AGENT-BASED SIMULATION OF HERDING IN FINANCIAL MARKETS

Xin Liu
Department of Management Science
Lancaster University Management School
Lancaster, LA1 4YX, U.K.
x.liu1@lancaster.ac.uk

Dr. Roger J. Brooks
Department of Management Science
Lancaster University Management School
Lancaster, LA1 4YX, U.K.
roger.brooks@lancaster.ac.uk

ABSTRACT

There are several models of financial markets which look at the herding effect. This is a situation where many market traders act as a herd in that they all behave in a similar way with their trading. This type of behaviour may explain certain observed characteristics (or ‘stylised facts’) in real markets. However, the various models have different herding mechanisms and market settings. This paper sets out the rationale of our approach and our initial work in trying to get a better understanding of herding in financial markets. Our research, though, is at an early stage. The basic methodology is to reproduce and compare some of the existing models, hopefully leading to a more general understanding and measure of herding and the relationship with market behaviour. One model has been investigated so far and this is described. A more general issue is the research importance of reproducing previous studies.

Keywords: Agent based simulation, Financial markets, Herding, Gurus, Reproducibility

1 INTRODUCTION

This introduction begins with a definition of agent-based simulation of financial markets, moving on to herding. Then the objective of the research and the chosen methodology are discussed. The initial herding model is described in section 2 with the outputs set out in section 3.

1.1 Agent-based simulation and financial markets

Agent-based simulation is a simulation technique based on the agent. There is no agreed definition of an agent. However, one possible definition is an entity which models a cognitive process such as an individual’s intention and belief (Edmonds and Mohring, 2005). For example, in the context of the economy, consumers, companies, regulators and governments could be agents.

Financial markets are markets where funds get exchanged. Stock and bond markets are examples of financial markets. Several empirical studies of financial markets have found common statistical features across financial instruments like cash or bonds, and across different time scales (Cont, 2001). These common statistical features are called stylized facts which financial market models are expected to produce. There are three well known stylized facts: real-returns follow a non-Gaussian distribution; absolute value or square value of real returns for each period are correlated; real returns for each period are not correlated (Taylor, 2005). Real returns are non-Gaussian in that compared to the normal, the distribution is more sharp and narrow in the middle with fat tails. This means that it has a positive kurtosis (Cont, 2005). Correlations of absolute value or square value of real returns is often described as volatility clustering. Big returns tend to be followed by big returns and small returns by small returns and so there are periods where the market is very volatile and periods where it is fairly stable (Cont, 2005). On the other hand there is little or no correlation in whether successive periods give a positive or negative return, hence real returns are not correlated.

Agent-based simulation in financial markets models the behaviour of the different traders with their interactions generating the market and determining the market price. The behaviour may include learning and adaptation. Agent based simulation in financial markets is mainly divided into two types: N-type models and autonomous agent models (Chen et al 2012). The N-type models (Lux, 1998;
Liu and Brooks

Brock and Hommes, 1998) divide the agents into N different types of beliefs although their individual characteristics with each group may vary. The autonomous agent models (Arifovic, 1996; LeBaron, 2001) divide individuals into different rules.

Unfortunately, there is no agent based modelling that generates all stylised facts. Herding has been used in agent based simulation of financial markets for explaining some of stylised facts and this is discussed in the next section.

1.2 Herding

Herding describes a behaviour where groups of people keep making similar decisions. This may be due to some type of interaction between them or just because they are using similar decision making rules. From a review of 20 models in the literature which focus on herding behaviour, there are neither general model mechanisms nor herding mechanisms. The agents in the market differ from N-type (Carro et al 2015; Lux 1995) to autonomous type (Lebaron and Yamamoto, 2007; Mauri and Tettamanzi, 2012). In addition, there are different rules of price setting like excess demand (Lux and Marchesi, 1999; Chowdhury and Stauffer, 1999; Alfarano et al 2005), and order-driven market (Lebaron and Yamamoto, 2007; Tedeschi et al 2012). In the Yamamoto (2011) and Kaizoji et al (2011) models, two stocks are considered in the market while in others only one stock is considered.

The most important part which decides the herding factor is also different in the literature. Some papers (Alfarano et al, 2005; Alfarano and Milakovíc, 2009; Carro et al., 2015) are based on Kirman’s idea of ant behaviour (1993) when facing two identical foods. The herding factor in these models is based on a discrete choice using a probability factor which usually affects the transition probability of two groups such as pessimistic and optimistic traders. Some papers (Lux, 1995; Lux, 1998; Lux and Marchesi, 1999) describe a herding factor with a continuous time discrete model which is also similar to Kirman’s idea but with different probability formulae. Unlike Kirman’s switching type model, there are several models mainly based on social imitation. The type of the agents in these models does not change, and the herd effect is modelled through impacting the agent’s decisions. The imitation rules in these models vary: the agents in the spin models (Chowdhury and Stauffer, 1999; Bornholdt, 2001; Kaizoji et al., 2002) imitate their nearest neighbours; the agents in the models (Markose et al., 2007; Kaizoji et al., 2011; Tedeschi et al., 2012; Mauri and Tettamanzi, 2012; Yang et al., 2012) are influenced by other agents such as the opinions of majorities or of successful agents; some autonomous models (Labaron and Yamamoto, 2007; Yamamoto, 2011; Chen and Yeh, 1999) set the herd mechanism through a genetic algorithm or genetic programming learning; Some models divide the agents into different groups with different size to replicate the herd by clustering (Chen et al, 2013; Lee and Lee, 2013; Manahov and Hudson, 2013).

1.3 Objective and Method of Research

The overall objective of this research is to improve the understanding of the nature and effects of herding in financial markets. Herding, as an important behaviour in financial behaviour studies, is a likely partial explanation of some of the observed stylized facts like volatility clustering and fat tails. As discussed in section 1.2, there are several papers modelling herding but with different mechanisms and also applied to different situations. It would be very useful to find a more general measurement of herding, irrespective of the specific herding mechanism, where the same herding level produces the same general market behaviour. The initial research approach therefore is to make a comparison of some of the previous models from the literature with a variety of herding mechanisms.

A ‘reproducing’, ‘repeating’ or ‘replicating’ method is taken by which we mean attempting to build a model from a previous study, as described in the paper of that study, and hopefully obtain similar results. Although our overall aim is an improved general understanding of herding we believe that reproduction of studies is a valuable research undertaking in itself and one that is not done enough in simulation and Operational Research (OR), and perhaps even across science in general.

Repeating experiments is a key aspect of the scientific method. Karl Popper’s (Popper, 2002) view was: “We do not take even our own observations seriously, or accept them as scientific observations, until we have repeated and tested them. Only by such repetitions can we convince ourselves that we are not dealing with a mere isolated 'coincidence', but with events which, on account
of their regularity and reproducibility, are in principle inter-subjectively testable. Every experimental physicist knows those surprising and inexplicable apparent ‘effects’ which in his laboratory can perhaps even be reproduced for some time, but which finally disappear without trace.” In a traditional natural science experiment, then, replication increases the confidence in the generalisability of the results. A single experiment may produce misleading results due to various factors such as errors, the specific conditions, or chance. A recent large scale project reproducing 100 psychology studies from top journals found that “only 39% of effects were subjectively rated to have replicated the original result” with the mean effect size being just half that of that of the original results (Open Science Collaboration, 2015). Very low levels of reproducibility have also been found in oncology studies in medicine (Begley and Ellis, 2012). One issue is a publication bias in that papers are more likely both to be submitted and published if they contain novel and unexpected (i.e., ‘exciting’) results. This can lead to bias in the method such as the selective reporting of results, and under certain assumptions the chance of a study being correct may be very low (Ioannidis, 2005). Of course, the peer review system does not involve checking the details of the study and its effectiveness is arguable (Smith, 2015).

Similar issues can potentially apply to a simulation or OR study. Results could be misleading due to errors, specific (perhaps implicit) assumptions, the specific conditions (e.g., parameter values or the scenario), and the experimentation and analysis process. Hence, one benefit and reason for reproducing a simulation study is a verification check of the model. Even high profile studies from famous scientists can contain basic errors. Reinhart and Rogoff (2010) published an economics paper on the relationship between debt and GDP for countries. The results have been quoted by key policymakers in the debates on austerity and economic policy. However, student Thomas Herndon found when attempting to reproduce the results that simple mean and median formulae in the original Excel spreadsheet referred to the incorrect cell range and so accidentally excluded five countries (Herndon et al., 2014). A second benefit is that reproducing a study should show up the assumptions being made. This also applied in the Herndon et al. (2014) reproducing study where they found that in the original study some data had been excluded and also that in the calculations for different debt / GDP categories a single average value was used for each country irrespective of the number of years of data. They were critical of both of these assumptions.

In addition to these benefits a simulation reproducing study can go further than simply repeating the original study. New results can be generated through different experiments, new analysis, additional outputs etc. This may also lead to new insights and to a greater depth of understanding of the reasons for the model behaviour.

The Wilensky and Rand (2007) paper has a good discussion of the nature and benefits of reproducing simulation models in the context of agent-based simulation in the social sciences. They list six dimensions which may differ between the original and reproducing study: “time, hardware, languages, toolkits, algorithms, authors.” Of course one problem is that if the study cannot be reproduced then it may not be clear whether there is an error or whether it is due to one of these dimensions. Another practical problem is that the original paper may not contain all the details of the model. What is required is the conceptual model – i.e., a complete software independent description of the model (Brooks and Robinson, 2001). In forecasting, Boylan et al. (2015) were unable to reproduce a previous study and one issue was insufficient information on the methods and data used.

Based on the above discussion, the five main steps planned for our research are as follows:

1. Selecting: several herding papers will be selected based on the level of model description and the model mechanism.
2. Checking: the models will be built based on the papers to check if the results can be reproduced or not.
3. Extending: then an extended model will be produced to obtain a more detailed understanding of the models through new experiments and analysis.
4. Comparing: after reproducing several papers, the results will be compared to see under what conditions similar results are obtained.
5. Generalising: attempt to generate a general herding measure or rule.

To date, one paper has been selected and the next section describes our current progress in reproducing it.
2 THE MODEL

This model is based on the Tedeschi et al. (2012) paper called ‘Herding effects in order driven markets: The rise and fall of gurus’. The model uses zero intelligent agents which stems from Gode and Sunder (1993) where agents trade according to their random behaviour. However, each agent is influenced by one other agent who is most likely to be a successful trader. The guru is the trader who has the most imitators. Each trader can view the current bid (buy) and ask (sell) price and submit a market order or limit order. A market order is an order which can be traded immediately fully or partially. A limit order is an order which cannot be traded immediately and is added to the order book. All parameter settings follow the original paper unless there is no information in the paper. The traders in the model have the same situation at the start. They all have 100 stock at price 1000 and 100,000 cash.

2.1 Reproducibility issues

To date, we have not managed to reproduce the results from the original paper. One issue is that although the paper contains a good level of detail in the description of the model, two of the parameter values are not specified. Another issue may be that the trader’s behaviour is based on a utility function which may make the model quite sensitive to the specific parameter values. In the results in the Tedeschi paper we also observe that the market price starts at 1000 but very quickly drops to about 500 and then fluctuates around that level. This initial drop is not particularly realistic if we take the initial conditions to be a stable situation. In view of these difficulties we decided to change to trading behaviour from the utility function to alternative rules that we considered are reasonably realistic of typical trading behaviour. This leaves the essential herding part of the model unchanged.

The model description below describes our current model. Sections 2.2 (network) and 2.3 (expectation) therefore follow Tedeschi et al. (2012), with section 2.4 (market mechanism) being our revised market trading rules.

2.2 The Network

The network is the starting point of this whole model to build a communication picture among the agents who are traders in this model. All agents are nodes in this network and the edges are the communication links. For each node, there is just one out-going link to keep it simple. This constrains that one agent can just get advice from one other agent.

In all the formulae below, the superscript $i$ is the particular agent $i$ from 0 to 149 (total of 150 agents) and the subscript $t$ is the time period $t$ from 0 to 1000. The wealth is equal to the current value of the stock plus the cash holding in equation (1):

$$W^i_t = S^i_t p_t + C^i_t$$  \hspace{1cm} (1)

In equation (1): $W$, $S$, $C$ and $p$ with just subscript are the symbols of wealth, stock, cash and price. Then, $W^i_t$ is the wealth for agent $i$ at time period $t$.

The fitness function measures the level of wealth for each agent in equation (2):

$$f^i_t = \frac{W^i_t}{W^\text{max}_t}$$  \hspace{1cm} (2)

In equation (2), $f$ represents the fitness. $W^\text{max}_t$ is the wealth of the agent who has the maximum wealth among all agents at time $t$.

Then the probability function based on the fitness forms the whole communication network. For each agent, there is one random assigned neighbour that is a potential newly formed link at the beginning of each period. Each agent is faced with a choice of keeping the existing neighbour or linking to the newly formed neighbour. The probability function (3) decides the probability of switching to the new link.

$$p^i_{t} = \frac{1}{1+e^{-\beta^i t(f^i_t-f^k_t)}}$$  \hspace{1cm} (3)

The $p^i_t$ with superscript in equation (3) is the symbol of agent $i$’s probability of switching to the newly formed link. $\beta^i$ is a random number that follows the uniform distribution ($5, 45$). This random number protects from locking to imitate the same guru. The superscript $k$ is the existing neighbour of agent $i$ and $j$ is the newly random formed neighbour of agent $i$. 
2.3 The Expectation

The agent’s expectation of future returns is based on the agent’s own idiosyncratic expectation and the neighbour’s expectation. The returns in these formulae are spot returns with time interval from time \(t\) to time \(t + \tau\). The agent’s idiosyncratic expectation is based on one volatility factor and one other normal noise. The volatility factor is:

\[
\sigma^i_t = \sigma^0_i (1 + l_{0t}^i (1 - w)) \tag{4}
\]

The \(\sigma\) is the volatility, \(w\) is the herding factor explained below and \(l_{0t}^i\) is the percentage of incoming links for agent \(i\) at time period \(t\). \(\sigma^0_i\) is a uniform distributed factor for agent \(i\) from 0 to \(\sigma^0\) where \(\sigma^0\) is a uniform distribution from 0 to 0.07. Then the idiosyncratic expectation is based on the result obtained from formula (4) multiplied by the normal noise:

\[
\hat{r}_{t,t+\tau}^i = \sigma^i_t \epsilon^i_t \tag{5}
\]

The \(\hat{r}\) is the symbol of agent’s idiosyncratic expected return and \(\epsilon\) is the normal noise with normal distribution \(\mathcal{N}(0,1)\) with mean 0 and standard deviation 1.

Then the return of each agent follows:

\[
r_{t,t+\tau}^i = w \hat{r}_{t,t+\tau}^i + (1 - w) \hat{r}_{t,t+\tau}^j \tag{6}
\]

In this formula, \(r\) is the symbol of return and \(\hat{r}_{t,t+\tau}^j\) is the neighbour \(j\)'s idiosyncratic return when agent \(j\) is the neighbour of agent \(i\). The herding factor \(w\) in equation (4) and (6) takes values between 0 and 1. Lower \(w\) means more herding.

2.4 The Market

The future price for each agent is also formed based on their expectation. The expected price at time \(t + \tau\) when the agents are currently at time \(t\) is:

\[
\hat{p}_{t,t+\tau}^i = p_i (1 + r_{t,t+\tau}^i) \tag{7}
\]

The \(\hat{p}\) here in the equation (7) is the future price. At each time, each agent always has just one order in the market. The discount factor in this model is ignored. They compare the current price with the expected price from equation (7) to submit a buy or sell order. Obviously, if the current price is greater than the expected price, traders sell now and hope to buy in the future to make a gain; otherwise, the traders buy now at a lower current price and hope to sell at the higher future price to make profits. When the difference between the current price and expected price is greater than or equal to 50, the buy order price for certain agent is the current price plus 50 and the sell order price is the current price minus 50 to ensure some gains. Otherwise, traders submit limit orders at buy order prices or sell order prices which are expected price minus 50 or expected price plus 50 to ensure at least 50 profit if the expectation is right. The amount of the order for a certain agent to submit is based on their cash, the probable gain they make and also the risk attitude. The probable gain is equal to:

\[
g_t^i = \frac{|p_o - \hat{p}_{t,t+\tau}^i|}{p_o} \tag{8}
\]

The \(g_t^i\) is the probable gain for trader \(i\) at time \(t\) and \(p_o\) is the price submitted into the order driven market. The equation for amount the agent to submit is:

\[
s_t^i = \frac{s_{Max}^i g_t^i}{p_o} (\text{when } g_t^i \leq Max) \tag{9}
\]

\[
s_t^i = \frac{c_t^i g_t^i}{p_o} (\text{when } g_t^i > Max)
\]

The \(s\) and Max stands for the stock amount to submit in the market and the acceptable percentage of gain to use the full cash. Max is a uniform distribution from 0.5 to 1.

The agent’s order is submitted on a rolling basis in a random sequence. Trading follow the rules in order driven market derived from Chiarella et al (2009). During time period \(t\), when agents are entering the market, the current price is equal to the average of best bid and best ask. If the price of the buy order is greater than the best ask or the price of the sell order is smaller than the best bid, there is a match. The new order is the market order and trades at best ask or bid. Once there is no match, and the new order still has some untraded amount of order, the remaining amount of new order goes into the order book and the best bid or best ask is updated accordingly.
3 RESULTS

The current model is not producing a strong herding effect. Changing the herding parameter does not produce a noticeable change in the pattern of the results or the output values. Also gurus have a short lifespan (i.e., the guru keeps changing from one agent to another) and only a little more wealth than the other agents. Therefore the model needs further development. Altering the parameters and model details such as the fitness function in appropriate ways should give a stronger herding effect. For example, by increasing the probability of an agent switching to a successful agent as their neighbour.

Since the model is not yet producing suitable results this section will focus on listing the main model outputs. The results reported in Tedeschi et al. (2012) include for a single illustrative run: diagram of the final network; time series charts of which agent is the guru, percentage of links to the guru, guru fitness, wealth of guru and imitators and others, market price. For 100 runs the results include guru average life, average wealth of guru and imitators and others, decumulative distribution functions for wealth and stocks, average values for various market values such as the volume of orders. Tedeschi et al. (2012) ran the model using herding parameter $w$ values of 0.1, 0.5 and 1. We have added other outputs to measure the stylised facts of price distribution (fat tails) and price correlations (volatility clustering).

Figure 1 shows an illustrative time series chart of market price for one run of our model with $w = 0.1$.

4 CONCLUSION AND RECOMMENDATIONS

The paper has set out the research objective of improving the understanding of herding in financial markets. The approach is based on reproducing and comparing previous herding models. One difficulty, however, is in getting enough detail of the original study and we have encountered such problems. Nevertheless, reproducing is important for increasing confidence in scientific results and assessing their generalisability. This approach should be given more emphasis and priority in simulation research.
REFERENCES


Lee S and Lee K (2013). Heterogeneous expectations leading to bubbles and crashes in asset markets: Tipping point, herding behavior and group effect in agent-based model. At press.

**AUTHOR BIOGRAPHIES**

**XIN LIU** received a BSc (Hons) Accounting and Finance from Lancaster University in 2013 and an MRes Management Science from the same university in 2014. She is now doing her PhD at the same university.

**ROGER J. BROOKS** is a lecturer in the Management Science department at Lancaster University. He has a B.A. (Hons) degree in mathematics from Oxford University and an MSc (Eng) and PhD in Operational Research from Birmingham University. He is co-author with Stewart Robinson of a textbook on simulation. His research areas include conceptual modelling, Boolean networks, agent-based simulation, and sports statistics.