A reexamination of modern finance issues using Artificial Market Frameworks

Iryna Veryzhenko

UNIVERSITÉ PARIS I - PANTHEON SORBONNE
Institut d'Administration des Entreprises de Paris
Ecole Doctorale "Sciences du Management" - ED 533
Equipe de Recherche GREGOR - EA 2474

P H D T H E S I S

to obtain the title of

Degree of Doctor of Philosophy

Specialty : FINANCE

Publicly defended on September 18, 2012
by

Iryna VERYZHENKO

A reexamination of modern finance issues using Artificial Market Frameworks

Jury :

Referees:	Fabrice RIVA	Professor, Université Lille 1
	Ollivier TARAMASCO	Professor, ENSIMAG Grenoble
Examiners:	Roland GILLET	Professor, Université Paris 1 - Panthéon Sorbonne
	Juan PAVÓN	Professor, Universidad Complutense Madrid (Spain)
	Paolo PELLIZZARI	Professor, Università Ca' Foscari di Venezia (Italy)
Supervisors:	Olivier BRANDOUY	Professor, IAE - Université Paris 1 - Panthéon Sorbonne
	Philippe MATHIEU	Professor, Université Lille 1

Acknowledgments

This dissertation would not have been completed without the participation and support of many people.

First of foremost, I would like to express my gratitude to my supervisors Olivier Brandouy and Philippe Mathieu. I am grateful for initialing the research proposal, for making me come to France, for hiring me as a Ph.D. student and trusting me with it. I thank my dear professors for their ideas, critics and questions challenged me to make good progress. I am thankful them for being beside me in any joyful as well as difficult moments of this dissertation. I am greatly indebted to them for reading repeatedly and thoughtfully my thesis, and for giving many constructive comments on it.

I was lucky to join the SMAC team of University Lille 1, where I have spent three years and where I have met great researchers and people. A special thanks goes to Antoine, Fabien, Tony, Maxime, Yann, Yoann. Our discussions, jokes and "repas d'équipe" made my studies more varied.

I would like to render thanks to the doctoral committee for offering their time to advice and guide me in this work. I am especially thankful to Fabrice Riva and Ollivier Taramasco for their valuable suggestions and criticisms during a preliminary thesis defense.

My loving spouse, Cyril, is the one who thoughtfully followed every single step of my Ph.D. thesis over these years. Dear Cyril, I thank you very much for your patience, for listening to me, and for encouraging me endlessly when I was distressed, for making me forget any problem and motivating me to go on. Cyril has also greatly contributed to this thesis by translating some part in French. Finally, I would like to thank my parents Olga and Boris for empowering me with an early sense of independence and responsibility; for supporting me from a great distance away; for being proud of my success.

Iryna Veryzhenko

Abstract

Agent-based modeling (ABM) is widely used to study economic systems under a complex paradigm framework. Within this research stream, financial markets have received a lot of interest from academics and practitioners these last years, notably in offering an alternative to mathematical finance and financial econometrics. The traditional approach to analyzing such systems uses analytical models. The latter make simplifying assumptions, for example about perfect rationality homogeneity of market participants. These limitations motivate the use of alternative tools. Thus, disciplines such as Computational Economics and Computational Finance have gained attention and earned their place in the scientific arena.

In this thesis we present an artificial stock market, called ATOM, and contribute to the understanding of some important issues regarding the construction of an abstract model of stock markets as well as a series of technical issues. In ATOM, we model a wide variety of trading strategies and market rules, that allows us to reexamine several traditional questions in finance within a totally different framework.

Firstly, we investigate different conditions under which the statistical properties of an artificial stock market resemble those of real financial markets. To the best of our knowledge, this research is the first to clearly reproduce set of price dynamics at different granularities (intraday and extraday over several simulated years). We argue that generating realistic financial dynamics that reproduce quantitative financial distribution is out-of-reach within the pure zero-intelligent traders framework.

Secondly, we increase agents' intelligence to address the problem of portfolio optimization at the level of individual strategies. We show that the higher relative risk aversion helps the agents earn higher Sharpe ratio and final wealth in the long-range. We also investigate the relation between rebalancing frequency and portfolio performance in low- and high-volatility market regimes with different transaction costs.

Thirdly, we renew the analysis of classical questions in finance, namely, the relative performance of various investment strategies. For that purpose we compare rational mean-variance portfolio optimization versus "naive diversification". We test the investors' performance, each of them following a specific strategy, scrutinizing their behavior in ecological competitions where populations of artificial investors coevolve. Some investment strategies, followed by artificial traders, are based on different variations of canonical modern Markowitz portfolio theory, others on the "Naive" diversification principles, and others on combinations of sophisticated rational and naive strategies.

Finally, we develop a new method for the determination of the upper-bound in terms of maximum profit for any investment strategy applied in a given time window. We first describe this problem using a linear programming framework. Thereafter, we propose to embed this question in a graph theory framework as an optimal path problem in an oriented, weighted, bipartite network or in a weighted directed acyclic graph.

Résumé

La modélisation multi-agent est à l'heure actuelle largement utilisée pour étudier les systèmes économiques complexes. Au sein de ce courant de recherche, les marchés financiers ont été étudiés avec intérêt ces dernières années aussi bien par les universitaires que les professionnels, principalement en apportant une alternative aux approches classiques: mathématique financière et économétrie. La méthode traditionnelle pour analyser un tel système se base sur l'utilisation de modèles analytiques. Cette dernière part d'hypothèses simplifiées du type comportement rationnel des traders ou homogénéité des investisseurs. Ces hypothèses peuvent maintenant être approfondies, ce quimotive l'intérêt pour des solutions alternatives. Ainsi, des disciplines telles que l'économie computationnelle et la finance computationnelle ont attiré l'attention des chercheurs et acquis leur légitimité auprès du milieu scientifique.

Les économistes constatent aujourd'hui une réelle mutation dans la simulation, passant des théories néo-classiques basées sur la résolution d'équations à des modèles alternatifs, approchant la finance par l'expérience, les facteurs comportementaux, ou le calcul. Les hypothèses traditionnelles néo-classiques sont basées sur des investisseurs rationnels et sur des marchés efficients introduites par Fama (1965), Fama et Blume (1966), Fama (1970). Fama a notamment démontré que si un marché est peuplé de nombreux investisseurs rationnels alors les investissements seront au juste prix et reflèteront les informations disponibles. Une hypothèse-clef de ce travail réside dans le fait que l'information est à la disposition de tous les acteurs du marché. Dans les modèles traditionnels, les investisseurs sont rationnels et font usage efficacement des informations; leurs prises de décision sont basées sur des fonctions d'utilité avec espérances, calculées via des procédures statistiques optimales. Ainsi, un tel investisseur représentatif détient toutes les connaissances les plus exactes et les exploite au maximum. Ces hypothèses ont inspiré des milliers d'études visant à déterminer si des marchés spécifiques sont "efficients" et si les acteurs du marché sont parfaitement rationnels. Pompian (2006) soutient que les économistes apprécient utiliser le concept d'acteurs boursiers rationnels, puisque cela leur permet aisément de réaliser une analyse économique et de présenter des résultats de façon élégante.

La théorie traditionnelle se fonde sur la rationalité, qui est un des postulats majeurs de l'Hypothèse d'Efficience des Marchés financiers (HEM). Selon le courant traditionnel, les investisseurs sont tous efficaces, et par conséquent, quand une nouvelle information devient disponible, ils ont la capacité cognitive

de l'interpréter correctement et d'y réagir avec justesse. Cependant, quelques-uns considèrent que ces théories classiques ne correspondent pas à la réalité. Certains chercheurs argumentent sur le fait que les investisseurs ne sont pas vraiment homogènes et pleinement informés. Ainsi, un large nombre d'écarts par rapport à cette rationalité espérée a été reconnu dans la littérature. Par exemple, des investisseurs irrationnels créent des écarts dans le prix des actifs par rapport à la valeur fondamentale, et les investisseurs rationnels devraient être capables de corriger ces mauvaises évaluations des prix à travers un processus d'arbitrage. Néanmoins, les stratégies d'arbitrage dans de véritables marchés financiers peuvent entraîner des coûts, des risques, ou des contraintes diverses, si bien que l'inefficience peut persister pendant une longue période de temps (Barberis et Thaler, 2003). La finance comportementale soutient que les traders sont hétérogènes et ont une rationalité limitée. La psychologie de l'investisseur y joue un rôle clé. Par exemple, un investisseur imparfaitement rationnel n'a pas une aversion uniforme au risque. Dans certaines circonstances, il peut vouloir rechercher volontairement davantage de risque. La littérature identifie plusieurs caractéristiques du trading psychologique tels que la rétroaction positive ou momentum trading, l'extrapolation de la tendance, noise trading, l'excès de confiance, la réaction excessive, la stratégie contradictoire (Cutler, Poterba et Summers, 1989; DeLong, Shleifer, Summers et Waldmann, 1990; Shleifer 2000; Barberis et Thaler, 2003).

Les discussions sur les hypothèses irréalistes utilisées par la recherche classique, telles que la capacité cognitive parfaite d'un investisseur rationnel amènent au souhait d'individualiser les décisions et donc, d'utiliser l'outil informatique pour étudier plus finement ces problèmes de finance. Il est aujourd'hui possible d'utiliser différentes techniques et outils comme la vie artificielle, l'informatique moléculaire, l'intelligence collaborative, les réseaux de neurones, la modélisation par agents, ou autres domaines de l'intelligence artificielle. L'approche multi-agents, et, subsidiairement, la microsimulation, se placent dans le cadre de la finance quantitative et mettent l'accent sur le besoin de représenter les traders non pas en tant que groupes, mais en tant qu'agents individuels hétérogènes. Ces approches visent à modéliser les marchés financiers dans un système évolutif concurrentieldans lequel les agents autonomes interagissent et développent des dynamiques d'apprentissage. Les applications de la modélisation multi-agents permettent le développement de meilleures explications des faits économiques observés. Les modèles multi-agents offrent ainsi l'opportunité de réaliser des expériences donnant la possibilité de produire un nombre important de simulations à partir d'un point de départ. Ils offrent aussi la possibilité d'incorporer des aspects comportemen-

taux du trading pour par exemple étudier les différents types de comportements possibles, leur influence sur les dynamiques de marché, les conséquences sur la mécanique des prix ou l'influence de la microstructure du marché sur les propriétés statistiques de rendement.

Dans cette thèse, nous présentons un marché financier artificiel développé à Lille1, nommé ATOM, offrant une grande souplesse dans l'individualisation des procédures, aussi bien au niveau de la microstructure qu'au niveau des comportements. Celui-ci contribue à la compréhension de certaines problématiques importantes concernant la construction d'un modèle abstrait de marchés boursiers ainsi que d'une série de questions techniques. Dans ATOM, il est possible de modéliser une large variété de stratégies de trading et de règles du marché, ce qui permet alors de réexaminer finement différentes questions traditionnelles en finance.

C'est le cas notamment des faits stylisés, les propriétés statistiques des dynamiques de prix, partagées par tous les titres quel que soit le marché où ils sont échangés. Ces faits stylisés sont généralement formulés en termes de propriétés qualitatives des rendements des titres et ne peuvent pas être assez précisément expliqués par des modèles analytiques. Aucune théorie n'explique de manière satisfaisante l'origine de ces phénomènes: il existe seulement des modèles capables de les reproduire ou de les identifier. Les principaux faits stylisés sont les suivants: l'absence d'autocorrélations sur les rendements sauf sur de très petites échelles intra-journalières, la distribution non-gaussienne mais leptokurtique des rendements des prix, la tendance vers la normalité avec une échelle de temps croissante, le clustering de volatilités des rendements, la corrélation entre volume et volatilité (Levy, Levy et Solomon 2000; Cont, 2001).

Savoir identifier l'origine de ces propriétés statistiques est donc un enjeu majeur, tant pour les théoriciens que pour les praticiens. L'étude des faits stylisés peut permettre l'élaboration de modèles mathématiques plus fiables sur la fluctuation des cours de prix, ce qui pourrait permettre de maximiser l'espérance des gains à long terme d'un portefeuille d'actions, et donc contribuer à une meilleure gestion de portefeuille.

L'existence de faits stylisés remet d'ailleurs en cause certaines théories fondatrices de la finance moderne. Par exemple, le fait que les valeurs absolues des rendements soient auto-corrélées remet en cause la théorie de la stochasticité des prix développée par Osborne (1959) et Samuelson (1965). Cette théorie proclame que si l'on considère un intervalle de temps large, les cours produits pendant ce temps par le marché sont indépendamment et identiquement distribués et que leurs distributions limites convergent vers celle d'une

variable aléatoire gaussienne ce qui implique que les cours de bourse suivent une marche aléatoire. Mais s'il existe des régularités dans les séries de prix (comme le montre l'analyse empirique), il est théoriquement possible de pouvoir obtenir des performances supérieures à la moyenne sans tenir compte de la valeur fondamentale d'un titre, sous réserve que les coûts de transaction n'annulent pas ce profit. Les marchés ne seraient donc pas totalement efficients.

Nous proposons dans cette thèse un modèle de marché adapté à la simulation des dynamiques de prix à l'intérieur d'une journée de cotation. Ce modèle est basé sur un carnet d'ordres à travers lequel les agents échangent des actions de manière asynchrone. Nous montrons que, sans émettre d'hypothèses particulières sur le comportement des agents, ce modèle exhibe de nombreuses propriétés statistiques qualitatives des marchés réels, mais par contre il génère des propriétés quantitatives éloignées de celles observées sur les marchés réels. Notre approche multi-agents essaye de mettre en lumière cette question en interrogeant les effets des comportements des investisseurs sur le prix des marchés. ATOM, par sa puissance et sa souplesse, nous autorise à analyser différentes dynamiques de prix émergents de différentes calibrations de microstructure du marché à l'aide de plusieurs comportements d'agents différents. Nous proposons ici une calibration minimale permettant de faire apparaitre les faits stylisés qualitatifs et quantitatifs. Cette calibration se base sur le flux des ordres de marché réel et le prix fixé: la distribution de prix, l'écart de prix des ordres envoyés par des acheteurs et des vendeurs, la distribution de volume des ordres et la part des ordres de type market, limit, cancel.

A l'aide du même outil, nous avons exploré les questions de la Théorie Moderne du Portefeuille (TMP). Nous contribuons à ces progrès à travers deux axes de recherche: i) nous proposons un modèle multi-agent capable d'aborder les questions d'optimisation du portefeuille ii) nous vérifions que les résultats théoriques sont confirmés dans le cadre des agents hétérogènes avec un mode de raisonnement dégradé. Nous abordons également des questions difficilement résolvables par l'analyse théorique ou empirique comme par exemple l'identification de la stratégie dominante.

Notre recherche se concentre ensuite sur la rationalité dans le corpus de la théorie moderne du portefeuille. Dans ce cadre, l'hypothèse de rationalité des investisseurs signifie que ceux-ci reçoivent et interprètent correctement toutes les informations pertinentes et qu'ils les utilisent pour faire les meilleurs choix. Les violations de cette hypothèse sont assez communes et découlent des éléments comportementaux. Elles donnent lieu à la rationalité limitée

et à de l'hétérogénéité dans la modélisation de la prise de décision des investisseurs. Souvent, les investisseurs, ayant des connaissances limitées sur les rendements espérés et les covariances ont tendance à tout simplement partager leur richesse uniformément sur l'univers des investissements, par exemple en répartissant des montants égaux aux N actifs financiers disponibles sur le marché. Par ailleurs, une étude récente menée par DeMiguel, Garlappi et Uppal (2009) montre qu'une telle stratégie naïve peut surpasser les modèles les plus complexes. Cela soulève donc clairement le débat de l'utilité pratique des modèles sophistiqués théoriques et d'étudier finement les travaux de chercheurs comme Kritzman, Page et Turkington (2010) ou Tu et Zhou (2011) qui proposent une série d'études empiriques en matière de défense de l'optimisation.

A notre avis, le principal problème avec toutes ces études précédentes réside dans leurs méthodologies de backtesting. Cette approche permet d'évaluer une stratégie d'investissement avec les données historiques, comme si leur mise en øeuvre n'aurait pas modifié ces prix. Cette hypothèse est en contraste avec l'analyse de Levy, Levy et Solomon (1995), Hommes (2006) qui montrent clairement que les prix peuvent bien être influencés par plusieurs paramètres (stratégies d'investissement, les compétences cognitives des investisseurs ou la microstructure du marché) qui sont négligés dans le cadre de la technique de backtesting. En outre, les coûts de transaction de pondération du portefeuille ne sont pas inclus dans les résultats. A l'aide de notre approche à base d'agents, nous testons la performance des investisseurs en scrutant leurs comportements dans les compétitions écologiques où les populations des investisseurs artificiels co-évoluent. Certaines stratégies d'investissement sont basées sur différentes variantes de la théorie moderne du portefeuille de Markowitz, d'autres sur les principes de diversification naïve, et d'autres encore sur des combinaisons de stratégies sophistiquées rationnels et naïves proposées dans Tu et al. (2011). Par ailleurs, nous réalisons un examen de l'influence de la fréquence de trading et de l'attitude des investisseurs à l'égard des risques sur la performance du portefeuille.

Dans ce travail nous abordons la question du rapport entre les préférences des agents au risque et leurs performances. Nous comparons la performance relative des stratégies d'investissement en fonction de leurs préférences pour le risque en utilisant une forme de compétition écologique entre les agents.

L'autre facteur important qui influe sur la performance est la fréquence de réajustement du portefeuille. En raison des coûts de transaction, un ajustement fréquent réduit la performance du portefeuille. Une faible fréquence d'ajustement cache un risque de ne pas réagir à temps aux changements im-

portants du marché. Une fréquence optimale permet non seulement de contrôler le risque, mais aussi d'améliorer le rendement du portefeuille. Nous proposons ici un modèle multi-agent permettant de mettre en évidence une fréquence d'ajustement optimale dans des conditions de marché différentes en présence de coûts de transaction.

La question de la performance en finance de marché est une question complexe qui encadre une série de questions méthodologiques. La démarche d'évaluation ex-post ne considère que les résultats statistiques d'une stratégie d'investissement au fil du temps, une fois les mouvements de prix parfaitement connus. Cette approche est largement utilisée dans la gestion des portefeuilles. La méthode dite de backtesting consiste à utiliser une stratégie sur la base des données historiques pour voir/étudier/simuler ce qu'aurait donnée une telle stratégie. Les résultats ne signifient pas que n'importe qui aurait pu atteindre les résultats indiqués. Fondamentalement, ces méthodes de backtesting sont utilisées pour tester la validité d'une stratégie spécifique sur une base théorique. La performance est généralement évaluée à l'aide d'une comparaison relative entre les fonds, car il est impossible de savoir ce qu'aurait été le meilleur comportement pendant la période pertinente. Nous proposons donc une nouvelle méthode pour la construction d'une stratégie capable d'apporter le gain maximum. Nous décrivons d'abord ce problème en utilisant un cadre de programmation linéaire. Ensuite, nous proposons d'intégrer cette question dans le cadre de la théorie des graphes en codant ce problème sous forme d'une recherche de de chemin optimal dans un graphe acyclique, orienté et pondéré.

Contents

General Introduction of the Thesis

Contents

Financial economists are in the midst of a debate about a paradigm shift, from the neoclassical equation-based one to alternative paradigms, such as the experimental, the behavioral, or the computational.
Many standard economic models are based on rational investors and on market efficiency assumptions. The Efficient Market Hypothesis (EMH) is introduced by Fama (1965). Fama demonstrates that if a securities market is populated by many well-informed rational investors, investments will be appropriately priced and will reflect all available information. A key assumption is that the information is publicly available to all market participants. In traditional models, rational investors make efficient use of this information; their decision making is based on utility functions with beliefs, calculated via optimal statistical procedures. Thus, such representative investor holds correct beliefs and is an expected utility maximizer. These assumptions inspire thousands of studies attempting to determine whether specific markets are in fact "efficient" and market participants are perfectly rational. Pompian (2006) argues that economists like to use the concept of perfect rational market participants, as it allows to make economic analysis relatively simple and to present results in more elegant way.

Rationality of market participants is one of the major assumptions behind the EMH. According to traditional stream investors are all rational, and

therefore, when new information becomes available, they all come to the same conclusion on what the fair price should be to reflect new information.

However, Levy, Levy and Solomon (2000), Farmer (2002), Hommes (2006) claim that several assumptions made by theories do not correspond to reality. Some question whether investors are homogeneous and fully informed. How does the market efficiency emerge through the interaction of investors? Thus, a large number of possible deviations from rationality is recognized in the literature. For example, irrational investors create deviation in asset prices from fundamental value, and rational investors should be able to correct this mispricing through the process of arbitrage. Nevertheless, arbitrage strategies in real financial markets can involve cost, risk, or various constraints, so that the inefficiencies may persist for a long period of time Barberis and Thaler (2003). *Behavioral finance* claims that traders are heterogeneous and bounded rational. Investor psychology plays a key role in behavioral finance. For example, imperfectly rational investors are not uniformly averse to risk. In some circumstances, they act as risk seekers. The literature identifies several features of psychology-based trading such as positive feedback or momentum trading, trend extrapolation, noise trading, overconfidence, overreaction, optimistic or pessimistic traders, overshooting, contrarian strategies (Cutler, Poterba and Summers, 1989; DeLong, Shleifer, Summers and Waldmann, 1990; Shleifer, 2000; Barberis and Thaler, 2003).

The desire to build financial theories based on more realistic assumptions lead to the application of computational approaches[1], that allow traders' heterogeneity, bounded-rationality and market non-equilibrium dynamics, to financial problems. New so-called *computational paradigm* bridges the gap between a human and computer systems. These disciplines use different computational techniques, such as artificial life, fuzzy logic, molecular computing, collaborative intelligence, neural networks, instant-based techniques, agent-based mod-

[1]It gives rise to new research field, Computational Finance

eling Chen, Wang and Kuo (2007), Llacay and Peffer (2010) and other areas of artificial intelligence to solve complex financial problems.

Agent-based Computational Economics (ACE) Tefatsion and Judd (2006), and, alternatively, *Microscopic Simulation (MS)* Levy *et al.* (2000) as the branches of Computational Finance emphasize the need to represent traders as heterogeneous individuals. These approaches attempt to model financial markets as evolving systems of competing, autonomous interacting agents and emphasize their learning dynamics. Agent-Based Modeling (ABM) applications are focused on the development of better explanations of observed economic facts. Agent-based models offer the opportunity to perform experiments which would not be possible in any other way since they provide the possibility to produce a large number of simulations from the same starting point. Agent-based models also offer the possibility to incorporate behavioral aspects of trading and to learn the behavioral effects on market dynamics. These models are usually built for the purpose of studying agent's behavior, price discovery mechanism, the influence of market microstructure on statistical properties of returns. Agent-based approach is strongly related to behavioral finance since the agents are bounded rational and can follow simple rules of thumb. This is a key characteristic of any behavioral model, and agent-based models have this characteristic. It is important to note that agent-based technologies are well suited for testing behavioral theories. The connections between agent-based approaches and behavioral approaches will probably become more intertwined as both fields progress.

Background

This section is devoted to the definition of the following terms that are actively used for this research: microscopic simulations, agent-based modeling, multi-agent systems, artificial stock markets. We discuss what they actually cover

and propose a brief overview of the central ideas behind these wordings.

Microsimulation (MS)is a category of computerized analytical tools that perform highly detailed analysis of system dynamics. This approach has been developed in the physical sciences as a tool for studying complex systems with many elements, which are generally intractable by analytical methods. The idea of microsimulations is to study complex systems by representing each of the microscopic elements individually and by simulating the behavior of the entire system, keeping track of all the individual elements and their interactions over time. With MS, complex dynamics are studied from the bottom-up. MS models are free of some modeling constrains and they allow to explore the effect of various parameters on the system. Agent-based simulations apply exactly the same paradigm as microsimulations.

Agent-based models contain multiple interacting agents within an environment (that may be a topological one or simply a framework allowing them to interact, like a market). An agent is a microscopic element of the model. A representation of an agent varies from a simple equation to complex software components with human-like artificial-intelligence, or even humans. An agent is capable of showing some degree of autonomy, communication with other agents, goal-directed learning, and adaptation to environmental changes.

Agent-based models in finance often refer to *Artificial Stock Markets (ASM)*. Artificial stock markets (or stock market simulations) represent a program or application geared at reproducing or duplicating some or all the features characterizing a real stock market (price formation mechanism, representation of market participants). The key property is that in ASM prices emerge internally as a result of trading interactions of market participants, i.e., the agents themselves.

Contrary to neoclassical models using a representative agent, agent-based artificial stock markets normally encompass a large variety of agents. Therefore, artificial stock markets are viewed as multi-agent systems. These models

allow for heterogeneous agents (with different attitudes towards risk, and different expectations about the future evolution of prices) to interact. Such heterogeneity can produce different aggregate behaviors for the system. It can result in equilibria, it can produce patterns and cycles, or bubble and crashes as well. As such, macro-level phenomena emerge from the micro-level, e.g. the interaction among agents. All technical issues of agent-based artificial market modeling are introduced in details in chapter 2.

The number of authors in mainstream finance applying microscopic simulations in their studies has significantly increased during the last decades. One of the earliest and prominent use of agent-based artificial stock markets is introduced by Kim and Markowitz (1989). Kim and Markowitz (1989) were interested in explaining the sudden crash of the U.S. stock market on October, 17th, 1987, when the stock market crashed for more than 20%. Since this market event could not be explained by the emergence of significant information, hedging strategies and portfolio insurance have been blamed as the main factors at play. Kim and Markowitz decided to test the destabilizing potential of computer-based dynamic hedging strategies, such as portfolio insurance, via Monte Carlo simulations in a particular *artificial financial market* (Markowitz, 1988). Namely, Kim and Markowitz tried to simulate a market populated by traders holding strategies found in real-life markets, and therefore, gave quite a detailed description of the real activity at the microscopic level.

This work has influenced scientists (Arthur, Holland, LeBaron, Palmer and Tyaler, 1997b; LeBaron, 2001a, 2006; Jacobs, Levy and Markowitz, 2004) use agent-based simulations so to include in their works more realistic assumptions. As a result, agent-based models became more complex, integrating more sophisticated behaviors, more complex information structures and networks in order to enrich artificial markets ontologies regarding their real life counterpart.

Research Objectives

Agent-based models can highlight the key role of several elements impacting market dynamics that are most difficult to incorporate in traditional models. However, this research does not claim that the representative agent approach, nor the analytical models are useless. In contrast, we use such models as the foundation on which multi-agent models are developed. We examine classical questions in finance within totally different frameworks. The main point of this thesis is to show that agent-based modeling is a good complement to traditional approaches developed in finance (for example econometric or statistical models).

We focus in this thesis on various market dynamics, decision-making processes of agent and the performance of their strategies in artificial stock markets. This approach delivers new and rich financial models, notably by introducing properties that would make analytical models intractable (or hardly tractable; e.g. a mix of agent heterogeneity, bounded rationality).

An artificial market can help in understanding the conditions under which the traditional finance approach remains valid, notably when additional complications such as bounded rationality and heterogeneous preferences are included in the landscape. What happens when the majority of investors are rational, but a minority of them make decisions rather randomly? Do the results of analytical models still hold? All these questions are addressed in this thesis.

First, we propose a discussion on how agent-based models have overcome some of the limitations of market microstructure and statistical models (Chapter 1). Based on the literature survey and the presentation of the different questions investigated with existing platforms, we compose our own artificial market framework, *ATOM*. The latter allows us to achieve the development of a new market model, and to investigate dynamics arising from different cali-

brations. We study the design of agent's minimal intelligence that is necessary to generate adequate financial stylized facts (Chapter 3). For that purpose, we propose a very simple model of trading only depending on very limited set of factors. As such, this model establishes the (absence of) relationship between agents' behavior and stylized facts (at a quantitative granularity). We then make this model more constrained by gradually including additional sophisticated mechanisms to make appear quantitative stylized facts.

Using the same kind of tools, we explore (relatively) classical questions in Modern Portfolio Theory (MPT). The latter, introduced decades ago, form the basis for most investment models, even if they constantly evolve to incorporate new advances, notably in portfolio optimization. We try to contribute to these advances along two lines: i) we develop an agent-based model geared at tackling portfolio optimization questions ii) we check whether theoretical results are confirmed in the framework of heterogeneous, bounded-rational, evolving agents; we also investigate the questions that are difficult to tackle through direct theoretical or empirical analysis, e.g. identifying of dominant strategy.

Next, this research focuses on rationality in the corpus of modern portfolio theory. Investor's rationality means that they receive and interpret all relevant information correctly and use them to make optimal choices. Notably, rational investors use unbiased expectations in forming and selecting mean-variance efficient portfolios. Violations of this assumption are quite common and stem from behavioral characteristics. It gives rise to bounded rationality and heterogeneity in modeling investors' decision making processes. Oftentimes, investors, having limited knowledge about expected returns and covariances tend to simply share their endowments evenly over the investment universe, e.g. allocate equal amounts to the n available financial assets. Moreover, a recent study by DeMiguel, Garlappi and Uppal (2009) shows that such a naive strategy can outperform more complex models. This raises the debate

on whether the sophisticated theoretical models are valuable in practice, and if they are of any use for practical portfolio construction as well. However, the debate is still open, and researchers like Kritzman, Page and Turkington (2010) or Tu and Zhou (2011) propose series of empirical studies in defense of optimization.

We use agent-based tools to shed some light on this debate. In our opinion, the main problem with all these studies, is the unrealistic "atomistic" assumption that underlies the backtesting methodology. Said simply, this assumption allows to gauge an investment strategy with historical data as if its true implementation would have not modified these prices. These assumptions are in sharp contrast to analysis of Levy, Levy and Solomon (1995), Hommes (2006) who clearly show that prices may well be influenced by several parameters (investment strategies, the cognitive skills of investors or the market microstructure itself) that are neglected in the backtesting approach. Moreover, transaction costs incurred by portfolio rebalancing are traditionally not included in the performance results but reported separately. We rely on ecological competition, where populations of artificial investors co-evolve, to test the strategies' performance. Some investment strategies are based on different variations of canonical modern Markowitz portfolio theory, others on the Naive diversification principles, and others on combinations of sophisticated rational and naive strategies proposed in Tu and Zhou (2011). Furthermore, a closer re-examination of the rebalancing effects and investor's attitude toward risk on portfolio performance is needed so to identify clearly what matters the most.

We address the question whether investors' survivability in a long run depends on their risk preferences. Agent-based tools help us take into account all these assumptions and to shed some new light on the relationship between the investors' individual risk preferences and their portfolios' final performance. Computational simulation tools allow us to demonstrate not only a single

value of measures of portfolio performance, but also to trace their evolution in the long-run. We compare the relative performance of investment strategies differ in their risk preferences using Ecological Competition.
The other important factor affecting strategy performance is rebalancing frequency. While a high rebalancing frequency reduces the portfolio performance due to transaction costs, a low rebalancing frequency hides a risk not to react in time to important market changes. Optimal rebalancing frequency helps not only to control the risk, but also to enhance the portfolio return. We construct the agent-based model to study an optimal portfolio rebalancing frequency in different market conditions in the presence of transaction costs.

Performance gauging in finance is a complex issue which unfolds a series of methodological questions. The ex-post evaluation approach considers only the statistical results of a given investment strategy over time, once price dynamics are perfectly known. This approach is widely used in professional asset management. For instance, backtesting involves using a strategy with historical data to determine how trading rules would have performed in the past. The results do not mean that anyone could have achieved the shown results. Basically, backtesting is used to test the validity of specific strategy on a theoretical basis. If the results were valid over long time frames, then they may work as well in the future.

Performance is usually evaluated using a relative comparison among funds, as it is impossible to know what would have been the best behavior during the relevant period, or how the best output compares with the performance upper bound. We develop a new method for *the determination of the upper-bound in terms of maximum profit* for any investment strategy applied in a given time window. We first describe this problem using a linear programming framework. Next, we propose to embed this question in a graph theory framework as an optimal path problem in an oriented, weighted, bipartite network or in a weighted directed acyclic graph.

The main goal of this research is to understand how the stock market operates and behaves, how to invest in the stock market, and to determine the best techniques to use in order to maximize earnings.

Organization of the thesis

The work presented in this thesis does not easily fit into a single research area. It addresses concerns in finance extensively using computer science tools. Being an inter-disciplinary research area, computational finance involves hard work by its very nature. This is because on the one hand researchers must gain sufficient knowledge in computing to know what their potential and limitations are. On the other hand, they also have to know enough about finance to know where computing techniques can be applied. This thesis is organized into two parts, and includes five chapters in total that cover a variety of financial topics investigated by computational tools.

Part I – Context and State of the art

The first part of this thesis introduces the context and the related literature, methodology and approaches used for this research, the important issues of financial market that we address in this thesis. This part is intended for those who are interested in modeling and application of agent-based artificial stock market architecture.

Chapter 1 is devoted to the introduction of agent-based approach as a "bottom-up" representation of stock markets. This introduction provides the *why*, *what*, and *whence*: some motivation for the uses of agent-based tools for financial market questions; an overview of the different problems in finance and methods for meeting them; and an indication of how the approach of multi-agent artificial markets fits into survey methods and into the general quest for scientific knowledge. We review the main advantages of this approach and compare it to the neoclassical framework. Obviously, this chapter also covers

some weaknesses of the agent-based research framework. *Chapter 1* deals with fundamentals, while *Chapter 2* with techniques.

In *Chapter 2* we review the literature on Artificial Stock Markets, what we consider to be the state-of-the-art in the development of agent-based models, but rather from the point of view of computer scientists. We introduce the important questions of market modeling and technical design in software engineering terms. We answer some questions of system calibration and validation. Later, we introduce ATOM, the agent-based computational platform, constructed to perform the experiments proposed afterwards. We detail the potential of our system for modeling a wide variety of trading strategies and market rules, which are crucial for research questions on financial stock market.

Part II – Results and Research Contributions

In this part of the thesis we report our results and contributions in the investigation of important issues for financial markets. This part contains five essays that address different problems in finance, from statistical properties of financial time series to relative performance of investment strategies, but these works all share the agent-based modeling research approach. Each essay is related to a paper that has been published (or will appear) in referred proceedings and collections, and that has been publicly presented at international workshops and conferences in economics and finance. The current state of work is summarized in table 1.

Chapter 3 presents our contributions in the understanding of qualitative and quantitative statistical properties of asset returns, also known as "stylized facts". We propose a minimal agent's intelligence calibration in order to generate realistic market dynamics.

Chapter 4 puts forward the investigation of a portfolio diversification problem within an agent-based computational finance framework. The discussion begins from the mathematical basis of the mean-variance optimization model.

Important facilities of ATOM, previously described in Chapter 4, are explored to implement this classical model within the agent-based framework. We model the agents that maximize their expected utility functions, however, there exists some degree of irrationality. For instance, the investor may deviate to some extent from the optimal rules, or he has limited memory capacities. We also introduce in our model the heterogeneity in expectations, risk preferences, holding periods. All these changes taken separately or simultaneously are impossible to solve analytically, but they can be integrated into agent-based models. We implement the extensive experiment based on the ATOM platform to shed some new light on the relationship between the investors' individual preferences, such as risk aversion and rebalancing frequency, and their portfolios final performance.

In this chapter we also renew the analysis of the relative performance of investment strategies, rational mean-variance portfolio optimization versus naive diversification. We show that the best possible strategy over the long run relies on a mix of Mean-Variance sophisticated optimization and a Naive diversification. These results reinforce the practical interests of the Markowitz framework.

Chapter 5 puts forward a new method for the determination of the *upper-bound* in terms of maximum profit for any investment strategy applied in a given time window. We show that, even in the "ex-post" framework, it is extremely complex to establish this upper bound when transaction costs are implemented. We first describe this problem using a linear programming framework. Next, we propose to embed this question in a graph theory framework and show that the determination of the best investment behavior is equivalent to the identification of an optimal path in an oriented, weighted, bipartite network or in a weighted directed acyclic graph. We illustrate this method using various real world data and make a new point on the notion of absolute optimal behavior in the financial world.

Finally, we summarize the major results and contributions of this work and give an outlook on future directions of research. Additionally, we describe some promising areas of research along the line of our artificial stock market.

Chapter	Publication/ Communication	Conference	Status
Chapter 2: Agent-based artificial stock market	On the Design of Agent-Based Artificial Stock Markets	Communications in Computer and Information Science	Forthcoming
	Key Points for Realistic Agent-Based Financial Market Simulations	International Conference on Agents and Artificial Intelligence (ICAART'2011)	Best student paper award. Referred Proceeding of the third International Conference on Agents and Artificial Intelligence, pages 74-83, 2011
Chapter 3: Minimal market calibration for realistic market simulation	A re-Examination of the "Zero is Enough" hypothesis in the Emergence of Financial Stylized Facts	Journal of Economic Interaction and Coordination	Forthcoming
	Why Zero Intelligence Traders are not smart-enough for Quantitative Finance	The 7th European Social Simulation Association Conference (ESSA'2011)	Publicly Presented, Montpellier France, September 19-23, 2011.
	Agent's minimal intelligence calibration for realistic market dynamics	Artificial Economics (AE'2010)	Referred Progress in Artificial Economics, Lecture Notes in Economic and Mathematical Systems, pages 3-14, 2010
Chapter 4 ABM: Portfolio Performance Gauging and Attitude Towards Risk Revisited			
Section 4.2 Optimal Portfolio Diversification? A multi-agent ecological competition analysis	Optimal Portfolio Diversification? A multi-agents ecological competition analysis	10th International conference on Practical Applications of Agents and Multi-Agents Systems (PAAMS'2012)	Highlihts on Practical Applications of Agents and Multi-Agent Systems, pages 323-331, 2012
		first Ph.D. Camp, ESSEC, Paris	Publicly presented, February 2012
Section 4.3 Impact of Trader's Preferences on Wealth Dynamic	Risk Aversion Impact on Investment Strategy Performance	Artificial Economics (AE'2012)	Forthcoming in Referred Proceeding, 2012
	A multi-agent ecological competition analysis of strategy performance: Does risk aversion matter?	17th Annual Workshop on Economic Heterogeneous Interacting Agents (WEHIA'2012)	Publicly presented, June 21-23, 2012
Chapter 5: Algorithmic determination of the maximum possible earnings for investment strategies	Algorithmic determination of the maximum possible earnings for investment strategies	CNRS 2* Referred International Journal	second round of review
	Gauging Agent-Based Trading of a Single Financial Asset. Definition of an Absolute Distance to the Best Behavior	Artificial Economics (AE'2009)	Lecture Notes in Economics and Mathematical Systems, 2009

Table 1: Thesis Organization and Research Contributions

Part I

Context and State of the Art

Chapter 1

From Traditional to Agent-based Modeling – Justification

Contents

This chapter focuses on agent-based modeling, a computational intensive method for developing and exploring new kinds of economic and finance models. The discussion begins with the representation of financial stock market as the complex evolving system. Then, we introduce an agent-based modeling approach to study complex systems. This chapter also covers the on-going debates and criticisms of agent-based markets. We also suggest some research directions, where agent-based models can be a useful complement to mainstream approaches.

1.1 Financial Stock Market as Complex Evolving System

"I think the next century will be the century of complexity"

Stephen Hawking (Complexity digest 2001/10, 5 March 2001)

This section is devoted to investigating stock markets as complex systems. There is no single precise definition of complex systems. Most authors, however, agree on the essential properties a system has to possess to be called complex. According to Weaver (1948), Simon (1962) the key features of complex system typically include the following aspects: i) the system is composed of interacting *agents* ii) their emergent behavior does *not* result from the existence of a *central controller* iii) the system's properties *emerge from the interaction* of its components iv) the system may show *unpredictable* behavior or lead to uncontrolled explosion (e.g. earthquake or stock market crashes) v) small change in the causes implies dramatic effects. Generally speaking, a complex system of connected agents exhibits an emergent global dynamic, resulting from the interactions between the agents.

Complex systems usually refer to those in the natural sciences, however, Rosser (2004), Arthur (2006) argue that *economy can also be viewed as a complex evolving system.* Particularly, financial markets exhibit properties that characterize complex systems. Traders (banks, brokers, mutual funds, individual investors) and assets (equities, bonds, futures, options, swaps, etc.) can be regarded as the interacting agents, the price and volume market dynamics are emergent phenomena. Thus, financial market might be understood much better as complex adaptive system than optimizing rational entity. The complexity of financial stock market also comes from the personal exchanges where perhaps there are few agents, but they are bounded rational, adaptive, purposeful, and strategic; additionally, they learn from other agents behavior.

Pellizzari (2005) identifies different levels of complexities in the stock market *Molecular, Organizational and Environmental.* The first refers to the complexity of agents: their decision making rules, evolutionary rules, memory span, optimizing and predictive capabilities.

Organizational complexity refers to agents interactions and their organization in groups. Some agents are insensitive to the action of others, the other ones are completely influenced by decisions made by their neighbors. In this case, it might be reasonable to create groups, to share knowledge, information and strategies.

Environmental complexity describes the market organization, a policy-making issue. Bottazzi, Dosi and Rebesco (2005) show the important impact of market organization on return distribution. Such effect can be explained by environmental and molecular level features.

According to McCauley (2004) the complexity of financial stock markets is hidden in the "missing theory of the expected return". The author states that it is an easy task to describe return dynamics by stochastic equation, but there is a big chance that this model is wrong empirically because such models cannot reflect the common sentiment of market participants which can be affected by political announcements and economical changes. The other part of complexity of financial markets is the unfixed empirical distribution. This distribution is influenced by agents' collective behavior. This sort of change cannot be anticipated or described by a simple stochastic theory.

Complex systems can be modeled using a pure *mathematical approach.* This is for example the case with the Lotka-Volterra prey-predator system (Lotka, 1910; Goel, Maitra and Montroll, 1971). However, the models should be as simple as possible for tractability reason. Thus, some of the system properties should be ignored, and only the relevant features that play an essential role in the emerging phenomena explanation, should be retained. With this simplification, a model is a simplified mathematical representation

of a system. Another difficulty of equation-based representation of complex systems is that most equations cannot be solved analytically.

Agent-based modeling (ABM) is a good alternative to equation-based models for studying complex systems, including financial stock markets. Érdi (2008) emphasizes that even if there exists mathematical tools to model and simulate spatiotemporal phenomena, agent-based computational modeling proposes a completely different philosophy and practice compared to equation-based modeling.

Axelrod and Tesfatsion (2005) describes simulation, and ABM in particular, as a third way of doing science, in addition to deduction and induction. *Deduction* is used to derive theorems from assumptions. *Induction* is used to find patterns are empirical data. The simulation also starts with assumptions, but does not prove the theorems with generality; at the same time simulation generates data suitable for analysis by induction. The simulation allows investigation of economical processes under controlled computational experiments and Axelrod and Tesfatsion (2005) points out four specific goals pursued by ABM:

1. Agent-based models can provide *empirical* understanding of macroscopic features nature without top-down control.

2. Agent-based models can provide *normative* investigation, testing the qualities of different designs, looking for one that gives desirable system performance.

3. Agent-based models can provide *heuristic* investigation of market phenomena, understanding of economic system behaviors under alternatively specified initial conditions. ABM sheds some new light on causal mechanisms in social systems.

4. Agent-based modeling can help researchers to get advance in *method-*

ological issues, as it provides the methods and tools needed to undertake the rigorous study of social systems through controlled computational experiments. This axis covers the necessity in testing of experimentally-generated theories against real-world data.

1.1.1 Weaknesses of standard approaches: how ABM can help?

Agent-based approach is an answer to highly centralized, top-down, deductive approach that is characteristics of mainstream, neoclassic economic theory. Most of the time, the neoclassic approach favors models where agents do not vary much in their strategies, beliefs or goals, and where a great effort is devoted to analytic solutions. By contrast, agent-based modeling considers decentralized, dynamic environments with populations of evolving, heterogeneous, bounded rational agents who interact with one another.

Moreover, assumptions made for reasons of tractability in theoretical models may miss many interesting phenomena. The ignored factor may have an important impact on investigated question. Computational modeling allows more complex and realistic assumptions (see table 1.1). Levy *et al.* (1995) illustrate with Sharpe-Lintner Capital Asset Pricing Model (CAPM) (Treynor, 1962) some unrealistic assumptions made for model tractability in order to obtain analytic results. The CAPM deals with rational investors with homogeneous expectations regarding future distribution of returns, they maximize expected utility and have the same holding period. Taxes and transaction costs are not incorporated in this model. Many of the assumptions made in the CAPM, as well as in most other models in finance, have been actively criticized. For instance, the assumption of no taxes and no transaction costs does not correspond to real market conditions. It is also clear that, in contrast to the homogeneous expectation assumption, investors differ in their

expectations, holding periods, decision-making processes, and so on. Stigler (1966) argues that a model with unrealistic assumptions is better than no model at all, and one should not reject a model unless a better one is found. According to Friedman (1953b) the model quality is measured by its explanatory power, and not by its assumptions. Nevertheless, Levy *et al.* (1995) report that several empirical tests, as well as several anomalies of financial stock market, contradict the theoretical results. In agent-based models the unrealistic assumptions can be relaxed one by one by one, this expands the realm of investigations. Epstein and Axtell (1996), Epstein (1999) introduce the key features of agent-based models allowing to relax some assumptions of theoretical models.

- *Heterogeneity.* Agents are not regarded as homogeneous pool of representative agents, every individual is explicitly represented.
- *Autonomy.* There is no central or "top-down" control that defines agents behavior. There exists the strong relation between agents' individual behavior and emergent macro phenomena.
- *Explicit space.* Agents act on the common environment under common restrictions and rules.
- *Local interactions.* Agents are able to interact, communicate, and form the groups with others.
- *Bounded rationality.* Agents have neither complete information nor infinite computational capacity. Agents are not absolute optimizers and they use local information.
- *Non-equilibrium dynamics.* ABM describes models' behavior without assumption about existence of equilibrium. ABM provides not only the information about the existence of equilibrium, but they also allow to access the entire solution trajectories, or how equilibrium was reached.

Neoclassical economics	**ABM**
Fully-informed	Limited access to information
Market participant are rational	Participant has bounded rationality
Participant interact only indirectly through markets	Participant interact directly with one another
Focus on equilibrium outcomes	Focus on dynamics The ability to learn about one's environment from gathered information, past experiences, social mimicry...

Table 1.1: The comparison of mainstream (neoclassical economical theory) and ABM approaches

Some comparison of agent-based models and equation-based or neoclassical models is summarized in the table 1.1 and is detailed in the next subsections.

1.1.1.1 Rationality

Jevons (1871), Menger (1871), Walras (1874) define economics as a problem of allocation of resources between competing forces. Regularities in economies derived from the uniform, simultaneous behavior of individuals optimizing their gains. Such individual, named *Homo economicus*, can be viewed as mathematically represented absolutely rational economic actor. From a neoclassical point of view, an individual tries to maximize his economic well-being and minimizing economic costs, selecting strategies maximizing the utility.

Most criticisms of the "Homo economicus" are based on three underlying assumptions: 1) *Perfect Rationality.* Rationality is not always the first driver of human decision making. As many psychologists believe, the human intellect is subservient to human emotions. 2) *Perfect Self-Interest.* This assumption is strongly connected with the previous one. Sometimes people is subject of impulses and emotions, hence they perform volunteering, helping the needy, even if it contradicts the wealth maximization objectives. However that might well fit with utility function integrating this altruistic dimension. Agents are

then no longer selfish as they are usually described. 3) *Perfect Information.* In the world of investment, there is nearly an infinite amount to know and learn, and even the most successful investors do not master all disciplines. Simon (1957), Kahneman and Tversky (1973), Tversky and Kahneman (1974) emphasize that the individuals are limited in their knowledge about their environment and in their computational abilities. They face costs to obtain sophisticated information that can be processed for rational decision making. Consequently, it is reasonable to describe the market participants as *bounded rational*, instead of perfectly rational entities with fully optimal decision rules. Bounded rational traders have been introduced to replace the standard expected utility theory and to represent human economic decision making in a more realistic manner. Bounded rationality assumes that individual is rational, but limited by general knowledge and cognitive capacity.

In contrast, Friedman (1953a) is one of the strongest defenders of the rational agent approach. He argues that non-rational agents will not survive in evolutionary competition and will therefore run out of the market. Alchian (1950) assumes that biological evolution and natural selection driven by realized profits may eliminate non-rational participants and leads to the situation where rational, profit maximizing firms dominate, and if price contain any predictable components the remaining rational investors will reduce this to zero.

The debates in finance about market efficiency and rationality are still unresolved. Hommes (2006) states that the perfect knowledge about the environment in a heterogeneous world implies that a rational agent has to know the beliefs of all other, non-rational agents. This assumption is highly unrealistic (Hommes, 2001). A bounded rational agent forms expectations based on available information and adapts them to new emergent information. *Rational expectations hypothesis* states that expectational model adopted by all participants leads to behaviors that produce patterns close to the expected ones.

It means that the expectational model produces the outputs consistent with the expectations. But the forecasts that are on average consistent with the outcome they predict do not exist and cannot be statistically deduced. This statement can be illustrated with the El Farol bar problem (Arthur, 1994b,a).

The El Farol bar problem

100 people should decide whether to go to a bar (El Farol bar) or to stay home. If person predicts that there will be more than 60 people in the bar, she will stay home; otherwise she will go there. All agents use the same rational expectations common rules that use available history to produce some prediction. If this weekend the rules predict the attendance higher than 60, everybody will stay home and will negate that forecasts. If rules predict the attendance lower than 60, everybody will show up the bar, and one more time will negate such rules. There is a self reference in such game. In the real world market, this game can be transformed into investor behavior. If there exists the rumors that the price will rise by 2%, it will attract the potential buyers, as a result the price will go down and expectations will fall away. In this case the theory of rational expectations fails. To resolve the anomaly, we should allow agents to start with a variety of rules and expectations, thus some of agents are no longer rational. Perfect rationality is related to heterogeneity issue. The problem of decision making of one single agent can be solved analytically. The same is true if all agents are identical, but when many heterogeneous agents are competing, the decision making process cannot be fully rational due to the complexity of this problem.

1.1.1.2 Heterogeneity and Investors Interactions

In neoclassical economics, market participants are independent decision makers, driven by observed prices and fundamental information. This assumption might appear unrealistic because market traders are not solely influenced by these factors. For instance, the intensive research around market volatility shows that high market volatility does not correspond to the period of high

changes in fundamental information and vice versa (Frankel and Froot, 1986; Cutler *et al.*, 1989). These empirical observations have played an important role in the increasing popularity of *bounded rational, heterogeneous and interacting agents* explanations of asset price movements. This means that agents are also influenced by other market participants. The statement is supported by Schelling (1978), who studies the variety of social phenomena, where the individual decision was determined by the behavior of the others in the group. The author reports that there exists a macrobehavior emerging from the micromotivations of individuals in the group. One of the mechanisms of decision-makers is based on the mimicking others behavior "go with winner". Hence, agents suppose to show some level of intelligence, ability to learn and adapt to environmental changes, to have decision making rules, be able to adjust these rules as a result of interaction and communication with others.

The confirmation of heterogeneity on the stock market can be found in empirical research of Vissing-Jorgensen (2003). The author reports that there exists heterogeneity in forecasting future asset prices on the stock market: 50% of individual investors consider the stock market to be overvalued, 25% believe that it is fairly valued, about 15% are unsure, and less than 10% believe that it is undervalued.

Milgrom and Stokey (1982), Fudenberg and Tirole (1991) reinforce the necessity of heterogeneous expectations, different opinion and trading rules on the market by introducing *the no trade theorems*. But the development of analytical models with heterogeneous agents is quite difficult[1], while agent-based models are more suitable for this purpose, as they easily incorporate a large number of interacting individuals with different rules, access to information and sources of information. Their choices are not deterministic and

[1]For example, Jouini and Napp (2012) conciliate heterogeneous agents level and representative agent in the corpus of mainstream finance.

predictable. Thus, agent-based modeling propose the powerful tool to test alternative models of decision-making which are more in line with the empirical record.

1.1.1.3 Equilibrium

According to Arthur (2006) agent-based computational tools allow to study wider questions compared with standard neoclassical economics methods. For instance, how the economy behaves out of equilibrium or how equilibrium is formed. Rational expectations economics focus on what forecasts are consistent with the outputs created by these expectations. Partial-equilibrium economics focuses on what local behaviors would produce larger patterns that would support those local behaviors. A behavior creates patterns and pattern in turn influences behavior. Neoclassical economics tends to describe the patterns in equation form. Consideration of economic patterns out of equilibrium introduces algorithmic updating and heterogeneity of agents.

In many stochastic, dynamic models, it is possible to characterize the equilibria and stability asymptotically, but little can be said about their out-of-equilibrium behavior. It is important to understand the behavior of out-of-equilibria system since such system may evolve a very long time to reach the asymptotic equilibrium (Axtell, 1999). ABM provides not only the information about the existence of equilibria, but they also allow to access the entire solution trajectories, or how equilibrium was reached. Epstein (2006a) affirms that the purpose of ABM is to provide new evidence on equilibriums and when they exist to generate them, without assuming the existence of these equilibriums.

Moreover, the participant should update their expectations in out-of-equilibrium system that can provoke the expectation changes among other agents. As the results we observe the cascades of changes in the system that

shows as low and high volatility market regimes. This phenomenon shows up in real market price series, but not in equilibrium models.

Market crashes is a good example when equilibrium models, which assume that all forces in the market balance, are ill-suited, and agent-based models can help gain deeper understanding of this phenomenon.

1.1.1.4 Model simplicity

The *modeling* of a system consists in the description of its elements, their behavior under different settings and conditions. The first goal of modeling is to shed some light on the impact of different factors on the behavior of the system. The second goal is to predict the behavior of system under different conditions. Mandelbrot (2006) suggests that the choice and implementation of modeling methodology (linear regression, artificial neural networks, etc.) can play an important role in the quality of the final model.

The modeling in the mainstream framework mandates that solutions of problems be based on theoretically defendable foundations with strong mathematical proofs that imply the series of underlying unrealistic assumptions. Assumptions made for reasons of tractability may miss many interesting phenomena. Epstein (2006b) says that any agent-based computational model can be expressed as an explicit set of mathematical formulas or recursive functions. Many agent-based models have been "mathematized", for example as stochastic dynamical system (Dorofeenko and Shorish, 2002). But even those formulas exist, they are intractable. Hence, the important question which of approaches equation-based or agent-based is most illuminating. So the opinion that agent-based modeling is just simulations for which no equation exist, is actually incorrect. The advantage of agent-based models is that they can be ran thousand times with different parameter values and easily produce the targeted outputs, while it is difficult to do in equation-based models.

1.1.1.5 Controlled replications

In econometrics framework each empirical observation contains some proportion of variation due to some proportion of noise assigned to chance or imperfect observations. In contrast, in agent-based framework researchers have perfect control over stochastic sources of variation. Thus, they have the capability to produce the effects of stochastic variation and simulate exact replicates of empirical samples. Gode and Sunder (1993) argue that agent-based modeling permits greater control over the preferences and information-processing capabilities. It is possible to specify a multi-agent complex adaptive system that generates the empirical phenomenon. These phenomena that emerge from simulations should be the result of multi-agent interactions and adaptation, and not the results of complex assumptions about individual behavior and/or the presence of "too many" free parameters. Thus, the ability to generate a particular empirical phenomenon facilitates understanding of the empirical phenomenon.

1.1.1.6 A mix of approaches

Agent-based modeling that generates a large variety of simulation outputs can be used in conjunction with econometric and statistic tools, which can be useful for risk analysis. Hence, these two streams can be mixed. Computation does not replace theory, but it allows to develop a theory.

Wooldridge (2002) reports the factors indicating the appropriateness of agent-based models. When the considered system is open, highly dynamic, uncertain, or complex, a model of flexible autonomous actors is often the only solution. In some systems, the distribution of data or control, the centralized solution is difficult or even impossible.

1.1.2 Critics of Agent-Based Modeling

Many economists recognize the problems of classical methods in finance, but at the same time they have serious doubts about taking agent-based approaches to address the fundamental issues. The criticisms of computational approaches in finance are summarized in (Tefatsion and Judd, 2006). Next subsections introduce the critics of ABM in finance and some possible methods to avoid these limitations.

1.1.2.1 Examples Vs. Theorems

One of the first criticism of computational methods applied to financial problems is that they produce only examples, while mainstream methods produce theorems. One single example shows only one element in an infinite set of possible outputs, but one single example cannot shed light in explanation of parameters importance in the model. Each run of simulations is a sufficiency theorem, but a single run does not provide any information on the robustness of such theorems. In mathematical economics the model sensitivity can be realized via inspection, simple differentiation, the implicit function theorem. The only way to deal with this problem in agent-based computing is through multiple runs of simulation with different initial conditions and parameters (Axtell, 1999). Hence, a few thousand well chosen examples generated in agent-based framework can be more convincing.

If a theorem is proved in a mathematical framework under the assumption that all agents have the same cognitive abilities or memory span, one can simply introduce the agent population in an agent-based model with some distribution of initial parameters, run the simulation and check whether the simulation still holds. If it does not, then the agent-based model represents a counter-example to the generalization of the theorem with respect to the assumption about agents heterogeneity.

1.1.2.2 Errors of Computations Vs. Errors of Specification

Critics point out that numerical results have errors. But these errors can be anticipated by the application of sophisticated algorithms and powerful hardware. Careful simulation methods can reduce simulation error by increasing the sample size and by exploiting variance reduction methods. More generally, careful numerical work can reduce numerical errors. The problem of numerical errors in agent-based computational models are no more difficult to handle (and more often much easier) than the analogous numerical problems that arise in maximum likelihood estimation and other econometric methods. Researchers face a trade-off between the numerical errors in computational work and the specification errors of analytically tractable models. Therefore, it is often argued that it is better to find a solution with some inaccuracy of correctly defined question, than to find exact solution to the wrongly defined question.

1.1.2.3 Parameters settings

A problem with computational economics and bounded rationality is that it leaves *many degrees of freedom.* Agent-based model should allow agents to evolve, to act and to interact with others overall experiment time without intervention from the modeler. The modeler cannot intervene to adjust system evolution. All initial specifications should be completely predefined, small changes in these specifications can significantly affect the output. The model should have the right parameters for the simulation to make sense. Therefore, sometimes it is difficult to justify the value taken for some parameters.

Another drawback of multi-agent models is that the decision making rules of the agents do not contain semantic specifications, modeler rather adapt ad-hoc functions of the decision making process without underlying cognitive processes (Grothmann, 2002). However, these limitations can be removed by

using feedforward neural networks (Beltratti, Margarita and Terna, 1996) or recurrent neural networks (Yang, 1999).

1.1.2.4 Agent-based models are hard sell

Axelrod (2006) describes his collaboration with Bill Hamilton on agent-based models in biological systems and how their paper "had a hard time getting published". He indicates that the absence of standards for testing the robustness of an agent-based model can make agent-based modeling a hard sell. It is difficult to validate outputs of agent-based models against empirical data. While ABM is useful in producing aggregate-level patterns from individual-level rules, the main issue in financial markets in the agent-based framework is calibration and validation: how can one evaluate the quality of a model from an econometric point of view. In many situations the fitness function of data fit cannot be practically formulated mathematically. Thus, simulations, due to the lack of rigorousness, are still not regarded as a science by many scientists. Robert Axelrod suggests that agent-based community should converge on standard tools for research, on a set of fundamental concepts and results. Without support from theory, the contribution of agent-based modeling can be rather limited. Hence, it is important to consolidate the ABM study with theoretical fundamentals. There exit actually some research tools, like game theory or econophysics, able to bridge a gap between theory and simulations.

Ashburn, Bonabeau and Ecemis (2004) argue that in order to overcome the problem of agent-based model validation, one can allow more subjective factors to guide the search for "good" models, by enabling ABM users to integrate financial economics expertise into their models. Such techniques, used in other fields (such as in the geosciences Boschetti and Moresi (2001)), rely on directed search evolutionary algorithm which requires human input to evaluate the fitness of how well the model reproduces the data qualitative and

uses common evolutionary operators to breed the individual-level rules that produce macro level patterns.

1.2 Survey of agent-based modeling research contributions

This section discusses simple models from an alternative approach in which financial markets are viewed as complex evolutionary systems, it also surveys the main areas in which agent-based models have been used. A range of important economical topics relevant to agent-based modeling is considered by Axelrod and Tesfatsion (2005): price distributions (Bak, Chen, Scheinkman and Woodford, 1993), price equilibrium in decentralized markets (Albin and Foley, 1990; Epstein and Axtell, 1996), trade networks (Tesfatsion, 1995; Epstein and Axtell, 1996), excess volatility in returns to capital (Bullard and Duffy, 1998), organizational behaviors (Prietula, Carley and Gasser, 1998), stock market price time series (Arthur *et al.*, 1997b), shape of the distribution of assets return (Cont, Potters and Bouchaud, 1997; Mantegna, 1991), higher-order statistical properties (Arnéodo, Muzy and Sornette, 1998; Cizeau, Liu, Meyer, Peng and Stanley, 1997). We compare the classical econometric approaches and agent-based modeling applied to financial issues in detail. First, we review the literature on stylized facts explanations. Second, we briefly describe the research on the agent-based explanation of market anomalies. Finally, we present the mainstream models of investment decision making, especially, the series of unrealistic assumptions is stressed. We contrast these models with computational models that help researchers overcome the drawbacks of traditional models.

1.2.1 Stock Market Volatility

Volatility is one of the most important characteristics of financial stock market. It is also an important parameter for portfolio optimization, asset pricing and risk management. Risk is the key problem in finance. Even if there exists different measure of risk, the definition of risk as the variance of logarithmic price series is still more popular. Nevertheless, volatility is not an easily observable parameter, and should be evaluated using different approaches presented in financial literature. Financial literature defines a series of *stylized facts* characterizing volatility.

The result of more than half a century of empirical studies on financial time series indicates the fact that all these series have common properties from a statistical point of view. Such statistical properties are known as stylized facts and have been reported for several types of financial data and for different types of financial markets (Cont, 2001).

There are several approaches to study and understand market dynamics and price series properties. Theoretical studies try to find explanations through analytically tractable models. Empirical studies analyze historical data. Experimental studies focus on analyzing trading behavior and its consequences on the market dynamics. Experimental studies are usually related to behavioral finance. We focus on the power of agent-based models to reproduce stylized facts. Important observed stylized facts in financial time series are as follows:

- Excess volatility. The fact that large (positive or negative) returns occur cannot always be explained by the arrival of new information on the market.

- No autocorrelation in raw asset returns

- Volatility clustering (slow decay of autocorrelations of squared returns

and absolute returns). This fact is often interpreted as a sign of long-range dependence

- Fat tails in the returns distribution. The distributions are approximately bell-shaped but assign more than normal probability with more peaked center (excess kurtosis) and at the extremes (heavy tails).

- Volume/volatility correlation: trading volume is correlated with all measures of volatility.

One of the main stylized facts observed in many financial time series is so called, "volatility clustering", where periods of high volatility are followed periods of low volatility, and so on. Mandelbrot (1963) was one of the first to observe this phenomenon. Applied researchers in finance use Autoregressive Conditional Heteroskedasticity (ARCH) models and its various extensions as an econometric tool, actively used to describe volatility clustering introduced by Engle (1982).

ε_t denotes the discrete time stochastic process $\varepsilon_t = z_t\sigma_t$, where $E(z_t) = 0$, $var(z_t) = 1$. In most applications ε_t corresponds to the innovation in the mean for some other stochastic process $y_t = g(x_{t-1}; b) + \varepsilon_t$ and $g(x_{t-1}; b)$ denotes a function of x_{t-1} and the parameter vector b, with x_{t-1} information set at the moment $t-1$. Engle (1982) suggests to express σ_t^2 as a linear function of past squared values of the process,

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-1}^2 \tag{1.1}$$

The equation 1.1 expresses $ARCH(q)$ model, the first attempt to capture volatility clustering (Bollerslev, Chou and Kroner, 1992). An alternative and more flexible structure is provided by Bollerslev (1986) with the

$GARCH(p,q)$.

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-1}^2 \tag{1.2}$$

It is necessary to note that the small numbers of parameters needed in these models seem sufficient to capture the variance dynamics over very long sample periods. For that reason, $GARCH(1,1)$, $GARCH(1,2)$ or $GARCH(2,1)$ are typically used. In the GARCH(1,1) model the squared volatility depends on last periods volatility.

It is also widely recognized that the return distributions tend to have fatter tails that the normal distribution (Mandelbrot, 1963). The unconditional distribution for ε_t in the $GARCH(p,q)$ have fatter tails than the normal distribution for many financial time series.

Stock returns tend to exhibit nonnormal unconditional sampling distributions, if one considers their skewness and excess kurtosis (Fama, 1965). The conditional normality assumption in ARCH generates some degree of unconditional excess kurtosis, but typically less than adequate to fully account for the fat-tailed properties of the data. Attempts to model the excess conditional kurtosis in stock return indices include the estimates of $EGARCH$ model with a generalized exponential distribution Nelson (1989, 1991). Skewness and kurtosis are important in characterizing the conditional density function of returns.

The flexibility of ARCH-GARCH models allow them to succeed in stylized facts replications. Nevertheless, there are still problems concerning these models. Empirical studies highlight the presence of fat tails and remaining asymmetry in the normalized residuals. Moreover, empirical results also show that variance measured by GARCH models is so large that it can induce explosive conditional variance. To capture this important characteristic Engle and Bollerslev (1986) introduce the Integrated Generalized Autoregressive Conditional Heteroskedasticity model (IGARCH) that tries to specify the second

moment of a time series.

In addition, the time-varying investors behavior may cause the structure of volatility. One of the critics of GARCH models is its impossibility to take into account structural changes or inflexibility of these models (see Lamoureux and Lastrapes (1990)). There are several attempts to overcome these critics. Some attempts to improve $GARCH$ models are inspired by heterogeneous agents literature. For example, Frijns, Lehnert and Zwinkles (2008) present the time variation in the coefficients of standard $GARCH(1,1)$ modeling with switching mechanism, often used in the agent-based modeling, where agents maximize a certain objective function and switch between different trading rules to achieve this (Brock and Hommes, 1998; Franke and Westerhoff, 2009). Such improved models can better capture the kurtosis and skewness observed in stock returns.

While GARCH models can mimic volatility clustering market properties, they provide only theoretical explanation of such phenomenon. Cutler *et al.* (1989) have shown that a substantial fraction of stock market fluctuations cannot be explained by macroeconomic news and a large part of price series properties are not due to major economic news or other events. Thus, a rational representative agent model has difficulty in explaining volatility clustering. Therefore, multi-agent artificial stock markets, allowing agent's heterogeneity, bounded rationality, and realistic microstructure, are actively applied to deal with these phenomena.

Next subsection discusses possible mechanisms proposed in the literature as the possible sources of volatility clustering:

1. Heterogeneous arrival rates of information (Andersen and Bollerslev, 1997b)

2. Evolutionary models. The Santa Fe market model replicates qualitatively some of the stylized facts (LeBaron, Arthur and Parlmer, 1999).

3. Behavioral switching. Lux (2000) constructs the behavioral switching model between fundamentalist and chartist. This model makes arising of the fat-tails and volatility clustering. Kirman and Teyssiere (2002) propose that the proportion of fundamentalist in the market follows a Markov chain. This modified model performs autocorrelation patterns in absolute returns.

An agent-based approach to stylized facts explanation

There is already a large literature on heterogeneous agent models replicating many of the important stylized facts of financial time series on short time scales, such as fat tails or volatility clustering (Brock and LeBaron, 1996; Arthur, Durlauf and Lane, 1997a; LeBaron *et al.*, 1999; Cont and Bouchaud, 2000; Farmer, 2002; Kirman and Teyssiere, 2002; Ladley and Schenk-Hoppe, 2009; Veryzhenko, Brandouy and Mathieu, 2010). Heterogeneity in agent's time scale has been regarded as a possible origin for various stylized facts in Guillaume, Dacorogna, Davé, Muller, Olsen and Pictet (1997), Andersen and Bollerslev (1997a). LeBaron (2001c) reports that the heterogeneity in horizons may lead to volume-volatility relationships similar to those of real market. The recent survey by Lux (2009) contains an extensive survey of behavioral interacting agent models mimicking the stylized facts of asset returns, in particular heavy tails of high-frequency data with convergence to the Normal distribution occurring only at relatively low frequencies. Such models with interacting agents appear to be quite robust generators of the stylized fact. Lux (2009) argues that this power of agent-based models to explain previously unexplained characteristics of financial market provides some credibility to this new approach. We provide a detailed description of the Lux model, as it has been successful in explaining of the series of stylized facts described in the previous subsection.

Lux model

Lux (2000) describes financial a market with a fixed number of *fundamentalist* and *chartists*. Fundamentalists' trading is based on fundamental (true) value. They buy (sell) when the current market price is below (above) the fundamental value. Chartists or technical traders pursue a combination of imitative and trend following strategy. The author presents a possible explanation for volatility clustering in multi-agent framework using a switching principle in strategy choice.

The model contains three key elements: i) switching between pessimistic and optimistic chartists ii) switching between fundamentalists and chartists iii) a price adjustment process based upon aggregate excess demand.

i) The total number of traders is $N = N_c + N_f$, where N_c – number of chartists, N_f – number of fundamentalists. There exists an intergroup separation of chartists: pessimistic and optimistic. Chartists can switch from the pessimistic to the optimistic type. Agents are allowed to switch between a chartist and a fundamentalist strategy after profits comparing; in addition, they can switch from pessimistic to optimistic strategy and vice versa. Interpersonal communications are also allowed.

   The chartists calculate the opinion index, representing the average opinion among non-fundamentalists, as $\phi = \frac{N_o - N_p}{N_c}$, where N_o is number of optimistic and N_p is the number of pessimistic chartists. This opinion index ϕ and price trend $P' = \frac{dP}{dt}$ define the probability of switching from the optimist to the pessimist depends on opinion index ϕ and price trend $P' = \frac{dP}{dt}$. The probability of switching from pessimistic to optimistic is defined as follows $p_{p \to o} = \nu_1 \frac{N_c}{N} e^{-U_1}$, and from optimistic to pessimistic respectively, $p_{o \to p} = \nu_1 \frac{N_c}{N} e^{U_1}$, where ν_1 is the switching frequency, $\frac{N_c}{N}$ is the part of chartist in the total population of agents, $U_1 = \alpha_1 \phi + \alpha_2 \frac{P'}{\nu_1}$, α_1 is the sensitivity parameter to the opinion index, α_2 – sensitivity to

price changes.

ii) Switching between the chartist and the fundamentalist is driven by price changes and current strategy profitability $\frac{(y+\frac{dP}{dt})}{P-r}$ where y are constant nominal dividends of the asset, r denotes return from other investments.

Fundamentalists believe that price will revert back around the fundamental value f, hence they buy (sell) when current market price lower (higher) the fundamental value. Hence fundamentalists' profit can be calculated as $s|\frac{P-f}{P}|$, where s is discounted factor, which reflects that the excess profits are realized only when current price reverts back to its' fundamental value.

$$
\begin{aligned}
p_{o\to f} &= \nu_2 \frac{N_o}{N} e^U_{2,1} & p_{f\to o} &= \nu_2 \frac{N_f}{N} e^{-U_{2,1}} \\
p_{p\to f} &= \nu_2 \frac{N_p}{N} e^U_{2,2} & p_{f\to p} &= \nu_2 \frac{N_f}{N} e^{-U_{2,2}} \\
U_{2,1} &= \alpha_3 (\frac{y+\frac{P'}{\nu_2}}{P} - R - s|\frac{P-f}{P}|) & U_{2,2} &= \alpha_3 (R - \frac{y+\frac{P'}{\nu_2}}{P} - s|\frac{P-f}{P}|)
\end{aligned}
$$

α_3 is the sensitivity of traders to differences in profits, ν_2 – chartist/fundamentalist switching frequency, $\frac{N_f}{N}$ is the part of fundamentalist in the total population and the probability for a fundamentalist to meet a chartist.

iii) Price changes are controlled by a market maker according to the aggregate excess demand of chartists and fundamentalists.

A chartist buys (sells) a fixed amount q_c of assets per period. Excess demand of chartist is $ED_c = (N_o - N_p)q_c = \phi N_c q_c \equiv \phi \frac{N_c}{N} T_c$ and $T_c \equiv N q_c$, where T_c – maximum trading volume of chartists.

Fundamentalists buy (sell) when $P < f$ ($P > f$), their excess demand is $ED_f = N_f \gamma (f - P) = (1 - \frac{N_c}{N}) N \gamma (f - P) = (1 - \frac{N_c}{N}) T_f (f - P)$, $T_f \equiv N\gamma$, where $\gamma > 0$ measures the reaction speed of fundamentalists to price deviation from fundamental value, T_f – trading volume of

fundamentalists.

Market makers estimate the price according to aggregate excess demand by $\frac{dP}{dt} = \beta[ED_c + ED_f] = \beta[\phi\frac{N_c}{N}T_c + (1 - \frac{N_c}{N})T_f(f - P)]$, where β is the speed of adjustments.

There are also some noise traders or liquidity traders in the market with the noise term $\mu \sim N(1, \sigma_\mu)$. The probabilities of increasing ($p^{\uparrow P}$) or decreasing ($p^{\downarrow P}$) are defined by

$$p^{\uparrow P} = min(max(0, \beta(ED + \mu)), 1)$$
$$p^{\downarrow P} = min(-min(0, \beta(ED + \mu)), 1)$$

This model is visualized in figure 1.1.

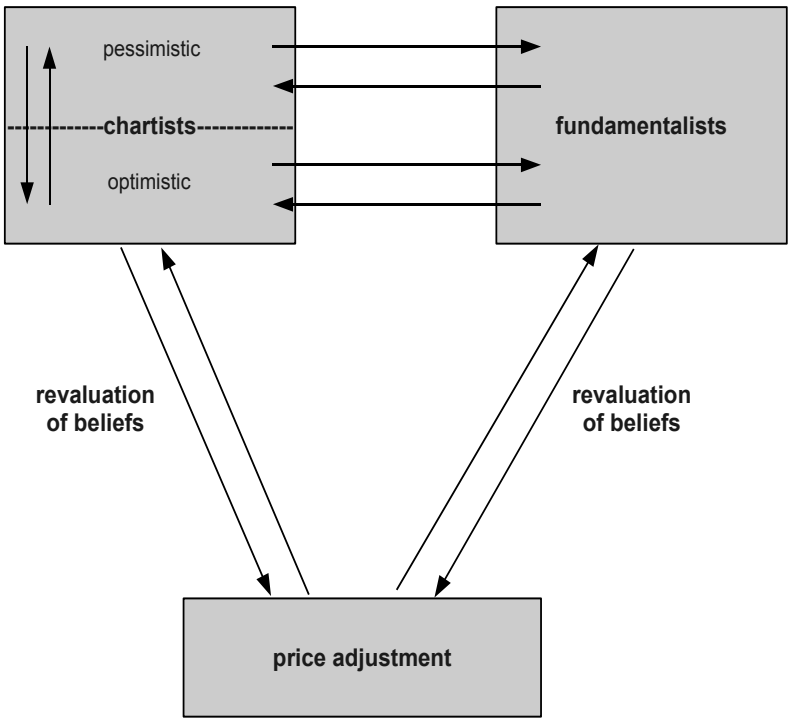


Figure 1.1: Lux-Marchesi model

To summarize, Lux's model generates all of the mentioned stylized facts of financial markets endogenously through the interaction of the agents. The author argues that the source of volatility clustering and leptokurtotic return distributions is the switching between chartist and fundamentalist strategies. Statistical investigation of the simulated time series showed that the main stylized facts can be found in the artificial market modeled by Lux (2000).

Other insights into the phenomena of stock returns are given by other agent-based model of Cont (2007). The author argues that many agent-based models are too complex to establish a simple relationship between the model's parameters and observed stylized facts. He questions whether all the ingredients of the model are indeed required for explaining empirical observations. Therefore, Cont (2007) proposes a simple enough agent-based model capable of generating time series returns with properties close to real data, so the origin of volatility clustering can be traced back to agents behavior. We now discuss a simple model (Cont, 2007) reproducing several stylized empirical facts, where the origin of volatility clustering can be explained by threshold response of investors to news arrivals.

Cont model

Cont (2007) proposes a model, able to explain some statistical properties of financial time series. N agents trade one single asset whose current price is denoted by P_t. Trading takes place at discrete periods $t = 1, 2...n$, where these periods are interpreted as trading days. In each period, agents have the possibility to buy or sell the asset, φ_t is the demand of the agent.

$$\varphi(x) = \begin{cases} 0 & \text{stay unchanged} \\ 1 & \text{buy} \\ -1 & \text{sell} \end{cases}$$

At each time period, agents receive a common signal of public information $\varepsilon_t \sim N(0, D^2)$. Each agent i compares this signal to her threshold θ_t^i and generates an order $\varphi(x)$ according to following rules:

$$\varphi(x) = \begin{cases} 0 & \text{if } \varepsilon_t > \theta_t^i \\ 1 & \text{if } \varepsilon_t < \theta_t^i \\ -1 & \text{if } |\varepsilon_t| \leq \theta_t^i \end{cases}$$

where θ_t^i is considered as the individual agent's subjective view on volatility ($\theta_t^i = |r_{t-1}|$). Excess demand is given by $Z_t = \sum_{i=1}^{N} \varphi_t^i$. It produces a change in the price $r_t = ln\frac{S_t}{S_{t-1}} = g\left(\frac{Z_t}{N}\right)$, where g is the price impact function. λ is the market depth $g'(0) = \frac{1}{\lambda}$. Any agent i has a probability s of updating her threshold θ_t^i. q – the fraction of agents updating their views at any period.

$$\theta_t^i = \begin{cases} |r_t| & \text{if } u_t^i < s \\ \theta_{t-1}^i & \text{if } u_t^i > s \end{cases}$$

where u_t^i is the uniformly distributed variable in $[0, 1]$, that determines whether agent i updates her threshold or not.

Compared to Lux (2000) model, in Cont (2007) model there is no exogenous fundamental value; fundamentalists and chartists are not considered in these simulations; additionally, the communications and interactions between agents are not allowed. The same public information is available to all agents, but they process this information in different ways, this provides the heterogeneity of the model. This simple model with very few parameters generates a time series of returns with properties similar to empirically observed properties of asset returns. The simulation results perform excess

volatility, leptokurtic distribution of returns with heavy tails, excess kurtosis $\simeq 7$, positive autocorrelation function of absolute returns. Cont (2007) argues that investor inertia provides an explanation of switching mechanism proposed in the econometrics literature as an origin of volatility clustering. In case of low volatility, agents become more sensitive to new arrivals, thus, generating higher excess demand and thus, increasing the amplitude of returns, as a result, increasing the volatility. Contrarily, in case of the high volatility agents become less reactive to news arrival, and such increasing agents' inertia provokes the decreasing of return volatility.

In this section we demonstrate that agent-based models, relying on behavioral aspects, can provide a useful complement to econometric analysis. Stylized facts, viewed as puzzles within the standard equilibrium modeling, emerge quite naturally in agent-based models. The behavior of heterogeneous agents, interaction and switching between them may lie at the heart of stylized facts explanation.

1.2.2 Market Anomalies

In this subsection we introduce a contribution of agent-based modeling on technical anomalies explanation. One of the market phenomenon is the profitability of technical trading that reveals inconsistencies with respect to the efficient market hypothesis. This is *technical anomalies*. Common technical analysis strategies are based on the relevant strength of the trend and moving averages. As technical trading techniques are mechanical, whether they can generate significant profit has been a long-debated issue since Fama and Blume (1966). Recent empirical studies find more and more supporting evidences for the profitability of technical analysis, including, among others, Sweeney (1986), Sweeney (1988), Brock, J.Lakonishok and LeBaron (1992), Blume, D.Easley and O'Hara (1994). These results suggest that technical

analysis is popular because it can "beat the market". Others argue that prices adjust rapidly in response to new stock market information and that technical analysis techniques are not likely to provide any advantage to investors who use them.

A wide variety of theoretical and empirical models have been developed to explain why technical trading is widespread in financial markets. Agent-based models provide new explanations of observed market anomalies and deeper insights into the dynamic of real-world financial markets. Joshi and Bedau (2000) propose an agent-based artificial model of a stock market to explore an explanation of this phenomenon. Agent expectations do not follow the fixed rules such as rational expectation rule. They choose the expectations among the evolving set of expectation rules to be the most successful predictors of recent stock-price changes. The authors use Santa Fe Artificial Stock Market LeBaron *et al.* (1999) to show that if the market is populated by fundamentalist agents, some individual using technical analysis for price prediction can take some advantages. As far as the majority of agents undertakes the technical trading rules (because the singular agent's decision is mirrored by other traders), the prediction becomes less accurate due to additional noise in the market provoked by technical traders. It drives the market to a symmetric Nash equilibrium in which the average final wealth of the agents in the market is lower than in the market in which only fundamentalists are trading. Obviously, it makes technical analysis not profitable anymore.

Chen and Yeh (2001a) study an artificial stock market with an evolving agent population. Agents use the learning mechanism, so-called business school, based on genetic programming. The authors show that the price series follows a random walk process in the long run. This finding is rather a confirmation of market efficiency (Fama and Blume, 1966). Additionally, the authors find that some agents can outperform the market in the short-term time scale, which is the evidence of short-term market anomalies.

1.2.3 Investment Decision Making

This subsection examines and confronts traditional models with agent-based modeling in investment topics. We present a brief review of mathematical basis as well as limits of classical models in investment decision making. The portfolio optimization theory provides a way to measure investor's preferences for risk they are willing to undertake in the hope of attaining greater wealth. Utility functions give a way to measure such wealth-risk relationships. The utility theory lies at the heart of Modern Portfolio Theory (MPT). Thus, we begin the discussion by Expected Utility Theory (EUT) von Neuman and Morgenstern (1947) overview. While utility functions are too simple to be directly relevant for real-life applications, they create the foundation for the development of the more complex theories. We introduce here Mean-Variance Portfolio Theory (Markowitz, 1952), Capital Asset Pricing Model (Sharpe, 1964), and Arbitrage Pricing Theory (Ross, 1976).

Expected Utility Theory (EUT). The modern economic theory of decision making under uncertainty is based on the expected utility framework developed by von Neuman and Morgenstern (1947). The authors provide a rational foundation for decision-making under risk according to expected utility properties. This theory was further developed by Samuelson (1950), Marschak (1950), Herstein and Milhor (1953) and others. Utility refers to the consumer's satisfaction from the consumption of goods and services. Utility can be applied to wealth as well as goods and services. Marginal utility function, decreasing with wealth, is actively used. This function reflects the fact that the every individual benefits from an additional unit of wealth, but the utility of this gain is less for someone who already has large wealth. Thus, it is common to maximize the expected utility of wealth rather than the expected wealth. Economic and financial literature introduce quadratic, logarithmic,

power, and negative exponential utility.

The expected utility assumption for modeling choice under risk and uncertainty has been discussed and disputed in (Allais, 1953; Ellsberg, 1961). It gave rise to alternative theories of investment decision making: weighted expected utility (Allais, 1979), rank-dependent expected utility Quiggin (1982), the cumulative prospect theory (Tversky and Kahneman, 1992), non-linear expected utility (Machina, 1982), regret theory (Loomes and Sugden, 1982), non-additive expected utility (Schmeidler, 1989), and state-dependent preferences (Karni, 1985).

Microscopic simulations (MS) (Levy *et al.*, 1995, 2000), alternatively agent-based modeling (ABM) (Tefatsion and Judd, 2006), provide the possibility to implement any form of utility functions. But the traditional utility function framework can be improved by heterogeneous risk preferences or beliefs about expected values.

Mean-Variance Portfolio Optimization Theory. The mean-variance formulation proposed by Markowitz (1952),Markowitz (1959) relies on Expected Utility Theory and provides a fundamental basis for portfolio selection. Significant visibility this theory gets after papers of Tobin (1958), Sharpe (1963), Sharpe (1964), Lintner (1965), and its analytic solution by Merton (1972). The fundamental lesson of the Markowitz analysis is to show that investors must care not only of the realized return, but also of the risk of their positions. Markowitz proposes to measure the risk of return by its standard deviation. Denote by ϖ the vector of weights of the n risky assets, R the vector of returns, R_f – the risk-free rate of return. The percentage of wealth invested inn this riskless asset is w_0. The optimization program is:

$$min\ \varpi' V \varpi \tag{1.3}$$

$$\varpi' R + (1 - \varpi' e) R_f = E(R_P) \tag{1.4}$$

The budget constraint is:

$$\varpi' e + \varpi_0 = 1 \Longleftrightarrow \varpi_0 = 1 - \varpi' e$$

where the variance-covariance matrix V is invertible.

There exist following assumptions in Markowitz Portfolio Theory:

- An investor seeks to maximize his final expected utility of wealth
- A single-period investment horizon
- An investor is risk-averse
- The investor chooses optimal portfolio on the basis of means and variance of assets
- Market are perfect: there are no transaction costs or taxes
- The model does not suit speculative traders
- Expected returns, variances, and covariances are known for all assets. Investors know the future values of these parameters
- Investors create optimal portfolios by relying solely on expected returns, variances, and covariances. No other distributional parameter is used.

The interesting insights provided by Markowitz (1952) arise from the interplay between the mathematics of return and risk. These simple and intuitive issues are at the heart of modern portfolio theory. Nevertheless, the large proportion of investment is not allocated on the basis of mean-variance optimization. What is the problem with "pure application" of mean-variance optimization, and what it makes it difficult to apply in practice?

The first complication is perhaps obvious. It is hard to quantify expected returns and covariances. These parameters can be estimated using historical

data, analytic models, analysts' forecasts, or other methods. When historical data is used to estimate model parameters, there are at least two main areas of concern: estimation errors and nonstationarity of the model parameters.

When security prices are determined within an efficient market structure, a probability distribution can be used to describe them. Normal probability distribution is assumed as an appropriate description of the return function. If returns are normally, identically and independently distributed (*niid*), with a constant population mean μ and variance σ^2, then estimates of the mean and standard deviation are given by

$$\overline{R} = \frac{\sum_{i=1}^{n} R_i}{n}$$

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n}(R_i - \overline{R})^2}{n-1}}$$

For example, using monthly returns data, we might find that $\overline{R} = 1\%$ and $\hat{\sigma} = 4\%$. It can be shown that the standard deviation of the estimate of $\overline{R}$ is $stdv(\overline{R}) = \frac{\sigma}{\sqrt{n}}$. Suppose we wanted to obtain an estimate of the population mean return of 1% that was accurate to $\pm 0.1\%$ with given $\sigma = 4\%$. This would require $n = \frac{4^2}{0.1^2} = 1600$ monthly observations, that is, 133 years of monthly data. Obviously, that 60 monthly observations (5 years) will be very poor. The standard deviation (for normally distributed returns) of $\hat{\sigma}$ is given by $stdv(\hat{\sigma}) = \frac{\sqrt{2}\sigma^2}{\sqrt{n-1}}$. We use the same 5 years sample ($n = 60$ monthly observations) of data to estimate $\hat{\sigma} = 4\%$. Using the above equation, we get $stdv(\hat{\sigma}) = \frac{\sqrt{2}4^2}{\sqrt{60-1}} = 0.38\%$. That is why, accuracy of $\hat{\sigma} = 4\%$ is relatively good at 0.38%. Hence, estimation of variance (covariance) using historical data is subject to much less error than estimates of the expected return. The problem with mean-variance approach aggravates, because the optimization algorithm is simply too sensitive to an inaccuracy of return estimation. Chan, LeBaron, Lo and Poggio (1999), Ledoit and Wolf (2004) propose methods that focus on

reducing the error in estimating the covariance matrix.

Beside estimation error, change in model parameters over time is another problem that estimation models face. On the one hand, when estimating the parameters from a long return series, the returns of an asset recorded several decades ago do not provide real insight into today's return properties. In other words, a long estimation window increases the possibility of nonstationarity in the parameters. On the other hand, only very recent historical data increase estimation error, because estimates of the unknown parameters will differ from their true values. Broadie (1993) proposes a trade-off between estimation error and stationarity when choosing the length of the estimation error.

The other drawback of modern portfolio theory is that this model deals with risk as measured by the variance-covariance matrix and not other forms of risk (political risk, business risk). It is the single factor risky model. This makes the theory inappropriate for use with multi-factor risk models. Furthermore, the variance is a squared term, thus it treats any deviation above the mean return as being as risky as any deviation below the mean return.

The final issue, which makes application of modern portfolio theory difficult, is that it is not easy to treat analytically. Implicitly assumed in MPT is that individuals act in an objective almost mechanical fashion in making decisions. Market participants are completely rational in use unbiased expectations in forming and then selecting mean-variance portfolios. However, investors often exhibit behavioral traits such as limited knowledge, bounded rationality, bounded self-control, and bounded self-interest. Rather than base investment decisions on sound mean-variance analysis, individuals use heuristic rules (rules of thumb). This process of making decisions with a combination of mean-variance analysis and heuristics is called bounded rationality. Such combination is easily realized in multi agent-based models. This is exactly the focus of this research. Using an artificial market can help us understand whether traditional finance approaches are still successful in a completely dif-

ferent framework of bounded rational and heterogeneous agents.

Capital Asset Pricing Model (CAPM). The Capital Asset Pricing Model[2] of Sharpe (Sharpe, 1964) implies that all investors should hold a broadly diversified market portfolio, combined with risk free asset according to their risk preferences. If there exists a risk-free asset with return R_f, any efficient portfolio P is a combination of the risk-free asset and the market portfolio M, that corresponds to the point of tangency between the two efficient frontiers (with and without the risk-free asset).

$$R_P = xR_f + (1-x)R_M \Longleftrightarrow R_P - R_f = (1-x)(R_M - R_f)$$

The choice of x depends on the risk aversion. The portfolio variance is $\sigma_P = (1-x)\sigma_M$, consequently

$$R_P = R_f + \sigma_P \frac{R_M - R_f}{\sigma_M}$$

$$\sigma_P = [x^2 + \sigma_i^2 + (1-x)^2\sigma_M^2 + 2x(1-x)\sigma_{iM}]^{\frac{1}{2}}$$

$$R_i - R_f = \beta_i(R_M - R_f) \text{ with } \beta_i = \frac{\sigma_{iM}}{\sigma_M^2}$$

where β represents the systematic risk which is due to exposition to the market variations. CAPM demonstrates that the prices of assets are such that the market portfolio is made up of all assets in proportion to their market capitalization. CAPM highlights the relationship between the excess mean return and the exposure coefficient beta. Additionally, at equilibrium the market portfolio is optimal.

The fact that CAPM relies to variance as the risk measure provokes the active discussion and gives rise to models in which exposure to more than a single market risk factor determines expected returns. Merton (1973), Long (1974)

[2]relies on Expected Utility Theory

introduce models where investors should not simply hold a broad market index and cash, but should also invest in hedge portfolios for other economically relevant risks, like interest rate changes and commodity price inflation.

According to Roll (1977) it is very difficult to determine the true market portfolio, because market portfolio should contain all risky assets, even those that are not traded. Using a portfolio, which is not the true market portfolio, may lead to estimation errors in the betas. Roll demonstrates that even if two potential proxies for M are correlated greater than 0.9, the beta estimates obtained using each may be very different. This suggests that the empirical implications of the model are very sensitive to the choice of proxy.

According to Levy *et al.* (2000), there are several unrealistic assumptions made in CAPM framework:

- All investors are risk averse expected utility maximizers
- There is no transaction costs nor taxes
- Investors are rational (they try to maximize their expected utility) and they are efficient (they know how to reach the goals)
- Investors never make errors and choose their portfolios from the efficient set on the capital line.
- Homogeneous beliefs about expected values are required
- Investors have the same holding period

Under these assumptions all investors hold the portfolios with an identical structure, a fraction of initial wealth is invested in the risk-free asset the rest in the tangency portfolio T called market portfolio. Because all investors acquire shares in the same risky tangency portfolio, and make no other risky investments, all existing risky assets must belong to T. All agents would be willing to buy or sell the same positions, thus, there would be no trade. Since

all investors hold all risky assets in the same proportion this implies that short sales cannot exist in equilibrium. Otherwise, an investor could make arbitrarily large profits by short selling large quantities of the more expensive of the securities and buying the cheaper in equivalent amounts. Such an arbitrage would have zero cost and be riskless.

What effects will have the relaxation of one of these assumptions? How are equilibrium asset prices determined when the majority of investors are indeed efficient and rational, but a minority of them are irrational or inefficient?

To address these issues in mainstream framework, Cvitanic, Jouini, Malamud and Napp (2011) introduce an equilibrium model where investors have three possible sources of heterogeneity. They may differ in their beliefs, their degrees of risk aversion, and their time preference rate. The aim of this work is to study the impact of heterogeneity on the equilibrium properties. The authors analyze agent interactions and heterogeneity impact at the individual (individual portfolio holdings and risk sharing rules) and aggregate levels (the market price of risk, the risk free rate, the bond prices, the stock price and volatility). Cvitanic *et al.* (2011) conclude that i) for very high level of aggregate endowment, the equilibrium Arrow-Debreu price is determined by the agent with the highest individual market price of risk ii) if there is heterogeneity in beliefs (risk aversion levels) only, the equilibrium Arrow-Debreu price for very low levels of aggregate consumption is given by the Arrow-Debreu price of the most pessimistic (most risk averse agent) iii) the agent with the lowest survival index[3] survives and dominates the market asymptotically.

However, agent-based modeling is more appropriate for relaxation of homogeneity and perfect rationality assumptions, it is a good complement

[3]The survival index of agent i is defined by $k_i \equiv \rho_i + \gamma_i \left(\mu - \frac{\sigma^2}{2}\right) + \frac{1}{2}\delta_i^2$, where ρ_i, γ_i, δ_i, μ, and σ respectively denote the individual level of time preference, risk aversion, optimism and the drift and volatility of aggregate endowment.

to theoretical models. ABM provides a deeper understanding of the impact of the agent heterogeneity on market equilibrium (the Arrow-Debreu prices). Whereas theoretical models describe price dynamics and uncertainty by stochastic processes, in ABM framework price dynamic is a direct result of agents interactions. Levy *et al.* (2000) test the effect of heterogeneous expectations on equilibrium prices, and the effect of the number of assets held in the portfolio on the CAPM's results. This experiment is considered below in this section.

Arbitrage Pricing Models. In the context of general equilibrium theory considered in the previous models, there exist several hypothesis. The first one is the rationality hypothesis leading to the specification of maximization problems under constraints. The other hypothesis relies on equilibrium price formation, meaning that demands equal offers in all markets under consideration. In equilibrium framework, there can be no arbitrage opportunities, in other words, there is no possibility to earn abnormal profits risklessly at zero cost. An arbitrage opportunity indeed implies that at least one agent can reach a higher level of utility without violating his budget constraint. The Arbitrage Pricing Theory (APT)(Ross, 1976) relaxes some of strong assumptions of the CAPM. The beauty of the APT is that it does not require any assumption about utility theory or that the mean and variance of a portfolio are the only two elements in the investor's objective function (Cuthbertson and Nitzsche, 2004). It does not assume normality of returns and supposes only that investors are risk-averse, without specifying a particular utility function. More precisely, the APT abandons the analytically powerful, but empirically complicated assumption on returns distribution. It replaces this assumption with the hypothesis that there exists a set of factors determining asset returns. Thus, the main building block of the APT is a factor model, also known as a

return-generating process:

$$R_{i,t} = E[R_i] + \sum_{k=1}^{K} b_{i,k} F_{k,t} + \varepsilon_{i,t}$$

where $b_{i,k}$ denotes the sensitivity of asset i to factor k, $F_{k,t}$ denotes the return of factor k with $E[F_{k,t}] = 0$, and $\varepsilon_{i,t}$ denotes the residual return of asset i.

The APT model assumes:

- The markets are perfectly efficient
- The factor model is the same for all investors
- The number n of assets is assumed to be very large

Arbitrage conditions lead to the existence of factor risk premia λ_k such that $E[R_i] - R_f = \sum_{k=1}^{K} \lambda_k b_{i,k}$ APT allows for several risk factors to determine assets expected returns. Denote δ_k, the expected return of a portfolio with a sensitivity to factor k equal to 1, and null sensitivity to other factors. Then:

$$\lambda_k = \delta_k - R_f \text{ and } E[R_i] - R_f = \sum_{k=1}^{K} (\delta_k - R_f) b_{i,k}$$

where $b_{i,k} = \frac{Cov(R_i, \delta_k)}{Var(\delta_k)}$ are the sensitivities to the factor loadings. The simplest one-factor market model, usually labeled the Market Model, equals to the CAPM. If it were empirically verified that a single factor model is sufficient, the CAPM would be the undisputed end point of asset pricing. But it is quite unlikely at an empirical level.

It is clear that the APT overcomes some weaknesses of the CAPM, nevertheless to derive Ross's APT one needs to assume a specific return generating process but it does not provide any information about how to select the factors for return generating process, no transactions costs, unlimited short sales are allowed, a borrowing interest rate is equal to the lending rate. A relaxation

of each of these unrealistic assumptions may either change the model results only slightly or change them very drastically. It is difficult to figure out analytically the effect of the relaxation of each of these assumptions on the model equilibrium results.

Heterogeneous expectations in equilibrium theory. All models reviewed above ignore many important options and consider only ideal variants. For instance, they assume that all traders are rational, they always make the most optimal choice in a given situation. The mechanism of trading (order execution, price formation, clearing mechanism) is ignored. Thus in theoretical market models equilibrium, reached when demands equal offers, is central. The currently observed asset prices are equilibrium ones. There is no attempt, however, to compute asset demand and offers functions explicitly. Theoretical models also rely on market efficiency. Proponents of efficiency claim that in financial markets it is not possible to earn abnormal profits (other than by chance) by exploring some set of information. The majority of standard models focus on risk management and portfolio optimization for individual investors. They are not interested in the effects of a large number of investors on the market overall when they use similar risk management methods. Many researchers question however, whether markets operate as described by modern finance theory. A number of approaches have been proposed to test this issue. These approaches include empirical studies, experimental economics, the market microstructure approach, micro simulations or agent-based modeling. For example, Kahneman, Solvic and Tversky (1982), Kahneman and Tversky (2000) argue that even if a majority of the models created in traditional finance contains the assumption about the constant degree of risk, people change their risk preferences according to current circumstances, hence decision making rules vary depending on conditions. To incorporate this fact in the modeling, Takahashi and Terano (2004) introduces

the agent-based model, where agents' decision making rules are affected by trading circumstances, and the degree of risk aversion varies depending on the amount of assets held.

One of the earliest analysis that confront classical models of investment decision making with (micro-)simulations approach can be found in Levy *et al.* (2000). In order to achieve analytic results, it is necessary to assume a decision framework and to make many specific assumptions, some of which are very unrealistic. What will be the effect if one of these assumptions is relaxed? It is possible that even a small deviation from the assumptions may completely reverse the theoretical results.

There is no analytic tools to measure the effect of deviation from expected utility theory assumption. There exist some attempts to explain the effect of assumption relaxation, for example, Lintner (1965) relaxes the homogeneity of expectations assumption, but results become difficult tractable . Levy (1978), Merton (1987), Markowitz (1990), and Sharpe (1991) relax the perfect market assumption. If some of the unrealistic assumptions are relaxed, the analytical results become complicated. Levy *et al.* (2000) conclude that the more the model is relaxed, the less the results are tractable. Since it is difficult to investigate analytically the effects of deviation from some assumptions in theoretical models, Levy *et al.* use microscopic simulation (MS) methodology to run such investigations. In MS models heterogeneous expectations, different holding periods can be introduced. With MS one can model investors as *bounded rational entities*, who maximize some expected utility but who also make some errors. For instance, they act on wrong signals, incorrect information, or use technical signals that are not related to fundamental information, they consider changes in wealth rather than the total wealth. Investment decision process can have two regimes: i) all investors maximize utility function and act exactly as implied by expected utility theory ii) all investors buy and sell randomly ignoring assets fundamental values. Bounded-rationality is be-

tween these two extremes scenarios, it best describes how individuals act on stock markets.

Levy *et al.* (2000) investigate the application of MS to price determination in the heterogeneous CAPM model. The authors assume that all investors agree on the estimation of correlation matrix V but disagree on R. The kth investor's estimate of the mean return of the ith asset is $R_{ik} = R_i(1 + \varepsilon_{ik})$, where $\varepsilon_{ik} \sim N(0, \sigma)$ is a disagreement (heterogeneity) factor, which provides the model heterogeneity. $E(\varepsilon_{ik} = 0)$, implying that on average the market is in agreement with the CAPM parameters. Thus, the investment proportions in the ith stock varies across investors, x_{ik} is the proportion of the wealth of agent k invested in stock i. Then, the clearing process is defined by the expression $P^*_{i0} = \sum_{k=1}^{K} \frac{W_k x_{ik}}{N_i}$, where N_i the number of outstanding shares of the $i$the firm, W_k the wealth of the kth investor, K the number of investors. Levy *et al.* using microsimulation model show that lower expected return stocks tend to be overpriced in the heterogeneous expectations, and higher expected return stocks tend to be underpriced with the small heterogeneity factor. The explanation of this effect is due to the nonlinear effect of the heterogeneity factor ε_{ik} on wealth invested in risky assets. But the magnitude of this effect decreases as the number of assets in the market grows. The authors also conclude that a large number of asset or investors increases the robustness of the CAPM vis-á-vis the relaxation of the homogeneous-expectations assumption. Thus, the heterogeneous beliefs are not crucial to CAPM results.

In the CAPM framework theoretically each investor holds all available risky assets, while in practice investors commonly hold a relatively small number of assets in their portfolios. How does the relaxation of this assumption affect the results? Levy (1978), Merton (1987), and Sharpe (1991) analyze this question theoretically. The equilibrium model in which investors hold limited number of assets is called General CAPM (GCAPM). Levy *et al.* (2000) examine this model in MS framework, they compare the classic CAPM and

GCAPM. They assume that agent k keeps in his portfolio n_k number of assets, where n is the total number of available assets. The investor constructs the efficient set of n_k stocks with a given risk free interest rate r_f. Thus, the investor should test all $\frac{n!}{(n-n_k)!n_k!}$ portfolios. Levy *et al.* consider two cases of asset selection: i) each asset has an equal chance to be included in each portfolio ii) some stocks have a lower chance to be selected, they are known in financial literature as small stocks. All investors share the same beliefs about expected values of stocks. As the CAPM and GCAPM market portfolio differ, beta is also not the same in these two equilibrium frameworks. The authors examine the small firm or neglected stock effect on equilibrium prices and the risk-return relationship.

The authors report a higher intercept that the riskless interest rate, which conforms with empirical data. They get a strong small firm effect.

1.3 Conclusion

The point of this chapter was to present the main principles of ABM and the relationship of the ABM methodology to more standard economic modeling. We have stressed that financial markets, as complex systems, are particularly well suited for agent-based explorations. Following the critiques of theoretical models, we have reviewed a large variety of articles reporting a power of agent-based methods to revive the theory. There is a long list of features that traditional approaches are not able to match. ABM proves an intriguing possibility for solving some of these puzzles. Then, we have focused on research directions, where ABM is regarded as a promising research tool. However, the list of topics, viewed in this thesis, is far from being exhaustive; such choice of covered topics has been made according to relevance to this research.

Chapter 2

Agent-based artificial stock market

Contents

2.1 State-of-the-art

In this chapter we give an overview of recent developments in agent-based artificial stock markets. The organization of this chapter is as follows: first, we introduce issues faced by modelers when designing an agent-based artificial stock markets. We provide an overview of design issues using the framework proposed in LeBaron (2001a), Martinez-Jaramillo (2007). The main design issues identified in LeBaron (2001a) are:

- Market Mechanism & Time Scheduling
- Traded Assets
- Agents & Learning
- Calibration and Validation

The next subsection describes existing platforms based on the same axes as those proposed in (LeBaron, 2001a). Then, we introduce our software platform **A**r**T**ificial **O**pen Market (ATOM) under the axes described above that allows us to show the common and outstanding features of our artificial market.

2.1.1 Market Mechanism & Time Scheduling

Martinez-Jaramillo (2007) indicates three ways to design market a environment. *A first solution* consists in creating a simple price response to the excess demand with a simple clearing mechanism (market impact function

as response to the observed market excess), like in the models of Cont and Bouchaud (2000), Chen and Yeh (2001b), Farmer (2002), Martinez-Jaramillo (2007). The market price level is adjusted as a response to the observed market imbalance between supply and demand.

$$p_{t+1} - p_t = \alpha(D_t - S_t) \tag{2.1}$$

where D_t and S_t are the demand and supply of the agents at time t, α – a positive constant, p_t – the actual market price level, and $p_{t+1} - p_t$ – the price adjustment. Farmer and Joshi (2000) proposes a price response on demand function of the form $p_{t+1} = p_t + \frac{1}{\lambda}\sum_{i=1}^{N} d_i$, where d_i is the number of shares bought by agent i, and λ is a constant.

The advantage of a market impact function is that it is computationally fast and it gives reasonable results. But this function is very sensitive to the value of the parameter α. Another disadvantage of this price formation mechanism is that the market is assumed to be symmetric.

The second solution consists in creating a simple market clearing mechanism (analytically or computationally) where local *equilibrium price* can easily be found Levy *et al.* (1995), Arthur *et al.* (1997b). The calculation of the temporary equilibrium price is either performed analytically, or in case of a complex nonlinear demand function, computationally (LeBaron, 2001a).

The third solution consists in explicitly implementing an auction mechanism that allows to issue real market orders, like in the models of Marchesi, Cincotti, Focardi and Roberto (2000), Maslov (2000), Yang (2002), Jacobs *et al.* (2004), Gordillo, Pardo-Guerra and Stephens (2004), Ponta, Raberto and Cincotti (2011). One of the most common example within this category of price formation mechanisms is a *double-auction market.* Agents may either submit bids or ask for the traded asset to the market. For example, Chakrabarti and Roll (1999), Chakrabarti (2000) applies the *double-auction*

concept to artificial inter-bank foreign exchange market. The current best bid-ask spread and transaction prices are known to all market participants. The result is the convergence to the competitive equilibrium price and quantity.

According to *execution systems* the market can be divided between *quote* and *order* driven. On quote-driven markets, the market makers post quotes (prices and volumes at which they are willing to conduct a transaction) and the investors may decide to accept those quotes. By posting those quotes, market makers provide liquidity to the market by allowing investors to execute transactions. For example, in Farmer (2002), the market maker trades based on his positions, all orders are market orders, and price is centrally set according to an automated mechanism. In Chan and Shelton (2001) the positions of the market maker influence his decision as well. In Lux (2000) the market maker determines changes in price by reacting on imbalances between demand and supply. Other examples of quote-driven markets can be found in Daniels, Farmer, Iori and Smith (2003), Darley and Outkin (2007), Boer-Sorban (2008).

On order-driven markets buyers and sellers directly trade together. Order driven markets often take the form of an auction. Trading orders are sent to a central order book, where they are matched. Most of real markets do not rely on a single execution system, but they combine quote and order driven markets. For example, the NASDAQ Stock Market is a quote-driven market, but sometimes traders can can trade directly. Such mixed executions systems is introduced in Boer-Sorban (2008).

According to *time* the market models can be divided between *discrete* and *continuous*. In discrete models, time advances in discrete increments, while in continuous models the system changes continuously over time.

Most agent-based artificial stock markets are organized as discrete-time (or call) models (Brock and Hommes, 1997; Raberto, Cincotti, Focardi and Marchesi, 2001; Challet, Marsili and Zhang, 2005). The discrete-time market

has a *discrete time* grid with $t = 1, ..., T$ time rounds. At each round an agent is asked to submit his orders. The specific characteristics of the simulation models determine how the players take turns and how they arrive at the decisions. In other words, a modeler can coordinate the actions of traders, parameterising waiting time for trade actions. In most platforms this round is considered to be a day, or even one year. So, at the beginning of each period t, new information arrives to the market. The agents process the information and convert this into trading decisions. At the end of trading period t, agents simultaneously submit their orders to the market maker. Before the next time period begins, the market maker computes the market price. The synchronous, discrete representation of time is suitable for the models containing just simple clearing rules for price formation, when all agents are inactive until the equilibrium price is determined. Levy *et al.* (2000) use a discrete synchronous model for their microscopic simulations model. In each period, a market equilibrium price is computed from the demand and supply curves of all investors.

However, most modern financial markets operate continuously. Continuous trading is introduced in several artificial market models of Shatner, Mushnik and Solomon (2000), Farmer, Patelli and Zovko (2005), Boer-Sorban (2008). In such models, agents issue the orders at any moment, producing an order flow, that result in new trades and a market price is continuously updated.

Boer, Kayamak and Spiering (2007) study a market dynamics with prices set by a learning market-maker. The authors report that the results are significantly influenced by the choice of model updating dynamics, discrete or continuous. They conclude that the continuous nature of trading should be explicitly introduced in the artificial stock market models for reproducing realistic price dynamics.

Two different trading schemes must also be distinguished: *synchronous* and

asynchronous (Farmer, 2002; LeBaron, 2001a). In synchronous time models, the time increases with a fixed increment like minute, day or month. In asynchronous models, time step advances by irregular increments, to the next scheduled event (Jacobs *et al.*, 2004). In synchronous dynamic model, traders are centrally selected for acting at specific moment. While, in asynchronous trading scheme investors may leave the market, or find a counterpart for transaction in any moment. The agent decide not only the price and volume of orders, but also exact timing. For instance, in the model of Shatner *et al.* (2000), a trader is "sleeping" between the events that make him to wake up. The trader defines the nature of such events: he can ask to be waken up after K time steps since he placed the last order, or after new information arrival, or after price changes by 5%. Once awake, the trader decides whether he wishes to change his positions.

Generally speaking, continuous asynchronous model, allowing action at any time by any participant, represents the markets in a more realistic way, than discrete synchronous model, which typically has to make special assumptions about the effect of everyone acting at the same time at assumed discrete times. The examples of synchronous trading are introduced in Levy *et al.* (1995), Arthur *et al.* (1997b), Zimmermann, Neuneier and Grothmann (2001), Farmer (2002), Markose, Tsang and Martinez-Jaramillo (2003), while, asynchronous trading is implemented in Shatner *et al.* (2000), Jacobs *et al.* (2004), Raberto and Cincotti (2005), Jacobs, Levy and Markowitz (2010)

2.1.2 Assets

There exists a large number of different financial securities traded in the money market and capital market. Examples of money market securities are treasury bills (T-bills), commercial papers, and Eurodollars. Capital market instruments may be subdivided into two categories, debt (Treasury Notes, Bonds,

Federal Agency Securities, and bonds) and equity. The fist category is referred to as the Fixed Income Market. Currencies are regarded as capital market instruments traded on foreign exchange market (FX-market). Derivative instruments are financial securities whose price is derived from the price of an underlying financial asset.

Agent-based models may incorporate all types of financial instruments. However, the majority of agent-based models is limited to simple financial securities, for instance, one risk-free asset (bond or cash) and several risky stocks. We are unaware of artificial stock markets allowing to trade derivatives. In the model of LeBaron (2001c) the agents allocate their wealth among a risk free asset and a risky security (equity). The risk free asset pays a fixed interest r. Equity pays a dividend at each time step, that follows a random walk:

$$log(d_{t+1}) = log(d_t) + \varepsilon_t \tag{2.2}$$

where ε_t is normally distributed. The presence of a dividend can provide a benchmark for actual share price by allowing to calculate rational expectations price.

Gordillo *et al.* (2004) uses a single risky asset and a risk free asset (cash). In the model of Cincotti, Ponta and Raberto (2005) there are 100 different stocks, each related to a particular firm, in total there are 10 sectors each constituted by 10 firms. Initially, all agents have an equal number of stocks 1000 and 10 millions EU of cash. In Loistl and Vetter (2000), Loistl, Schossmann and Vetter (2001), Loistl and Veverka (2004) the number of assets is limited to 50 units.

2.1.3 Agents

The aim of this section is to give an overview of economic agents and some issues associated with their development. One of the most important design

issues is the modeling of *agents*. This issue covers *heterogeneity*, *decision making*, *utility function*, and *learning* of agent(Grothmann, 2002).

Since there is no universally accepted definition of the term "agent", we propose a definition by Wooldridge and Jennings (1995) as *"computer system that is situated in some environment, and that is capable of autonomous in this environment in order to meet its design objectives"*. The agent also has a set of actions, with which he affects the environment. Wooldridge and Jennings (1995) also provides a mathematical formalization of the agent. Assume that the agent's environment can be characterized as a set of *environment states* $S = \{s_1, s_2, ...\}$ that the agent can influence only partially. The influence of agents is a set $A = \{a_1, a_2, ...\}$ through which agent can affect the environment

$$action : S \longrightarrow A$$

Wooldridge and Jennings (1995) provides the important notions related to agent that we can interpret in the financial market context: The first one is *autonomy*, it means that an agent is not a passive subject to a global, external flow of control in its actions. An agent has his own objectives, abilities to accept information, then to analyze it and based on these results to make decisions about further actions. The second one is *situatedness*, it means that the agents act in a particular environment. In our case, the market is the environment, where the agents perform their trading. This environment presents set of constraints, rules, regulations. Finally, *proactivity* means that the agents act in order to achieve their objectives or goals.

To summarize, an agent uses a set of rules for making decisions to achieve his trading goals. This set of rules can be regarded as a trading strategy. Generally, the trading strategy styles vary from zero-intelligent agents to complex evolving strategies using neural networks, genetic algorithms and genetic programming.

Decision making rules. The principal role of agents in the stock market is to analyze information and to use own decision making rules to convert this knowledge into buy and sell actions. Agents collect information and act fairly sensibly on the basis of that information. Their choices may not be the entirely deterministic and predictable. Hence, the most typical classification of agents is based on their input information. Generally, we distinguish i) random, ii) technical and iii) fundamental traders.

Random traders ignore any information and send orders randomly. Gode and Sunder (1993) initially introduced this kind of trader and labeled them "Zero intelligence trader" (ZIT). Gode and Sundders use these agents to investigate the source of rapid convergence to competitive equilibrium in double auction markets. They argue that this equilibrium is due to the double auction rules alone. The authors report that zero-intelligence agents, under a simple constraint not to make unprofitable trades, produce the price paths close to those produced by a human subject. The prices remain close to the competitive equilibrium price, the volatility declines. At the same time, unconstrained zero-intelligent agents produce extremely volatile prices, and there is no evidence of convergence to the competitive equilibrium.

The model of Gode and Sunder (1993) provoked an active reaction and their original work has been extended by other authors to investigate related market mechanism questions. Thus, the ZIT methodology is actively applied to explain economic phenomena in various environments. For example, Bolleslev and Domowitz (1993) employs the ZIT to analyze the effect of varying or restricting the size of the order book.

Cliff and Bruten (1997a), Cliff and Bruten (1997b) examine the sensitivity of Gode and Sunder's findings to the elasticity of supply and demand. They propose an alternative algorithm, which they call "zero-intelligence plus" (ZIP). ZIP agents aim for a particular profit margin on each unit bought or sold, and this profit margin dictates the bid or ask they submit.

Jamal and Sunder (2001) examine the case of markets with imperfect information and uncertainty of state using three variants of the ZIT. The ZIT has also been employed in analyzing bubbles and crashes in asset markets.

Wu and Bhattacharyya (2003) introduce speculators into standard double-auction markets. They find that ZIP traders can no longer guarantee market efficiency when there is a large number of speculators comparing to normal traders.

Strategies that rely on historical price series as the main source of information are called *technical trading rules*. This is the most common type of trading strategy in some markets like the foreign exchange markets (Neely, 1997; Neely, Weller and Dittmar, 1997). Technical traders use charts and graphs to make decisions, in other words, they analyze price movements and chart patterns from the past to draw conclusions about future buying and selling. Technical analysts focus on generating trading signals that provide a higher investment return. A large variety of trading rules are proposed by the tenants of technical analysis: filter rules, moving average, support-and-resistance, channel break-outs, on-balance volume averages, momentum strategies, head-and-shoulders, broadening tops and bottoms, triangle, rectangle, and double tops and bottoms. By using charts and graphs to determine whether or not to be in or out of the market at any given time, agents are actually practicing *market timing*.

This type of trading strategy is largely represented in the artificial stock markets, like in Shatner *et al.* (2000), Daniels, Farmer, Iori and Smith (2002), Daniels *et al.* (2003), Boer-Sorban (2008), Brandouy and Mathieu (2007), Veryzhenko *et al.* (2010). These agents use N last prices to extrapolate the next market dynamics, e.g. future price $P_{predicted}$. Then, they compare the current market price $P_{current}$ with the predicted one $P_{predicted}$. If the $P_{current} > P_{predicted}$, the agent sells stocks. If $P_{current} < P_{predicted}$, the agent buys stocks. The main advantage of these simple trading rules is a clear tractability of

results produced by such behaviors.

Fundamental traders are driven by the "true" value of assets. The true value is estimated based on fundamental information like financial reports, information about the management of the company, earning per share, revenue, cash flow, earning announcements, dividends, analyst upgrades. This information allows to identify returns, hence the fundamental value when discounted rates are available. So it is an important task to model the fundamental value in agent-based models. Once this value has been obtained, the investor is able to compare the "true" value with the current security price. If the latter is higher its true value, the fundamental traders sell the stocks, and vice versa.

Some examples of fundamental information definitions in agent-based models of financial stock market can be found in Glosten and Milgrom (1985), Palmer, Arthur, Holland and LeBaron (1998), Shatner *et al.* (2000), LeBaron (2001a).

In Santa Fe stock market (Palmer *et al.*, 1998), the risky stock pays a stochastic dividends that follows the autoregressive AR(1) process:

$$d_t = \hat{d} + o(d_{t-1} - \hat{d}) + \mu_t \tag{2.3}$$

where $\hat{d} = 10$, $p = 0.10$ and $\mu_t \sim N(0, \sigma_\mu^2)$.

In the Baron's Model (LeBaron, 2001a,c,b) the dividends follow a random walk

$$log(d_{t+1}) = log(d_t) + \varepsilon_t \tag{2.4}$$

where d_t and ε_t are normally distributed.

The model of Glosten and Milgrom (1985) introduces the traded asset with true value toward which the trading price converges. The fundamental value of the stock evolves according to a jump process

$$V_t = V_{t-1} + \triangle \tag{2.5}$$

where $\triangle \sim N(0, \sigma)$.

The fundamental value can reach its low $\underline{V}$ with a probability σ or its high level $\overline{V}$ with a probability $1 - \sigma$. Additionally, authors introduce noisily informed traders that observe a slightly modified fundamental value $W_t = V_t + \varepsilon$, where $\varepsilon \sim N(0, \sigma_W^2)$. Noisily informed traders buy if the fundamental value is higher than the market maker's ask price ($W_t > P_a$), they sell if the observed fundamental value is below the bid price ($W_t < P_b$). They do not issue the order in case if $P_b \leq W_t \leq P_a$.

Objective functions. Concerning the objective function, there are two ways to design this important element of decision making. As mentioned in Grothmann (2002), it can be designed implicitly or explicitly. In case of an implicit objective function, the decision making process incorporates indirectly the agents' objectives. For example, profit maximization agents take advantages of the price fluctuations in order to maximize their gains. Such type of a function is introduced in Farmer (2002).

An other example of profit maximizers is given in Zimmermann *et al.* (2001). $\hat{r}_{t+1} = E[ln\frac{p_{t+1}}{p_t}]$ denotes the expected price shifts with $t = 1, ..., T$. The objective profit maximization function of agent is described as follow:

$$\frac{1}{t}\sum_{t=1}^{T} \hat{r}_{t+1} a_t^i \longmapsto max_{a_t^i} \tag{2.6}$$

where a_t^i is the trading decision of agent i at the moment of time t. If the price is declining $\hat{r}_{t+1} < 0$, the agent sells the stocks. On the other hand, the agent keeps long positions (buy assets) if he predicts a price increasing $\hat{r}_{t+1} > 0$.

Kaizoji (2000) introduces an interesting approach to maximize the profits. The agent tries to predict the trading decisions of a majority of market participants. This approach is motivated by the idea that the price is significantly influenced by majority. If a majority of agents sells stocks the price declines.

If the majority buys, the price increases. So some agents can make extra profit by anticipating the behavior of the crowd.

The explicit representation of objective function is the utility maximization function.

Utility function. Utility refers to the consumers' satisfaction from the consumption of goods and services. Utility can be applied to wealth as well as goods and services. The marginal utility function is actively used in the literature. This function reflects the fact that the every individual benefits from an additional unit of wealth, but the utility of this gain is less for someone who already has large wealth. Thus, it is common to maximize the utility of wealth rather than the expected wealth. The utility function provides a relative measure of investor's preferences for wealth and the amount of risk they are willing to undertake in order to maximize their wealth. There are different types of investors: risk averters, risk lovers or risk neutral that differ by the shape of their utility functions.

Usually, in agent-based models, the utility function is expressed in terms of wealth or risk management. One of the most frequent utility functions is the Constant Absolute Risk Aversion (CARA) function

$$U(W_{t+1}) = -e^{-AW_{t+1}} \tag{2.7}$$

where A is the absolute risk-aversion degree and W_{t+1} is the agent's expected wealth level for the next period. CARA is used in the models of Arthur *et al.* (1997a), Palmer *et al.* (1998), Yang (1999), Chen and Yeh (2001b), Hommes (2001), Yang (2002). In the Santa Fe stock market agents use a classifier system to make predictions on the first and second moments of stock returns.

Agents maximize their utility function with respect to their budget constraint:

$$W_{i,t+1} = x_{i,t}(p_{t+1} + d_{t+1}) + (1 + r_f)(W_{i,t} - p_t x_{i,t}) \tag{2.8}$$

where $x_{i,t}$ is the number of risky assets held by agent i in t, d_{t+1} is the dividend attributed to risky assets, and r_f is a risk free interest rate.

Thus, the objective of the agent is to maximize the expected utility $E(U_i(W_{i,t+1}))$:

$$E(U_i(W_{i,t+1})) = E(-e^{-AW_{i,t+1}}|u_{i,t}) \tag{2.9}$$

Given a set of external influences $u_{i,t}$ the agent's conditional utility expectation $E(U_i(W_{i,t+1}))$ depends on his future wealth $W_{i,t+1}$. One of the most important design issues of the CARA utility function concept is the construction of the price and dividend expectations $E(p_{t+1} + d_{t+1})$. Several solutions exist for doing so, for instance, Yang (1999) use recurrent neural networks for this task.

There is a large variety of agents-based models implementing the utility maximization trading rules (Levy *et al.*, 1995; Arifovic, 1996; LeBaron, 2001c; Chiarella and He, 2001). For example, Levy *et al.* (1995) introduce logarithmic utility function in their Levy Levy Solomon Model, this function is expressed as follows $U(W) = ln(W)$.

In the Baron's Model (LeBaron, 2001c) agents use Constant Relative Risk Aversion (CRRA) utility function to maximize their wealth at each time step. CRRA is expressed as follows $U(W) = \frac{W^{1-\gamma}}{1-\gamma}$, where $\gamma > 0$, $\gamma \neq 1$.

In the Business school agent-based multi-asset model (Chen and Yeh, 2001b; Chen, Yen and Liao, 2002) agents use Quadratic utility $U(W_{t+1} = W_{t+1} - \frac{b}{2}W_{t+1}^2)$ for their decision making.

Mean-variance portfolio optimization rules are implemented in the models of Chen *et al.* (2002) Cincotti *et al.* (2005).

2.1.4 Learning

Learning is an important element of the design of an artificial financial market. Agents should be able to update their trading strategies in response to changing market conditions. Brenner (2006) provides an explicit overview of the learning techniques used in economic models. He proposes to classify the learning models according to their origin: psychology-based models, rationality-based models, adaptive models, belief learning models, and models inspired by artificial intelligence and biology. The second classification relies on the main economic fields in which the models are applied. Bayesian learning and least-squares learning is used in macro-economic; reinforcement learning, fictitious play and direction theory are applied in experimental economics; evolutionary programming and genetic programming are frequently used in agent-based computational economics. Artificial Intelligence techniques are the main tool to design agent's learning: the examples including *genetic algorithm* are implemented in the agent-based models of Palmer, Arthur, Holland, LeBaron and Tayler (1994), Arifovic (2001), LeBaron (2001c); *learning classifier systems* in Arthur *et al.* (1997a), LeBaron *et al.* (1999), Schulenburg and Ross (2002); *artificial neural networks* in Yang (2002), Zimmermann *et al.* (2001); *genetic programming* in Chen and Yeh (2001a), Chen and Yen (2002), Edmonds (1999), Martinez-Jaramillo (2007); *reinforcement learning* Chan and Shelton (2001).

In this thesis we propose an overview of *genetic algorithms* principles (Holland, 1975), as they are regarded as a key component in many agent-based financial market for modeling of learning and adaptation (Palmer *et al.*, 1994; Arifovic, 2001; LeBaron, 2001c). While there is a large variation in the specific details of genetic algorithms, there are some general principles and procedures that are regarded as relatively standard:

– The objective function of genetic algorithm should be specified. The

parameter that is going to maximize (or minimize) the objective should also be indicated.

- The vector of parameters, representing candidate solutions are encoded as string of finite length.
- The performance of each string in the population is evaluated using the objective criterion – the string's fitness.
- A new generation of N strings is determined using operations that mimic natural selection that occurs in biological processes.

For instance in the Santa Fe model (Arthur *et al.*, 1997b; LeBaron *et al.*, 1999) agents predict the future return and dividend of the traded asset. The sets of trading rules evolves on the basis of genetic algorithms. The GA is invoked every K period for each agent and replaces the 20 worst rules of the set of 100. GA uses the genetic operators of mutation and crossovers to generate new rules that allow to adapt the set of strategies to the changing market conditions. Mutation is an important part of any evolutionary algorithm, which helps maintain a diverse population. It could be interpreted as learning by experiment. The performance of the rules is used as a fitness criteria. Thus, inefficient trading rules are replaced by the best ones. Initially, the agents are limited in their rationality and their knowledge about the market. During the trading, the agents learn and thus, become reasonable experts in their domains. The probability of GA activation by each agent is an important parameter that determines the "speed of learning" of the agents.

2.1.5 Interactions

The next key aspect of agent-based models are the agent interactions that are at the heart of the explanation of many statistical properties of stock markets. There are two important points of interaction in the financial market:

- Who is interacting with whom: market-market, agent-market, agent-agent, agent-external world, etc.

- What is the result of the interactions: monetary payoffs, the commodity, money, etc. For example, agent A interacts with agent B, provides information to agent C and pays money to agent B.

Many existing empirical studies put forward the *market-market relation*. For instance, there exist long-run linkages and co-movements between the markets, such as Canadian, Mexican and United States. But usually this type of interaction is ignored in agent-based models.

Market organization is an important factor affecting the strategies of participants and their profitability that is *agent-market interaction*. Often, the outcome of a strategy is not uniquely associated with any particular feature of the model or behavior, it's up to a set of market rules and market participants. An example of such relation is the influence of tick size on strategies profitability.
This sort of influence fits Interaction Movement Computation MIC^* model (Gouachin, Michel and Guiraud, 2004). This model is actively used for multi-agent systems engineering. In this model the environment defines a set of actions for autonomous agents to achieve their goals. The environment plays a fundamental role in order to guarantee the autonomy property of agents. The agents interact with one another in order to achieve either a common or an individual objective. Moreover, existing models of advance traders interactions fit well the MIC^* architecture.

The simplest *agent-agent interactions* can be "communication" through the environment, via market microstructure. In such interactions, agents have a different access to external information and a different interpretation of it. They estimate their own positions, make decisions about perspective bid-ask prices, and do not share the trading strategy. To demonstrate the interactions

through order book, we consider the following simple example: agent i sends bid order β_i or ask orders α_i, with bid-ask spread $\alpha_i - \beta_i$. A trade is concluded between agent i and agent j, if $\beta_i > \alpha_j$. In such a way, the traders' decisions are influenced by decisions of other market participants.
To keep simulations fairness and to avoid the biases in the internal information access, agents should be simultaneously informed about changes in the order book. All orders, as influences, are collected in the order book, once, all agents have sent their orders, the price is fixed as a reaction. This model fits the Influence Reaction Model for Simulation (IRM4S) concept (Ferber and Muller, 1996; Michel, 2007). Trading strategies influence general system state (price formation) through "collection of influences" (order book), at the same time, they use current environment information (historical price) for further decisions. In other words, the agents can also interact through the common variable of the past price history, but they are not directly affected by the actions of others.

More *complex communication model* is introduced by Bouchaud and Potter (2000). Agents have three choices of market actions: buy, sell, or remain positions. Traders can form coalitions with others who share some choices of actions. N agents are assumed to be located at the vertexes of a random graph, and agent i is linked to agent j with a probability p_{ij}. A coalition is simply the set of connected agents (a cluster) with a given action $\triangle\Theta$. Agents in a cluster share the same actions and do not trade among themselves.

In the MIC^* model (Gouachin *et al.*, 2004) two agents are regarded as interacting when the perceptions of an agent are influenced by the emissions of another. This type of interaction is realized in the *adaptive population model* of Lux and Marchsi (1999), where the agents are divided into two groups: fundamentalists and noise (chartist) traders. Noise traders are divided into an optimistic and a pessimistic group. The important feature of this model is possibility to switch strategy between optimistic and pessimistic patterns,

moreover between the noise and the fundamentalist agents groups, based on the profit difference in such groups.

Generally speaking, cross-group movements change the traders group's proportion, hence change the market state that is influenced also by individuals' behavior (market price dynamic depends on agents' strategies).

Influence-opinion formation model of Kirman (1993) represents the *direct interaction scheme*. The agent may hold one of two views. In each time step, the two agents may randomly meet, and there is a fixed probability that one agent may convince the other agent to follow his opinion. In addition, there is also a small probability that an agent changes his opinion independently. Applied to a financial market setting, one may observe such interaction between technical traders and fundamentalist. Note, that agents may change rules due to direct interactions with other agents, but switching probabilities are independent of the performance of the rules.

2.1.6 Calibration and validation

While agent-based models are able to represent the market structure and trading rules in a very realistic manner and are capable to reproduce many real market patterns, most models may be not easily calibrated to real-world data.

Uncalibrated models may be used to investigate the periods of high and low volatility, or agents learning and adaptation over time. Even if each of these scenarios is designed to represent real world situations, it does not rely on real data for calibrating the arrival rate of new information, the real proportion between fundamentalists and chartists on the real market, the assets held by each trader. The *calibration* of model parameters is one way to connect artificial market model to the real world (LeBaron, 2001a). A wide discussion on calibration is presented in Windrum, Fagiolo and Modeta

(2007). The approach of incorporating parameters borrowed from real or experimental market to calibrate a model can be found in Zimmermann *et al.* (2001), Boswijk, Hommes and Manzan (2003).

In calibrating the parameters and behaviors in the simulations, Darley and Outkin (2007) identify two interrelated areas: 1) calibrating to the real-world trading volume distribution, and 2) calibrating behaviors, especially, market makers behaviors. Darley and Outkin (2007) use a Nasdaq Inc. data set for calibrations to identify an interval of the trading day and associate it with an interval in the simulation, the total traded volumes in real market and artificial market is statistically similar. Then, they create dealers and investors orders with different size that approximately reproduce the distribution of traded volumes observed in the real world. Next, the authors use data mining techniques to calibrate the agents' behavior. With the information provided by time series data of real market transactions, the strategies are calibrated to generate the price paths that are the most consistent with the real world Nasdaq market.

Calibration consists in setting parameters to help the model best fit empirical data, while *validation* consists in verifying the hypothesis about the ability of the model to fit real data. Validation is needed to select the model which best fits real market data or data properties. An agent-based model is validated if the generated data and the real data belong to the same distribution. Calibration methodologies are necessary and crucial in validation.

2.2 An overview of existing platforms

In this subsection we introduce two "classical" the most cited and the most used artificial stock markets, the Santa Fe Artificial Stock Market and the Genoa Artificial Stock Market. Then, we overview recently developed platforms, Agent-Based Simulation of Trading Roles in an Asynchronous Contin-

uous Trading Environment (ABSTRACTE) and the NASDAQ Market Simulator. In these platforms the developers made design choices close to those in ATOM. The outflow of each subsection is as follows: first, we describe the choice of market mechanism and traded assets made in the systems; second, different types of agents and their learning capabilities; third, the technical features including programming languages and libraries used for platform development; finally, we review the research contributions in finance made using concerned platforms.

2.2.1 Santa Fe Artificial Stock Market (SF-ASM)

The Santa Fe Artificial Stock Market is a well known agent-based system. It is cited in many articles Arthur *et al.* (1997b), LeBaron *et al.* (1999), Palmer *et al.* (1994), LeBaron (2002), Ehrentreich (2002), Ehrentreich (2006). The Santa Fe market was not the first. There were several early simulations that tried to answer some concrete questions. Cohen, Maier, Schwartz and Whitcomb (1983) investigate the impact of randomly behaving agents on various market structures, Kim and Markowitz (1989) look at the interactions of specific trading strategies.

It was originally designed to investigate the dynamics of a market in which bounded rational agents form endogenous expectations by means of inductive reasoning (Arthur, 1994c). It helps study the emergence of trading patterns as agents learn over time.

Market mechanism. Double-auction mechanism. Market makers announce the price to all the traders.

Agents. There are N traders in the market, fundamentalists and technical traders using moving average trading signals. Agents take their decisions synchronously and send the orders with direction $a_{i,t}$ (buy $a_{i,t} = 1$ or sell $a_{i,t} =$

-1), and the market maker calculates the imbalance between demand and supply ($I_t = \sum_i a_{i,t}$). The price of orders is computed as $p_{t+1} = p_t(1 + \beta \times I_t)$

Traders use constant absolute risk-aversion expected utility function $CARA\ U(W_{t+1} = -e^{-\lambda W_{t+1}})$, where λ is their degree in risk-aversion, and W_{t+1} is an agent's expected wealth level for the next time period. Agents use a classifier system to make predictions on the first and second moments for stock returns. Agents maximize their utility function with respect to their budget constraint:

$$W_{i,t+1} = x_{i,t}(p_{t+1} + d_{t+1}) + (1 + r_f)(W_{i,t} - p_t x_{i,t}) \tag{2.10}$$

Traders are homogeneous with respect to utility function and risk aversion, they are heterogeneous with respect to expectations about future price and dividends. The agents buy/sell stocks and receive interest/dividends on their investments.

SF-ASM allows the interactions between heterogeneous agents as described in Lux and Marchsi (1999)

Assets. There are only *two assets.* First, there is a risk free bond with a constant interest rate $r = 0.10$. The second asset is the risky stock , paying a stochastic dividend that follows the autoregressive AR(1) process : $d_t = \overline{d} + p(d_{t-1} - \overline{d}) + \mu_t$, where $\overline{d} = 10$, $p = 0.95$, $\mu_t \sim N(0, \sigma_\mu^2)$. The fundamental share value is unknown to the traders. It depends on the price, the dividends and the risk-free interest rate. There are no complex instruments, such as options.

Learning. SF-ASM uses a generic algorithm (GA) to modify the trading rules. Each trader $i = 1, \cdots, N$ updates his set of forecasting rules with probability P in each period t using generic algorithm (GA). The updating of

forecasting rule sets happens in different time periods for different traders. P is an important parameter determining the speed of learning. For each agent, GA is invoked every K period and replaces S worst rules of the rule set.

Technical features. The SF-ASM market was programmed in the C programming language under UNIX, later it was modified to objective-c. One important feature of the empirical results presented in LeBaron *et al.* (1999) is that they use a cross section of 25 different market runs. The second key result is that these features are very sensitive to the learning speeds of agents, or the frequency with which they run the generic algorithm. The platform programmed based on the Santa Fe Artificial Stock Market and its modification were introduced in the papers of LeBaron (2002), Johnson (2002), who discuss platform design issues. A current objective-C version using the Swarm package is currently hosted by Paul Johnson at `http://ArtStkMkt.sourceforge.net`. It was also reprogrammed using Java and RePast library.

Research questions. SF-ASM is a relatively simple platform actively used to address several important and controversial questions in financial economics. SF-ASM is used to examine whether the introduction of trader learning helps explain empirical observations. Depending on the generic algorithm invocation interval, LeBaron *et al.* (1999) report two different regimes: 1) the so-called rational expectations regime emerges when agents have a slow learning rate 2) the so-called rich psychological or complex regime arises when agents have a fast exploration rate. When the agents frequently update their rules the market is more likely to generate the patterns common to actual financial time series. When the agents update their rules more slowly the market is very close to what would be predicted in the homogeneous rational expectations equilibrium. The price series exhibit bubbles and crashes, fat tails in return distribution, trading volume exhibits GARCH-behavior and

is auto-correlated. Trading volume is also strongly persistent and correlated with price volatility. The authors report that the emergence of these statistical properties is due to the interactions between many heterogeneous agents.

2.2.2 Genoa Artificial Stock Market (GASM)

The Genoa is the artificial financial stock market presented by Marchesi *et al.* (2000), Raberto *et al.* (2001), Cincotti *et al.* (2005). The Genoa platform contains several modules: i) assets ii) trading iii) clusters iv) simulation part implements market organization with trading rules, agents who follow these rules. The authors claim that this platform has been developed not as a stand-alone optimization application for present model, but as an evolving system able to be continuously modified and updated. For instance, the platform can be extended to an unlimited number of different kinds of securities, it can be used as an engine for a trading game and, moreover, for implementing real online trading.

Market mechanism. The trading mechanism of the GASM is based on a realistic double-auction order matching mechanism. The clearing price is defined as the crossing of demand and supply functions.

Assets. Initially this platform has been introduced as a mono-asset market in Marchesi *et al.* (2000), but it has been recently updated to a multi-asset system by Ponta *et al.* (2011).

Agents. Each agent is an autonomous entity that uses a specific number of assets and cash, which are specified as the initial parameters. There is a variety of trading strategies: random trading, fundamental analysis, technical analysis, mean-variance portfolio optimization, and mean-reversion traders.

- *Random Traders (R)* have zero intelligence and simply issue random orders. They issue a buy or a sell limit order with equal probability. The order size is computed at random with a uniform probability, but there are budget constraints, meaning that the issued sell order depends on available cash and stocks.

- *Fundamentalist traders (F)* believe that each asset has a fundamental price P_f related to external factors.

- *Momentum Traders (M)* follow the past price trends. They buy (sell) when the price goes up (down).

- *Contrarian Traders (C)* are trend followers too, but they speculate, if the stock price is rising, it will stop rising soon and fall, so it is better to sell near the maximum, and vice versa. Contrarian Traders compute the order size in the same way as Random Traders.

Learning. The Genoa platform uses a generic algorithm (GA) to modify the trading rules.

Technical features. The system is developed using the Smalltalk language.

Research questions. The Genoa artificial stock market has been employed to investigate asset price dynamics from a microscopic perspective. Within a multi-asset artificial stock market, zero-intelligence traders generate realistic price series (see figure 2.1) with returns exhibiting volatility clustering, fat tails and reversion to the mean. The authors report that only the restriction on agents' allocation strategies produce such realistic stylized facts. They also analyze the impact of stock option trading on the market of the underlying security, and the influence of hedging strategies on the long-run wealth distribution of traders and on price volatility.

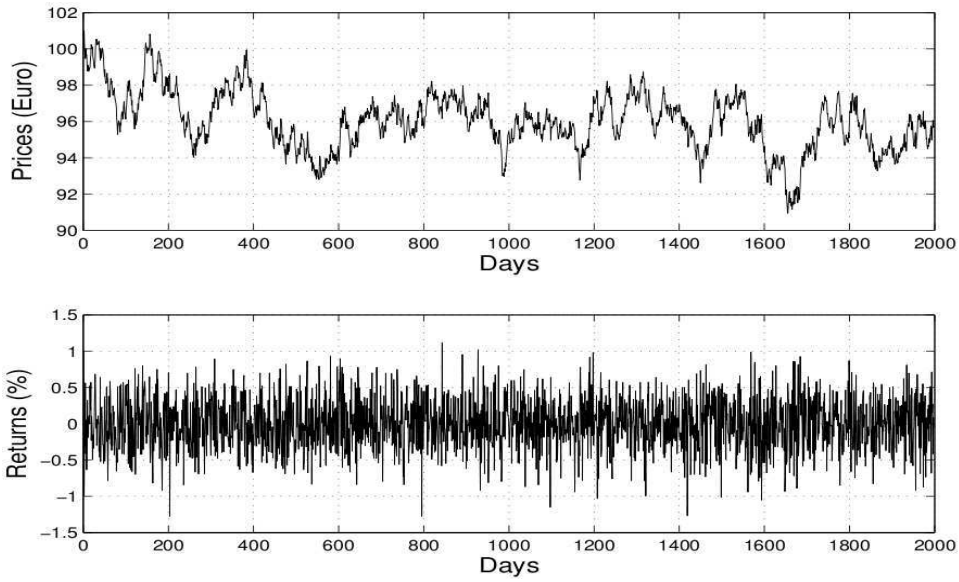


Figure 2.1: Prices and Returns in simulations without dividends paying with ZIT.

The question of the long-run wealth distribution of agents with different trading strategies is studied using the Genoa Artificial Stock Market (Raberto, Cincotti, Focardi and Marchesi, 2003). The results show that a trading strategy cannot be judged solely on the basis of the strategy itself. Its success depends also on market conditions. The authors conclude that in an artificial market with finite resources, the average price level and the trends are defined by the amount of available cash that is injected in the market. Different populations of traders with simple but fixed trading strategies cannot coexist in the long run. One population prevails and the others progressively lose weight and disappear.

Mannaro, Marchesi and Setzu (2008) use the Genoa framework to assess the impact of a Tobin-like transaction taxes on market volatility and trading mechanism. For this research they introduce several types of traders: ZIT, fundamentalist ans two kinds of chartists. The authors report that the price volatility increases consistently with the tax rate, but only when the chartist

traders are present in the market.

2.2.3 Agent-Based Simulation of Trading Roles in an Asynchronous Continuous Trading Environment (ABSTRACTE)

ABSTRACTE is a modular agent-based trading environment introduced and described by Boer-Sorban (2008). This is a modular tool for representing and studying several types of markets and trading strategies. The main purpose of ABSTRACTE developers is the understanding and explanation of market dynamics. ABSTRACTE is based on JADE (Java Agent DEvelopment Framework). JADE agents communicate through asynchronous message passing. The agents use Agent Communication Language (ACL) format to write their messages.

The framework consists of three main components. *The market place* models the institutional structure behind price formation. This component integrates financial traders (market makers and brokers). Their role is determined by the market rules of specific markets. *Investors* are not an internal part of a market organization. They observe markets, make trading decisions, and send their orders to financial traders. *The information source* component is designed to generate news related to the stock traded on the market, such as dividends or fundamental value. The information source component generates news about a stock. Based on this information, the investors send their orders to brokers or market makers on the marketplace. The process of order executing depends on the specific execution system and on the strategy applied by financial traders. The figure 2.2 introduces the structure of the system.

Market mechanism. Primarily, the authors focus on continuous quote-driven system, where bid and ask quotes are placed by market makers. Later,

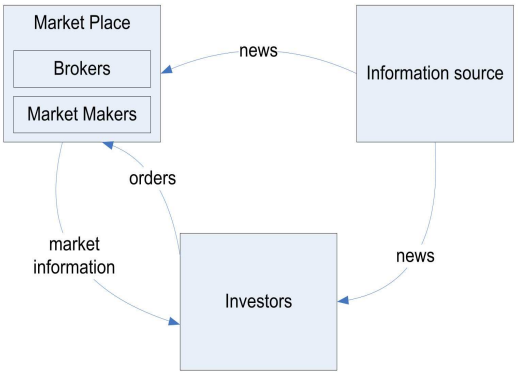


Figure 2.2: The components of the ABSTRACTE framework

the auction-based execution mechanism has been also incorporated into the environment.

Continuous and call trading sessions are implemented within the platform. It is up to the user to decide with which form to experiment. In continuous trading sessions, the orders are executed as far as there is a possibility of matching. If call sessions are used, the orders are placed only at designated times. Traders are notified whenever a call session opens or closes.

Limit and market orders can be placed within this framework. Orders are described by a stock name, size, side, quoted price and timestamp.

Assets. Like the majority of artificial stock markets ABSTRACTE developers focus on experiments where one risky stock is traded Boer-Sorban (2008). Cash is regarded as a risk free instrument.

Agents. The ABSTRACTE incorporates three types of market participants. *Market makers* execute the orders of other market participants. They are responsible for providing market liquidity, perceiving the environment, determining bid and ask quotes, receiving orders and their executing.

Brokers are primarily entitled to execute orders on behalf of investors. Theoretically a broker has three ways to carry out a trading instruction: a)

match orders internally: if there are other earlier received orders in the order book of the broker that clears at a price close to the current market price; b) try to negotiate with other brokers within the market maker's quoted spread. c) submit the order to a third party (such as a market maker or a central matching system) for execution.

There are two classes of *investors* in the system. Trader agents represent the market participants, who hold a list of stocks and trade them on the market. Manager agents control the market by keeping track of the time, creating and managing a given list of traders, getting information, and diffusing news.

In the ABSTRACTE framework traders are not centrally selected for their tradings, but are individual autonomous elements. They decide when to place an order. Autonomy results from the chosen agent-based implementation. Moreover, agents can carry out different tasks at the same moment.

Technical features. ABSTRACTE is built using the Java Agent DEvelopment Framework (JADE) environment.

Research questions. Boer *et al.* (2007) study the behavior of a learning market maker in a market with information asymmetries, and observe the differences caused by market dynamics (discrete and continuous). They show that the market price is significantly influenced by the choice of market dynamic. The authors report that the main difference in the outcomes of the discrete-time and continuous-time simulation is the fluctuation of bid-ask quotes, and consequently prices. Prices tend to fluctuate more often and with larger amplitude in continuous, asynchronous setting. The magnitude of the fluctuations tends to increase with the increasing number of investors. Boer *et al.* explain this phenomenon by the interaction of agents, who act all at the same time, and market maker, who overreact to changes in order queues.

2.2.4 NASDAQ Market Simulator

The NASDAQ Market Simulator (Darley and Outkin, 2007) is an agent-based model that has been initially developed to explore the effects of the changes in market microstructure or market rules on the behavior of participants such as market makers and traders in the Nasdaq market.

At the highest level, the simulator consists of four types of objects: Price, Dealers, Investors, and Market.

Market mechanism. This platform mimics the architecture of the NASDAQ real market.

Assets. A single risky security is traded. The "true" value of the stock follows a stochastic dynamics.

Agents. The players in the market, *Dealers* and *Investors*, are represented as "agents" in the Nasdaq market simulator. Each investor has access to the current "true" value of the underlying security adjusted by an "error", that depends on the individual agent's "informedness". Investors decide whether to purchase or to sell stocks by comparing their current price with the "true" value. Dealers, on the other hand, do not have any prior information about the "true" value. They post their current bid and ask prices on the public board. If a dealer's current prices are below (above) the investor's perception of the value, then the Investor will buy (sell) a pre-determined number of shares of the security.

Several Dealer strategies have been implemented in the Nasdaq market simulator:

- A *Basic Dealer* maintains his quotes until he receives a certain number of trades on the buy (or sell) side, and then adjusts the quote appropriately.

- A *Price Volume Dealer* tries to deduce from the market data the current demand and supply schedules and thus decides whether the current price is above or below of the "true" value.

- A *Volume Dealer* looks at the past discounted volume only, without explicitly taking the price into account. If the Volume Dealer observes more buys in the past, he concludes that the price is probably above the "true" value. Thus, he increases the ask price.

- A *Parasitic Dealer* waits until a sufficiently narrow spread with sufficient volume appears on the board to be able to realize a transaction without price discovery.

- A *Matching Dealer* learns the connections between observations and actions which are profitable.

- A *Classifier Dealer* is similar to a Matching Dealer, except that he learns over patterns of observations.

- A *Dynamic System Dealer* uses a discrete dynamical system to set his bid-ask spread and mean price.

Technical features. Nasdaq simulator is developed using the Java programming language. There are tree main classes: the Market class, the Dealer class, and the Investor class. Both Dealers and Investors have strategy functions that determine their trading desires. Investors decide whether to buy or sell the security, while the dealer decides what bids and asks to post on the Market's Quote Montage.

Research questions. One of the questions investigated using NASDAQ Market Simulator is the effect of decimalization and reduction of tick size on price discovery, market volatility, and behavior of market participants Darley

and Outkin (2007). Decimalization means a change to expressing prices in a decimal system, rather than in dollars and fractions of dollars. The authors conclude that the decimalization may have a significant negative impact on price discovery in the presence of parasitic strategies. A parasitic strategy is defined as a strategy which attempts to make a profit without contributing to the process of price discovery, or as a strategy that is to take advantage of other players' actions.

2.3 ArTificial Open Market (ATOM)

We have presented in the section 2.1 an extensive overview of the existing methodology to design an agent-based models of financial stock markets. A series of existing artificial stock market platforms has also been introduced. Their entities' organization, agents' strategies and price setting mechanisms have been detailed. As concluded, there is a vast number of markets. Most of present artificial market platforms suffer from a lack of flexibility and must be viewed as software rather than as Application Programming Interface (API). They are oriented for specific problem solving, but cannot be used to explore a wide range of financial issues due to some structural choices made by the developer during the coding phase. These observations have motivated the development of a new agent-based artificial trading environment in order to overcome some limits of other platforms and to provide a new research tool.

ATOM is a robust and reliable platform, on which researchers can run the thousands of sophisticated evolving agents (Mathieu and Brandouy, 2012). It has been developed at Lille 1 University[1] . This software architecture is currently realized based on an architecture which is close to the Euronext-NYSE Stock Exchange, described in the Backgrounds Section. We with start the description by giving the architecture features.

[1] `http://atom.univ-lille1.fr`

Execution system ATOM implements both *quote-driven* and *order-driven* systems. In the quote-driven framework, bid and ask quotes are placed by market makers. The investors decide whether to accept or to refuse those offers and demands.
In the order-driven framework, an order book is the core of the application. Incoming orders are gathered and trades are cleared whenever there exist the conditions for price setting. Orders are sorted according to price-time priority. In an order-driven market the price is set as the result of matching the orders of buyers and sellers.

In ATOM, there is also a possibility to implement several order books, each for specific stock, agents send their orders to the market, that forwards these orders to concerned order books. As far as a price is set, it is sent to the market and becomes a public information. As a result, agents will react to such market changes, which creates feedback loops in the system.

Discrete and continuous time trading regimes In ATOM, one has a possibility to switch between discrete and continuous trading regimes. *Discrete time* trading (or call session) means that agents should act at the moment of time t before any action can be performed at $t + 1$. In a call session all traders trade at the same time when the market is called. This regime is realized with controlled scheduling that implies the consequent action collection and their execution.
If a *continuous trading* session is applied, orders can be continuously placed by investors, and trades are arranged whenever possible. Trading is continuous in the sense that traders may continuously attempt to arrange trades. Continuous trading is provided in ATOM by controlled scheduling with multiple processes.

Many continuous order-driven exchanges open their trading session with call market auctions and then switch over to continuous trading. The number

of call auctions and the time interval between two call-auctions can be easily specified in the ATOM framework. Such parameterization is at the heart of day time-frame realization in artificial stock markets.

Intraday and extraday time grain The choice of the time step of the model is an important design question. Contrary to most existing artificial markets that consider one trading round as one day or one year time period (Jacobs *et al.*, 2004; Martinez-Jaramillo, 2007), ATOM reproduces intraday (with openning, continuous trading and closing process) as well as extraday trading. Many continuous order-driven exchanges open their trading sessions with discrete market auctions or pre-opening session and then switch over to continuous trading. During the pre-opening session orders are sent to the central order book without any transaction taking place. It allows to determine the opening price and to provide initial market liquidity. One trading round corresponds to one intraday trading tick. We provide this market functionality in ATOM, the number of ticks of discrete and continuous trading is configuration parameter.

Orders In ATOM, most of NYSE-Euronext orders are allowed: limit, market, cancel or update orders, as well as sophisticated combinations such as stop-limit orders or limit orders with *"iceberg"* execution (see table 2.1).[2] Orders always specify which instrument to trade, how much to trade, and whether to buy or sell. Additionally, each order has a validity duration. The order is active in the order book until it has generated a trade, or it has been canceled by an agent or automatically due to the validity duration. Each order has a time-stamp and a unique number providing the possibility to monitor the order state. Order execution depends on the execution mechanism applied.

[2]See NYSE-Euronext rule-book, at `http://www.euronext.com`.

Order Types	Pre-Opening	Continuous Trading	Closing	Price Required
Limit	Yes	Yes	Yes	Yes
Market	Yes	Yes	No	No
Iceberg	Yes	Yes	Yes	Yes
Stop Limit	Yes	Yes	No	Yes
Stop Loss	Yes	Yes	No	Yes

Table 2.1: Orders and trading scheduling implemented in ATOM

Multi-asset market Most artificial markets deal with one single asset. This choice is made for software solution simplicity. At the same time this choice excludes the possibility of portfolio optimization and risk management question of research conducted based on such market. Today, there are only a few artificial stock markets presenting multi-asset order book (Chen *et al.*, 2002; Jacobs *et al.*, 2004; Ponta *et al.*, 2011). ATOM is one of them. It allows to simulate more advanced agents' behavior with risk-return management and wealth optimization utility functions (for example, see section 2.3.2). Cash supply expresses an asset with the risk free rate. The information, return and risk of the risky assets can be regularly generated on the market and is publicly available information. Moreover, agents have different memory span length, they can store their own relevant information and can elaborate various kinds of strategies based on these data.

Endogenous and exogenous sources of information Usually, the artificial agents aim to maximize their wealth by investing in different assets. To make profitable decisions they rely on different information for their decision making process. We propose different sources of information. First, we can impose to agent an exogenous information, such as dividend distributions or moments of return distribution. Second, the agents can use the past price history (endogenous information) to estimate an 'endogenous' (ex-post) rate of return and volatility of securities or to predict future trends and turning points in stock prices. These options can help understand the origin of the

price fluctuations exogenous or endogenous (i.e. reaction to external events or caused by the trading activity itself).

Unrestricted trading volume Other unrealistic assumption made in some artificial markets is the restriction to "one share per trade". This is a common simplification in order to reduce the implementation time, but a questionable one. Volume plays an important role in financial market stylized facts emergence. Unrestricted trading volume is also essential for portfolio diversification questions. In ATOM, traders decide not only the trading direction (buy or sell) and price, but also volume. This functionality allows the ATOM users to investigate volume-volatility relation.

Replay-engine All information about trading (identification of stock, identification of agents, fixed price, etc.) on the real stock market or on the artificial stock market is usually logged in the special log-file. ATOM takes less than few seconds to replay an entire day of trading from such log-file, that contains 400,000 activities. This tool is extremely important for policy-oriented experiments focusing on the technical features of the market microstructure (tick size, price setting protocol) and its influence on the price dynamic. It also allows the perfect reproductibility of any experiment run on ATOM which is a major advantage for both scientific and technical research.

We can, for example, use the trading data from a given day or week, replay them in the artificial market framework, and then answer questions about how the market would have behaved during that time period "if" the tick size or "if" transaction costs had been different. This ability also allows to re-produce real-world order flow.

Flexibility One of the critics of ABM is that the modeler can not intervene to adjust system evolution. All initial specifications should be completely pre-

defined. To overcome this weakness the platform should be flexible and should easily incorporate the evolving parameters settings. For example, it has been shown that as far as agents easier undertake a risk. Thus the system should allow to update the risk aversion parameter of agents in accordance with their available wealth (Iglesias, Goncalves, Abramson and J.L.Vega, 2004). ATOM introduces the possibility to have evolving agent populations. ATOM is able to remove or introduce agents into the market during the simulations. Moreover one can easily control the proportion of sent orders. For instance, we can generate or not the market or stop-limit orders.

Transaction costs For the sake of simplicity, the majority of multi-agent models does not incorporate transaction costs. However, the latter are essential for several questions in finance, such as portfolio rebalancing, because profit maximizing agent will trade only if the expected return of the transaction is higher than the expenses generated by their orders for example. We implement transaction costs in our platform.

2.3.1 Technical design issues

Even if the objective of this work is to concentrate on the financial applications more than on implementation details, we still illustrate some technical solutions for architecture development. This section is intended for those who are interested in the software development of agent-based artificial stock markets.

The ATOM platform offers the implementation of multi-agent systems abstract design issues. For instance, *agent autonomy* and its behavior mean that each agent asses his position and makes decisions individually. Price history and orders sets are *emergent phenomenon* of market activities. Extremely fast simulations are realized based on *distributed simulations* with many computers interacting through a network as well as local-host. Possibility to design experiments mixing *human being and artificial traders* are implemented in our

platform.

2.3.1.1 Modular organization

From the very beginning, this API was designed having in mind the desired properties of "scalability". That implies the possibility to easily modify the system in order to incorporate some new features and functionalities. Hence during the development of an artificial stock market many MAS design principles are employed. For example, "modularity" and "encapsulation", suggesting the division of the system into different sub-organizations.

There are different attempts to organize artificial stock markets (modular organization). For example, LeBaron (2001c) propose following components of the system: *Agents*, *Rules*, and *Securities*. Separate binary modules of *Autonomous engine* (platform core), *Simulation User Interface*, and *Agents* are implemented in Muchnik and Solomon (2006); Boer-Sorban (2008) organizes her platform using *Traded instruments – Market participants –Market microstructure* parts.

The choice of the ATOM organisation is results from the intention to introduce fairly tractable markets, to be close to the NYSE-Euronext stock exchange organization and to ground modelling on MAS design main concepts. Thus, the ATOM architecture can be viewed as a system with interacting components: i) *Market* is defined in terms of microstructure ii) *Agents*, and their behaviors, and iii) *Bank* reflects intermediaries and monetary financial institutions iv) the *Artificial Economic World* provides economical indicators. We link each system entity with the sets of *Responsibilities* in order to cover all functionality and complexity of real world market.

The *Market* is the main class that incorporates and maintains all agents and order books. "Market" is responsible for the generation of market scenar-

ios. It presents the set of constraints, rules, regulations, leading participants to activities. Moreover it generates the files that gather the simulation outputs. The content of these files is easily parametrized. One can include the information about sent orders (order book Id, agent's identification, direction (Bid/Ask), volume and price), the information about fixed prices (order book Id, price, volume, direction, Id of counterparts), the information about agents (unique identification, available cash, the number of hold assets with its' current price), and the information of trading sessions (opening, continuous trading or closing) as well the price of opening and closing.

Agents may play roles of either buyer or seller, on both roles with different trading objectives. Agents directly initiate transactions. An agent is represented as an object with either private or public attributes.

The *Bank* component represents all intermediaries; it maintains information exchange between buyers and sellers. At the same time, Bank can be considered as a special type of buyer or seller, that has unlimited wealth, hence take active part in stock trading. We propose to consider Bank as a trading and intermediary agent.

Artificial Economic World provides external information about respective corporate developments, dividends and coupons, tax policies modification and so on. This information may influence agent's decisions. The artificial market architecture (system elements and interaction between them) is presented in figure 2.3.

We now turn to a technical question, which is crucial in ABM development: how time is handled in our platform.

2.3.1.2 Time handling and scheduling system

A crucial question for the design of distributed systems is the way one deals with time. There are two aspects for this problem: the modeling choice (se-

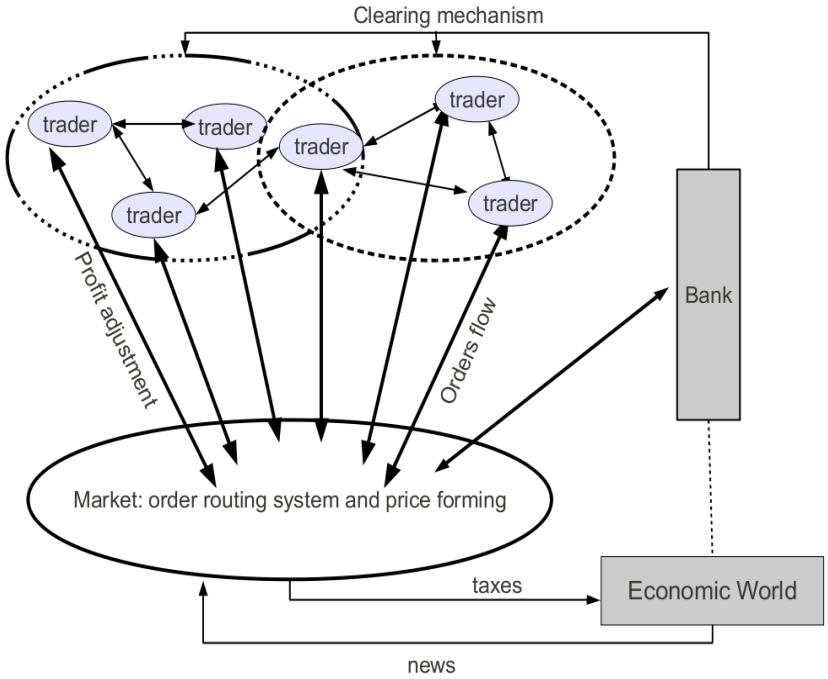


Figure 2.3: Market organizations and interactions

quential vs. parallel evaluation of agents) and the architecture choice (single stream or multiple streams processes). In ATOM, the scheduling system is parametric, thus one can choose between four possible configurations (see figure 2.4). In each case, one can also decide if order issuing will be balanced among agents or not. These possibilities give the researcher a real power for answering a wide range of problems through an appropriate design for his experiments.

The main difference between the sequential and the parallel simulation principle consists in how the tick time is managed.
In MAS philosophy, one considers that time changes along environments changes:

- In the *sequential* mechanism, the tick time changes after each order sent by an agent. This means that an agent is always able to see other agents' decisions before making his own decision.

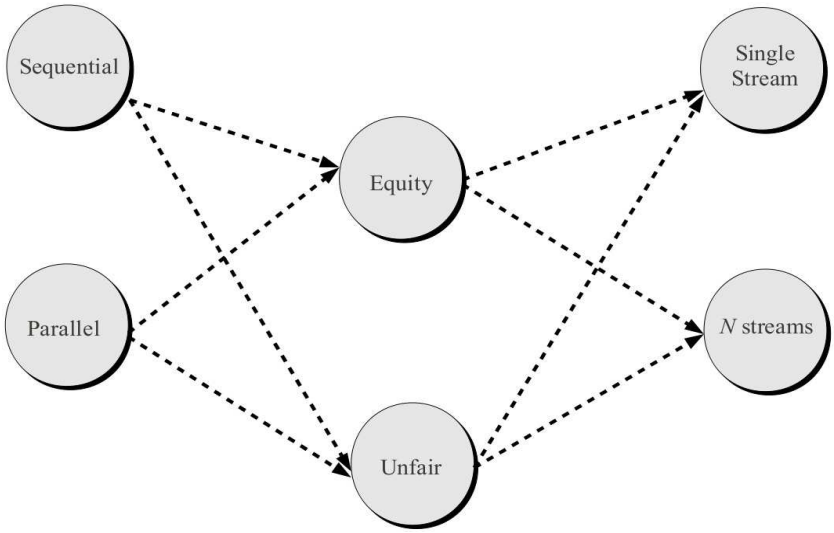


Figure 2.4: Modelling and architecture choices in time handling problem

– In the *parallel* mechanism, the tick time changes after one decision round processed over all agents. It is a way to simulate parallelism between agents. This is the principle followed by the Conway's game of Life (Garder, 1970).

However choosing one of these arrangements is just a modeling choice. Any of them can be obtained in a single stream architecture as well as in a multiple stream architecture.

In a single stream architecture, one needs a specific software engineering pattern to code the parallelism. The easiest way is to let the Market collect all the orders before their execution. ATOM uses a *"equitable round table"* for acting to ensure fairness among agents.

It is a way to simulate simultaneity in agents' decisions. If one needs to ensure fairness among agents, ATOM uses a loop to give the talk to all agents – *"equitable round table"*. An agent is allowed to act only once in each talk round. Of course, if one wants to depart from this fairness, it is sufficient to

pick randomly an agent and to offer him the possibility to decide again.

One feature of real stock market is that at each time step there is only a small fraction of the market population which is involved in trading. The assumption that all investors trade at each time step is unrealistic due to high transaction costs. Even professional fund managers realize only few trades per day in order to get closer to target portfolio weights. So it is important to control trading frequency of agents. Note that in ATOM an agent can decline the possibility to issue a new order, hence, even if we have a single stream, we can easily simulate different talking frequencies.

A possibility to express an intention does not necessarily imply that a new order is issued. Since agents are autonomous, they can evaluate their positions every round, modify trading rules according to new market conditions. Developing an agent that sends twice less orders than the others can be made by programming his behavior such as he will decline word on odd turns (keep unchanged position), while others accept to talk each time they have the possibility to do so.

The main advantage of the single stream architecture is that the designer can reproduce perfectly all the experiments. He keeps control on agent's talk. We consider that it is the best way to build and test experiments.

In a multiple stream architecture, parallelism is obvious, but the designer does not have control over the talking order of agents. This order is defined by the Operating System, and of course, it can produce biases in simulations. Nevertheless, one particularity of this approach is that the time is given in seconds – real time. It is also easy to express the different trading frequencies for different agents, similarly with what is described above. If there is no synchronization mechanism between the streams, the simulation is unfair, an agent can talk twice more than another one. In a fair simulation, one just has to put a synchronization pattern like a "Cyclic Barrier" to grant this property. This architecture is preferable if one wants to include humans in

the loop. This statement is detailed in the subsection 2.3.2.1

As ATOM is a multi-asset artificial market platform, we have implemented a *"one order for one book"* rule: during a talk round, agents are just allowed to send at most one single order to a given order book (i.e. one order at most per stock) within the same "round table discussion". This principle helps keep fairness in agents actions. However, note that agents have the possibility to send several orders within the same "round table discussion" to several order books: this ability is simply constrained by the *"one order for one book"* rule. If the ASM is settled such as it runs a multi-stock experiment, an agent can therefore rebalance his portfolio using one order per category of stocks he holds. The proper system scheduler provides this possibility.

It is necessary to stress again that ATOM can govern all combinations between sequential and parallel mechanisms, equity or unfairness in agents' actions, one stream or multiple streams processes. The combination of single stream, parallel mechanism, equity in actions is used for most of the experiments concerning financial problems. Parallel mechanism, multiple streams processes are used to allow the human investor to trade in equitable conditions with artificial agents within the platform.

2.3.2 Agents

In this subsection we consider the trading strategies implemented in ATOM. In ATOM, there exists a large variety of agents' trading strategies, from Zero Intelligence Traders (ZIT), whose behavior is merely based on stochastic choices, to complex Sophisticated Intelligence Traders (SIT) with memory, information analysis, and expectations.

In ATOM, artificial traders are characterized by their a) available set of *actions* (buy, sell, or remain positions) and possibility to switch between these activities (from buyer to seller) b) *decision making rules*: the simplest example

of decision making is when buyers cannot buy if he does not have enough cash for transaction c) *scheduling* of actions: how often an agent is able to send orders in response to market request, some agent participate one time per week, while others trade every minute d) *information* consideration: agent requires specific information from the market or external world, he can also share the information with others e) possibility to describe current *state* of each trader: number of held assets, available cash, or budget constrains. In such a way, each agent is represented by it's own object with different number of stocks, memory span, risk preferences, trading frequency, utility, target weights. The platform allows an implementation homogeneous as well as heterogeneous population depending on the purposes of research.

Zero Intelligence Traders (ZIT). This behavior is merely based on stochastic choices: there are equal possibilities to send ask or bid order, ZITs do not observe and do not ask any information to set up prices and quantities, that are random variable. Concerning scheduling, such traders respond to every market request. This kind of behavior has been popularized in economics by Gode and Sunder (1993). The discussion introduced in section 2.1.3 shows that despite their extreme simplicity, ZIT agents are widely used because more sophisticated forms of rationality appear to be useless to explain the emergence of the main financial stylized facts at the intraday level. The possible calibrations and applications of ZITs in ATOM are introduced in section 3.

Technical Traders. Technical analysis is an important tool for decision making in investment. Technical traders is a trend followers making decisions depending on the trend of past prices. In the literature this type of traders is usually referred as noise or chartist traders (Brianzoni, Mammana and E.Michetti, 2010; Ponta *et al.*, 2011; Brianzoni, 2012). As shown in section 2.1.3, these agents try to identify patterns in past prices (using charts or statistical signals) that could be used to predict future prices and henceforth

send appropriate orders. One can find an example of such behavior in Arthur (1994c). From a software engineering perspective, these agents need to have some feedback from the market and some kind of learning process as well (reinforcement learning for large sets of rules is generally used). At the same time, technical traders ignore the actual nature of the company, currency or commodity. This leads to some complex algorithmic issues. For example, if one considers a population of a few thousand Technical Traders, it is highly desirable to avoid that each agent computes the same indicators, or simply stores the whole price series.

In ATOM, the agents are given the possibility to use Moving Average Crossovers, Exponential Moving Average (EMA), Momentum, Relative Strength Index (RSI) and Period techniques to identify when the market is overbought or oversold, and to generate Ask and Bid orders.

Sophisticated Intelligence Traders (SIT). *Cognitive Agents* generally have a full artificial intelligence, such as memory, information analysis processes, expectations, strategies and learning capacities. For example, an agent buying at a specific price and sending immediately a "stop order" to short her position if the price drops under θ% times the current price, will fall in this category. Agents using strategic order splitting (see for example Tkatch and Alam (2009)) or exploiting sophisticated strategies (for instance, Brandouy, Mathieu and Veryzhenko (2009)) can also be considered as Cognitive Agents. The other example of agents belonging to this group is the traders that allocate their funds across different risky assets and riskless assets according to their risk preferences. These agents at each time step confront two decisions: how much of their wealth to invest and how mush to save in cash. Usually, they are guided by their utility functions.

ATOM contains agents characterized by different utility functions (all mentioned in the section 2.1.3): Constant Relative Risk Aversion (CRRA), Constant Absolute Risk Aversion (CARA), Logarithmic Utility Function,

Quadratic Utility Function.

We have also implemented the agents that try to minimize risk for a given target return following the mean-variance optimization rules introduced by Markowitz (1952). This type of agents is called *mean-variance* traders (or optimizers). The features of this trading strategy and its implementation for portfolio optimization issues investigation is described in detail in the section 4.1.

2.3.2.1 Human in the loop

Nowadays, software agents are commonly used to replace human traders in making decision and taking action in the electronic trading. Thus, advanced software platforms should combine artificial agents and human. ATOM can include human-beings in the simulation loop. This is an important feature that is seldom offered in multi-agent artificial stock markets, if simply possible with respect to the algorithmic choices made in other platforms. A human agent is an interface allowing for human-machine interaction. Through this interface one can create and send orders. Notice that human agents do not have any artificial intelligence: they just embed human intelligence in a formalism that is accepted by the system.
To allow the introduction of humans in the loop, ATOM has been designed to deal with communications over the network. Human agents can be run on different machines and the system allows client-server configurations. This approach is particularly fruitful for a pedagogic use of the platform during finance class for example. In this latter case, several students have their own trading interface on their computers. In other terms, each of them runs a human agent linked to the ATOM server through the network. However, the presence of human agents does not alter the way the scheduler operates.

Two kinds of human agents can co-exist in ATOM: Modal Human Agents

(MHA) and Non Modal Human agent (NMH). MHA can stop the scheduling system. As long as a human-entity does not express her intentions (to issue a new order or to stay unchanged), the simulation is temporary frozen. In a classroom, this aspect is particularly important and leaves time for students to estimate current position and to make decision.

NMH cannot freeze the simulation, which means that human agents compete in real time with artificial traders. Even if human agents can have a hard time in this situation, it remains realistic in a financial world where algorithmic trading is more and more frequent.

2.3.3 Validation tests

As mentioned previously, every artificial stock market should succeed in processing a given order flow collected from a real-world stock market at a specific date. The conclusion about system validity is made by comparing prices delivered by the real stock market and prices generated by artificial one using the same set of orders. It should also generate relevant "stylized facts" with regard to their real-world counterpart: these stylized facts are statistical characteristics of financial time series that prove to be systematically observed in various contexts (different assets, periods of time, countries).

This section presents how ATOM fulfills these requirements. Moreover, we also introduce here a series of performance tests.

ATOM reality-check

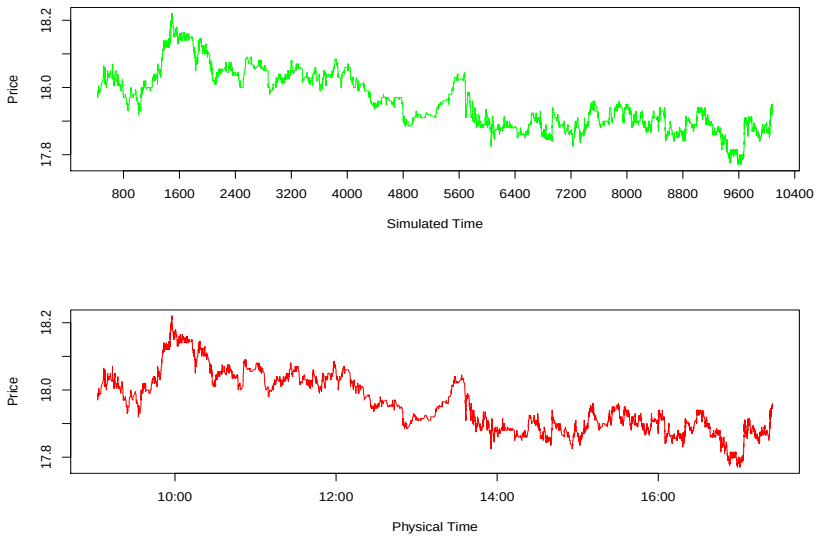
In this section, we report a series of tests conducted to check whether ATOM can generate financial dynamics in line with those of the Euronext-NYSE stock-exchange or not. The first series of test is devoted to the ability of ATOM to generate unbiased prices when it deals with a real-world order-flow. We check whether the agents are able to re-processes real orders submitted to

the market on a given day with the prices similar to those of real market. Figure 2.5 illustrates such an experiment with a set of 83616 real-world orders of the French blue-chip France Telecom (FTE) submitted to the NYSE-Euronext market on June 26th 2008 between H9.02'.14".813"' and H17.24'.59".917"'. Figure 2.5(a) presents the price series produced with ATOM and the real-order flow while Figure 2.5(b) presents the corresponding exchanged volumes. In each of these figures, two set of data are plotted. The upper set corresponds to the series generated by ATOM in processing the real-world order flow. The bottom part displays those actually observed on NYSE-Euronext market. One clearly sees that the replay-engine included in our artificial platform can process the real-world order flow in the same way it is treated by the NYSE-Euronext engine. Handling time in simulations is particularly complex and may lead to unsolvable dilemma. We cannot guarantee an exact matching of waiting times but rather a coherent distribution of these values delivered by the simulator engine with regard to the observed waiting times.

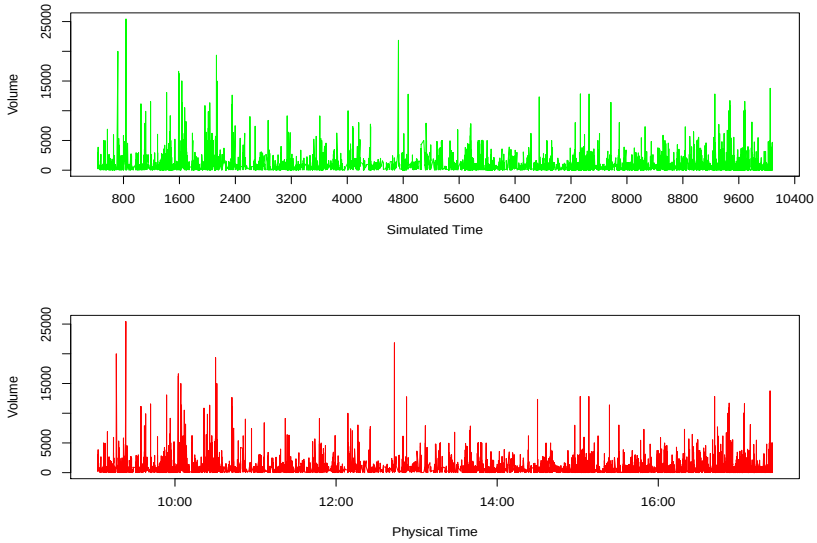
Notice that ATOM performs rather decently in satisfying the first reality check procedure.

Stylized facts

The second test focuses on ATOM's ability to generate realistic artificial prices when populated by artificial agents. We run a series of simulations to verify if ATOM can generate major stylized facts that are usually reported in the literature (Cont, 2001). In this subsection, we report only the classical departure from Normality of asset returns at the intraday level (Figures 2.6(a) and 2.6(b)). Real data are those used previously for the reality check, artificial data is generated using a population of ZITs. ATOM produces qualitative and quantitative stylized facts in line with real market data, that is quite difficult task for most of artificial market platforms (Veryzhenko *et al.*, 2010). Even if

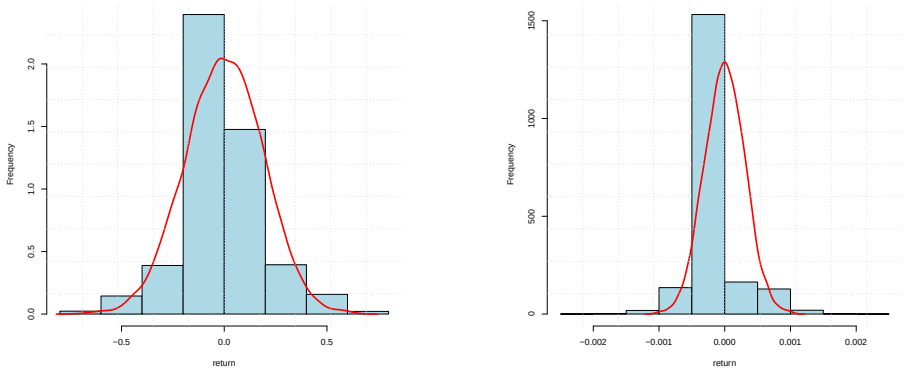


(a) Prices : ATOM vs. Euronext-NYSE



(b) Volumes : ATOM vs. Euronext-NYSE

Figure 2.5: Results of the "Reality Check" procedure. The population of artificial "hollow" agents treats a set of 83616 real-world orders of the French blue-chip France Telecom (FTE) submitted to the NYSE-Euronext market on June 26th 2008 between H9.02'.14".813"' and H17.24'.59".917"'



(a) Departure from Normality, ATOM

(b) Departure from Normality, Euronext-NYSE

Figure 2.6: Stylized facts, ATOM vs. Euronext-NYSE. ATOM returns are generated by Zero Intelligence Traders (ZIT) (see subsection 2.3.2). Real data correspond to France Telecom (FTE) prices settled to the NYSE-Euronext on June 26th 2008)

some artificial markets are able to reproduce the main stylized facts such as the non Gaussian return distribution or volatility clustering, the corresponding quantitative characteristics (basic statistics) do not fit real ones. ATOM can be easily calibrated to match specific quantitative market features (moments). This calibration facility is described in detail in the paper Veryzhenko *et al.* (2010). Additionally, section 4.1 puts forward ATOM's ability to perform realistic price dynamics in the multi-asset framework.

Performance test

We run several experiments to demonstrate running time for realistic price series generation and existing order-flow execution.

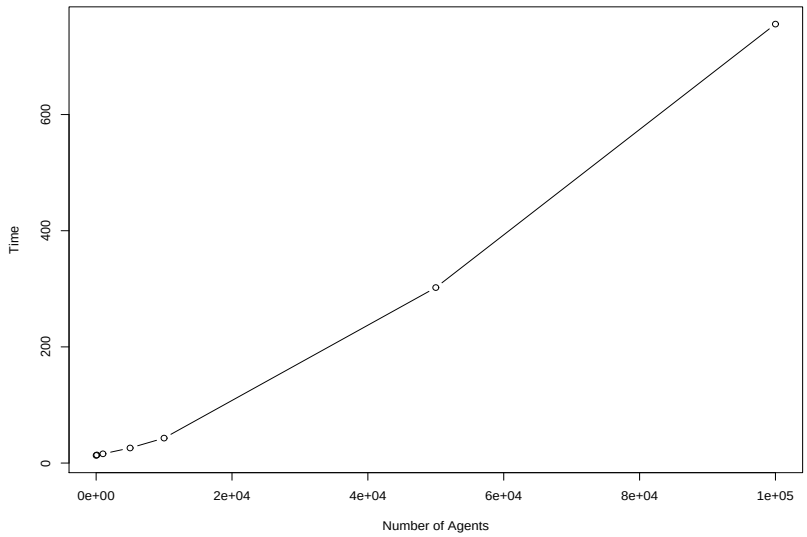
To demonstrate the ATOM price setting ability, we use a group of heterogeneous agents. The population consists of Zero Intelligence Traders (ZIT) and Technical Simple Moving Average Traders (in equal proportions), described in section 2.3.2. The number of set prices is 10^5 (on the Euronext Stock

Exchange the number of fixed prices for different stocks varies from 1000 to 5000 per day). The number of agents varies from 10 to 10^5. The results are reported in figure 2.7(a). It takes about 12 minutes to run 10^5 agents for price setting.

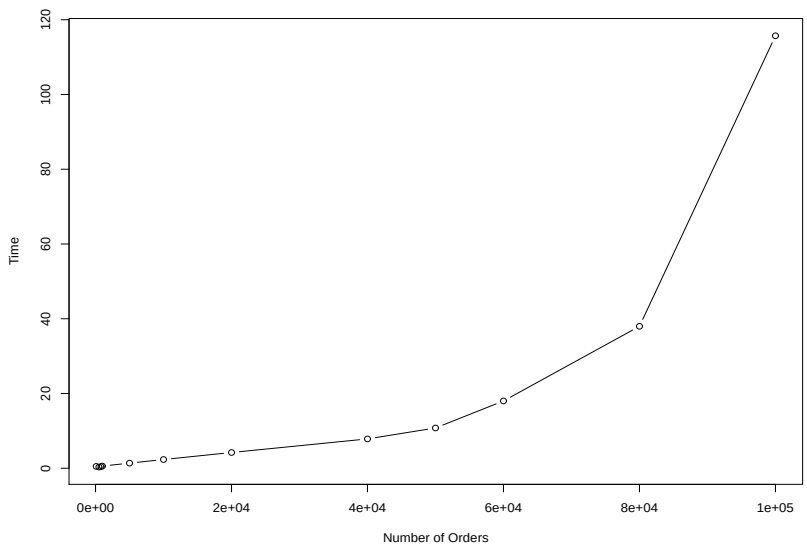
To test the running time of the replaying engine, we use real market order flow. The same population of agents is used to read all variety of orders (limit, market, stop-limit, iceberg, etc.) and send them to the order book. It is up to the market to set prices in a proper way (according to a price setting protocol). The number of orders varies from 100 to 10^5. It takes 2 minutes to replay 10^5 orders (see figure 2.7(b)).

2.4 Conclusion

This chapter has given an overview of the current state of research in agent-based computational finance along with some ideas concerning the design and construction of working simulations. We have stressed that the development of artificial stock market platforms puts forward a series of complex issues in terms of computer science. Probably, the most important question is the design of the economic environment itself. What type of market microstructure, quote-driven or order-driven, will be implemented? What types of securities will be traded? Will there be some kind of fundamental value? How information will be presented and how the agents process it? This chapter has provided some answers to these questions. We have started with some insight concerning the construction and design of trading environment (market mechanisms, number of assets, types of trading securities, types of orders, trading sessions etc). We have also focused on modeling of traders' behavior. The aspects that differentiate investors are: investment objectives, investment constraints, attitude to risk, investment strategy, portfolio maintenance, trading frequency, and memory span. We have introduced some existing plat-



(a) Price fixing time. The number of agents varies for 10 to 10^5. The agents set 10^5 prices.



(b) Orders flow execution time. 10 agents. The number of sent orders varies from 100 to 10^5.

Figure 2.7: Results of the performance testing

forms. It has been impractical to mention all or even most existing models, we have focused on the pioneering and the most successful efforts to design and develop artificial stock markets. Finally, we have introduced the ATOM platform, designed to deal with market dynamics and agents' decision making in an absolutely different way, compared to equilibrium representative agent models. This artificial stock market generates the necessary and reasonably realistic market dynamics that allow us to use it as a testbed for evaluating of trading strategies in the next chapter.

Part II

Results and Research Contributions

Chapter 3

Minimal market calibration for realistic market simulation

Contents

In this chapter we investigate the question of the sophisticated level of agents' behavior and intelligence to obtain with realistic market microstructure both qualitative and quantitative stylized facts. We show that qualitative stylized-facts can be generated with ZIT, but they are without any quantitative predictive power. In this chapter we report that at coarse grain, in most of the cases, such qualitative stylized facts hide unrealistic price dynamics at the intraday level and ill-calibrated return processes as well. Generating "realistic" financial dynamics that reproduce quantitatively financial distributions is out-of-reach within the pure ZIT framework.

In addition we show that even with highly constrained ZIT agents, one cannot reproduce real time series. Except in a few cases, none of the first order moments of ZITs versus real data will be equal. We therefore claim that stylized facts produced by means of ZIT agents are useless for financial engineering.

Our results lie in the strand of literature of ZIT, which we review in section 3.1. We are interested not only in coarse grain, qualitative empirical regularities, but also in the actual ability of ZIT at generating quantitatively acceptable stylized facts. For that purpose, we produce several families of ZIT agents (introduced in section 3.2) similar to those found in the literature, but calibrated using real market data. Section 3.3 puts forwards an introductory case study, presents the core results and proposes a sensitivity analysis of the latter.

3.1 Literature survey

Stylized facts are difficult to explain by the mainstream theory and the effort in empirical research to describe data lacks of a convenient theoretical foundation of these facts. However, an alternative approach has emerged from econophysics which described the same financial facts as scaling laws. Indeed, this approach considers that physical systems which consist of a large number of interacting particles obey universal scaling laws that are independent of the microscopic details and that economics could be considered in the same way as Amaral, Cizeau, Gopikrishnan, Liu, Meyer, Pend and Stanley (1999). The approach seems to imply that the rational individual choice seems not important in explaining facts. But, as Lux (2009) points out, it might be the heterogeneity of market participants together with a few basic principles of interactions which may exert a dominating influence on the macroscopic

market behavior in more or less the same way in different institutional settings. Indeed a large literature has emerged which aims at discerning among two major explanations: market micro-structure and agents' behavior and heterogeneity.

Papers in favor of the explanation by the ***microstructure*** initiate with the research on Zero Intelligent Traders (here-after ZIT) of Gode and Sunder (1993) who show that traders acting randomly but within a budget constraint act comparably well in terms of convergence to equilibrium price and efficiency as compared to human subjects in experimental economics (see Smith (1962)). Dave Cliff (1997) show that for special demand or supply function this may not be the case and a slightly more complex behavioral assumption is required to achieve equilibrium in continuous double auction markets. Ladley and Schenk-Hoppe (2009) in a similar framework as Gode and Sunder's one, but with a constant flux of traders entering the market, reproduce price movements and show that aspects of the order book such as size of spreads and conditional probabilities of order submissions can be obtained by the interplay of ZITs and the book. The observed frequency of the different order types submitted seems however related to strategic behavior based on the observed book state.

Maslov (2000) introduces a model where traders randomly choose to trade either at the market price or by placing a limit order. Maslov shows that fat tails, long range correlation in the volatility and non-trivial Hurst exponents arise in such framework. One paper by Farmer *et al.* (2005) shows that a simple ZIT model working as a continuous double auction with both market and limit orders predicts well bid-ask spreads, price diffusion rates and market impact function related to the supply and demand of 11 stocks in the London Stock Exchange. Their conclusion is that the price formation mechanism strongly constrains the market, playing a more important role than strategic behavior. They adjust their model to real data by making simple assumptions about order placement, canceling process of limit orders and ticks (price

increments). These quantities are estimated on a daily basis from the real stocks and serve to make predictions. Nevertheless, their framework exhibits odd assumptions such as an order issuing based on log prices rather than raw prices and various Poisson laws.

Other studies have shown that stylized facts can be reproduced by the ***behavioral*** assumptions of agents and heterogeneity of behavior. Most of these models (LeBaron *et al.*, 1999; Hommes, 2006; Lux, 2009) suggest that the aggregation of simple interactions at the micro level leads to complex non linear behavior at the macro level. Typically, the heterogeneity of behavior is due to different types of rationality (informational and computational) and heterogeneity in preferences as well.

The experimental economics literature also tackles some issues concerning stylized facts. Early experiments have already shown how easily it is to reproduce bubbles in double auction markets (Smith, Suchanek and Williams, 1988). Bubbles are resilient to market conditions such as short selling , margin buying opportunities, limit price-change rules, informed insider trading (King, Smith, Williams and Bening, 1993).

Plott and Sunder (1982) are the first to report some stylized facts within the lab without however providing any explanation about it. They show excess kurtosis and lack of autocorrelation of returns in prices. Kirchler M. (2007) presents an explanation based on asymmetric information between traders in a double auction markets. Typically they introduce a market characterized by heterogeneous traders concerning information. They show that the heterogeneous fundamental information is the source of fat tails and absolute returns whereas higher noise trading (trading not based on fundamentals) does not explain absolute returns.

To the best of our knowledge, only one paper (Liu, Gregor and Yang, 2008) tries to reconcile two streams of literature by suggesting that different elements can be at the heart of stylized facts emergence in different time hori-

zons. Liu *et al.* (2008) show that a market microstructure and zero intelligent agents are responsible for reproducing leptokurtosis and heavy-tailed distributions, autocorrelation and excess volatility for intradays data. However,the microstructure is only responsible for excess volatility in daily returns. Therefore, behavioral assumptions are required to explain other facts.

3.2 Simulation methodology

In this section we introduce four different types of ZITs. For the sake of possible replication of the results presented in this research, the pseudo-code describing each agent is available in Appendix A.2.

3.2.1 Calibration elements: agent's behavior

The most basic ZITs we use (called hereafter "Unconstrained ZIT") are directly inspired by the work of Maslov (2000). From this starting point, we progressively add constraints to their allowed behavior in accordance with the real market data with the aim of reproducing quantitatively some stylized facts. Notice that we denote, in the following developments, by capital letters all real market data (P, V, R, for example *resp* prices, volumes and returns) and by small letters all simulated data.

Except when a cancel instruction is issued, the trading activity consists in sending to the market an order made of a direction, a price limit (except when this order is a market order), and a quantity. The common characteristics for all of our ZITs, whatever their level of constraints, are as follows:

- The proportion between different order types is 80% of limit orders, 15% of market orders and 5% of cancel orders which reflects realistic characteristics of real markets. These figures were obtained from a data set of real order flows gathering around 36000 observations of intraday

trading (courtesy of Calyon SA, here-after "CDS", for Calyon Data Set). They may vary from a given stock to another and across categories of investors (Foucault, 1999; Handa and Schwartz, 2006). Furthermore, this proportion is modified in the sensitivity analysis (see page 127).

- Each agent can submit both orders, Buy and Sell.
- Buy and Sell orders arise with equal probability ($p = 0.5$).
- A single asset is traded.
- Budget constraints are implemented: agents cannot make a trade that yields a negative profit, *i.e.*, buyers cannot buy at a price higher than their reservation value and sellers cannot sell for a price below their marginal cost.
- Within each ZIT family, we define two subcategories of agents with respect to the real market average volume observed for a given stock. *"Big fishes"* draw volumes between the mean and the maximum real volume , while *"Small fishes"*, draw this value between the minimum and the mean real volume. The reason for this choice is that it may generate a realistic picture of contemporary markets, where, in a typical experiment, the ratio of big to small fishes is 1 over 5. This proportion is also obtained from the "CDS".

Note that all the parameters chosen for the initial settings of the simulations are calculated based on real market data. These parameters are inferred for each stock on each day.

In our simulations, we denominate ZIT artificial agents that mostly use random number generators to determine prices and volumes in their orders. These agents are more or less sophisticated but none of them use artificial

intelligence methods such as classifiers or learning mechanisms to adapt their behavior and/or to evolve. We thus do not claim to investigate all possible extensions of ZITs but a series of models that are common in the literature.

- *Unconstrained ZIT ($UZIT$) with two types of price-generating processes:*
  - *Uniform price distribution ($UZIT_U$,* see algorithm 7, page 244*):* p_t is drawn from a Uniform distribution in $[P_{min}, P_{max}]$ where, for a reminder of the convention, P_{min} and P_{max} are the minimum, respectively maximum price observed on the real market. These values are determined at the beginning of trading day.
  - *Normal price distribution ($UZIT_N$, see algorithm 8, page 245):* p_t is drawn from a Normal $N(P_{mean}, P_{sd})$.
- *Statistically calibrated ZIT ($SZIT$, see algorithm 9, page 245)* are kind of bounded $UZIT_U$, meaning that i) they still perform a random draw from a Uniform distribution ii) the price range is limited by $[P_{min}, P_{max}]$ and iii) the range of admissible price is different between Sellers and Buyers, since we took for Sellers a simulated range for ask prices a_{min} and a_{max} and for bid prices b_{min} and b_{max}. These simulated boundaries are obtained from the "CDS" as mentioned previously. To go into the details, we first separate Bids and Asks and then find for each of these subsets the minimum and the maximum values observed on a given day. Similarly, the same procedure was applied to volume data.
- *Trend calibrated agents ($TZIT$, see algorithm 10, 246)* are $SZIT$ with the following additional feature: when they issue a new order, they pick a price that is formed using two additional parameters γ_t and δ. γ_t is geared at reproducing the tendency of a given series. δ_t generates some additional randomness. More precisely, $TZIT$ agents are modeled in the following way:

- We divide the day into n sub-intervals. The number n influences the accuracy of results. The larger the number of sub-intervals the greater the fitting accuracy of simulations.
- We compute min and max prices within each sub-intervals
- We choose the time period t (by fixing the parameter θ) and apply equations (3.1) to (3.4) to estimate prices.

We next consider the following example: on a given period, one observes for a given stock a slow decay from a maximum to a minimum price. This slow decay can be described using equations (3.1) to (3.4).

$$\gamma_0 = 0 \tag{3.1}$$

$$\gamma_t = \gamma_{t-1} + (1/\theta) \text{ with } t \in [0, \theta] \tag{3.2}$$

$$\delta_t \sim logN(0, 1) \tag{3.3}$$

$$P_t = P_{max} + (P_{min} - P_{max}) \times \gamma_t \times \delta_t \tag{3.4}$$

Note that in equation 3.2, γ_t is an increasing step function from zero to one with step size equal to θ. θ allows to track the underlying trend with a certain level of accuracy. The bigger θ, the more accurate this fitting.

Thus, the behavior of these agents necessitates to specify within a given day n sub-periods and, for each of these sub periods, the maximum and the minimum price. The precision in this procedure can be tuned so as to track more or less closely the underlying dynamics. $n = 50$ is the value arbitrarily chosen for the experiments. Note that equation 3.4 implies that the first price set by this category of agents is close to P_{max}, for this example of a slow decay in prices.

Based on these information, agents can track the global price tendency

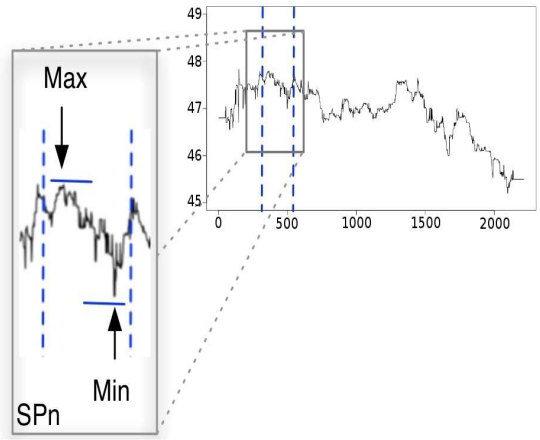


Figure 3.1: Example of calibration procedure, Renault SA. SP_n: Subperiod "n"

so as to algorithmically grasp the underlying dynamics (see Figure 3.1). So to speak, we endow TZITs with limited foreseeing capabilities in the very short run.

- Inspired by the procedure introduced by Farmer *et al.* (2005), the fourth category of ZIT agents is characterized by their relative aggressiveness *(AZIT, see algorithm 11, 247)*:

  - Patient agents ($AZIT_P$): send limit buy orders with prices drawn from a uniform distribution between zero and a_{min} and limit sell orders from b_{max} to ∞. a_{min} and b_{max} are inferred as described for $SZIT$. These agents are parameterized in such a way that they cannot trade one with other. Note that in Farmer *et al.* (2005) log prices and not prices where drawn from a Uniform distribution. Trades cannot occur within this population. The quantity is determined randomly using a uniform distribution between two integers

$min(V)$ and $max(V)$.

- Impatient ZIT ($AZIT_I$): send market orders using the same distribution for quantities as Patient ZIT.

There are 85% of $AZIT_P$ and 15% of $AZIT_I$ in a typical simulation, these figures corresponding roughly to the proportion of market *vs.* limit orders in real markets. The quantity posted by $AZIT_P$ is twice the quantity posted by $AZIT_I$.

3.3 Empirical design and results

We introduce a case study to illustrate the main concepts and statistics used in the empirical part of this research and then generalize our results.

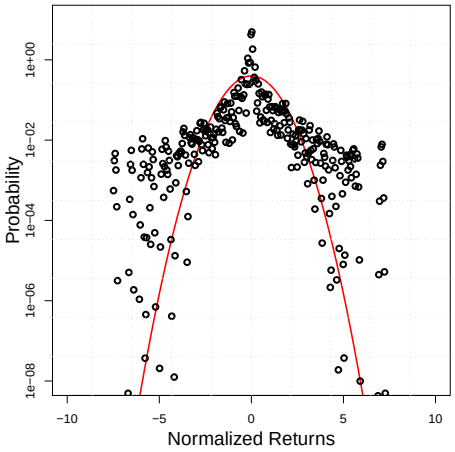
3.3.1 An introductory case study

In this section, we focus on a single stock (Renault SA) for a single day (August 1, 2002) and compare simulated *vs.* real data in terms of price dynamics and stylized facts. It is clear that the return series coming from Renault S.A. prices is not normal. Moreover, it exhibits fat tails and the autocorrelation in absolute returns decays slowly in log-log scale (see Figures 3.2(a) to 3.2(d)).

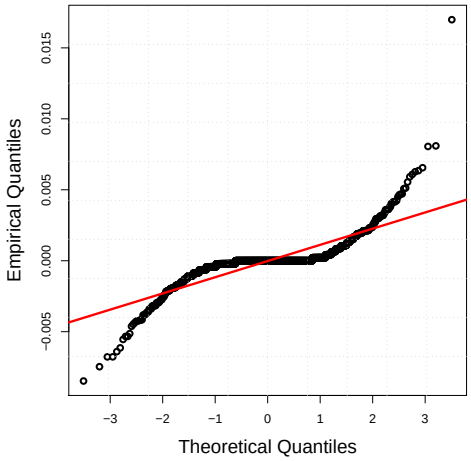
To assess the Volume-Volatility relationship we use the following, simplified framework:

1. We first slice the series in non overlapping time windows containing 300 observations. Depending upon the length of the series under investigation, the number of time windows may vary slightly.

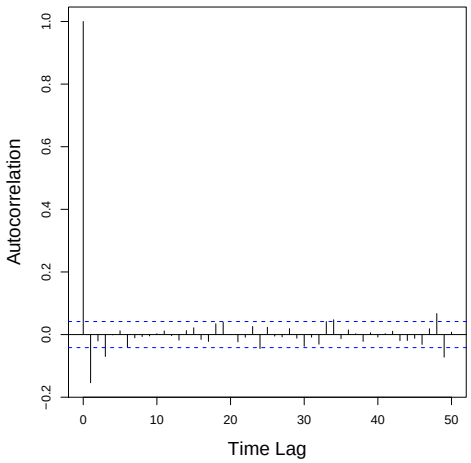
2. We calculate for each of these slices and each artificial series produced by the agents, two indicators: the mean volume over this period (mv_t) and the standard deviation of these returns (sd_t).



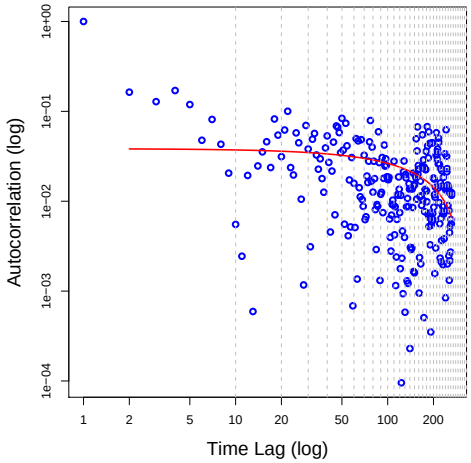
(a) Fat Tails (Normalized Returns)



(b) QQ-Plot

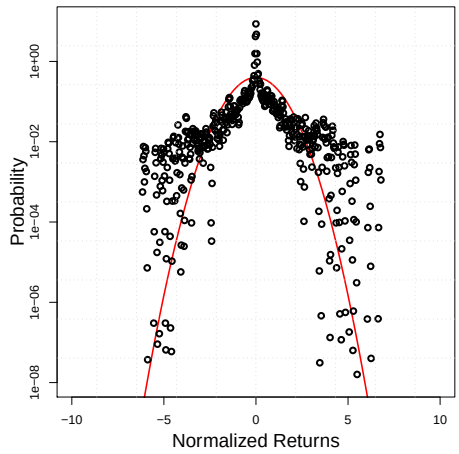


(c) Auto-correlations, Raw returns

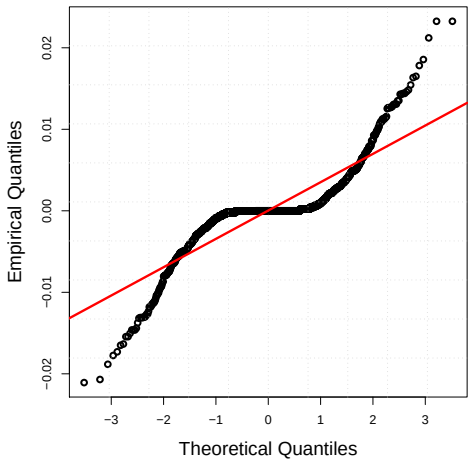


(d) Auto-correlations, Absolute returns (log-log scale)

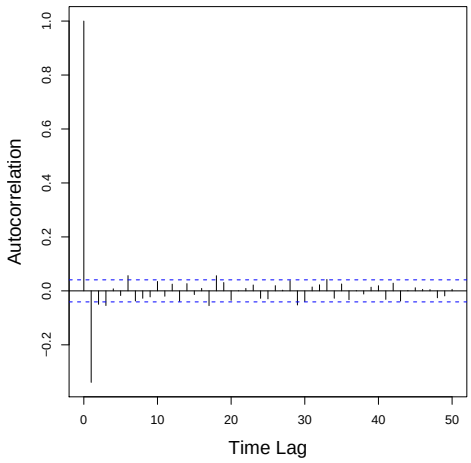
Figure 3.2: Renault SA intraday returns, departure from Normality. Intraday data on August 1, 2002



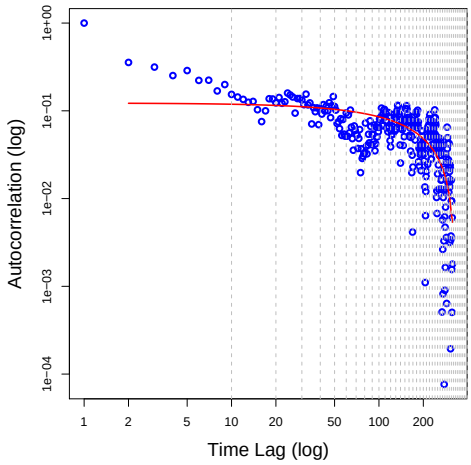
(a) Fat Tails (Normalized Returns)



(b) QQ-Plot

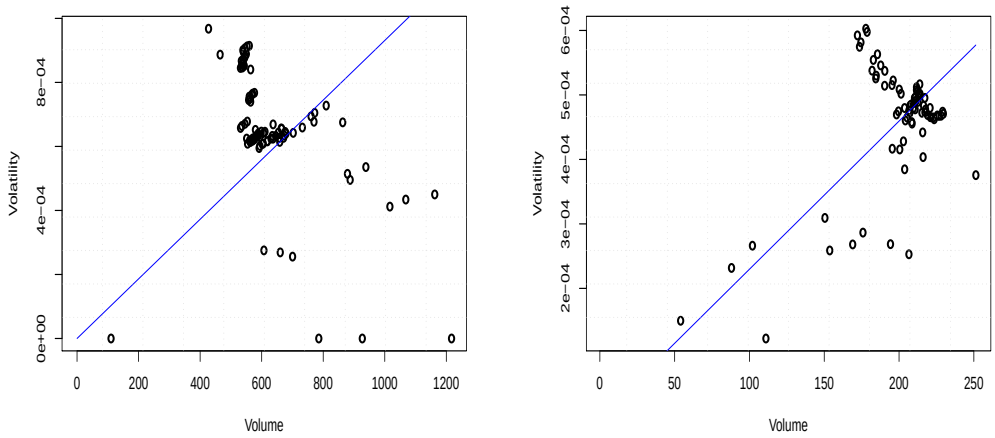


(c) Auto-correlations, Raw returns



(d) Auto-correlations, Absolute returns (log-log scale)

Figure 3.3: Qualitative Stylized facts, Uncalibrated ATOM ZIT ($UZIT_U$). ATOM generated intraday returns with calibration to Renault SA on August 1, 2002. Experiment settings: $P_{min} = 45.2$, $P_{max} = 47.83$, $V_{min} = 1$, $V_{max} = 8075$



(a) Renault S.A (b) $UZIT_U$

Figure 3.4: Intraday volume–volatility relationship

3. We then estimate a linear model:

$$sd'_t \sim \alpha + \hat{\phi} mv'_t \tag{3.5}$$

We use the value of $\hat{\phi}$ for comparing the volume volatility relationship between simulated and real data. In doing so, and for the sake of simplicity, we explicitly exclude from the analysis the value of the constant α (said in different terms, we consider this constant as similar for both models).

This process is used in this article whenever the volume-volatility relationship is investigated. For example, this relationship for Renault is presented in Figure 3.4(a). The value of the regression slope is 1.3086E-6. The same Figure, (see Figures 3.4(b)) is also produced using Unconstrained ZIT_U agents.

In Table 3.1 we report a series of univariate statistics illustrating some of the stylized facts analyzed in this research.

Even if several stylized facts are reproduced by $UZIT_U$, $TZIT$ or even

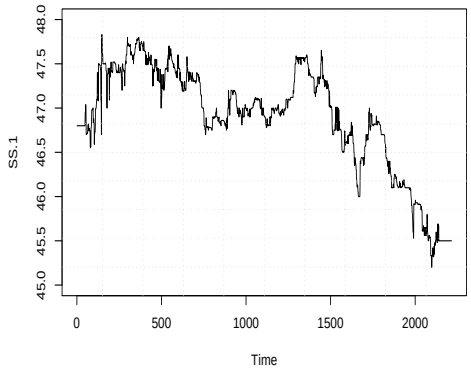
	Real	$UZIT_U$	$UZIT_N$	$SZIT_U$	$TZIT$	$AZIT$
nobs$^{(1)}$	2214	2302	2273	2222	2219	2005
Minimum$^{(2)}$	-0.0085	-0.0211	-0.0156	-0.0232	-0.0151	-0.0900
Maximum$^{(3)}$	0.0170	0.0233	0.0153	0.0279	0.0157	0.0900
Mean$^{(4)}$	-1.30E-05	6.00E-06	-2.00E-06	-8.00E-06	-1.40E-05	-2.64E-04
Stdev$^{(5)}$	0.0011	0.0035	0.0036	0.0050	0.0019	0.0163
Skewness$^{(6)}$	1.5192	0.1071	-0.0216	0.1130	0.1471	0.2351
Kurtosis$^{(7)}$	33.1606	10.0905	3.6706	5.2325	12.1035	9.8945

Table 3.1: Descriptive statistics. $^{(1)}$: number of observations; 2214 means for example 2214 returns for the given trading day. $^{(2)}$,$^{(3)}$,$^{(4)}$,$^{(5)}$,$^{(6)}$,$^{(7)}$: minimum, maximum, and moments for observed returns.

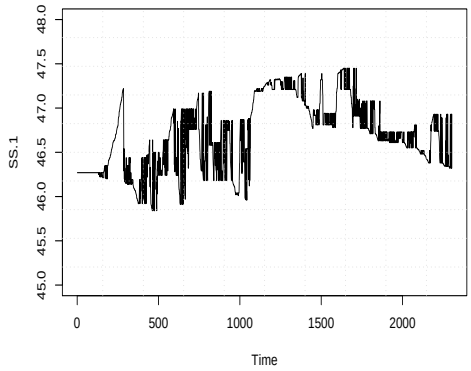
$AZIT$, one should also consider the underlying price dynamics from which these statistics are computed. In Figure 3.5, we report the prices coming from the real market (Subfigure 3.5(a)) and for the five families of ZIT (Subfigure 3.5(b) to 3.5(f)). It is clear that most of the ZIT families, beyond their apparent ability to generate congruent stylized facts, have a hard time at producing "realistic" price dynamics. The only "possible" exception, even if this can be discussed, is the $TZIT$ example that produces a price dynamics that looks like, at coarse grain, the real one.

In summary, we point out two shortcomings in the use of ZIT for financial markets simulations: i) The adequacy of stylized facts generated by ZIT with regard to the quantitative values of observed stylized facts coming from real data is questionable. ii) The underlying price motion remains most of the time totally unrealistic from a qualitative point o view.

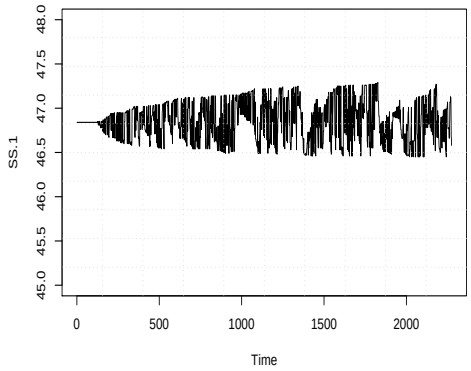
In the following section, we present a procedure enabling us to make a comparison between a set of 37 real data and different simulations from the five ZIT families described section 3.2.1.



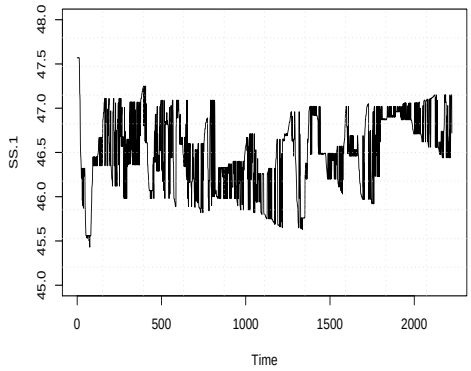
(a) Renault S.A., real



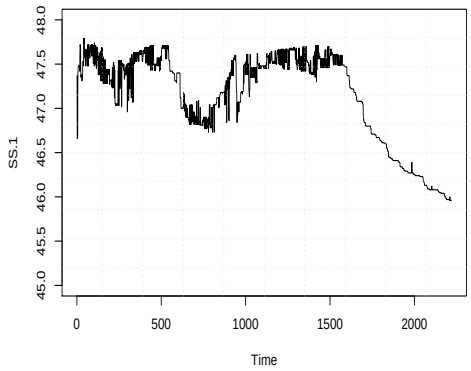
(b) $UZIT_U$



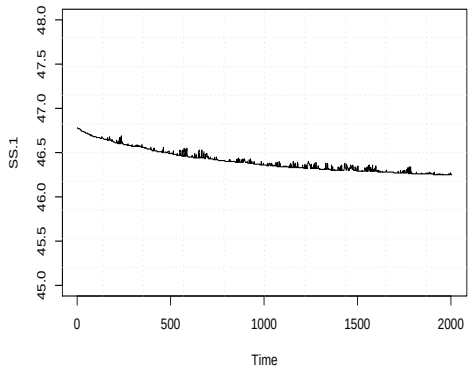
(c) $UZIT_N$



(d) $SZIT$



(e) $TZIT$



(f) $AZIT$

Figure 3.5: Intraday price dynamics, real *vs* simulated. ATOM generated intraday returns with calibration to Renault SA on August 1, 2002. Experiment settings: $P_{min} = 45.2$, $P_{max} = 47.83$, $V_{min} = 1$, $V_{max} = 8075$, $P_{mean} = 46.88915$, $P_{sd} = 0.6015639$, $a_{min} = 45.2$, $a_{max} = 50$, $b_{min} = 30.3$, $b_{max} = 47.83$, the number of prices is 2215, the number of sliced windows is 50.

3.3.2 Beyond the case study: *"zero is not enough"*

Our data consists in intraday prices collected from the Paris Euronext Stock Exchange covering 37 stocks in August 2002 (22 trading days). These stocks are components of the CAC 40 index and are therefore amongst the most traded within the French market. In table 3.2 we show a summary of the data including information regarding the trading activity (number of trades per stock family and exchanged volumes).

For each day and each of the $i = 37$ stocks in the sample we produce a set of 5 simulations with the 4 different ZIT families (2 uncalibrated ZITS, statistically calibrated, Trend ZITs, Agressive ZITs) described in section 3.2.1. Notice again that ZIT agents are calibrated using real values calculated from the sample.

For each simulation, we produce one concatenated return series based on the 22 simulated days. To avoid closing-to-opening jumps due to overnight information accumulation, we exclude returns that can be computed using closing prices at date t and opening prices at date $t + 1$. In other terms, the 22 days for each stock are summarized in $j = 1$ long time series. Note that the same concatenation procedure is run over the real dataset. We thus have $k = 6$ subsets, 1 from real data, 5 from simulated ones. The dataset is therefore equivalent to a $i = 37 \times j = 1 \times k = 6$ tensor.

For each of these simulations, we estimate, over the 37 observations, the distribution for the following statistics calculated on returns:

1. Mean

2. Standard deviation

3. Skewness

4. Kurtosis

	Prices				Volume			
Name	Mean	SD	Min	Max	Mean	SD	Min	Max
Accor	1587	447	938	2345	525	144	357	889
Air France-KLM	822	269	416	1426	332	101	182	623
Air Liquide	2279	736	1161	3685	150	29	112	251
Alcatel	6294	2178	3252	10488	1616	349	1059	2678
Axa	6440	1806	3394	9596	1583	231	1182	2107
BNP	4236	1041	2277	6387	848	122	639	1205
Bouygues	1411	368	728	1962	671	468	357	2719
Cap Gémini	1801	414	1087	3041	508	99	355	687
Carrefour	3782	1119	1849	5912	634	150	455	1090
Credit Agricole	2187	607	1108	3523	1116	1960	390	9828
Danone	2236	549	1208	3274	266	40	208	358
Dexia	1212	410	599	2290	1465	289	931	2009
EADS	1579	379	1165	2459	1051	284	722	1898
Essilor Int.	450	135	189	761	485	205	246	1052
France Télécom	7044	2271	3442	12554	1034	180	699	1524
L'Oréal	3068	814	1800	5628	437	83	288	602
Lafarge	1674	468	883	2388	397	347	221	1924
Lagardère Groupe	1024	231	578	1506	439	285	260	1671
LVMH	2706	852	1620	4942	572	254	392	1622
Michelin	1030	263	567	1666	555	522	284	2822
Pernod Ricard	498	167	260	885	333	258	136	1425
Peugeot S.A.	1712	497	1093	2921	605	381	338	2240
PPR	2518	1001	1180	6246	290	125	184	758
Publicis Groupe SA	770	215	464	1188	772	260	449	1639
Renault	1785	283	1408	2299	534	145	356	1084
Saint Gobain	3166	1089	1679	5363	720	166	477	1138
Sanofi Synthé labo	3023	771	1941	4718	594	101	434	760
Schneider	1361	438	597	2297	482	88	327	663
Société Générale	3342	922	1633	5311	560	115	417	795
STMicroelectronics	4160	1058	2060	6404	1185	161	861	1533
Suez Lyonnaise	4139	1165	1857	6051	812	77	635	993
Thomson CSF	730	180	362	1154	461	457	263	2488
Total	5269	1321	2613	7667	535	164	391	1230
Unibail-Rodamco	251	102	63	444	234	78	83	421
Veolia Environnement	1930	925	937	4355	898	262	510	1400
Vinci	1150	287	513	1605	309	88	198	482
Vivendi Universal	9400	5956	3755	23318	1220	269	803	2012

Table 3.2: Data: Univariate Statistics Summary

5. ρ_1 and ρ_2, the first two values for autocorrelation coefficients computed using raw returns

6. The slope of the decay function for autocorrelation coefficients calculated over absolute returns.

7. The value of $\hat{\phi}$ (see equation 3.5) indicating the direction and the strength of the volume volatility relationship.

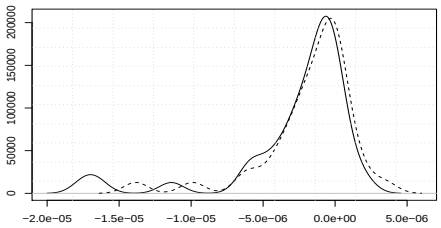
We then run two series of non-parametric tests using one simulation and real data as a benchmark to: i) test equality in population distribution (two-sample KS test). For two distributions D_1 and D_2, the null is that D_1 *and* D_2 *come from the same distribution.* ii) test equality in means (Fligner-Policello test, which is the equivalent to the Mann-Whitney test but without assuming equality in variance, and paired Wilcoxon test). For two distributions D_1 and D_2, the null is that D_1 *and* D_2 *have the same mean value.*

These tests are geared at appreciating whether the quantitative stylized facts are reproduced or not by means of ZITs.

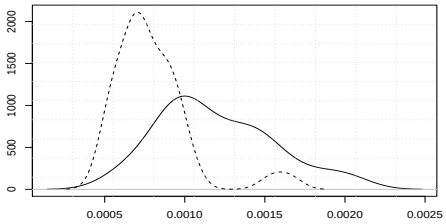
For illustration purposes, a limited example of the distribution over the 37 samples of each characteristic value for stylized facts calculated on $TZIT$ is plotted against the corresponding values calculated with real data in Figure 3.6.

We first report a series of non-parametric tests geared at examining whether the whole distribution of each stylized fact (summed-up with a single parameter) is similar to the distribution of the corresponding stylized facts calculated from the real world distributions.

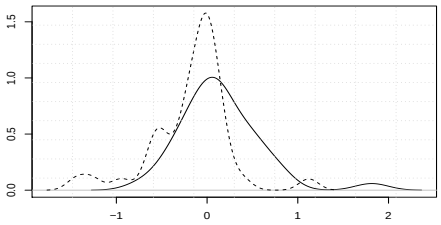
Results of the Kolmogorov-Smirnov test of equality in distribution is presented in Table 3.3. The results are rather unambiguous: except for the single case of the distribution of mean returns generated by Trend Calibrated ZIT ($TZIT$) (see figure 3.6(a)) and for the volume/volatility relationship for



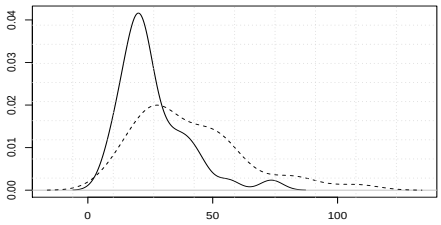
(a) Mean



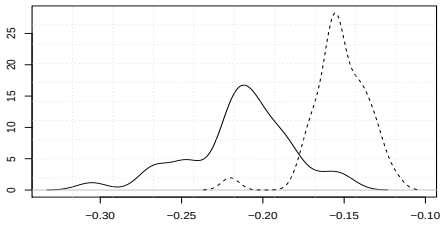
(b) SD



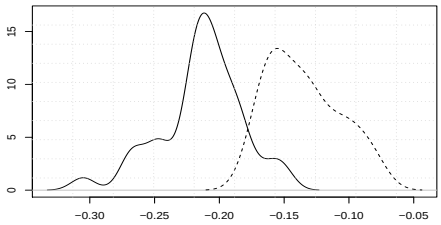
(c) Skewness



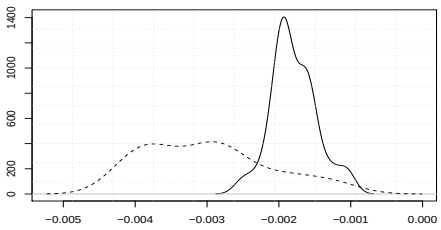
(d) Kurtosis



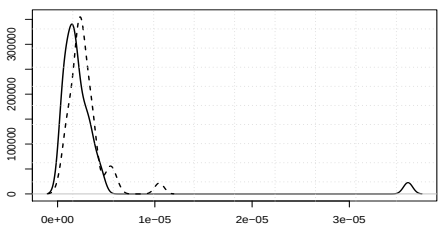
(e) ρ_1



(f) ρ_2



(g) Slope



(h) Vol/Vol

Figure 3.6: Stylized facts distributions for $TZIT$, real (solid lines) *vs* simulated (dash lines). The distribution is based on 37 statistics of ATOM simulated series (generated with 4 families of ZITs) and intraday concatenated data of 37 real stocks.

$UZIT_U$ and $UZIT_N$, the two-sample Kolmogorov-Smirnoff test can be rejected with high levels of confidence. The interpretation is straightforward: neither higher moments nor autocorrelation-based stylized facts can be matched by any of our ZIT families. However, unconstrained ZITs, probably due to their high level of freedom, generate a realistic relationship between the average volume traded and the resulting average volatility observed within the same time window. One can also notice that when ZITs are more and more constrained, this stylized fact, if still noticeable, does no longer fit the values of our benchmark.

If we restrict our attention on first order moments, and check equality of means for these moments between simulated and real data by the use of a Fligner-Policello test (see Fligner and Policello (1981)), we get the following results (see table 3.4).

Here again, the tests reject the ability of our ZIT families at reproducing quantitative stylized facts. The only cases where these tests cannot be rejected are those of $UZIT_U$, $UZIT_N$, $SZIT$ for the skewness distribution and for the value of the volume/volatility relationship, and $TZIT$ for the mean. Said differently, the first three categories of agents might do a relative good job at generating realistic third-order moments for the return distribution and, as mentioned previously, in delivering a realistic volume/volatility interplay. However, in these tables only the mean of each parameter distribution is tested against its real-world counterpart. If the test cannot be rejected the only think we can conclude is that the central values of the distribution are not so different. In our opinion, this is a necessary but not sufficient condition to accept a family of Agents. For example, if one considers the distribution of the Skewness for $UZIT_U$ against the real series, the equality of the means cannot be rejected (see Table 3.4 and 3.5) although it is clear that the whole distribution is different (see figure 3.7).

Rejection for $TZIT$ may be more surprising since they were constrained

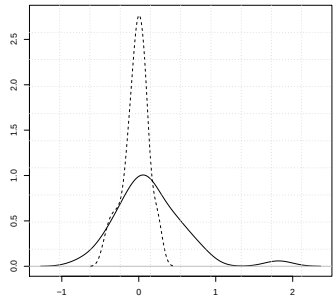


Figure 3.7: Distribution of Skewness for $UZIT_U$ –dash line– vs. Real World series –solid line–

to generate more realistic price dynamics, nevertheless, they were not able to reproduce first orders moments beyond the mean. Even if skewness is an important feature of financial distributions (notably important for asset managers), given the overall negative conclusions drawn on other moments and correlations of the distributions, this is a rather weak result.

If we go further in the analysis with a paired Wilcoxon rank test, $TZIT$ is now rejected while none of the skewness-related tests are rejected at the 5% threshold.

		Mean	SD	Skewness	Kurtosis	ρ_1	ρ_2	Slope	$\hat{\phi}$
$UZIT_U^a$	D	0.4865	0.9189	0.4324	0.4865	0.9459	1.0000	1.0000	0.2162
	p.value	0.0002	0.0000	0.0017	0.0002	0.0000	0.0000	0.0000	0.3565
$UZIT_N^b$	D	0.5676	0.6757	0.3784	0.4595	0.9459	1.0000	1.0000	0.1892
	p.value	0.0000	0.0000	0.0094	0.0007	0.0000	0.0000	0.0000	0.5279
$SZIT^c$	D	0.4324	0.9459	0.3243	0.5135	0.9189	1.0000	0.9730	0.2973
	p.value	0.0017	0.0000	0.0402	0.0001	0.0000	0.0000	0.0000	0.0757
$TZIT^d$	D	0.1892	0.6486	0.3784	0.4324	0.8919	0.9189	0.7838	0.3514
	p.value	0.5279	0.0000	0.0094	0.0017	0.0000	0.0000	0.0000	0.0201
$AZIT^e$	D	0.9459	0.3514	0.5676	0.9189	0.7027	0.7027	0.3243	0.6757
	p.value	0.0000	0.0201	0.0000	0.0000	0.0000	0.0000	0.0402	0.0000

Table 3.3: Two-Sample Kolmogorov Smirnov Tests. Significant at the $p < 0.05$ level. ATOM generated intraday returns with calibration to Renault SA on August 1, 2002. Experiment settings: ^{a} basic design of $UZIT_U$, $P_{min} = 45.2$, $P_{max} = 47.83$, $V_{min} = 1$, $V_{max} = 8075$. ^{b} basic design of $UZIT_N$, $P_{mean} = 46.88915$, $P_{sd} = 0.6015639$, $V_{min} = 1$, $V_{max} = 8075$. ^{c} basic design of $SZIT$, $a_{max} = 50$, $b_{min} = 30.3$, $b_{max} = 47.83$. ^{d} basic design of $TZIT$, the number of sliced windows is 50. ^{e} basic design of impatient agents $AZIT_I$, order volume is 250; basic design of patient agents $AZIT_P$, order volume is 500. The number of impatient agents is 30, the number of patient agents is 170. The number of prices is 2215.

		Mean	SD	Skewness	Kurtosis	ρ_1	ρ_2	Slope	$\hat{\phi}$
$UZIT_U$	Mean	1.276E-7	0.0035	-0.0001	17.7071	-0.3015	-0.0781	-0.0035	4.288E-6
	U*	-5.6787	-40.3570	1.5459	3.6420	25.6899	$-\infty$	∞	-0.3682
	p.value	0.0000	0.0000	0.1221	0.0003	0.0000	0.0000	0.0000	0.7127
$UZIT_N$	Mean	1.754E-7	0.0022	-0.0119	19.6185	-0.3001	-0.0746	-0.0035	1.912E-6
	U*	-5.9756	-8.9075	1.6285	3.0100	25.4974	$-\infty$	∞	-0.9570
	p.value	0.0000	0.0000	0.1034	0.0026	0.0000	0.0000	0.0000	0.3385
$SZIT$	Mean	-1.395E-6	0.0040	-7.5817	2221.9776	-0.2809	-0.0800	-0.0033	5.725E-6
	U*	-4.2283	-39.8479	0.9062	3.6821	12.4847	$-\infty$	17.5064	0.0049
	p.value	0.0000	0.0000	0.3648	0.0002	0.0000	0.0000	0.0000	0.9960
$TZIT$	Mean	-1.747E-6	0.0008	-0.1787	40.6920	-0.1526	-0.1349	-0.0030	2.781E-6
	U*	-0.9957	7.3446	3.3761	-3.8624	-16.1067	-32.4546	8.1609	2.3564
	p.value	0.3194	0.0000	0.0007	0.0001	0.0000	0.0000	0.0000	0.0184
$AZIT$	Mean	-4.035E-5	0.0010	-16.9185	2739.8962	-0.1996	-0.1997	-0.0017	6.623E-7
	U*	155.6127	2.4828	4.3272	-20.2753	-2.6619	-2.6619	-2.4313	-8.8679
	p.value	0.0000	0.0130	0.0000	0.0000	0.0078	0.0078	0.0150	0.0000
Real	Mean	-2.656E-6	0.0012	0.1441	25.8839	-0.2137	-0.2137	-0.0018	2.816E-6

Table 3.4: Fligner Policello Test. Significant at the $p < 0.05$ level. ATOM generated intraday returns with calibration to Renault SA on August 1, 2002. Experiment settings: [a] basic design of $UZIT_U$, $P_{min} = 45.2$, $P_{max} = 47.83$, $V_{min} = 1$, $V_{max} = 8075$. [b] basic design of $UZIT_N$, $P_{mean} = 46.88915$, $P_{sd} = 0.6015639$, $V_{min} = 1$, $V_{max} = 8075$. [c] basic design of $SZIT$, $a_{max} = 50$, $b_{min} = 30.3$, $b_{max} = 47.83$. [d] basic design of $TZIT$, the number of sliced windows is 50. [e] basic design of impatient agents $AZIT_I$, order volume is 250; basic design of patient agents $AZIT_P$, order volume is 500. The number of impatient agents is 30, the number of patient agents is 170. The number of prices is 2215.

		Mean	SD	Skewness	Kurtosis	ρ_1	ρ_2	Slope	$\hat{\phi}$
$UZIT_U$	Mean	1.276E-7	0.0035	-0.0001	17.7071	-0.3015	-0.0781	-0.0035	4.288E-6
	W	69.0000	5.0000	458.0000	576.0000	702.0000	0.0000	703.0000	720.000
	p.value	0.0000	0.0000	0.1105	0.0004	0.0000	0.0000	0.0000	0.7068
$UZIT_N$	Mean	1.754E-7	0.0022	-0.0119	19.6185	-0.3001	-0.0746	-0.0035	1.913E-6
	W	76.0000	44.0000	468.0000	536.0000	702.0000	0.0000	703.0000	774.0000
	p.value	0.0000	0.0000	0.0800	0.0046	0.0000	0.0000	0.0000	0.338
$SZIT$	Mean	-1.395E-6	0.0040	-7.5817	2221.9776	-0.2809	-0.0800	-0.0033	5.726E-6
	W	115.0000	4.0000	449.0000	584.0000	665.0000	0.0000	702.0000	684.0000
	p.value	0.0002	0.0000	0.1448	0.0003	0.0000	0.0000	0.0000	1
$TZIT$	Mean	-1.747E-6	0.0008	-0.1787	40.6920	-0.1526	-0.1349	-0.0030	2.781E-6
	W	154.0000	677.0000	532.0000	98.0000	12.0000	0.0000	679.0000	473.0000
	p.value	0.0023	0.0000	0.0056	0.0001	0.0000	0.0000	0.0000	0.0219
$AZIT$	Mean	-4.035E-5	0.0010	-16.9185	2739.8962	-0.1996	-0.1997	-0.0017	6.623E-7
	W	703.0000	516.0000	616.0000	1.0000	183.0000	184.0000	239.0000	1191.0000
	p.value	0.0000	0.0121	0.0000	0.0000	0.0101	0.0106	0.0913	0.0000
Real	Mean	-2.656E-6	0.0012	0.1441	25.8839	-0.2137	-0.2137	-0.0018	2.816E-6

Table 3.5: Paired Wilcoxon Tests. Significant at the $p < 0.05$ level. ATOM generated intraday returns with calibration to Renault SA on August 1, 2002. Experiment settings: [a] basic design of $UZIT_U$, $P_{min} = 45.2$, $P_{max} = 47.83$, $V_{min} = 1$, $V_{max} = 8075$. [b] basic design of $UZIT_N$, $P_{mean} = 46.88915$, $P_{sd} = 0.6015639$, $V_{min} = 1$, $V_{max} = 8075$. [c] basic design of $SZIT$, $a_{max} = 50$, $b_{min} = 30.3$, $b_{max} = 47.83$. [d] basic design of $TZIT$, the number of sliced windows is 50. [e] basic design of impatient agents $AZIT_I$, order volume is 250; basic design of patient agents $AZIT_P$, order volume 500. The number of impatient agents is 30, the number of patient agents is 170. The number of prices is 2215.

3.3.3 Sensitivity analysis: the importance of model parameters

A variety of factors may have an effect on the results presented above. This section is dedicated to a discussion on two factors that may affect notably our results: i) the impact of the proportion of "Big fishes" *vs.* "Small fishes" and ii) the role of the ratio "Limit" to "Market" orders. To address these points, we first vary only one of these parameters while the other stays constant. We then report the impact of these variations on the market.
Note that $TZIT$ are not significantly influenced by the variation in the proportions of the two factors, but rather by the price trend itself. Among the two remaining families, $SZIT$ is the more driven by the dynamics of order flows, and thus is the best potential candidate for an efficient calibration through both factors. We have chosen to restrict the presentation of the sensitivity analysis to the latter and to the $AZIT$ inspired by the Farmer model (Farmer *et al.*, 2005). For the latter, we vary the proportion of "patient" *vs.* "impatient" agents in the population.

In table 3.6, we report the correlations between the model parameters and the statistical properties tracked throughout this essay.

i) Big fishes / Small fishes : The results in table 3.6 clearly show that except for the mean, which is insensitive to the modification of the model parameters, all other stylized facts do react in some way to the latter. Even if the correlation coefficients may be relatively small (ranging from -0.5781 for the Kurtosis to 0.5876 for the "slope"), all of them are significant at the 1% level. However, if volatility tends to increase with this ratio (the correlation coefficient being equal to 0.5744), kurtosis tends to decrease (coef = -0.5781). *Big fishes*, due to the large trading volume they can generate, provide a siz-

Statistics		*Big/Small fishes* proportion[a]	Limit/Market proportion[b]	Impatient Agents	Patient Agents
Mean	*corr.*	0.3216	0.1408	-0.1551	0.1848
	$p-value$	0.1250	0.2414	0.2493	0.1942
SD	*corr.*	0.5744	-0.9759	0.2259	-0.6677
	$p-value$	0.000	0.000	0.0911	0.000
Skewness	*corr.*	-0.05648	0.1688	0.0990	-0.0597
	$p-value$	0.000	0.1593	0.4637	0.000
Kurtosis	*corr.*	-0.5781	-0.1872	-0.5669	0.2283
	$p-value$	0.000	0.1179	0.000	0.1070
ρ_1	*corr.*	0.4663	-0.7495	0.6469	-0.7771
	$p-value$	0.000	0.000	0.000	0.000
ρ_2	*corr.*	0.1470	0.0580	-0.7164	0.6859
	$p-value$	0.000	0.000	0.000	0.000
Slope	*corr.*	0.5876	-0.5572	0.0339	-0.2688
	$p-value$	0.000	0.000	0.8021	0.0564
$\hat{\phi}$	*corr.*	-0.3429	0.0571	0.0084	-0.3503
	$p-value$	0.000	0.6358	0.9502	0.0117

Table 3.6: Linear correlation coefficient of model parameters and stylized facts. 100 replications for each parameter set have been conducted. Each value is an averaged result. [a] Experiment settings: basic design of $SZIT$. The number of Small Fishes is fixed as 200, while the number of Big Fishes varies from 0 to 200. [b] Experiment settings: basic design of $SZIT$. The proportion of *Cancel* orders is fixed, $\Delta_C = 5$. The proportion of *Limit* orders Δ_L varies from 60 to 95 percent, and proportion of *Market* orders is defined according to formula $\Delta_M = 100 - \Delta_L - \Delta_C$. Significant at the $p < 0.05$ level.

zling amount of liquidity that directly benefits to *Small fishes*. In other words, *Big fishes* "feed" *Small fishes*. Thus, *Small fishes* can easily buy or sell stocks with a price close to the current market price, until a big order is completely executed. When the number of *Big fishes* significantly increases, these big players loose their role of liquidity providers, as they trade more frequently within their own group.

Increasing the proportion of *Big fishes* has also a positive effect on the slope of the decay function for autocorrelation coefficients calculated over absolute returns (correlation coefficient equals 0.5876). This result may suggest that the more "Big fishes" in the market, the more likely the emergence of volatility clusters.

In summary the fact that we observe significant correlations in that case indicates that finding the appropriate proportion of *Big fishes* vs. *Small fishes* might be a route to quantitatively fit stylized facts observed on real markets using ZITs. However, we were unable in our simulations to find that "ideal mix" which would have led to the perfect emergence of the whole set of studied stylized facts. In fact it seems a rather impossible task to fit most of the stylized fact by varying only one parameter. Moreover, the point is not to find the proportion which would fit theses stylized facts but rather to show that within a reasonable range corresponding to what is commonly observed on the market, this is not the case.

ii) Limit / Market orders : From the same table 3.6, one can notice that the proportion of "limit" to "market" orders has a significant effect on most of the studied stylized facts except the Skewness, the Kurtosis and the volume/volatility relationship.

Concerning the Standard deviation (SD), the correlation coefficient is close to -1. This result was expected: on the one hand, traders supply liquidity by posting limit orders and, on the other hand, demand liquidity by submitting

market orders that yield immediate partial or full execution. Thus, a large proportion of limit orders provides an important liquidity on both sides of the order book (Bids and Asks). On the contrary, a market order is immediately executed against order(s) standing in the limit order book: it moves the market by walking up or down the limit order book. Clearly, the proportion of Limit/Market orders has a significant impact on market volatility. A higher proportion of limit orders stabilizes the market by decreasing the standard deviation (correlation coefficient = -0.9759).
The negative coefficient for ρ_1 (-0.7495) suggests that the more Limit orders, the lower the auto-correlation of raw returns: this is a well-known fact related to the Bid-Ask bounce.
To summarize, some important stylized facts (Skewness, Kurtosis and volume/volatility relationship) seem to be insensitive to the modifications of the ratio Limit to Market orders. This indicates that fitting with accuracy real-world stylized facts is probably out of reach using this ratio alone.

iii) Patient/Impatient agents One can observe that increasing the number of these two categories of agents has a similar impact on various stylized facts as the variation of the proportion between Limit and Market orders does (see table 3.6). This result was expected since "patient" agents provide liquidity while "impatient" agents demand liquidity.

We now explore the sensitivity of the studied stylized facts to a modification of both parameters simultaneously using a regression approach. The results are reported in Table 3.7. In this table, we report, for each dependent variable (a stylized facts parameter), the estimate of the ordinary least squares (OLS) parameter for our two independent variables (proportion of Big *vs.* Small fishes and proportion of Limit *vs.* Market orders). A panel data compiling the outcomes of the 1000 replications of our analysis was collected.

We then run multiple linear regressions on the parameter values to observe their impact on empirical market outcomes. The general form of the model is:

$$SF \sim \alpha + \beta_1 BvS + \beta_2 LvM \tag{3.6}$$

In equation 3.6, SF is one of the stylized facts we study in this article, BvS is the proportion of Big *vs.* Small fishes and LvM the proportion of Limit *vs.* Market orders.

One first observes that all the values for β_1 and β_2 have non-zero values (at the 1% level). A second observation is that all R^2 coefficients have decent levels, ranging from 0.137 to 0.931. These results indicate that the levels of each stylized facts can be explained to some extent by a combination of the two ratios.

Unfortunately, the use of such relations to calibrate one ZIT model to the real-world stylized facts appears rather complicated: for example in model 2 (SD), the effects of β_1 and β_2 are similar whereas they are opposite for model 4 (Kurtosis). In other terms, fine tuning SD may result in an incorrect level of Kurtosis. The same problem can be identified comparing models 5 and 6 (ρ_i).

However, all these models make the assumption of multiple linear relations between independent and dependent variables although non-linearities between the latter may exist.

In summary, here again we conclude to the high level of complexity for fitting real-world stylized facts only using the two ratios BvS and LvM within a ZIT population.

A first conclusion can be drawn from this sensitivity analysis : simulations using ZIT should probably use the two leverages of the size of the traders and the type of orders to calibrate the level of volatility. However, the relation-

Model	*Statistics*	α	β_1	β_2	R^2
1	Mean	$1.859E-4$	$1.573E-8$	$-2.137E-6$	0.301
		$< 2E-16$	$< 2E-16$	$< 2E-16$	
2	SD	0.231	$-2.300E-5$	-0.0025	0.931
		$< 2E-16$	$< 2E-16$	$< 2E-16$	
3	Skewness	2.3831	0.0006	-0.0252	0.642
		$< 2E-16$	$< 2E-16$	$< 2E-16$	
4	Kurtosis	5.7459	0.0143	-0.0469	0.137
		$3.09E-11$	$< 2E-16$	$1.59E-5$	
5	ρ_1	-0.2864	-0.0002	0.0015	0.324
		$< 2E-16$	$< 2E-16$	$< 2E-16$	
6	ρ_2	-0.0951	0.0002	-0.0011	0.259
		$< 2E-16$	$< 2E-16$	$< 2E-16$	
7	Slope	$1.542E-4$	$-4.750E-6$	$-6.467E-6$	0.247
		0.4180	$< 2E-16$	0.0074	
8	$\hat{\phi}$	$2.558E-2$	$-1.130E-5$	$-2.542E-4$	0.266
		$< 2E-16$	$1.86E-11$	$< 2E-16$	

Table 3.7: Regression results for effects of model parameters on stylized facts. [a]Significant at the $p < 0.05$ level. 1000 runs. $A = 500$ – total number of agents in population, $A_{Big} \sim D(0, 500)$ – proportion of *Big fishes*, $A_{Small} = A - A_{Big}$ proportion of *Small fishes*. The proportion of *Cancel* orders is fixed, $\Delta_C = 5$. The proportion of *Limit* orders $\Delta_L \sim D(60, 95)$, and proportion of *Market* orders is defined according to formula $\Delta_M = 100 - \Delta_L - \Delta_C$. Significant at the $p < 0.05$ level.

ship within the higher moments set seems to rely on a more subtle interplay: increasing the volatility may well decrease the kurtosis of the series, and alter in a rather complicated way the Skewness and other stylized facts. Said differently, generating "calibrated" stylized facts all together using ZITs seems hard to realize, if simply possible.

3.4 Conclusion

In the current work, we used a methodology based on artificial stock markets. Typically, we ran experiments on an agent-based platform which microstruc-

ture is close to the Paris NYSE-Euronext stock exchange to simulate the price distribution of 37 stocks with the use of Zero Intelligence Traders (ZIT). Five families of agents deriving from the ZIT definition were considered and all were calibrated with the use of 37 real stocks data listed on the CAC 40 index. Notably, agents' choices were constrained by volume and prices exchanged on the real market, and we tested for different parametrizations either related to existing procedures in the literature or by increasingly constraining agents with respect to parameters of the real stocks.

The overall conclusion of this research is that none of the ZIT families can really be considered as a candidate to duplicate quantitative stylized facts. Furthermore, in most of the case (without calibration, $TZIT$ for example) price dynamics are completely unrealistic. This implies that these results discard the use of such agents in quantitative finance.

So to speak, ZITs help in understanding at coarse grain what drives the main stylized facts in actual price dynamics and can explain patterns in orders submissions (notably in high frequency trading schemes). They also qualitatively highlight that beyond "behaviors", market microstructure does matter. However, they are poor candidates for examining the dynamics of the price time series. The amount of information on data necessary to reproduce stylized facts (as for example with TZITs) seem outrageous and probably a better avenue will be to explore behavioral issues. Thus we need more sophisticated behaviors for artificial agents and/or strong calibrations processes as, for example, by Johnson system (see Johnson (1949)), that could lead to good replication of real data and predictive power of simulations. This is a necessary step if one wants to use agent-based simulations in quantitative finance.

Chapter 4

ABM: Portfolio Performance Gauging and Attitude Towards Risk Revisited

Contents

Despite the common understanding that the market is populated by traders with different tastes, skills, and beliefs, most asset pricing models are based on the so called "representative agent" models (see section 1.2.3 for details). Multi-agent simulations of financial markets frequently address investment problems by describing markets as complex systems of bounded rational and heterogeneous agents (Hommes, 2006; LeBaron, 2006; Tefatsion and Judd, 2006). This approach seeks to improve the traditional framework by introducing more flexible, robust, and realistic assumptions and to provide more powerful and sophisticated analysis tools for investment decision making.

The contributions of agent-based models can be divided into two partially overlapping classes (Anufriev, Bottazzi and Pancotto, 2006). The first class contains models where results come from a strict analytical investigation, like in Chiarella and He (2001), or Chiarella and He (2002). The second class consists of models based on the presentation and discussion of extensive computer simulations, such as in Lux and Marchsi (1999), Farmer and Joshi (2000), and Bottazzi *et al.* (2005). The research results presented in this chapter belong rather to the second class of research contributions.

In this chapter we describe the implementation of mean-variance model

within agent-based framework. That allows us to introduce the heterogeneity of the traders and to study their relative performance. In this work we use dollar wealth and the Sharpe ratio as portfolio performance measures. Despite different opinions about good sides and drawbacks of the Sharpe ratio, it appears to be a standard measure for evaluating portfolio performance. We also use the Sharpe ratio as a performance measure for better comparability of results with previous research.
This thesis sheds some new light on the question: whether mean-variance optimization can be outperformed by simpler allocation strategies. Furthermore, we examine what factors in the optimization are prominent for portfolio performance. The study takes into account transaction costs incurred by portfolio rebalancing. The results indicate that it is possible to achieve notably higher wealth and the Sharpe ratio with portfolio optimization rules than by using a naive strategy. Moreover, we find that the risk aversion as well as rebalancing frequency (or tolerance to deviation from target weights) significantly affect portfolio performance.

The natural question is whether this heterogeneous population evolves into homogeneous one, or whether one strategy dominates the market. To answer these questions, we use computational simulation techniques allowing the evolution of populations and their performance indicators. We compare the relative performance of investment strategies using an ecological competition (Lotka, 1925; Volterra, 1926) where populations of artificial investors co-evolve. This research methodology is widely used to understand nonlinear dynamical systems in which two or more species or agents interact through competition for resources.

This chapter is organized as follows. We first describe in section 4.1 the implementation of mean-variance model in the ATOM framework. In section 4.2, we propose a new analysis of the relative performance of investment strategies, rational mean-variance portfolio optimization versus naive diversi-

fication. Section 4.3 investigates the effect of risk preferences on the survival of agents in a long run. Section 4.4 presents how the rebalancing frequency affects both the final wealth and the Sharpe ratio in the presence of transaction costs.

4.1 Implementation of Mean Variance Optimization Model using ATOM

The mathematical basis of mean-variance optimization is detailed in the appendix A.3. As it is shown in the appendix A.3, the solution of mean variance optimization problem can be derived analytically. The result depends on the agent's expectations about the mean and the variance of the returns for the next period, and individual agent's risk preferences. In this section, we show how theoretical mean-variance optimization model can be implemented in the artificial stock market framework.

4.1.1 Simulation Model

We consider a securities market populated by a finite number of traders with *heterogeneous* preferences, indexed by $i \in 1, 2, ...I$. Time is discrete and indexed by $t = 0, 1, 2, ...$ There is also a finite number of assets $j \in 1, 2, ...J$. Traders come to the market with an initial amount of assets: trader i holds the following combination of assets $(q_{i,1}, q_{i,2}, ...q_{i,J})$, $i = 1...I$.

We denote investor $i's$ wealth at time t by W_t^i:

$$W_t^i = \sum_{j=1,J} p_{j,t} \cdot q_{j,t}^i + C_t^i \tag{4.1}$$

In equation 4.1, $p_{j,t}$ is the current market price of asset j at time t, and $q_{j,t}^i$ is the quantity of assets j held by trader i at dime t, C_t^i is an available

cash holding by agent i at t.

Agents, called mean-variance optimizer (or trader), endowed with this strategy try to minimize risk for a given target return following the mean (μ) – variance (σ^2) optimization rules.

$$min\frac{1}{2}\sigma_p^2 = min_{\varpi}\varpi^T V\varpi \tag{4.2}$$

$$\mu_p = \varpi^T\mu \tag{4.3}$$

$$\sum_{i=1}^{n}\varpi_i = 1,\ \varpi = (\varpi_1, \varpi_2, \ldots, \varpi_n) \tag{4.4}$$

where n – number of assets, μ_p – expected return of the portfolio, σ_p – standard deviation of the portfolio, V – correlation matrix, ϖ – target weights defined according to Markowitz rules. This program can deliver solutions outside the range $[0, 1]$ for the portfolio weights, which means that *shorting* is allowed.

This problem can be reformulated since each agent maximizes the mean-variance utility of the next period total return

$$U(\varpi) = \mu_p - \frac{1}{2}A\sigma_p^2 = \varpi^T\mu - \frac{1}{2}A\varpi^T V\varpi \tag{4.5}$$

An important parameter in this process is the investor's risk aversion A. Note, that in ATOM there exists also risk free asset yielding a zero percent interest rate, actively bought by conservative agents.

Agent i computes the optimal allocation of wealth of the risky assets $\varpi_t^{i,*} = (\varpi_{1,t}^{1,*}, \varpi_{2,t}^{i,*}, \varpi_{j,t}^{i,*})$. This allocation is a reflect of the risk preferences or the *rate of risk aversion* of an agent. Based on this information, a mean-variance trader calculates a desired quantity of stock j required to adjust its portfolio

to the "ideal" one.

$$q_{j,t}^{i,*} = \frac{\varpi_{j,t}^{i,*} \cdot W_t^i}{p_{j,t}} \tag{4.6}$$

Thus, to get as close as possible to the target weights defining his "optimal portfolio", a trader i issues at date t "buy" or "sell" orders following these rules: if the difference between the desired amount of stocks $q_{j,t}^{i,*}$ and the amount $q_{j,t-1}^i$ of stocks he actually hold is negative, he has to issue a sell order ("ask"). Conversely, if this difference is positive, he has to issue a buy order ("bid"). If there is no difference, the agent let his position unchanged. These rules are described in Algorithm 1. Transaction costs are incurred in the

if $q_{j,t}^{i,*} - q_{j,t-1}^i > 0$ **then**
  Send a *Bid* order
else if $q_{j,t}^{i,*} - q_{j,t-1}^i < 0$ **then**
  Send a *Ask* order
else
  Remain unchanged

Algorithm 1: Decision making process. $q_{j,t}^{i,*}$ – desired amount of asset j at the moment of time t held by agent i, $q_{j,t-1}^i$ – real amount of stock j held by agent i at the moment of time $t-1$

purchase and sale of each security. The costs are proportional to the value of each transaction $v_{j,t} = c \times p_{j,t} \times |q_{j,t}^i - q_{j,t-1}^i|$, where $p_{j,t}$ denotes the price of the jth security at time t; $q_{j,t-1}^i$ is the current number of stocks j held by agent j and the date $t-1$; $q_{j,t}^i$ is the desired number of stocks j defined according by optimization rules by agent i at the period t; c are the transaction costs of buying of selling. The total costs of portfolio rebalancing consisting m security are $T(v_1, \ldots, v_m) = \sum_{j=0}^m v_{j,t}$, $j = 1, \ldots, m \leq J$.

Another important question is how the limit price is determined. We propose the procedure inspired by price setting principle described in Jacobs

et al. (2004).

1. Bid price

$$P_{Bidt} = P_{Bidt-1} + \beta_t \tag{4.7}$$

   where P_{Bidt-1} is the best bid price in the order book in $t-1$; β_t is a random value in the range $[1; 10]$: it means that best bid price at the moment t will be increased by value from 1 to 10 cents. P_{Bid0} is equal to the previous day closing price.

2. Ask price

$$P_{Askt} = P_{Askt-1} + \alpha_t \tag{4.8}$$

   where P_{Askt-1} is the best ask price in the order book in $t-1$; α_t is a random value with the range $[1; 10]$: it means that best ask price at the time t will be decreased by value from 1 to 10 cents. P_{Ask0} is previous day closing price.

In the condition of double auction market, a profit-oriented buyer sets up the price lower his limit price because there would be a seller willing to accept this low bid price. Similarly, a seller sets a price higher his limit price, expecting that there would be a bidder ready to accept a high ask price. In condition of competitive market, the price comes closer to the market equilibrium price. As long as the buyer can undercut a competitor and still make a profit, he will add some insignificant amount to the last best bid price, similarly, seller will decrease the last best ask price by insignificant value, if it does not exceed his limit price.

These trading rules lead to bid-ask spread. Moreover, if a trade is not executed, the bids will tend upwards, and the asks will tend downwards until a trade occurs. Such situations will arise in sharp increasing or decreasing price dynamics. If there are some groups of agents with the same preferences, they will issue orders to buy or sell the same assets. If these participants want

to purchase a particular asset, they will send the bid orders one by one with slightly increasing price. That will move price upwards. This is exactly the case in the simulation of Jacobs *et al.* (2010). For this reason, they propose an *"anchoring rule"*, limiting the emergence of such side effect trends. These rules become effective when the security's price deviates too far from its recent level. Jacobs *et al.* (2010) provide the possibility for users to set up the "recent" value and the measure of maximum deviation from the recent level. For ask orders, the anchoring rule limits the minimum price according to the following formula $P_{O,min} = P_R - cP_L$, where P_R is a recent price, c is the user-specified parameter, and P_L is the average recent price or the standard deviation of recent prices (as defined by the user).

Inspired by the *anchoring rules* defined by Jacobs *et al.* (2010) and observing the same phenomena in our early experiments, we introduce the *anchoring rules* in our simulation in the following way (see algorithm 2).

if $\frac{P_{Bid_t} - P_{init}}{P_{init}} > \Lambda$ **then**

  $P_{Bidt} = P_{Bidt} \times (1 - \Delta)$

else if $\frac{P_{Askt} - P_{init}}{P_{init}} < -\Lambda$ **then**

  $P_{Askt} = P_{Askt} \times (1 + \Delta)$

end

Algorithm 2: Anchoring rule of price adjustment. At the beginning of each day we set up the initial price P_{init}, equals to the previous day closing price. Λ is the possible *"deviation"* of the current price from the initial one. Δ is an *adjustment* parameter, $\Delta > \Lambda$, Δ.

After a series of simulations, we have observed that if all the market participants are heterogeneous, then the market is active and stable. However, the results change dramatically if agents are homogeneous. Mean-variance optimizers are heterogeneous with respect to several characteristics (or attributes). The first one is the *coefficient of risk aversion* A. It determines agents' at-

titude toward risk. Agents with $A = 0$ are risk lovers or aggressive traders while agents with $A \geq 1$ are absolute risk averters (or conservative traders). The literature provides a large range for risk aversion parameter estimation. For example, the lowest risk aversion measure producing a profit is found in Mankiw, Rotemberg and Summers (1985) and is equal to $A = 0.3$. Hansen and Singleton (1982) define the possible ranges of risk aversion as 0.3502 and 0.9903. Gordon, Oaradis and Rorke (1972) use a risk aversion value between 0.6 and 1.4. Chen *et al.* (2007) define risk aversion in the range $[0.5, 5]$ with CRRA utility. Levy *et al.* (1995) investigate two groups of agents with risk aversion measures equal to 0.5 and 3.5 (in a CRRA framework). Risk aversion is equal to 18 in Obstfeld (1994), 30 – in Kandel and Stambaugh (1991). Kallberg and Ziemba (1983) define the ranges for the risk aversion parameter for a quadratic utility function as $0 \to \infty$. In this research, we consider the range $[0.1, 10]$ as representative of the different levels mentioned in the literature.

The second component of agents heterogeneity that maintains market liquidity and long-therm trading, is trading (or rebalancing) frequency Θ for different investors, said differently: agents rebalance their portfolios every Θ rounds. This condition helps to avoid empty order books, which could result from all orders cancellation by all agents at the same time. As noted in several articles (Shatner *et al.*, 2000; Hommes, 2006), the heterogeneity of time scale for agents actions is the important feature for emergence of stylized facts.

The next parameter of heterogeneity is a source of information about future asset returns and correlations. In ATOM there are two possibilities to get this information. The first one is to distribute exogenous information over agents. This option is important for initialization of simulations. The other way is to allow traders to calculate assets properties individually based on historical closing prices. At the end of the day (or trading period) t, investor i observes the time series (history) of assets $S^j_{t-1} \equiv \{P^j_n\}^{t-1}_{n=s}$, where $s < t - 1$, P^j_n is the closing price of stock j at day n. Based on this information, the agent

recalculates the expected return and the correlation matrix of assets that are necessary for the weights calculation procedure. The length of historical data set $|(t-1)-s|$ used for statistical calculation is an individual agent parameter.

4.1.2 Basic example of trading

To demonstrate the application of previously described rules we propose to consider following basic example.

1. **Assets properties**

   Weekly data

$$E(R) = \begin{pmatrix} A_0 & 0.0001818194 \\ A_1 & 0.001121336 \\ A_2 & 0.0002112892 \end{pmatrix} \sigma = \begin{pmatrix} A_0 & 0.01565114 \\ A_1 & 0.03523479 \\ A_2 & 0.02448014 \end{pmatrix}$$

$$Corr = \begin{pmatrix} & A_0 & A_1 & A_2 \\ A_0 & 1.00000000 & 0.1473994 & -0.02493462 \\ A_1 & 0.14739937 & 1.0000000 & 0.14394131 \\ A_2 & -0.02493462 & 0.1439413 & 1.00000000 \end{pmatrix} P = \begin{pmatrix} A_0 & A_1 & A_2 \\ 53 & 33 & 30 \end{pmatrix}$$

   The data for *agents parametrization* is described in the table 4.1

Name	$Markowitz_0$	$Markowitz_1$	$Markowitz_2$	$Markowitz_3$
Assets quantity	$(44; 62; 68)$	$(74; 92; 61)$	$(96; 46; 17)$	$(58; 90; 74)$
Risk aversion	7.8	4.3	1.4	1
Initial cash	0	0	0	0
Frequency	1	1	1	1
Optimal weights	$\begin{pmatrix} 0.596087 \\ 0.159774 \\ 0.244139 \end{pmatrix}$	$\begin{pmatrix} 0.540475 \\ 0.239173 \\ 0.220352 \end{pmatrix}$	$\begin{pmatrix} 0.442361 \\ 0.379257 \\ 0.178384 \end{pmatrix}$	$\begin{pmatrix} 0.131488 \\ 0.823101 \\ 0.045411 \end{pmatrix}$

Table 4.1: Agents Initialization

   The degree of risk aversion is defined in the interval $[0, 10]$. The agent $Markowitz_3$ has a risk aversion equal to 1 and is thus a risk lover. He

has an intention to invest a relatively large part of his wealth (0.8231009) into high-risk $Asset_1$. While, the agent $Markowitz_0$ is risk reverter, his degree of risk aversion is 7.8. This agent avoids risk and invests less into high-risk $Asset_1$.

2. **Trading.**

   First of all, agents estimate the quantities for their optimal portfolios according to equation 4.6. The ATOM scheduler randomly chooses an agent to act. Since simulations are asynchronous, the agent should take into account any price changes at any moment and the action of other traders during their decision making.

   The total wealth and the optimal allocation of the agent, named $Markowitz_0$, can be calculated in the following way:

$$W_0^0 = 44 \times 53 + 62 \times 33 + 68 \times 30 + 0 = 6418$$
$$q_{0,0}^0 = \frac{0.5960871 \times 6484}{53} = 73$$
$$q_{1,0}^0 = \frac{0.1597735 \times 6484}{33} = 31$$
$$q_{2,0}^0 = \frac{0.24413943 \times 6484}{30} = 53$$

   Since $Markowitz_0$ initially has 44 units of $Asset_0$ and the optimal quantity is 73, he should buy 29 units of $Assets_0$, sell 31 units of $Assets_1$ and 15 of $Assets_2$.

   We repeat the same procedure for other traders. The total wealth and the optimal asset allocations of the agent $Markowitz_3$ are as follows:

$$W_0^3 = 58 \times 53 + 90 \times 33 + 74 \times 30 + 0 = 8264$$
$$q_{0,0}^3 = \frac{0.13148814 \times 8264}{53} = 21$$
$$q_{1,0}^3 = \frac{0.8231009 \times 8264}{33} = 206$$
$$q_{2,0}^3 = \frac{0.04541096 \times 8264}{30} = 13$$

Since $Markowitz_3$ initially holds 58 units of $Asset_0$, 90 units of $Asset_1$, and 74 units of $Asset_2$, he should sell 37 units of $Asset_0$ and 61 units of $Asset_2$, and buy 116 of $Asset_1$.

As simulations are continuous, prices change in any moment, even if not all agents have sent their orders during the current "round". This is inspired from the functionality of real markets, where the environment can change while participants make decisions. This is why agent $Markowitz_1$ should take into account the new prices of $Asset_0$ and $Asset_1$, that stem from the orders issued by $Markowitz_0$, $Markowitz_3$ (see tables 4.2, 4.3). The total wealth and the optimal asset allocations of agent $Markowitz_1$ are as follows:

$$W_0^1 = 74 \times 53.4 + 92 \times 32.6 + 61 \times 30 + 0 = 7159.1$$
$$q_{0,0}^1 = \frac{0.540475 \times 7159.1}{53.4} = 72$$
$$q_{1,0}^1 = \frac{0.23917317 \times 7159.1}{32.6} = 53$$
$$q_{2,0}^1 = \frac{0.22035183 \times 7159.1}{30} = 53$$

From these equations and initial information in table 4.1, we conclude that $Markowitz_1$ should sell 2 units of $Asset_0$, 39 units of $Asset_1$, and 8 units of $Asset_2$.

$Markowitz_2$ calculates his current wealth and optimal asset allocations, taking into consideration price changes resulting from orders issued by $Markowitz_0$, $Markowitz_1$, $Markowitz_3$:

$$W_0^2 = 96 \times 53.4 + 46 \times 32.6 + 17 \times 30 + 0 = 7136$$
$$q_{0,0}^2 = \frac{0.64913344 \times 7136}{53.4} = 87$$
$$q_{1,0}^2 = \frac{0.0840371 \times 7136}{32.6} = 18$$
$$q_{2,0}^2 = \frac{0.26682946 \times 7136}{30} = 64$$

$Markowitz_1$ should sell 9 units of $Asset_0$ and 28 of $Asset_2$, and buy 47 units of $Asset_1$.

The prices of the orders are defined according to the formulas 4.7 and 4.8. For the moment, all order books are empty, and an initial set of prices should be used for initializing the process. In this example, $\beta_t \sim U(0, 10)$ and $\alpha_t \sim U(0, 10)$: the agents can modify the current price from 0 to 10 cents. There are two realized trades concerning $Asset_0$ and $Asset_1$ (see tables 4.2, 4.3) resulting from the first round of trading.

Direction	Quantity	Price	Agent ref.	Fixed price	Fixed quantity
~~B~~	~~29~~	~~53.4~~	~~$Markowitz_0$~~	~~...~~	~~...~~
A	37	53.0	$Markowitz_3$	...	...
...	...	...	...	53.4	29
A	8	53.0	$Markowitz_3$	...	...
A	2	52.7	$Markowitz_1$	...	...
A	9	53.0	$Markowitz_2$	...	...
~~B~~	~~6~~	~~53.6~~	~~$Markowitz_0$~~	~~...~~	~~...~~
...	...	...	...	53.0	6
~~B~~	~~3~~	~~53.2~~	~~$Markowitz_3$~~	~~...~~	~~...~~
...	...	...	...	52.7	2
...	...	...	...	53.0	1
~~B~~	~~4~~	~~53.9~~	~~$Markowitz_1$~~	~~...~~	~~...~~
...	...	...	...	53.0	4
A	9	52.8	$Markowitz_2$	...	...

Table 4.2: $Asset_0$

The second round runs with updated prices; moreover, the agents should either update or remove their orders pending in the order books. For this reason, the agents recalculate their wealth, and figure out how close they are to the targets.

The agent $Markowitz_0$ is quite close to his target weights, however, he should make some adjustments:

$$W_1^0 = 73 \times 53.4 + 31 \times 32.6 + 68 \times 30 + 0 = 6948.8$$

Direction	Quantity	Price	Agent ref.	Fixed price	Fixed quantity
~~A~~	~~31~~	~~32.6~~	~~$Markowitz_0$~~	~~...~~	~~...~~
B	116	32.6	$Markowitz_3$	...	...
...	...	...	...	32.6	31
B	85	32.6	$Markowitz_3$	...	...
A	39	33.0	$Markowitz_1$	...	...
B	47	32.9	$Markowitz_2$	...	...
~~B~~	~~3~~	~~33.1~~	~~$Markowitz_0$~~	~~...~~	~~...~~
...	...	...	...	33.0	3
~~A~~	~~36~~	~~33.0~~	~~$Markowitz_1$~~	~~...~~	~~...~~
B	61	33.2	$Markowitz_3$	...	...
...	...	...	...	33.0	36
B	25	33.2	$Markowitz_3$	...	...
~~B~~	~~19~~	~~33.5~~	~~$Markowitz_1$~~	~~...~~	~~...~~
~~A~~	~~28~~	~~32.9~~	~~$Markowitz_2$~~	~~...~~	~~...~~
...	...	...	...	33.5	19
...	...	...	...	33.2	9
B	16	33.2	$Markowitz_3$	...	...

Table 4.3: $Asset_1$

$$q^0_{0,1} = \frac{0.5960871 \times 6948.8}{53.4} = 78$$
$$q^0_{1,1} = \frac{0.1597735 \times 6948.8}{32.6} = 34$$
$$q^0_{2,1} = \frac{0.24413943 \times 6948.8}{30} = 57$$

During the second round, $Markowitz_0$ should try to buy 6 units of $Asset_0$ and 3 units of $Asset_1$, sell 9 units of $Asset_2$.

The current portfolio of $Markowitz_3$ is far from optimal weights.

$$W^3_1 = 21 \times 53.0 + 121 \times 33.0 + 71 \times 30.6 + 0 = 7278.6$$
$$q^3_{0,1} = \frac{0.13148814 \times 7278.6}{53.0} = 18$$
$$q^3_{1,1} = \frac{0.8231009 \times 7278.6}{33.0} = 182$$
$$q^3_{2,1} = \frac{0.04541096 \times 7278.6}{30.6} = 11$$

$Markowitz_3$ will buy 3 units of $Asset_0$ and 61 units of $Asset_1$, sell 60 units of $Asset_2$

Direction	Quantity	Price	Agent ref.	Fixed price	Fixed quantity
A	61	29.5	$Markowitz_0$	...	...
A	15	29.1	$Markowitz_3$	...	...
A	8	30.2	$Markowitz_1$	...	...
A	28	30.6	$Markowitz_2$	...	...
~~B~~	~~9~~	~~30.5~~	~~$Markowitz_0$~~	~~...~~	~~...~~
...	...	...	...	30.6	9
A	19	30.6	$Markowitz_2$	...	...
A	60	30.3	$Markowitz_3$	...	...
~~A~~	~~6~~	~~29.8~~	~~$Markowitz_1$~~	~~...~~	~~...~~
~~B~~	~~45~~	~~30.4~~	~~$Markowitz_2$~~	~~...~~	~~...~~
...	...	...	...	29.8	6
...	...	...	...	30.3	39
A	21	30.3	$Markowitz_3$	...	...

Table 4.4: $Asset_2$

$Markowitz_1$ recalculates his optimal allocations since the prices have been changed.

$$W_1^1 = 74 \times 53.0 + 56 \times 33 + 61 \times 30.6 + 0 = 7636.6$$
$$q_{0,1}^1 = \frac{0.540475 \times 7636.6}{53.0} = 78$$
$$q_{1,1}^1 = \frac{0.23917317 \times 7636.6}{33} = 55$$
$$q_{2,1}^1 = \frac{0.22035183 \times 7636.6}{30.6} = 55$$

$Markowitz_1$ should buy 4 units of $Asset_0$ and 19 of $Asset_1$, sell 6 units of $Asset_2$.

$Markowitz_2$ updates his positions in the following way:

$$W_1^2 = 87 \times 53.0 + 46 \times 33.0 + 17 \times 30.6 + 477 = 7126.2$$
$$q_{0,1}^2 = \frac{0.64913344 \times 7126.2}{53.0} = 87$$
$$q_{1,1}^2 = \frac{0.0840371 \times 7126.2}{33.0} = 18$$
$$q_{2,1}^2 = \frac{0.26682946 \times 7126.2}{30.6} = 62$$

$Markowitz_1$ should sell 9 units of $Asset_0$ and 28 of $Asset_1$, buy 45 units of $Asset_2$.

The agents will repeat the same procedures over next trading rounds, until all of them get their target portfolio weights and all traders are satisfied. Hence, we can assume that after several trading rounds all traders will be satisfied, thus there will be no trade any more. To avoid such unrealistic situation, one should introduce heterogeneous agents' populations with different beliefs, risk aversion degrees, rebalancing frequencies, trading strategies, and so on.

4.1.3 Model validation

In this subsection we test the ability of mean-variance traders, introduced in section 4.1.1, to maintain a long-run trading on the system, and to reproduce main stylized facts observed in real financial time series in intraday as well as in extraday time-frame. For this purpose, we investigate the presence of known stylized facts of financial time series, i.e., volatility clustering and fat tails in the distribution of returns. We run the experiments within the three-asset limited framework. The following simulation assumptions are considered. Hundred traders populate the market. At the beginning of the simulations the agents get the information about expected returns, variances and correlation matrices of stocks. Based on this information, agents calculate the target weights and try to keep their portfolios closer to these target weights over the next 200 days, that is regarded as trading period. After this period of trading, the agents use recently generated prices to estimate the information concerning the traded stocks. For sake of simplicity, we assume that the information concerning the underlying probability distribution of securities prices as well as current securities prices is perfect information that is available continuously and costlessly to all investors.

One "simulation day" contains 50 rounds of preopening session, 1000 rounds of continuous trading, and 50 rounds of the closing session. Such time

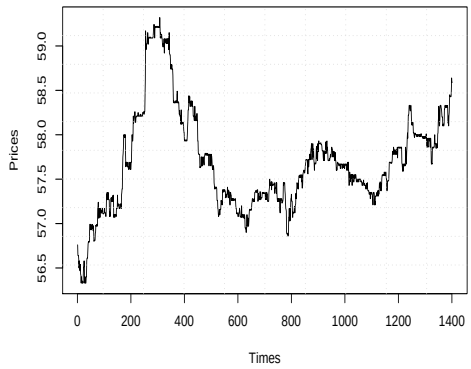
organization helps us control the rebalancing frequency of agents. Each agent is attributed by parameter, trading frequency, $\theta \sim U(1, 10000)$. It means that there are some agents trading every day, while others reoptimize their portfolios every 10 days. The agents are also heterogeneous with respect to their attitude toward risk. The degree of risk aversion is uniformly distributed in $[0.1, 10]$. 50 agents out of 100 are allowed to hold short position.

We reproduce the years of trading on the fine grain level through intraday trading. This feature provides us with an outstanding possibility to investigate the stylized facts that require different time granularity. For example, *aggregational gaussianity* is observed when one increases the time scale $\triangle t$ over which returns are calculated Cont (2001).

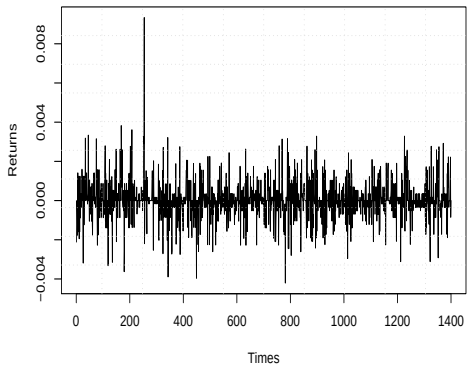
Figure 4.1 displays the intraday price series. We quantitatively (table 4.5) and qualitatively (figure 4.2(b) and 4.2(c)) show the nongaussian behavior in intraday returns. Figure 4.2(b) puts forward the deviation from normality, especially sharp peaked distribution (for comparison, the solid line represents the distribution of standard normal distribution); figure 4.2(c) clearly exhibits fat tails. Kurtosis and skewness displayed in table 4.5 confirm the deviation from normality, as all kurtosis are far from their Gaussian values and positive skewness exhibits gain/loss asymmetry in return distributions.

In figures 4.2(d) and 4.2(e), we present the autocorrelation $C(\tau)$ of the raw returns and the absolute returns at different time lags τ. While the autocorrelation of raw returns exhibits rapid decay, the autocorrelation of the absolute value of returns shows the presence of long-range correlations with a very slow exponential decay. We can conclude that the simulated time series exhibit the well known stylized fact of volatility clustering observed in real-world markets.

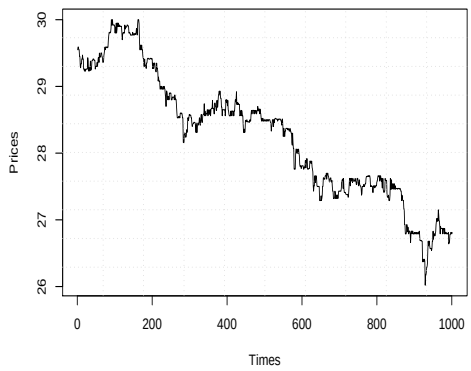
Next, our research is conducted to examine the ability of mean-variance artificial investors to generate the prices with realistic statistics in accordance with theoretical and empirical researches on statistical return properties in



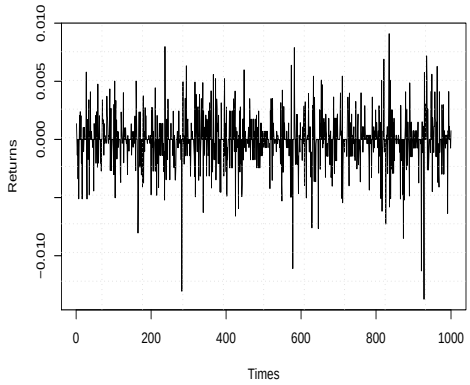
(a) *Asset 0: Prices*



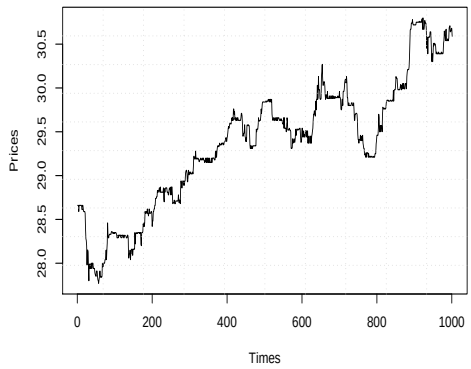
(b) *Asset 0: Returns*



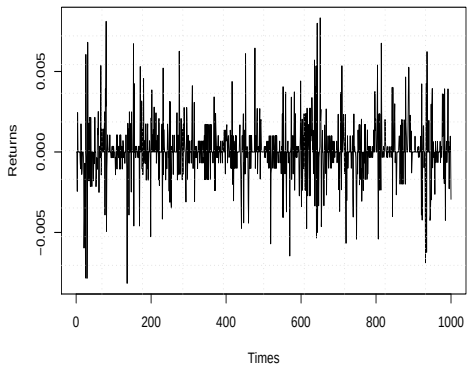
(c) *Asset 1: Prices*



(d) *Asset 1: Returns*

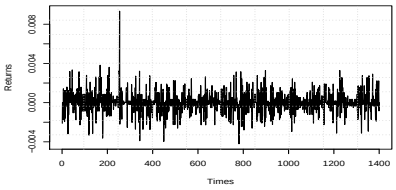


(e) *Asset 2: Prices*

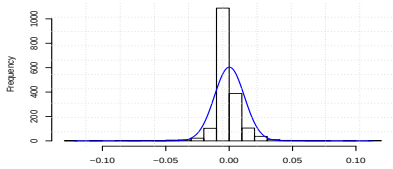


(f) *Asset 2: Returns*

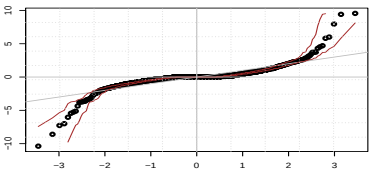
Figure 4.1: Intraday price dynamics. Prices are generated by 100 mean-variance optimizers, heterogeneous with respect to their beliefs, risk aversion $A \sim U(0.1, 10)$, trading frequency $\theta \sim U(1, 10000)$.



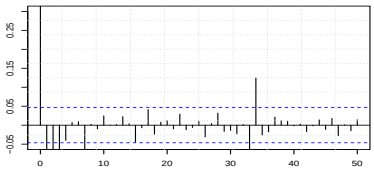
(a) *Returns*



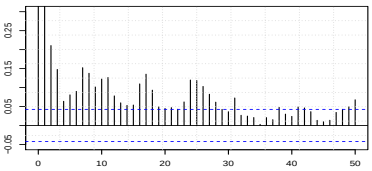
(b) *Histograme*



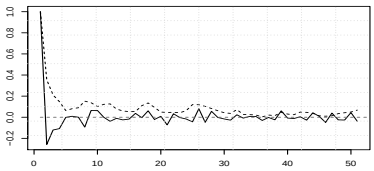
(c) *QQ-Plot*



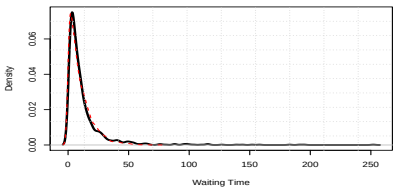
(d) *Autocorrelation*



(e) *Autocorrelation of absolute returns*



(f) *Autocorrelation of row vs absolut returns*



(g) *Price setting waiting time*

Figure 4.2: Stylized facts for intraday prices. Prices are generated by 100 mean-variance optimizers, heterogeneous with respect to their beliefs, risk aversion $A \sim U(0.1, 10)$, trading frequency $\theta \sim U(1, 10000)$.

Statistics	$Asset_0$	$Asset_1$	$Asset_2$
Minimum	-0.03008936	-0.1920777	-0.0800427
Maximum	0.03008936	0.282567	0.08701138
Mean	4.182412e-05	-0.000438054	0.0001725323
Variance	0.0001120417	0.003788261	0.0004502855
Stdev	0.01058498	0.06154885	0.02121993
Skewness	0.07756422	0.07483113	0.1398351
Kurtosis	17.802075	11.49896	15.49548
ρ_1	-0.0780217	-0.1637344	-0.1571716
ρ_2	-0.07361454	-0.1166177	-0.1195904
Slope	-0.0008539391	-0.001849647	-0.001875495

Table 4.5: Basic Statistics for indraday returns. Prices are generated by 100 mean-variance optimizers, heterogeneous with respect to their beliefs, risk aversion $A \sim U(0.1, 10)$, trading frequency $\theta \sim U(1, 10000)$.

extraday time frame.

Extraday trading

On the real market, investors characterized by mean-variance optimization rules, limit their tradings only to a few orders per day. They use daily, weekly or even annual returns data to estimate the volatility and correlations of assets. In this subsection we examine the ability of mean-variance optimizers to generate extraday price dynamics in line with real market. For this purpose we extend previously described example to thousand of days. The price series as well as long-returns are shown in figure 4.3.

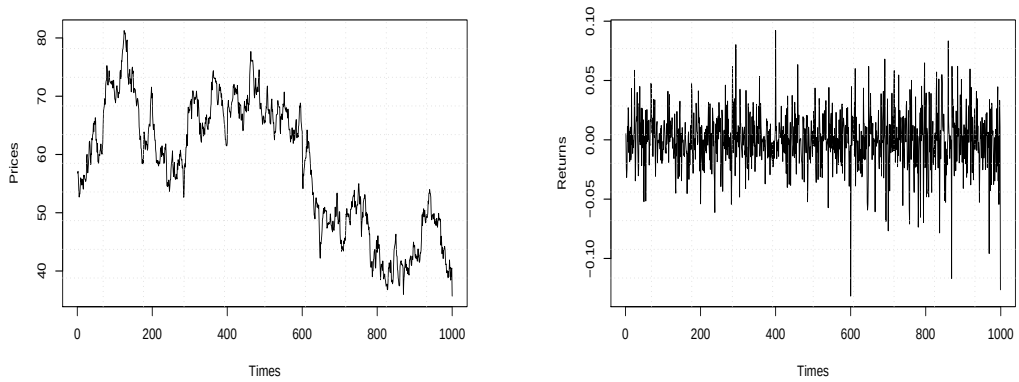
We now test how the population of artificial traders is able to reproduce the price dynamic with statistical properties close to the real one at a daily horizon. First, we analyze the qualitative stylized facts. In Figures 4.4(i) and 4.4(j) we compare the autocorrelation functions of the absolute returns, these results show the presence of long-range nonlinear correlations. The autocorrelation of raw returns decays immediatly 4.4(g) and 4.4(h). The computational experiments performed for this work show a number of important results.

Figures 4.2(b) and 4.4(d) exhibit aggregational Gaussianity: as we increase the time scale over which returns are calculated from intraday to daily, the returns distribution is approximately Gaussian, the shape of distribution is not the same. Figure 4.4(i) puts forward a positive autocorrelation of absolute returns over several days. From these results we can conclude that the returns of extraday price series exhibit the main features of real market: fat tail, zero autocorrelation or raw returns, slow decay of the autocorrelation of the absolute values.

Finally, we run a series of extensive simulations to check whether the artificial mean-variance agents generate prices in accordance to their "expectations". In this work, this ability is called "predictive power". In other words, we report that the mean-variance optimizers are able in some degree to generate the price series with moments in line with those imposed to them for optimization at the beginning of simulations. This test is detailed in Appendix A.4.

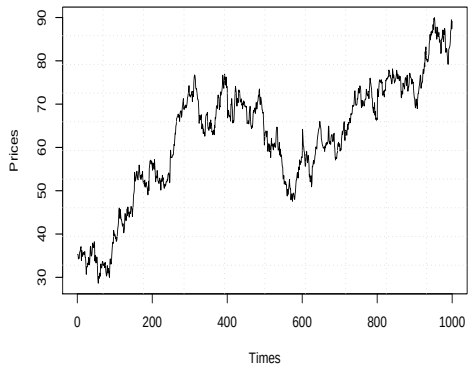
In summary we can say that the artificial agents are able to perform realistic price dynamics with quantitative and qualitative stylized facts at intraday as well as daily horizons in line with those from real stock markets.

Next sections are devoted to study the relative performance of portfolio optimization strategies and factors affecting them in the artificial stock market framework.

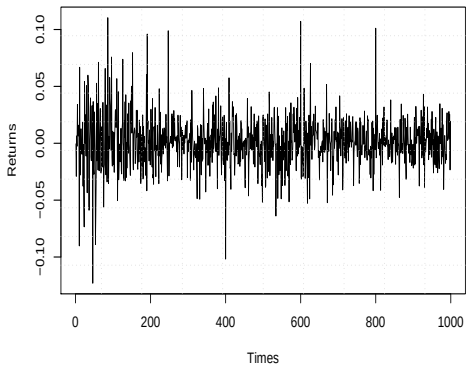


(a) *Asset 0: Prices*

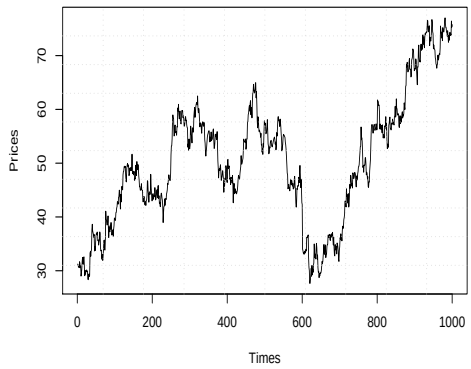
(b) *Asset 0: Returns*



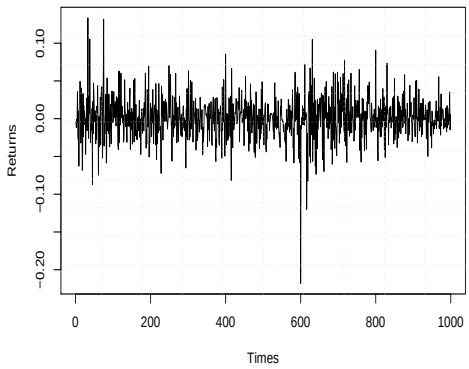
(c) *Asset 1: Prices*



(d) *Asset 1: Returns*

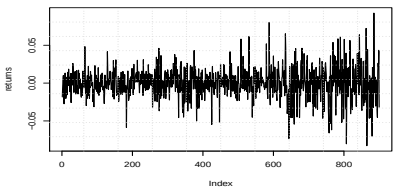


(e) *Asset 2: Prices*

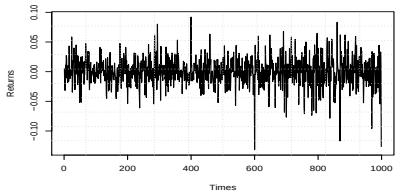


(f) *Asset 2: Returns*

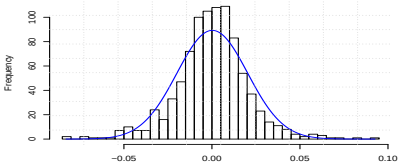
Figure 4.3: Extraday price dynamics. Prices are generated by 100 mean-variance optimizers, heterogeneous with respect to their beliefs, risk aversion $A \sim U(0.1, 10)$, trading frequency $\theta \sim U(1, 10000)$.



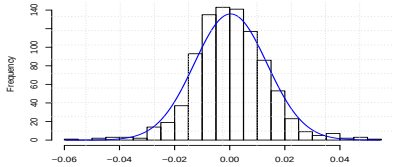
(a) *Real returns*



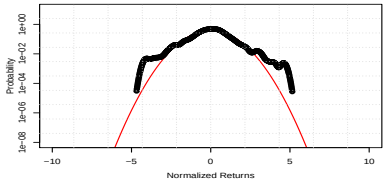
(b) *ATOM returns*



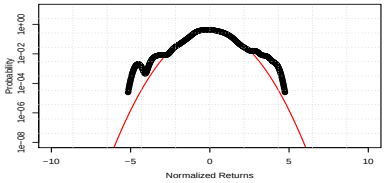
(c) *Real Returns Distribution*



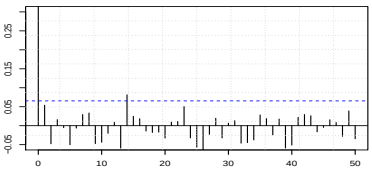
(d) *ATOM Returns Distribution*



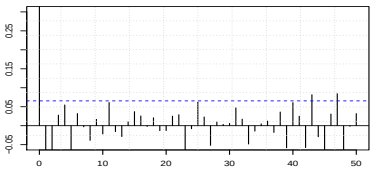
(e) *Real Fat tails*



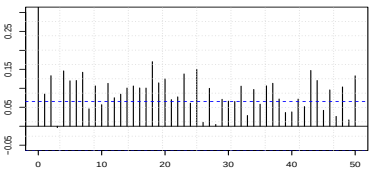
(f) *ATOM Fat Tails*



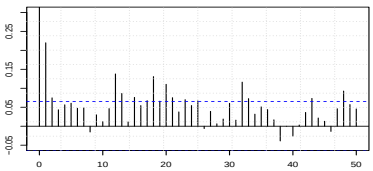
(g) *Real Autocorrelation of row returns*



(h) *ATOM Autocorrelation of row returns*



(i) *Real Autocorrelation of absolute returns*



(j) *ATOM Autocorrelation of absolute returns*

Figure 4.4: Comparison of Stylized facts for Extraday Price Series (1000 days). Prices are generated by 100 mean-variance optimizers, heterogeneous with respect to their beliefs, risk aversion $A \sim U(0.1, 10)$, trading frequency $\theta \sim U(1, 10000)$.

4.2 Optimal Portfolio Diversification? A multi-agent ecological competition analysis

4.2.1 Introduction

In this section, we renew the analysis of a classical question in finance, namely, *the relative performance of investment strategies.* We try to figure out whether a rational mean-variance portfolio optimization can be outperformed by naive diversification. This research is motivated by the contradictory and controversial findings of DeMiguel *et al.* (2009), Kritzman *et al.* (2010), and Tu and Zhou (2011) who did the same kind of research but within the traditional finance philosophy (no agents, no co-evolution, no complexity, no heterogeneity). DeMiguel *et al.* (2009) compare several investment strategies using a *backtesting methodology.* They evaluate the sample-based mean-variance strategy and its extensions designed to reduce estimation errors. The authors conclude that non of these strategies is consistently better than the naive diversification rule in terms of the Sharpe ratio. This result can be explained by the errors in estimating means and covariances.

Kritzman *et al.* (2010), as practitioners, argue that by relying on longer-term samples for estimating expected returns, optimized portfolios outperform equally weighted portfolios out of sample. Kritzman et al. estimate expected volatility and correlations, using the monthly 5-, 10-, and 20-year data, while DeMiguel *et al.* (2009) state that for a portfolio with 50 assets, the estimation window should be more than 6000 months. However, the minimal-variance portfolio generates superior out-of-sample performance compared with equally weighted portfolio in the simulations conducted by Kritzman *et al.* (2010). The authors suggest that investors should not rely solely on naive extrapolations of long historical samples. Instead, investors

may benefit by adjusting optimization inputs on subsamples of high-volatility and low-volatility regimes in accordance with their expectations.
Tu and Zhou (2011), extending the backtesting methodology of DeMiguel *et al.* (2009), suggest that a combination of the 1/N strategy with the sophisticated diversification can each of its constituents taken separately. This result is proposed in an empirical framework which is extremely similar to the one of De Miguel and al.

Understanding the characteristics of winning and losing market strategies is an important question for investors and regulators. But in our opinion, the main problem with the researches mentioned above is the unrealistic *"atomistic"* assumption that underlies the backtesting methodology. Said simply, this assumption allows to gauge an investment strategy with historical data as if its true implementation would have not modified these prices. This assumption is in sharp contrast with analyses of Levy *et al.* (1995), Hommes (2006) who clearly show that prices may well be influenced by several parameters (investment strategies, the cognitive skills of investors or the market microstructure itself) that are neglected in the backtesting approach. It seems obvious that different investors are characterized by different investing behaviors that are, at least partially, responsible for the time evolution of market prices. We argue in this research that a convincing answer to the question *"among this set of investment strategies, which one outperforms the others?"*, overcoming the previously mentioned limitations, can be delivered by a multi-agent system allowing to implement ecological competitions among these strategies. We show, among others, that the best possible strategy over the long run relies on a mix of mean-variance sophisticated optimization and a naive diversification. This result reinforces the practical interests of the Markowitz framework that is strongly discussed in DeMiguel *et al.* (2009) for example.

4.2.2 Agents behavior

One of the advantages of ABM is that the agents are autonomous. In a mathematical model, all market participants are defined as equal-power rational entities facing homogeneous constraints. Agent actions are predetermined by strict equations describing their reaction in response to particular market conditions. ABM allows to overcome the limitations imposed homogeneity. In this research, we design 8 agent populations, each of them following a generic strategy. These strategies are presented in subsection 4.2.2.1.

4.2.2.1 Trading strategies

We start by introducing how a portfolio of assets is modelled and what kind of decision agents must make in a simulation. The purpose of each strategy is to allow agents to manage a diversified portfolio of financial assets over time in different ways.

A portfolio is defined as a vector of weights over the investment universe. This vector is denoted ϖ^{xx}, xx allowing to identify the generic strategy determining this vector. Depending upon the strategy definition or the empirical design, these weights can be negative or not. If this is the case, one will refer to this situation as "shorting allowed", which means that agents are allowed to sell borrowed assets and they will repurchase them later on.
Each time a new portfolio is computed, the current weight vector ϖ_t^{xx} is compared to the previous one ϖ_{t-1}^{xx} to adjust the number of stocks to hold. This adjustment take into account the weight vectors and the corresponding assets current prices. As a result agents decide to buy or to sell certain assets they hold to reach their new (weight vector) target (see algorithm 1). These decisions must be practically implemented, that means *"translated into buy or sell orders"*, with quantities and prices in accordance to the target. One must remember that each strategy implies different parameters that may have dif-

ferent values within the same agent population; thus each agent has his own weight vector holding during a simulation.

This process being the same whatever the behavior, we can now describe at fine grain the 8 generic strategies (see Table 4.6).

Name	Short Name	Basic definition & particularities
Naive	N	Equal weights, no sophisticated behavior
Mean Variance 1	$MLong$	Markowitz optimization, long positions only
Mean Variance 2	$MShort$	Same as $MLong$, shorting allowed
Market Portfolio Holders	MP	Weights according to assets capitalisation on the market
Bayesian Traders 1	$BLong$	Based on Markowitz, estimation of moments co-moments of asset returns improved, long positions only
Bayesian Traders 2	$BShort$	Same as $BLong$, shorting allowed
Strategy Combinators 1	$CLong$	Mix of N and $MLong$
Strategy Combinators 2	$CShort$	Mix of N and $MShort$

Table 4.6: Strategies description

Population 1: Naive diversification investors The agents endowed with the naive strategy (N) ignore all information about risk and return of assets. Naive investor i allocates his funds equally among the J risky assets in equal proportions $\varpi_{j,t}^{i,N} = \frac{1}{N}\ \forall j = \overrightarrow{1, J}$ the weights of wealth allocated to stock j of agent i at the moment of time t. In contrasts to sophisticated rules that are usually asymptotically unbiased but have a large (variance) estimation error in small samples, the 1/N rule is biased, but has zero estimation error.

Populations 2 and 3: Mean-variance optimizers Agents endowed with this strategy define optimal allocation weights $\varpi^M = (\varpi_1, \varpi_2, \ldots, \varpi_J)$ by applying the mean (μ) variance (σ^2) optimization rules (Section 4.1). This optimization problem provides the solutions outside the range $[0, 1]$, that allows

shorting. From its definition, we create two agents population, one allowed to use short selling (MShort), the other not allowed to do so (long only, MLong)

Population 4: Market portfolio holders Market portfolio (MP) holder is the type of agent with a portfolio consisting of all the assets in the market with weights proportional to asset capitalization (Treynor, 1962). In a more realistic context, if an investor has no special insight about expectation returns and volatility of individual stocks he is supposed to hold the market portfolio (portfolio of all available stocks).

$$\varpi_{j,t}^{i,MP} = \frac{P_{j,t} \times Q_{j,t}}{\sum_{n=0}^{J} [P_{n,t} \times Q_{n,t}]} \tag{4.9}$$

$P_{j,t}$ price of asset j at moment t, $Q_{j,t}$ number of asset j traded on the market at the moment t, C_t total market capitalization.

Population 5 and 6: Bayesian traders Agents within this population have a behavior that extends the Markowitz rules. The Markowitz approach has been criticized due to measurement errors in the estimation of assets moments and co-moments. To overcome these problems authors like Klein and Bawa (1976) or Brown (1979) propose to improve the co-moments estimation by using a factor equal to $1+\frac{1}{M}$ that reduces the estimation error and leads to more reliable investment weights. Moments and co-moments being estimated following this rule, agents the same rules as Markowitz agents to determine the target weights.

From this logic we define two different populations, one in which short selling is allowed, Bayesian Short Selling *(BShort)* and one in which it is forbidden, Bayesian Long Only strategies *(BLong)*.

Population 7 and 8: Strategy combinators The last population has the ability to combine the naive 1/N strategy with the sophisticated mean-

variance optimization strategy. It has been studied by some authors who thought it could improve the overall performance of investors (Brown, 1976). Mathematically the combination of two strategies can be described as follows:

$$\varpi_{j,t}^{i,C} = (1-\delta)\varpi_{j,t}^{i,N} + \delta\varpi_{j,t}^{i,M} \qquad \delta = \frac{\varphi_1}{\varphi_1+\varphi_2}$$
$$\varphi_2 = \frac{1}{A^2}\left[\frac{(T-2)(T-J-2)}{(T-J-1)(T-J-4)}\right] \tag{4.10}$$

where $\varpi_{j,t}^{i,C}$ – weights defined by strategies combination, $\varpi_{j,t}^{i,N}$ are the weights defined according to the naive diversification rule, $\varpi_{j,t}^{i,M}$ are the weights defined according to the Markowitz rule, δ is the combination parameter $0 \leq \delta \leq 1$, J is the number of assets, and T is the memory span or the length of available historical data. "Markowitz Shorting allowed" and "Markowitz Long-only" are used for combinations, hence Combination Short *(CShort)* and Combination Long *(CLong)* populations are studied in this research.

4.2.3 Simulation settings

We compare the relative performance of investment strategies using Ecological Competition (EC) (Lotka, 1925; Volterra, 1926), where agents change their strategies between the trading periods using their historical performances. This research approach is widely used to understand nonlinear dynamical systems in which two or more species or agents interact through competition for resources. Stock market can be regarded as an environment where agents compete for the value of traded stocks. Traders occur losses and change strategies that performed well during the last round. The agents populations compete each against the others in order to get higher wealth or the Sharpe ratio. This approach not only allows us to track a particular performance measure, but also to follow its evolution in the long-run. Additionally, ecological com-

petitions show the effects of each strategy on the others. For instance, one population of agents can take advantage from the presence of the others.

Initially, we populate the ATOM environment with our 8 populations of agents. The size of each of these populations x_k for $k = \overrightarrow{1,8}$ is the same $\forall k$. The total number of agents is $X = \sum_{k=1}^{8} x_k$. Populations are updated every simulation round according to their performance $x_k = X\frac{P_k}{P_T}$, where P_k the performance of population k and P_T the overall performance of the whole soup of populations. The performance can be measured as i) the total wealth (cash + market capitalization of the stocks of all the agents in each population) or ii) the average Sharpe ratio (Sharpe, 1966) of the population, during the previous round. A population is said to be extinct if $x_k = X\frac{P_k}{P_T} < 1$

Each strategy is encoded in an initial population of NNN agents. These populations are mixed and compete in the same market, trading the same stocks. Prices are the direct result of the flow of orders sent by the agents to the central order books ruling the artificial stock exchange. A time step in our ecological competitions is made of several *rounds*, each of them encompassing 1000 trading days.

Simulation settings We study the 8 populations of agents presented in table 4.6. Each population of agents starts with 100 agents, that have costless access to all information concerning the underlying probability distribution of security prices as well as current security prices. The agents are homogeneous with respect to their initial budget (they enter the market with 50 units of each type of stocks and 1000$ in cash). Contrary to DeMiguel *et al.* (2009) and Tu and Zhou (2011) who set the risk aversion parameter to 1 or 3 for Markowitz strategies and its extensions, risk aversion in our simulations is uniformly drawn between [0.5,5] $A \sim U(0.5, 5)$ in order to test a larger variety of behaviors, from risk averse agents to risk takers. The agents trade 30 different stock classes (like the 40 different families of stocks listed in the CAC40

index). Such relatively large number of stocks allows us to have significant difference in portfolio composition of heterogeneous mean-variance agents. For the first trading period, we provide to agents an initial information about assets, then they rely on historical price series generated by the trading activity itself. Contrary to DeMiguel *et al.* (2009), that use monthly return data, we deal with significantly more information-rich daily data. Each dataset consists of 500 observations. Monthly data would require an investment period of nearly 40 years to include as many observations as are presented here in the daily return data.

4.2.4 Results and Discussions

We present here the results of two different ecological competitions. In the first one, the reproduction rate of each population is linked to dollar earnings (see subsection 4.2.4.1) while in the second one, it is a function of the Sharpe ratio (see subsection 4.2.4.2).

4.2.4.1 Ecological competition 1: wealth

The simulations results (figure 4.5(a)) show that all the constrained (long-only) strategies ($MLong$, $BLong$, $CLong$), the naive (N) and the market portfolio strategies (MP) quickly disappear from the market at the end of 50 rounds. According to Levy and Ritov (2011), a possible explanation of this phenomenon could be linked to the large positions (positive or negative) implied by short selling, when it is allowed : the long-only strategies have zero-positions ($\varpi_{j,t}^{i,*} = 0$) in about 50% of the traded assets. Thus, the agents with long-only strategies trade only half of the investment set to maintain their target weights. At the same time, agents with short-selling strategies trade the whole set of assets and increase their wealth more efficiently.

In addition, we observe that the population $CShort$ are better than their

individual component rules ($MShort$ and N) which is clearly in line with the results of Tu and Zhou (2011).

We also investigate a possible effect of the initial size of the population in its survival time. We therefore changed population initial distributions dynamically ($\sim U(20, 200)$) so to get a majority of certain types of agent in the whole population soup at the beginning of each experiment. Our results indicate that even if the initial proportion of naive agents (≈ 200 individuals) is much bigger than the proportion of others (100 individuals), they cannot survive much longer in the ecological competitions where wealth rules the reproduction rate.

To explain these results we should rely on experiment initial settings. The population of mean-variance traders is heterogeneous with respect to their risk preferences, that define the composition of their optimal portfolios. Kallberg and Ziemba (1983) provide guidance regarding the significance of the changes in risk aversion for optimal portfolio composition. Thus, mean-variance traders have different preferences for different stocks. On the other hand, the population of naive agents (as well as market portfolio holders) is completely homogeneous. All investors from this group invest the same amount of wealth in the same assets $\varpi_{j,t}^{0,N} = \varpi_{j,t}^{1,N} = \cdots = \varpi_{j,t}^{I,N}$. No trade can occur within this group. The trading success of naive strategies directly depends on the desire to trade of agents characterized by rational diversification rules.

Suppose that, on a particular day, naive investors try to reoptimize their portfolios. For this purpose, they all should buy security A, sell security B, and buy security C. At various times during the day, the investors from population x_k would try to purchase the desired numbers of shares A and send bid orders. Once the supply offered on the order book related to stock A is used up (because there is no the same number of traders willing to sell stock A), a naive strategy will raise the current bid a bit, and the next a bit more,

the next a bit more, increasing bid-ask spread. It makes the buyers to accept a price proposed by sellers and to conclude the unprofitable transactions.

4.2.4.2 Ecological competition 2: Sharpe ratio

We measure the Sharpe ratio in order to estimate the agents ability to hedge the portfolio risk with many assets. Figure 4.5(b) reports the average evolution of agent proportions based on this indicator. Note that the Sharpe ratio is not consistent measure if the mean return of portfolio has a negative value. In this case high standard deviation improves the Sharpe ratio, which is exactly the opposite of what the investor prefer. In our simulations, as far as an agent gets negative return of his portfolio, he is regarded as "ran out" of the market. The evolution of the population is guaranteed by the holders of portfolio with positive return. The unconstrained strategies outperform the constrained ones. These results confirm those of Levy and Ritov (2011), who stress the importance of short selling in markets with many assets. At the same time, our results are not congruent with those of DeMiguel *et al.* (2009) who report that the Sharpe ratio of sample-based mean-variance strategy is much lower than that of naive strategy. There are several issues that explain such discrepancy of the results. One reason is that DeMiguel *et al.* (2009) use diversified portfolios with low volatility in their numerical simulations while our simulations rely on individual assets with more volatility. The other reason is that mean-variance traders in our model use more information-rich daily return data. The final issue that plays a key role in explaining of simulation results is "predictive power" of mean-variance traders (see section A.4 for technical details). As the agents introduced in the current model are price takers, they are able to produce price dynamics with realized statistics close the expected statistics imposed at the beginning of simulations. It increases their estimation accuracy, and makes their portfolios more stable.

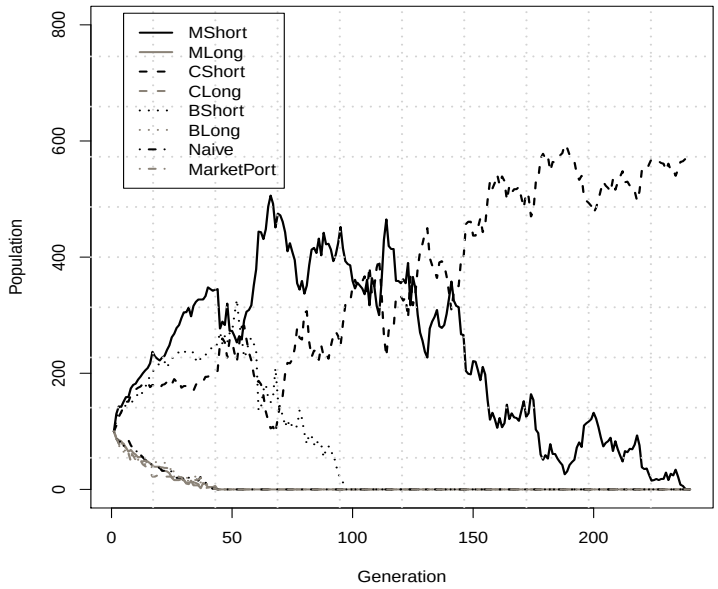
In addition to the Sharpe ratio, we also report the portfolios turnover. This indicator provides evidence about the portfolio "stability". The total portfolio turnover is calculated as follows:

$$\vartheta^i = \sum_{t=1}^{T} \sum_{j=1}^{j=J} |\varpi_{j,t}^{i,*} - \varpi_{j,t-1}^{i,*}| \ , \forall i = 1, I \tag{4.11}$$

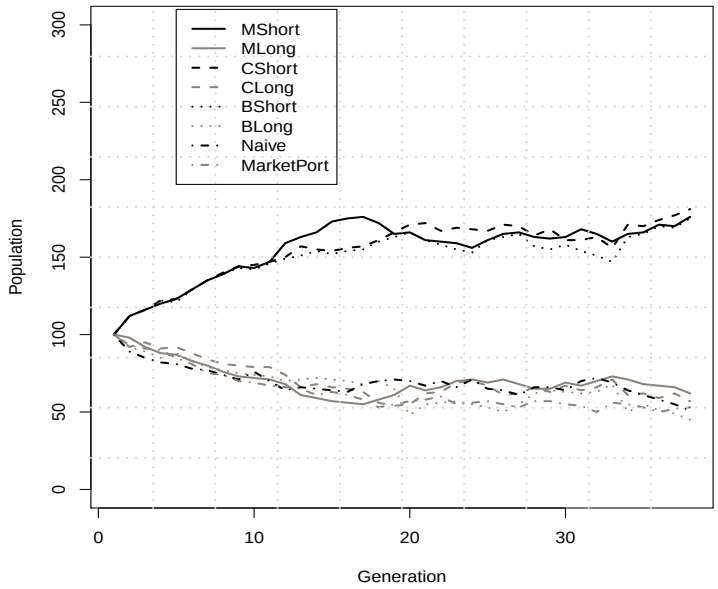
where $\varpi_{j,t-1}^{i,*}$ is agent i's portfolio weight in asset j at time $t-1$ (or target weight); $\varpi_{j,t}^{i,*}$ is the portfolio weight before rebalancing at time t; T is the number trading days in the investment period (one year); J denotes the total number of stocks.

Table 4.7 reports the average Sharpe ratio and the average portfolio turnover; it allows to have a deeper understanding about the effects of transaction costs on portfolio performance. The highest portfolio turnover, the highest reallocation volume and as a consequence, the highest transaction costs. The reason why naive strategies perform poorly is visible in the turnover amount. The portfolio turnover of agents characterized by following a naive diversification strategy is two times larger compared with those of rational portfolio optimizers. Naive diversification strategies appear to produce more unstable portfolio weights, requiring larger trading volume when rebalancing, as a results, naive agents incur more transaction costs.

To sum up, we report that classical mean-variance optimization rules still outperform the naive rules in artificial market framework where the price dynamics is a direct result of agents' trades. Our findings are consistent with those of Tu and Zhou (2011) and Levy and Ritov (2011). The performance of unrestricted portfolio strategies outperforms the long-only and naive strategies in both ecological competitions where the Sharpe ratio or the earnings rule the reproduction rate of the populations. The reason behind this performance



(a) Wealth



(b) Sharpe ratio

Figure 4.5: Ecological competition.

Strategy	Avg. Sharpe Ratio[a]	Avg. Portfolio Turnover
N	0.690896	**7.022831**
MP	0.241616	6.275086
$MLong$	0.957809	4.361393
$MShort$	0.909716	4.326295
$BLong$	**1.062356**	4.361645
$BShort$	0.954551	4.347188
$CLong$	0.95105	4.384113
$CShort$	0.892726	4.321861

Table 4.7: Table represents the average Sharpe ratios and average turnover amounts (calculated according to equation 4.11) of 8 trading strategies, held by 100 agents each over 250 days ($\simeq$ 1 year trading), 3 assets. [a] the average Sharpe ratio is calculated as follows $= \frac{1}{I} \times \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{I} Sharpe_t^{i,*}$ The highest Sharpe ratio and the highest turnover are in bold.

can be at least partly attributed to the "predictive power" of mean-variance agents, their heterogeneity with respect to risk aversion, and higher stability of their portfolios, resulting in less trading costs. These results suggest that with appropriate combinations of degree of risk aversion and rebalancing frequency, an investor can significantly improve his portfolio performance. We search for optimal degree of risk aversion (section 4.2) and reoptimization policy (section 4.4) that maximize investment earnings. Our analysis also reports that even though the ex-ante parameters estimation of moments and co-moments involves estimation errors due to the small size of sample, the combination of mean-variance sophisticated rules and naive rules can improve the performance of their individual counterparts.

4.3 Risk Aversion Impact on Investment Strategy Performance

4.3.1 Introduction

The degree of risk aversion determines portfolio holdings and subsequently the distribution of wealth. In financial markets there is a trade-off between the risk involved and the expected returns. Risky financial securities should generate, in equilibrium, a return higher than the one of the safer investments such as Treasury Bills (Mehra and Prescott, 1985). For example, the CAPM (Sharpe, 1964) assumes a linear risk-return relationship $\mu_{P,t} = r_f + \beta_P \sigma_{i,t}$, where μ_P is a portfolio expected return, r_f the risk-free rate, β_P the portfolio beta and $\sigma_{i,t}$ the market risk premium. Risk preferences of investors have a direct impact on their investment decisions. A risk-averter (or conservative) investor tends to hold more Treasury Bills than a risk-lover (or aggressive) investor who will tend to invest in riskier stocks with higher expected return. Thus, risk aversion affects the portfolio composition of investors and therefore the distribution of future wealth. In other words, each trader invests his capital in a portfolio reflecting his risk-aversion.

In this section we address the question whether investors' survivability depends on their risk preference. This work is motivated by empirical studies focusing on the relation between risk aversion and wealth dynamics (see for example, Levy (2005)). Several agent-based simulations researches have also investigated this question. In fact, some simulation-based works lean towards a framework where investor optimal decisions depend on their wealth, which is in line with the assumption of CRRA utility function (Levy *et al.*, 1995, 2000). Chiarella and He (2001) investigate the characteristics of asset prices and wealth dynamics arising from the interaction of heterogeneous agents with CRRA utility. Levy *et al.* (1995) study the effect of heterogene-

ity of preferences, expectations and strategies on wealth and price dynamics with CRRA and logarithmic utility functions using a microscopic simulation approach. Chen and Huang (2004b), Chen and Huang (2004a), Chen *et al.* (2007) investigate relative risk aversion (later RRA) relation to wealth dynamics (CRRA utility function) and relationship between RRA and survival dynamics (CRRA, CARA, Logarithmic, CAPM). They find that only the CRRA investors with the RRA coefficient close to one can survive in the long-run time framework.

4.3.2 Simulation settings

We consider a securities market populated by a finite number of mean-variance traders with *heterogeneous* preferences, indexed by $i \in 1, 2, ...I$ ($I = 1000$). These agents have an open access to information concerning the underlying probability distribution of security prices. This population heterogeneous only with respect to the degree of risk aversion, which is uniformly distributed in $[0.1, 10]$. They enter the market with 50 units of each class of assets and \$1000 in cash. These agents rely on the same estimation window length for estimating the covariance matrix and returns and use the same rebalancing frequencies.

As the mean-variance traders send only a few orders daily to rebalance their portfolios, we implement a special type of agents, which provide liquidity: the *liquidity providers*. These agents do not seek to increase their wealth or decrease their portfolio risk, they trade rather randomly. These traders allocate in the risky assets a random fraction of wealth (uniform distribution).

$$\varpi^i_{j,t} \sim U(x|0,100)\ \forall j = \overrightarrow{1,J} \tag{4.12}$$

$$\varpi^i_{j,t} = \frac{\varpi^i_{j,t}}{\sum_{n=1}^{J} \varpi^i_{n,t}}\ \forall j = \overrightarrow{1,J} \tag{4.13}$$

$$\sum_{n=1}^{J} \varpi^i_{n,t} = 1 \tag{4.14}$$

The equations 4.12–4.13 guarantee a random distribution of weights in [0,1], with a total sum equal to 1. Thus, we do not allow the liquidity providers to have short positions or an aggregate negative wealth. The desired quantity of stocks is then calculated based on the formula (4.6).

As liquidity providers do not seek to optimize their positions in the market, we do not compare their performance with the other agents' results.

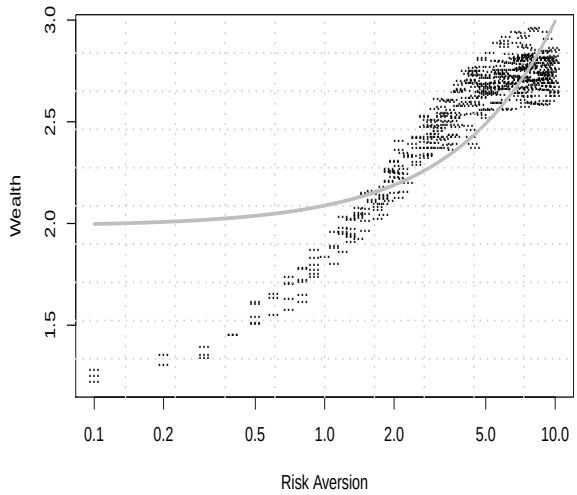
In addition, Milgrom and Stokey (1982) or Fudenberg and Tirole (1991) stress the necessity of heterogeneous expectations, different opinions and trading rules in the market. As mentioned in several publications Shatner *et al.* (2000), Hommes (2006), heterogeneity of time scale for the agents' actions is an important feature to obtain realistic price dynamics. We thus introduce different trading frequency $\Theta \sim U(x|10,1000)$ for different investors; said differently, agents rebalance their portfolio every Θ rounds. This condition helps avoid an empty order book, which could result from a general cancellation of orders by all the agents at the same time.

Results and Discussions First of all, we estimate the performance of trading strategies based on the end-of-the-period values like in most models dealing with similar research question. Then, we put the agents in a competitive market such that the populations of investors co-evolve: agents change their strategy between the trading periods based on their historical performance. Finally, we compare the results.

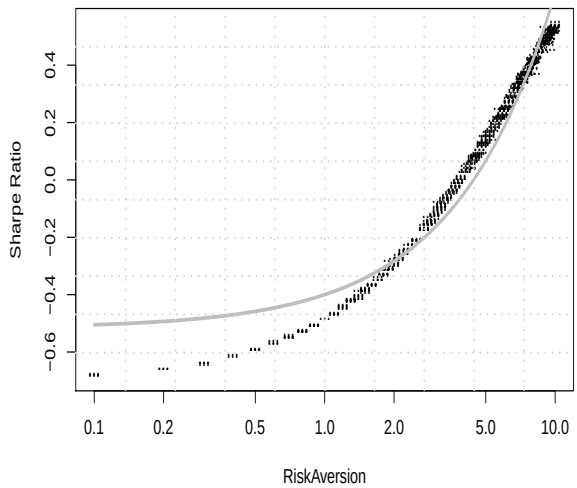
We run 1500 days of trading (which corresponds to 6-year or 15 trading periods, 100 days each). For the first trading period (100 days) we provide the initial statistics for the traded assets to the mean-variance traders. During the next periods, agents calculate assets statistics themselves, based on the generated price series. The traders do not change their risk preferences and their trading strategies between periods (in an ecological competition framework this constraint will be relaxed). We run 100 simulations with different initial asset statistics. We also test short-selling and long-only cases. We begin by discussing the *3-asset case*.

Figure 4.6(a) we depict the relationship between agents' risk preferences and their wealth distribution. On the horizontal axis we set out the different initial parameters – risk aversion between 0.1 and 10, with 0.1 as an increment in log-scale. The vertical axis shows the final wealth corresponding to these different initial parameters. A great difference between the wealth distribution and its linear regression fitting (a gray solid line) indicates that the wealth increases sharply for agents with risk aversion from 0.1 to 3.5. Thereafter, it increases smoothly. This behavior can be explained by the composition of the optimal portfolio. Kallberg and Ziemba (1983) provide guidance regarding the significance of the changes in risk aversion for optimal portfolio composition. Agents with $A > 4$ are very risk averse and prefer portfolios with low variance. If the degree of risk aversion A is superior of 4, the portfolio composition does not vary even for large changes in A. Range $2 \leq A \leq 4$ yields moderately risky portfolios with a modest degree of change in the optimal portfolio with changes in parameter A. The range $0 \leq A \leq 2$ yields risky portfolios and there are dramatic changes in the target weights for even small changes in A.

We also investigate risk-adjusted reward to the volatility of individual portfolios, also known as the Sharpe ratio. Observing the Sharpe ratio dynamic over different risk aversion frameworks (see figure 4.6(b)), we get similar results as Chen *et al.* (2007). Even if high-risk-averse agents choose assets with



(a) Wealth Distribution for Agents with Different Risk Aversions. *Regression coefficient*= 0.10065

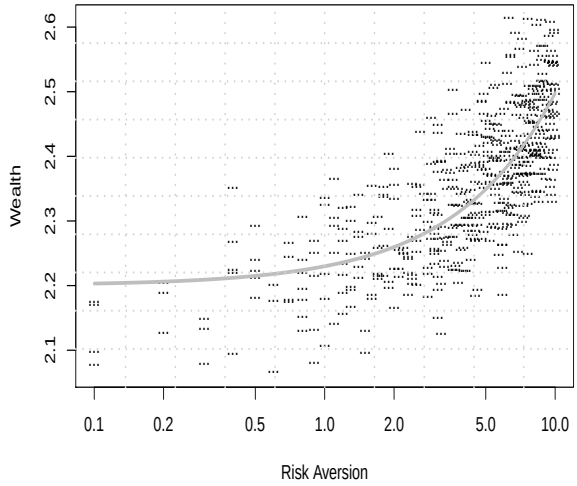


(b) The Sharpe ratio Distribution for Agents with Different Risk Aversions. *Regression coefficient* = 0.11678

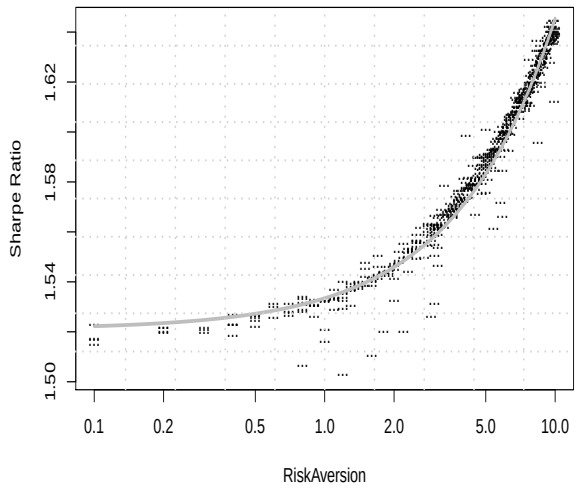
Figure 4.6: 3-asset long-only case. Each point is the averaged value of 100 simulations. X axis is in log-scale

low risk and low return, they earn a higher the Sharpe ratio and a higher final wealth. This effect can be explained by the mathematical properties of the efficient frontier. The first derivative of portfolio return μ_p with respect to portfolio risk σ_p indicates that large values of A (the minimal variance portfolio) correspond to a big slope on the efficient frontier. Hence, conservative investors get significant increase in portfolio returns by bearing a small amount of extra risk. The slope becomes smaller when A decreases. The second derivative of μ_p with respect to σ_p is negative, which means that the efficient frontier is concave. For large values of A, the second derivative has a large negative magnitude, so the slope is sharply decreasing. With $A \to 0$ the slope decreases much more slowly. Contrary to Chen and Huang (2004b) and Chen *et al.* (2007), in our simulations, less risk-averse agents ($A < 1$) do not run out of the market, even if, on average, they obtain a lower gain than risk averters ($A > 1$). If the number of assets remains relatively small and short selling is allowed, the Sharpe ratio distribution in relation to risk aversion is close to that received with long-only constraint. Thus, the 3-asset short-selling case is not considered in current work.

We continue to increase the number of trading assets. We now consider a 20-asset long-only case. The simulation results are presented in figures 4.7(a) and 4.7(b). Wealth has not such a sharp increase as in the 3-asset case (the linear regression coefficient now equals 0.02977): it rather increases smoothly with the increasing risk aversion. This behavior can be explained by the fact that the portfolio composition is affected differently by the changes in A for different number of asset classes that constitute the optimal portfolio Kallberg and Ziemba (1983). The Sharpe ratio has an increasing dynamics when risk aversion increases, but the difference between the maximum and the minimum values is relatively small ($1.644210 - 1.502465 = 0.141745$). Thus, we can conclude that risk aversion has a relatively small effect on the variations of the Sharpe ratio.

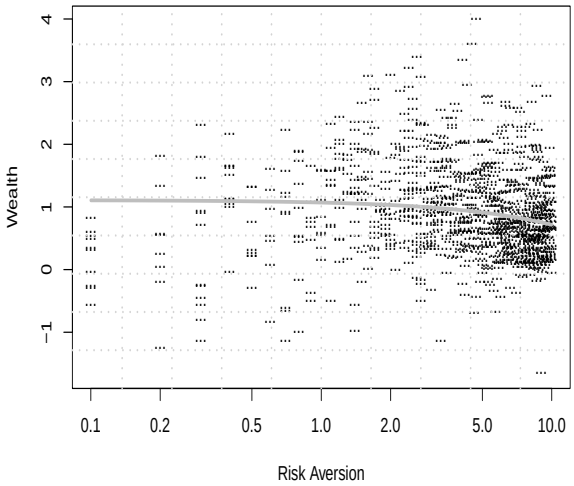


(a) Wealth Distribution for Agents with Different Risk Aversions. *Wealth Increasing* $= \frac{W_t}{W_0}$. *Regression coefficient*= 0.02977

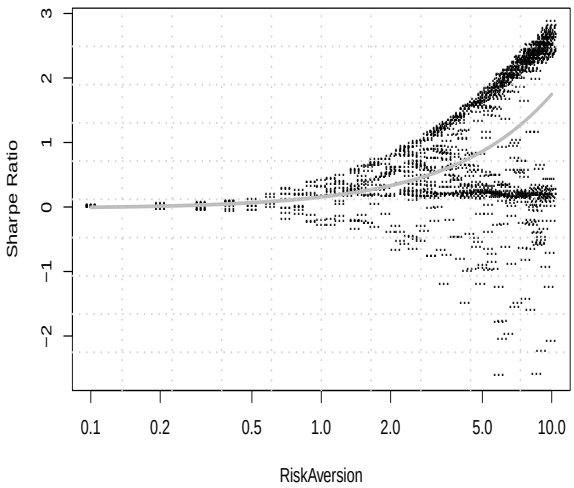


(b) the Sharpe ratio Distribution for Agents with Different Risk Aversions. *Regression coefficient*= 0.01244

Figure 4.7: 20-asset long-only case. Each point is the averaged value of 100 simulations. X axis is in log-scale



(a) Wealth dynamic for agents with different risk aversions. *Wealth Increasing* $= \frac{W_t}{W_0}$. *Regression coefficient*$= -0.03942$



(b) the Sharpe ratio dynamics for agents with different risk aversions. *Regression coefficient*$= 0.17709$

Figure 4.8: 20-asset short-selling case. Each point is the average value of 100 simulations. X axis is in log-scale

As soon as short-selling is allowed, one part of the risk taking agents runs out of the market, while other agents with the same risk preferences $A < 1$ obtain a much higher wealth than in the constrained-portfolio case. Thus, there are two possibilities for the risk taking agents: either they lose their initial endowment, or they increase their wealth by a factor much higher than the one of risk-averters. The conservative agents (risk-averters). On the one hand have a moderate wealth increase factor, on the other hand, they have few chances to lose their initial wealth (see figure 4.8(a)).

Figure 4.8(b) as well as the regression coefficient (0.17709) show that, contrary to the constrained portfolio situation, risk aversion has a significant effect on the Sharpe ratio when short selling is allowed. Even if the Sharpe ratio distribution exhibits a higher variance when the risk aversion increases, conservative agents tend to considerably improve their the Sharpe ratio. We can conclude that in the unconstrained portfolio framework it is better to be risk averse and to invest in risk-free assets.

4.3.3 Ecological Competition Analysis of Strategy Performance

Next, we compare the relative performance of investment strategies using ecological competition. Initially, we consider an environment with $N = 5$ families of traders[1] with populations x_i $i = \overrightarrow{1, N}$ in equal proportions (200 agents each) who interact through trading to obtain the *highest possible wealth.* The only difference in population strategies is the risk preferences. $2 \times (i-1) < A_i \leq 2 \times i$ is the risk aversion measure for the population x_i, $i = \overrightarrow{1, 5}$.
The total number of agents is thus equal to 1000 and remains constant over the simulations.
The strategy proportions of each family within this population are updated

[1]Liquidity providers are not included in competitions

every n step corresponding to a "simulation generation" according to their wealth ($x_i = X\frac{W_i}{W_T}$, where W_i is the wealth earned by agent population i, W_T is the total wealth). One generation of competition corresponds to 100 trading days.
A family of agents is said to be run-out of the market if $x_i = X\frac{W_i}{W_T} < 1$. An analogous allocation principle is used with the Sharpe ratio instead of the wealth criterion for the second set of simulations.

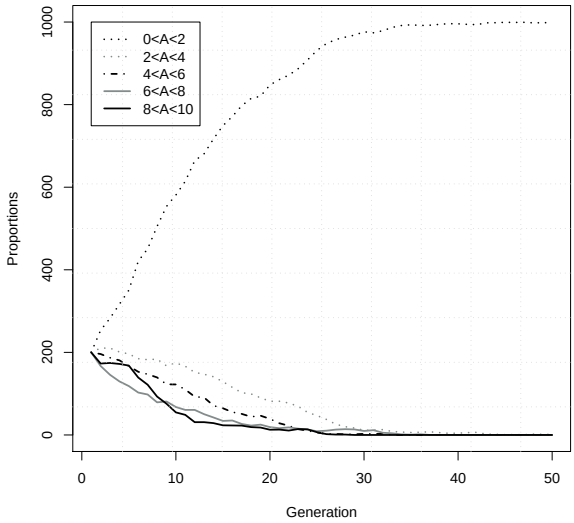
Ecological Competition Analysis: Short Selling Allowed

Figures 4.9(b) and 4.9(a) confirm the results highlighted only with the end-of-period results (see figures 4.8(a) and 4.8(b)). When short selling is allowed, the risk lovers compete with other agents in terms of wealth but quickly run out-of the market in the competitions where the Sharpe ratio is used as a performance measure.

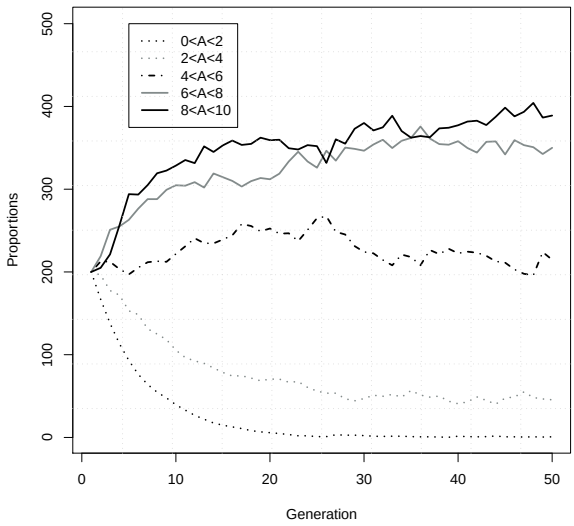
Ecological Competition Analysis: Long-only

In the case of long-only constrained portfolios, the figure 4.10(a) shows that the highest (as well as the lowest) risk aversion values do not guarantee the highest earnings. Risk lovers ($0 < A \leq 2$) as well as absolute risk averters ($8 < A \leq 10$) run quickly out-of the competition (in ≈ 100 rounds). Only the traders with a moderate level of risk aversion $4 < A < 6$ survive in the long run (> 500 rounds).

In the competition based on risk adjusted returns, the conservative traders slightly outperform the aggressive ones (figure 4.10(b)). These results are consistent with those previously presented in figure 4.7(b). A similar conclusion emerges: risk aversion does not have a significant impact on the Sharpe ratio improvement.

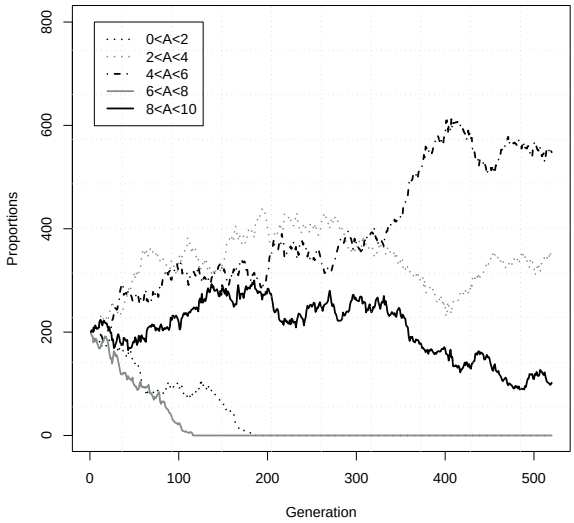


(a) Wealth dynamics for agents with different risk aversions.

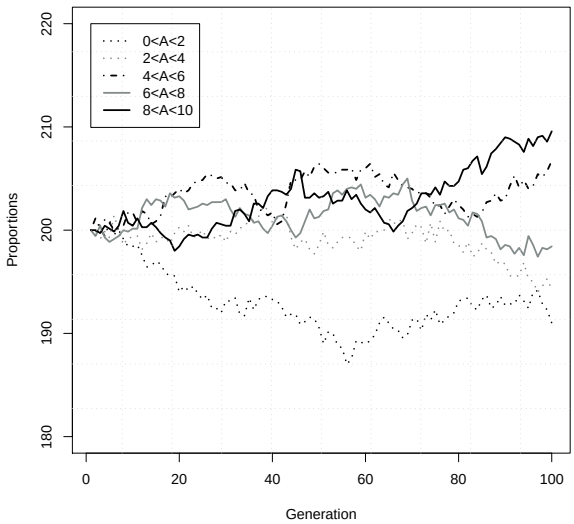


(b) the Sharpe ratio dynamics for agents with different risk aversions.

Figure 4.9: Ecological competitions: 20 assets, short-selling allowed. Strategies are grouped by two for the sake of results tractability



(a) Wealth dynamics for agents with different risk aversions.



(b) The Sharpe ratio dynamics for agents with different risk aversions.

Figure 4.10: Ecological competitions: 20 assets, long only. Strategies are grouped by two for the sake of results' tractability

If the market is populated by agents with constrained portfolios and agents with unconstrained portfolios, traders using short-selling easily win the competition for wealth. A possible explanation for these phenomena is that the portfolio performance is improved because traders sell the assets that outperform ("sell overprices assets") and buys the assets that underperform during the trading period ("buy underpriced assets"). According to Levy and Ritov (2011) the long-only strategies have zero-positions ($\alpha_{j,t}^{i,*} = 0$) in about 50% of the traded assets. Thus, the agents with long-only strategies rebalance only half of their investment set to maintain their target weights. At the same time, the agents with short-selling strategies trade the whole set of assets, and increase their wealth more efficiently.

4.3.4 Results and Discussions

The computational experiments performed in this research show a number of important results. First, we show that the degree of risk aversion significantly affects the survivability of agents and their portfolio performance. However, we cannot identify a unique and absolute "winner of game". We also highlight that the agents' profits depend on market conditions and other market participants as well.

By overcoming models based on fixed proportions of agents, we conclude that the final wealth as well as agents' risk adjusted return not only depend on their accuracy to predict expected returns and covariances of assets, but also on their risk preferences.

Our model based on ecological competition characterizes the evolution of agent populations when traders switch from the old strategy to a new one (by adjusting their risk preferences) according to its performance in the past. The main assumption is that all agents belonging to a group share the same risk preferences (risk aversion range $[A_{min}, A_{max}]$), but are allowed to change

groups between the trading periods. In such a way, the fraction of agents using the same strategy characterizes its success in the past.

We report that when random traders and unconstrained mean-variance traders populate the market, risk lovers ($A < 2$) outperform others when wealth is used as the basis for ruling reproduction within the agent population. However, they quickly run out of the market in the competitions based on the Sharpe ratio. In that last case, only conservative traders survive in the long run.

Furthermore, when short selling is forbidden (long-only case), the highest, as well as the lowest risk aversion rates do not guarantee the highest earnings. Aggressive ($A < 2$) and strongly conservative ($A > 8$) traders run quickly out of the competition for wealth. Conservative traders beat aggressive traders in the competition for a higher the Sharpe ratio of the portfolio.

4.4 Reabalancing Frequency Impact on Investment Strategy Performance

A high rebalancing frequency reduces the portfolio performance due to transaction costs, whereas a low rebalancing frequency hides a risk not to react in time to important market changes. Optimal rebalancing frequency helps not only to control the risk, but also to enhance the portfolio return. In the absence of transaction costs and when risky asset prices follow geometric Brownian motions, the optimal investment policy is to constantly trade in order to keep a constant dollar amount in each risky asset (Merton, 1971). In the presence of transaction costs, trading continuously incurs infinite transaction costs. There exists a series of works modeling an optimal tradeoff between rebalancing benefits and rebalancing costs. Akian, Menaldi and Sulem (2004) consider an optimal investment problem with proportional transaction costs and use numerical simulations to compute the no-transaction region. Liu (2004) solve numerically the problem of the optimal transaction policy when the risky asset returns are uncorrelated. The author shows that the optimal investment policy in each risky asset is to keep the dollar amount invested in the asset between two constant levels. Once the amount reaches one of these bounds, the investor trades to the corresponding optimal targets. Walter, Ayres, Chen, Schouwennars and Albota (2006) propose the dynamic programming-based approach to construct a policy to trade only when the costs of rebalancing is less than the cost of doing nothing.

There are two common methods for portfolio rebalancing: periodic (calendar) (Donohue and Yip, 2003) and tolerance bands (Masters, 2003). The calendar based approach is actively criticized in literature Walter *et al.* (2006). This method relies on the fact that, on average, the portfolio becomes less and less optimal, but it does not take into account the real portfolio and market state. For this reason, in this research we consider only tolerance band (or

drift-based) rebalancing method.

Our first contribution in this section is to derive the rebalancing policy using agent-based simulations. An agent rebalances his portfolio in order to maintain a long-term goal for asset allocation. The choice of rebalancing frequency is an essential for reaching long-term objectives. Transaction costs as well as taxes make frequent rebalancing highly unattractive. At the same time, without rebalancing, the portfolio becomes less diversified and is subject to greater volatility.

4.4.1 Tolerance band rebalancing

When the market is stable, the portfolio is reasonably close to its target allocations, hence, an investor can rebalance his portfolio only when the current weights run far from the targets. But how far is too far? To determine when rebalancing is efficient, we should compare the benefits from rebalancing and rebalancing costs. Moreover, the costs of rebalancing are increasing as the drift out of target weights increasing. There are two main benefits of rebalancing: performance improvement and risk control. Portfolio performance is improved because the trader sells the assets that outperform (overpriced) and buys the assets that underperform over current trading period (underpriced).

Tolerance band rebalancing is the method to trade only when the weight of an asset class drifts outside the tolerance ranges. It can be more effective than calendar-based rebalancing (rebalancing quarterly or annually) (Overway and John, 2006). The optimal transaction policy is to trade only when the current weights are far from the targets. Hence, the optimal way is to rebalance only when the current portfolio weights run out the tolerance ranges. Portfolio return tends to increase as tolerance is widened, but once it reaches the certain limits return declines. We try to find an optimal tolerance range of deviation around the target weights.

There exit two types of tolerance ranges:

1. *Absolute range* is the fixed band for all assets. A commonly used range is ±5%, that is the reasonable band.

2. *Relative range* is the relative band around each of target weights.

Tables 4.8 and 4.9 demonstrate the possible bands in case of absolute and relative tolerance ranges. Algorithm 3 and Algorithm 4 aim at clarifying the difference between absolute and relative range rebalancing.

Target weight	min weight	max weight
30	20	40
20	10	30
50	40	60

Table 4.8: 10% absolute tolerance range

for $j = 1$ **to** J **do**
  $\alpha^i_{j,t} = \frac{\sum_{j=1,J} p_{j,t} q^i_{j,t} + C^i_t}{p_{j,t} q^i_{j,t}}$
  if $\alpha^i_{j,t} - \alpha^i_{j,0} > T^i$ **then**
    Rebalance all portfolio
  end
end

Algorithm 3: Absolute tolerance band. $\alpha^i_{j,0}$ target weight of wealth hold by agent i invested in the stock j. $\alpha^i_{j,t}$ current weight of wealth hold by agent i invested in the stock j at the moment t. T^i tolerance band for weights deviation of agent i

Target weight	min weight	max weight
30	27	33
20	18	22
50	45	55

Table 4.9: 10% relative tolerance range

```
for j = 1 to J do
    α_t = (Σ_{j=1,J} p_{j,t} q^i_{j,t} + C^i_t) / (p_{j,t} q^i_{j,t})
    if (α^i_t − α^i_{j,0}) / α^i_{j,0} > T^i then
        Rebalance all portfolio
    end
end
```

Algorithm 4: Relative tolerance band. $\alpha^i_{j,0}$ target weight of wealth hold by agent i invested in the stock j. $\alpha^i_{j,t}$ current weight of wealth hold by agent i invested in the stock j at the moment t. T^i tolerance band for weights deviation of agent i

Additionally, several research studies indicate that it is not cost-effective to return a portfolio completely the way back to its initial allocations. The best way to minimize transaction costs is to do halfway back to initial weights. Masters (2003) finds the optimal tradeoff between the geometrically (quadratically) increasing benefit of rebalancing and the linearly increasing transaction costs. Masters (2003) argues that it is optimal to rebalance back only to halfway between the target weight and boundary limit. It has been shown that this strategy reduces transaction costs by approximately 50%.

Stewart (2005) empirically shows that it is better to take the portfolio back to the target weights, but it is more profitable to take it beyond the targets weights. The results of the period he examines and the assets he considers show the benefit to overweight the assets that had underperformed and underweight assets that had outperformed.

In the current work we test both methods of rebalancing *complete way back* and *halfway back* to the target weights; *absolute* and *relative* bands of deviation from targets (see figure 4.11).

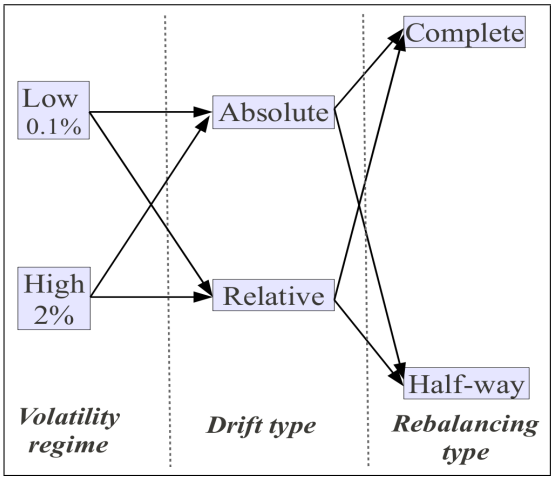


Figure 4.11: Scenarios of simulations

4.4.2 Simulation settings

We introduce 1000 mean-variance traders holding tree-asset portfolios. Each of them has tolerance for deviation from the target allocation $T \in \{0\%, 1\%, 2\%, ..., 100\%\}$. Whenever a weight of an asset class drifts outside its tolerance for deviation range, a portfolio is rebalanced.

The mean-variance agents are homogeneous with respect to their degree of risk aversion A, that equal to 3. This choice is made based on the simulations described in section 4.3. The degree of risk aversion 3 corresponds to the moderate level of risk preferences. This parameter is constant during the simulations.

We specify two market regimes: low volatility and high volatility (see table 4.10). Decreasing of tick size results in narrower bid-ask spreads. This directly affects market liquidity and volatility. When the tick size is made narrower and possible increments are finer, then potential price changes may be smaller, thereby resulting in less variable price changes. Hence, to increase market volatility, we increase a tick size. In periods of high volatility transaction costs are higher than those of the low volatility regime. We consider 0.1 %

	Low volatility			
	μ	σ^2	*skew.*	*kurt.*
S_1	0.0015	0.0181	0.1524	5.7657
S_2	0.0026	0.0418	-0.0209	4.1524
S_3	0.0009	0.0264	-0.4681	3.4666
	High volatility			
	μ	σ^2	*skew.*	*kurt.*
S_1	0.0021	0.12257	0.02051	3.8434
S_2	0.0014	0.08399	-0.2971	2.5589
S_3	0.0012	0.06221	-0.0946	2.6621

Table 4.10: Statistical properties of each asset class. Volatility regimes are parametrized by tick size.

transaction costs in low volatility regime, and 2% for high volatility regime. Transaction costs are not included in the budget constraints. The capital C is used for both purposes to buy the securities and to pay the transaction costs. Thus, agents can run out of market (get negative total wealth).

Finally, as in section 4.1, we introduce the liquidity providers that behave rather randomly and simplify price settings.

4.4.3 Results and Discussions

The averaged results of 100 runs of each scenario are summarized in table 4.11. In the low-volatility regime, the agents characterized by $0 < T \leq 20$ absolute range and halfway rebalancing rules, slightly outperform the others. The wealth distribution of these agents exhibits high kurtosis (≈ 18) and high skewness (1.004011), indicating that majority of agents get average wealth, while the minority tend to outperform this average level. We can conclude that the absolute drift range ($0 < T \leq 20$) and halfway-back rebalancing rules is an optimal tradeoff between rebalancing costs and rebalancing benefits. Moreover, the distribution of wealth in figure 4.12 reports that even though 5% tolerance band is actively used by market practitioners, in low-volatility

conditions, they can tolerate until 20% deviation from the targets without significant losses.

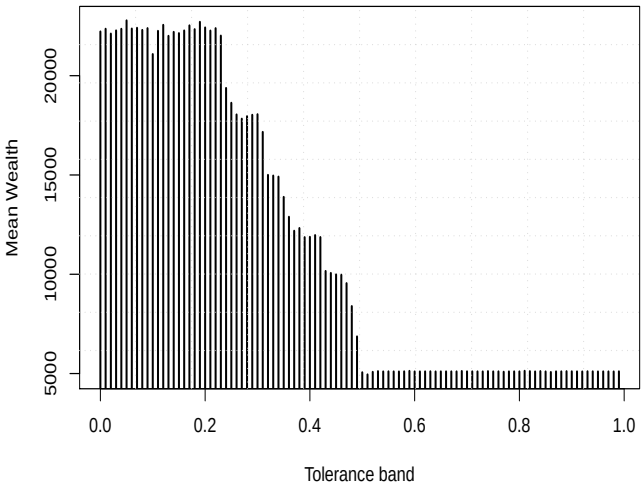


Figure 4.12: Wealth distribution. 0.1% transaction costs, low-volatility regime, absolute tolerance band, halfway back rebalancing

In high-volatility regime, a frequent rebalancing or a large tolerance band is not an optimal solution for gain maximization or even for risk adjusted return improvement. A portfolio held by agents with low tolerance for deviation from targets is a subject to high transaction costs (see figure 4.13).

The portfolio of agent with high tolerance for deviations is the subject to high risk (see figure 4.14). The agent with low tolerance for deviation reduces his portfolio risk by frequent rebalancing. However, he incurs wealth losses due to the high transaction costs (2%). The results in the table 4.11 basically point out that the halfway back rebalancing helps improve the Sharpe ratio but it has a rather small impact on wealth increase.

The findings in this section basically point out the optimal rebalancing policy in low- and high-volatility market in the presence of transaction costs. Our simulations results confirm the suggestions of Masters (2003) that the ab-

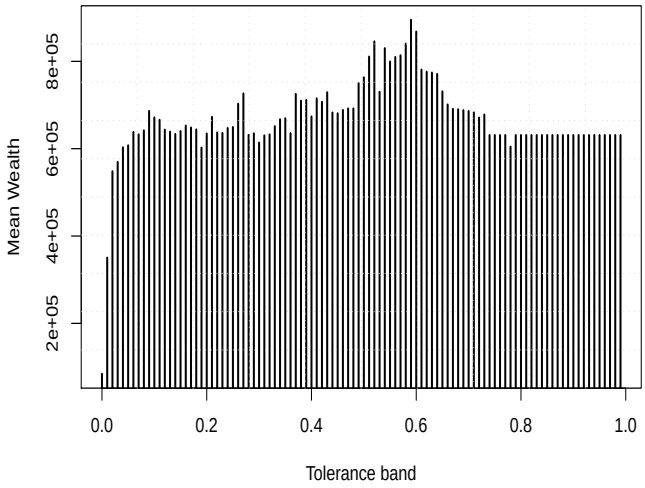


Figure 4.13: Wealth distribution. 2% transaction costs, high volatility regime, absolute tolerance band, complete rebalancing

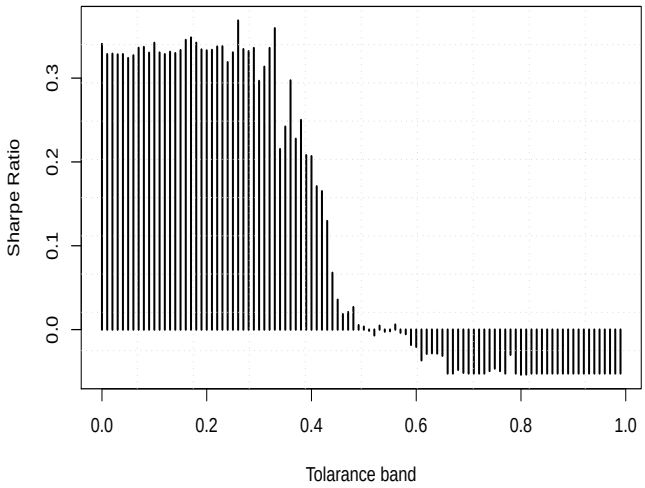


Figure 4.14: the Sharpe ratio distribution. 2% transaction costs, high volatility regime, absolute tolerance band, complete rebalancing

solute drift range ($0 < T \leq 20$) and the halfway back rebalancing is a tradeoff between small repeated tradings and large costs for high volume rebalancing of largely declining portfolio from its targets in low volatility conditions. However, the same rules have rather negative effect in high volatility regime. The agents with low tolerance for deviations from targets loose easily their initial wealth, even if they improve the Sharpe ratio of portfolio. These observations suggest a higher flexibility in investment strategies: to rebalance less frequently, but look more frequently at price dynamics to find the best rebalancing opportunities.

	Complete			Halfway				
	$0 < T \leq 20$	$20 < T \leq 40$	$T > 40$	$0 < T \leq 20$	$20 < T \leq 40$	$T > 40$		
min	4480.55	1451.53	1383.35	**11055.09**	4732.24	3746.65	Absolute	Low
max	91831.1	4974.11	5010.28	**53248.75**	50417.45	25106.58		
mean.	9142.87	4138.10	4081.17	**22291.08**	16578.84	7349.90		
skew.	4.03639	-1.36271	-1.2016	**1.004011**	1.362383	1.671348		
kurt.	14.7653	0.32726	-0.43695	**18.86696**	9.86747	2.662102		
Sharpe	1.3000625	1.3109093	1.3171058	**1.431137**	1.424201	1.408475		
min	4181.64	4736.14	3935.22	9638.94	9633.82	1761.14	Relative	
max	5105.05	5245.77	5185.38	31773.78	31493.45	32387.11		
mean.	5060.69	5082.15	5029.46	21949.21	21918.38	11958.63		
skew.	4.70747	-1.509016	-3.369181	-0.5582034	-0.5535263	0.9118452		
kurt.	21.5114	2.521223	12.12310	-1.354962	-1.366551	-0.7340671		
Sharpe	1.1779515	1.1853119	1.1845493	0.9884437	0.986747	0.954649		
min	593.24	5201.98	5861.35	1223.87	2007.83	4579.79	Absolute	High
max	8161.01	8453.41	11321.10	2137.45	5831.42	6740.29		
mean.	6391.73	6576.76	7453.12	1380.12	4901.85	5702.5		
skew.	-2.350237	0.6809661	1.105914	1.886898	-1.246212	-0.3178618		
kurt.	5.466913	-0.3249746	-0.2448154	1.691031	-0.3578788	-0.7341588		
Sharpe	0.2989989	0.2006190	-0.01141703	0.2618616	0.2808233	0.1582175		
min	963.35	5804.17	6369.36	936.38	928.06	937.75	Relative	
max	6586.71	7081.81	7028.13	1453.61	1446.12	5731.11		
mean.	4773.17	6457.12	6581.23	1134.12	1136.81	2680.99		
skew.	-0.871587	-0.1288835	1.209252	0.6130973	0.5794839	0.4508615		
kurt.	-1.055114	-0.4300516	0.02682	-1.517962	-1.559853	-1.554189		
Sharpe	0.05324615	0.05304157	0.05468743	0.5380607	0.5378979	0.4924253		

min – minimum value, max – maximum value, mean – mean value, kurt. – kurtosis, skew. – skewness, Sharpe – the average Sharpe ratio

Table 4.11: Statistics of wealth distribution for different tolerance bands

4.5 Conclusion

In this section we have introduced the mean-variance model within the artificial stock market framework. One of the critics of agent-based models is its calibration and validation. For this reason, we have focused on these difficult issues. We have examined the ability of artificial agents to produce realistic market dynamics in different market conditions and time horizons. The re-

sults have pointed out statistical properties of the multi-asset artificial market similar to those of real market.

Next, we have applied this model to answer some important questions of portfolio optimization. The major benefit of our approach is its flexibility compared with backtesting, actively used in financial studies. Indeed, the ABM approach allows (i) any number of traders on the market (ii) combination of large variety of strategies (iii) any number of risky assets. This flexibility provides a distinct advantage over alternative approaches to the portfolio optimization problems.

Contrary to research works claiming the uselessness of the Markowitz theory, we have reported that this classical rule still outperforms the naive rules in the artificial market framework where the price is a direct result of investors' trade and transaction costs are incurred whenever security is traded. Our result has shown that naively diversified portfolios are sub-optimal. This conclusion has stressed the importance of transaction costs and individual preferences in portfolio optimization model and has motivated the studies introduced in section 4.3 and 4.4. Special attention has been paid to the effects of degree of risk aversion and rebalancing frequency on portfolio performance in the presence of transaction costs.

The importance of attitude toward risk of agents was evident in the results. However, extensive simulation results could not provide us the best behavior in all market conditions. Likewise, the simulation results have indicated that the rebalancing frequency is an important factor affecting portfolio performance. However, we could not identify a unique winner of game. This speaks for the importance of market conditions and other market participants on strategy performance. Thereby, for thorough evaluation a strategy needs to be tested in different circumstances: during periods of high and low volatility, in the presence of other heterogeneous traders. We have shown that the ATOM platform perfectly suits for this sort of experiment.

Several intriguing topics for future research arise from the results. For example, we can establish the relationship between the length of the estimation window (memory span) and portfolio performance. Consequently, it could be interesting to investigate an optimal combination of all parameters: memory span, rebalancing frequency, and risk aversion. Furthermore, the sensitivity analysis of parameter changes could be examined in more detail when various market conditions prevail.

Chapter 5

Algorithmic determination of the maximum possible earnings for investment strategies

Contents

Strategy profitability analysis is an important question in finance. For instance, there is a long debate about profitability of technical trading techniques since Fama and Blume (1966). Recent empirical studies (Brock *et al.*, 1992; Blume *et al.*, 1994) report evidence for the profitability of technical analysis, that is able to "beat the market". It is conceivable that, by repeatedly examining different trading rules using the same data set, some rules would appear to be profitable, yet such profitability may simply be due to luck. This concern is shared by academic and market professionals (Bass, 1999; LeBaron *et al.*, 1999). In order to clarify the debate, White (2000) proposes a formal test, *White's Reality Check*, on whether there exists a superior model (rule) in a "universe" of models (rules). White's Reality Check requires constructing a "full" universe of trading rules.

In this chapter we estimate the complexity of constructing a "full" universe of strategies. We also introduce a new method for the determination of the *upper-bound in terms of maximum profit* for any investment strategy applied in a given time window $[0, T]$. This upper bound is characterized once all the prices are known at time T and therefore stands for an "ex-post" maximum efficiency to any investment strategy determined during the relevant time interval. This approach, later called $S*$, allows gauging in absolute terms behaviors defined with atomic "buy" and "sell" actions and can be extended to more complex strategies. We show that, even in the "ex-post" framework, it is extremely complex to establish this upper bound when transaction costs are implemented.

This approach is also useful in agent-based framework. $S*-$*determination* approach can provide the information concerning allocative efficiency observed in artificial market. Following Smith (1962), one can define allocative effi-

ciency of markets as the total profit actually earned by all agents divided by the maximum total profit that could have been earned by all the traders (total surplus extracted). $S*$ can be a measure of allocation performance of individual traders, as the profit actually earned by the agent divided by the maximum profit that he could earn.

Sections 5.1 and 5.2 describe the context of this problem, provide simple illustration, and introduce some important terminology used later. We first describe this problem using a linear programming framework in section 5.3. Thereafter, we propose to embed this question in a graph theory framework and show that the determination of the best investment behavior is equivalent to the identification of an optimal path in an oriented, weighted, bipartite network or in a weighted directed acyclic graph in section 5.3.4. Section 5.5 illustrates this method using various real world data and makes a new point on the notion of absolute optimal behavior in the financial world.

5.1 Introduction

Performance gauging in Finance is a complicated issue that generates a series of methodological questions (Sharpe, 1991; Elton, Gruber and Blake, 1996; Malkiel, 2004). In assessing the performance of a sequence of investment/divestment actions relating to a financial asset over time (for example a particular tracker fund), two frameworks can be considered.

The first option is to adopt an ex-ante evaluation point of view, answering the following question: *"Were the choices of the investor, given his knowledge of the future at that time, optimal or not when they were realized?"*. This point of view acknowledges that investment occurs in a stochastic context and that a poor ex-post result does not necessarily indicate that bad decisions were made ex-ante, or during the decision process. Notice that this ex-ante performance assessment requires an awareness of the investor's conception of the future at

each stage in the process, and is therefore difficult to achieve in practice.

The second option is to adopt an ex-post evaluation approach, which considers only the statistical result of a given investment strategy over time, once price dynamics are perfectly known. This approach is widely used in professional asset management. For example, the performance of various *investment styles* is gauged using this technique. Financial journals use this ex-post approach to create yearly rankings and to report on the performance of asset managers and funds. In the latter case, performance is evaluated using a relative comparison among funds, as it is impossible to know what would have been the best behavior during the relevant period, or how the best output compares with the performance upper bound.

This paper can provide, in the ex-post framework previously described, the upper bound to any investment strategy in a given time window, for the trading of a single financial asset. We do not address strategic/tactical allocation or the operational process that allows fund managers to identify states in the market where buying or selling is particularly appropriate (for example in exploiting results delivered by neural network forecasting). Neither do we propose a method that ranks various strategies in terms of risk-return performance (although our approach might be extended to this bicriteria framework). Instead, we offer a computational characterization of the profits upper bound that might have been reached, by chance or skill, in trading a single financial asset during a given time-window.

Computing this limit allows the determination of an ex-post optimal strategy $S*$ that actually delivers the upper bound. We call this problem the $S*-$*determination*, and show that it is far from trivial, despite its similarity to many popular models that have frequently proved completely inefficient. Our new method delivers *an absolute performance indicator* geared towards the ex-post evaluation of a wide range of trading strategies.

This upper bound can be characterized using a linear programming frame-

work and solved with a simplex approach or with dynamic programming formalism. Nevertheless, if these methods are theoretically correct, they suffer from severe limitations in terms of computability (in the worst case, the underlying algorithm being non-polynomial for the simplex). We therefore propose to embed this question in a graph theory framework and to show that determining the best investment behavior is equivalent to identifying an optimal path in an oriented, weighted, bipartite network. We illustrate these results with real data as well as simulated algorithmic trading methods.

5.2 Elements of the game, formalizations and examples

5.2.1 Elements of the game

Consider the situation in which one investor has realized a sequence of investments/divestments for a given financial asset (a stock, an index or a portfolio) during a given time window $[t = 0, t = n]$. At time $t = n$, his actions (for example Buy, Sell) and the prices at which they were undertaken (that is, the historical price series $\overrightarrow{p} = \{p_t | t \in [0, n]\}$) are perfectly known. We do not focus on "how" the investor behavior has been formed (for example, this investor should have generated trading rules with genetic programming), or on the relevant information that are needed to do so. We rather focus on the decisions it delivered *as data* and that lead to a specific profit (or loss) at time $t = n$.

This investor has the opportunity to assess his performance with respect to the best possible behavior in this time window. This assessment can be made checking whether or not his behavior matches the absolutely optimal set of actions that could have had realized. Notice this optimal set can theoretically be computed at time $t = n$ since all the prices are known.

This comparison requires some hypotheses to be respected. The following "rules of the game" present these hypotheses and describe a formal framework in which the actual set of undertaken, compared actions can be matched against any other set of trades pertaining to the same conditions, and specifically, to the absolutely optimal set of actions.

Market liquidity: Let's assume that the prices in $[t = 0, t = n], n \in \mathbb{N}$ are those at which this investor has had the opportunity to rebalance his portfolio. We posit a price-taker framework, *i.e.*. The agent's decisions cannot affect these prices; sufficient liquidity at these prices is assumed.

The "all or nothing" general constraint: We now define a set of "rules" for this investor, in other words, a series of constraints on his behavior. These simplifications are useful in allowing rigorous comparisons between sets of actions (strategies) undertaken during a given period. In this article, these rules define an "all or nothing" behavior: whether the investor is totally invested in the risky asset or has realized all his wealth in cash:

- At the initialization stage (*i.e.* at $t = 0$), the initial wealth W_0 of the investor is composed of a certain amount of cash (C_0) and no stock ($A_0 = 0$): $W_0 = A_0 + C_0$. At date $t = 1$ (the beginning of the game) we posit C_1 to be equal to the first price of the considered time series.

- The investor must decide for each $t \in (1, n)$ one specific action with regard to the composition of his portfolio: Buy, Sell or Remain unchanged (respectively coded `B`, `S` and `U`). In other terms, the investor has to compose a "sentence" of size n using characters in `B,S,U`. The interpretation of each of these actions is as follows:

  - Buy: One can write `B` *if and only if* $W_{t-1} = C_{t-1}$. If `B` is written at date t; all the investor's cash is converted into assets (delivering a

new quantity for $A_t \neq 0$), assuming transaction costs at a $c\%$ rate,

$$A_t = \frac{W_{t-1}}{p_t \times (1+c)}$$

Additionally, the first character in any sentence must be a `B`.

- Sell: *if and only if* $A_{t-1} \neq 0$, the investor can write `S` and convert his position into cash. Considering an identical rate of transaction costs c,

$$C_t = A_{t-1} \times (p_t \times (1-c))$$

- Remain Unchanged: Whatever the nature of W_{t-1} (cash or assets), he can also decide to write `U` and let his position remain unchanged at date t: $W_t = W_{t-1}$.

- This "sentence" is one investment strategy S_i over $\overrightarrow{p}$ chosen in a set of strategies $\{S\}$. Notice, that in this framework $Card\{S\} = 2^n$.

Note that these "rules of the game" can be used by the investor without knowing the future prices (he performs ex-ante decisions by definition) and will deliver different results: each instance of S_i can be gauged in terms of relative performance with respect to any other strategy $S_{j,j\neq i}$ (and reciprocally). Among these strategies, the best possible one in terms of maximum profit, denoted by $S*$, can be determined ex-post the realization of the price sequence (when $t = n$). Consequently, the objective function is:

$$S* \rightarrow max(W_{t+n} - W_t) \tag{5.1}$$

Thus, it can be generated by an investor acting in the "ex-ante" framework by chance or skill (the latter alternative is not discussed here). In any case, S^* is the upper bound in terms of absolute performance in $\{S\}$ and therefore a much more interesting parameter for gauging any strategy S_i. As we will show

later, the best strategy is relatively easy to identify when transaction costs are not implemented. When transaction costs alter profits, this identification is far more complex.

5.2.2 Basic illustration

Let's consider the following (arbitrarily chosen) price series (see Table 5.1 and Figure 5.1):

t	1	2	3	4	5	6	7	8	9	10	11	12
p_t	100	120	90	160	126	150	140	160	110	170	168	180

Table 5.1: Basic Artificial Time Series

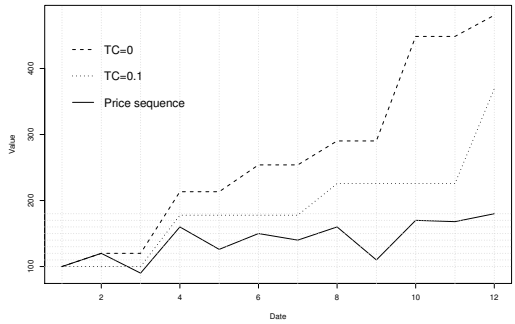


Figure 5.1: Basic artificial time series & some strategies

This example illustrates simply that, when transaction costs are minor (or absent), the best strategy consists in accumulating all *positive spreads* (*i.e. positive slopes*) observed in Figure 5.1. This strategy is denoted $S1^*$ in Table 5.2 (see also Figure 5.1). When transaction costs are implemented, the same strategy becomes far less interesting (see $S3$, Table 5.2). Some trades are simply not profitable in the context of high transaction costs. The optimal strategy when such costs are supported is $S2^*$ (see the same Figure and Table). It does not consist of realizing all profitable trades as soon as they are observed

in the price sequence (for example "Buy" in position 9 and "Sell" in position 10). It is clearly different from the situation in which there are no transaction costs, and does not match trivial formulations such as the following, which would lead to $S5$ in Table 5.2:

"capture the biggest spread in the price sequence (thus, here "Buy at time 3" and "Sell at time 12"), *then eliminate all impossibilities in further trades implied by the rules of the game* (thus, it remains one potential trade between time 1 and 2 ...), *and repeat this loop until all net positive trades are realized* (trade between times 1 and 2 would not be realized here because it is not profitable with 10% TC)*".*

	t	1	2	3	4	5	6	7	8	9	10	11	12	$W_{12} - W_1$
	p_t	100	120	90	160	126	150	140	160	110	170	168	180	
TC=0.0	$S1^*$	B	S	B	S	B	S	B	S	B	S	B	S	480.61
	$S4$	U	U	B	S	B	U	U	S	B	U	U	S	369.41
TC=0.1	$S2^*$	U	U	B	S	B	U	U	S	B	U	U	S	202.33
	$S3$	B	S	B	S	B	S	B	S	B	S	B	S	144.17
	$S5$	U	U	B	U	U	U	U	U	U	U	U	S	163.64

Table 5.2: Some Strategies among all 2^{12} potential *sentences*

Note that $S2$, which is similar to $S3^*$ in a transaction-cost free framework, is not as interesting as $S1^*$. An easy way to solve this problem when transaction costs are implemented is to generate all possible sentences and to use these to compute the net earning and identify $S*$. This set is of finite size 2^n and thus exponential. As we will now show, there are at least two ways to improve efficiently the computation of the optimal strategy $S*$, whatever the level of transaction costs. One is based on a simplex method, another relies on locating an optimal path in an oriented bipartite network.

5.3 Mathematical models: linear programming method and search in graphs

In this section, we show that the identification of $S*$ can be described as a linear programming problem with a classical simplex solution. Unfortunately, this approach is relatively inefficient since the simplex algorithm is non-polynomial in the worst case (*i.e.*; one may lack the necessary computing resources to obtain a result immediately, as soon as the size of $\overrightarrow{p}$ becomes important. We also present a solution that uses the Bellman (1957) dynamic programming approach.

5.3.1 Initial simplification

Before formal results are presented, we introduce the two theorems necessary for solving the problem. These preliminary elements aim to simplify the solution we propose.

First simplification: ***filtering the price sequence.***

Let's consider the price vector $\overrightarrow{p}$ consisting of three consecutive prices p_t, p_{t+1}, p_{t+2} and the function

$$R(x, y) = y(1 - c) - x(1 + c) \tag{5.2}$$

In equation 5.2, the $R(x, y)$, function computes the net earnings of successive buy and sell actions with $c\%$ transaction costs. In this equation, x denotes the price at which one buys and y the price at which one sells. By definition, y appears later in the time sequence than x. We show that $S*$ in $\overrightarrow{p}$, as defined on page 202, can be identified in a subset of $\overrightarrow{p}$ denoted $\overrightarrow{fp}$, consisting of the extreme points in the price sequence (peaks and troughs) and

ignoring any intermediary points (here, p_{t+1}). We assume $p_{t+2} \geq p_{t+1} \geq p_t$. Therefore $R(p_t, p_{t+2}) > R(p_t, p_{t+1})$ and $R(p_t, p_{t+2}) > R(p_{t+1}, p_{t+2})$. In this latter case, p_{t+2} is a peak while p_t is a trough.

Theorem 1 *Ignoring intermediary points: Identifying $S*$, p_{t+1} can be ignored.*

Proof 1 *Reductio ad absurdum* / proof by contradiction*:*
If it were not the case, since buying and selling on the same date is not allowed:
$R(p_{t+1}, p_{t+2}) > R(p_t, p_{t+2})$
Therefore: $p_{t+2}(1-c) - p_{t+1}(1+c) > p_{t+2}(1-c) - p_t(1+c)$
Which can be simplified: $-p_{t+1} > -p_t$
Thus, $p_{t+1} < p_t$ *since, by definition* $p_{t+1} > p_t$
Q.E.A

Note that an analogous demonstration can be made in the case where $p_{t+2} \leq p_{t+1} \leq p_t$. As a consequence, if p_{t+1} is an intermediary point, as revealed previously; it is unnecessary to identify $S*$. In other words, if one considers a complete price sequence $\overrightarrow{p}$, only peaks and troughs should be used to identify $S*$ (that is, $\overrightarrow{fp}$).

Lemma 1 *No inclusion of losses: To identify $S*$, one can ignore all situations in which $R(x, y) < 0$.*

In other words, no trade with negative net earnings can be included in the best strategy, which also excludes situations in which the so-called "buy and hold" strategy is unprofitable.

Determining two subsets of prices for *potential* "buy" and "sell" actions.

From Theorem 1 we know that it is necessary and sufficient for determining $S*$ to focus on extreme points in the price sequence. We now show that $\overrightarrow{fp}$ can itself be divided into two separate sub-vectors of peaks and troughs corresponding to two independent potential buy and sell positions in $\overrightarrow{p}$ (*resp.* denoted $\overrightarrow{fp_B}$ and $\overrightarrow{fp_S}$).

Let's consider four consecutive prices $p_t, p_{t+1}, p_{t+2}, p_{t+3}$ such as $p_{t+1} > p_t$, $p_{t+3} > p_{t+2}$ and $p_{t+2} < p_{t+1}$. In the latter case, we do not consider a situation in which $p_{t+2} > p_{t+1}$, as it is equivalent to the initial simplification case discussed previously.

Theorem 2 *To identify $S*$, none of the $\overrightarrow{fp_B}$ can be associated with a decision* `S` *and none of the $\overrightarrow{fp_S}$ can be associated with a decision* `B`*.*

Proof 2 *(i) Since $p_{t+1} > p_t \Rightarrow R(p_t, p_{t+3}) > R(p_{t+1}, p_{t+3})$. Then $p_t \leftarrow$ `B` $\succ p_{t+1} \leftarrow$ `B` with "$\leftarrow$" denoting "can be associated with a decision ..." and "$\succ$" the preference operator.*

(ii) Similarly, since $p_{t+2} < p_{t+1} \Rightarrow R(p_{t+2}, p_{t+3}) > R(p_{t+1}, p_{t+3})$. Then $p_{t+2} \leftarrow$ `B` $\succ p_{t+1} \leftarrow$ `B`

From Lemma 1 we know that the situation in which $p_{t+3} < p_t$ can be omitted. Therefore, from (i), (ii) and Lemma 1:

–whether $p_t \leftarrow$ `B` and $p_{t+1} \leftarrow$ `U` from (ii); thus $p_{t+2} \leftarrow \{$`U`$\}$ and $p_{t+3} \leftarrow \{$`U` or `S`$\}$

– or $p_t \leftarrow$ `U` and $p_{t+1} \leftarrow$ `U`; thus $p_{t+2} \leftarrow \{$`U` or `B`$\}$ and $p_{t+3} \leftarrow \{$`U` or `S`$\}$

$(p_t, p_{t+2}) \leftarrow \{$`U` *or* `B`$\}$; $\overrightarrow{fp_B} = \{p_t, p_{t+2}\}$

$(p_{t+1}, p_{t+3}) \leftarrow \{$`U` *or* `S`$\}$; $\overrightarrow{fp_S} = \{p_{t+1}, p_{t+3}\}$

Q.E.D

This theorem does not state where to buy or to sell in the subsets $\overrightarrow{fp_B}$ and $\overrightarrow{fp_S}$ to identify $S*$. It uniquely states that it is not worth buying in any element of $\overrightarrow{fp_B}$ or selling in any element of $\overrightarrow{fp_S}$.

5.3.2 A linear programming method for the identification of $S*$

A first way to solve the $S*$ determination problem is to use a linear programming method. The basic idea here is to maximize an objective function subject to a set of constraints formalizing the rules in which this problem is embedded. We now expose how this program should be written.

Let $a(i,j)$ denote the potential benefit one can obtain if $p_i \in \overrightarrow{fp_B}$ and $p_j \in \overrightarrow{fp_S}$. Notice $a(i,j)$ is computed using equation 5.2. To be more explicit:

$$a(i,j) = p_j(1-c) - p_i(1+c), \text{ with } p_i \in \overrightarrow{fp_B} \text{ and } p_j \in \overrightarrow{fp_S} \tag{5.3}$$

Let $x(i,j)$ be a dummy variable coding 0 or 1 that will be used to ignore (resp. to identify) transitions between any two prices p_i and p_j. If $p_i \leftarrow \mathtt{U}$ or $p_j \leftarrow \mathtt{U}$ then $x(i,j) = 0$, else $x(i,j) = 1$. The $S*$ strategy consists in increasing an initial wealth W_t to obtain the maximum terminal wealth W_{t+n} in selecting an optimal set of trading actions at p_i and p_j. Using the notations defined above, the identification of $S*$ can be done solving the following linear problem:

$$max \sum_{(i,j)\in\overrightarrow{fp_B}\cup\overrightarrow{fp_S}} a(i,j)x(i,j) \tag{5.4}$$

$$(s.t.)$$

$$\sum_{(i,j)\in S*} x(i,j) \leq n \tag{5.5}$$

$$\sum_{(i,j)\in \overrightarrow{fp_B}\cup\overrightarrow{fp_S}} x(j,i) + x(i,j) \leq 1, \forall i \in \overrightarrow{fp_B} \tag{5.6}$$

$$\sum_{i,k\in fp_B} x(i,k) + \sum_{j,k\in fp_S} x(j,k) = 0 \tag{5.7}$$

$$\sum_{i>j} x(i,j) + \sum_{j>i} x(j,i) = 0, \forall i \in \overrightarrow{fp_B}, j \in \overrightarrow{fp_S} \tag{5.8}$$

$$0 \leq x(i,j) \leq 1, \forall i \in \overrightarrow{fp_B}, j \in \overrightarrow{fp_S} \tag{5.9}$$

Literally, the objective function (5.4) states one seeks to maximize the total benefits in trading (that is, to identify $S*$). Note that the program in Equation 5.4 to Equation 5.9 is equivalent to Equation 5.1.

Constraint (5.5) implies that $S*$ cannot be composed of more than n prices (if the graph has n nodes) while constraint (5.6) requires that the number of matching edges incident to vertex i not exceed one.

Constraint (5.7) guarantees the absence of any connection inside buy and sell subsets.

Constraint (5.8) does not allow backwards in the price series with respect to their sequential ordering.

Constraint (5.9) requires that $x(i,j) = 1$ if a trade occurs between position i and j in $\overrightarrow{f_p}$, otherwise, $x(i,j) = 0$ (This constraint requires that each edge (i,j) not be used in the matching more that once). This latter constraint means that the problem can be solved by simplex method.

However, it is virtually impossible to explicitly enumerate all these constraints when $\overrightarrow{f_p}$ is of moderate size. It is also recognized that the simplex algorithm is exponential even if it can be solved for certain cases in polynomial time.

Provided the problem does not involve integers,[1] an underlying matrix of dimension (m, n) (where n is number of variables and m, the number of constraints) will lead to an exponential computational time $O(n^m)$ which means that any computation of such an algorithm for large price sequences will have a significant computing elapsed time.

5.3.3 An alternative dynamic programming method

The $S*-determination$ problem can also be described in the terms of the Bellman (1957) equation.

We first need to define an objective utility function, which, in the present case, is simply the investor's final wealth maximization: $W_n = Q_n \times p_n + C_n \rightarrow max$. Let's consider that at moment $k-1$ we are in the state x_{k-1}, then following control:

$$u_k = \begin{cases} buy \begin{cases} complete \\ partial \end{cases} & k = 1, 3, \cdots n-1, \\ sell \begin{cases} complete \\ partial \end{cases} & k = 2, 4, \cdots n \end{cases}$$

leads the system to state $x_k = f_k(x_{k-1}, u_k)$ and future controls $u_{k+1}, u_{k+2}, ..., u_n$ should be defined with respect to the optimality in the state x_k. The difficulty of this problem is that we do not try to maximize the value $\sum_{i=1}^{n} f_i(x_{i-1}, u_i)$, but are rather interested in the maximization of the final wealth. Therefore we can rewrite the problem as a recursive definition of the value function:

$$W_n = Q_n \times p_n + C_n \rightarrow \max \tag{5.10}$$

[1] In this last case, the problem is clearly NP-Hard, see for example Garey and Johnson (1979)

$$W_k = Q_k \times p_k + C_k \tag{5.11}$$

$$C_k = \begin{cases} Q_{k-1} \cdot p_k & \text{complete} \quad \text{if } k = 2, 4, \cdots n, \\ \Re \cdot p_k & \text{partial} \quad \text{if } k = 2, 4, \cdots n, \\ 0 & \text{complete} \quad \text{if } k = 3, 5, \cdots n-1, \\ C_{k-1} - \Re \cdot p_k & \text{partial} \quad \text{if } k = 3, 5, \cdots n-1. \end{cases} \tag{5.12}$$

$$Q_k = \begin{cases} 0 & \text{complete} \quad \text{if } k = 2, 4, \cdots n, \\ Q_{k-1} - \Re & \text{partial} \quad \text{if } k = 2, 4, \cdots n, \\ \frac{C_{k-1}}{p_k} & \text{complete} \quad \text{if } k = 3, 5, \cdots n-1, \\ \Re & \text{partial} \quad \text{if } k = 3, 5, \cdots n-1. \end{cases} \tag{5.13}$$

$$C_1 = 0 \tag{5.14}$$

$$Q_1 = X \tag{5.15}$$

$$\Re < X \tag{5.16}$$

$$k = 2, \cdots n \tag{5.17}$$

This problem can be solved with backward induction method with running time $O(n^3)$ which is an improvement with respect to the simplex method but remains more complex than the S^* algorithm described further.

We now propose to develop an alternative approach for this problem allowing an efficient solution. We tackled the $S*$ determination problem as the identification of an optimal path in an oriented bipartite network.

5.3.4 Embedding the identification of $S*$ in a graph structure

Let each price in $\overrightarrow{fp}$ be depicted as a vertex in a network. The cardinality of this subset is equal to k. Each vertex is indexed with an integer with respect to its place in the price series. We show now how to construct a bipartite, oriented and weighted network $\mathcal{N}\big(E, \overrightarrow{fp_B}, \overrightarrow{fp_S}\big)$ connecting points in $\overrightarrow{fp_B}$ and $\overrightarrow{fp_S}$.

Definition: Let $\aleph_X$ represent the subset of vertices succeeding vertex X. The network $\mathcal{N}$ is defined by the successors of each vertex.

Graph construction: The initial situation from which we start is: $\forall X \in \overrightarrow{fp}, \aleph_X = \emptyset$. From this situation, two different kind of edges can be built:

- **Trading edge** ($TE_{i,j}$)**:** for any two vertices $i \in \overrightarrow{fp_B}$ and $j \in \overrightarrow{fp_S}$, vertex $j \in \aleph_i$ if and only if:

  1. $j > i$ (to ensure temporal consistency)
  2. c being the rate of transaction costs,

$$R_{i,j} = p_j(1-c) - p_i(1+c) \geq 0 \tag{5.18}$$

- **Forward edge** ($FE_{m,n}$)**:** for any two vertices $m \in \overrightarrow{fp_S}$ and $n \in \overrightarrow{fp_B}$, $n \in \aleph_m$ if and only if:

  1. $n > m$ (which ensure temporal consistency)
  2. $\aleph_n \neq \emptyset$

Notice we impose a time consistency rule, similar to equations 5.7 and 5.8, to avoid backward connections in this bipartite oriented graph. This means

that a starting vertex p_{t+k} cannot be connected to an ending vertex p_{t+l} with $k \geq l$.

The rule presented in equation 5.18 obviously determines a profit as in equation 5.2. For any two vertices, these profits can be analyzed as weights for the corresponding edges of $\mathcal{N}$.

Consequently, we receive a balanced quasi-bipartite, weighted and directed network. We propose to interpret weights computed with 5.18 as distances between two vertices.

In the construction of $\mathcal{N}$, one can see that the number of edges depends upon the level of transaction costs c:

- The greater c makes the network $\mathcal{N}$ sparser and the solution of the problem easier.

- When $c \rightsquigarrow 0$, the number of edges increases and makes the network dense. For a specific threshold, θ, $\mathcal{N}$ is a complete antisymmetric network (with respect to the time consistency rule). θ can be computed linearly; for any two consecutive prices in fp, $p_i \in \overrightarrow{fp_B}$ and $p_j \in \overrightarrow{fp_S}$:

$$\theta = min(p_j - p_i)/(p_j + p_i), \forall(i, j) \tag{5.19}$$

  In the example provided in section 5.2.2 (see table 5.1), this threshold is 3%.

Proposition 1 *If $c < \theta$, then the $S*-$determination problem is the maximum number of edges appearing in the path.*

When $c < \theta$, $\mathcal{N}$ is completely antisymmetric. In this situation, we can derive Theorem 3.

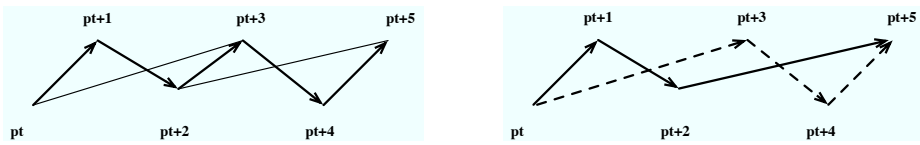
Theorem 3 *If $c < \theta$ and any 4 consecutive prices p_t, p_{t+1}, p_{t+2}, p_{t+3} in a filtered price series such as $\overrightarrow{fp}$ (see section 5.3.1) with $R(t, t+1) > 0, R(t, t+$*

$3) > 0, R(t+2, t+3) > 0$ *then:*

$$R(t, t+1) + R(t+2, t+3) > R(t, t+3)$$

Proof 3 *We make the difference between* $R(t, t+1) + R(t+2, t+3)$ *and* $R(t, t+3)$ *to show that this difference is positive.*
$-p_t(1+c) + p_{t+1}(1-c) - p_{t+2}(1+c) + p_{t+3}(1-c) + p_t(1+c) - p_{t+3}(1-c) =$
$p_{t+1}(1-c) - p_{t+2}(1+c) =$
$(p_{t+1} - p_{t+2}) - c(p_{t-1} + p_{t+2})$
From that point it is clear that if: $c = \frac{p_{t+2}-p_{t+1}}{p_{t+2}+p_{t+1}} \Rightarrow p_{t+1}(1-c) - p_{t+2}(1+c) = 0$ *and if* $c < \frac{p_{t+2}-p_{t+1}}{p_{t+2}+p_{t+1}}$ *or* $c < \theta$, $\Rightarrow p_{t+1}(1-c) - p_{t+2}(1+c) > 0$, *thus* $R(t, t+1) + R(t+2, t+3) > R(t, t+3)$
Q.E.D

Thus, if $c < \theta$, computing the longest path taking into account the profits made at each Trading Edge is similar to computing the longest path in terms of number of edges appearing in the path: $\forall c < \theta, S* = \sum_{i=1}^{k-1} TE_{i,j=(i+1)}$. In other terms, when $c < \theta$, it is proved that $S*$ is the path connecting all the edges as they appear in sequential order (see figure 5.2(a)). $S*$ connect all the vertices.



(a) Complete Bipartite Network (b) Incomplete Bipartite Network

Figure 5.2: Different paths related to different levels of c regarding Θ

When $c > \theta$, this result cannot be established and the longest path taking into account the profits made at each Trading Edge is not similar to computing the longest path in terms of the number of edges on it. For example, in Figure 5.2(b), we posit c such as $R(t+2, t+3) < 0$; one cannot follow a path in the

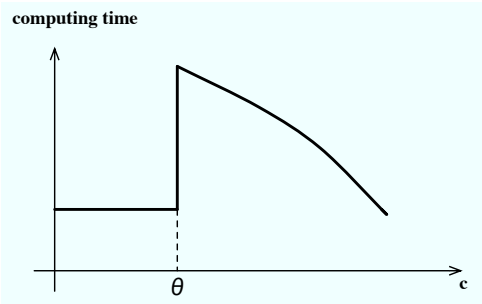


Figure 5.3: Evolution of complexity and computing time

price series connecting all vertices; several potential and interesting paths can be discovered (see Figure 5.2(b)) and therefore must be compared to determine $S*$. One way to tackle this problem might be to compute all possible paths, thus delivering an exponential algorithm.

Notice (i) that the maximum complexity of the task appears when $c = \theta + \varepsilon$ and decreases gradually beyond this threshold (see Figure 5.3); (ii) a numerical illustration of the graph construction is provided in Annex 1.

We now show how to solve this computational problem using algorithms to determine $S*$ in this graph formalism.

5.4 The $S*-determination$ algorithms

In order to make this paper self-contained, we present two different algorithmic solutions for the $S*-determination$ problem. The first derives from a technique demonstrated by Floyd (1969); the other is an algorithm for searching of the longest paths in a *directed acyclic graph* (DAG). We have chosen to emphasize the first algorithm, as it is very efficient, simply programmed, and widely used (Papadimitriou and Steigleitz, 1998). Although the Floyd algorithm is outperformed by the DAG longest path algorithm in running time, its pedagogical benefits outweigh its drawbacks. The other advantage of the DAG longest path algorithm is that it takes into account the number of both

vertices and edges as parameters for characterizing algorithmic complexity. This assumption is very important as the number of edges may vary greatly with changes of transaction costs, making the graph dense ($c \rightsquigarrow 0$) or sparse ($c \rightsquigarrow 1$).

5.4.1 Floyd algorithm approach

In the Floyd algorithm for the shortest-path problem, most steps consist of pairwise comparisons and additions of integers. When the Floyd algorithm is performed with a maximization instead of a minimization procedure, it produces the *maximum longest path* that corresponds, in our formalism, to $S*$. However, we must consider that the absence of an edge between two vertices must be interpreted as a length $-\infty$, whereas in the shortest-path problem this absence is interpreted as a length of $+\infty$. To simplify the algorithm, we can convert a multisource problem into a single-source problem by adding zero-weight edges between the first vertex from subset $\overrightarrow{fp_B}$ and other elements in this subset. This convention is needed to find the longest path, not between every pair of vertices in the graph, but between the first vertex in $\overrightarrow{fp_B}$ and every other vertex. The other modification introduced is the prohibition of backward loops in the network (see Annex 1 for a numerical illustration). The pseudo-code of the $S*-determination$ algorithm is presented in Algorithm 5.

```
for k = 1 to n do
    for j = k to n do
        path[0][j] = max(path[0][j], path[0][k] + path[k][j])
    end
end
```

Algorithm 5: $S*-determination$ algorithm: Floyd algorithm approach

The complexity of the $S*-determination$ with modified Floyd algorithm must now be established. The longest path from vertex 1 to every other

vertex is searched. During the first iteration one must go over $n-1$ vertices. Hence, $n-1$ additions and $n-1$ minimizations have to be processed; the first iteration consists of $2(n-1)$ operations. Similarly, it is possible to show that the second iteration consists of $2(n-2)$ operations, the third $2(n-3)$, and so on. The following formula defines the total number of operations carried out by the $S*-determination$ algorithm:

$$\sum_{i=1}^{i=n} 2(n-i) = n(n-1) \tag{5.20}$$

Thus the $S*-determination$ algorithm based on Floyd has a $O(n^2)$ running time and belongs to the $PSPACE$ group (algorithms necessitating a memory of polynomial space). Note that the complexity of the classical formulation of the Floyd algorithm for the shortest-path problem comprises $O(n^3)$ arithmetic operations. As there is no reverse in the graph studied in this research, the main loop of the Floyd can be ignored, decreasing complexity to a level of $O(n^2)$.

We can build other longest-path algorithms able to take into account the sort of constraints presented by the solutions proposed by Dantzig (1966) and Shier (1973). The first solution resembles Floyd (1969), although the order in which the calculations are performed is different. The second algorithm, known as the *double-sweep algorithm*, finds the k shortest path lengths between a specified vertex and all other vertices in the graph and can be applied to our problem. The longest path in a directed acyclic network can be easily found using a suitable modification of the Dijkstra shortest-path algorithm (Dijkstra, 1959).

5.4.2 DAG longest path algorithm

In this subsection, we consider $S*-determination$ as a longest-paths problem in a directed weighted acyclic graph (DAG) $G = (V, E)$, $V = fp_B \cup fp_S$. We then apply the linear-time DAG longest path algorithm (Sedgewick and Wayne, 2011) to solve this problem. The first task is to confirm that a given DAG has no directed cycles. A depth-first search can be used to formally analyze cycle existence. If a directed graph has a cycle, then a back edge will always be encountered in any depth-first search of the graph. Since the considered graph has no back edges, cycles are excluded by the graph construction assumptions.

The key element for effective solution of the longest-paths problem in DAG is *topological ordering*, which allows us compute the longest path for each vertex without having to revisit any decisions. We pass just once over the vertices in topologically sorted order. As we process each vertex, we relax each edge that leaves the vertex. By relaxing the edges of a weighted DAG $G = (V, E)$, according to a topological ordering of its vertices, we can compute the longest paths from a single source in $O(|V| + |E|)$ time (Sedgewick and Wayne, 2011).

In line with the assumptions described, we also consider the single source longest path problem. We add a dummy source vertex s plus a dummy sink vertex t, such that for all $i \in \overrightarrow{fp_B}$, $(s, i) \in E$ and the weight $R(s, i)$ is zero; and for all $j \in \overrightarrow{fp_S}$, $(j, t) \in E$, $R(j, t) = 0$. Hence, the $S*-determination$ problem can be defined as longest path from s to t in $G(V, E)$. Vertices must be numbered in such a way that an edge (i, j) is always directed from a vertex numbered i to a higher numbered vertex j. The source s is then numbered 0 and the sink t numbered $n + 1$. Vertex j is associated with $l(j)$, the longest path from 0 to j, where $l(j) = max_{i:(i,j)\in E}[l(i) + R(i, j)]$. Vertex $j + 1$ can be labeled using the same equation, and so on until the final vertex

$n+1$ is labeled with $l(n+1)$. Initially $l(0)$ is set to zero. The label $l(n+1)$ represents the length of the longest path from 0 to $n+1$. The algorithm 6 shows the pseudocode of this algorithm. This method requires the vertices to be processed in topological order. Thus, any topological ordering algorithm can be adapted to solve the longest path problem in DAGs.

```
// topologically sort the vertices of G
l[s] = 0
forall the j ∈ V \ {s} do
|  l[j] = −∞
end

foreach i ∈ V \ {s} do
|  // in topological order
|  foreach j : (i, j) ∈ E do
|  |  if l[i] + R(i, j) > l[j] then
|  |  |  // relax each outgoing edge from i
|  |  |  l[j] = l[i] + R(i, j)
|  |  end
|  end
end
```

Algorithm 6: $S*-determination$ algorithm: DAG longest path.

The running time of this algorithm is easy to analyze. Assuming that the DAG is represented using an adjacency list, we can process each vertex in constant time, with an additional time proportional to the number of its outgoing edges. The topological ordering of the vertices in G can be carried out in $O(|V|+|E|)$ time. Thus, the entire algorithm runs in $O(|V|+|E|)$ time. DAG longest path algorithm is faster than Dijkstra algorithm by a factor proportional to the cost of priority-queue operations in Dijkstra algorithm (Sedgewick and Wayne, 2011).

5.4.3 Extension of the $S*-determination$ problem

Risk-free rate

Let's now consider a (more realistic) situation where the investor retains cash because the potential buy positions all lead to negative or zero profits. This investor is offered an opportunity to invest his cash in a risk-free asset (such as short-term US Treasury Bills) delivering interests at the rate r between t and $t+1$. When Remain Unchanged is chosen after a Sell action, for a given timeframe ranging between t and $t+k$, his wealth increases according to the following formula:

$$\triangle W = W_{t+k} - W_t = W_t(1+r)^k - W_t \tag{5.21}$$

This new cash reinvestment rule substantially modifies the graph construction.

Graph modifications First of all, edge selection should be adapted; according to Lemma 1, all situations where $R(x,y) < 0$ are ignored. But when the risk-free rate is available, the condition of Lemma 1 is no longer relevant. Each trade must provide higher profits in relation to the one-step, risk-free interest r. Consequently, only situations with

$$\frac{y(1-c) - x(1+c)}{x(1+c)} > r \tag{5.22}$$

are accepted. In this formula, x denotes the price at which one buys and y – the price at which one sells. After such modifications, the $S*-determination$ approach provides the same kind of information as in its initial formulation (without a risk-free rate): when to enter the market and when to leave it. According to this new reinvestment rule, the forward edges (defined above), leading from $\overrightarrow{fp_S}$ to $\overrightarrow{fp_B}$, will no longer be zero-weighted. Weights for such edges must be calculated according to the formula 5.21.

Short selling

The $S*-determination$ problem can be also extended to short selling by modifying the graph structure. Since in our framework price series are perfectly known, one can identify future decreases in the price of the financial asset, and beyond selling the current asset holdings, one can additionally sell B_t assets that may be borrowed from a third party (*i.e.* short selling). When the price drops, the investor should repurchase B_t assets to settle short positions in the market, but at a lower price. The lender can thus recover his initial holdings.

In the graph theory terminology, short selling and borrowing mean that negative-weight edges are allowed in the graph. Since the graph construction rules exclude any cycle, the DAG algorithm can easily find the path delivering the maximum profits while containing negative-weight edges.

To tackle this particular case, we first modify the *"elements of the game"* (see section 5.2). The investor must decide for each $t \in (1, n)$ one specific action with respect to the current composition of his portfolio. The investor sells all his shares A_{t-1} bought in $t-1$, converting his position into cash. Additionally, he sells B_t borrowed shares. His total cash and wealth ain t is defined as follows:

$$
\begin{aligned}
C_t =& (A_{t-1} + B_t) \times p_t \times (1-c) \\
A_t =& -B_t \\
W_t =& C_t + A_t \times p_t
\end{aligned}
$$

If the investor decides to buy at date $t+1$, he converts all cash into assets. The trader returns the borrowed stocks to the lender.

$$\begin{aligned} A_{t+1} =& \frac{C_t}{p_{t+1} \times (1+c)} - B_t \\ C_{t+1} =& 0 \\ W_{t+1} =& C_{t+1} + A_{t+1} \times p_{t+1} \end{aligned}$$

Additionally, we have to modify the graph construction rules. The weight for a trading edge between two vertices $i \in \overrightarrow{f_{p_B}}$ and $j \in \overrightarrow{f_{p_S}}$ is defined as follows :

$$R_{i,j} = p_j(1-(1+\beta)c) - p_i(1+c) \tag{5.23}$$

In equation 5.23, $j > i$, and $\beta = \frac{B_t}{A_t}$ is the fraction of short selling in the total volume. For simplicity sake, we assume that this parameter is constant. The forward edges, connecting $\overrightarrow{fp_S}$ to $\overrightarrow{fp_B}$, are no longer zero-weighted. The weight of the forward edge from vertex $j \in \overrightarrow{f_{p_S}}$ to $i \in \overrightarrow{f_{p_B}}$ $(i > j)$, is a positive value $R_{j,i} = |p_i(1+(1+\beta)c) - p_j(1-(1+\beta)c)|$

Short selling incurs some costs, such as a fee for borrowing and repayment of any dividends that may be obtained from the borrowed assets. Thus, the profits from transaction $R_{j,i}$ must be corrected by the amount of these costs.

5.5 Numerical Illustrations

A simple example

We consider again the basic price series $\{100, 120, 90, 160, 126, 150, 140, 160, 110, 170, 168, 180\}$ and $c = 10\%$ transaction costs (see Table 204). In order to construct the bipartite graph, we first slice $\overrightarrow{p}$ in two subsets $\overrightarrow{fp_B}$ and $\overrightarrow{fp_S}$ as explained in section 5.3.1 (see also

Table 5.3).

$$\overrightarrow{fp_B}$$

$$\left\{ \begin{array}{cccccc} 1 & 3 & 5 & 7 & 9 & 11 \\ 100 & 90 & 126 & 140 & 110 & 168 \end{array} \right\}$$

$$\overrightarrow{fp_S}$$

$$\left\{ \begin{array}{cccccc} 2 & 4 & 6 & 8 & 10 & 12 \\ 120 & 160 & 150 & 160 & 170 & 180 \end{array} \right\}$$

Table 5.3: Initial price series sliced in two subsets

We compute the incidence matrix of $\mathcal{N}$ with the rules presented above (see page 213 and Table 5.4). The absence of trading edge between two vertices, due to the violation of the constraint expressed in equation 5.18 is interpreted as a weight (or length) of size $-\infty$. In this matrix, the absence of a transition edge due to the backward interdiction rule is also denoted $-\infty$. Transition edges between $\overrightarrow{fp_S}$ and $\overrightarrow{fp_B}$ systematically receive a weight of 0. The graphical representation of $\mathcal{N}(E, \overrightarrow{fp_B}, \overrightarrow{fp_S})$ is presented in Figure 5.4.

Plain bold arrows, $S* = \{UUBSBUUSBUUS\}$ with $c = 0.1$

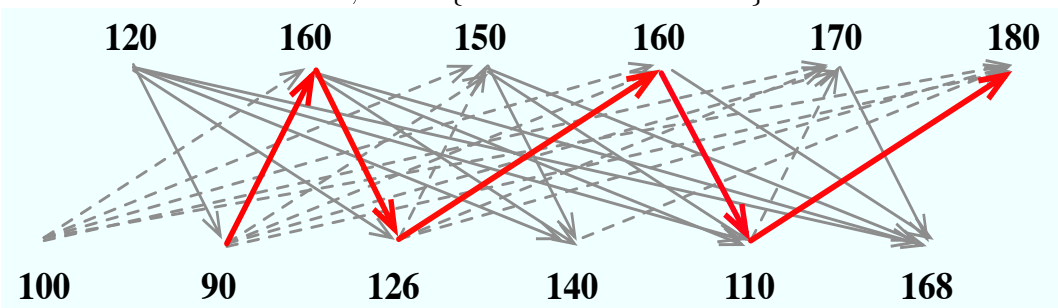


Figure 5.4: Bipartite Network from the incidence matrix B

As mentioned above, the $S*-determination$ problem can be formulated in graph theory framework. The special distinguishing feature of this graph is that its nodes can be linearized as shown in Figure 5.5. We can find the longest path from s to t, that represents the optimal solution, by comparing the paths (algorithm 6):

$l(s) = 0$

$l(100) = -\infty$

	100	120	90	160	126	150	140	160	110	170	168	180
100	$-\infty$	$-\infty$	0	34	0	25	0	34	0	43	0	52
120	$-\infty$	$-\infty$	0	$-\infty$	0	$-\infty$	0	$-\infty$	0	$-\infty$	0	$-\infty$
90	$-\infty$	$-\infty$	$-\infty$	45	0	36	0	45	0	54	0	63
160	$-\infty$	$-\infty$	$-\infty$	$-\infty$	0	$-\infty$	0	$-\infty$	0	$-\infty$	0	$-\infty$
126	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	0	5.4	0	14.4	0	23.4
150	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	0	$-\infty$	0	$-\infty$	0	$-\infty$
140	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	0	$-\infty$	0	$-\infty$	0	8
160	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	0	$-\infty$	0	$-\infty$
110	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	32	0	41
170	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	0	$-\infty$
168	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$
180	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$

Table 5.4: Modified Incidence Matrix of $\mathcal{N}$ for the $S*-determination$ with Floyd modified algorithm

$l(100) = max\,\{l(100), R(s,100)\} = max\,\{-\infty, 0\} = 0$

$l(120) = -\infty$

$l(90) = -\infty$

$l(90) = max\,\{l(90), R(s,90)\} = max\,\{-\infty, 0\} = 0$

$l(160) = -\infty$

$l(160) = max\,\{l(160), l(100) + R(100,160), l(90) + R(90,160)\} = max\,\{-\infty, 34, 45\} = 45$

$l(126) = -\infty$

$l(126) = max\,\{l(126), R(s,126), l(160) + R(160,126)\} = \{-\infty, 0, 45\} = 45$

.....

$l(t) = -\infty$

$l(t) = max\,\{l(t), l(160) + R(160,t), l(150) + R(150,t), l(160) + R(160,t), l(170) + R(170,t), l(180) + R(180,t)\} =$
$max\,\{-\infty, 45, 36, 50.4, 82.4, 91.4\} = 91.4$

Real-data Example

We now propose one application of the $S*-determination$ method with a real-world financial series consisting in the daily Dow-Jones index. This index is observed each day at the closing of the New-York Stock Exchange (NYSE) from Dec., 2nd, 1980 to Feb. 20th 2009 (*i.e.* 7156 observations). The unpredictability of future price changes, one of the cornerstones of modern finance, can be observed from the high randomness of financial returns. No one can seriously defend the idea that one particular economic agent could be able to predict with accuracy the next 7156 closing prices of the Dow-Jones Index by Dec., 2nd, 1980. Notice that even if it were possible (which is most improbable), taking advantage of this knowledge under the constraints enumerated in

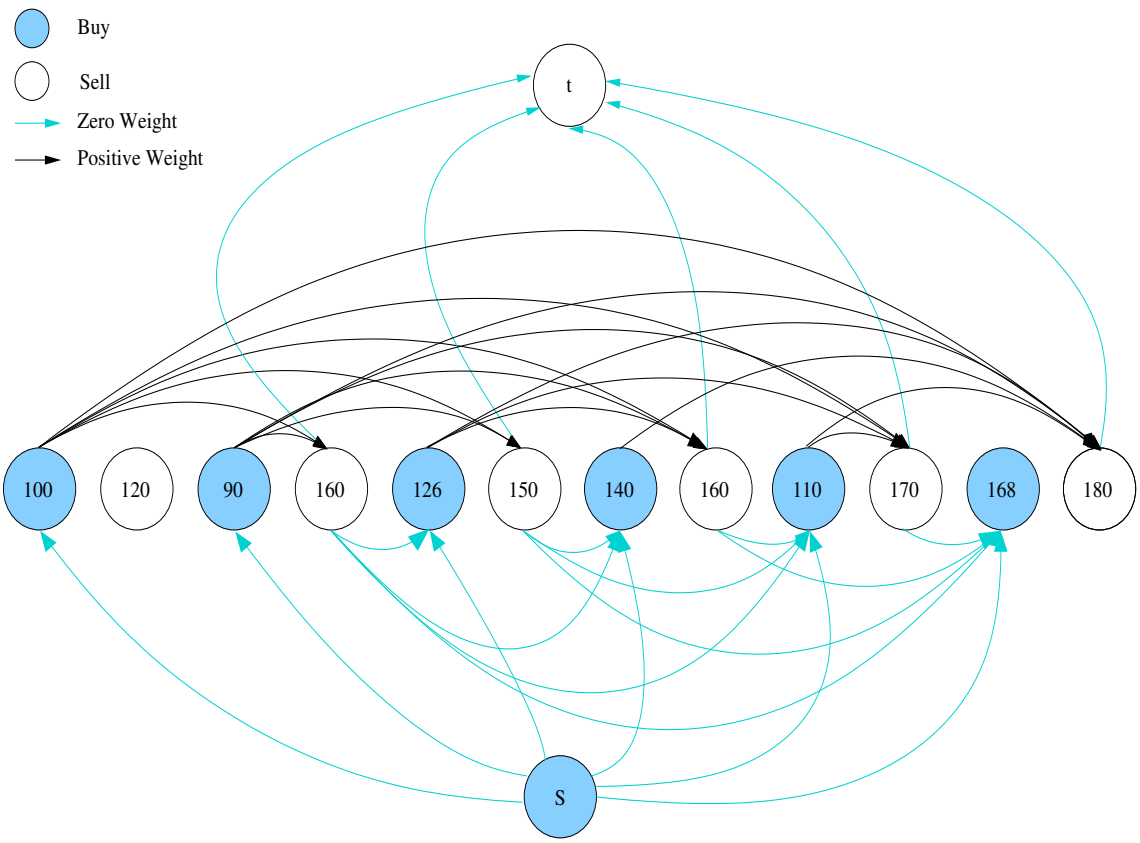


Figure 5.5: A DAG and its topological ordering. Dummy source s and dummy sink t are separated in order to distinguish them.

section 5.2 would also be extremely difficult if not simply possible without using the $S*-determination$ algorithm. Any additional element of uncertainty (unpredictable prices for example) simply increases this initial complexity.

Nevertheless, any strategy where a Dow-Jones Index tracker would be traded in this time window could be matched against the optimal set of actions that can be identified with the $S*-determination$ method.

With approaches based on Floyd algorithm (subsection 5.4.1) or DAG-longest past algorithm (subsection 5.4.2), we determine the best behavior with transaction costs c respectively at 0% and 5%. The maximum wealth one should obtain in these two cases is bigger than 1.10E+015 in the first case and bigger than 1.83E+010 in the second case. These figures seem extraordinarily

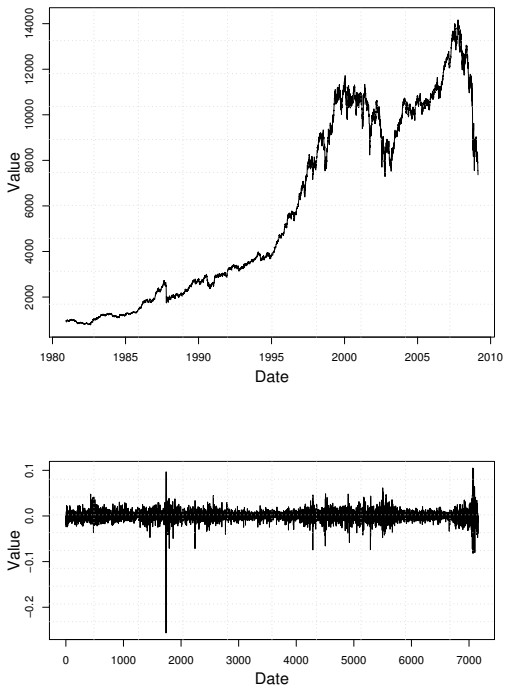


Figure 5.6: Dow Jones Index and its counterpart returns

high: one must keep in mind they are simply impossible to obtain because of the global unpredictability of the market dynamics at date t with regards to the available information at this date. In figure 5.7 we present the evolution of an investor's wealth that would have found (by chance or skills) the $S*$ set of actions in both contexts.

Nevertheless, on shorter horizons, some agents claim they have skills to predict future prices with some accuracy or at least detect specific dates where it is worth entering the market or shorting their positions. For example, technical traders claim they can detect signals in past prices (based on patterns) associated with potential market reversals. If the perceived signals indicate at date t a further increase in stock prices, these investors will try to buy stocks immediately until they receive a new signal, in date $t + T$, associated with a next decrease in prices. Then technical traders will short their positions to

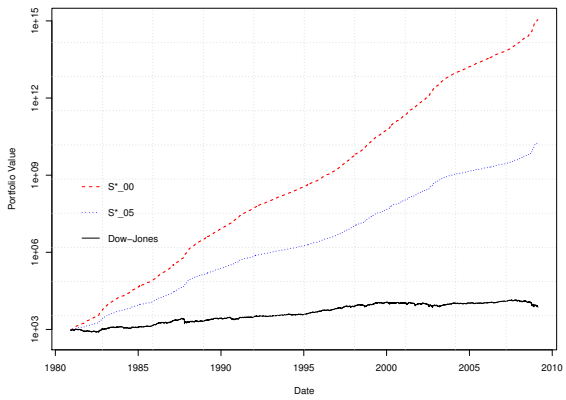


Figure 5.7: $S*$ with resp. $c = 0\%$ and $c = 0.5\%$ and Dow-Jones Index (y axis in log scale)

avoid losses.

Among others, one popular model for technical traders consists in comparing two moving averages based on past prices. The moving average with i lags MM_i is equal to $(1/i)\sum_i(p_{(t-i+1)})$. One is computed over a long range period L, the other on a short time window s. If MM_s crosses MM_L from the top to the bottom, technical traders would predict a further decrease in stock prices and try to sell immediately their holdings. On the contrary, if these moving average cross from the bottom to the top, the signal will be interpreted as "buy" signal.

In Figures 5.8 we generated such signals using the same data as previously; we also computed portfolios managed with respect to the signals. For this purpose the artificial investor is endowed with an amount of cash equal to the Dow-Jones index value at date 1 (974.40). Notice that the "moving averages" strategies provide an example of the "rules of the game" presented in section 5.2. Concerning the signals sub-figures, we only present a limited time window for graphical clarity reasons. The portfolio subfigures report the evolution of an investor's wealth using these signals in context of 0% transaction costs.

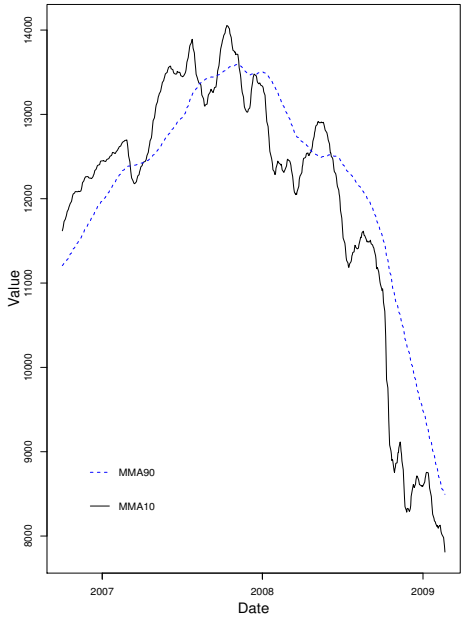
In Figure 5.8(a), MM_s is based on 10 days while MM_L is based on 90 days.

With these values we can generate 135 signals in the complete time window, which delivers the portfolio evolution. In Figure 5.8(c), these moving averages are respectively based on 5 and 20 trading days which delivers 469 signals. Notice none of these strategies is interesting in any manner

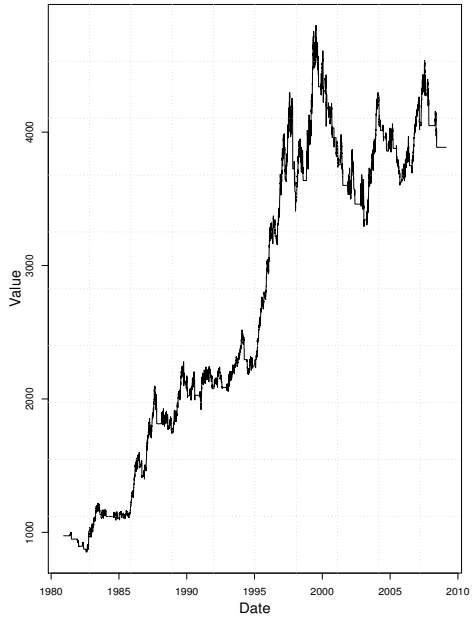
One can easily rank these strategies in term of overall profitability: MM_{10} v.s. MM_{90} seems to perform better than MM_5 v.s. MM_{20} in this price sample since the first one bears an overall profitability of $+299\%$ (terminal value of the portfolio $= 3886.36$) against $+70\%$ for the second (terminal value: 1657.18). In any case, one can also measure how far these two strategies are from the optimum $S*$. In other terms, whatever the relative performance of any trading strategy, $S*$ can be used to gauge its absolute performance. In our example, both MM_{10} v.s. MM_{90} and MM_5 v.s. MM_{20} were poor performing strategies. A simple Buy and Hold behavior "buying" the market at date 1 and selling it at date 7156 performs far better than these two moving average techniques. Nevertheless, one can suppose some automatic trading strategies could outperform this B&H strategy, especially in the context of high frequency data.

In a similar manner, an important literature dealing with the forecasting of future market trends (for example with Neural Networks, see Motiwalla and Wahab (2000), Chen, Leung and Daouk (2003), Chen and Leung (2004), or with Support Vector Machines, see Huang, Nakamori and Wang (2005)) is geared at delivering investment tools that can directly be assessed adopting the same steps.

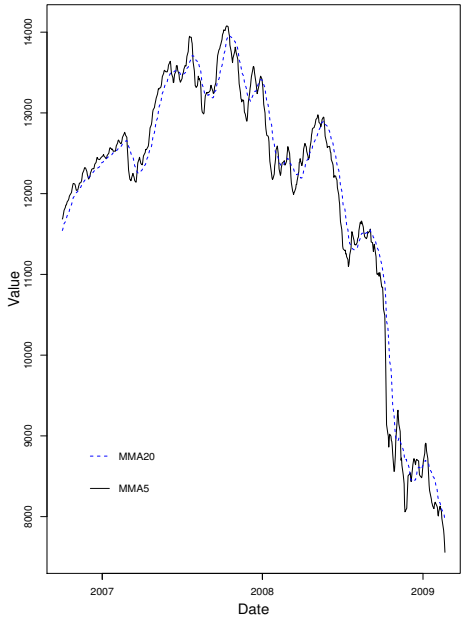
Resolving the $S*-determination$ problem does not give insights on the kind of signals one should feed automatic trading systems with, nor indicate a plausible behavior for any real-world investor. It simply establishes a boundary that was, to the best of our knowledge, largely unknown, and proposes a reference in terms of maximum-profit trajectory against which any population of investment trajectories can be gauged.



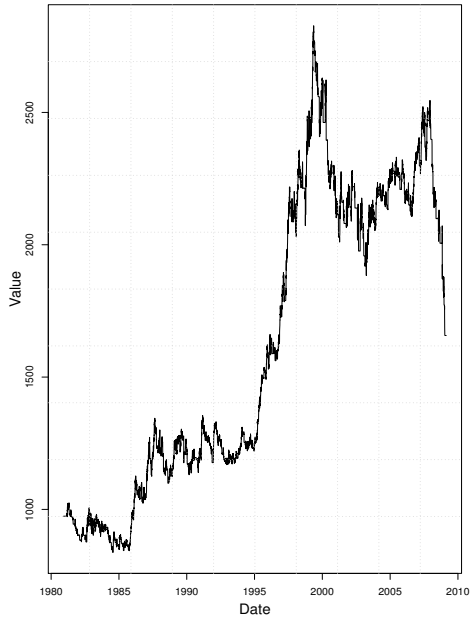
(a) Excerpt of the signals, MM_{10} v.s. MM_{90}



(b) Counterpart portfolio, MM_{10} v.s. MM_{90}



(c) Excerpt of the signals, MM_5 v.s. MM_{20}



(d) Counterpart portfolio, MM_5 v.s. MM_{20}

Figure 5.8: Two investment strategies based on moving averages techniques

5.6 Conclusion

This chapter offers a rather different approach compared to the others. While a large part of this thesis is devoted to the application of *agent-based approaches* for investigating classical financial questions, this chapter offers an *algorithmic approach* for determining the upper bound for the profits of any investment strategy. Knowing this upper bound may be used to gauge the absolute performance of an artificial agent within a population, not only to verify if it outperforms the others but also to evaluate how far it is from the optimum. We thus propose a method for determining the ex-post optimal strategy $S*$ that actually delivers the ex-post evaluation of a wide range of trading strategies. We have embedded this question in a graph theory framework and proposed the linear-time solution.

General Conclusion

This chapter brings the thesis to a conclusion. We begin by summarizing the key points of this work, what guided us in the direction and what can be learned from our models and experiments. We then review our contributions, academic achievements, and future work.

In this thesis, various financial issues have been studied using computational approaches. Unlike traditional methods in economic modeling, the approaches applied in this thesis do not rely on the assumptions on agents' rationality and homogeneity. Traditionally, economic models are mostly of a static nature and have a strong focus on deriving equilibrium. Such models are solved analytically. Relaxing of these assumptions provides a new way in which economic reality is modeled. In models of bounded rationality, dynamic elements play a much more prominent role and such models do not put particular emphasis on equilibrium. These models are analyzed via extensive computational simulations. Initially, these models were built for the purpose of studying agent's behavior, price discovery mechanisms, the influence of market microstructure on price dynamics, or the understanding of the nature of stylized facts.

To explain the general interest in using agent-based simulations for financial research questions, we have devoted Chapter 1 to discussing some alternative methodologies that can be used for studying financial topics. By confronting different approaches, we focused attention on the advantages of agent-based models in finance, and we also pointed out some weaknesses of this approach.

Our results fall into a number of different areas. First, in order to contribute to the recent advances in the field of computational finance, we have developed an agent-based artificial stock market, ATOM, implementing a re-

alistic double auction mechanism. This model has been introduced in Chapter 2. Special attention was paid to the design of the market architecture and to model validation. This research has pursued multiple aims and objectives for artificial stock market application: firstly, to verify different behaviors and the environment's calibration ability to recover qualitatively and quantitatively the main stylized facts exhibited by real markets; secondly, to implement classical model of mean-variance portfolio optimization and to investigate how individual investors allocate their portfolios under heterogeneous preferences.

Chapter 3 studied market dynamics as emergent properties of individual agent tradings. An agent-based approach provided more flexibility than standard analytical models. By flexibility we refer here to the possibility to gradually include parameters (or calibration settings) the effect of which we have studied. Further, the computational aspect of the approach enabled us to observe explicitly the effect of the agents' decision on prices. We have progressively modified agents' behavior and observed the appearance of stylized facts close to the real ones. Since the behavior of the agents is completely under our control, such model helps us relate experiment parameters to observed phenomena. Hence the contribution of Chapter 3 consists of pointing out the importance of proportion between limit, market and cancel orders, as well as order volume via Big Fishes/Small Fishes proportion on the appearance of realistic quantitative and qualitative stylized facts.

We next explored the topics of modern portfolio theory. For this purpose we have implemented heterogeneous artificial agents characterized by mean-variance optimization rules. We have established the relationship between investors' individual preferences (risk aversion, rebalancing frequency, optimization methods) and their performance in a long run. The agent's heterogeneity helps us identify a dominant strategy.

We have carried out a study of relative performance of different investment strategies, from naive diversification to some extensions of mean-variance

portfolio optimization designed to reduce estimation errors. Our findings were consistent with those of Tu and Zhou (2011) and Levy and Ritov (2011). The performance of unrestricted portfolio strategies outperforms the long-only and naive strategies with respect to the Sharpe ratio and wealth. Thus, our result has showed that naively diversified portfolios are sub-optimal.The reason behind this performance can be at least partly attributed to the "predictive power" of mean-variance agents and higher stability of their portfolios, resulting in less trading costs. Our analysis has also suggested that even though the ex-ante parameter estimation of moments and co-moments involves estimation errors due to the small size of sample, the combination of mean-variance sophisticated rules and naive rules can improve the performance of their individual counterparts.

We have also conducted the extensive simulations to find out the importance of individual agents' preferences for degree of risk aversion and rebalancing frequency on their performance. We have shown the strong correlation between risk aversion degree and survivability in a long run. Our extensive simulations demonstrated that in case of allowed short selling, the risk lovers compete others on the wealth basis, on the other hand, they quickly run out of the market in the competitions based on the Sharpe ratio. In case of long-only constrained portfolio, the highest as well as lower risk aversion do not guarantee the highest earning. Aggressive and strongly conservative traders drive quickly out of competition for wealth. However, conservative traders beat the aggressive traders in the competition for risk adjusted return of portfolios.
The extensive simulations have basically pointed out the optimal type of portfolio rebalancing in low- and high-volatility market. Our simulation results confirm the suggestions of Masters (2003) that the halfway back rebalancing is the compromise between small repeated tradings and large costs for high volume rebalancing of largely declining portfolio from its targets.
The conducted experiments have suggested that despite the critique mean-

variance optimization has received, it still is an attractive choice over a simple allocation strategy. However, the varying traders behaviors and market conditions imply that a universal model or an optimal combination of parameters that perform the best in all situations might not exist.

Finally, in Chapter 5 we proposed a new algorithm to construct a unique absolute optimal strategy, moreover we used this technique to estimate technical trading rules performance in the experimental parts. We introduced a new method for *the determination of the upper-bound in terms of maximum profit* for any investment strategy applied in a given time window. We first described this problem using a linear programming framework. Thereafter, we proposed to embed this question in a graph theory framework as an optimal path problem in an oriented, weighted, bipartite network or in a weighted directed acyclic graph.

In this thesis we have confirmed the added value of agent-based artificial market models in studying the financial topics. The usefulness of agent-based simulations stems from their ability to integrate the evolving heterogeneous population of bounded rational agents, to relate individual traders strategies with aggregate market dynamics. In such a way price fixing depends directly on trading strategies of market participants, at the same time, agents can change their behavior according to market conditions. We showed that agent-based research methodology should be seen as a complementary to other approaches, such as experiments and empirical research, for studying portfolio optimization or market microstructure topics.

Future Work

There are several directions for agent-based computational model application. First, data simulated in agent-based artificial stock market can be a good supplementary to real data, when there is no enough empirical results or statistics,

for instance, about investors individual preferences or hedge-funds managers' strategies. In order to get this information, researchers usually rely mostly on experiments, and there are many studies in which subjects participated in various investment tasks. One of the common critiques of these experimental studies is that they are often conducted only with student participants, rather than actual traders and investors. In contrast, to laboratory experiments with humans, in pure computational experiments, the simulated behavior of artificial traders is completely observable. The reasoning behind decision making, the relation between cause and effect is easily tractable. Agent-based artificial stock markets can facilitate the understanding of the relationship between individual investor strategy and aggregated market phenomena, by allowing the modeler to specify the investor behavior, to implement different market microstructure, and to analyze the resulting asset prices. In such a way, the artificial stock market can help investigate the scenarios for which empirical data do not exist, or are difficult to obtain.

The other important and often not easily obtained data are an order flow of intraday trading. When the bids and asks are available. Some traders make their profits by buying and selling within the same day. That is intraday trading. When the bids and asks are available, they should help improve forecasting accuracy and to make intraday trading more profitable. Data about orders flow provide the analysis with information that the available prices alone do not. For example, the analysis of millions of tick-by-tick data points uncovered dynamics of price, volume, volatility, order book dynamics (Mantegna and Stanley, 1999; Bouchaud and Potter, 2000; Dacorogna, Gencay, Muller, Olsen and Pictet, 2001). Agent-based artificial stock markets give insight into such topics, by providing the order flows from different scenarios.

Another field that could benefit greatly from using advanced computational agent-based modeling is the study of *Market Microstructure*. Sometimes it is desirable to compare different trading mechanisms, for example,

it is interesting to compare price formation of the continuous double auction with Walrasian equilibrium price fixing. The role played by the market makers or specialists in price formation can be also investigated using agent-based modeling.

In such a way, there are some possible extensions of our work: scenario investigation, risk analysis, macroscopic phenomena explanation.

We actually have several projects in progress. First, we extended the results in Chapter 3 to minimal market calibrations for realistic *extraday* price dynamics. The simulated time series that are results of agents trading and not returns from a given distribution, are a good supplement for real market data for *model risk analysis* Henaff and Martini (2011).

The other challenging direction for our future research is inspired from Mathieu and Brandouy (2011), investigating an optimal order flow between brokers and clients or *order cost of execution*. Brokers hold a central position where they have the possibility to influence price dynamics in ordering the flow of orders received from their clients. They can arrange the pending orders from their client in order to realize maximum benefits for themselves. Thus, the possible scenarios of orders posting and their affection of price should be tested before sending the orders to the markets. This problem can be solved only using an agent-based decision support systems.

As it has been shown, the different types of price fixing mechanisms, a large variety of learning mechanisms and agents' strategies, possible to be implemented, suggest a rich field of research using artificial stock markets. The number of research opportunities to explore is countless.

Appendix A

Appendix

A.1 NYSE Euronext Stock Exchange Overview

Below is a brief overview of the NYSE Euronext microstructure that aims to provide a description of the market features that we implement in our platform. For a more comprehensive overview, one can refer to the official website `http://www.euronext.com`, which provides a significant amount of information on the recent changes.

In 2000, the Amsterdam, Brussels and Paris exchanges (later Portuguese Stock Exchange as well) become Euronext, the first pan-European stock exchange. The main characteristic of Euronext is to be a pure electronic order-driven trading system. The "order driven" specification means that traders send their orders directly to an order book. A single order book for each security or financial product is introduced in Euronext for greater transparency and liquidity. All products are traded electronically on the NSC system adopted by all of the Euronext members. Transactions are cleared through LCH.Clearnet [1] , acting as the central counterparts and thus guaranteeing payment and delivery for all market transactions.

Traders send instructions (orders) to exchanges that arrange their trades. Orders explain how agents want their trades to be arranged. There are different *types of orders* introduced on the NYSE Euronext stock exchange. Traders choose orders that have properties that allow them to reach trading targets. *Limit orders* indicate the price that an investor wants to pay or receive for buying or selling shares. The trade will not take place until the limit price is reached. The limit order is considered as conditional because it is executed only if the limit price or a better price can be obtained. A buy limit order indicates that a stock can be purchased only at a specific price or lower. A sell limit order authorizes the stock to be sold at a specific price or higher. The

[1] The world's leading independent clearing house based in Europe that serves major international exchanges and platforms

danger of a limit order is that there is no guarantee that the order will ever be executed. The set of unexecuted limit orders held by the system constitutes the book. Limit orders can be canceled or modified at any time – hence, the book is dynamic.

Market orders are executed at the best available price on the market. A market order simply means buying or selling at the current market price. Market orders have priority on the trading floor and thus ensure maximum immediacy.

Security markets are sometimes characterized by their *liquidity*. In a liquid market, a small shift in demand or supply does not result in a large price change. On the Euronext stock exchange, limit orders are regarded as a source of liquidity in the market as they provide the necessary pools of supply and demand. Market orders consume the liquidity, because they get executed when they arrive.

Stop Orders. A stop order is considered conditional because it specifies that a trade will not be executed until the market moves to a designated price. At that time, the order becomes a market order.

Hidden or iceberg orders are limit orders specifying a disclosed quantity which refers to the number of shares the trader wishes to be displayed on the market screens. The difference between the disclosed quantity and the total order size represents the hidden part of the order. If the displayed quantity is executed, it is refreshed from the reserve quantity. The procedure mimics a human trader who might execute a large order by splitting it up into smaller quantities.

The order of execution is determined by two factors:

- A limit buy order is sorted before all others with lower limits and a limit sale order is sorted before all others with higher limits. This procedure determines the best bid and best ask prices. Traders cannot accept bids

or asks at any inferior price. Buyers can accept only the lowest priced asks and sellers can accept only the highest priced bids.

- Orders of the same kind and at the same limit are filled in the same order as they were entered into the order book.

Trading takes place in several stages:

- Pre-opening from 7:15 a.m. to 9:00 a.m.: orders are sent to the central order book without any transaction taking place.

- Opening at 9:00 a.m.: on the basis of all orders recorded in the book, the central computer automatically calculates the opening price or call auction price that allows to match the largest number of bids and asks. Orders that are not compatible with opening price remain in the book, pending for a possible future matching against new opposite orders.

- Continuous trading from 9:10 a.m. to 5:25 p.m.: the arrival of a new order immediately triggers one or several transactions if the central order book contains an order or several orders at the opposite side at a compatible price. If there are no such orders, the incoming order is recorded, remaining on the order book at the specified limit.

- Pre-closing from 5:25 p.m to 5:30 p.m.: as with the pre-opening, orders are entered without any transactions taking place.

- Closing call auction at 5:30 p.m.: all remaining orders are compared and trades executed where possible.

- Last, trading from 5:30 p.m. to 5:40 p.m.: trading members may enter orders at the closing price, which is the price set at the closing call auction except in exceptional cases.

For a normal double-auction market, the best (highest) bid price is always less than the best (lowest) ask price. The difference between the two is called the spread of the market.

A.2 Minimal intelligence calibration. Algorithms

Data: $P_{min}, P_{max}; V_{min}, V_{max}$
Result: Order
`/* initialisation */`
$\Lambda \sim D(0,1)$
`/* equal possibilities to buy or sell */`
if $\Lambda > 0.5$ **then**
| $Direction = "Ask"$
else
| $Direction = "Bid"$
end
`/* price and quantity definition */`
$P \sim D(P_{min}, P_{max})$
$Q \sim D(V_{min}, V_{max})$

return $(Direction, P, Q)$

Algorithm 7: Uniform price distribution $UZIT_U$

```
Data: P_mean, P_sd; V_min, V_max
Result: Order
/* initialisation */
Λ = D(0,1)
/* equal possibilities to buy or sell */
if Λ > 0.5 then
|  Direction = "Ask"
else
|  Direction = "Bid"
end
/* price and quantity definition */
P ~ N(P_mean, P_sd)
Q ~ D(V_min, V_max)

return (Direction, P, Q)
```

Algorithm 8: Normal price distribution $UZIT_N$

```
Data: a_min = P_min, a_max, b_min, b_max = P_max; V_min, V_max
Result: Order
/* initialisation */
Λ ~ D(0,1)
/* equal possibilities to buy or sell */
if Λ > 0.5 then
|  Direction = "Ask"
|  P ~ D(a_min, a_max)
else
|  Direction = "Bid"
|  P ~ D(b_min, b_max)
end
/* quantity definition */
Q ~ D(V_min, V_max)

return (Direction, P, Q)
```

Algorithm 9: Statistically calibrated $SZIT_U$

Data: $\{P^k{}_{min}, P^k{}_{max}\}_{k=0}^{N}$; V_{min}, V_{max}; N_{fix} is number of total fixed prices, N is number of foreseen frames, $N < N_{fix}$
Result: Order
$\triangle = \frac{[P_{max} - P_{min}]N}{N_{fix}}$
$\Phi[0] = 0$
$\forall t \in [1; \frac{N_{fix}}{N}]$
$\Lambda \sim U(0, 1)$
`/* equal possibilities to buy or sell */`
if $\Lambda > 0.5$ **then**
| $Direction = "Ask"$
else
| $Direction = "Bid"$
end
`/* price and quantity definition */`
$\delta_t \sim logN(0, 1)$
$\gamma_t = \Phi[t-1] + \triangle$
$P_t = P^k{}_{min} + (P^k{}_{max} - P^k{}_{min}) \times \gamma_t \times \delta_t$
$Q_t \sim U(V_{min}, V_{max})$
$\Phi[t] = \gamma_t$

return $(Direction, P_t, Q_t)$

Algorithm 10: Trend calibrated agent $TZIT_N$

```
Data: Agent indicator = {Patient, Impatient}; a_min, a_max, b_min, b_max;
      V_min, V_max
Result: Order
/* initialisation */
Λ ~ D(0, 1)
if Patient then
    if Λ > 0.5 then
        Direction = "Ask"
        log(P) ~ D(b_max, ∞)
    else
        Direction = "Bid"
        log(P) ~ D(0, a_min)
    end
    Q ~ D(V_min, V_max)
    return (Limit, Direction, P, Q)

else if Impatient then
    if Λ > 0.5 then
        Direction = "Ask"
    else
        Direction = "Bid"
    end
    Q ~ D(V_min, V_max)

    return (Marekt, Direction, Q)
end
```

Algorithm 11: Patient and Impatient Agressive calibrated agent $AZIT$

A.3 Mathematics of the mean-variance model

The goal of portfolio analysis is finding the efficient set of portfolios in the mean-variance framework with minimal variance and some constraints on portfolio weights and desired return. The mathematics of mean-variance problem is the trade off between desired return and level of risk reached using probability theory and optimization theory. Let consider application of different utility functions in the single-period mean-variance optimisation framework. Let define α is the part of initial wealth W_0 invested in the single risky asset with expected return R, other part $1-\alpha$ is invested in the risk-free asset with return r. The expected wealth is $W_1 = (\alpha W_0)(1+R) + [(1-\alpha)W_0](1+r)$. In such way we can define the return and variance of such portfolio: $R_p = \frac{W_1 - W_0}{W_0} = \alpha(R-r) + r$, $\sigma_p = \alpha\sigma_R$. The investor tries to maximise

$$ER_p - \frac{c}{2}\sigma_p^2 \xrightarrow{\alpha} max$$

, where c is a measure of risk aversion. Hence utility function is:

$$f(\alpha) = \alpha(R-r) + r - \tfrac{c}{2}\alpha^2\sigma_R^2$$
$$f'(\alpha) = R - r - c\alpha\sigma_R^2 = 0$$
$$\alpha^* = \tfrac{R-r}{c\sigma_R^2}$$

where α^* is the optimal part invested to the risky asset. The absolute amount invested in the risky asset is $A_0 = \alpha^* W_0$.

$$\frac{A_0}{W_0} = \frac{R-r}{c\sigma_R^2}$$

absolute amount is proportional to initial wealth $A_0 \sim W_0$, and inversely related to the risk aversion coefficient $A_0 \sim c$ and to volatility of risky asset $A_0 \sim \sigma_R^2$

If an investor maximises expected utility of end-of-period portfolio wealth,

then it can be shown that this is equivalent to maximising a function of expected portfolio returns and portfolio variance providing: a) either utility is quadratic or b) portfolio returns are normally distributed.

If initial wealth is W_0, R_p is portfolio return, R is the return of risky asset, r risk-free return, then the end-of-period wealth and utility are

$$W = W_0(1 + R_p)$$
$$W = \alpha W_0(1 + R) + [(1 - \alpha)W_0](1 + r)$$
$$U(W) = U[W_0(1 + R_p)]$$

Expanding $U(R_p)$ in a Taylor series around the mean of $R_p(= \mu_p)$ gives

$$U(R_p) = U(\mu_p) + (R_p - \mu_p)U'(\mu_p) + \frac{(R_p - \mu_p)^2}{2}U''(\mu_p) + \frac{(R_p - \mu_p)^3}{6}U'''(\mu_p)...$$

since $E(R_p - \mu_p) = 0$ and $E(R_p - \mu_p)^2 = \sigma_p^2$

$$E[U(R_p)] = U(\mu_p) + \frac{1}{2}\sigma_p^2 U''(\mu_p) + ...$$

If utility is quadratic, then higher-order terms other than U'' are zero. If return are normally distributed, then $E[(R_p - \mu_p)^n] = 0$ for n odd, and $E[(R_p - \mu_p)^n]$ for n is a function only of the variance σ_p^2. Hence for cases (a) and (b), $E[U(R_p)]$ is a function of onlly the mean μ_p and the variance σ_p^2. These results will help us to show that indifference curves in (μ_p, σ_p) are convex. The definition of optimal weights (with short selling) is a standard quadratic programming problem with an analytic solution. If we want to solve the portfolio allocation problem with constraints ($w_i > 0$), then, there is no analytic solution and a numerical optimisation routine is needed. In this section mean-variance model is defined in the terms of nonlinear optimisation problem. Let

I is number of assets

ϖ_i is proportion of portfolio invested in assets $i, 1 \leq i \leq I$

ϖ is column vector of proportions w_i

R_i expected return of asset $i, 1 \leq i \leq I$

R is column vector of expected returns α_R, $R = [R_1, ...R_I]^T$

σ_i is standard deviation of the return of asset i

ρ_{ij} is correlation coefficient of the returns of assets i and j

p_{ij} is covariance of asset i with $j, 1 \leq i \leq I$ and $1 \leq j \leq I$, $p_{ii} = \sigma_i^2$, $p_{ij} = \rho_{ij}\sigma_i\sigma_j$ for $i \neq j$

$V = I \times I$ is matrix of covariances $p_{i,j}$

R_P is expected return of portfolio

σ_P is standard deviation of portfolio

A denotes coefficient of relative risk aversion The problem is to maximize $f(w)$:

$$\begin{aligned} f(\varpi) &= R_P - \frac{1}{2}A\sigma_P^2 \\ &= \varpi^T R - \frac{1}{2}A\varpi^T V\varpi \\ = \sum_{i=1}^{I} \varpi_i R_i - \frac{1}{2}A \sum_{i=1}^{I}\sum_{j=1}^{I} w_i w_j pij \end{aligned}$$

subject to the constraint:

$$\sum_{i=1}^{I} \varpi_i = 1$$

To deal with the such constraint, we use Lagrange multiplier λ and new ob-

jective function $\hat{f}$

$$\begin{aligned}\hat{f}(w,\lambda) &= f(\varpi) + \lambda(1 - \sum_{i=1}^{I} \varpi_i) \\ &= w^T x - \frac{1}{2} A \varpi^T V \varpi + \lambda(1 - \sum_{i=1}^{I} \varpi_i) \\ &= \sum_{i=1}^{I} \varpi_i R_i \frac{1}{2} A \sum_{i=1}^{I} \sum_{j=1}^{n} \varpi_i \varpi_j p_{ij} + \lambda(1 - \sum_{i=1}^{I} \varpi_i)\end{aligned}$$

To solve such problem, we should take the $n+1$ partial derivatives of $\hat{f}$ and send them equal to 0.

$$\frac{\partial \hat{f}}{\partial \varpi_i} = R_i - A \sum_{j=1}^{I} p_{ij} \varpi_j - \lambda = 0 \tag{A.1}$$

$$\forall 1 \leq i \leq n \tag{A.2}$$

$$\frac{\partial \hat{f}}{\partial \lambda} = 1 - \sum_{i=1}^{I} \varpi_i = 0 \tag{A.3}$$

Rewrite these equations as:

$$\sum_{j=1}^{I} p_{ij} \varpi_j - \frac{\lambda}{A} = \frac{\alpha_i}{A} \tag{A.4}$$

$$\forall 1 \leq i \leq I \tag{A.5}$$

$$\sum_{i=1}^{I} \varpi_i = 1 \tag{A.6}$$

This system of equations we can solve using linear algebra. Define vectors

and matrices as follows:

$$\hat{V} = \begin{pmatrix} \rho_{11} & \dots & \rho_{1n} & 1 \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \rho_{n1} & \dots & \rho_{nn} & 1 \\ 1 & \dots & 1 & 0 \end{pmatrix}$$

$$\hat{\varpi} = \begin{pmatrix} \varpi_1 \\ \dots \\ \dots \\ \varpi_n \\ \lambda/A \end{pmatrix} \hat{x} = \begin{pmatrix} R_1 \\ \dots \\ \dots \\ R_I \\ 0 \end{pmatrix} \hat{y} = \begin{pmatrix} 0 \\ \dots \\ \dots \\ 0 \\ 1 \end{pmatrix}$$

We can rewrite equations A.3 in the terms of matrices.

$$\hat{V}\hat{\varpi} = \frac{1}{A}\hat{x} + \hat{y}$$

We assume for the moment that the matrix $\hat{V}$ is non-singular and hence has an inverse. In this case the solution of this problem is:

$$\hat{\varpi} = \frac{1}{A}\hat{V}^{-1}\hat{x} + \hat{V}^{-1}\hat{y} \tag{A.7}$$

For an infinitely risk-averse investor $A = \infty$, the solution becomes simply $\varpi_i = \hat{V}_{ij}\hat{y}_i$, for all $1 \leq j \leq n$.

Lagrange multiplier deals with only the budget constraint which says that the sum of the asset proportion must equal 1, hence all all values for the asset proportions are permitted, even those outside the range $[0, 1]$, that allows the short-selling. In case of no short sales the optimisation problem is exactly the same but with an additional constraint $\varpi_i \geq 0 (i = 1, 2 \dots I)$, in this case we should also introduce additional Kuhn-Tucker conditions. In the next

subsection we work out all the mathematical details of constraint problem with the Lagrange multiplier approach to deal with that constraint.

One riskless asset Denote by R_f the risk free return. $R_f < R_{min}$, where R_{min} is the return of portfolio with minimal variance of risky assets. That is natural since the risky portfolio has a positive risk associated with it while the riskless asset does not. The proportion of wealth invested in riskfree asset is ϖ_0.

$$\varpi_0 = 1 - \varpi$$

where w is the vector of weights of risky assets in optimal portfolio. Then, the portfolio return and variance are defined as follow:

$$\alpha_P = \varpi\alpha + (1 - \varpi)R_f \tag{A.8}$$

$$\sigma_P = \varpi\sigma \tag{A.9}$$

By definition, a risk-free asset has standard deviation of 0, a variance of 0 and covariance of 0 with all other assets, hence it does not contribute to the portfolio general risk. Correlation matrix should be enhanced by column and low of 0.

Therefore, the new optimization program is:

$$f(\varpi) = R_P - \frac{1}{2}A\sigma_P^2 = \tag{A.10}$$

$$= \varpi\alpha + (1 - \varpi)R_f - \frac{1}{2}A\varpi^2\sigma^2 \tag{A.11}$$

using the same notions as defined above $\hat{\varpi} = \frac{1}{A}\hat{c} + \hat{V}^{-1}\hat{y}$. Let put $\hat{d} = \hat{V}^{-1}\hat{y}$

$$\hat{V}\hat{d} = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & \rho_{22} & \dots & \rho_{2n} & 1 \\ \dots & \dots & \dots & \dots & \\ \dots & \dots & \dots & \dots & \\ 0 & \rho_{n2} & \dots & \rho_{nn} & 1 \\ 1 & 1 & \dots & 1 & 0 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ \dots \\ \dots \\ d_n \\ d_{n+1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \dots \\ \dots \\ 0 \\ 1 \end{pmatrix} = \hat{y}$$

The solution of optimal weights: $\varpi_1 = \frac{1}{A}V^{-1}x_1 + 1$ for risk free asset, and $\varpi_i = \frac{1}{A}V^{-1}x_i$ $(2 \leq i \leq I)$ for risky assets.

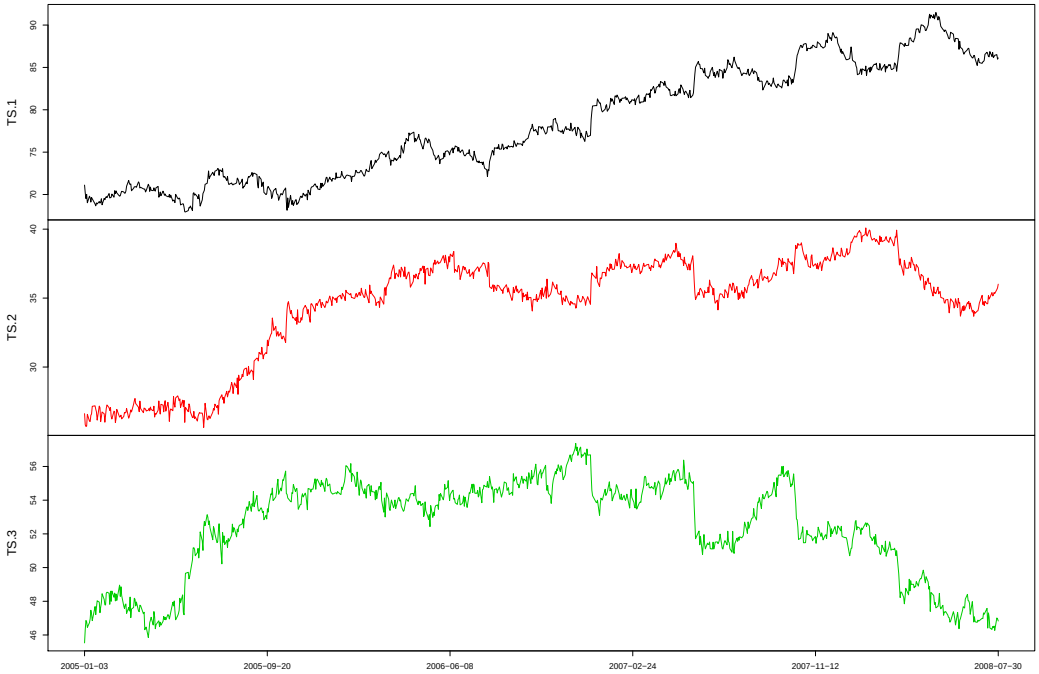
The efficient frontier is a straight line that contains the risk-free asset and the optimal risky portfolio.

A.4 Predictive power of mean-variance traders

In this section we test the ability of artificial mean-variance traders to generate stock returns in accordance with the expected moments they receive at the beginning of a simulation. In this experiment, the agents follow the trading rules described in section 4.1.1. We use different combinations of stocks listed on the CAC40 index form January 2005 to July 2008, for model calibration. To simple demonstrate the test, we relay on daily data of 3 stocks from January 3, 2005 to July 30, 2008 (900 observations per stock), as this result is typical output of test. The whole series is divided in 3 sub-sets of 300 daily observations. For each of these subsets, we calculate $\bar{R}_n$, σ_n^2 and $Corr_n$, $n = \overrightarrow{1,3}$. We use these statistics as proxies for the assets expected moments. The agents trade over 300 days using these proxies. Then this information is updated every 300 days (a new set of expected moments is used). Hence we propose a sliding-window adaptation scheme to approximate the benchmark dynamics in artificial stock market. In other words, throughout the investment period, the agents rebalance their portfolios at specific intervals to update old portfolio weights to new ones calculated from data in the new estimation window.

We run 1000 times the process described above to check if the artificial agents generate prices in accordance to their “expectations”. Figure A.2 compares the distribution of statistics generated using the 1000 runs and the benchmark price series (in that case, we only use the "Carrefour" stock – ticker, CA.PA is time series from January 3, 2005 to July 30, 2008). Figures A.2(b) and A.2(d) succeed in reproducing the second (standard deviation) and fourth (kurtosis) statistical moments. However, they produce mean values higher than imposed values of real time series (see A.2(a)). The third moments (skewness), on average, are also higher than those of real price series.

(a) *Real data*



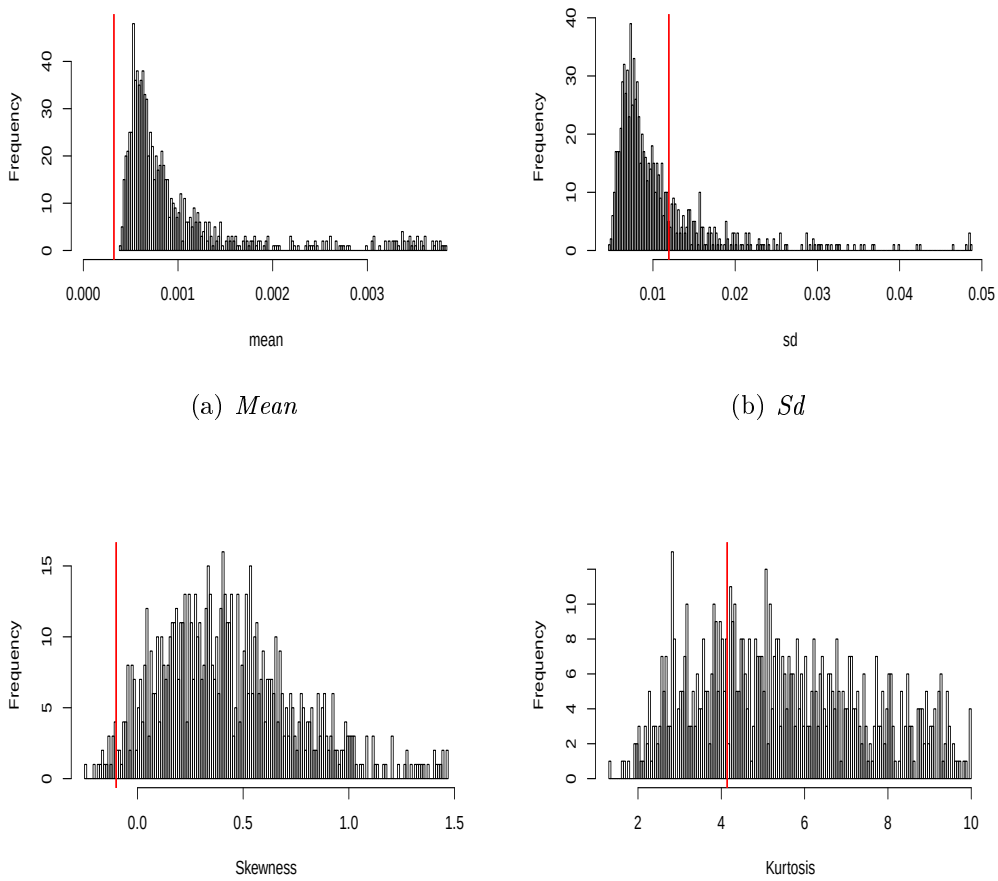
(b) *One of the possible outputs generated by ATOM with real data initialisation*

Figure A.1: Real CAC40 stocks Vs. ATOM simulated series over 900 days. The exogenous information is imposed every 250 days. A sliding-window adaptation scheme to approximate the original dynamics is used. Prices are generated by 1000 mean-variance optimizers, heterogeneous with respect to their risk aversion $A \sim U(0.1, 10)$, and trading frequency $\theta \sim U(1, 10000)$

Next, we measure similarity between distributions of generated and benchmark price series. A measure of similarity between $P(X)$ and $Q(X)$ is provided by the Kullback-Leibler (KL) divergence (Kullback and Leibler, 1951), which is defined as:

$$\begin{aligned} D(Q|P) &= E_Q\left(ln\left(\frac{dQ}{dP}\right)\right) \\ &= E_P\left(\frac{dQ}{dP} ln\left(\frac{dQ}{dP}\right)\right) \end{aligned}$$

where X is a vector of random variables generated by the stochastic processes, $P(X)$ and $Q(X)$ are probability distributions. This divergence is expressed non-parametrically, making no assumptions, and directly from the samples, without explicit estimation of the underlying probability density functions. Using this approach, we can identify the list of series that are within certain "distance" of the benchmark series. According to KL-divergence measure ($k = 40$, $\theta = 0.01$), we select 360 ATOM generated price series out of 1000 with CA.PA daily data as a benchmark. We repeat the same procedure with BNP.PA and AC.PA as benchmark, and select respectively 308 and 432 ATOM generated price series.



(a) *Mean* (b) *Sd*

(c) *Skewness* (d) *Kurtosis*

Figure A.2: Histogram exhibits the distribution of 1000 statistics generated by artificial mena-variance traders, the red line corresponds to the moment of CA.PA from January 3, 2005 to July 30, 2008.

Bibliography

Akian, M., Menaldi, J. and Sulem, A. (2004). On an investment-consumption model with transaction consts. *SIAM Journal of Control and Optimization*, **34**, 329–364. 185

Albin, P. and Foley, D. (1990). Decentralized, dispersed exchange without an auctioneer: A simulation study. *Journal of Economic Behavior and Organization*, **18(1)**, 3–30. 19

Alchian, A. (1950). Uncertainty, evolution, and economic theory. *Journal of Political Economy*, **58**, 211–221. 10

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine. *Econometrica*, **21**, 503–546. 33

— (1979). *Expected Utility Hypothesis and the Allais Paradox*, Dordrecht: D. Reidel, chap. The so-called Allais paradox and rational decisions under uncertainty. 33

Amaral, L., Cizeau, P., Gopikrishnan, P., Liu, Y., Meyer, M., Pend, C.-K. and Stanley, J. (1999). Econophysics: can statistical physics contribute to the science of economics? *Computer Communications*, **121-122**, 145–152. 104

Andersen, T. and Bollerslev, T. (1997a). Heterogeneous information arrivals and returns volatility dynamics. *Journal of Finance*, **52**, 975–1005. 24

— and — (1997b). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, **4**, 115–158. 23

Anufriev, M., Bottazzi, G. and Pancotto, F. (2006). Equilibria, stability and asymptotic dominance in a speculative market with heterogeneous agents. *Journal of Economic Dynamics and Control*, **30**, 1787–1835. 136

Arifovic, J. (1996). The behavior of the exchange rate in the generic algorithm and experimental economics. *Journal of Political Economy*, **104**, 510–541. 60

— (2001). Evolutionary dynamics of currency substitution. *Journal of Economic Dynamics & Control*, **25**, 395–417. 61

Arnéodo, A., Muzy, J.-F. and Sornette, D. (1998). "direct" causal cascade in the stock market. *The European Physical Journal B*, **2**. 19

Arthur, W. (1994a). Bounded rationality and inductive behavior (the el farol problem). *American Economic Review Papers and Proceedings*, **84**, 406–411. 11

— (1994b). *Increasing Returns and Path Dependence in the Economy*. University of Michigan Press. 11

— (1994c). Inductive reasoning and bounded rationality. *American Economics Review*, **84(2)**, 406–411. 67, 91

— (2006). Out-of-equilibrium economics and agent-based modeling. In L. Tesfatsion and K. Judd (eds.), *Handbook of Computational Economics*, vol. 2, Elsevier B.V., pp. 1552–1564. 4, 13

—, Durlauf, S. and Lane, S. (1997a). *The Economy as an Evolving Complex System II*. Addison-Wesley, Reading, MA. 24, 59, 61

—, Holland, J., LeBaron, B., Palmer, R. and Tyaler, P. (1997b). Asset pricing under endougenous expectations in an artificial stock market. *Economic Notes*, **26**, 297–330. v, 19, 49, 52, 62, 67

Ashburn, T., Bonabeau, E. and Ecemis, I. (2004). Interactive inversion of agent-based models. In Japan (ed.), *Agent-Based Simulation: from Modeling Methodologies to Real-World Applications. Third International Workshop on Agent-Based Approaches in Economic and Social Complex Systems*, vol. 1, Springer, pp. 896–922. 18

Axelrod, R. (2006). Agent-based modeling as a bridge between disciplines. In L. Tesfatsion and K. Judd (eds.), *Handbook of Computational Economics*, vol. 2, Elsevier B.V., pp. 1565–1584. 18

— and Tesfatsion, L. (2005). A guide for newcomers to agent-based modeling in the soccial sciences. *Appendix in Handbook of Computational Economics. Agent-Based Computational Economics*, **2**. 6, 19

Axtell, R. (1999). Why agents? on the varied motivations for agent computing in the social sceinces. *Agent Simulation: Application, Models, and Tools*, pp. 1–23. 13, 16

Bak, P., Chen, K., Scheinkman, J. and Woodford, M. (1993). Aggregate fluctuations from independent sectoral shocks: Self-organized criticality in a model of production and inventory dynamics. *Ricerche Economiche*, **47**, 3–30. 19

Barberis, N. and Thaler, R. (2003). *Handbook of the Economics of Finance*, Elsevier, chap. A survey of behavioral finance, pp. 1051–1121. ii

Bass, T. (1999). Ny:henry holt and company. *The Predictors*. 198

Bellman, R. E. (1957). *Dynamic Programming*. Princeton University Press. 206, 211

Beltratti, A., Margarita, S. and Terna, P. (1996). *Neural Networks for economic and financial modeling*. London, UK: International Thomson Computer Press. 18

Blume, L., D.Easley and O'Hara, M. (1994). Market statistics and technical analysis: The role of volume. *Journal of Finance*, **49**, 153–181. 30, 198

Boer, K., Kayamak, U. and Spiering, J. (2007). From discrete-time models to continuous-time, asynchronous models of financial markets. *Computational Intelligence*, **23**. 51, 75

Boer-Sorban, K. (2008). *Agent-Based Simulation of Financial Markets. A modular, continuous-time approach.* Ph.D. thesis, RSM Erasmus University, Erasmus School of Economics. 50, 51, 56, 73, 74, 84

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31**, 115–144. 21

—, Chou, R. and Kroner, K. (1992). Arch modeling in finance. *Journal of Econometrics*, **52**, 5–59. 21

Bolleslev, T. and Domowitz, I. (1993). *The Double Auction Market: Institutions, Theories, and Evidence. Santa Fe Institute Studies in the Scinces of Complexity.* Perseus Publishing, Cambridge MA. 55

Boschetti, F. and Moresi, L. (2001). Interactive inversion in geoschences. *Geophysics*, **66**, 1226–1234. 18

Boswijk, H., Hommes, C. and Manzan, S. (2003). Behavioral heterogeneity in stock prices. In *IFAC symposium Computational Economics, Finance and Engineering*, Seattle, USE. 66

Bottazzi, G., Dosi, G. and Rebesco, I. (2005). Institutional architectures and behavioral ecologies in the dynamics of financial markets. *Journal of Mathematical Economics*, **41**, 197–228. 5, 136

Bouchaud, J.-P. and Potter, M. (2000). *Theory of Financial Risk.* Cambridge University Press. 64, 237

Brandouy, O. and Mathieu, P. (2007). A conceptual framework for the evaluation of agent-base trading and technical analysis. *Artificial Markets Modeling. Methods and Applications. Lecture Notes in Economics and Mathematical Systems*, **599**, 63–79. 56

—, — and Veryzhenko, I. (2009). Ex-post optimal strategy for the trading of a single financial asset. *SSRN eLibrary.* 91

Brenner, T. (2006). Agent learning representation advice in modelling economic leanrning. In L. Tesfatsion and K. Judd (eds.), *Handbook of Computational Economics*, vol. 2, Elsevier B.V., pp. 896–922. 61

Brianzoni, S. (2012). Wealth distribution in an asset pricing model: the role of the switching mechanism. *Applied Mathematical Sciences*, **6**, 423–442. 90

—, Mammana, C. and E.Michetti (2010). Updating wealth in an asset pricing model with heterogeneous agents. *Discrete Dynamics in Nature and Society*, **2010**, 27. 90

Broadie, M. (1993). Computing efficient frontiers using estimated parameters. *Annals of Operations Research*, **45**, 21–58. 36

Brock, W. and Hommes, C. (1997). A rational route to randomness. *Econometrica*, **65**, 1059–1095. 50

— and — (1998). Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control*, **22**, 1235–1274. 23

—, J.Lakonishok and LeBaron, B. (1992). Simple technical trading rules and the stochastic properties of stock returns. *Journal of Finance*, **47**, 1731–1764. 30, 198

— and LeBaron, B. (1996). A structural model for stock return volatility and trading volume. *Review of Economics and Statistics*, **78**, 94–110. 24

Brown, S. (1976). *Optimal portfolio choice under uncertainty: A Bayesian approach*. Ph.D. thesis, University of Chicago. 163

— (1979). The effect of estimation risk on capital market equilibrium. *Journal of Financial and Quantitative Analysis*, **14**, 215–220. 162

Bullard, J. and Duffy, J. (1998). *Learning and Excess Volatility*. Tech. Rep. Macroeconomics Seminar Paper, Federal Reserve Bank of St. Louis. 19

Chakrabarti, R. (2000). Just another day in the inter-bank foreign exchange market. *Journal of Financial Economics*, **1**, 29–64. 49

— and Roll, R. (1999). Learning from others, reacitng and market quality. *Journal of Financial Markets*, **2**, 153–178. 49

Challet, D., Marsili, M. and Zhang, Y.-C. (2005). *Minority Games: Interacting Agents in Financial Markets*. Oxford University Press. 50

Chan, N., LeBaron, B., Lo, A. and Poggio, T. (1999). Agent-based models of financial markets. comparision with experimental markets. *MIT Artificial Markets Project*, **124**. 35

— and Shelton, C. (2001). An electronic market maker, working Paper AI Memo 2001-005, Massachusetts Institute of Technology. 50, 61

Chen, A.-S. and Leung, M. T. (2004). Regression neural network for error correction in foreign exchange forecasting and trading. *Computers & Operations Research*, **31** (7), 1049 – 1068. 229

—, — and Daouk, H. (2003). Application of neural networks to an emerging financial market: forecasting and trading the taiwan stock index. *Computers & Operations Research*, **30** (6), 901 – 923. 229

CHEN, S.-H. and HUANG, Y.-C. (2004a). Risk preference and survival dynamics. *Proceedings of the 3rd International Workshop on Agent-based Approaches in Economic and Social Comples Systems*, pp. 9–16. 172

— and — (2004b). Risk preference, forecasting accuracy and survival dynamics: Simulations based on a multi-asset agent-based artificial stock market. *Working Paper Series.* 172, 176

—, WANG, P. and KUO, T.-W. (2007). Computational intelligence in economics and finance: Shifting the research frontier. In *Computational Intelligence in Economics and Finance*, vol. II, Springer-Verlag Berlin Heidelberg. iii, 143, 172, 174, 176

— and YEH, C.-H. (2001a). Evolving traders and the business school with genetic programming: A new architecture of the agent-based artificial stock market. *Journal of Economic Dynamics and Control*, **25**, 363–393. 31, 61

— and — (2001b). Evolving traders and the business school with genetic programming: A new architecture of the agent-based artificial stock market. *Journal of Economic Dynamics & Control*, **25(3-4)**, 363–393. 49, 59, 60

— and YEN, C.-H. (2002). On the emergent properties of artificial stock markets. *Journal of Economic Behavior and Organization*, **49**, 217–239. 61

—, — and LIAO, C.-C. (2002). On aie-asm: a software to simulate artificial stock markets with genetic programming. *in: S.-H. Chen (Ed.) Evolutionary Computation in Economics and Finance, Physica-Verlag*, pp. 107–122. 60, 81

CHIARELLA, C. and HE, X. (2001). Asset price and wealth dynamics under heterogeneous expectations. *Quantitative Finance*, **1**, 509–526. 60, 136, 171

— and — (2002). Heterogeneous beliefs, risk and learning in a simple asset pricing model. *Computational Economics*, **19**, 95–132. 136

CINCOTTI, S., PONTA, L. and RABERTO, M. (2005). A multi-assets artificial stock market with zero-intelligence traders. *WEHIA*. 53, 60, 70

CIZEAU, P., LIU, Y., MEYER, M., PENG, C.-K. and STANLEY, H. (1997). Volatility distribution in the s&p500 stock index. *Physica A*, **245**. 19

CLIFF, D. and BRUTEN, J. (1997a). *Minimal-intelligence agents for bargaining behaviors in market-based environments.* Tech. rep., Hewlett-Packard Research Labs, Bristol, England, technical report HP-97-91. 55

— and — (1997b). *More than zero intelligence needed for continuous double-auction trading.* Tech. rep., Hewlett-Packard Research Labs, Bristol, England, technical report HP-97-157. 55

COHEN, K., MAIER, S., SCHWARTZ, R. and WHITCOMB, D. (1983). A simulation model of stock exchange trading. *Simulation*, **41**, 181–191. 67

Cont, R. (2001). Empirical properties of asset returns: stylized facs and statistical issues. *Quantitative Finance*, **1(2)**, 223–236. 20, 94, 151

— (2007). Volatility clustering in financial markets: empirical facts and agent-based models. *Journal of Financial Economics*, **56**, 29–64. 28, 29, 30

— and Bouchaud, J.-P. (2000). Herd behavior and aggregate fluctuations in financial markets. *Macroeconomic Dynamics*, **4**, 170–196. 24, 49

—, Potters, M. and Bouchaud, J.-P. (1997). Scaling in stock market data: Stable laws and beyond. *Lecture given at Les Houches Workchop on Scale Invariance.* 19

Cuthbertson, K. and Nitzsche, D. (2004). *Quantitative Financial Economics. Stocks,Bonds and Foreign Exchange.* John Wiley & Sons, Ltd, 2nd edn. 40

Cutler, D., Poterba, J. and Summers, L. (1989). What moves stock prices? *Journal of Portfolio Management*, **15**, 170–196. ii, 12, 23

Cvitanic, J., Jouini, E., Malamud, S. and Napp, C. (2011). Financial markets equilibrium with heterogeneous agents. *Review of Finance*, **Fothcoming**. 39

Dacorogna, M., Gencay, R., Muller, U., Olsen, R. and Pictet, O. (2001). *An Introduction to High-Frequency Finance.* London: Academic Press. 237

Daniels, M., Farmer, J., Iori, G. and Smith, E. (2002). A quantitative model of trading and price formation in finacial markets. *arXiv:cond-mat.* 56

—, —, — and — (2003). Quantitative model of price diffusion and market friction based on trading as mechanistic random process. *Physical Review Letter 90*, **10/108102/1-4**. 50, 56

Dantzig, G. (1966). All shortest routes in a graph. In *Theory of Graphs, International Symposium, Rome*, Gordon and Breach, New York, pp. 91–92. 218

Darley, V. and Outkin, A. (2007). *A NASDAQ market simulation. Insights on a major market from the science of complex adaptive systems*, *Complex Systems and Interdisciplinary Science*, vol. 1. Singapore: World Scientific Publishing Co. Pte. Ltd. 50, 66, 76, 77

Dave Cliff, J. B. (1997). Zero is not enough: On the lower limit of agent intelligence for continuous double auction markets, technical Report HPL-97-141, Hewlett-Packard Laboratories Bristol. 105

DeLong, J., Shleifer, A., Summers, L. and Waldmann, R. (1990). Positive feedback investment strategies and destabilizing rational speculation. *Journal of Finance*, **45**, 379–395. ii

DeMiguel, V., Garlappi, L. and Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Review of Financial Studies*, **22**, 1915–1953. vii, 158, 159, 164, 165, 167

Dijkstra, E. (1959). A note on two problems in connection iwth graphs. *Numerische Mathematik*, **1**, 269–271. 218

Donohue, C. and Yip, K. (2003). Optimal portfolio rebalancing with transaction costs. *The Journal of Portfolio Management*, **Summer**, 49–63. 185

Dorofeenko, V. and Shorish, J. (2002). Dynamical modeling of the demographic prisoner's dilemma. In *Economic Series*, 124, Vienna: Institute for Advanced Studies. 14

Edmonds, B. (1999). *Computational Techniques for Modelling Learning in Economics*, Kluwer, chap. Modelling bounded rationality in agent-based simulations using the evolution of mental models. 61

Ehrentreich, N. (2002). A corrected version of the santa fe institute artificial stock market model. *Social Science Computer Review*, **2** (20), 174–196. 67

— (2006). Technical trading in the santa fe institute artificial stock market revisited. *Journal of Economic Behavior & Organization*, **61**, 599–616. 67

Ellsberg, D. (1961). Risk, ambiguity and the savage axioms. *Quarterly Journal of Economics*, **76**, 643–669. 33

Elton, E. J., Gruber, M. J. and Blake, C. R. (1996). The persistence of risk-adjusted mutual fund performance. *Journal of Business*, **69** (2), 133–57. 199

Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of u.k. inflation. *Econometrica*, **50**, 987–1008. 21

— and Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Reviews*, **5**, 81–87. 22

Epstein, J. (1999). Agent-based computational models and generative social science. *Complexity*, **4(5)**, 41–57. 8

— (2006a). *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton University Press. 13

— (2006b). Remarks on the foundations of agent-based generative social science. In L. Tesfatsion and K. Judd (eds.), *Handbook of Computational Economics*, vol. 2, Elsevier B.V., pp. 1582–1602. 14

— and Axtell, R. (1996). *Growing Artificial Societies: Social Science from the Bottom Up*. MIT Press, Cambridge, MA. 8, 19

Érdi, P. (2008). *Complexity Explained*. Springer-Verlag Berlin Heidelberg. 6

Fama, E. (1965). The behavior of stock-market prices. *Journal of Business*, **38** (1), 34–105. i, 22

— and Blume, M. (1966). Filter rules and stock-market trading. *Journal of Business*, **39**, 226–241. 30, 31, 198

Farmer, J. (2002). Market force, ecology and evolution. *Industrial and Corporate Change*, **11**, 895–953. ii, 24, 49, 50, 52, 58

— and Joshi, S. (2000). The price dynamic of common trading strategies. *SFI Working Paper 00-12-069*. 49, 136

—, Patelli, P. and Zovko, I. (2005). The predictive power of zero intelligence in financial markets. *Proceedings of the National Academy of Scence*, **102**, 2254–2259. 51, 105, 111, 127

Ferber, J. and Muller, J.-P. (1996). Influence and reaction: a model of situated multiagent systems. In *Second International Conference on Multiagent Systems ICMAS-96*, pp. 72–79. 64

Fligner, M. and Policello, G. (1981). Robust rank procedures for the benrens-fisher problem. *Journal of the American Statistical Association*, **76**, 162–168. 122

Floyd, R. (1969). Algorithm 97, shortest path algorithms. *Operations Research*, **17**, 395–412. 216, 218

Foucault, T. (1999). Order flow composition and trading costs in a dynamic limit order market. *Journal of Financial Markets*, **2**, 99–134. 108

Franke, R. and Westerhoff, F. (2009). Estimation of a structular stochastic volatility model of asset pricing. *Computational Economics*, **38**, 53–83. 23

Frankel, J. and Froot, K. (1986). Understanding the us dollar in the eighties: the expectations of chartists and fundamentalists. *Economic Record*, **1(2)**, 24–38. 12

Friedman, M. (1953a). The case of flexible exchange rates. In *Essay in Positive Economics*, University of Chicago Press. 10

— (1953b). *Essays in Positive Economics*, Chicago: University of Chicago Press, chap. The Methodology of Positive Economics. 8

Frijns, B., Lehnert, T. and Zwinkles, R. (2008). Modeling structural changes in the volatility process. *Journal of Empirical Finance*, **21**, 1187–1222. 23

Fudenberg, D. and Tirole, J. (1991). *Game Theory*. Cambridge: MIT Press. 12, 173

Garder, M. (1970). Mathematical games: The fantastic combinations of john conway's new solitaire game 'life'. *Scientific American*, **October**. 87

GAREY, M. R. and JOHNSON, D. S. (1979). *Computers and Intractability: A Guide to the Theory of NP-Completeness.* W.H. Freeman. 211

GLOSTEN, L. and MILGROM, P. (1985). Bid, ask and transactions prices in a specialist model with heterogeneously informed traders. *Journal of Financial Economics*, **14**. 57

GODE, D. and SUNDER, S. (1993). Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. *Journal of Political Economy*, **101**, 119–137. 15, 55, 90, 105

GOEL, N., MAITRA, S. and MONTROLL, E. (1971). On the volterra and other nonlinear models of interacting populations. *Reviews of Modern Physics*, **43**, 231–277. 5

GORDILLO, J. L., PARDO-GUERRA, J. and STEPHENS, C. (2004). Adaptation in the presence of exogeneous information in an artificial financial market. *MICAI 2004: Advances in Artificial Intelligence, Lecture Notes in Compute Science*, **2972**, 342–351. 49, 53

GORDON, M., OARADIS, G. and RORKE, C. (1972). Experimental evidence on alternative portfolio rules. *American Economic Review*, **62**, 107–118. 143

GOUACHIN, A., MICHEL, F. and GUIRAUD, Y. (2004). Mic: A deployment environment for autonomous agents. In *E4MAS 2004*, Springer-Verlag, pp. 109–126. 63, 64

GROTHMANN, R. (2002). *Multi-agent market modeling based on neural networks.* Ph.D. thesis, Faculty of Economics, University of Bremen. 17, 54, 58

GUILLAUME, D., DACOROGNA, M., DAVÉ, R., MULLER, U., OLSEN, R. and PICTET, O. (1997). From the birds eye view to the microscope: A survey of new stylized facts of the intraday foreigh exchange markets. *Finance and Stochastics*, **1**, 95–131. 24

HANDA, P. and SCHWARTZ, R. A. (2006). The limit order effect. *Journal of Finance*, **125** (5), 094302. 108

HANSEN, L. and SINGLETON, K. (1982). Genaralized instrumental variables estimation of nonlinear rational expectation models. *Econometrica*, **50**, 1269–1286. 143

HENAFF, P. and MARTINI, C. (2011). Model validation: theory, practice and perspectives. *The Journal of Risk Model Validation*, **5**, 3–15. 238

HERSTEIN, I. and MILHOR, J. (1953). An axiomatic approach to measurable utility. *Econometrica*, **21**, 291–297. 32

HOLLAND, J. (1975). *Adaptation in Natural and Artificial Systems.* University of Michigan Press. 61

Hommes, C. (2001). Financial markets as nonlinear adaptive evolutionary systems. *Quantitative Finance*, **1**, 149–167. 10, 59

— (2006). Heterogeneous agent models in economics and finance. In L. Tesfatsion and K. Judd (eds.), *Handbook of Computational Economics*, vol. 2, Elsevier B.V., pp. 882–892. ii, viii, 10, 106, 136, 143, 159, 173

Huang, W., Nakamori, Y. and Wang, S.-Y. (2005). Forecasting stock market movement direction with support vector machine. *Computers & Operations Research*, **32** (10), 2513 – 2522. 229

Iglesias, J., Goncalves, S., Abramson, G. and J.L.Vega (2004). Correlation between risk aversion and wealth distribution. *Physica A*, **342**, 186–192. 83

Jacobs, B., Levy, K. and Markowitz, H. (2010). Simulating security markets in dynamic and equilibrium modes. *Financial Analysts Journal*, **66** (5), 42–53. 52, 142

Jacobs, B. I., Levy, K. N. and Markowitz, H. M. (2004). Financial market simulation. *The Journal of Portfolio Management*, **30th Anniversary Issue**, 142–151. v, 49, 52, 80, 81, 140

Jamal, K. and Sunder, S. (2001). Why do biased heuristics approximate bayes rule in double action? *Journal of Economic Behavior & Organisation*, **46(4)**, 431–435. 55

Jevons, W. (1871). *The Theory of Political Economy*. Macmillan and Co. 9

Johnson, N. (1949). Systems of frequency curves genrated by methods of translation. *Biometrika*, **1-2**, 149–176. 133

Johnson, P. (2002). What i learned from the artificial stock market. *Social Science Computer Review*, **2** (20), 174–196. 69

Joshi, S. and Bedau, M. (2000). Technical trading creates a prisoner's dilemma: Results from an agent-based model. In *Proceedings of the 6th International Conference Computational Finance*, MIT Press, pp. 465–479. 31

Jouini, E. and Napp, C. (2012). Behavioral biases and the representative agent. *Theory and Decision*, **Fothcoming**. 12

Kahneman, D., Solvic, P. and Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press. 42

— and Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, **80**, 237–251. 10

— and — (2000). *Choices, values, and frames*. Cambridge University Press. 42

KAIZOJI, T. (2000). Speculative bubbles and crashes in stock marketsl an interacting-agent model of speculative acrivity. *Physica A, Statistical Mechanics and its Applications*, **Special issue, 287** (3-4), 493–506. 58

KALLBERG, J. and ZIEMBA, W. (1983). Comparison of alternative utility functions in portfolio selection problems. *Management Science*, **29**, 1257–1276. 143, 166, 174, 176

KANDEL, S. and STAMBAUGH, R. (1991). Asset returns and intertemporal preferences. *Journal of Monetary Economics*, **27**, 39–71. 143

KARNI, E. (1985). *Decision Making under Uncertainty: the Case of State-Dependent Preferences*. Cambridge MA: Harvard University Press. 33

KIM, G. and MARKOWITZ, H. (1989). Investment rules, margin, and market volatility. *Journal of Portfolio Management*, **16**, 45–52. v, 67

KING, R., SMITH, V., WILLIAMS, A. and BENING, M. (1993). *Nonlinear Dynamics and Evolutionary Economics*, Oxford University Press, chap. The Robustens of Bubbles and Crashes in Experimental Stock Markets, pp. 183–200. 106

KIRCHLER M., H. J. (2007). Fat tails and volatility clustering in experimental asset markets. *Journal of Economic Dynamics and Control*, **31** (6), 1844–1874. 106

KIRMAN, A. (1993). Ants, rationality, and recruitment. *Quarterly Journal of Economics*, **108**, 137–156. 65

— and TEYSSIERE, G. (2002). Microeconomic models for long memory in the volatility of financial time series. *Studies in Nonlinear Dynamics & Economics*, **5**, 675–702. 24

KLEIN, R. and BAWA, V. (1976). The effect of estimation risk on optimal portfolio choice. *Journal of Financial Economics*, **3**, 215–231. 162

KRITZMAN, M., PAGE, S. and TURKINGTON, D. (2010). In defese of optimization: The fallacy of 1/n. *Financial Analysts Journal*, **66**, 31–40. viii, 158

KULLBACK, S. and LEIBLER, R. (1951). On information and sufficiency. *Annuals of Mathematical Statistics*, **22**, 79–86. 257

LADLEY, K. and SCHENK-HOPPE, K. R. (2009). Do stylised facts of order book markets need strategic behaviour? *Journal of Economic Dynamics and Control*, **33**, 817–831. 24, 105

LAMOUREUX, C. and LASTRAPES, W. (1990). Heteroskedasticity in stock return data: Volume versus garch effects. *The Journal of Finance*, **45**, 221–229. 23

LEBARON, B. (2001a). A builder's guide to agent based financial markets. *Quantitative Finance*, **1(2)**, 254–261. v, 48, 49, 52, 57, 65

— (2001b). Empirical regularities from interacting long- and short- memory investors in an agent based stock market. *IEEE Transactions on Evolutionary Computation*, **5(5)**, 442–455. 57

— (2001c). Evolution and time horizons in an agent based stock market. *Macroeconomic Dynamics*, **5**, 225–254. 24, 53, 57, 60, 61, 84

— (2002). Building the santa fe artificial stock market. *Working Paper*. 67, 69

— (2006). Agent-based computational finance. In L. Tesfatsion and K. Judd (eds.), *Handbook of Computational Economics*, vol. 2, Elsevier B.V., pp. 882–892. v, 136

—, Arthur, W. and Parlmer, R. (1999). Time series properties of an artificial stock market. *Journal of Economic Dynamics and Control*, **23**, 1487–1516. 23, 24, 31, 61, 62, 67, 69, 106, 198

Ledoit, O. and Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, **88**, 365–411. 35

Levy, H. (1978). Equilibrium in an imperfect market: A constraint on the number of securities in the portfolio. *American Economic Review*, **68**. 43, 44

Levy, M. (2005). Is risk-aversion hereditary? *Journal of Mathematical Economics*, **41**, 157–168. 171

—, Levy, H. and Solomon, S. (1995). Microscopic simulation of the stock market: the effect of microscopic diversity. *Journal de Physique I (France)*, **5**, 1087–1107. viii, 7, 8, 33, 49, 52, 60, 143, 159, 171

—, — and — (2000). *Microscopic Simulation of Financial Markets: From Investor Behaviour to Market Phenomena*. Academic Press. ii, iii, 33, 38, 40, 43, 44, 51, 171

— and Ritov, Y. (2011). Mean-variance efficient portfolios with many assets: 50% short. *Quantitative Finance*, **11** (10), 1461–1471. 165, 167, 168, 183, 235

Lintner, J. (1965). Security prices, risk and maximal gains from diversification. *Journal of Finance*, **XX** (4), 587–615. 33, 43

Liu, H. (2004). Optimal consumption and investment with transaction costs and multiple risky assets. *The Journal of Finance*, **LIX**, 289–338. 185

Liu, X., Gregor, S. and Yang, J. (2008). The effects of behavioral and structural assumptions in artificial stock market. *Pysica A Statistical Mechanisms and its Applications*, **387**, 2535–2546. 106, 107

Llacay, B. and Peffer, G. (2010). Realistic agent-based simulation of financial crises: the 1998 turmoil. In A. Gil-Lafuente and J. Merigo (eds.), *Computational Intelligence in Business and Economics. Proceeding of the MS'10 Internation Conference*, vol. 3, pp. 189–197. iii

Loistl, O., Schossmann, B. and Vetter, O. (2001). Xetra efficient evaluation and nasdaq modeling by kapsyn. *European Journal of Operational Research*, **135**, 270–295. 53

— and Vetter, O. (2000). *KapSyn Computer Modeled Stock Exchanges, User Manual, Ver.3.02.* 53

— and Veverka, A. (2004). *An integration of current markets microstructure results into a consistent picture catallactic modeling of capital market's micro structure.* Tech. rep., Vienna University of Economics and Business Administration. 53

Long, J. (1974). Stock prices, inflation, and hte term structure of interest rates. *Journal of Financial Economics*, **41**, 131–170. 37

Loomes, G. and Sugden, R. (1982). Regret theory: an alternative theory of rational choice under uncertainty. *Economic Journal*, **92**, 805–824. 33

Lotka, A. (1910). Contribution to the theory of periodic reaction. *Journal of Physical Chemistry A*, **14**, 271–274. 5

— (1925). *Elements of Physical Biology.* Baltimor: Williams and Wilkins. 137, 163

Lux, T. (2000). Volatility clustering in financial markets: a microsimulation of interacting agents. *International Journal of Theoretical and Applied Finance*, **3**, 675–702. 24, 25, 28, 29, 50

— (2009). *Handbook of Financial Markets: Dynamics and Evolution*, Academic Press, chap. Stochastic behavioral asset pricing models and the stylized facts, pp. 161–215. 24, 104, 106

— and Marchsi, M. (1999). Scaling and criticality in a stochastich multi-agent model of a financial market. *Nature*, **397**, 498–500. 64, 68, 136

Machina, M. (1982). Expected utility analysis without the independence axiom. *Econometrica*, **50**, 227–232. 33

Malkiel, B. (2004). Can predictable patterns in market returns be exploited using real money? *Journal of Portflio Management*, **30**, 131–141. 199

Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*, **XXXVI**, 392–702. 21, 22

— (2006). *Biologically Inspired Algorithms for Financial Modelling.* Springer-Verlag Berlin Heidelberg. 14

Mankiw, G., Rotemberg, J. and Summers, L. (1985). Intertemporal substitution in macroeconomics. *Quarterly Journal of Economics*, **100**, 225–251. 143

Mannaro, K., Marchesi, M. and Setzu, A. (2008). Using an artificial financial market for assessing the impact of tobin-like transaction taxes. *Journal of Economic Behavior & Organization*, **67**, 445–462. 72

Mantegna, R. (1991). Lévy walks and enhanced diffusion in milan stock exchange. *Physica A: Statistical Mechanism nad its Applications*, **179**, 232–242. 19

— and Stanley, H. (1999). *An Introduction to Econophysics: Correlations and Complexity in Finance*. MA: Cambridge University Press. 237

Marchesi, M., Cincotti, S., Focardi, S. and Roberto, M. (2000). Development and testing of an artificial stock market. *Model Dynamic in Economic Finance*. 49, 70

Markose, S., Tsang, E. and Martinez-Jaramillo, S. (2003). The red queen principle and the emergence of efficient financial markets: an agent based approach. In P. Angeline, Z. Michalewicz, M. Schoenauer, X. Yao and A. Zalzala (eds.), *Workshop on economics and heterogeneous interacting agents (WEHIA)*, Kiel, Germany: Springer, vol. 2, pp. 1253–1259. 52

Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, **7**, 77–91. 32, 33, 34, 92

— (1959). *Portfolio Selection: Efficient Diversification of Investment*. New York: John Wiley & Sons. 33

— (1988). Stock market simulator sms1. program description. *Baruch College Working Paper Series 88-24*. v

— (1990). Risk adjustment. *Journal of Finance, Accounting and Auditing*. 43

Marschak, J. (1950). Rational behaviour, uncertain prospects and measurable utility. *Econometrica*, **18**, 111–141. 32

Martinez-Jaramillo, S. (2007). *Artificial Financial Markets: An Agent Based Approach to Reproduce Stylized Facts and to study the Red Queen Effect*. Ph.D. thesis, Centre for Computational Finance and Economic Agents, University of Essex. 48, 49, 61, 80

Maslov, S. (2000). Simple model of a limit order-driven market. *Physica A*, **278**, 571–578. 49, 105, 107

Masters, S. (2003). Rebalancing. *The Journal of Portfolio Management*, pp. 5–10. 185, 188, 191, 235

Mathieu, P. and Brandouy, O. (2011). Efficient monitoring of financial orders with agent-based technologies. In *9th International conference on Practical Applications of Agents and Multi-Agents Systems (PAAMS'2011)*, Salamanca (Spain), pp. 277–286. 238

— and — (2012). Introducing atom. In *The 10th International conference on Practical Applications of Agents and Multi-Agent Systems (PAAMS'2012)*, pp. 127–128. 78

McCauley, J. (2004). *Dynamics of Markets: Econophysics and Finance*. University of Houston: Cambridge University Press. 5

Mehra, R. and Prescott, E. (1985). The equity premium: a puzzle. *Journal of Monetary Economics*, **15**, 145–161. 171

Menger, C. (1871). *Principles of Economics*. reprinted by Ludwig von Mises Institute in 2007. 9

Merton, R. (1971). Optimum consumption and portfolio management with fixed transaction costs. *Journal of Economic Theory*, **3**, 373–413. 185

— (1972). An analytical derivation of the efficient portfolio frontier. *Journal of Financial and Quantitative Analysis*, pp. 1851–1872. 33

— (1973). An intertemporal capital asset pricing model. *Econometrica*, **41**, 867–887. 37

— (1987). A simple model of capital market equilibrium with incomplete information. *Journal of Finance*, **42**. 43, 44

Michel, F. (2007). The irm4s model: The influence/reaction principle for multiagent based simulation. In *AAMAS'07. Sixth International Joint Converence on Autonomous Agents and Multiagent Systems*, pp. 908–910. 64

Milgrom, P. and Stokey, N. (1982). Information, trade and common knowledge. *Journal of Economic Theory*, **26**, 17–27. 12, 173

Motiwalla, L. and Wahab, M. (2000). Predictable variation and profitable trading of us equities: a trading simulation using neural networks. *Computers & Operations Research*, **27** (11-12), 1111 – 1129. 229

Muchnik, L. and Solomon, S. (2006). Markov nets and the natlab platform; application to continuous double auction. *New Economic Windows*. 84

Neely, C. (1997). Technical analysis in the foreign exchange market: A laymans guide. *Review*, **September**. 56

—, Weller, P. and Dittmar, R. (1997). Is technical analysis in the foreign exchange market profitable? a genetic programming approach. *Financial and Quantitative Analysis*, **32**, 405–426. 56

Nelson, D. (1989). Commentary: price volatility, international markets links, and their implications for regulatory policies. *Journal of Financial Services Research*, **3**, 247–254. 22

— (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, **59**, 347–370. 22

Obstfeld, M. (1994). Risk taking, global diversification, and growth. *American Economic Review*, **84**, 1310–1329. 143

Osborne, M. (1959). Brownian motion in the stock market. *Operations Research*, **7**, 145–173. 5

Overway, C. and John, B. (2006). Rebalancing multi-asset portfolios. *Natixis Global Associates Whitepaper*. 186

Palmer, R., Arthur, W., Holland, J. and LeBaron, B. (1998). An artificial stock market. *Artificial Life and Robotics*, **3**, 27–31. 57, 59

—, —, —, — and Tayler, P. (1994). Artificial economic life: A simple model of stock market. *Physica D*, **75**, 264–274. 61, 67

Papadimitriou, C. and Steigleitz, K. (1998). *Combinatorial Optimization: Algorithms and Complexity*. Mineola. New York: Dover Publications, Inc. 216

Pellizzari, P. (2005). Complexities and simplicity: a review of agent-based artificial markets. In *Giornata di studi in onore di Giovanni Castellani*, Venice. 4

Plott, C. and Sunder, S. (1982). Efficiency of epxerimental sycirity markets with insider information: an application of rational expectation models. *Journal of Political Economy*, **90**, 663–698. 106

Pompian, M. (2006). *Behavioral Finance and Wealth Management. How to Build Optimal Portfolios That Account for Investor Biases*. John Wiley & Sons, Inc. i

Ponta, L., Raberto, M. and Cincotti, S. (2011). A multi-assets artificial stock market with zero-intelligence traders. *A Letter Journal Exploring the Frontiers of Physics*, **93**, 28002–p1–28002–p7. 49, 70, 81, 90

Prietula, M., Carley, K. and Gasser, L. (1998). *Simulating Organizations*. MA: MIT Press. 19

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behaviour and Organization*, **3**, 323–343. 33

Raberto, M. and Cincotti, S. (2005). Modeling and simulation of a double auction artificail financial market. *Physica A: Statistical Mechanics and its Applications*, **355**, 34–45. 52

—, —, Focardi, S. and Marchesi, M. (2001). Agent-based simulation of a financial market. *Physica D. Nonlinear phenomena*, **299**, 319–327. 50, 70

—, —, — and — (2003). Traders' long-run wealth in an artificial financial market. *Computational Economics*, **22**, 255–272. 72

Roll, R. (1977). A critique of the asset pricing theory's tests. *Journal of Financial Economics*, **4**, 129–176. 38

Ross, S. (1976). The arbitrage pricing theory. *Journal of Economic Theory*, **1**, 341–360. 32, 40

Rosser, J. (2004). *Complexity in Economics: The International Library of Critical Writings in Economics*. Edward Elgar, Aldergate. 4

Samuelson, P. (1950). Probability and attempts to measure utility. *Economic Review*, **1**, 167–173. 32

— (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, **6**, 41–49. 5

Schelling, T. (1978). *Micromotives and Macrobehavior*. Norton. 12

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, **57**, 571–587. 33

Schulenburg, S. and Ross, P. (2002). Explorations in lcs models of stock trading. In P. Lanzi, W. Stolzmann and S. Wilson (eds.), *Advances in Leraning Classifier Systems: 4th Internation Workshop, IWLCS 2001*, San Francisco CA, USA: Springer Berlin/Heidelberg, vol. 2321, Lecture Notes in Computer Science, pp. 150–179. 61

Sedgewick, R. and Wayne, K. (2011). *Algorithms*. Pearson Education, Inc., 4th edn. 219, 220

Sharpe, W. (1963). A simplified model for portfolio analysis. *Management Science*, **9** (2), 277–293. 33

— (1964). Capital asset prices - a theory of market equilibrium under conditions of risk. *Journal of Finance*, **XIX** (3), 425–442. 32, 33, 37, 171

— (1966). Mutual fund performance. *Journal of Business*, **39(S1)**, 119–138. 164

— (1991). Capital asset pricing with and without negative holdings. *Journal of Finance*, **June**. 43, 44, 199

Shatner, M., Mushnik, L. and Solomon, S. (2000). A continuous time asynchronous model of the stock market. *arXiv:cond-mat*. 51, 52, 56, 57, 143, 173

Shier, D. (1973). Iterative methods for determining the k shortest paths in a network. *Networks*, **6**, 205–230. 218

Shleifer, A. (2000). *Inefficient markets*. Oxford: Oxford University Press. ii

SIMON, H. (1957). *Models of Man*. New York, NY: Wiley. 10

— (1962). The architecture of complexity. *Proceedings of the American Phylosophical Society*, **106**. 4

SMITH, V. (1962). An experimental study of competitive market behavior. *Journal of Political Economy*, **70**, 111–137. 105, 198

—, SUCHANEK, G. and WILLIAMS, A. (1988). Bubbles, crashes, and endogeneous expectations in experimental spot asset markets. *Econometrica*, **56**, 1119–1151. 106

STEWART, J. (2005). Rebalancing equity allocation. *The Journal of Investing*, **Summer**. 188

STIGLER, G. (1966). *The Theory of Price*. New York: Macmillen. 8

SWEENEY, R. J. (1986). Beating the foreign exchange market. *Journal of Finance*, **41**, 163–182. 30

— (1988). Some new filter rule tests: Methods and results. *Journal of Financial and Quantitative Analysis*, **23**, 285–300. 30

TAKAHASHI, H. and TERANO, T. (2004). Analysis of micro-macro structure of financial markets via agent-based model: Risk management and dynamics of asset pricing. *Electronics and Communications in Japan*, **87**, 618–628. 42

TEFATSION, L. and JUDD, K. (eds.) (2006). *Handbook of Computational Economics. Agent-Based Computation Economics*, vol. 2. The Netherlands: Elsevier B.V. iii, 16, 33, 136

TESFATSION, L. (1995). *A Trade Network Game with Endogenous Partner Selection*. Tech. Rep. Economic Report 36, Iowa State University, Department of Economics. 19

TKATCH, I. and ALAM, Z. S. (2009). Strategic order splitting in automated markets. *SSRN eLibrary*. 91

TOBIN, J. (1958). Liquidity preference as behavior towards risk. *The Review of Economic Studies*, **25** (2), 65–86. 33

TREYNOR, J. (1962). Toward a theory of market value of risky assets, fall 1962. A final version was published in 1999 in Asset Pricing and Pertfolio Performance. Robert A. Korajczyk (editor) London: Risk Books, pp.15-22. 7, 162

TU, J. and ZHOU, G. (2011). Markowitz meets talmud: A combination of sophisticated and naive diversification strategies. *Journal of Financial Economics*, **99**, 204–215. viii, 158, 159, 164, 166, 168, 235

TVERSKY, A. and KAHNEMAN, D. (1974). Judgment under uncertainty: heuristics and biases. *Science*, **185**, 1124–1131. 10

— and — (1992). Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, **5**, 297–323. 33

VERYZHENKO, I., BRANDOUY, O. and MATHIEU, P. (2010). Agent's minimal intelligence calibration for realistic market dynamics. *Progress in Artificial Economics Computational and Agent-Based Models. AE 2010. Lecture Notes in Economics and Mathematical Systems 645*, pp. 3–14. 24, 56, 94, 96

VISSING-JORGENSEN, A. (2003). *NBER Macroeconomics Annual*, Cambridge: MIT Press, chap. Perspective on behavioral finance: does "irrationality" disappear with wealth? Evidence from expectations and actions. 12

VOLTERRA, V. (1926). *Animal Ecology*, New York: McGraw-Hill, chap. Variations and fluctuations of the number of individuals of animal species living together. 137, 163

VON NEUMAN, J. and MORGENSTERN, O. (1947). *Theory of Games and Economic Behavior*. Princeton NJ: Princeton University Press. 32

WALRAS, L. (1874). *Éléments d'économie politique pure, ou théorie de la richesse sociale*. reprinted by Routledge in 2003. 9

WALTER, S., AYRES, F., CHEN, L.-W., SCHOUWENNARS, T. and ALBOTA, M. (2006). Optimal rebalancing for institutional portfolios. *The Journal of Portfolio Management*, **Winter**. 185

WEAVER, W. (1948). Science and complexity. *American Scentist*, **36**, 536. 4

WHITE, H. (2000). A reality check for data snooping. *Econometrica*, **68**, 1097–1126. 198

WINDRUM, P., FAGIOLO, G. and MODETA, A. (2007). Empirical validation of agent-based models: Alternatives and prospects. *Journal of Artificial Societies and Social Simulation*, **10**, 137–156. 65

WOOLDRIDGE, M. (2002). *An Introduction to MultiAgent Systems*. John Wiley & Sons, LTD. 15

— and JENNINGS, N. (1995). Intelligent agents: Theory and practice. *The Knowledge Engineering Review*, **10** (2), 115–152. 54

WU, S. and BHATTACHARYYA, S. (2003). Minimal intelligence agents in double auction markets with speculators. In *Proceedings of 7th Joint Conference on Information Sciences*, pp. 1092–1095. 56

YANG, J. (1999). Heterogeneous beliefs, intelligent agents, and allocative efficiency in an artificial stock market. *Society for Computation Economics Seies of Computating in Economics and Finance*, **612**. 18, 59, 60

— (2002). The efficiency of an artificial double auction stock market with neural learning agents. *Evolutionary Computation in Economics and Finance*, pp. 85–106. 49, 59, 61

ZIMMERMANN, H., NEUNEIER, R. and GROTHMANN, R. (2001). *IEEE Transaction on Neural Networks*, IEEE Press, vol. Special issue, 12, chap. Multi-Agent Modeling of Multiple FX-Markets by Neural Networks, pp. 735–743. 52, 58, 61, 66

Abstract

The aim of this thesis is to contribute to the understanding of market dynamics and the decision making of investors by extending an agent-based computational approach. Agent-based modeling (ABM) studies stock market as complex evolving system by representing each of the microscopic elements individually and by simulating the behavior of the entire system, keeping track of all the individual elements. We, first, explore the framework of zero-intelligence traders (ZITs), that puts forward the role of market microstructure, for understanding at coarse grain what drives the main qualitative and quantitative stylized facts in price dynamics and patterns in order submissions. The results of extensive simulations indicate that realistic price dynamics are out-of-reach within the pure ZIT's framework, only the elements of strategic behavior and strong calibration improve this situation.
Next, this research focuses on the questions on rationality in the corpus of modern portfolio theory. Scientists still debate about ability of naive strategies to outperform the more complex models. The current research sheds new light on the topic. We test the investors' performance, each of them following a specific strategy, scrutinizing their behavior in ecological competitions. Some investment strategies considered in this thesis are based on different extensions of canonical modern Markowitz portfolio theory, others on the naive diversification principles, and others on combinations of sophisticated rational and naive strategies. Furthermore, we perform closer examination of the effects of rebalancing frequency and investor's attitude toward risk on portfolio performance in order to identify clearly what matters the most.
Finally, we explorer the computational tools for algorithmic determination of an absolute performance measure geared towards the ex-post evaluation of a wide range of trading strategies of investors (agents in our case).
Experimental results confirm a real added value of agent-based artificial market models for studying various financial topics. Notably, ABM allows to go beyond traditional approaches which may suffer from implementation drawbacks or absence of tractable result in some cases.

Résumé

Cette thèse apporte une contribution à la compréhension des dynamiques de marché et à la prise de décision des traders à l'aide d'une plateforme de simulation de marchés multi-agents. La modélisation multi-agents permet notamment d'étudier le système boursier comme un système complexe évolutif dans lequel chaque trader artificiel possède son propre comportement et qui, par ses prises de décision, influence l'ensemble des autres acteurs du système. Dans une première partie, nous mettons en évidence à l'aide de "traders à intelligence zéro" (ZIT), le rôle de la microstructure pour comprendre la nature des principaux faits stylisés de l'évolution des prix. Les résultats issus de nombreuses simulations, indiquent que l'usage des ZIT n'est pas suffisant pour reproduire de façon convaincante les évolutions de prix réels, car ceux-ci doivent être appréhendés à la fois de manière qualitative mais aussi quantitative. Nous montrons que seuls des éléments de stratégies de trading et une forte calibration peuvent améliorer cette réplication par simulation, suggérant que les aspects comportementaux importent tout autant que les aspects microstructurels.

Dans une seconde partie, nous concentrons notre recherche sur la problématique de la rationalité dans le corpus de la théorie moderne du portefeuille. Le marché artificiel nous permet de tester si des stratégies naïves peuvent surpasser, en terme de performance, des modèles plus complexes. Diverses stratégies d'investissement sont implémentées dans le système artificiel et mises en interaction afin d'observer leur survie dans des compétitions écologiques basées sur leurs performances relatives. Certaines de ces stratégies d'investissements sont fondées sur des variations du modèle canonique de la théorie de portefeuilles de Markowitz, d'autres suivent des principes de diversification naïfs, d'autres encore obéissent à des combinaisons de stratégies rationnelles sophistiquées et de stratégies naïves.

Enfin, de manière à mieux saisir les facteurs qui influent sur la performance du portefeuille, nous montrons les effets de la fréquence de pondération et des préférences pour le risque des investisseurs sur l'issue de ces compétitions.

Pour finir, afin de fournir une mesure de performance absolue orientée vers l'évaluation ex-post d'un large éventail de stratégies de trading des investisseurs (agents dans notre cas) nous proposons un nouvel algorithme de complexité polynomiale permettant de déterminer la borne supérieure absolue des profits atteignables pour n'importe quelle stratégie sur une période de temps donnée. Cet algorithme met en contact deux champs a priori éloignés: la théorie des graphes d'une part et la finance computationnelle d'autre part.