Event Study Approach for Validating Agent-based Trading Simulations

Shih-Fen Cheng
School of Information Systems
Singapore Management University
Republic of Singapore
sfcheng@smu.edu.sg

***Abstract*—In this paper, we introduce how one can validate an event-centric trading simulation platform that is built with multi-agent technology. The issue of validation is extremely important for agent-based simulations, but unfortunately, so far there is no one universal method that would work in all domains. The primary contribution of this paper is a novel combination of event-centric simulation design and event study approach for market dynamics generation and validation. In our event-centric design, the simulation is progressed by announcing news events that affect market prices. Upon receiving these events, event-aware software agents would adjust their views on the market and act accordingly. Their actions would be based on their roles and also their private information, and collectively the market dynamics will be shaped. The generated market dynamics can then be validated by a variant of the event study approach. We demonstrate how the methodology works with several numerical experiments and conclude by highlighting the practical significance of such simulation platform.**

***Keywords*-Trading Simulation; Agent-based Computational Economy; Simulation Validation**

I. Introduction

Simulations and serious games have been widely used in various domains to enable more effective knowledge transfer. Of all the disciplines, trading is probably the one that has benefited most from the rapid development of serious games and simulations. This is so because trading games are becoming more and more like real trading platforms thanks to the aggressive push for markets to become fully electronic in recent years. As rich web technologies become ubiquitous, a large number of sites are being set up to teach people how to trade via highly interactive web games. For example, UMOO (http://www.umoo.com/) is a popular fantasy stock trading site that allows people to practice day trading on selected real stocks. In commodity trading, FACTSim (http://www.factsim.org/) developed at University of Florida allows participants to trade in a simulated environment that tracks real commodity prices over several weeks or even months.

All the above mentioned systems track real market prices. Although tracking real markets has the appeal of realism, it is not always the most preferred approach in constructing training scenarios or simulations, for the following three reasons. Firstly, real markets move slowly most of time, and a meaningful scenario might require several weeks (or longer) of simulation time. Secondly, since the simulation tracks real markets, we have little control over the kind of scenario to be constructed. Finally, financial markets are known to demonstrate fat-tailed behavior and experiences for these cases are extremely valuable; however, since these events are rare, it is highly unlikely that we will be able to guarantee having such events in our simulation.

To address these concerns, during the past few years our research group has developed a simulation platform that allows a wide variety of trading simulation scenarios to be created (the platform is generic, however, we mainly use it in simulating commodity futures; for details, see [2], [3]). The major difference between our system and most existing trading simulation platforms is that we allow the creation of arbitrary market scenarios that would be more intensive and dramatic than the paces of real markets. For example, instead of having to wait for days or even weeks for market-changing events, a scenario could be easily designed to contain a year worth of important events within an hour of simulation time. Extreme market conditions that are rare but important could also be easily designed to develop participant's crisis management ability or test emergency market rescue measures [8]. As demonstrated by the case studies presented in [3] and [8], this platform enables us to design highly effective training programs for novice traders and effective simulations for policy analysis.

We achieved this flexibility with an event-centric design powered by the multi-agent framework. Event-centric design allows us to design arbitrary market scenarios by introducing sequences of events. The multi-agent framework allows us to build the simulated market from the bottom-up. By designing individual building blocks carefully and introducing appropriate market mechanisms, close-to-reality market dynamics can be generated.

One major challenge we face is the validation of the simulation outcome. Unlike other modeling techniques, real-world data necessary for the validation usually is not available for the scenarios designed within our platform. This is because introduced events could be completely fictitious, and the sequence in which they appear could also be arbitrary. The major contribution of this paper is the use of the *event study approach* as a way to validate event-centric agent-based market simulations.

II. BACKGROUND

For the interest of readers, we provide a brief overview on the system design. A more comprehensive description of the system design can be found in [3].

A. Software Architecture

Three important components are included in our commodity trading simulation: a) human trader terminals, b) market server (servicing market mechanism and dispatching events), and c) market agents (to be described later).

The software architecture is distributed in nature and all components are network-connected. Unlike most classical market simulations that usually run under strict theoretic assumptions on price dynamics, the price dynamics in our simulation is entirely determined by agent interactions in ways that are similar to real exchanges. All agents are allowed to buy and sell at any time and transactions are matched continuously; these requirements can be easily met by a standard *continuous double auction* (CDA). All human traders are also subject to a hard limit on their standing positions, and they are expected to clear all their positions before simulation ends. To keep the simulation compact and focused, most non-critical exchange features, such as daily settlements and margin calls are not explicitly modeled.

B. Event-Centric Design

As mentioned earlier, one critical design that enables artificial scenarios to be created is the introduction of events in the simulation. Events don't alter price stream directly, however, market agents that are part of the simulation will *read* the news events and react to them following a rational model. We do want to highlight that the agents we design do not actually read the news events in textual format; instead, we request that when an event is being introduced, the designer should quantify the impact of the event in numerical form. These impact values are only accessible to market agents and are invisible to human traders.

Formally speaking, an event is defined by the following four parameters:

- **Title and content**: This qualitative information is meant for human participants only.
- **Impact**: This integer number describes how strong an event is: 1 being the weakest and 5 being the strongest. A positive (negative) value corresponds to a bullish (bearish) event.
- **Arrival time**: This is the time when an event is sent to both human and market agents.
- **Effective time window**: This is the interval in which the event is effective. Note that consecutive time windows could overlap with each other to emulate a chain of events.

For all agents, an event only comes into existence when its designated arrival time has passed. Once an event arrives, only the "title and content" will be made available to the human traders. Both "impact" and "action time" are only available to the market agents. Of course, besides qualitative information, latest market information is also available as price quotes.

By putting events with different impact values in a sequence, the scenario designer can then realize the kind of market movement she wants to have in the simulation.

III. DESIGNING MARKET AGENTS

Using agents in modeling complex economic or financial systems is not new, in fact, a large number of literature has been devoted to the subject of "Agent-based Computational Economics" (ACE) (e.g., see [7], [12]). The ACE models is probably best explained in Tesfatsion's own words [12]: "*the defining characteristic of ACE models is their constructive grounding in the interactions of agents, ... Starting from an initially specified system state, the motion of the state through time is determined by endogenously generated agent interactions.*" Our model follows similar constructive principle, creating various agent roles that interact to generate market dynamics. Hedgers and speculators are two primary agent roles implemented in our trading simulations. In this section we provide a very concise overview on major market agent types that are included in our simulation.

A. Hedger Model

Hedgers are the original users of the futures market. They are usually producers or consumers of the commodity who would like to lock in at some specific prices and quantities well before the time of production (for producers) or usages (for consumers). These hedgers represent fundamentalists in the market.

To properly incorporate producers and consumers in our model, we assume that they exhibit stationary behaviors, i.e., the rate of their production and consumption will be stationary. Since we assume that only one futures market exists for this commodity, this assumption implies that all producers and consumers have to constantly establish new hedges in this market, and their collective actions will create the market dynamics accordingly. We further assume that all producers and consumers will employ a simple *hedge-and-forget* strategy, meaning that they will establish new hedges based on their own needs (new produces or usages), the current market condition, and their expectation; once the hedges are established, they will hold them to the end (in other words, no *dynamic hedge* will be considered).

B. Speculator

While "mean-reverting" hedgers constitute the "fundamental" part of the simulated market, most of the market volatility, on the other hand, is generated by the speculator agents. In our simulation, we adopt the classical *zero intelligence* (ZI) strategy [6] in constructing our speculator agent. To prevent ZI agents from destroying the market trend generated by producer and consumer agents, we limit the price range to fall in the bid-ask spread.

After the price is randomly decided, the ZI agent will decide to long or short equally likely. Since each ZI agent is granted the same trading limit as human traders, it will also randomly decide how much remaining position allowance it would devote to the new trade. Again, just like human traders, ZI agents are required to exit all positions at the end, and they are programmed to gradually exit their positions when the end draws near.

IV. Validating Simulation with Event Study Approach

A. Introduction

The modern event study approach, introduced by Fama et al. [4], is widely applied in economics and finance in measuring the effects of an event on the value of firms using financial data. As MacKinlay [9] puts it: "*the usefulness of such a study comes from the fact that, given rationality in the marketplace, the effects of an event will be reflected immediately in security prices*." This feature is important since the impact of an event could be measured in a relatively shorter time periods of several weeks or days, as opposed to several months or even years if we use other more lagging indicators (e.g., production levels, revenues). This description also explains our choice in picking the event study as the tool in validating the event-based trading simulation we built, since our simulation progresses by letting market agents react to events.

The event study approach has already been applied in detecting a wide-variety of events, e.g., mergers and acquisitions, earning announcements (both examples are discussed in [9]), or even macroeconomic news [11]. In most applications, common equity price of the studied firm is used; however, the event study approach could also be applied to other type of securities with little modifications [5].

Roughly speaking, event studies try to statistically test for *abnormal returns* from the security prices within a predetermined time window. Due to practical limitations resulting from data (could be related to both security prices and events), many event study variants have been suggested. Binder [1] reviewed a wide variety of event study methodologies, and discussed some frequently encountered empirical issues.

B. Event Study Procedure

These past researches on event study approaches provide a sound analytical framework for us to analyze whether artificial events generate consistent market dynamics in our simulation. Despite differences in the statistical techniques applied, most event study methodologies have the following general procedures. To stay focused, we only include steps that are relevant to our analysis (complete coverage on the methodology can be found in [9]):

1) First, the events of interest are identified, and for each identified event, a time window surrounding that event is defined so that security price information could be collected. For event j, let the beginning and the ending of the time window be τ_j^1 and τ_j^2 respectively.
2) Second, determine the firms to be included as data samples. The selection criteria usually involve data availability and the characteristics of the firm such as market capitalization and industry, so that an unbiased set of samples could be constructed.
   In our case, since the target we are studying is the price of the commodity derivative itself, the concept of firms does not apply here. Alternatively, we will construct the sample set by executing multiple simulation instances for the same event series (this is conceptually identical to the reactions of multiple securities to the same event series).
3) To understand the impact of the event, we need a measure on the event-induced abnormal return (AR), which is simply the actual return minus the *normal return* over the event window. The normal return is the expected return when no event is introduced. In our study, the normal return is computed by finding the mean price of the commodity before the occurrence of any event. To accommodate events with multiple periods, we define the cumulative abnormal return (CAR) for event j as:
   $$\mathrm{CAR}(\tau_j^1, \tau_j^2) = \sum_{\tau=\tau_j^1}^{\tau_j^2} \mathrm{AR}_\tau, \tag{1}$$
   where AR_τ represents the abnormal return in time τ.

With the computed CAR, various statistical tests could be administrated for different purposes. In our study, we are interested in testing the occurrence of an event and asserting that proper response strengths are generated. To detect the occurrence of event j, we define the *null hypothesis* to be no event occurrence and compute the test statistics as:

$$\theta = \frac{\overline{\mathrm{CAR}}(\tau_j^1, \tau_j^2)}{\sqrt{\mathrm{var}(\overline{\mathrm{CAR}}(\tau_j^1, \tau_j^2))}}, \tag{2}$$

where $\overline{\mathrm{CAR}}(\tau_j^1, \tau_j^2)$ represents the average CAR from all experiment instances[1], and $\mathrm{var}(\cdot)$ represents the variance of all results. θ computed in (2) should follow the standard normal distribution of $N(0,1)$. To test the occurrence of a bullish or bearish event, we should define a one-tailed positive or negative alternative hypothesis (i.e., $H_1 : \overline{\mathrm{CAR}}(\tau_j^1, \tau_j^2) > 0$ for bullish events, $H_1 : \overline{\mathrm{CAR}}(\tau_j^1, \tau_j^2) < 0$ for bearish events). In either case, the rejection of the null hypothesis could lead

[1] A scenario can be repeatedly simulated, generating a number of price streams. A separate $\mathrm{CAR}(\cdot)$ is computed for every price stream, and $\overline{\mathrm{CAR}}(\cdot)$ is then obtained by taking the average.

us to the conclusion that a bullish or a bearish event has occurred.

To assert that impact levels from 1 to 5 indeed generate appropriate price dynamics in the market, we would like to establish that the events with higher impact levels indeed produce larger CAR. To establish this result statistically, we compare mean CAR for consecutive levels in pairs using t-tests, i.e., comparing levels 1 and 2, 2 and 3, and so on. The null hypothesis will be no difference in mean CAR. The alternative hypothesis is similarly defined to be one-tailed, stating that the mean CAR from the stronger event is greater than that of the weaker event, i.e., $H_1 : \overline{\text{CAR}}(\tau_j^1, \tau_j^2) > \overline{\text{CAR}}(\tau_k^1, \tau_k^2)$, assuming that j is stronger than k.

C. Validating Occurrences of Events

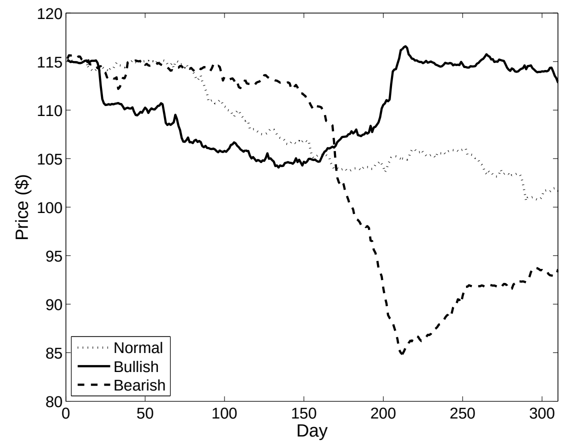
To simplify the simulation and avoid clustering effects from overlapping events, we create a special scenario with only one event. For the market agents, we include 12 producers, 13 consumers, and 2 ZI agents (to emulate human trader's actions). Both producers and consumers are constructed following the hedger model. The length of a simulation day is defined to be 1 second, and the length of the simulation is just over 370 days. The event occurs in day 160 (which is known to all agents), and the event window is defined to be 20 days before the event occurrence and 20 days after the event occurrence[2]. In other words, $\tau^1 = 140$ and $\tau^2 = 180$. For bullish, bearish, and no event scenarios, the impact levels are set to +5, -5, and 0 respectively. Sample price evolutions of these three scenarios are shown in Figure 1(a).

To collect enough sample data points, the same scenario is executed 15 times. Following the event study procedures described in Section IV-B, we test the null hypothesis for bullish, bearish, and no event cases. One sample CAR from each case is plotted in Figure 1(b). Note that in Figure 1(b), we use -20 and 20 to represent 20 days before and after the event occurrence respectively.

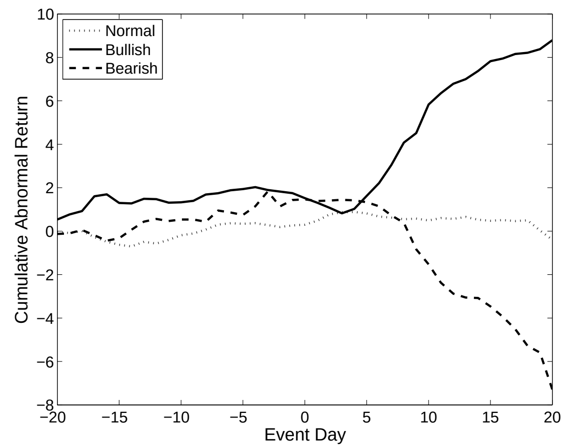
For both the bullish and the bearish cases, the p-values $\sim$ 0, implying that positive/negative abnormal returns are statistically significant. For the normal (no-event) case, p-value $\sim$ 0.065, indicating that no significant abnormal return is detected during the event window.

D. Validating Relative Strength of Events

The events in our simulation are labeled from 1 to 5, indicating their respective strengths. It would be impractical to try to validate the absolute response for each impact level, since many other factors (e.g., number of agents, percentage of human traders, size of price tick, just to name a few) also affect the absolute return. Alternatively, we would focus on comparing the relative magnitudes of events with difference strengths. Regardless of the absolute return levels, events with higher strength level should consistently generate greater market return magnitudes than events with lower strength level.



(a)



(b)

Figure 1: (a) Sample price dynamics of the one-event scenarios. The event is announced in day 160. (b) The cumulative abnormal returns (CAR) for bullish, bearish, and normal (no event) scenarios.

Table I: Summary statistics of CAR and paired comparisons for all impact levels.

Lv	Mean	Std. dev.	Sample size	p-value against previous level
1	6.042	3.067	20	–
2	10.562	2.897	20	0
3	14.206	4.661	60	0.0008
4	16.661	5.962	60	0.0067
5	21.138	7.895	40	0.0009

[2]As discussed earlier, the 20-day period before the event occurrence is used for estimating normal return. On the other hand, the 20-day period after the event is required since in the simulation and the real-world market alike, the impact of an event usually take a number of days to be fully realized.

To validate relative event strengths we construct a similar one-event scenario as specified in the previous subsection and vary its impact level from 1 to 5. The setup of the experiment is almost identical except that we now have five scenarios, each with different impact level.

For each impact level, 20 to 60 samples are generated depending on how variable the samples are. For impact level 3, 4, and 5, the results are particularly noisy, thus we have executed more simulations. The experiment results are summarized in Table I. From Table I, we can see that the impact level of an event indeed dictates the strength of response from market agents (all comparisons are statistically significant).

V. Conclusions

This paper attempts to address the validation issue in constructing agent-based trading simulations. Without proper validation mechanisms, it would be difficult to claim the the trustworthiness of the simulation, and hence void any prediction or analysis derived from the simulation. As recently argued by Marks [10], the lack of formal validation protocol is probably one of the major reasons why agent-based simulation is still not widely accepted by the mainstream economic, finance, and social science research.

The primary contribution of this paper is the introduction and the application of the event study approach as one candidate in performing micro-level validation for agent-based trading simulations. The event study methodology is ideal and natural for this application since the trading simulation is built with the event-centric principle, and the fundamental market movement is generated by event-aware agents. The novelty of our approach thus lies in the combination of the event-centric design in generating market dynamics and the use of event study approach in validating generated market dynamics.

We do not attempt to claim that our methodology would be universally applicable (for other types of agent-based simulations where there are no event, our methodology will not be applicable), however, for agent-based trading simulation that relies on events in scenario generation, the use of event study approach is indeed very effective.

With this validated simulation platform, a wide spectrum of applications are granted a sound foundation. For example, in [3], the platform allows human traders to participate, and the simulation platform is basically a training ground: exposing human traders to controlled and realistic market scenario so that they can practice their trading skills. A very different application is discussed in [8], where the same platform is used for policy analysis, allowing policy maker to evaluate the effectiveness of different regulation changes under different market composite. In this application, there is no human trader and the market only includes autonomous software agents with different strategies. For both cases, the validation procedure is extremely important since we need to assure the simulation users (human traders and policy makers) that the platform generates reliable dynamics.

VI. Acknowledgements

This work was supported in part by the International Trading Institute and Wharton-SMU Research Centre (under grant C220/MSS7C008) at Singapore Management University. The author would like to acknowledge the dedicated efforts by Chao-Chi Liu and Sharom Lim in running simulation experiments. The author would also like to thank Samuel Owens and Anil Shamdasani for lending their expertise in commodity trading.

References

[1] J. J. Binder. The event study methodology since 1969. *Review of Quantitative Finance and Accounting*, 11(2):111–137, 1998.

[2] S.-F. Cheng. Designing the market game for a commodity trading simulation. In *IEEE/WIC/ACM International Conference on Intelligent Agent Technology*, pages 445–449, San Jose, CA, USA, 2007.

[3] S.-F. Cheng and Y. P. Lim. An agent-based commodity trading simulation. In *Twenty-First Annual Conference on Innovative Applications of Artificial Intelligence*, pages 72–78, Pasadena, CA, USA, 2009.

[4] E. F. Fama, L. Fisher, M. C. Jensen, and R. Roll. The adjustment of stock prices to new information. *International Economic Review*, 10(1):1–21, 1969.

[5] T. R. Fortenbery and D. A. Sumner. Tthe effects of USDA reports in futures and options markets. *The Journal of Futures Market*, 13:157–173, 1993.

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

[7] B. LeBaron. A builder's guide to agent-based financial markets. *Quantitative Finance*, 1:254–261, 2001.

[8] B. Lee, S.-F. Cheng, and A. Koh. An analysis of extreme price shocks and illiquidity among systematic trend followers. *Review of Futures Market*, 18(4):385–419, 2010.

[9] A. C. MacKinlay. Event studies in economics and finance. *Journal of Economic Literature*, 35(1):13–39, March 1997.

[10] R. Marks. Validating simulation models: A general framework and four applied examples. *Computational Economics*, 30(3):265–290, Oct. 2007.

[11] G. McQueen and V. V. Roley. Stock prices, news, and business conditions. *Review of Financial Studies*, 6(3):683–707, 1993.

[12] L. Tesfatsion. Agent-based computational economics: A constructive approach to economic theory. In L. Tesfatsion and K. L. Judd, editors, *Handbook of Computational Economics, Volume 2: Agent-based Computational Economics*. Elsevier, 2006.