i,t} / n \tag{19}$$

$$\boldsymbol{\sigma}_{\boldsymbol{S}_t} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\boldsymbol{S}_{i,t} - \bar{\boldsymbol{S}}_t)^2} \tag{20}$$

We average these over time for each simulation run in order to summarise the steady-state contribution of each of the fundamentalist, chartist and noise-trader signals, giving the vectors $\bar{\boldsymbol{S}} = \sum_{t=1}^{D} \boldsymbol{S}_t / D$ and $\boldsymbol{\sigma}_{\boldsymbol{S}} = \sum_{t=1}^{D} \boldsymbol{\sigma}_{\boldsymbol{S_t}} / D$ as additional independent variables which we can use for later analysis.

We run the simulation under a total of three different experimental treatments, as summarised in Table 1. The basic treatment, static (S), involves no learning. We compare this with two learning treatments: learning (L) in which agents have the ability to imitate any other agent in the market, and learning-networked (LN) in which agents can only imitate their nearest neighbours on the graph Q_t.

For each treatment we run a total of 2×10^2 independent simulations, recording the randomly–chosen free parameters for later analysis. Through each simulation, we sample the trade-network Q_t at intervals of 2×10^2 time steps, and we record the relatative frequencies of motifs 4 to 16 for those nodes that have at least one connection, along with the value of EA calculated over the corresponding window.

Table 3: Adjusted R^2 values

	S	L	LN
without network metrics	0.57	0.69	0.67
with network metrics	0.80	0.75	0.86
percentage-point increase	0.23	0.06	0.19

Table 4: Summary of OLS model over all treatments

Parameter	Estimate	Std Error	t value
$\hat{S}_1$ (fundamentalist)	9.249e-01	3.184e-03	290.501
σ_{S_1}	-2.889e-01	2.813e-03	-102.700
$\hat{S}_2$ (chartist)	3.228e-01	5.955e-03	54.218
σ_{S_2}	-6.480e-03	3.436e-04	-18.857
$\hat{S}_3$ (noise)	-1.011e-01	7.780e-03	-12.988
σ_{S_3}	-6.696e-02	5.040e-04	-132.859
Clustering coefficient	1.471e+00	1.225e-02	120.037
Network diameter	2.373e-02	7.069e-04	33.574

4. RESULTS

Figure 2 shows the median value of EA over all simulation runs according to each experimental treatment. The learning treatment (L) shows a significantly higher efficiency compared with any other experimental condition. However, when imitation occurs over the trade network (treatment LN) this effect is reduced.

Whilst it is intuitive that learning might play a role in maintaining prices closer to fundamentals, there is still considerable variance within each treatment. In order to try and explain the factors which contribute to efficiency we run a multiple linear regression with EA as the dependent variable, and all free parameters listed in Table 3 together with the first two moments of the actual population strategy coefficients ($\bar{\boldsymbol{S}}$ and $\boldsymbol{\sigma_S}$) as independent variables

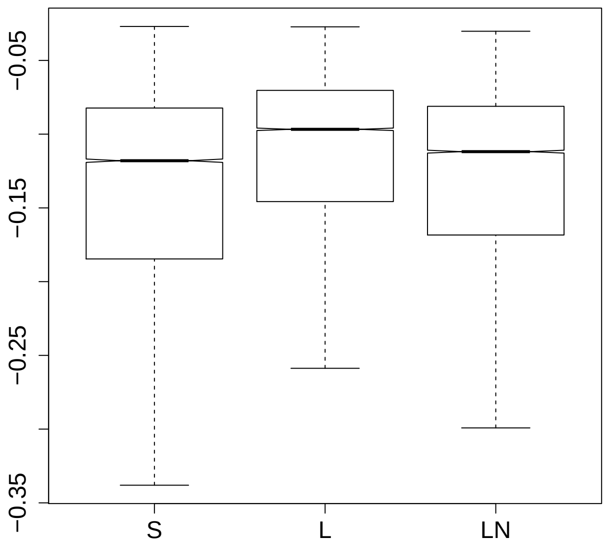
As expected, under all treatments, the OLS regression analysis shows that the configuration of strategies in the population $\bar{\boldsymbol{S}}$ and $\boldsymbol{\sigma_S}$ explains much of the variance in the final value of EA with significance levels for these estimates less than 2×10^-16. On the other hand, the free parameters governing initial conditions are not found to play a role. What is highly interesting, though, is that when we additionally include the values of the clustering coefficient and network diameter averaged over time, we significantly increase the explanatory power of the statistical model. Table 3 summarises the adjusted R^2 values that we obtain when we regress with and without the time-averaged global network metrics of the trade network defined by equation 18.

Thus the global network metrics show a remarkable ability to predict the efficiency of the market. Figure 3 shows the clustering coefficient plotted against efficiency as defined by equation 16 on a log-log scale.

In order to further understand how the network topology influences the efficiency of the market we analyse how each individual network motif enumerated in Figure 1 contributes to EA. The calculated clustering coefficient measures the proportion of closed triangles, in which relations are undirected, relative to the total number of possible triads. It is possible that particular configurations of closed triads influence the efficiency of the market more than others.

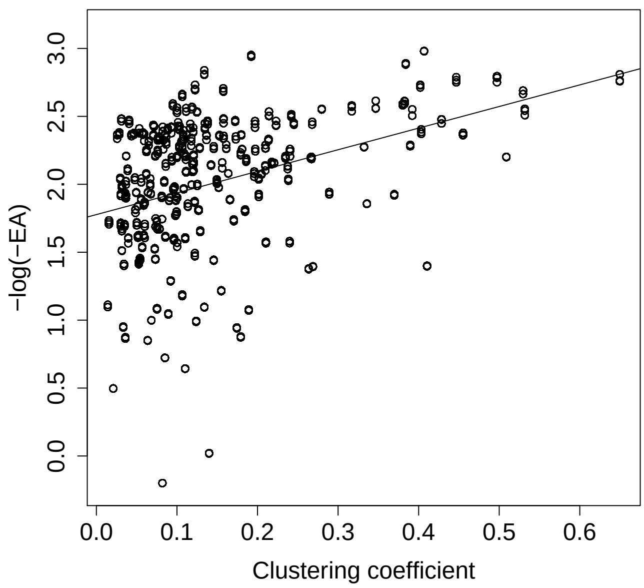
The prevalence of particular motifs is relatively low to those found in, for example, biological networks, as the trade networks are very sparse. The comparative rarity of particu-

Figure 2: Efficiency (EA) by treatment



The above boxplots show the median value of EA under each experimental treatment summarised in Table 1. The centre line of each boxplot shows the median surrounded by the 25th and 75th percentiles. The inward notches show the 95% confidence intervals for the median, and the whiskers show the minimum amd maximum values, excluding outliers.

Figure 3: Clustering coefficient against market efficiency (log scale).



Each observation corresponds to an independent run of the simulation with initial parameters drawn i.i.d. from the distributions in Table 3. Because the graphs are changing over time, we take the mean clustering coefficient averaged over the entire duration of the simulation as the representative value of the clustering coefficient. The regression line shows the best linear univariate model ($R^2 = 0.14$, intercept = 1.78, slope = 1.59).

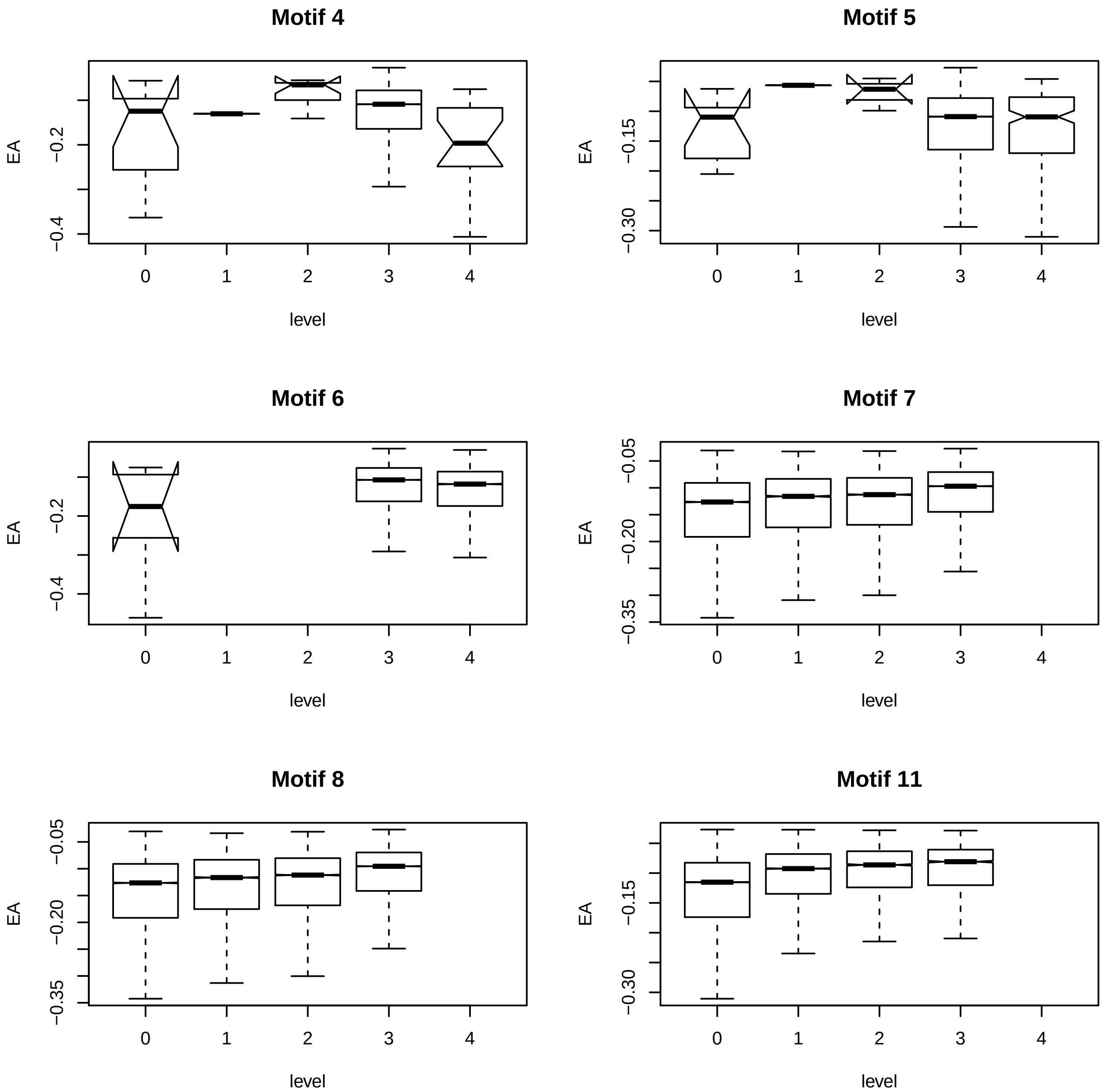


Figure 4: EA **against motif level (non-triangular)**
Each graph above shows the how the prevalence of a particular network motif affects the distribution of EA. The prevalence of each motif is a scalar in the interval $(0, 1)$ representing the proportion of all possible triads that contain the named motif. These are binned into the following levels: 0, $[0, 0.01)$, $[0.01, 0.02)$, $[0.02, 0.5)$, and $[0.5, 1.0)$. The boxplot for each level shows the median value of EA in the centre surrounded by the 25th and 75th percentiles. The inward notches show the 95% confidence interval for the median.

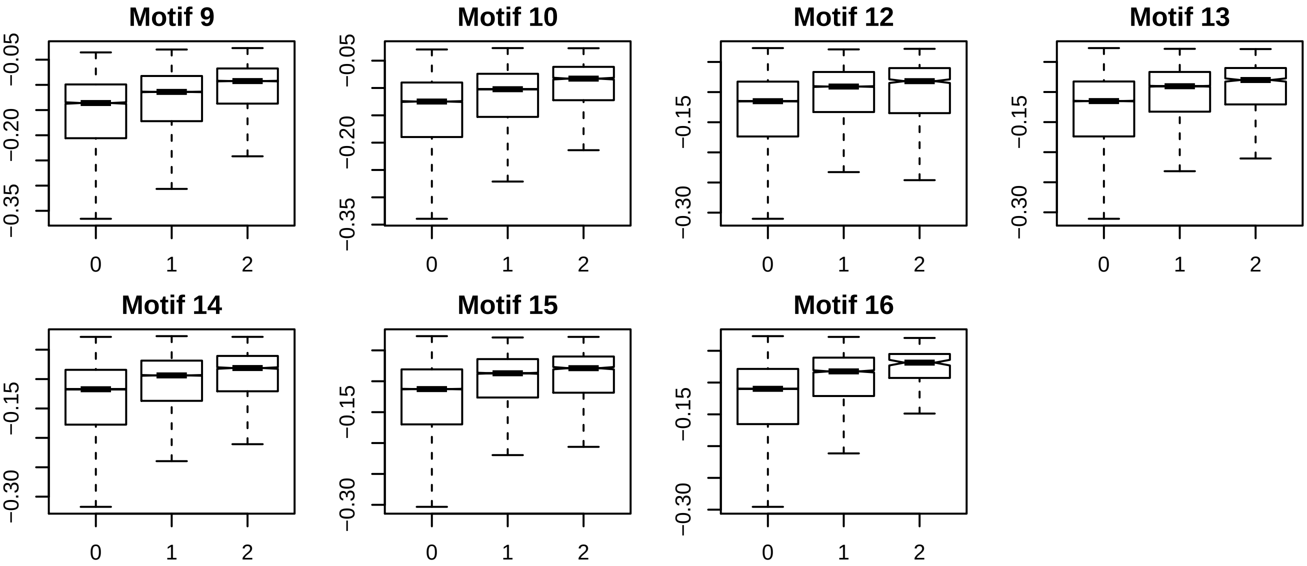


Figure 5: EA **against motif level (triangular motifs)**
Each graph above shows the how the prevalence of a particular network motif affects the distribution of EA. The prevalence of each motif is a scalar in the interval $(0, 1)$ representing the proportion of all possible triads that contain the named motif. These are binned into the following levels: 0, $[0, 0.02)$, $[0.02, 1.0)$. The boxplot for each level shows the median value of EA in the centre surrounded by the 25th and 75th percentiles. The inward notches show the 95% confidence interval for the median.

lar motifs makes an OLS regression against EA impractical. For this reason, we treat the prevalence of a particular motif as a categorical variable by binning into five discrete levels so that we can see whether a high prevalence of particular motif has any statistically significant effect compared to its absence. Figures 4 and 5 summarise the distribution of EA values obtained for each level of the corresponding motif, across all experimental treatments.

5. DISCUSSION

Motifs 4 and 5 appear to have a negative effect on efficiency at high-levels. This can be intuitively explained since in each of these motifs trade is occurring in a unilateral direction. This could could occur either in a strong bull or bear market where we might expect speculative strategies to come into play causing bubbles. Alternatively, high prevalence of these motifs could also indicate situations in which there is large competition on one side of the market with very little liquidity on the other side.

As might be expected given the correlation between the clustering coefficient and EA closed triads clearly have a positive effect on efficiency. This is particularly surprising in the case of Motif 10, which corresponds to closed loop. In other types of network the presence of this motif can indicate feed-forward loops. Our original hypothesis was that Motif 10 would be an indication of feedback in the market corresponding to speculative trading. In contrast, our results show that this motif actually corresponds to an increase in the efficiency of the market.

In addition to motifs corresponding to closed-triads, motifs 7, 8 and 11 *also* appear to contribute to efficiency when present at sufficiently high levels.

The common feature shared by motifs which have a positive effect on market efficiency is that they indicate the presence of agents who are trading in both directions within the same time window (2000 time steps). This bilateral trading can either be with the same agent (as in motifs 7, 8, 11, 12 and 16), or with different trading partners *who are themselves actively trading* (motifs 9, 10). These patterns suggest the presence of liquidity-provision behaviour, also known as market-making, whereby traders simultaneously buy low and sell high (or vice versa) within the market-quote over short time horizons, resulting in additional liquidity provision to the market. Although there are no explicit market-making strategies in our agent-based model, a combination of chartist and noise expectations can serve to provide liquidity on the other side of the market when fundamentalist expectations indicate to buy or sell in order to correct the price back to fundamentals. Indeed, in a market populated solely by agents using the fundamentalist rule, we see a highly *inefficient* market: in a similar manner to the Grossman-Stiglitz paradox [7], all agents have the same expectations about the future price and therefore there are no counter-parties on the other side of the market with which to trade. This also explains why the chartist co-efficients have a small positive effect on efficiency (see Table 4). Chartist expectations are necessary for liquidity provision, *so far as they enable fundamentalists to correct the price when it is both above and below the fair value.* It is this latter condition that is captured by our motif analysis, and which explains the additional explanatory power of the trade network metrics in addition to the expectations of the agents.

6. CONCLUSION

In this paper we have analysed the emergent dynamic trade-network that is formed when pairs of agents trade with each other in a continuous double-auction market. We have shown that both global and local properties of this emergent network can be used to explain the efficiency of the price discovery process in an agent-based model incorporating both speculative and fundamental expectations.

A key part of our analysis has been the use of network motifs to study the microscopic behaviour of the trade network. To the best of our knowledge, this is the first application of the motif analysis to the study of emergent microeconomic networks.

We found that the motif analysis can detect patterns of trading indicative of market-making and liquidity-provision. By analysing the prevalence of such trading patterns, we found that we can better explain the efficiency of the market.

Our work highlights the potential of network and motif analysis for studying the transactions that occur in real financial markets. Although empirical high-frequency tick data is readily available, it is usually anonymised in such a way that individual identities of traders are not distinguishable. However, some researchers do have access to datasets with hashed membership codes [13], and an interesting possibility for future research would be to apply our analysis to such data.

Acknowledgements

We would like to thank Wing Lon Ng for contributing many helpful discussions and ideas.

7. REFERENCES

[1] S. Alfrano and M. Milakovic. Network structure and N-dependence in agent-based herding models. *Journal of Economic Dynamics and Control*, 33(1):78–92, Jan. 2009.

[2] D. Cliff and J. Bruten. Minimal-Intelligence Agents for Bargaining Behaviors in Market-Based Environments. Technical report HPL-07-91, HP Labs, 1997.

[3] R. Cont and J.-P. Bouchaud. Herd behavior and aggregate fluctuations in financial markets. *Macroeconomic Dynamics*, 4:170–196, 2000.

[4] D. Friedman and J. Rust, editors. *The Double Auction Market: Institutions, Theories, and Evidence (Proceedings of the workshop on double auction markets held June 1991 in Santa Fe, New Mexico)*. Westview, 1993.

[5] D. K. Gode and S. Sunder. Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. *Journal of Political Economy*, 101(1):119–137, 1993.

[6] T. Gross and H. Sayama. *Adaptive networks: Theory, models and applications.* Springer-Verlag, Heidelberg, Germany, 2009.

[7] S. Grossman and J. Stiglitz. On the impossibility of Informationally Efficient Markets. *American Economic Review*, 70:393–408, 1980.

[8] G. Iori and C. Chiarella. A Simulation Analysis of the Microstructure of Double Auction Markets. *Quantitative Finance*, 2:346–353, 2002.

[9] S. Itzkovitz, N. Kashtan, G. Ziv, and U. Alon. Subgraphs in random networks. *Physical Review E.*, 68, 2003.

[10] N. Kashtan, S. Itzkovitz, R. Milo, and U. Alon. Efficient sampling algorithm for estimating subgraph concentrations and detecting network motifs. *Bioinformatics*, 20(11):1746–1758, 2004.

[11] D. Ladley and K. R. Schenk-Hoppe. Do Stylised Facts of Order Book Markets Need Strategic Behaviour ? *Journal of Economic Dynamics and Control*, 33(4):817–831, 2009.

[12] B. LeBaron and R. Yamamoto. Long-memory in an order-driven market. *Physica A: Statistical Mechanics and its Applications*, 383(1):85–89, Sept. 2007.

[13] F. Lillo, E. Moro, G. Vaglica, and R. Mantegna. Specialization and herding behavior of trading firms in a financial market. *New Journal of Physics*, 10(043019):1–15, 2008.

[14] A. W. Lo. Reconciling Efficient Markets with Behavioural Finance: The Adaptive Markets Hypothesis. *Journal of Investment Consulting*, 7(2):21–44, 2005.

[15] T. Lux and M. Marchesi. Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature*, 397(February):498–500, 1999.

[16] S. Mangan, A. Zaslaver, and U. Alon. The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks. *Journal o. Molecular Biology*, 334:197–204, 2003.

[17] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chlovskii, and U. Alon. Network motifs: simple building blocks of complex networks. *Science*, 298:824–827, 2002.

[18] S. Phelps. Emergence of social networks via direct and indirect reciprocity. *Journal of Autonomous Agents and Multiagent Systems*, (forthcoming), 2012.

[19] S. Phelps, S. Parsons, and P. McBurney. An Evolutionary Game-Theoretic Comparison of Two Double-Auction Market Designs. In P. Faratin and J. A. Rodriguez-Aguílar, editors, *Agent-Mediated Electronic Commerce VI*, pages 101–114. Springer Verlag, 2005.

[20] N. Rayner, S. Phelps, and N. Constantinou. Testing adaptive expectations models of a double auction market against empirical facts. In *Lecture Notes on Business Information Processing: Agent-Mediated Electronic Commerce and Trading Agent Design*, Barcelona, 2011. CCFEA, Springer (in press).

[21] A. Vazquez, R. Dobrin, D. Sergi, Z. N. Eckmann, J.-P.and Oltvai, and A. Barabasi. The topological relationship between the large-scale attributes and local interaction patterns of complex networks. *Proc. of the National Academy of Science, USA*, 101(17):940, 2004.

[22] S. Wasserman and K. Faust. *Social Network Analysis: Methods and Applications.* Cambridge University Press, New York, 1994.

[23] E. Young-Ho, L. Soojin, and J. Hawoong. Exploring local structural organization of metabolic networks using subgraph patterns. *Journal of Theoretical Biology*, 241:823–829, 2006.