Babeş-Bolyai University of Cluj-Napoca
Faculty of Economics and Business Administration
Business Information Systems Department

Ph.D. Thesis Summary

Contributions to the Development of Intelligent Trading Applications

Ph.D. Supervisor:
Ghişoiu Nicolae, Ph.D. Professor

Ph.D. candidate:
Stan Alexandru-Ioan

November 27, 2013

Ph.D. Committee:

TOMAI Nicolae, Ph.D. Professor, *president*

COCIANU Cătălina-Lucia , Ph.D. Professor, *member*

NEGREA Bogdan, Ph.D. Professor, *member*

SILAGHI Gheorghe Cosmin, Ph.D. Professor, *member*

Contents of the Thesis

Contents of the Summary

Keywords. The 2010 Flash Crash, multi-agent artificial markets, flash crash simulation, artificial intelligence, computational finance

Nomenclature

ABE Agent-based Economics

ABM Agent-based Models

ACE Agent-based Computational Economics

ACF Agent-based Computational Finance

ATOM Artificial Open Market Framework

BAS Bid-Ask Spread

CFTC United States Commodity Futures Trading Commission

ETF Exchange-traded Funds or Index Trackers

MM Market Makers

PIN Probability of Informed Trading

SEC United States Securities and Exchange Commission

SIT Sophisticated Intelligence Traders

SVM Support Vector Machines or Support Vector Networks

TF Trend Followers

VPIN The Volume Synchronized Probability of Informed Trading

ZIT Zero Intelligence Traders

Chapter 1

Introduction of the Thesis

1.1 Context

There are very few research domains not to have been impacted or utterly remoulded by the computational revolution which emerged in the second half of the last century and has ever since been gaining momentum. Economics, in general, and finance, in particular, make no exception in this respect. The advent of high-performance computers has allowed modern-day economists to push the boundaries of what social scientists can do. The computer-based approach to economics gave rise to an alternative *computational paradigm* [113, 166] to the rigid neoclassical theory, which heavily relies upon complex mathematical models and pays little regard to whether these truthfully describe the real economy.

The new computational class of models challenge and redefine almost all the core assumptions which are sacrosanct to the neoclassic economic theory, and which make the associated mathematical models appealing and tractable. Thus, standard fundamental concepts such as, for instance, *perfect rationality* or *homogeneity of economic agents* are merely dismissed and replaced by flexible hypotheses of *bounded rationality*[139] or, respectively, *heterogeneity* in attributes and behaviour. The new types of bounded-rational computational agents no longer attempt to perfectly optimize the utility of their choices but rather to have a more realistic behaviour, which aims at making satisfactory choices.

The need for closer to reality financial models lead to the employment of computational Agent-based Models (ABM) [86, 171], which set up trading environments characterised by heterogeneity, bounded-rationality and market non-equilibrium dynamics, and where the artificial agents are abstractions of real-world market participants, humans and computer systems alike. The agents can be endowed with full artificial intelligence, learning and analyses capabilities, complex strategies and expectations, and, more recently, adaptive evolutionary aptitudes [81].

The encounter between economics and Agent-Based Models gave birth to a new branch of economics which has recently received a lot of consideration in economic and financial research under the *Agent-based Computational Economics*(ACE)[80] or ,more specifically for the financial context, *Agent-based Computational Finance*(ACF)[32] denominations . Agent-based Computational Economics studies from a computational perspective economic processes by modelling them as dynamic systems of interacting agents. Over the past two decades, ACF, in particular, has increasingly been seen as an effective approach to financial analysis by both researchers and practitioners as it does not abide by the constricted and simplifying assumptions of neoclassical theoretical models.

ACF models apply numerical methods to analyse data stemming from computer-based market simulations [96] of complex market dynamics where standard theoretic foundation is not easy applicable. Adopting a constructive philosophy, they can simulate intricate financial market environments characterized by unstable equilibria [68] where market crashes and peaks can be easily reproduced. Therefore, ACF models does not require simplifying assumptions as, through repeated computer simulations, the modeller can easily investigate the diversity of emerging patterns and accurately discern the dynamics of very specific financial markets configurations which are virtually impossible to approach in the classical way. For these reasons, we intend to use them in our thesis to reproduce and comprehend flash crash market events.

1.2 Motivation

On May 6th 2010, the S&P 500 e-mini contract suddenly lost about 9% of its value in a matter of seconds only to recover its losses within minutes, drawing attention on the potential repercussions of ever-more complex and sophisticated stock markets. After nearly five months of investigation, the Securities and Exchange Commission (SEC) and Commodity Futures Trading Commission (CFTC) filed a joint report September 30, 2010 entitled "Findings Regarding the Market Events of May 6, 2010"[1], tracing the sequence of events that led this *Flash Crash*. Many possible explanations were expounded by both the official inquiry and the academic research [46, 91] and high-frequency trading was hinted as one of the possible causes for this market event, but without clearly identifying a single responsible class of factors.

If the official inquiry did not hold responsible the high-frequency operators and their diverse and sophisticated new types of trading practices, it nevertheless did identify a large automated sell execution program in the S&P 500 market as the main accelerator of the flash crash. This large automated sell was soon after followed by a liquidity shock at both the composite index and individual stocks level which directly resulted in the flash crash.

While the governmental regulators vaguely correlated the flash crash occurrence to the growing dominance of HFT orders in the markets, the academic community[46] offered a single viable theoretical model able to satisfactorily ex-

[1]http://www.sec.gov/news/studies/2010/marketevents-report.pdf

plain the liquidity shock and based their argument on a cornerstone *information asymmetry* assumption which[46] coined *order flow toxicity*: the market makers as liquidity providers were forced out of the market by informed traders through large high frequency orders which eventually triggered the price crash.

Figure 1.1: The flash crash of 6 May 2010

Despite the considerable amount of attention received from various think-thanks, economists, media groups, researchers and U.S. governmental regulators, the specific cause of the flash crash remains in dispute. Emergence of high amplitude flash crashes is very bad news for investors. Worse though, is the general attitude which focused the corrective measures on mitigating the flash crash effects rather than elucidating its cause and preventing other reappearances.

Although crashes do appear in the financial markets, experts worryingly note that, in the current marketplaces dominated by automatic trading, crashes keep occurring dangerously often. For instance, a smaller amplitude mini flash crash took place just a few months ago on April 23, 2013. Nonetheless, their average amplitude is much smaller than the amplitude of the 2010 mega Flash Crash; the flash crash of 2013 provoked a price fell of "only" two percent.

The worries of the investors are fuelled by increasingly frequent technical incidents. Weeks ago, on August 26, 2013, trading on the Nasdaq stock exchange halted in the middle of the trading session for an unprecedented three-hour period following a computer related problem. The disruption pushed brokers to apprehend what went wrong, and rose additional concerns about the pitfalls of computer-driven stock trading. Fortunately enough, the Nasdaq freeze arose in an orderly fashion, did not cause panic or perturb other segments of the stock market, and, thereby, did not lead to a flash crash. Nasdaq officials announced that the problem was situated in its price-disseminating software system. A thorough investigation is under way.

These episodes do have concrete financial repercussions. Some analysts [76] estimate temporary market value losses going up to as much as one trillion dollars ($ 1.000.000.000.000) as a result of the temporary market price fall on May 6, 2010. Therefore, these hazardous market incidents are obviously extremely harmful and damaging for the confidence of the investors in the overall reliability of the market pricing mechanisms. They can induce a sense of concern amongst market participants that repeated flash crash occurrences could affect their investment strategies.

Facing the uncertainty of flash crash causes and the adverse economic consequences they can induce, we must redouble our efforts to utterly unravel the configurations which might provoke these highly poisonous market events. As these events still remain statistically rare, agent-based artificial markets can provide a very useful framework to reproduce, analyse and comprehend the flash crash characteristics and dynamics.

1.3 Research Objectives

Agent-based models provide very powerful simulation platforms which can accurately capture all essential aspects of market dynamics, which are generally difficult to observe and fully analyse in traditional analytical models. We use such models as basis for our research undertaking.This approach results in new financial models, by introducing agent heterogeneity, bounded rationality and other properties that would make analytical models hardly tractable. Within these agent-based frameworks, we examine one of the most recent research question in market finance: the emergence and unfolding of flash crashes. Our agent-based model analysis of flash crash events and forecasting metrics is complemented with traditional mathematical and statistical instruments.

First, we offer an overview of the flash crash market events and the main forecasting model proposed so far by the academic community *Volume Synchronized Probability of INformed Trading* (VPIN) (Chapter 2).

Then, we present how agent-based models have invested the financial modelling world through the advent of artificial agent-based financial markets, and we list some of technical and implementation issues related with their emergence. We also provide a succinct presentation of the the artificial market platform we use in our experiments (Chapter 3).

In the context offered by the literature survey and the presentation of the platform, we develop our own market models within the presented framework in order to investigate the the dynamics of *artificially generated flash crashes* (Chapter 4). While the main research stream of academic literature [46, 47, 48] is directed towards a better assessment of the likelihood of flash crash occurrences, to our knowledge, little or no attention is paid to fully understand the unfolding of a such market events. Therefore, our approach aims to elucidate the dynamics and the key features of a market flash crash, by relying upon the theoretical developments stated above, including the information asymmetry assumption, and a complementary hypothesis of predatory/prey behaviour of the informed traders towards the uniformed ones. This assumption states a *prey/predatory nature of financial markets*[40] where any type of advantage is immediately exploited by the market participants that possess it. These behavioural traits could set in motion the underlying forces behind the flash crash events as prey/predatory ecosystems frequently have unstable dynamics where small perturbations or disruptions prevent the convergence towards equilibria, whereby inducing cyclic movements, which can explain to a certain extent the large swing in assets prices during flash crash unfolding.

Given the fairly modest amount of empirical evidence which asserts VPIN as a trustworthy flawless informed trading proxy, we also attempt to evaluate its consistency as a flash crash predicting metric through a specific experimental set-up placed within an artificial agent-based market (Chapter 5). Concretely, we want to verify whether the VPIN metric detects significant levels of informed trading within settings where all market participants are uniformed. In this context, we also plan to observe VPINs behaviour and dynamics. Another objective we have is to offer, based on the observations, a better theoretical basis for some specific VPIN dynamics identified by critics [7].

Finally, as the existing literature lacks, besides the VPIN metric, high frequency estimators of the *probability of informed trading* [55, 133], we also aspire to develop some alternative procedures for quantifying informed trading. We aim at developing an alternative metric using some an artificial intelligence techniques. We would also like to find VPIN substitutes obtained through algorithmic based strategies. Naturally, we have to evaluate the relative effectiveness of these new crash-predicting metrics against the VPIN benchmark (Chapters 6 and 7).

1.4 Organization of the thesis

The work presented in this thesis pertains to a single research area. It addresses several aspects of the flash crash topic in financial markets by extensively using computer science simulation tools. This is primarily a thesis of *Computational Finance* [2] with all the involved complex interdisciplinary aspects ranging from financial theory and computer science to mathematical finance, numerical methods and econometrics [32].

Computational finance, as a research area, has expanded into virtually every branch of finance, especially in the financial markets field [33] over the last two decades. It represents a powerful tool-kit for analysing problems of practical interest in finance and, thereby, support or invalidate market models and hypothesis. This thesis aims to use the capabilities of this type of analyses in order to reproduce and further understand statistically rare market crash events.

This thesis is organized into two parts, and includes eight chapters in total which introduce and , then, investigate through computational tools several points of the flash crash topic.

1.4.1 Part I - Introduction, Context and State of the Art

The first part of the thesis introduces the context and the related literature used throughout this research. It presents financial markets flash crash related facts and theories which of use for our dissertation. This part also provides a succinct state of the art of agent-based artificial stock market systems.

Chapter 1. In this first chapter, we present the general introduction of the thesis. Here we convey the motivation and the research objectives of our work.

Chapter 2 presents the some important points of interest of the 2010 Flash crash unfolding. We refer to the conclusions of the official enquiry as well as to the academic research related to this topic. We conclude our undertaking by providing the reader with a state of the art of the VPIN model which claim to be the best predictor flash crash events.

Chapter 3 offers a short overview of agent-based artificial stock market frameworks, the motivation for using them in financial market simulation and research area. This chapter also points out the advantages of this approach as compared to the classical approaches stemming from the economic theory. We continue by presenting some important concepts for artificial market agent-based realm. We conclude the chapter by a short presentation of the simulation framework that we use throughout our thesis : ArTificial Open Market (or ATOM).

1.4.2 Part II - Results and Research Contributions

The second part of the thesis presents our results and contributions in the investigation of several flash crash related issues. This part contains four chapters that address three research aspects ranging from the limitations and advantages of the VPIN metric under specific market modes and configurations to attempts at providing theoretical explanation to some of the VPIN dynamics observed shortly after the flash crash unfolding. All the simulations are realized using artificial agents implemented in ATOM. The chapters contain information and research results closely or more loosely related to personal papers which has been published or accepted for publication [149, 150, 151, 152, 153]. A list of accepted papers and conference presentations can be found at the beginning of the thesis.

Chapter 4 elucidates the dynamics and the features of a market flash crash by describing the asset price processes in volume time units. Relying on the information asymmetry assumption and considering a stochastic prey-predator Lotka-Volterra model [167], we develop the deterministic and stochastic differential equations for the price development at flash crash time. The main novelty of our strategy resides in modelling the price stochastic process by a *volume-time* Ito process. The theoretical foundation of our approach relies primarily upon the hypothesis that the flash crash is the immediate repercussion of growing levels of order flow toxicity in asymmetric markets. The main ingredient of the model is the assumption that the set of interacting agents which generate at every moment informed transactions behave in a predatory manner towards the uniformed trading population at flash crash time [40].

In chapter 5 we gauge the reliability of the VPIN metric as an order flow-toxicity estimator within an artificial agent-based stock market populated with zero-intelligence traders and trend-followers. We verify whether the VPIN wrongly detects informed trading in a very specific experimental set-up where all market participants are uniformed. We evaluate the expected behaviour of the metric through Monte Carlo simulation. We also put forth a brief theoretical argument aiming to explain a particular feature of the VPIN dynamics observed during simulation phase. The argument relies on slightly modified assumptions for the VPIN model.

Chapter 6 outlines a *Support Vector Machine* based discrete volume-time approach to estimating the *Probability of Informed Trading*. We aim to offer an alternative metric of the extent to which liquidity providers are prone to systematically engage into losing transactions within a high-frequency trading environment characterized by systematic overbuying or overselling conditions. Our theoretical model supposes the existence of two homogeneous classes of markets participants: the informed traders who attempt to take advantage of the private information set they possess and the uninformed traders who provide liquidity on the market. We assume that the expected price depends on the excess demand generated by informed traders. Under this assumptions, we derive a time point formula for the PIN and offer a matrix based method for estimating it.

Chapter 7 tests whether simple volume-time algorithmic approaches combined with elements of optimization theory can produce effective estimators of the PIN metric in environments which replicate much of the real-world markets price series dynamics. The assessment of this algorithmic approach is done by using different types of settings and informationally asymmetric classes of trading agents.

Chapter 8 concludes our dissertation and summarises all the research contributions obtained throughout the thesis. Finally, it points to several possible future directions of research.

Chapter 2

2010 Flash-Crash and VPIN

2.1 The 2010 Flash Crash

The 2010 Flash Crash was a stock market crash that took place on Thursday May 6, 2010 at 2:45 pm eastern time, on the equity indices throughout the United States and which rapidly propagated to virtually all open equity markets around the world. For instance, the Canadian indices sharply fell approximately two minutes after[1]. The American stock indices lost, on average, around fife to six percent [120] of their value, only to recover most of their losses within half an hour. Notable hight declines appeared on the Dow Jones Industrial Average, which suddenly dropped more than nine percent (or around 1000 points), witnessing the biggest intra-day point decline in its history, and S&P future contracts, which lost about eight percent of their value. But even more spectacular were the swings that appeared in the individual composing securities like Accenture PLC stock which, for instance, just before the crash, was trading at a greater than 95 percent discount to the price and was quoted at around 39 dollars. Then, it had suddenly saw its price plunging at an incredible 1.84 dollars per share level, before recovering to close the trading session at 41.09 dollars. And this was not by any means a singular case [1].

Figure 2.1: An example of a recent small amplitude flash crash that took place on April 23, 2013 on Dow Jones Industrial Average

Although crashes are not uncommon in the financial markets, experts worryingly note that, in the current marketplaces dominated by automatic trading, High-Frequency Liquidity-Shortage Crashes keep repeating dangerously often. For instance Fig. 2.1 illustrates the mini Flash crash that took place on April 23, 2013. Nonetheless, their average amplitude is much smaller than the amplitude of the

Flash Crash. For instance, the flash crash of 2013 implied a price fell of only two percent on fake report of explosions at the White House disseminated into the market after hacking an influential twitter account and posting market-moving news. The 2010 flash crash is peculiar very peculiar in two respects:

- Its amplitude is antecedently high for so short a time period;
- As quickly as the market dropped, it suddenly and dramatically reversed itself reaching within minutes the preceding price level;
- It seems to be essentially liquidity-shortage driven.

In the May 6 case, "the market makers were buying stocks and it became very hard for them to turn the portfolio around to sell.They accumulated losses and at some point had to shut their portfolios down and vanish from the market. The crash occurred because there was no liquidity"[75].

Some financial analysts [76] estimate temporary market value losses going up to as much as one trillion dollars ($ 1.000.000.000.000) as a result of the temporary market price fall. Therefore, this type of hazardous behaviour is obviously extremely harmful and damaging for the confidence of the investors in the overall reliability of the market pricing mechanisms. It can induce a sense of concern among market participants that repeated occurrences of these incidents could affect their investment strategies. As a result of this concern, the U.S. Securities and Exchange Commission and the Commodity Futures Trading Commission established on May 7, 2010, in the wake of the crash, the Joint CFTC-SEC Advisory Committee on Emerging Regulatory Issues (also known as *The Flash crash Committee*) to develop recommendations on emerging and ongoing issues relating to both agencies. The founding chart of the joint committee states that it will serve an essential role in addressing the challenges of identifying how events in one market can adversely impact investors and markets elsewhere.

The first assignment of the committee was to conduct a review of the market events of May 6, 2010 and to make recommendations related to market structure issues that may have contributed to the volatility explosion, as well as harmonize disparate trading conventions and regulations across various markets.

The joint CFTC-SEC official inquiry was not very conclusive while correlating the flash crash occurrence to the growing dominance of HFT orders in the markets.It primarily identified a large automated sell execution program in the S&P 500 market as the main accelerator of the flash crash. This large automated

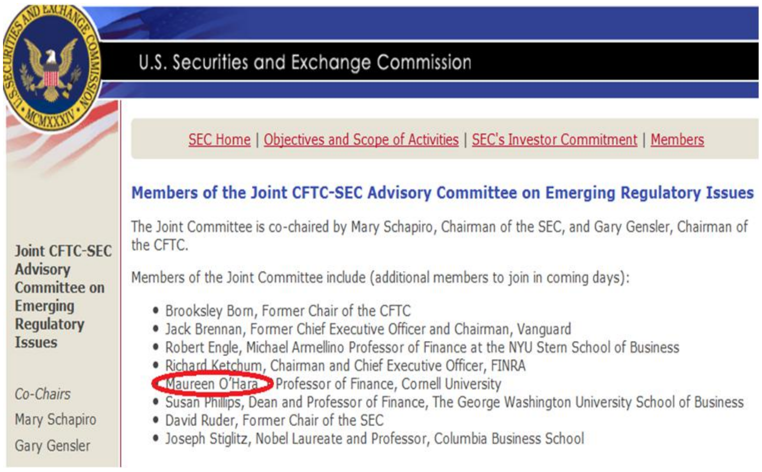


Figure 2.2: The Board of the Flash crash committee as of 10 May 2010 on the U.S. Securities and Exchange Commission's (SEC) website

sell was soon after followed by a liquidity shock at both the composite index and individual stocks level which directly resulted in the flash crash.

The Joint CFTC-SEC Advisory Committee indicated that although many factors contributed to events of May 6, and "different observers place different weights on the impact of each factor, the net effect of that day was a challenge to investors' confidence in the markets." [23].

A far more accurate analysis and better retracement can be found in the academic literature. The current literature on the subject indicates the next steps during the crash unfolding [120]:

- A large sell order of 75,000 E-mini contracts was gradually placed on the market [144];
- The order was placed in an already very toxic market environment which, as a result, became even more toxic [49];
- The high-frequency traders ended up flooding the market with sell orders causing the E-mini price to plunge [91];
- In an increasingly interrelated and interdependent market system, this induced price declines, first in the index trackers (ETFs, for Exchange-traded Funds) and soon after in the index composing securities [16];
- for lack of liquidity induced the collapse of the ETFs [22]. The loss was considerably higher in markets were the relative size of high frequency trading in the overall volume was higher [109].

For more details, [120] offer an in-depth anatomy of the Flash crash at the milliseconds level offering further insight into causal links and propagation velocity of the crash.

In the next section, we will succinctly present the Volume Synchronized Probability of INformed Trading model, the one viable response proposed, to this date, by the academic community to explain the occurrence of the 2010 Flash Crash.

2.2 The Volume Synchronized Probability of INformed Trading (VPIN)

So far, VPIN is the most significant theory proposed by the academic community to clarify the causes of the flash crash [46]. Using well-known capital market microstructure models, it presents the crash as a direct result of information asymmetry in the market. The most prominent of its authors is professor Maureen O'Hara from Cornell University, who is chairman of the board at Investment Technology Group, and is also a member of the Joint CFTC-SEC Advisory Committee on Emerging Regulatory Issues.

Easley et al. [46] highlight the effect of systematic informed trading on the liquidity of the market and propose to capture its pernicious effect with a metric which accurately assesses the probability of informed trading. Their work received notable consideration in the empirical finance literature and has been pursued by its promoters, notably [48, 51]. According to Easley at al. [46], the probability that informed traders adversely select uninformed traders can be seen as a proxy for what they named *order flow toxicity*. To this end, they developed the *Volume Synchronized Probability of Informed Trading* (VPIN) metric, which is a real-time estimate of the manner in which the liquidity is provided on the market under informationally asymmetric conditions.

Order flow becomes toxic when it is initiated by counterparts possessing reliable private information on the direction the prices are heading. This creates trade imbalances that eventually produce considerable losses for market makers unless they reduce the size of their positions. As order flow toxicity increases, the less informed market participants have to withdraw in order to avoid further losses, consequently drawing even more liquidity out of the market and increasing the overall level of flow toxicity in the traded volume. This self-sustaining mechanism eventually results in driving all the market making activity out of the market, and causes liquidity-induced crashes; the May 2010 flash crash being one major example according to them. Easley et al. [46] indicate that one hour before the flash crash VPIN reached its highest value in recent history.

Easley et al. [46] put forward a possible market-based solution to prevent the market makers from being forced out of the market. The solution they advocate would suppose setting up a futures contract with pay-offs based on VPIN. Thus, as soon as the VPIN goes up, they can hedge their risks and don't have to withdraw from the market.

2.2.1 Succinct Description of the Theoretical Model

The VPIN model stems from the seminal work of Easley and al. [55], published in the Journal of Finance in 1996. Using a Bayesian strategy to infer the fraction of informed orders, the study puts forth a market micro-structure theory which explains the degree to which market makers are willingly liquidity providers and defines a metric known as Probability of Informed Trading (PIN) to estimate the magnitude of informed orders.

Concretely, the PIN model assumes that arrival of private information makes the security's price (here denoted by S) evolve to 2 possible values S_B or S_G according to the nature of the news (bad or good). The model also supposes that information arrives in the market at a homogeneous intensity rate α. The news are either bad, with the probability δ, or good, with the probability $1-\delta$. Easley and al.[55] prove that the expected value of the security's price can then be computed at time t as

$$\mathrm{E}[S_t] = (1-\alpha_t)S_0 + \alpha_t\left[\delta_t S_B + (1-\delta_t)S_G\right] .$$

Informed and uninformed orders arrive at constant Poisson rates of μ and ϵ respectively. Then, in order to avoid losses from informed traders, market makers reach break-even at a bid level

$$\mathrm{E}[B_t] = \mathrm{E}[S_t] - \frac{\mu\alpha_t\delta_t}{\epsilon + \mu\alpha_t\delta_t}(\mathrm{E}[S_t] - S_B) .$$

while the break-even ask level at time t is

$$\mathrm{E}[A_t] = \mathrm{E}[S_t] + \frac{\mu\alpha_t(1-\delta_t)}{\epsilon + \mu\alpha_t(1-\delta_t)}(S_G - \mathrm{E}[S_t]) .$$

Thus, the break-even bid-ask spread has the following expression

$$\mathrm{E}[A_t - B_t] = \frac{\mu\alpha_t(1-\delta_t)}{\epsilon + \mu\alpha_t(1-\delta_t)}(S_G - \mathrm{E}[S_t]) + \frac{\mu\alpha_t\delta_t}{\epsilon + \mu\alpha_t\delta_t}(\mathrm{E}[S_t] - S_B) .$$

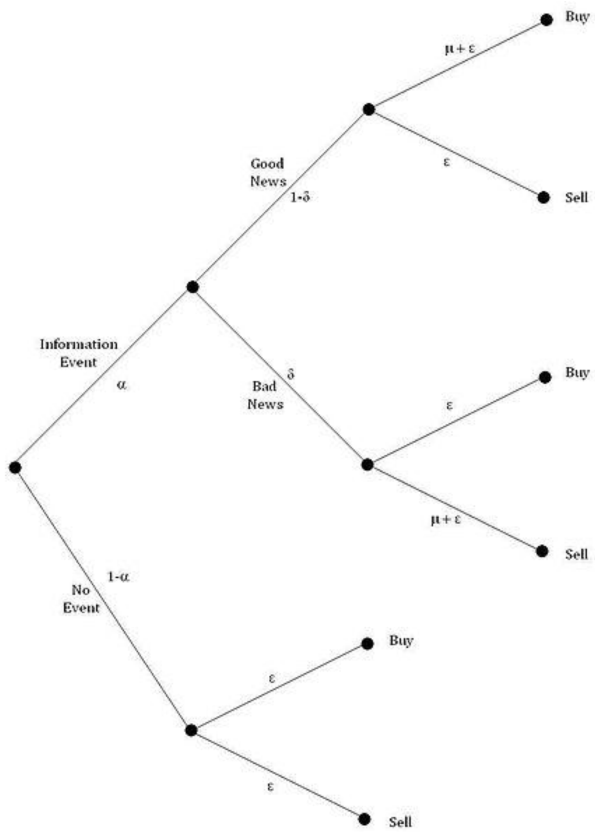


Figure 2.3: The PIN trading diagram [55]

Therefore,

$$\delta_t = \frac{1}{2} \Rightarrow \mathrm{E}[A_t - B_t] = \frac{\alpha_t \mu}{\alpha_t \mu + 2\epsilon} \left(S_G - S_B \right),$$

which tells us that the determinant of the range at which market makers provide liquidity is

$$PIN_t = \frac{\alpha_t \mu}{\alpha_t \mu + 2\epsilon}.$$

The probabilities α and δ are updated at a time point t through a Bayesian estimating strategy [27] in order for them to incorporate information after each trade arrives to the market.

In the aftermath of the 2010 flash crash, Easley and al. [46] proposed VPIN as a high-frequency estimate of PIN. This metric is a volume clock which realizes a volume sampling of the market activity at regular volume buckets volume-time intervals.

$$VPIN = \frac{\sum_{\tau=1}^{n} |V_\tau^S - V_\tau^B|}{nV},$$

where τ is a bucket of traded volume, V_τ^S and V_τ^B are the sell and buy volumes (traded against the Bid and Ask respectively), V is the total volume per bucket.

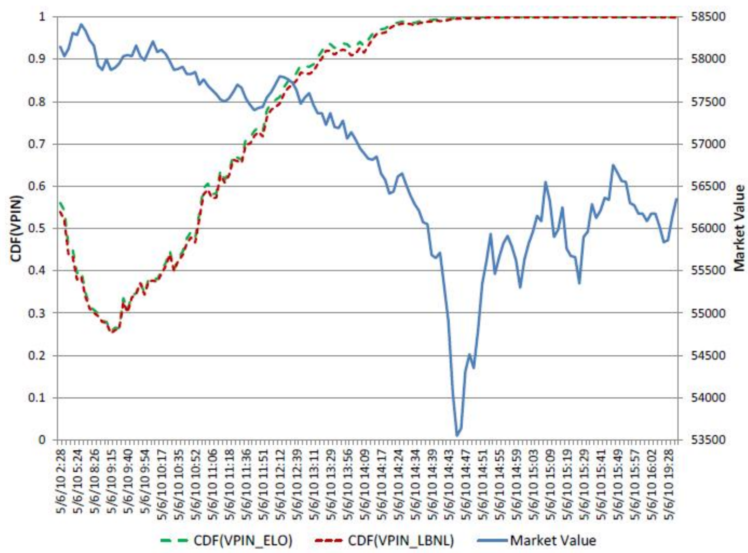


Figure 2.4: VPIN evolution just before and during Flash crash computed on E-mini S&P 250 [17]

2.2.2 VPIN as an Order Flow Toxicity Estimator

The VPIN model is used to determine in real-time the extent to which the Market Makers are subject to systematic hostile market conditions induced by informed trading. The systematic adverse selection of market makers by informed traders could eventually have deleterious effects in the overall market liquidity provision and induce liquidity crises similar to the 2010 Flash crash. If the order flow becomes highly toxic, market makers gradually need to leave the market and, by doing so, increase even further the relative toxicity of the order flows, which eventually triggers a cascading effect and results into a liquidity driven crash. Under these circumstances, the capital advantage of the VPIN metric is that it monitors and predicts in real-time the dynamics of these poisonous market phenomena.

The high-frequency VPIN model gained in popularity as a crash predictor on May 6, 2010, when it reached, with as much as one hour before the crash, unprecedented high levels in recent years prefiguring the emergence of a liquidity driven market crash. This prompted, in the succeeding period, the academic community and regulatory public authorities to further study its overall dynamics and features.

For instance, the governmental Lawrence Berkeley National Laboratory, in an independent study [17] of the May 6, 2010 events realised on behalf of the U.S. Securities and Exchange Commission (SEC), stressed out the viability of the VPIN metric as a liquidity crash-predictor in hight-frequency trading environments concluding that it is, by far, the most accurate early warning signal known to them at the time. [3] realized a second independent study only to reassert VPIN's capacity to predict toxicity-induced volatility on cash stocks. This study underlines the fact that the VPIN field of action does not restrain to the high-frequency trading environments, although it was conceived to fit their specifications. They indicate that empirical results indicate its ability to be a proxy for adverse selection risk for longer periods of time. The main parameter of the VPIN calculation is the granularity of the volume bucket. For specific granularities VPIN can be helpful beyond the well confined high-frequency trading realm.

The interest in understanding the relationship existing between VPIN dynamics and disrupted liquidity provisions becomes of paramount relevance as these studies show that liquidity crises can become more and more frequent in a high-frequency world dominated by algorithmic market-making strategies.

Another by-product brought about by the VPIN theory is the effect the Order Flow Toxicity has on volatility levels as VPIN explains the spread widening between the bid and ask quotations as result of increases toxicity. A discussion on the causal relationship between VPIN and volatility levels can be found in [49]. A good practical illustration of applying VPIN concepts and theory is provided in [40].

2.2.3 VPIN as a potential instrument for Financial Markets Supervision

[17] showed that VPIN would be a useful metric to monitor in real-time the probability of a liquidity crisis. Academics as well as a some U.S. officials considered a possible solution to averting liquidity induced market crashes by using a regulatory systems dedicated to computing time estimates of crash probabilities. If values of the metric reach dangerously high levels, stock exchanges could decrease volume exchange rates, giving the market time to clear and avoid a crash generated by sudden flooding with buy or sell orders. Some SEC officials deem that a viable VPIN based system able to monitor financial markets cannot be put into place as long as a certain number of theoretically related aspects are not fully understood and clarified.

[120] examined the VPIN evolution during the flash Crash and found a negative correlation between the relative proportion of large selling programmes in the market and the order flow toxicity. The study confirms the usefulness of the VPIN as an Order flow toxicity estimator [3]. [120] indicate the reason of this

negative correlation resides in the fact that informed traders sell passively during upturns and aggressively just after while they do not trade in downturns.

As pointed out in the previous sections, VPIN can be used as an order flow toxicity estimator beyond the high frequency world and can hint adverse selection the same the PIN metric does. The estimation accuracy parameter would depend the number of volume buckets used. This findings opens the perspective for VPIN to be a regulatory instrument in longer horizon contexts [3].

More interesting, [169] show that different types of securities have different VPIN characteristics. Furthermore, VPIN as an order flow toxicity estimator can explain and predict future several intra-day trading factors such as :

- Price volatility;
- Quote imbalance;
- Volume bucket duration;
- Trade intensity.

They also point out that causality flows the other way around: "similarly, it is not surprising that these intra-day factors are also able to explain VPIN to a certain extent and the reverse causality is also true" [169]. This findings reassert VPIN as a powerful regulatory instrument.

2.3 Conclusions

In this chapter we briefly discuss the flash crash of 2010 and the model the academic community put forward in order to explain its emergence Volume Synchronized Probability of INformed Trading. We start by exposing several important figures which characterise the market event, and present the reaction of the market regulatory authorities. We also list the chronology of the pivotal points of flash crash unfolding and give some insight into the causal links amongst them and the propagation velocity. In the second section, after a short summary of the VPIN model, we enumerate the most important aspects which recommend the metric as an interesting flash crash predictor.

Chapter 3

Agent-based Artificial Financial Markets

3.1 Agent-based models

Agent-based models (ABM)[86, 171] are computational models which simulate the actions and interactions of autonomous agents in order to evaluate a set of characteristics of the system they are operating into. Recently, their usage has became increasingly widespread and popular in social sciences including economics [124].

ABM reproduce the actions of numerous agents and the interdependency that exists among them, in order to emulate and forecast the evolutions of complex systems. By agent we understand an entity part of the computationally constructed world endowed with specific characteristics, attributes and behavioural methods [83]. Starting from initial conditions specified by the modeller, the model evolves as a bottom-up construction where lower level constituents, sub-systems or individual agents, aggregate their actions in order to give rise to the top level system dynamics. The agents may be highly dynamic as they adapt to their environment, learn, die or reproduce [21]. [21] indicate the general composition of an ABM:

- A mix-up of various types of agents;
- The specific behaviour rules for each type of agent;
- A set of learning or adaptive rules;
- An interaction topology;
- The operating environment.

In the ABM the agents are purposeful. They act and interact in order to simulate the functioning of a complex system. ABM are fit to describe rich environments that include various features with greater fidelity than do mathematical and/or statistical techniques. They represent a powerful analysis framework, in-between unambiguous mathematics and potentially inconsequential descriptions.

One of the pivotal aspects of such architectures is K.I.S.S ("Keep it simple and short"). This frequently encountered tenet, states in this context that complex top level behaviour and patterns emerge as a combination of rather simple and nave individual behaviours. Another governing characteristic of these models is that the system as a whole is greater than the sum of its parts, since it encompasses the whole complexity that stems from the interactions taking place among the subcomponents.

3.2 Agent-Based Models in Economy and Finance

Two of the main applications of Agent-based models are the economic and financial research fields [80]. The encounter between economics and Agent-Based Models gave birth to a new branch of economics which has received a lot of consideration in economic and financial research recently under the *Agent-based Computational Economics* denomination. Agent-based Computational Economics (ACE) studies from a computational perspective economic processes by modelling them as dynamic systems of interacting agents.

Over the past two decades the ACEs has increasingly been seen as an effective approach to economic and financial analysis by both researchers and practitioners [71, 114, 150, 159] as they do not abide by the constricted and simplifying assumptions found in the neoclassical theoretical models characterised by *homogeneous economic agents*. These are replaced by *agents with heterogeneous, dynamic, and interdependent behaviour*. Through their significant flexibility and ability to integrate a large variety of assumptions on the behaviour, configuration and interaction topology of the market participants, they complete and, to a certain extent, replace the existing mathematical and econometric toolboxes. They certainly further the understanding of highly complex and seemingly unpredictable conditions and configurations present in the economies, in general, and financial markets, in particular. For these very reasons, we employ them in our thesis so as to gain insight and validate the mathematical modelling of the flash crash events.

ACE models apply numerical methods to analyse data stemming from computer-based simulations of complex dynamics where standard theoretic foundation is not easy applicable [89]. Adopting a constructive philosophy, they can simulate, for instance,intricate market environments characterized by unstable equilibria [68] where market crashes and peaks are very frequent. ACE models, unlike the classic models, does not require top-down assumptions as, through repeated computer simulations, the modeller can easily investigate the diversity of emerging patterns. The top-down assumptions in the classical models are generally made for mathematical *tractability* tractability reasons, but these restrictions often result in an oversimplified less interesting abstractions of the real-world phenomena.

Although the ACE models can always be formalized as systems of mathematical equations [61], these are sometimes intractable. The ACE modelling is therefore deemed superior to general equilibrium models or stochastic dynamical systems. [65, 122] indicate that ABMs represent more accurately the markets complex environment than classical models which do not incorporate complex information sets about the financial sector. The *Nature* journal editors [122] also indicate that "economic modellers should consider adopting the modular architecture used in many climate models. This approach makes it easy to aggregate smaller models into more comprehensive simulations, while still allowing steady improvement in each piece". Thus, modellers should be able to accurately simulate the dynamics of very complex computational economies which are very difficult to approach in the classical way.

In this perspective [122, 159], ACE models are primarily conceived to emulate computational economies as *complex dynamic systems* [11, 136] with a plethora of micro-agents repeatedly engaged into local interactions, and with other complicated aspects encompassing micro-behaviours, interaction patterns, and global regularities. These global regularities after being created feed back into the determination of local interactions. This back-and-forth diffusion mechanism result in an intricate system of interdependent feedback loops connecting micro-behaviours, interaction patterns, and global regularities.

Therefore, according to [157], the ACE main subject of study is to decipher the phenomena that underlie the natural emergence of regularities in economic processes unfolding. Such an example is the unplanned coordination of trading activities in decentralized market economies that economists associate with Adam Smith's invisible hand [157]. ACE aim at explaining how these global regularities arise from the local interactions of autonomous agents rather than through unrealistic oversimplified mathematical mechanisms.

There have been observed many similarities between ACE and game theory when applied to modelling agent interactions [77, 138, 145], but also significant differences as, within ACE, many events are solely determined by initial conditions, whether or not equilibriums exist or are tractable. Another difference stems from the ability of the agents to enhance their autonomy and continuous learning and adaptation [159].

Economic processes are perceived by ACE as agent-based complex adaptive systems [158] where the actors are human-like computational objects that attempt to make the most credible or profitable choices at action time. The rules that map situations to actions try to reproduce human behaviour and social interactions based on reinforcement learning [154]. The agents must discover which actions maximize a reward function by trying them. The rule can also be discovered through AI techniques.

ACE replaced the classical economic *market equilibrium* by the notion of *bounded rational agents* adapting to market conditions [141, 142]. [11] affirms that ACE are able to model a wider variety of situations than the neoclassical equilibrium models (out-of-equilibrium and convergence to equilibrium configurations). In the classical economic models regularities stem from the *perfectly rational* (and unrealistic) behaviour of all economic agents, which simultaneously seek to maximize their utility functions in the context of *scarce resources allocation to competing ends* [87, 119, 168].

Bounded rationality implies that agents are approximatively rational by making the "best" choice within a set of simple choices available. Their decision-making is impaired by lack of cognitive ability and resources [163, 164](frequently, information and/or time). The access to additional resources of any sort, when available, generally induces supplementary costs. Therefore, the agents chose a satisficer's (and not a optimizer's or a maximizer's) behaviour: they only look for satisfactory solutions as opposed to optimal ones [90]. *Bounded rationality* is, therefore, a concept meant to provide an acceptable solution to the critic that finite computational resources prevent agents from making perfectly rational choices and behaving fully rationally. The theory of bounded rationality was pioneered by [146]. Generally, in these models, the market agents are not heterogeneous with large diversity of behaviours and attributes, but rather, instances of classes where all agents are, at best, slightly modified replications of a class representative.

In financial research, rationality, full or bounded, is an important topic as it has non-trivial implications [9, 64, 81, 82, 102] on the overall market efficiency [62, 63, 110] (the inability to consistently predict the evolution of asset prices at different time horizons starting from current and past information).

[8] separate the research contributions of ACE models into two partially overlapping classes:

- A class composed of models where results proceed from a analytical investigation that results in tractable closed-form mathematical formulas [35, 36];
- A class containing models where research contributions stem from analysis of computer simulations, such as in [25, 66, 108]. This intuitive simulation strategy leaded [157] to describe the process as a bottom-up approach to the study of economic systems. The data is the output of repeated simulations performed with different parameters values corresponding to different contexts. The contributions we bring in this thesis pertain primarily to this second class.

For the ACE models from the second class simulation is very important as it provides the dataset to which the modellers will resort during their analysis. The simulation output allows researchers to track the evolution of the computational economy, the dynamics, adaptation and successive states of agents as well as the interactions occurring amongst them.

3.3 Agent-based artificial financial markets

The Agent-based artificial markets are computational platforms which try to constructively simulate the functioning of real world financial markets through agent-based trading panels and various auction mechanisms. The cornerstone guide, to this date, on the building and designing such complex systems is provided by [96]. Nevertheless, additional aspects can be found throughout the literature [140]. These markets use various specific models of price creation as bid and ask order streams proceed from the market participants. The market actors are artificial trading agents which are most often characterized as *bounded rational* [139, 161] acting to optimize some profit and risk function, using simple decision-making rules or more sophisticated full artificial intelligence methods.

The main convergence point between real-world markets and artificial ones is provided by the *stylized facts* in price and volume time series [38]. Stylized facts are simplified characterisations of empirical findings and essentially refer to market data statistical properties which hold true most of the time. An exhaustive list of the stylized facts known to this date is provided by [32], and we present it in table 3.1. They were first observed in the real-world financial markets and can be correctly reproduced, through sometime fine tuning *calibration*[24, 96, 170, 173] and *validation*[5, 114, 170] methods, within the artificial multi-agent market frameworks. Calibration refers to choosing the parameters which make a model to best fit the empirical data. Validation refers to verifying the ability of a model to fit or not the data. This ability of the artificial markets to mimic the statistical properties of real-world data to the point of not being able to tell them apart, is essential to the market research undertaking.

Interestingly enough, the literature indicates that many stylized facts and market dynamics do not derive from sophisticated agents behaviours. For instance, interesting conclusions are drown from the early work of [73]. They have demonstrated that even highly uninformed Zero-Intelligence Traders (ZIT) can perform well during double-auction experiments where they observed that high market efficiency was generally obtained as long as the traders acted within their budget constraints. [73] conclude that the market efficiency they observed, derived from the structural aspects of the auction and not from the learning capabilities of the agents. The basis of these early findings was later confirmed by developments of [34, 74] which came to the conclusion that good market performance should not automatically be attributed to trader learning and rationality. Nevertheless, they also drew some cautions about the generality of these early findings.

A interesting development by [32] presents a econometric viewpoint on the development of ACE in an attempt to build the econometric foundation of ACE; they attempt to explain the emergence mechanisms for some of the stylized facts. They reflect on providing an agent-based foundation for financial econometrics. They focus their analysis more specifically on agent-based modelling of financial markets.

In the artificial markets platforms, an important role is played by the simulation capabilities. For the artificial financial market case, multi-agent simulations frequently address investment and asset price dynamics problems. This approach seeks to provide more conclusive analysis tools for investment decision making through more realistic and flexible assumptions. The simulations can be repeated multiple times using the Monte Carlo method [129] in order to derive the main statistical properties of the significant variables time series.

Over the last twenty years, the simulation process has continuously enhanced its performance and effectiveness by significant advances in the artificial intelligence techniques and increased computer capabilities. The final scientific goal of the simulation process is to "test theoretical findings against real-world market data in ways that permit empirically supported theories to cumulate over time."[159] Financial market simulations focus on a multitude of contexts and specificities of real markets, and try to grasp their intricate dynamics and causes [24, 89, 172].

The subject has been first applied to research in asset pricing by [73] and continues to be deemed as a promising path forward in the domain. The fundamental question that goes along with simulation is the type of market microstructure, agents' behaviour and intelligence needed to reproduce realistic stylized facts. Extensive valuable discussion on the subject can be found, for instance, in [6, 95, 103, 107, 115].

The multi-agent ecosystems of financial markets can be very rich [26]. [32] identifies an n-type formulation of heterogeneity for ACF systems containing n different types of trading agents (see Fig. 3.1. The majority of the most frequently encountered models in the literature are weakly heterogeneous.

In our dissertation, we will also resort to rather weakly heterogeneous trading environments, containing only three types of agents (3-type models). The market

Table 3.1: Stylised facts encountered in financial markets[32]. "The stylized facts are separated in 6 blocks in the table. The first two refer to the stylized facts pertaining to return and trading volume , using low-frequency data. The next four refer to the stylized facts of return, trading duration, transaction size ,and bid-ask spread , using high-frequency data."[32]

No.	Code	Stylized Facts	Reference
1	AA	Absence of Autocorrelations	[38]
2	AG	Aggregational Gaussianity	[38]
3	BC	Bubbles and Crashes	[137]
4	CE	Calendar Effect	[156]
5	CHT	Conditional Heavy Tails	[38]
6	EPP	Equity Premium Puzzle	[93]
7	EV	Excess Volatility	[39]
8	FT	Fat Tails	[38]
9	GLA	Gain/Loss Asymmetry	[38]
10	LE	Leverage Effect	[38]
11	LM	Long Memory	[38]
12	PLBR	Power Law Behavior of Return	[69]
13	PLBV	Power Law Behavior of Volatility	[67]
14	VC	Volatility Clustering	[38]
15	VVC	Volatility Volume Correlations	[39]
16	PLBTV	Power Law Behavior of Trading Volume	[69]
17	VLM	Long Memory of Volume	[60]
18	AA-H	Absence of Autocorrelations	[156]
19	FT-H	Fat Tails of Return Distribution	[156]
20	LM-H	Long Memory	[156]
21	PE	Periodic Effect	[156]
22	BU	Bursts	[156]
23	CTD	Clustering of Trade Duration	[130]
24	DLM	Long Memory	[130]
25	DO	Overdispersed	[130]
26	PLBT	Power Law Behavior of Trades	[69]
27	US	U Shape	[162]
28	SCPC	Spread Correlated with Price Change	[162]
29	TLS	Thinness and Large Spread	[121]
30	TD	Turn-of-the-year Declining	[121]

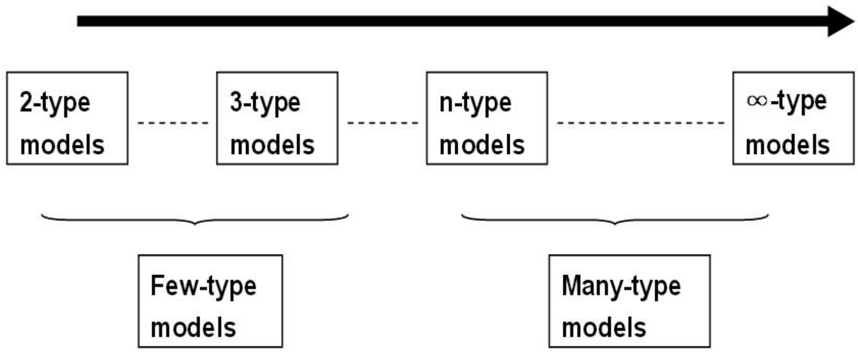


Figure 3.1: A spectrum of ACF models: Heterogeneity [32]

micro-structure we use is quote driven. As our case studies are essentially meant to evaluate the characteristics of the VPIN metric as a volume flow-toxicity estimator or to compare it to other metrics, we populate the marketplace with the following types of participants with different behaviours and functions [40]:

- Zero-intelligence[84] traders which contribute through their buying and selling activities to the price creation; they are along with the market maker liquidity providers. In the VPIN model [46, 55] they would correspond to the uniformed traders which generate streams of buy and sell orders at some Poisson rate;

- One market maker[31, 125] which stands ready to buy/sell asset at its posted bid and ask prices at any time;

- Some other homogeneous class of artificial traders which are all carrying out a common strategy as soon as they detect the emergence of a trading signal and which make the VPIN metric evolve through the relative magnitude of their market activities and the correlation of their behaviour. Their correlated behaviour affect the VPIN dynamics. This third class of agents can also have predatory behaviour towards the agents of the other two classes generally based upon an additional private information set. The study of prey/predatory ecosystems has received a huge amount of research. A rather complete overview of the main aspects can be found in [15]. The main point we want to stress here is the *prey/predatory nature of financial markets*[40] where any type of advantage is immediately exploited by the market participants that possess it. More important, these behavioural aspects are the main cause behind the flash crash events as these systems have unstable dynamics where small perturbations or disruptions can prevent the convergence towards equilibria inducing cyclic movements, which can explain the large swing in assets prices during flash crash unfolding, along with their amplitude and timing.

In the next section we will offer a short presentation of the Agent-based artificial financial market that we make use of throughout our dissertation, and within which we define and run our models.

Figure 3.2: An interface window in ATOM illustrating the simulation of a simple market configuration (Minimum market Model)

3.3.1 ATOM

There are significant and interesting artificial market software implementations in the literature:

- Santa Fe Artificial Stock Market. It is cited in a number of significant publications as [12, 58, 59, 98, 131]. [88, 98] present technical and architectural features of the platform;
- Genoa Artificial Stock Market [37, 111, 112];
- NASDAQ Market Simulator [42];
- Agent-Based Simulation of Trading Roles in an Asynchronous Continuous Trading Environment,described or cited, for instance, in [19, 20];
- Frankfurt Artificial Stock Market [79];
- Artificial Open Market [116];

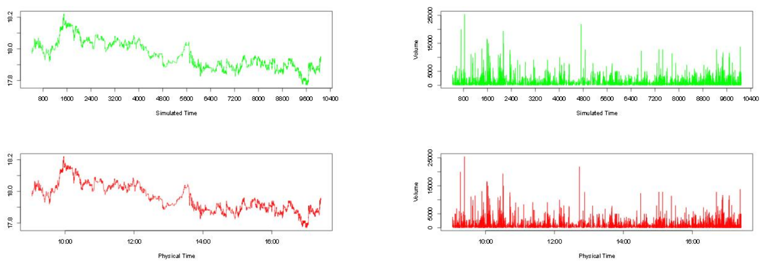
As giving a detailed description of all the platforms enumerated above would be impractical and request a large amount of space, in this section we limit ourselves to provide a short overview of the Artificial Open Market framework as we make extensive usage of it as our simulation platform in this thesis.

3.3.1.1 General Overview

Throughout the thesis we make use of a specific artificial stock market, called ATOM. We populate this framework with artificial trading agents who, through their trading strategies and interactions, allow us to reproduce market flash crashes and examine their dynamics as well as the validity and the accuracy of flash crash forecasting VPIN metric proposed by the academic community in the aftermath of the 2010 Flash Crash.

ATOM is a simulator of an *order driven* financial market in which buy and sell orders are matched against each others through double auction books. In our thesis we used a derived experimental version of the framework, which is *quote driven* and makes use of a single market maker which continually posts BID and ASK prices.

ATOM has been developed by researchers at Lille 1 University. Although, ATOM can accurately simulate the behaviour of most order-driven markets, it can be considerer more specifically as a NYSE-Euronext stock exchange emulator as illustrated in Fig. 3.3.



(a) Prices : ATOM vs. Euronext-NYSE

(b) Volumes : ATOM vs. Euronext-NYSE

Figure 3.3: ATOM can generate financial dynamics in line with the ones of the Euronext-NYSE stock-exchange [116]

Since ATOM is delivered as a java-based development library, it can be executed on all operating systems (Windows, Linux, Unix etc.) supporting java virtual machines.

ATOM can efficiently be used in both intra and extra-days trading sessions where it can generate, play or replay simple and complex experimentations. The first class allows us to create time series of prices or orders, the second allows

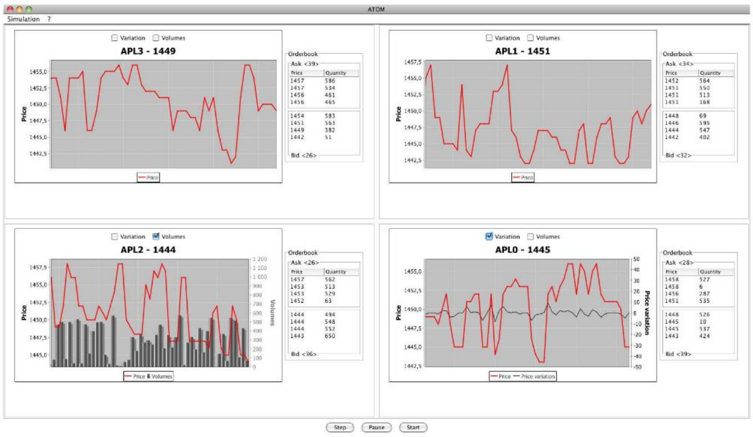
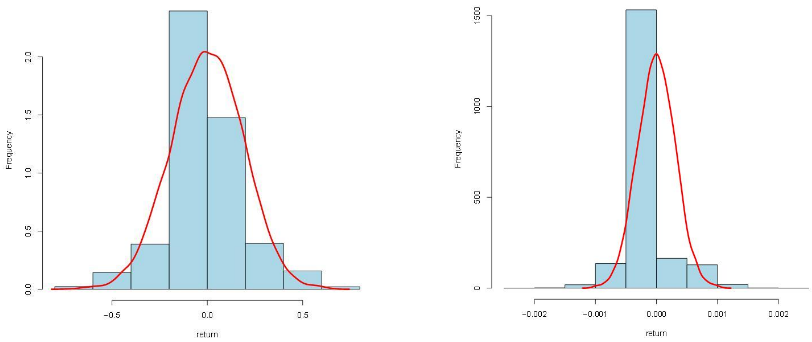


Figure 3.4: An interface window in ATOM representing the price evolution of several assets during simulation

to replay an order flow. These two classes generally suffice for a large number of experimentations/learning activities. The trading agents can be either all artificial or a mix of humans and artificial traders.



(a) Departure from Normality, ATOM

(b) Departure from Normality, Euronext-NYSE

Figure 3.5: Stylized facts, ATOM versus Euronext-NYSE [116]

A interesting set of investigations, performed through ATOM simulations, on different circumstances under which the econometric properties of artificial stock markets accurately reproduce the real world price dynamics at various granularities, for booth intra-day and extra-day periods, can be found in [166]. [166] argues that generating realistic financial dynamics which reproduce quantitative financial distribution is "out-of-reach" within a ZIT only framework and proposes more complex trading configurations in order to achieve that purpose.

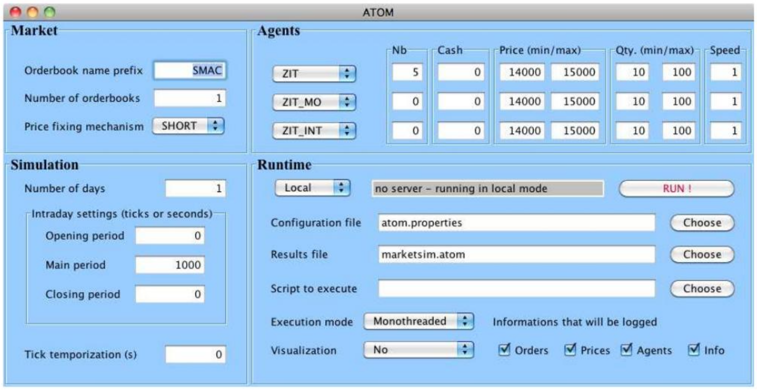


Figure 3.6: An configuration interface window in ATOM

3.3.1.2 Short overview of the ATOM concepts

In ATOM, the market place is represented by either a set of order books or a single market maker. The trading agents must be listed with the market place in order to trade. When agents, artificial or human, send orders at a given price, they try to be executed if possible.

The framework offers the most current types of orders used in electronic markets:

- Market orders - orders to buy/sell at the current price;
- Limit orders - orders to buy at no more than a specific price, or to sell at no less than a specific price;
- Cancel orders - to cancel an order which has not yet been executed;
- Update orders - to set new pricing to existing orders;
- Stop-loss orders - orders to sell a security when it reaches a certain price. A stop-loss order is designed to limit an investor's loss on a security position. They are also known as a "stop orders" or "stop-market orders";
- Stop-limit orders - they combine the features of stop order with those of a limit order. A stop-limit order will be executed at a specified price (or better) after a given stop price has been reached. Once the stop price is reached, the stop-limit order becomes a limit order to buy (or sell) at the limit price or better;
- Iceberg orders - large single orders that has been divided into smaller lots, usually through the use of an automated program, for the purpose of hiding the actual order quantity.

In ATOM, an agent is an entity that sends orders to the market place or posts bid and ask prices if it is a market maker. Agents may posses no or very sophisticated behaviours, in accordance with the complexity of the experiments. Agents with no behaviour, stand for order flow sources either systematic or exogenous (real-world order flows). Agent with endowed with specific behaviour, are generally in full control of the size, frequency or periodicity at witch they place the orders. The market place frequently uses a pop-order system when it comes to interacting to them : at specific trading times they are asked to decide whether or not they want to place a new order.

The framework offers three main types of trading agents :

- Zero Intelligence Traders (ZIT) - Their behaviour is merely stochastic as they essentially reproduce a Bernoulli discrete process. In other words, at every market request, an independent identically distributed Bernoulli trial Binomial$(1, \frac{1}{2})$ of parameter $\frac{1}{2}$ takes place. According to the outcome of the trial (0 or 1), the ZIT agent sends either an ask or a bid order; the size and the price of the order are randomly chosen. The ZITs [73] are particularly useful in financial simulation of intra-day trading sessions, as there is a considerable amount of empirical evidence that ZIT-only trading panels suffice to reproduce the main stylized facts of intra-day price time-series (the main stylized facts seem not to be the result of rational behaviour);

- Technical Traders - Technical analysis is a methodology for predicting the direction and the future value of prices through the study of past market data. It tries to identify patterns, primarily, in price and volume time series, without paying any attention to the economic and/or financial fundamentals of the traded security. The technical analysis can be applied to any type of security and is often used as a complementary toolbox in investment decision making. It provides the chartist traders [29, 30, 132] with different types of buying or selling signals; these signals can be either directly used [10] or combined and filtered in specific ways to elaborate more sophisticated trading strategies/algorithms [150, 151]. The trading algorithms often make use of learning strategies meant to optimize some specific profit and risk related objective functions.

  In ATOM, the agents can use such technical indicators as Simple Moving Average (SMA), Exponential Moving Average (EMA), Moving Average Crossovers, Momentum, Relative Strength Index (RSI), etc. in order to decide the timing, the size and the direction of their orders;

- Sophisticated Intelligence Traders (SIT) - There are three types of SIT [116]:
  - Finite-State Agents and Hollow Agents - as finite-state machine models offer simple and natural representations for basic economic and financial features, the framework offers this useful programming modelling possibility for the trading agents. Finite-State agents posses a small level of sophistication in their behaviours;
  - Cognitive Agents - can be endowed with full artificial intelligence, learning and analyses capabilities, complex strategies and expectations. Multi-asset investors can be seen as a SIT examples while allocating their wealth in assets with different risk levels in accordance with their specific utility functions. [28] [160] and [153] are other examples of cognitive agents that implement order chunking or complex trading strategies;
  - Evolutionary Agents - these are agents based on Evolutionary Algorithms(EA) . EAs [13], citeeib2003, [43] are generic population-based stochastic optimization algorithms inspired by biological evolution processes that allow populations of individuals to adapt to their environment. The adaptation implies the survival of the fittest individuals. The evolution mechanisms encompass a set of stochastic operators [41], known as reproduction, mutation, recombination, and selection, and are iteratively applied on every individual of the population. Evolutionary agents outperform cognitive agents in terms of complexity since they evolve with their environment. Each and every type of evolutionary agent can be implemented in ATOM.

  ATOM offers sophisticated intelligence agents with different utility functions [166] :

  - Constant Relative Risk Aversion (CRRA)
  - Constant Absolute Risk Aversion (CARA)
  - Logarithmic
  - Quadratic

A simulation instance encompasses a market, its associated trading panel and the definition of a trading day. The simulation object enables us to modify fine tuning properties such as, for instance, which of the two types of execution mono or multi-threaded we want to use.

The relative simplicity with which ATOM allows programmers to construct simulation is illustrated in Fig. 3.7 where as few as a dozen lines of java code are only necessary in order to build a fully functional simulation.

```java
import v12.*;
import v12.agents.*;
class SimpleAgent extends Agent {

    public SimpleAgent(String name) {
        super(name);
    }

    public Order decide(String obName, Day day) {
        // Alterning ASK and BID orders
        if (day.currentTick() % 2 == 0) {
            // Logging a message in the generated log file
            market.log.info("Sending an ASK order [tick="+day.currentTick()+"]");
            return new LimitOrder(obName, "nop", LimitOrder.ASK, 500, (long) 14000);
        }
        market.log.info("Sending a BID order [tick="+day.currentTick()+"]");
        return new LimitOrder(obName, "nop", LimitOrder.BID, 500, (long) 14000);
    }

    public static void main(String[] args) {
        Simulation sim = new MonothreadedSimulation();
        // Sending all logs to standard output
        sim.setLogger(new Logger(System.out));
        sim.addNewOrderBook("AAPL");
        sim.addNewAgent(new SimpleAgent("Alan"));
        sim.run(Day.createSinglePeriod(MarketPlace.CONTINUOUS, 10), 1);
        sim.market.printState();
    }
}
```

Figure 3.7: A very simple java programming example of simulation with ATOM

The authors indicate, in technical documentation of the product, the following steps to build a simulation with ATOM:

- Chose either a mono or a multi-threaded simulation;
- Define the output formats by using specific logging modalities;
- Create order-books;
- Create the trading agent panel and add the agents to the simulation;
- Launch the simulation over a trading period encompassing one or multiple trading days.

3.4 Conclusions

In this chapter we have given an succinct summary of the agent-based artificial markets along with some insight into the most important related concepts such as bounded-rationality, agent's heterogeneity and some technical aspects as simulation, calibration or validation. We have started with a section dealing with important notions such as Agent-based Models, Agent-Based Computational Economics and Finance then, we offered a short presentation of the artificial agent-based financial markets. We concluded the chapter by an overview of the Artificial Open Market platform. This artificial market environment accurately emulates realistic market dynamics and, thereby, enables us to use it as a basis for assessing the market models studied in second part of the thesis.

Chapter 4

Contributions and Conclusions

This chapter begins by summarizing the most important aspects of the thesis, then, we review the contributions and drown conclusions.

In this thesis, we discuss **flash crash** market circumstances reproduced within **agent-based artificial frameworks** [96]. In the beginning of the thesis (Chapters 2 and 3), we present the context and state of the art of our research, the Flash crash of May 2010 and computer simulation of market behaviour through agent-based artificial financial markets. For specific **prey-predator** [167] **informationally asymmetric** market environments, we propose a new model to predict the dynamics of flash crashes and test its overall efficiency (Chapter 4). As these market events start unfolding, we assess the effectiveness of the newly proposed **VPIN** flash crash detecting metric [46, 55] and observe some of its limitations(Chapter 5). In the following two chapters of the thesis (Chapters 6 and 7), we propose two alternative strategies to compute point estimates of VPIN and compare their evolutions at flash crash time to the ones of VPIN.

We tested all our models within java-based artificial financial framework named ATOM [116, 117]. This simulation platform provided us with a general environment for agent-based simulations of stock markets encompassing such meta-entities as the market micro-structure, the trading agents and their behaviours and the asymmetric information diffusion. The artificial agents in the simulation setting do not rely on the assumptions of rationality and homogeneity as our modelling encompass multiple types of agents with different functions, purposes and behaviours.

Chapter 4 elucidates the dynamics and the features of a market flash crash by describing the asset price processes in time units. Relying on the information asymmetry assumption and considering a stochastic prey-predator Lotka-Volterra [167] model, we developed the deterministic and stochastic equations for the price development at flash crash time. The theoretical foundation of our contribution relies primarily upon the hypothesis that the flash crash is the immediate repercussion of growing levels of order flow toxicity in asymmetric markets. The main ingredient of the model is the assumption that the set of interacting agents which generate at every moment informed transactions behave in a predatory manner towards the uniformed trading population at flash crash time [40].

From our price modelling we determined the fraction of informed trading volume PIN. We theoretically showed that the PIN peak represents a lagging indicator of the flash crash bottom expected price, this being a limitation for any metric estimating PIN. More, the lag between the PIN peak value and the flash crash bottom price realization times represents an indicator of the intelligence of the uninformed traders, i.e. the smaller the lag the higher is the diffusion of the private information within the uninformed trading population. This result theoretically proves the power of PIN in highly efficient markets.

By modelling the price as a function of $volume \cdot PIN$ we showed that PIN is fast to increase but slow to decrease after the flash crash, which theoretically explains empirical findings of Easley et al. [46] and their opponents [7].

Based on our stochastic modelling, using VPIN as a proxy for PIN, we developed an estimation methodology for inferring the model parameters during the flash crash unfolding. This allows one to early predict the amplitude of the crash, the expected time of the flash crash bottom price, flash crash duration and the expected price trajectory.

We tested our model within an artificial market simulator and showed that, for un-leveraged markets, we were able to accurately estimate flash crash dynamics and characteristics during its unfolding.

We also approached the general case of a competition amongst the informed traders and showed how our modelling can fit such a market, making our model of general usage. We showed that competition reduces the lag of the PIN metric both punctually and asymptotically for low levels of learning rates in the uniformed population.

We provided a straightforward interpretation of the interchangeable and multiplicative effects that both increased competition and higher intelligence levels have on the PIN effectiveness. A higher intelligence level of the uniformed trading population can substitute for a lack of competitiveness among the informed traders and vice versa. Both of them have compounded multiplying effects that enhance the effectiveness of the PIN as an early flash crash estimator. The substitution rates depend primarily on the ability of informed traders to over exploit the private information.

Obviously, all the desirable properties and unsatisfactory limitations that we related to the *Probability of Informed Trading* [55] variable in the context of a informationally driven predatory flash crash, hold true for any of its good estimators including VPIN.

In chapter 5 we gauge the reliability of the VPIN metric as an order flow-toxicity estimator within an artificial agent-based stock market populated with zero-intelligence traders and trend-followers. We verify whether the VPIN wrongly detects informed trading in a very specific experimental set-up where all market participants are uniformed. We evaluate the expected behaviour of the metric through Monte Carlo simulation. We discover that it may erroneously "disclose" information asymmetry in environments where all the operating traders are uniformed.

Furthermore, a considerable surge in percentage of trade-followers can induce unpredictable and volatile evolutions of the VPIN. We also notice behavioural asymmetry when prices are trending upwards and downwards as VPIN is more likely to rise during downward trends. We conclude that, in a high frequency trading environment, VPIN gauges not only the level of order flow-toxicity, but more generally, the degree to which a significant minority fraction of the market participants share a collective alternative vision of price development which makes them function and operate as a unique aggregate and potent market actor. This correlated behaviour ultimately results in significant imbalances between overall supply and demand, and assimilates the VPIN metric to a by-product proxy of systematic excess demand or supply regardless of the underpinning cause.

In chapter 5 we also put forth a brief theoretical argument aiming to explain the VPIN lingering effect observed during simulation phase in accordance with the findings of [7]. The argument relied on slightly modified assumptions for the VPIN model.

We proved that the modified VPIN model can be assimilated to a standard VPIN model as long as the ratio of informed volume yet to flow is proportionally significant. When the ratio of informed volume yet to flow is proportionally small the intensity of informed volume flowing is decreasing logarithmically in informed-volume time. Therefore, this slightly modified VPIN model of chapter 5 seems to offer a better theoretical interpretation of the VPIN lingering behaviour and can provide the basis for future developments in the field.

The theoretical models of chapters 6 and 7 rely mainly upon binomial and trinomial lattice approaches which accurately describe the martingale philosophy behind the market making activity [123]: the hypothesis that, at the local level, the evolution of publicly traded asset prices follows random motions, where the probability, at any moment, of an uptick or downtick movement in the price are identical and, thus, the economic viability of the market makers stems from positive gain expectations provided by the spread they perceive on each transaction.

Chapter 6 outlines a *Support Vector Machine* [101] based discrete volume-time approach to estimating the *Probability of Informed Trading*. We aim to offer an alternative metric of the extent to which liquidity providers are prone to systematically engage into losing transactions within a high frequency trading environment characterized systematic overbuying or overselling conditions. Our theoretical model supposes the existence of two homogeneous classes of markets participants: the informed traders who attempt to take advantage of the private information set they possess and the uninformed traders who provide liquidity on the market. We assume that the expected price depends on the excess demand generated by informed traders. Under this assumptions, we derive a time point equation for the PIN and offer a matrix based method for estimating it. Just as VPIN our metric is a good estimator of sheer systematic imbalances between supply and demand. Through causality tests and correlation performed under various market conditions we prove that our metric is roughly equivalent to VPIN. We show that our estimator is more stable displaying a lesser amount of volatility and a smaller percentage of outliers.

Chapter 7 confirms that simple volume-time algorithmic approaches combined with elements of optimization theory can produce effective estimators of the PIN metric in environments which replicate much of the real-world markets price series dynamics. It might, therefore, be used outside of this rather simplistic and confined simulation framework, in a solving a substantial range of financial problems which require PIN estimations [18, 44, 45, 53, 85].

The assessment of this algorithmic approach was done by using different types of settings and informationally asymmetric classes of trading agents. Under these conditions the algorithmic computation of PIN was close enough of the VPIN metric in most of the cases including situations with high VPIN values indicating possible occurrence of a crash. Through causality tests and correlation performed under various market conditions we empirically prove that the our metric approaches the quality of VPIN. As the accumulation of inventory and the overall volatility of the prices have considerable impact on the market makers ability not to desert the market, we also discussed very briefly the optimization aspects that the algorithmic approach may take into account.

The contributions brought about by this thesis deepen the understanding of statistically rare flash crash events and validate the usefulness of agent-based artificial market models in studying this important financial topic. We make use of extensive simulations in order to relate the evolution of heterogeneous bounded-rational [139] agent populations with asset price dynamics during flash crash unfolding. Through direct application of the Monte-Carlo methods we identify expected behaviours under flexible assumptions. Repeated simulations exhibit the direct relationships which exist between specific proportions of market participants applying correlated trading strategies, and high values of flash crash predicting metrics. We use agent-based research methodology along with other analytical instruments stemming from mathematical and econometric fields in order to provide rigorous explanations for the observed market dynamics.

List of Publications

STAN, A. Assessing inbound call centers scheduling through bootstrapping and gdp based monte carlo. *Review of Economic Studies and Research Virgil Madgearu*, 2 (2011), 135–147.

STAN, A. An automata based approach to modeling real-time trading applications. *Review of Economic Studies and Research Virgil Madgearu*, 2 (2011), 135–147.

STAN, A. Day trading the emerging markets using multi-time frame technical indicators and artificial neural networks. In *Advanced Intelligent Computational Technologies and Decision Support Systems*, B. Iantovics and R. Kountchev, Eds., vol. 486 of *Studies in Computational Intelligence*. Springer International Publishing, 2014, pp. 191–200.

STAN, A., AND MOLDOVAN, D. Assessing the dependability of the vpin metric as an order flow-toxicity estimator through a specific high frequency trading experimental setup. In *in Proceedings fo the 12th International conference on Practical on Informatics in Economy (IE2013)* (2013), ASE, Bucharest, Romania, pp. 589–593.

STAN, A., AND MOLDOVAN, D. A neural network vwap algorithm for artificial financial markets populated with zero-intelligence trading agents. In *in Proceedings fo the 12th International conference on Practical on Informatics in Economy (IE2013)* (2013), ASE, Bucharest, Romania, pp. 609–612.

References

[1] Review of the market events of may 6, 2010, 2010. 11

[2] *Simulation in Computational Finance and Economics: Tools and Emerging Applications*. 2013. 7

[3] Abad, D., and Yage, J. From {PIN} to vpin: An introduction to order flow toxicity. *The Spanish Review of Financial Economics 10*, 2 (2012), 74 – 83. 18, 19

[4] Akaike, H. On entropy maximization principle. *Applications of Statistics* (1977), 27–41.

[5] Alfarano, S., Lux, T., and Wagner, F. Empirical validation of stochastic models of interacting agents. *The European Physical Journal B 55*, 2 (2007), 183–187. 25

[6] Amaral, L., Cizeau, P., Gopikrishnan, P., Liu, Y., Meyer, M., Peng, C.-K., and Stanley, H. Econophysics: can statistical physics contribute to the science of economics? *Computer Physics Communications 121122*, 0 (1999), 145 – 152. ¡ce:title¿Proceedings of the Europhysics Conference on Computational Physics {CCP} 1998¡/ce:title¿. 26

[7] Andersen, T. S., and Bondarenko, O. VPIN and the Flash Crash. *SSRN* (2011). http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1881731. 6, 39, 40

[8] Anufriev, M., Bottazzi, G., and Pancotto, F. Equilibria, stability and asymptotic dominance in a speculative market with heterogeneous traders. *Journal of Economic Dynamics and Control 30*, 9-10 (2006), 1787–1835. 24

[9] Arthur, W. *Increasing Returns and Path Dependence in the Economy*. Ann Arbor Paperbacks. University of Michigan Press, 1994. 24

[10] Arthur, W. Inductive reasoning and bounded rationality. *American Economics Review*, 93 (2011), 1–7. 33

[11] Arthur, W. B. Out-of-Equilibrium Economics and Agent-Based Modeling. vol. 2. 2006, ch. 32, pp. 1551–1564. 23

[12] Arthur, W. B., Holl, J. H., Lebaron, B., Palmer, R., and Tayler, P. Asset pricing under endogenous expectations in an artificial stock market. In *Economic Notes* (1997), no. 26, Addison-Wesley, pp. 15–44. 29

[13] Ashlock, D. *Evolutionary Computation for Modeling and Optimization*. Springer, 2006. 34

[14] Badegruber, T. *Agent-based Computational Economics: New aspects in learning speed and convergence in the Santa Fe Artificial Stock Market*. PhD thesis, Universitat Graz, Graz, Austria, 2003.

[15] Barbosa, P., and Maryland, C. *Ecology of Predator-Prey Interactions*. Oxford University Press, USA, 2005. 28

[16] Ben-David, I., Franzoni, F., and Moussawi, R. Etfs, arbitrage, and contagion. Working Paper Series 2011-20, Ohio State University, Charles A. Dice Center for Research in Financial Economics, December 2011. 13

[17] Bethel, E. W., Leinweber, D., Rübel, O., and Wu, K. Federal market information technology in the post flash crash era: roles for supercomputing. In *Proceedings of the fourth workshop on High performance computational finance* (New York, NY, USA, 2011), WHPCF '11, ACM, pp. 23–30. 17, 18

[18] Bharath, S. T., Pasquariello, P., and Wu, G. Does asymmetric information drive capital structure decisions? *Review of Financial Studies 22*, 8 (2009), 3211–3243. 41

[19] Boer, K., Kaymak, U., and Spiering, J. From discrete-time models to continuous-time, asynchronous modeling of financial markets. *Computational Intelligence*, 23 (2007). 29

[20] Boer-Sorban, K. *Agent-Based Simulation of Financial Markets: A Modular, Continuous-time Approach*. PhD thesis, Erasmus University Rotterdam, Rotterdam, The Netherlands, January 2008. 29

[21] Bonabeau, E. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America 99*, Suppl 3 (2002), 7280–7287. 21

[22] Borkovec, M., Domowitz, I., Serbin, V., and Yegerman, H. Liquidity and price discovery in exchange-traded funds. one of several possible lessons from the flash crash. Tech. rep., Investment Technology Group, 2010. 13

[23] Born, B. E., Brennan, J. J., Engle, R. F., Ketchum, R. G., OHara, M., Philips, S. M., Ruder, D. S., and Stiglitz, J. E. Recommendations regarding regulatory responses to the market events of may 6, 2010, 2011. 13

[24] Boswijk, H. P., Hommes, C. H., and Manzan, S. Behavioral heterogeneity in stock prices. *Journal of Economic Dynamics and Control 31*, 6 (June 2007), 1938–1970. 25, 26

[25] Bottazzi, G., Dosi, G., and Rebesco, I. Institutional architectures and behavioral ecologies in the dynamics of financial markets. *Journal of Mathematical Economics 41*, 1-2 (February 2005), 197–228. 24

[26] Bottazzi, G., Dosi, G., and Rebesco, I. Institutional architectures and behavioral ecologies in the dynamics of financial markets. *Journal of Mathematical Economics 41*, 1-2 (February 2005), 197–228. 26

[27] Brahma, A., Chakraborty, M., Das, S., Lavoie, A., and Magdon-Ismail, M. A bayesian market-maker. In *Proceedings of the 13th ACM Conference on Electronic Commerce EC '12* (2012), ACM New York, NY, USA, pp. 215–232. 16

[28] Brandouy, O., Mathieu, P., and Veryzhenko, I. Ex-post optimal strategy for the trading of a single financial asset. In *The Generative Method in Economics*, C. Hernandez, M. Posada, and A. Lopez-Paredes, Eds., no. 93 in Lecture Notes in Economics and Mathematical Systems. Springer, 2009, pp. 171–184. 34

[29] Brianzoni, S. Wealth distribution in an asset pricing model: the role of the switching mechanism. *Applied Mathematical Sciences*, 6 (2012), 423–442. 33

[30] Brianzoni, S., Mammana, C., and E.Michetti. Updating wealth in an asset pricing model with heterogeneous agents. *Discrete Dynamics in Nature and Society*, 27 (2010). 33

[31] Chan, N., and Shelton, C. Technical report. In *An electronic market-maker*. MIT, Cambridge, MA, 2001. 28

[32] Chen, S.-H., Chang, C.-L., and Du, Y.-R. Agent-based economic models and econometrics. *The Knowledge Engineering Review 27* (June 2012), 187–219. 2, 7, 25, 26, 27, 28

[33] Chen, S.-H., Kampouridis, M., and Tsang, E. Microstructure dynamics and agent-based financial markets. In *Multi-Agent-Based Simulation XI*, T. Bosse, A. Geller, and C. Jonker, Eds., vol. 6532 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2011, pp. 121–135. 7

[34] Chen, S.-H., Tai, C.-C., and Chie, B.-T. Individual rationality as a partial impediment to market efficiency. In *Genetic Algorithms and Genetic Programming in Computational Finance*, S.-H. Chen, Ed. Springer US, 2002, pp. 357–377. 26

[35] Chiarella, C., and He, X.-Z. Dynamics of beliefs and learning under al processes - the heterogeneous case. Research Paper Series 55, Quantitative Finance Research Centre, University of Technology, Sydney, 2001. 24

[36] Chiarella, C., and Iori, G. A simulation analysis of the microstructure of double auction markets. *Quantitative Finance 2*, 5 (2002), 346–353. 24

[37] Cincotti, S., Ponta, L., and Raberto, M. A multi-assets artificial stock market with zero-intelligence traders. In *WEHIA* (June 15 2005). 29

[38] Cont, R. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance 1* (2001), 223–236. 25, 27

[39] Cont, R. Volatility clustering in financial markets: Empirical facts and agent-based models. In *Long Memory in Economics*, G. Teyssire and A. Kirman, Eds. Springer Berlin Heidelberg, 2007, pp. 289–309. 27

[40] Corcoran, C. *Systemic Liquidity Risk and Bipolar Markets: Wealth Management in Today's Macro Risk On / Risk Off Financial Environment.* Wiley Finance, February 2013. 6, 8, 18, 28, 38

[41] da Costa Pereira, C., Mauri, A., and Tettamanzi, A. Cognitive-agent-based modeling of a financial market. In *IAT* (2009), IEEE, pp. 20–27. 34

[42] Darley, V., and Outkin, A. *A NASDAQ Market Simulation: Insights on a Major Market from the Science of Complex Adaptive Systems.* Complex systems and interdisciplinary science. World Scientific, 2007. 29

[43] De Jong, K. A. *Evolutionary Computation: A unified approach.* Springer, 2002. 34

[44] Duarte, J., Han, X., Harford, J., and Young, L. Information asymmetry, information dissemination and the effect of regulation FD on the cost of capital . *Journal of Financial Economics 87*, 1 (2008), 24 – 44. 41

[45] Duarte, J., and Young, L. Why is pin priced? *Journal of Financial Economics 91*, 2 (2009), 119–138. 41

[46] Easley, D., de Prado, M. L., and O'Hara, M. The microstructure of the flash crash: Flow toxicity, liquidity crashes, and the probability of informed trading. *The Journal of Portfolio Management 37*, 2 (2010), 118–128. 3, 4, 6, 14, 15, 16, 28, 38, 39

[47] Easley, D., de Prado, M. L., and O'Hara, M. The exchange of flow toxicity. *The Journal of Trading 6*, 2 (2011), 8–13. 6

[48] Easley, D., de Prado, M. L., and O'Hara, M. Bulk classification of trading activity. *Johnson School Research Paper Series 8* (2012). 6, 14

[49] Easley, D., de Prado, M. L., and O'Hara, M. Flow toxicity and liquidity in a high-frequency world. *Review of Financial Studies 25*, 5 (2012), 1457–1493. 13, 18

[50] Easley, D., de Prado, M. L., and O'Hara, M. The volume clock: Insights into the high-frequency paradigm. *The Journal of Portfolio Management 39*, 1 (2012), 19–29.

[51] Easley, D., de Prado, M. L., and O'Hara, M. Optimal execution horizon. *Mathematical Finance, to appear* (2013). 14

[52] Easley, D., Engle, R. F., O'Hara, M., and Wu, L. Time-varying arrival rates of informed and uninformed trades. *Journal of Financial Econometrics 6*, 2 (2008), 171–207.

[53] Easley, D., Hvidkjaer, S., and O'Hara, M. Is information risk a determinant of asset returns? *The Journal of Finance 57*, 5 (2002), 2185–2221. 41

[54] Easley, D., Kiefer, N. M., and O'Hara, M. One day in the life of a very common stock. *Review of Financial Studies 10*, 3 (1997), 805–35.

[55] Easley, D., Kiefer, N. M., O'Hara, M., and Paperman, J. B. Liquidity, information, and infrequently traded stocks. *The Journal of Finance 51*, 4 (1996), 1405–1436. 6, 15, 16, 28, 38, 39

[56] Easley, D., and O'Hara, M. Price, trade size, and information in securities markets. *Journal of Financial Economics 19*, 1 (September 1987), 69–90.

[57] Easley, D., and O'Hara, M. Microstructure and ambiguity. *The Journal of Finance 65*, 5 (2010), 1817–1846.

[58] Ehrentreich, N. A corrected version of the santa fe institute artificial stock market model. *Social Science Computer Review 20*, 2 (September 2002), 174–196. 29

[59] Ehrentreich, N. Technical trading in the santa fe institute artificial stock market revisited. *Journal of Economic Behavior & Organization 61*, 4 (December 2006), 599–616. 29

[60] Engle, R. F., and Russell, J. Analysis of high frequency and transaction data. In *Handbook of Financial Econometrics*, Y. Ait-Sahalia and L. Hansen, Eds., Lecture Notes in Financial Econometrics. North Holland, 2010. 27

[61] Epstein, J. M. Remarks on the foundations of agent-based generative social science. In *Handbook of Computational Economics*, L. Tesfatsion and K. L. Judd, Eds., vol. 2 of *Handbook of Computational Economics*. Elsevier, 2006, ch. 34, pp. 1585–1604. 22

[62] Fama, E. F. The Behavior of Stock-Market Prices. *The Journal of Business 38*, 1 (1965), 34–105. 24

[63] Fama, E. F., and Blume, M. E. Filter Rules and Stock-Market Trading. *The Journal of Business 39*, 1 (1966), 226–241. 24

[64] Farmer, J. D. Market force, ecology and evolution. *Industrial and Corporate Change 11*, 5 (2002), 895–953. 24

[65] Farmer, J. D., and Foley, D. The economy needs agent-based modelling. *Nature 460*, 7256 (August 2009), 685–686. 22

[66] Farmer, J. D., and Joshi, S. The price dynamics of common trading strategies. *Journal of Economic Behavior & Organization 49*, 2 (October 2002), 149–171. 24

[67] Farmer, J. D., and Lux, T. Introduction to special issue on applications of statistical physics in economics and finance. *Journal of Economic Dynamics and Control 32*, 1 (2008), 1 – 6. Applications of statistical physics in economics and finance. 27

[68] Follmer, H., Horst, U., and Kirman, A. Equilibria in financial markets with heterogeneous agents: a probabilistic perspective. *Journal of Mathematical Economics 41*, 1-2 (February 2005), 123–155. 3, 22

[69] Gabaix, X., Gopikrishnan, P., Plerou, V., and Stanley, H. E. A theory of power-law distributions in financial market fluctuations. *Nature*, 6937 (2003), 267270. 27

[70] Gandolfo, G. Giuseppe palomba and the lotka-volterra equations. *Rendiconti Lincei 19*, 4 (2008), 347–357.

[71] Gevel, A. v. d., and Noussair, C. The nexus between artificial intelligence and economics. Discussion Paper 2012-087, Tilburg University, Center for Economic Research, 2012. 22

[72] Glosten, L. R., and Milgrom, P. R. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics 14*, 1 (March 1985), 71–100.

[73] Gode, D., and Sunder, S. Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. *Journal of Political Economy*, 101 (1993), 119–137. 25, 26, 33

[74] Gode, D. K., and Sunder, S. What makes markets allocationally efficient? *The Quarterly Journal of Economics 112*, 2 (May 1997), 603–630. 26

[75] Grant, J. An algo that prevents crashes. Advanced Trading, March 18 2011. 12

[76] Grocer, S. Senators seek regulators report on causes of market volatility. WallStreet Journal, May 7 2010. 5, 12

[77] Halpern, J. Y. computer science and game theory. In *The New Palgrave Dictionary of Economics*, S. N. Durlauf and L. E. Blume, Eds. Palgrave Macmillan, Basingstoke, 2008. 23

[78] Hatemi-J, A. Asymmetric causality tests with an application. *Empirical Economics 1*, 43 (2012), 447–456.

[79] Hein, O., Schwind, M., and Spiwoks, M. Frankfurt artificial stock market: a microscopic stock market model with heterogeneous interacting agents in small-world communication networks. *Journal of Economic Interaction and Coordination 3*, 1 (2008), 59–71. 29

[80] Holland, J. H., and Miller, J. H. Artificial adaptive agents in economic theory. *American Economic Review 81*, 2 (May 1991), 365–71. 2, 22

[81] Hommes, C. Financial markets as nonlinear adaptive evolutionary systems. *Quantitative Finance 1*, 1 (2001), 149–167. 2, 24

[82] Hommes, C. Heterogeneous agent models in economics and finance. *Elsevier* (2005). 24

[83] Hughes, H. P. N., Clegg, C. W., Robinson, M. A., and Crowder, R. M. Agent-based modelling and simulation: The potential contribution to organizational psychology. *Journal of Occupational and Organizational Psychology 85*, 3 (2012), 487–502. 21

[84] J. Farmer, P. P., and Zovko, I. The predictive power of zero intelligence models in financial markets. vol. 102, National Academy of Sciences of the United States of America, pp. 2254–2259. 28

[85] Jayaraman, S. Earnings volatility, cash flow volatility, and informed trading. *Journal of Accounting Research 46*, 4 (2008), 809–851. 41

[86] Jennings, N. R. Agent-based computing: promise and perils. In *Proceedings of the 16th international joint conference on Artificial intelligence - Volume 2* (San Francisco, CA, USA, 1999), IJCAI'99, Morgan Kaufmann Publishers Inc., pp. 1429–1436. 2, 21

[87] Jevons, W. *The Theory of Political Economy*. Macmillan and Company, 1871. 24

[88] Johnson, P. What i learned from the artificial stock market. *Social Science Computer Review 20*, 2 (September 2002). 29

[89] Judd, K. L. Computationally intensive analyses in economics. In *Handbook of Computational Economics*, L. Tesfatsion and K. L. Judd, Eds., vol. 2 of *Handbook of Computational Economics*. Elsevier, January 2006, ch. 17, pp. 881–893. 22, 26

[90] Kahneman, D. Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review 93*, 5 (September 2003), 1449–1475. 24

[91] Kirilenko, A. A., Kyle, A. S., Samadi, M., and Tuzun, T. The flash crash: The impact of high frequency trading on an electronic market. *Capital Markets: Market Microstructure eJournal 5* (2011). 3, 13

[92] Kleinert, H. *Path Integrals in Quantum Mechanics, Statistics, Polymer Physics, and Financial Markets, 5th edition*. World Scientific Publishing Company, 2009.

[93] Kocherlakota, N. R. The equity premium: it's still a puzzle. 27

[94] Kyle, A. S. Continuous auctions and insider trading. *Econometrica 53*, 6 (November 1985), 1315–35.

[95] Ladley, D., and Schenk-Hoppe, K. R. Do stylised facts of order book markets need strategic behaviour? *Journal of Economic Dynamics and Control 33*, 4 (2009), 817 – 831. 26

[96] LeBaron, B. A builder's guide to agent-based financial markets. *Quantitative Finance 1*, 2 (2001), 254–261. 3, 25, 38

[97] LeBaron, B. Evolution and time horizons in an agent based stock market. *Macroeconomic Dynamics 5*, 2 (2001), 225–254.

[98] Lebaron, B. Building the santa fe artificial stock market. working paper, graduate. In *School of International Economics and Finance, Brandeis* (2002), pp. 1117–1147. 29

[99] LeBaron, B. Empirical regularities from interacting long and short memory investors in an agent based stock market. *IEEE Transactions on Evolutionary Computation 5*, 2 (2004), 442–455.

[100] LeBaron, B., Arthur, W., and Palmer, R. Time series properties of an artificial stock market. *Journal of Economic Dynamics and Control 23*, 2 (1999), 1487–1516.

[101] Lee, Y., Lin, Y., , and Wahba, G. Multicategory support vector machines, theory, and application to the classification of microarray data and satellite radiance data. *Journal of the American Statistical Association 465*, 99 (2004), 67–81. 40

[102] Levy, H., Levy, M., and Solomon, S. *Microscopic Simulation of Financial Markets: From Investor Behavior to Market Phenomena.* Elsevier Science, 2000. 24

[103] Liu, X., Gregor, S., and Yang, J. The effects of behavioral and structural assumptions in artificial stock market. *Physica A: Statistical Mechanics and its Applications 387*, 11 (2008), 2535 – 2546. 26

[104] Lotka, A. J. Analytical note on certain rhythmic relations in organic systems. *Proceedings of the National Academy of Sciences of the United States of America 6* (1920), 410–415.

[105] Lotka, A. J. Undamped oscillations derived from the law of mass action. *Journal of the American Chemical Society 42* (1920), 1595–1599.

[106] Lotka, A. J. *Elements of Physical Biology.* Williams and Wilkins Company, 1925.

[107] Lux, T. {CHAPTER} 3 - stochastic behavioral asset-pricing models and the stylized facts. In *Handbook of Financial Markets: Dynamics and Evolution*, T. Hens and K. R. Schenk-Hopp, Eds., Handbooks in Finance. North-Holland, San Diego, 2009, pp. 161 – 215. 26

[108] Lux, T., and Marchesi, M. Scaling and Criticality in a Stochastic Multi-Agent Model of a Financial Market. *Nature 397* (November 2 1999), 498–500. 24

[109] Madhavan, A. Exchange-Traded funds, market structure and the flash crash. *Social Science Research Network Working Paper Series* (September 2011). 13

[110] Malkiel, B. G., and Fama, E. F. Efficient capital markets: A review of theory and empirical work*. *The Journal of Finance 25*, 2 (1970), 383–417. 24

[111] Marchesi, M., Cincotti, S., Focardi, S., and Raberto, M. The genoa artificial stock market: Microstructure and simulations. In *Heterogenous Agents, Interactions and Economic Performance*, R. Cowan and N. Jonard, Eds., vol. 521 of *Lecture Notes in Economics and Mathematical Systems*. Springer Berlin Heidelberg, 2003, pp. 277–289. 29

[112] Marchesi, M., Cincotti, S., Focardi, S., and Roberto, M. Development and testing of an artificial stock market. *Model Dynamic in Economic Finance* (2000). 29

[113] Markose, S. M. The new evolutionary computational paradigm of complex adaptive systems: Challenges and prospects for economics and finance. Tech. rep., 2001. 2

[114] Marks, R. Validating simulation models: A general framework and four applied examples. *Computational Economics 30*, 3 (October 2007), 265–290. 22, 25

[115] Maslov, S. Simple model of a limit order-driven market. *Physica A: Statistical Mechanics and its Applications 278*, 3-4 (Apr. 2000), 571–578. 26

[116] Mathieu, P., and Brandouy, O. A generic architecture for realistic simulations of complex financial dynamics. In *Advances in Practical Applications of Agents and Multiagent Systems*, Y. Demazeau, F. Dignum, J. M. Corchado, and J. B. Perez, Eds., vol. 70 of *Advances in Intelligent and Soft Computing*. Springer, 2010, pp. 185–197. 29, 30, 31, 34, 38

[117] Mathieu, P., and Brandouy, O. Introducing atom. In *Advances on Practical Applications of Agents and Multi-Agent Systems*, Y. Demazeau,

J. P. Muller, J. M. C. Rodriguez, and J. B. Perez, Eds., vol. 155 of *Advances in Intelligent and Soft Computing*. Springer, 2012, pp. 269–272. 38

[118] Mazliak, L., and Shafer, G. The splendors and miseries of martingales. *Electronic Journal for History of Probability and Statistics. Entire issue dedicated to Martingale probability theory. 1*, 5 (June 2009).

[119] Menger, C. *Principles of economics.* reprinted by University Press 1983, 1871. 24

[120] Menkveld, A. J., and Yueshen, B. Z. Anatomy of the Flash Crash. *SSRN* (April 2013). http://ssrn.com/abstract=2243520 or http://dx.doi.org/10.2139/ssrn.2243520. 11, 13, 14, 18

[121] Muranaga, J., and Ohsawa, M. Measurement of liquidity risk in the context of market risk calculation. the measurement of aggregate market risk. Tech. rep., Bank for International Settlements, 1998. 27

[122] Nature. A model approach. *Nature 460* (August 2009). 22, 23

[123] Nevmyvaka, Y., Sycara, K., and Seppi, D. J. Electronic Market Making: Initial Investigation. *Electronic Carnegie Mellon University Database* (2004). 40

[124] Niazi, M., and Hussain, A. Agent-based computing from multi-agent systems to agent-based models: a visual survey. *Scientometrics 89*, 2 (2011), 479–499. 21

[125] Nuti, G. *An electronic market-making algorithm.* PhD thesis, UCL (University College London), 2008. 28

[126] Oksendal, B. *Stochastic Differential Equations. An Introduction with Applications, 6th edition.* Springer, 2010.

[127] O'Sullivan, A., and Sheffrin, S. *Economics: Principles in Action.* Prentice Hall, 2005.

[128] Othman, A. M. *Automated Market Making: Theory and Practice.* PhD thesis, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213 USA, 2012. http://repository.cmu.edu/dissertations/159.

[129] Pachamanova, D. A., and Fabozzi, F. J. *Simulation and Optimization in Finance.* John Wiley & Sons, Inc., 2010. 26

[130] Pacurar, M. Autoregressive conditional duration models in finance: A survey of the theoretical and empirical literature. *Journal of Economic Surveys 22*, 4 (2008), 711–751. 27

[131] Palmer, R. G., Arthur, W. B., Holland, J. H., LeBaron, B., and Tayler, P. Artificial economic life: a simple model of a stockmarket. *Physica D 75*, 1-3 (Aug. 1994), 264–274. 29

[132] Ponta, L., Raberto, M., and Cincotti, S. A multi-assets artificial stock market with zero-intelligence traders. A Letter Journal Exploring the Frontiers of Physics. *Applied Mathematical Sciences 84*, 2 (2011), 406–411. 33

[133] Popescu, M. *Two Essays on the Probability of Informed Trading*. UMI, 2009. 6

[134] Rao, C. R., and Toutenburg, H. *Linear models: least squares and alternatives*. Springer, 1999.

[135] Ravshanbek, D., and Gappar, N. Limit cycle, trophic function and the dynamics of intersectoral interaction. *Current Research Journal of Economic Theory 2*, 2 (2010), 32–40.

[136] Rosser, J. *Complexity in Economics*. The international library of critical writings in economics. Elgar, 2004. 23

[137] Rosser, J.Barkley, J. Speculations on nonlinear speculative bubbles. *Nonlinear Dynamics, Psychology, and Life Sciences 1*, 4 (1997), 275–300. 27

[138] Roth, A. E. The economist as engineer: Game theory, experimentation, and computation as tools for design economics. *Econometrica 70*, 4 (2002), 1341–1378. 23

[139] Rubinstein, A. *Modeling Bounded Rationality*. Zeuthen lecture book series. MIT Press, 1998. 2, 25, 41

[140] Samanidou, E., Zschischang, E., Stauffer, D., and Lux, T. Agent-based models of financial markets. *Reports on Progress in Physics 70*, 3 (2007), 409. 25

[141] SARGENT, T. J. *Bounded rationality in macroeconomics / Thomas J. Sargent*. Clarendon Press ; Oxford University Press Oxford [England] : Oxford ; New York, 1993. 23

[142] SCHELLING, T. C. *Schelling Micromotives and Macrobehavior (Fels lectures on public policy analysis)*. W W Norton & Co Ltd, April 1978. 23

[143] SCHWARZ, G. E. Estimating the dimension of a model. *Annals of Statistics 2*, 6 (1978), 461–464.

[144] SEC, AND CFTC. Findings regarding the market events of may 6, 2010. Tech. rep., U.S. Securities and Exchange Commission and Commodity Futures Trading Commission, New York, NY, United States, 2010. 13

[145] SHOHAM, Y. Computer science and game theory. *Commun. ACM 51*, 8 (Aug. 2008), 74–79. 23

[146] SIMON, H. A. Rational choice and the structure of the environment. *Psychological Review 63*, 2 (1956), 129–38. 24

[147] SINDHYA, K., DEB, K., AND MIETTINEN, K. A local search based evolutionary multi-objective optimization approach for fast and accurate convergence. *Parallel Problem Solving from Nature. Lecture Notes in Computer Science*, 5199 (2008), 815–837.

[148] SINDHYA, K., RUIZ, A. B., AND MIETTINEN, K. A preference based interactive evolutionary algorithm for multi-objective optimization: Pie. *Evolutionary Multi-Criterion Optimization. Lecture Notes in Computer Science*, 6576 (2011), 212–224.

[149] STAN, A. Assessing inbound call centers scheduling through bootstrapping and gdp based monte carlo. *Review of Economic Studies and Research Virgil Madgearu*, 2 (2011), 135–147. 8

[150] STAN, A. An automata based approach to modeling real-time trading applications. *Review of Economic Studies and Research Virgil Madgearu*, 2 (2011), 135–147. 8, 22, 33

[151] STAN, A. Day trading the emerging markets using multi-time frame technical indicators and artificial neural networks. In *Advanced Intelligent Computational Technologies and Decision Support Systems*, B. Iantovics

and R. Kountchev, Eds., vol. 486 of *Studies in Computational Intelligence*. Springer International Publishing, 2014, pp. 191–200. 8, 33

[152] Stan, A., and Moldovan, D. Assessing the dependability of the vpin metric as an order flow-toxicity estimator through a specific high frequency trading experimental setup. In *in Proceedings fo the 12th International conference on Practical on Informatics in Economy (IE2013)* (2013), ASE, Bucharest, Romania, pp. 589–593. 8

[153] Stan, A., and Moldovan, D. A neural network vwap algorithm for artificial financial markets populated with zero-intelligence trading agents. In *in Proceedings fo the 12th International conference on Practical on Informatics in Economy (IE2013)* (2013), ASE, Bucharest, Romania, pp. 609–612. 8, 34

[154] Sutton, R. S., and Barto, A. G. *Introduction to Reinforcement Learning*, 1st ed. MIT Press, Cambridge, MA, USA, 1998. 23

[155] Tavella, D. *Quantitative Methods in Derivatives Pricing: An Introduction to Computational Finance*. Wiley, 2002.

[156] Taylor, S. *Asset Price Dynamics, Volatility and Prediction*. Princeton University Press, 2005. 27

[157] Tesfatsion, L. Agent-based computational economics: Growing economies from the bottom up. *Artificial Life 8*, 1 (March 2002), 55–82. 23, 24

[158] Tesfatsion, L. Agent-based computational economics: Modeling economies as complex adaptive systems. *Information Sciences 149* (2003), 263–269. 23

[159] Tesfatsion, L. Agent-based computational modeling and macroeconomics. Staff General Research Papers 12402, Iowa State University, Department of Economics, April 2006. 22, 23, 26

[160] Tkatch, I., and Alam, Z. S. Strategic order splitting in automated markets. *SSRN* (2009). http://ssrn.com/abstract=1400307 or http://dx.doi.org/10.2139/ssrn.1400307. 34

[161] Tsang, E. Computational intelligence determines effective rationality. *International Journal of Automation and Computing 5*, 1 (2008), 63–66. 25

[162] Tsay, R. S. *Analysis of Financial Time Series*, 1 ed. Wiley-Interscience, October 2002. 27

[163] Tversky, A., and Kahneman, D. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology 5* (1973), 207–232. 24

[164] Tversky, A., and Kahneman, D. Judgment under uncertainty: Heuristics and biases. *Science 185* (1974), 1124–1131. 24

[165] Vapnyarskii, I. Lagrange multipliers. In *Encyclopaedia of Mathematics* (1994), M. Hazewinkel, Ed., Springer.

[166] Veryzhenko, I. *A reexamination of modern nance issues using Artificial Market Frameworks.* PhD thesis, Universite Paris I - Pantheone Sorbonne, 2012. 2, 31, 34

[167] Volterra, V. *Lecons sur la theorie mathematique de la lutte pour la vie.* Gauthier-Villars, 1931. 8, 38

[168] Walras, L. *Elements of Pure Economics: Or the Theory of Social Wealth.* Elements of Pure Economics, Or the Theory of Social Wealth. reprinted by Routledge 2003, 1874. 24

[169] Wei, W. C., Gerace, D., and Frino, A. Informed trading, flow toxicity and the impact on intraday trading factors. *Australasian Accounting Business and Finance Journal 7*, 2 (2013), 3–24. 19

[170] Windrum, P., Fagiolo, G., and Moneta, A. Empirical validation of agent-based models: Alternatives and prospects. *Journal of Artificial Societies and Social Simulation 10*, 2 (2007), 8. 25

[171] Wooldridge, M., and Jennings, N. R. Pitfalls of agent-oriented development. In *Proceedings of the second international conference on Autonomous agents* (New York, NY, USA, 1998), AGENTS '98, ACM, pp. 385–391. 2, 21

[172] Yamamoto, R., and LeBaron, B. Order-splitting and long-memory in an order-driven market. *The European Physical Journal B - Condensed Matter and Complex Systems 73*, 1 (January 2010), 51–57. 26

[173] Zimmermann, G., Neuneier, R., and Grothmann, R. Multi-agent fx-market modeling based on cognitive systems. In *Artificial Neural Networks*

ICANN 2001, G. Dorffner, H. Bischof, and K. Hornik, Eds., vol. 2130 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2001, pp. 767–774. 25