

Diss. ETH No. 22580

Detecting Bubbles in Financial Markets: Fundamental and Dynamical Approaches

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES OF ETH ZÜRICH
(Dr. sc. ETH Zürich)

presented by

ZALÁN FORRÓ

Msc. in Physics, École Polytechnique Fédérale de Lausanne

born on April 30th, 1987

Citizen of Ecublens, VD

accepted on the recommendation of

Prof. Dr. Didier Sornette, examiner

Prof. Dr. Thorsten Hens, co-examiner

2015

Abstract

The subject of this thesis has been the detection of bubbles in financial markets through two fundamentally different approaches.

The first approach consisted in quantifying bubbles as the difference between the market value and fundamental price of an asset, taking Zynga, a social networking company, as a case study. By recognizing the centrality of its users in its revenue structure, we were able to determine an upper bound for the value of the company by forecasting its future user base as well as the revenue each of them would generate with non-linear dynamical models. Our valuation led to our diagnosis of a bubble which in turn, combined with a financial event known as the end of the lock-up period, allowed us to successfully predict Zynga's downward price trajectory within a specific time-window. In addition, Zynga's long-term price dynamics have been consistent with our valuation performed shortly after its IPO. The generality of our methodology was emphasized by its successful application to a different problem : the forecasting of the future oil production of Norway and the U.K. The strength of this approach lied in the absence of a need to postulate a mechanism for the emergence of a bubble : while this made the detection of bubbles possible, forecasting their burst was difficult, except in rare cases such as the one described here.

Our second approach was based on the log-periodic power law model (LPPL), a model taking its roots in critical phenomenas, defining bubbles as transient super-exponential regimes punctuated by phase transitions. Our contribution has been to show in a systematic way and on a large scale that LPPL's predictive power was robust across a wide range of strategies, assets and time periods. This predictive power was defined as the persistent deviation between LPPL-based strategies and their random counterparts. We went on to generalize the importance of super-exponential regimes in the pricing of assets : we showed, in a factor regression model, that the first difference of past returns, in other words the difference between two consecutive growth rates, yielded significant predictive power over future returns. This phenomenon was captured by a new factor called Γ^s that outperformed and explained the momentum factor, a pillar of classical finance. This suggests that price acceleration (Γ^s) clearly dominates price velocity (momentum) in the pricing of assets, supporting the LPPL paradigm.

In summary, the work presented in this thesis supports a view of markets out-of-equilibrium, permeated by unsustainable regimes in which the detection of bubbles and the forecast of their crashes is possible.

Résumé

Le sujet de cette thèse a été la détection de bulles financières à travers deux approches très différentes.

La première approche a consisté en la quantification des bulles en tant que différence entre la valeur fondamentale et le prix du marché d'un actif, à travers l'exemple de Zynga, une compagnie de réseau social. Reconnaisant la centralité des utilisateurs dans les revenus de l'entreprise, nous avons pu déterminer une limite supérieure à sa valoration en modélisant l'évolution de sa base d'utilisateurs ainsi que les revenus générés par ceux-ci, via des modèles dynamiques non-linéaires. Notre valoration nous a conduit à conclure que Zynga était dans une bulle, information qui, combinée avec la fin de la période de blocage, nous a permis de prédire la chute de son cours à court terme. De plus, notre valoration a été cohérente avec l'évolution de Zynga à long terme. La généralité de notre méthodologie a été démontrée en l'appliquant avec succès à un problème de nature différente : la prédiction de la production pétrolière de la Norvège et du Royaume-Uni. La force de notre méthodologie a résidé dans l'absence de recours à un modèle spécifique pour l'émergence de bulles : bien que rendant la détection de bulles possibles, cela a rendu la prévision de leur crash plus difficile, à l'exception de celui décrit ci-dessus.

Notre deuxième approche s'est basée sur un modèle venant du monde des phénomènes critiques, le modèle de log-periodic power law (LPPL). Ce modèle définit les bulles comme des régimes transitoires super-exponentiels ponctués de transitions de phase. Notre contribution principale a été de montrer de manière systématique la robustesse du pouvoir prédictif de LPPL appliqué à un large éventail de stratégies, d'actifs et de périodes temporelles. Ce pouvoir prédictif a été défini en tant que déviation entre les stratégies basées sur LPPL et leurs homologues aléatoires. Nous avons ensuite généralisé l'importance des régimes super-exponentiels dans la valoration des actifs : nous avons montré, dans le cadre d'un modèle de régression de facteurs, que la différence première des rendements passés avait un pouvoir prédictif sur les rendements futurs. Ce phénomène a été exprimé à travers un nouveau facteur, Γ^s , qui a surpassé en performance et a expliqué le momentum, un pilier de la finance moderne. Ceci suggère que l'accélération des prix (Γ^s) domine la vitesse des prix (momentum) dans la valoration d'actifs, soutenant ainsi le paradigme de LPPL.

En résumé, notre travail soutient une vision des marchés hors d'équilibre, pénétrés de régimes dans lesquels la détection de bulles ainsi que celle de leur crash est possible.

Acknowledgements

A PhD program is like a sailing boat: there are many vectors which act on it in order to arrive to its destination. The speed of the wind, the muscles of the sailors, the force of the waves, the position of the rudder, they all count. I have to thank many people whose help was very precious during these three years.

First of all, I want to express my gratitude to my supervisor, Prof. Didier Sornette, for having offered me a position in his group. His enthusiasm, optimism and creativity were contagious. I am also grateful to him for the freedom he gave me, and for always having had an open door for discussions, despite his busy schedule. I also wish to thank Prof. Thorsten Hens for accepting to be my co-referee and Prof. Antoine Bommier for being the president of the thesis committee.

I am also grateful to Prof. Albert-László Barabási. It is in his group that, as a master student, I had my first exposure to research. It was a formative experience that gave me the necessary confidence to apply to the group of Prof. Sornette. I am more than grateful to him.

Every research topic presented here has been performed with collaborators. As such, I would like to thank my colleague and friend Peter Cauwels: it has been a pleasure to work with him. I learned valuable skills from him, not least how to write a convincing scientific paper. This thesis has definitely benefited from his valuable feedback.

I also wish to thank Ryan Woodard for his friendship and support during my PhD. He has always been very generous in sharing his knowledge and has had a great impact on this thesis by, among others things, introducing me to the magical world of Python, Pandas and Django. The three years spent here would not have been the same without those legendary poker nights!

It has also been a pleasure to work with Lucas Fiévet. I have very much enjoyed the various discussions we have had together.

Diego Andrés Ardila Alvarez (or just Diego) has been a great colleague; his Stakhanovite work power and deep knowledge of the literature made it a real pleasure to work with him.

I am grateful to Vladimir Filimonov who has been an outstanding office mate and friend during my years in the group. Being in the same office offered many advantages, one of which was the possibility of enjoying the delicious pastries of his wife, Rita.

I also would like to thank Donnacha Daly for having given me thorough and valuable feedback on all the chapters of this thesis. I hope to continue having those exciting chess games with him in the future!

I would not want to forget Hyun-U “Haya” Sohn, the personal Apple Store of the group! Being kept informed on all the latest Apple products was as simple as walking into his office. On a more serious note, I really enjoyed our philosophical, political, economical discussions during the past three years!

Thanks also to Afina Inina for the pleasant coffee breaks after lunch and for having been a more reliable badminton partner to me than I was to her.

I wish to extend my thanks to Tatyana for the Russian dictionary, Dorsa for managing Diego masterfully, Susanne for her nice laugh, Guilherme for bringing his South American cordiality to the group, Dmitry for that great bottle of vodka, Spencer for the Asturian chorizo, and Qunzhi for sharing his LPPL-library.

I also wish to thank the two administrative assistants Isabella Bieri and Heidi Demuth who took care of the administrative side of my everyday’s life with the great kindness that characterizes them.

Last but not least, I wish to thank my family:

I thank my mother, Livia, for her efforts in keeping my brother and I on the right track during our childhood; checking our homework, bringing us to our piano lessons, to karate and more. I would not be where I am today were it not for her.

I thank my father, László, for having showed me, from a young age, how much fun science could be! I also want to thank him for his unconditional support during my whole life.

I thank Csaba for being an exceptional brother. In addition to his willingness to always offer a helping hand, I value his intelligence more and more everyday.

I finally thank my wonderful wife, Anahí: I am afraid that the English language is not rich enough to express all my love and gratitude to her. Thank you for always standing by me during all these years!

Contents

Abstract/Résumé	i
Acknowledgements	iii
Contents	v
1 Introduction	1
1.1 Theories	4
1.1.1 Value in economics	5
1.1.2 Rational bubbles	5
1.1.3 Heterogenous beliefs bubbles	7
1.1.4 Behavioral bubbles	8
1.2 Approach of this thesis	10
2 Bubbles as the difference between fundamental and market value	12
2.1 The Zynga Mania	12
2.1.1 Introduction	13
2.1.2 Valuation methodology	15
2.1.3 Hard Data	16
2.1.3.1 General approach	16
2.1.3.2 The tails of the DAU decay process	17
2.1.3.3 Innovating process	18
2.1.3.4 Predicting the future DAU of Zynga	21
2.1.4 Soft Data	21
2.1.5 Valuation	23
2.1.6 Historic evolution until April 18 th 2012	24
2.1.7 Arbitraging Zynga's bubble	27
2.1.7.1 End of lock-up and its implication	27
2.1.7.2 Prediction for Zynga	28
2.1.8 Post-mortem analysis, on May 24 th , 2012, of the proposed strategy	31
2.1.9 Outlook. Added on January 23 rd , 2015.	32
2.1.10 Conclusion	33
2.2 Forecasting future oil production	35
2.2.1 Introduction	35
2.2.2 Methodology	36

2.2.2.1	Extending the oil production of individual fields	36
2.2.2.2	Discovery rate of new fields	37
2.2.3	Results	39
2.2.4	Validation	41
2.2.5	Conclusion	41
3	Bubbles as price dynamical processes	45
3.1	The predictive power of LPPL	46
3.1.1	Introduction	46
3.1.2	The log-periodic power law model	47
3.1.3	Conceptual approach	49
3.1.4	The signal	51
3.1.5	Strategies	52
3.1.6	Results	54
3.1.7	Conclusions	56
3.2	The Γ factor	57
3.2.1	Introduction	58
3.2.2	Methodology	61
3.2.2.1	Data	62
3.2.3	Results	62
3.2.3.1	Γ on different time horizons	62
3.2.3.2	Γ vs the standard factors	62
3.2.3.3	Performance by industry	64
3.2.3.4	The Fama-French + Γ factor model	65
3.2.3.5	The correlation between Γ and momentum	65
3.2.3.6	The Γ^s factor	68
3.2.3.7	The Fama-French + Γ^s factor model	68
3.2.4	Discussion	69
4	Conclusion and outlook	74
	Bibliography	78

Chapter 1

Introduction

Financial bubbles have always been a subject of fascination reaching far beyond the financial community. They are arguably as old as the stock market itself. Documented examples go back as far as the 17th century with the so-called Tulip mania in the Netherlands (then, United Province), when tulip bulbs, originating from the Ottoman Empire, were introduced to Europe. The bulbs started to experience a price increase due to growing demand from the aristocratic milieu, but eventually turned into a purely speculative object, where the distinction between buyers and sellers became largely irrelevant. Indeed, many individuals were buying those bulbs for an ever increasing price tag, hoping to resell them even more expensively to a greater fool.

These expectations of future price increase led to the development of a spectacular bubble and ended, due to the unsustainable nature of the process, in a massive crash. While some economists find rational justifications for the price increase in the years preceding the crash

([Garber, 1989](#)), there is

a general consensus about the fact that the price increase in the months preceding the turning point had no rational justification. The bulbs, some of them selling for the equivalent of the yearly revenues of a skilled craftsman, lost more than 95% of their peak value within a period of several months ([Thompson, 2007](#)). Another example of a famous



FIGURE 1.1: 17th century satirical painting by Dutch painter Brueghel the Younger depicting monkeys trading tulip bulbs. Source: Wikimedia Commons

bubble and its subsequent crash happened a century later, in 1720, with the South Sea Company. The South Sea Company was founded as a public-private venture in Britain as a way to reduce the burden of government debt. To that effect, stocks were issued, backed by government bonds, delivering a yearly dividend of 5%. In addition, and probably one of the major causes of the bubble, the South Sea company was given monopoly to trade with South-America, then under Spanish control. The prospect of future profits triggered the self-fulfilling expectation of future price increase that ended with the crash of 1720, the stock losing 80% of its value in a matter of months, after having increased by a factor of 8 over 6 months. It should be noted that the likelihood of any trade taking place between Britain and the Spanish colonies was slim from the beginning, given the geopolitical situation at the time. Nonetheless, the “guaranteed future trade profits” were used to rationalize the spectacular price increase. Closer to our time, we have the example of the stock market crash of October 29, 1929 also called Black Tuesday. On that event, Galbraith (2009) writes that “an increasing number of persons were coming to the conclusion - the conclusion that is the common denominator of all speculative episodes - that they were predestined by luck, an unbeatable system, divine favour, access to inside information, or exceptional financial acumen to become rich without work”. As a consequence, the stock market experienced, in the years preceding the crash, a super-exponential increase in price that was unsustainable, the Dow Jones more than doubling over 2 years. The euphoria ended with Black Tuesday. The stock market lost approximately 40% of its value during the crash and continued its fall until 1933, having lost 90% of its peak value by then. Once again, the crash had been the consequence of unsustainable price dynamics that had started well before Black Tuesday.

Financial bubbles are not limited to the distant past; in fact, a major bubble developed throughout the 1990’s during the Internet Revolution. The introduction of a major new technology, and the potential economic implications associated with it, quickly sparked over-



FIGURE 1.2: Headline of the Brooklyn Daily Eagle on Black Tuesday. Source: www.wikispaces.com

optimism in the stock market participants. As in the previous examples, they soon started to trade internet related stocks in a purely speculative manner, counting on an everlasting price increase to buy low and sell high.

This lasted until the price ceased to increase, and the Dot-Com Bubble burst in mid-March 2000. The super-exponential price increase of the technology companies - the NASDAQ more than decupled between 1990 and 2000 - ended in a major stock market crash, the same index losing more than 75% of its value during the next two years. This small account of some of history's most famous bubbles wouldn't be complete without mentioning the recent housing crash of 2008. This was the proximate cause of the ensuing crisis coined the Global Financial Crisis. Its ultimate cause was rooted in the US housing market bubble (Sornette and Cauwels, 2014) that as a consequence of Cheap credit, low interest rates following the burst of the Dot-Com Bubble, as well as complex financial instruments acted as catalysts for the development of the housing bubble. The housing bubble was once again the product of self-fulfilling expectations of future price increase that led the investors to act in a purely speculative manner, the housing market experiencing an explosive growth in price. The growth not being sustainable, the housing market crashed, bringing down in cascade the highly-leveraged financial sector which in turn lead to a global recession, not unlike the Great Depression.

These different bubbles share some similarities.

An important element is the over-enthusiasm of the market participants leading to explosive price dynamics that due to their unsustainable nature end in crash. This over-enthusiasm is usually

rooted in an innovation seen as a gamechanger in the marketplace (Kindle-

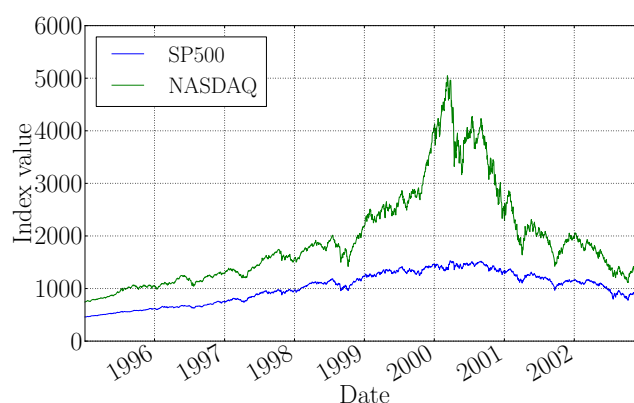


FIGURE 1.3: Dot-Com bubble: the evolution of the NASDAQ and SP500 are shown. The NASDAQ's explosive behavior driven by the technology companies is even more evident in comparison to the evolution of the SP500. Source of the data: Yahoo Finance.

berger and Aliber, 2005): tulips in the case of the Tulip Mania, prospects of trade with the Spanish colonies for the South Sea Company, utilities for the bubble leading to Black Tuesday, the Internet for the Dot-Com bubble and financial innovation for housing bubble and its subsequent crash. Also, in each case, the dominant view was negating the existence of any overpricing, always finding a rationale to justify the price increase. In fact, even bright minds like Isaac Newton suffered substantial losses during stock market crashes; according to the accounts, after investing about £3'500 in the South Sea Company, and making a 100% profit on it during the maturation phase of the bubble, he lost £20'000 in the crash. He then famously declared: "I can calculate the motion of

the heavenly bodies, but not the madness of people.” In 1929, three days before Black Tuesday, Irvine Fisher, one of the most influential economists at the time, declared: “Stock prices have reached what looks like a permanently high plateau.” This inability to see the bubble cost him most of his personal wealth. As a final example, let us cite Alan Greenspan, the chairman of the Federal Reserve who in 2005, shortly before the housing crisis, declared during a speech: “The use of a growing array of derivatives and the related application of more-sophisticated approaches to measuring and managing risk are key factors underpinning the *greater resilience* of our largest financial institutions, which was so evident during the credit cycle of 2001-02 and which seems to have persisted. Derivatives have permitted the *unbundling of financial risks*.” These failures by the public, in general, and the economic authorities, in particular, to see the bubbles, speak about the difficulty to detect them. This is precisely the issue that this thesis will address; *how to detect bubbles?* After giving a short literature review of the theories aiming to describe the mechanisms behind these strange phenomena, the approach taken in this work and its added value to the field will be highlighted.

1.1 Theories

Bubbles are still a controversial subject in classical economics; they imply that there is sizeable and persistent deviation between the fundamental value of an asset and its market value. But the dominant paradigm in economics, the Efficient Market Hypothesis (EMH) (Fama, 1970), states that all information about the fundamentals of an asset are reflected in the market price through the action of the rational market participants. As such, the fundamental and the market value are the same: if there were a difference between the two, there would be an arbitrage opportunity (an opportunity to make a profit without any risk) and that difference would quickly be traded away. An important challenge to this paradigm is the empirical evidence on bubbles, such as the historical examples cited above. And while economists who take the EMH perhaps a little too literally argue that even large bubbles, such as the Tulip Mania, can be explained fundamentally, a growing number of models have been proposed to explain these phenomena. The theories can be classified into three broad categories: rational bubbles, heterogenous belief bubbles and behavioral bubbles. Before giving an overview of the main bubble related literature, we will explain the crucial concept of (fundamental) value in economics. Let us note that the overview of the existing literature doesn't aim to be exhaustive, but rather to give the reader a basic understanding of some of the main theories. What follows has been based in part on three review papers on the topic from Brunnermeier and Oehmke (2012), Kaizoji and Sornette (2008), Scherbina (2013).

1.1.1 Value in economics

The notion of value in economics is based on two fundamental concepts. The first one can be enunciated as follows: \$100 today is worth more than \$100 tomorrow. The intuition is simple: receiving \$100 today and putting it into the bank for a 1% interest rate (or return) is the same as receiving \$101 in a year. It is said that the present value of \$101 one year from now is \$100, given a 1% annual return. Generalizing the concept, the present value of \$C in t years with an annual return of r is:

$$PV = \frac{C}{(1+r)^t} \quad (1.1)$$

where the exponent t comes from compounding. Applying equation 1.1 to value an asset (with infinite maturity) distributing dividends C_t every year is:

$$PV = \sum_{t=0}^{t=\infty} \frac{C_t}{(1+r)^t} \quad (1.2)$$

Equation 1.2 is also called the Discounted Cash Flows equation.

The second fundamental concept relates to the magnitude of r , also called *risk premium*. Economics postulate a positive relationship between risk and return. In other words, a higher risk has to be remunerated by a higher return. r has a lower bound called the risk-free rate (r_f) which is the remuneration that one gets for a riskless investment. It embodies time value and is usually proxied by the return on US T-bills. Determining the value of the additional component of r associated with the riskiness of an asset is the major problem of economics. Theories such as the Capital Asset Pricing Model (CAPM) (Sharpe, 1963) or the 3-factor model (Fama and French, 1992) propose, with more or less success, specific relationships between risk and return.

Although this economic definition of value is very narrow, since it ignores other dimensions of value such as the social or tactical dimensions, it offers a well-defined framework to postulate and test hypotheses.

1.1.2 Rational bubbles

Rational bubbles are probably the dominant approach taken by economists to explain the emergence of bubbles. These models examine the conditions under which bubbles can form given that agents are rational.

The foundation of the rational bubble models is the one proposed by Blanchard and Watson (1983). By decomposing the value of an asset into its fundamental value and a

bubble component,

$$P_t = P_t^{fund} + B_t = E_t \left[\sum_{\tau=t+1}^{\infty} \frac{C_{\tau}}{(1+r)^{\tau-1}} \right] + \lim_{T \rightarrow \infty} E_t \left[\frac{B_T}{(1+r)^{T-t}} \right] \quad (1.3)$$

where the fundamental component is the present value of the discounted future cash-flows (equation 1.2) and the bubble component is the present value at $T \rightarrow \infty$ (the bubble component doesn't deliver dividends). The first consequence of this model is that, for the bubble to exist, it has to grow with rate r . Indeed, if the bubble would grow with growth rate $r_B < r$, the present value of the bubble component would be 0 and as such the price of the asset would be equal to its fundamental value (replace B_T with $B_0(1+r_B)^{T-t}$ in equation 1.3). If on the contrary, $r_B > r$, the bubble component would be infinite and, as a consequence, so would the price. The second implication of the model is that the asset has to be infinitely lived: suppose that the asset has a maturity date T . At that time, the asset would be liquidated for its fundamental value and $B_T = 0$. But if $B_T = 0$, nobody would be willing to buy the asset for more than its fundamental value at $T - 1$, knowing that one time step later, the bubble component would be 0. This backward induction argument implies that a bubble can only exist for infinitely lived assets.

One of the many problems of the model is that, as the bubble component of the price grows exponentially, the price-to-cash-flow ratio becomes infinite, i.e. $\lim_{T \rightarrow \infty} \frac{P_T}{C_T} = \infty$, which is unrealistic. To remedy this problem, [Froot and Obstfeld \(1989\)](#) proposed to make the bubble component depend on the cash flows rather than on time. This choice was motivated by the observation that investors are bad at predicting future cash flows. By doing so, the authors show that under the hypothesis of no bubbles, $\frac{P_T}{C_T} = k$, k being a constant. Applying this criteria on the SP500 over the 1900-1988 period, they reject the hypothesis of no bubbles.

By relaxing the assumption about common knowledge, i.e. the fact that everybody knows that everybody knows, and limiting short-selling, [Allen et al. \(1993\)](#) show that bubbles can form on finitely lived assets. The intuition behind this result is that the absence of common knowledge eliminates the backward induction argument since a rational agent can hope to resell the asset to another one who might not know. Furthermore, constraining short-selling limits the ability of the agents to learn other agents' private information from market prices.

Another class of models departs from the assumption that all agents are rational and achieves bubbles by introducing a second class of traders that are behavioral feedback traders. It is the interaction between the rational arbitrageurs and the behavioral traders

that creates the bubbles. [DeLong et al. \(1990\)](#) show that, in their setup, rational arbitrageurs buy the asset after a good news in order to bait the behavioral feedback traders into pushing the price further up, allowing the rational agents to sell their shares profitably at the expense of the behavioral ones.

In [Abreu and Brunnermeier \(2003\)](#), all the rational agents know that there is a bubble, but it is the difference of opinions about the timing of the start of the bubble, or in other words their estimate of the asset's fundamental value that leads to a synchronization problem, each agent not being able to burst the bubble on its own. As such, rational arbitrageurs ride the bubble until they can synchronize, due to a piece of exogenous information for example. These models offer a powerful argument against the Efficient Market Hypothesis which claims that even if there are irrational agents in the market, rational arbitrageurs will prevent any possibility of mispricing.

[Lin and Sornette \(2011\)](#) on the other hand argue that it is the difference of opinions about the timing of the end of the bubble leading to the persistence of bubbles. They support their claim by applying their model on real data and derive an operational procedure that allows them to diagnose bubbles and forecast their termination.

1.1.3 Heterogenous beliefs bubbles

Heterogenous belief bubbles take place in a setup where agents don't agree on the fundamental value of the asset. This can be due to psychological biases, or simply to the difficulty to make predictions in an uncertain future. [Miller \(1977\)](#) shows in a very simple framework that given a limitation on short selling, the divergence in opinion between the agents about the asset's return lead to an equilibrium where the price is higher than the average estimate. Put simply, the optimists push the price higher than the average estimate because the pessimists stay out of the market, not being able to fully reflect their opinion by shorting the asset. Moreover, the author shows that the price of the asset increases with the diversity of the opinions about its future returns.

In a dynamic framework, [Harrison and Kreps \(1978\)](#) show that not only can bubble arise when agents have different opinions about the fundamentals of an asset, but the price of the asset can even surpass the valuation of the most optimistic agent. This happens because the optimistic agent choses to pay a premium to buy the asset in the hope of reselling it later when he will be pessimistic (and other agents will be optimistic). We should note that short-selling is also restricted in this model.

[Scheinkman and Xiong \(2003\)](#) build on [Harrison and Kreps \(1978\)](#) model and extend it into continuous time. They interestingly conclude that bubbles are characterized by higher trading volumes, a fact that can be observed empirically.

1.1.4 Behavioral bubbles

Some financial economists agree that psychological biases must play a fundamental role in the formation of bubbles, departing completely from the claim that agents are rational, and frontally attacking the Efficient Market Hypothesis. [Shiller \(2002\)](#) cites several behavioral mechanisms to be at the origin of bubbles. Among the most relevant ones are positive feedback loops between price and investor's enthusiasm, as well as herding, i.e. the fact that people tend to imitate each other. In light of the historical bubbles and crashes, some of which were described in the introduction, these mechanisms seem very convincing in generating bubbles. It is also worth mentioning the work of [Hens and Schenk-Hoppé \(2004\)](#) who show in an evolutionary framework, where strategies implemented by heterogenous agents with different opinions and behaviors compete for market capital, that seemingly irrational strategies can outperform seemingly rational ones.

Models coming from physics, in particular the Ising model, have been very successful at describing how imitation between agents can lead to bifurcation in their aggregate opinions (see [Sornette \(2014\)](#) for a complete review). The Ising model consists of agents influencing each other. In a nutshell, if an agent is surrounded by agents willing to sell, it is likely that he will start selling as well. Two opposite forces are at play; the ordering force of social imitation and the disordering force of idiosyncratic noise. In an Ising-like framework, [Orléan \(1989\)](#) shows that Bayesian opinion lead to two qualitatively different dynamics. When the agents' estimates of the group's opinion, i.e. the estimation of the fraction of agents that would buy, are heterogenous, the stationary distribution of opinions is peaked around 50%: half of the agents would buy and the other half would sell. However, when the same estimate is homogenous, the interaction among agents leads to a stationary distribution with two peaks: most agents would buy and few would sell, or most agents would sell and few would buy. These situations can be interpreted as bubble and crash states respectively.

[Lux \(1995\)](#) sets up a framework with two kind of agents: speculative traders and fundamentalists. Speculative traders form their decisions based on the opinion of other traders of their kind, as well as on price dynamics (momentum). Fundamentalists buy or sell based on the difference between the asset's market and fundamental values. The interaction between the Ising-like speculative traders and fundamentalists gives rise to

a rich phenomenology of price dynamics, with prices moving around an equilibrium as well as boom and bust cycles, depending on the different parameters. This shows that simple mechanisms are enough to generate a variety of dynamical regimes.

Contrary to the previous works, where every agent was interacting with every other, [Cont and Bouchaud \(2000\)](#) impose a random graph topology on how the agents interact. A random graph is a network where an agent has a probability of $\frac{c}{N}$ to be connected to another agent, N being the total number of agents and c the average number of connections of per agent. In their setup, the authors propose that every agent of a component, i.e. the set of agents connected through a path, would take the same action of buying, selling or staying out of the market. The components could be thought of as organizations such as hedge funds, individual traders etc. . . The main result of the paper is that for $0 < c < 1$, the fat-tails of the distribution of returns, a well-known stylized fact, can be recovered. $c = 1$ corresponds to a critical value where a giant component emerges, encompassing a finite fraction of the system. This can be interpreted as a bubble or a crash, since the giant component contains agents with the same action.

Although all of the models described so far offer an explanation as to what mechanisms could be at the origin of the formation of bubbles, they suffer from a major limitation: either they cannot be calibrated to real data and as such are not testable, or they lack any predictive power. One major model that distinguishes itself by proposing a functional form for the price dynamics up to the crash is the Log-Periodic Power Law model and was first proposed by [Sornette et al. \(1996\)](#) and was later formalized by [Johansen et al. \(2000\)](#). The model is based on the behavioral and well-documented phenomenon of positive feedback between demand and price, i.e. the fact that an increase in price leads to an increased demand in speculative periods which itself leads to an increase in price. The historical cases cited at the beginning of the introduction are good examples. Translating the positive feedback mechanism into mathematical terms, one can show that the price dynamics obey a super-exponential law with a finite time singularity, beyond which the solution doesn't exist. This finite-time singularity can be interpreted as the time of the crash. In addition, imposing a hierarchical topology on the networks of traders (much more realistic than the random graph) leads to the super-exponential price dynamics being decorated by log-periodic oscillations of decreasing amplitude. Other explanations for the log-periodic oscillations include the competition between non-linear trend followers and non-linear value investors [Ide and Sornette \(2002\)](#).

1.2 Approach of this thesis

As mentioned before, the main objective of this thesis is the detection of bubbles in financial markets. Many approaches were possible as suggested by the short literature review. The way bubbles have been quantified in this work revolves around two fundamentally different approaches which were the guidelines of our research. These two approaches correspond, in a sense, to both ends of the bubbles literature spectrum and are quite complementary.

The first approach consisted in quantifying the bubbles as the difference between the market price and the fundamental value of an asset. The predictive nature of this difference on the dynamics of prices was then investigated. Even though this may seem like a simplistic method to tackle the problem, the real difficulty lay in the determination of the fundamental value. Defining bubbles in this way has an important upside; the absence of a need to postulate any mechanism to explain the bubble formation process. The first research topic to fit this paradigm was the valuation of Zynga, a social networking company making games on Facebook. The choice of a social networking company was intentional because of the simplicity of its business model: social networking companies derive the vast majority of their revenues from their user base (through advertising). As such, if one is able to model the dynamics of users and how the revenues scale with users, it is possible to give a good forecast of the future cash flows of the company and discount them to get the fundamental value (equation 1.2). This would be much more difficult with companies having various comparable sources of revenues. The second research topic, seemingly unrelated to financial bubbles, was the forecasting of future oil production. The methodology for forecasting future oil production turned out to be strikingly similar to forecasting Zynga's future user base. On a more conceptual level, the similarity with Zynga lies in the fact that the difference between our forecast of future oil production (validated through backtesting) and the standard estimates may be informative of long time scale oil price dynamics.

The second approach consisted in quantifying the bubble state of an asset through its price dynamics, without resorting to the determination of a fundamental value. In our first attempt, we studied the predictive power of the Log-Periodic Power Law model, based on the mechanism of herding among agents. We tested the power of the model by using trading strategies as the analogs of a market spectrometer: by comparing the outcome of trading strategies based on the log-periodic power law model (LPPL) with random strategies, we were able to show the robustness of LPPL's predictive power. The study was performed on a large-scale to underline the significance of our results. Our second study went on to generalize the importance of super-exponential price dynamics on the pricing of assets. This was done by looking at the explanatory power of the first

difference of returns, the simplest form encapsulating the idea of a super-exponential price increase. Indeed, the first difference (or change) of returns, i.e. the growth of the growth rate itself, is a concept at the core of the super-exponential price dynamics, one of the fundamental signatures of LPPL. We studied its impact on the return of assets within the framework of factor regression models, a standard tool in financial economics. Our findings reinforce the notion that amplification mechanisms that take the form of price-to-price positive feedback loops are a crucial component of asset price dynamics.

The structure of the thesis is be straightforward; these two different approaches are laid out in the following chapters and exemplified by the four research topics mentioned. A conclusion will tie the different research results together and an outlook will be given.

Chapter 2

Bubbles as the difference between fundamental and market value

In this chapter, we investigate how bubbles, defined as persistent deviations from fundamental value, impact price dynamics through Zynga’s example. We then show how the methodology developed for Zynga can be extended to the forecasting of future oil production.

2.1 The Zynga Mania

¹ This section is based on the paper co-authored with Peter Cauwels and Didier Sornette: “When games meet reality: is Zynga overvalued?” (Forró et al., 2012b). The context in which this research was performed is quite interesting: On December 16th 2011, after some prominent initial public offerings (IPO) such as those of Groupon, LinkedIn and Pandora, Zynga itself went public on the NASDAQ. Entering the market with a \$7 billion market capitalization, Zynga became one of the biggest web-IPOs since Google. Around the same time, the vast majority of the estimates surrounding Zynga seemed excessive, as exemplified by JP Morgan Chase who was issuing a price target of \$15 per share, corresponding to a valuation north of \$10 billion. But more importantly, no rigorous methodology was presented to back up those claims. This motivated us to develop a transparent methodology based on a two-tiered approach, giving an upper bound to Zynga’s valuation. The first tier consisted in introducing a new model to forecast Zynga’s user base, based on the individual dynamics of its major games. The second modeled its revenues per user using a logistic function, a standard model for

1. This section is based on the paper co-authored with Peter Cauwels and Didier Sornette: “When games meet reality: is Zynga overvalued?”

growth in competition. Combining these two elements, we were able to bracket Zynga's upper value using three different scenarios: 4.2, 5.2 and 7 billion USD in the base case, high growth and extreme growth scenario respectively. We published these conclusions in a first paper [ref] on December 27th, 2011, less than two weeks after the IPO. We then consolidated our methodology and updated our computations taking into account the new financial data point of the 4th quarter (December 31st) of 2011 that became available beginning of 2012. This resulted in an important decrease of the uncertainty in our valuation with our new computation yielding 3.4, 4.0 and 4.8 billion USD in the three scenarios. These results were published on April 2nd and led to our diagnosis of Zynga being in a bubble. In addition, on April 18nd we formulated a short-term trading strategy to take advantage of Zynga's overvaluation. The future proved us right...

2.1.1 Introduction

The pricing of IPOs, and companies in general, has been extensively studied. [Ibbotson and Ritter \(1995\)](#), [Ritter and Welch \(2002\)](#), among others, reviewed well-known stylized facts when companies go public. We can cite the underpricing of new issues, ie, the fact that underwriters often underprice the IPO leading to high returns on the first day of trading, or the long-term underperformance of the underpriced IPOs compared to their "fairly" priced counterparts. During the dot-com bubble, the rapid rise of the internet sector contrasting with the modest growth of the "old economy" raised a lot of interest. [Bartov et al. \(2002\)](#) showed that there were differences in the valuation of internet and non-internet firms. Notably, for the latter, profits were rewarded (positively correlated with the share value) and losses were not (as is usually the case). However, the reverse was true for Internet companies, where losses were rewarded and profits were not. This somewhat paradoxical situation arose from the perception that losses were not the result of poor company management but rather investments that would later pay off. [Demers and Lev \(2001\)](#), [Hand \(2001\)](#) further showed that web-traffic was an important factor in the market value of Internet companies. Indeed, in the case of web-traffic intensive companies, while losses were being rewarded before the peak of the bubble and profits were not, the situation reversed after the peak of the bubble: profits became rewarded and losses were not anymore. This phenomenon was not observed for internet companies without web-traffic.

We should notice that the studies, mentioned so far, tried to explain the market price of companies using different explanatory variables, such as revenues, type of company, amount of web-traffic, difference between IPO price and first day closing price and so on, making the implicit assumption that the market is efficient and reflects the intrinsic value of the company. While in the long-run this may be a good approximation, this is

not true for shorter time-scales, during a bubble typically. As such, these methods, often based on linear relationships between the market price and the explanatory variables, are not meant to reveal the fundamental value of a company or make long-term predictions.

[Ofek and Richardson \(2002\)](#) tackled the problem from a different angle. They assumed that, in the long-run, the price-to-earnings ratio of internet companies would converge to their “old economy” counterparts, and computed the growth in earnings necessary to achieve that. They found unrealistic growth rates making an argument against market rationality. [Schwartz and Moon \(2000\)](#) used a real-option approach to value Amazon, the company having the option to go bankrupt (thus limiting their losses). Their model relied on the future growth rate of revenues combined with the discounted cash flows method. The upside of coming up with a valuation for the company was somewhat mitigated by the important sensitivity of the valuation to variations in the model parameters. [Gupta et al. \(2004\)](#) extended a methodology developed by [Kim et al. \(1995\)](#) to value Amazon, Ameritrade, E-bay and E*Trade (internet companies). Their model used the discounted cash flows analysis where the future revenues are computed based on the prediction of the company’s user base combined with an estimation of the revenues generated by each user in time. They obtained robust valuations of the internet companies, allowing for a quantitative assessment of the discrepancy between the market capitalization of the companies and their fundamental values. Adopting a similar approach, [Cauwels and Sornette \(2012\)](#) showed that Facebook and Groupon were overvalued. The main insight of the aforementioned works was to recognize that, for companies deriving their value directly from their users, such a simple approach could give much better estimate of the intrinsic value of a company than methods employed so far.

The present work adds to the existing literature by extending the methodology of [Cauwels and Sornette \(2012\)](#) to Zynga, a company where the user dynamics are very different and more complicated from the ones observed in Facebook, Groupon, Amazon and so on. Indeed, the evolution of Zynga’s user base is a result of the individual dynamics of each of its individual games and couldn’t therefore be modeled by a single function. Moreover, we found that the revenues per user had entered a saturation phase. This limited Zynga’s ability to increase its revenues much further, as their user base was already in a quasi-stationary phase. Finally, we found that Zynga had been greatly overvalued since its IPO. This analysis, however, gave no indication of the short-term price dynamics, our model making no assumption about the mechanisms responsible for the overvaluation. But by combining our fundamental analysis with a financial event, in the form of the end of lock-up period, we were able to forecast Zynga’s short-term price trajectory.

Some precision about how this work is organized is in order since different parts of it have been performed/published at different times. In addition to laying out the content of the different sections, we believe that specifying their timeline is particularly enlightening in the context of a work dealing with prediction. Section 2.1.2 gives a brief summary of the methodology used to value Zynga. Section 2.1.3 describes the dynamics of the number of its daily active users (DAU). Section 2.1.4 analyzes the financial data relevant to the valuation of the company. Section 2.1.5 gives its estimated market capitalization. All these previous sections are based on the results obtained on April 2nd, 2012. Section 2.1.6 analyzes the evolution of Zynga until April 18th, 2012, in the light of its valuation. Section 2.1.7 discusses possible strategies to arbitrage the over-valued stock of Zynga and was also added on April 18th. Section 2.1.8 gives a post-mortem analysis of our proposed strategy and was performed on May 25th, 2012. Section 2.1.9 analyzes the evolution of Zynga from its post-mortem analysis until the time of writing of this thesis, to see how our prediction fared on a longer timescale. Finally, section 4 concludes.

2.1.2 Valuation methodology

The major part of the revenues of a social networking company is inherently linked to its user base. The more users it has, the more income it can generate through advertising. From this premise, the basic idea of the method proposed by [Cauwels and Sornette \(2012\)](#) was to separate the problem into 3 parts:

1. First, we needed to forecast Zynga's user base. This is what we call the part of the analysis based on hard data and modeling. Because Zynga uses Facebook as a platform, and one does not need to register to have access to its games (a facebook account is sufficient), there is no such thing as a measure of total registered users (U). These registered users were used by [Cauwels and Sornette \(2012\)](#) in their valuation of Facebook and Groupon. Because it takes an effort to unregister, the number of registered users is an almost monotonically increasing quantity. As such, [Cauwels and Sornette \(2012\)](#) were able to model and forecast Facebook and Groupon's user dynamics with a logistic growth model (equation 2.1).

$$\frac{dU}{dt} = gU\left(1 - \frac{U}{K}\right) \quad (2.1)$$

Here, g is the constant growth rate and K is the carrying capacity (this is the biggest possible number of users). This is a standard model for growth in competition. When $U \ll K$, U grows exponentially since $\frac{dU}{dt} \approx gU$ (this is the unlimited growth paradigm) until reaching saturation when $U = K$ (and $\frac{dU}{dt} = 0$). This model is a good description of what happens in most social networks: the number

of users starts growing exponentially and eventually saturates because of competition/constrained environment. For Zynga, a different approach had to be worked out. Here, the analysis was based on the number of Daily Active Users (DAU), a more dynamical measure. DAU can fluctuate (up and down) and as such couldn't be modeled with a logistic function. Moreover, Zynga's users formed an aggregate of over 60 different games at the time the first version of the paper was published (more than 120 games by the end of 2014). Therefore, to understand the dynamics of Zynga's user base, we had to examine the user dynamics of its individual games. Figure 2.1 gives the total number of Zynga users and the DAU of two of its most popular games. We decided to model each of its top 20 games individually, this approach accounting for more than 98% of the recent total number of Zynga users. The specifics of this analysis are further elaborated in section 2.1.3.

2. The second part of the methodology was based on what we consider as "soft" data: this part used the financial data available in the S1/A Filing to the SEC (2011). These were used to estimate the revenues that are generated per daily active user in a certain time period. It also revealed information on the profitability of the company. Due to the limited amount of published financial information, we had to rely on our intuition and good-sense to give our best estimate of the future revenues per user generated by the company. This is why we call this part the soft data part. It is further elaborated in section 2.1.4.
3. The third part combined the two previous parts to value the company. With an estimate of the future daily number of users (DAU) and of the revenues each of them would generate (r), it was possible to compute the future revenues of the company. These were converted into profits using a best-estimate profit margin (p) and were discounted using an appropriate risk-adjusted return (d). The net present value of the company (PV) was then the sum of the discounted future profits (or cash flows), given by equation 2.2 (which is simply a reformulation of equation 1.2).

$$PV = \sum_{t=0}^{t=\infty} \frac{p \cdot r(t) \cdot DAU(t)}{(1+d)^t} \quad (2.2)$$

1.2. We hereby optimistically assumed that all profits were distributed to the shareholders.

2.1.3 Hard Data

2.1.3.1 General approach

We have taken the following steps to forecast Zynga's DAU:

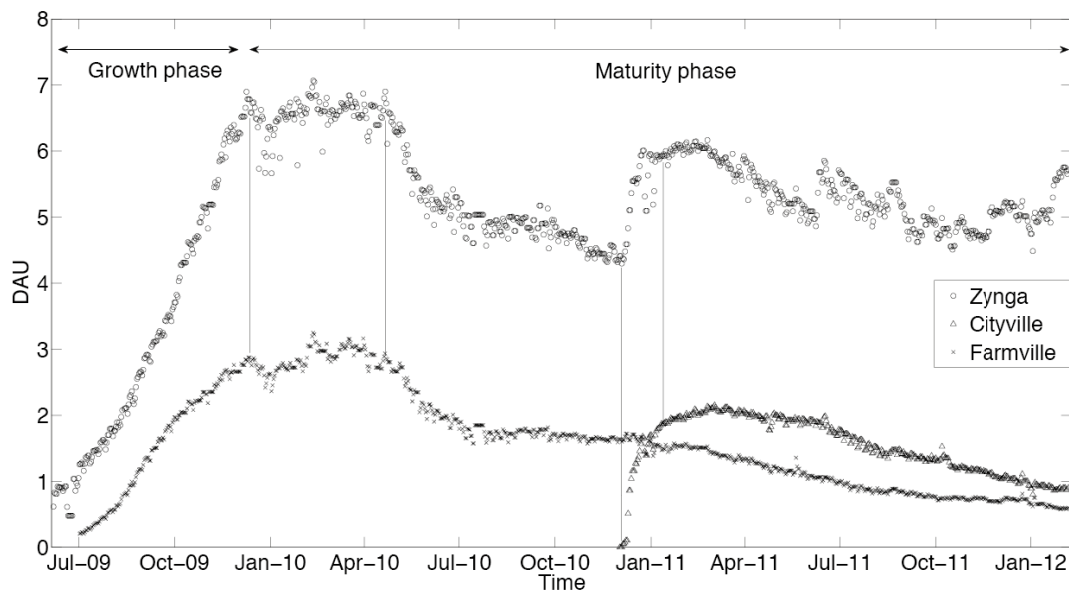


FIGURE 2.1: The number of DAU as a function of time for Zynga and two of its most popular games, Cityville and Farmville. After an initial growth period, Zynga entered a quasi-stable maturity phase since January 2010. A typical feature of the games can be seen in Cityville and Farmville: after an initial rapid rise, the DAU of the games entered a slower decay phase. The black vertical lines show that the total DAU of Zynga depends strongly on the performance of the underlying games. Notice that, even though Zynga exists since mid-2007, we do not have DAU data since its beginning. (Source of the data: <http://www.appdata.com/devs/10-zynga>)

1. We have used a functional form for the DAU of each of the top 20 games to forecast the future DAU evolution of the company. This was done as follows:
 - The data that were available (until December 2011) were used as it is.
 - The user dynamics were extended into the future by extrapolating the DAU-decay process with an appropriate tail function.
2. Because Zynga relies on the creation of new games in order to increase or even maintain its user base, it was important to quantify its rate of innovation. This was done by using $p(\Delta t)$ the probability distribution of the time between the implementation of 2 consecutive new games (restricted to the top 20).
3. Finally, a future scenario was simulated as follows: for the next 20 years, each Δt days, Δt being a random variable taken from $p(\Delta t)$, a game was randomly chosen from our pool of top 20 games. The DAU of Zynga over time was then simply the sum of the simulated games. A thousand different scenarios were computed.

2.1.3.2 The tails of the DAU decay process

The functional form of the DAU of each game was composed of the actual observed data and a tail that simulated the future decay process. We used a power law, $f(t) \propto t^{-\gamma}$,

for that purpose. This resulted in a slow decay process and as such didn't give rise to any unnecessary devaluation of the company by underestimating its future user base. Figure 2.2 shows the power law fits (left) and the extension of the user dynamics into the future (right) for the games Farmville and Mafia Wars.

Such power law was a reasonable prior, given the large evidence of such time dependence in many human activities (Sornette, 2006), which includes the rate of book sales (Deschâtres and Sornette, 2005, Sornette et al., 2004), the dynamics of video views on YouTube (Crane and Sornette, 2008), the dynamics of visitations of major news portal (Dezsö et al., 2006), the decay of popularity of internet blogs posts (Leskovec et al., 2007), the rate of donations following the tsunami that occurred on December 26, 2004 (Crane et al., 2010) and so on.

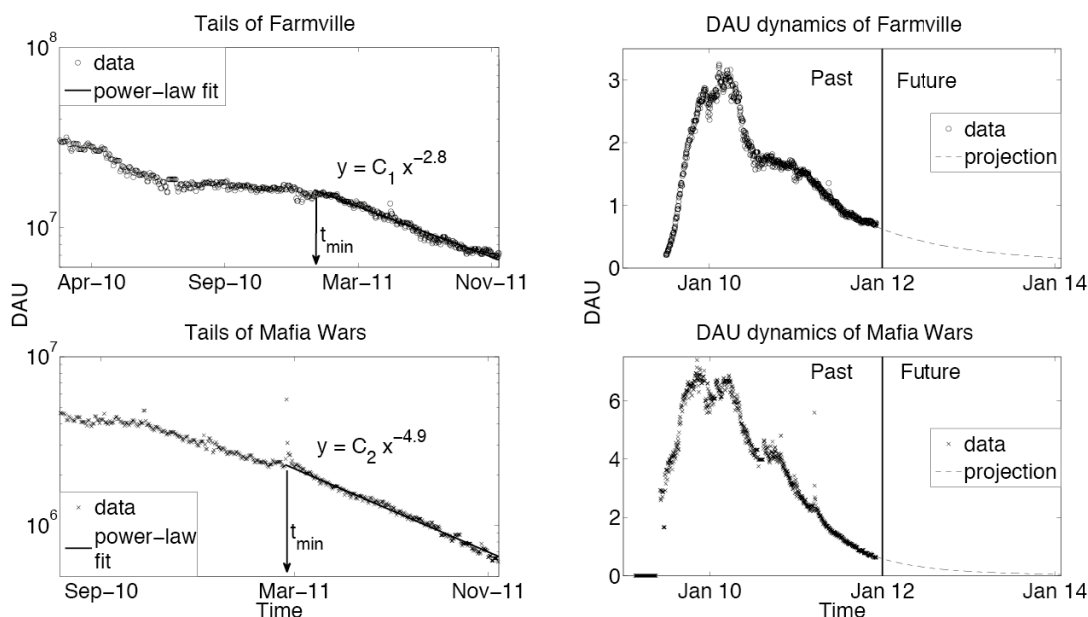


FIGURE 2.2: Left: Decay of Farmville (circles) and Mafia Wars (crosses), 2 representative games out of Zynga's top 20. It can be seen that a power law was a good fit for the tails from t_{min} onwards. Right: Simulated dynamics of the games based on the power law parameters of the left panel. (Source of the data: <http://www.appdata.com/devs/10-zynga>)

2.1.3.3 Innovating process

To be able to realistically simulate Zynga's rate of innovation, it was important to understand the generating process underlying the creation of new games. The simplest process that could be used for that purpose was the Poisson process. To understand its meaning, consider the Bernoulli process, its discrete counterpart. It has a very intuitive meaning and can be thought of as follows: at each time step, a game is introduced with a probability of p (and no game is introduced with a probability of $1 - p$). For a large

enough number of time steps and a small enough p , the Bernoulli process converges to the Poisson process. The Poisson process has 3 important properties:

1. It has a constant innovation rate.
2. It has independent inter-event durations.
3. The inter-event durations have an exponential distribution: $p(\Delta t) = e^{-\lambda\Delta t}$.

To assess whether this was a suitable process to model the innovation rate, we measured the time between the introduction of two consecutive new games, $\Delta t_{(1,2)}, \Delta t_{(2,3)}, \dots, \Delta t_{(n-1,n)}$, and tested for the above mentioned 3 properties.

To test whether the innovation rate was constant, different approaches could be adopted. One possibility was to test for the stationarity of the DAU of Zynga since it entered its maturation phase. Stationarity in the number of users would imply a constant innovation rate. Indeed, figure 2.1 suggests that the user dynamics of Zynga were stationary, its number of DAU being comprised between 43 and 70 million user since the end of the growth phase. However, due to the short time-span of the data, it was hard to implement rigorous statistical tests such as unit-root tests. Instead, we adopted a different approach. If the rate of creation of new games was constant, then the number of new games created as a function of time had to lie around a straight line with slope λ , the intensity of the Poisson process. Figure 2.3 shows the counting of new games as a function of time.

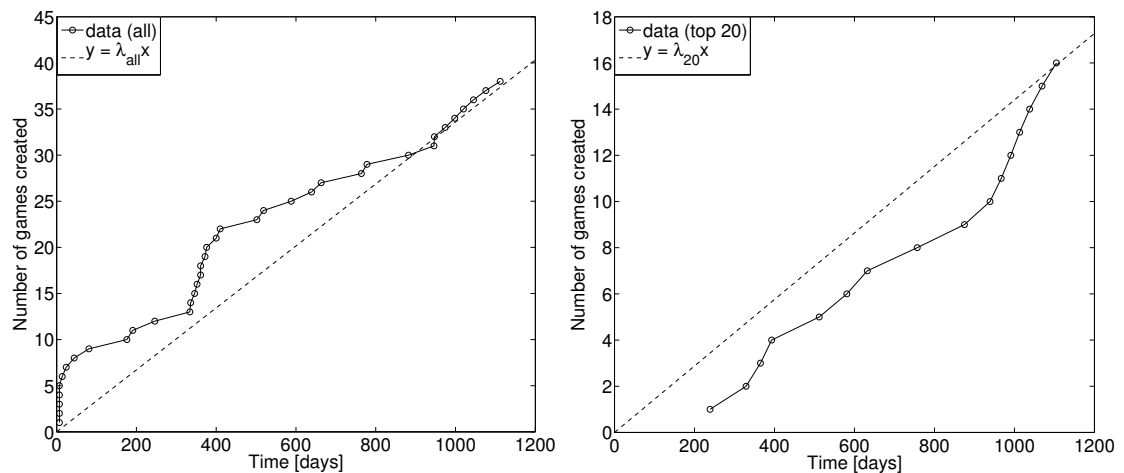


FIGURE 2.3: Left: number of newly created games as a function of time for all the games. The empirical innovation rate was at most equal to the theoretical rate coming from the Poisson process (dashed line). The parameter λ_{all} was obtained by using maximum likelihood (assuming a Poisson process). Right: number of newly created games as a function of time for the top 20. The empirical innovation rate seemed to be higher for the last games than λ_{20} , the theoretical rate from the Poisson process. This was most likely due to insufficient statistics, this phenomenon being absent when all games were taken into account.

As we can see from figure 2.3, the constant innovation rate was a good approximation. Our main concern was to discard the possibility of an important increase in the frequency

of creation of new games towards the end of the time period, which would have led to an underestimation of the number of new games created in the future and hence of the future number of users. When all games were taken into account, the innovation rate was at most equal to the one coming from the Poisson process. As such, the Poisson process with constant intensity would not have led to an underestimation of the value of the company.

To test for the independence of the measured Δt , we study the autocorrelation function $C(\tau)$, the correlation between $\Delta t_{(i,i+1)}$ and $\Delta t_{(i+\tau,i+\tau+1)}$. The result is given in figure 2.4.

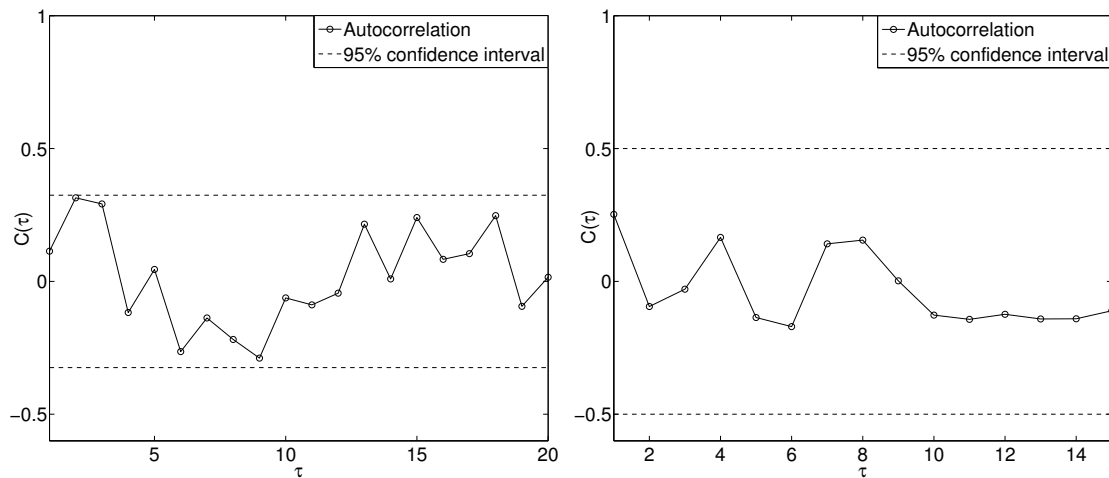


FIGURE 2.4: Left: Autocorrelation function for the inter-event times of all the games. The confidence interval (CI) indicates the critical correlation needed to reject the hypothesis that the inter-event times are independent. It was computed as $CI_{.95} = \pm \frac{2}{\sqrt{N}}$, N being the sample size (Chatfield, 2003). Right: Autocorrelation function for the inter-event time of the top 20 games. In both cases, the independence hypothesis couldn't be rejected.

As can be seen, $C(\tau) = 0$ was within the confidence interval for $\tau > 0$ in both cases, so the independence hypothesis couldn't be rejected.

To test for the distribution of inter-event times, we used a Q-Q plot. The Q-Q plot is a graphical method for comparing two distributions by plotting their quantiles against each other. If the obtained pattern lies on a straight line, the distributions are equal. In our case, the two distributions to be compared were the empirical one (from data) and the theoretical one, an exponential with parameters obtained from a maximum likelihood fit to the data. The results of this analysis are presented in figure 2.5.

As we can see, in both cases the Q-Q plots show a reasonable agreement between the empirical and the exponential distribution given the number of data points, so that the exponential distribution for the inter-event times (Δt) couldn't be rejected. The innovation process was thus modeled as a Poisson process.

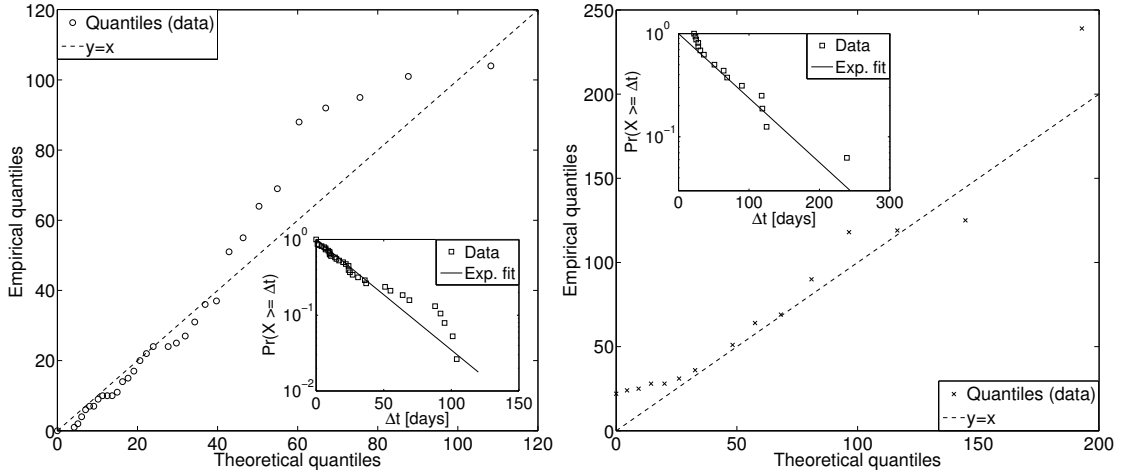


FIGURE 2.5: Left: Q-Q plot of the distribution of time intervals Δt between the introduction of new games, for all the games. The theoretical (assuming a Poisson process) and data quantiles (circles) agree well as can be seen from the proximity to the dashed black line ($y=x$). The subpanel shows the CDF of the data (squares) and the exponential fit (black line) on which the quantiles were built. Right: Q-Q plot of Δt for the top 20 games. We can see a deviation for small values of Δt between exponential theoretical and data quantiles. This could be attributed to insufficient statistics.

2.1.3.4 Predicting the future DAU of Zynga

Starting from the present (April 2012) up to the next 20 years, a top 20 game was randomly sampled each Δt days, with Δt drawn from its theoretical exponential distribution $p(\Delta t)$. For each of these sampled games, the DAU was calculated using its functional form. Summing the DAU of all these games, the user's dynamics of Zynga were computed. This process was repeated a thousand times, giving a thousand different scenarios. As can be seen from figure 2.6, the evolution of the user base between scenarios could be quite different. That was the reason why a wide range of scenarios were needed. The valuation of the company was then computed for each of those scenarios (using equation 2.2). This allowed us to give a probabilistic forecast of the market capitalization of Zynga.

2.1.4 Soft Data

The next step in order to calculate Zynga's value was to estimate the revenues per DAU per year. We based our analysis on the S1/A Form (2011) complemented with the 8-K Form of Filings to the SEC (2012), to add the last quarter of 2011 results. The yearly revenues were then computed each quarter as a running sum over the four previous quarterly revenues:

$$R_i = R_{i-3}^q + R_{i-2}^q + R_{i-1}^q + R_i^q. \quad (2.3)$$

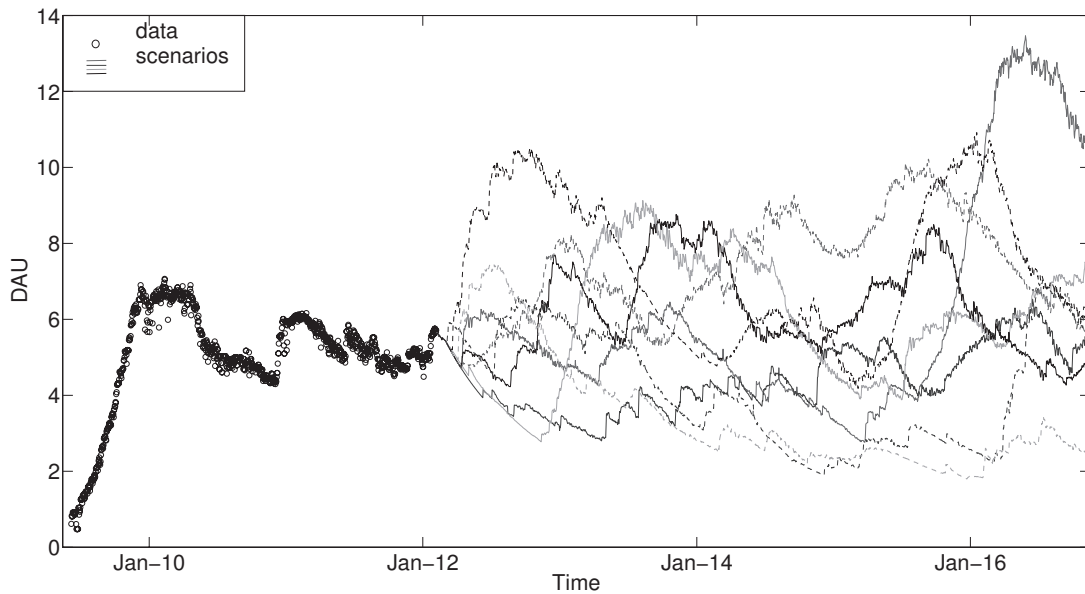


FIGURE 2.6: 8 different scenarios of the DAU evolution of Zynga for the next 4 years. The company was to be valued for each of these scenarios (section 2.1.5). This was to give a range for its expected market capitalization.

Here, R_i and R_i^q are respectively the yearly and quarterly revenues at quarter i with $i \in (4, \text{last})$. The yearly revenues per DAU at each quarter, r_i , were then obtained by dividing R_i by $\langle DAU_i \rangle_{year}$, the realized DAU at time i averaged over the preceding year. Figure 2.7 gives the historical evolution of the revenues per DAU. Initially, this followed an exponential growth process. However, as can be clearly seen in the right panel, this growth entered a saturating phase and the process followed the trajectory of a logistic function. This implies that the revenues per DAU were reaching a ceiling. As such, different logistic functions were fit to the dataset. Each of these corresponded to a different scenario: a base case, a high growth and an extreme growth scenario (as defined in [Cauwels and Sornette \(2012\)](#)). They can be seen on the left panel.

From the IPO of Zynga up to April of 2011, both the exponential and the logistic fits performed similarly. Indeed, when $r_i \ll K$, when the revenues per user were still far away from saturation, the logistic function could be approximated by an exponential. April 2011 was a turning point in the sense that the growth of the revenues per user slowed down, hence the deviation from the exponential (growth at constant rate). This saturation in r_i is easy to explain: there have to be constraints on how much money can be extracted from a user. Under spatial constraints (there is a limited number of advertisements that can be displayed on a webpage), time constraints (there are only so many advertisements that can be shown per day) and ultimately the economic constraints (there is only so much money a user can spend on games or an advertiser is willing to spend), the revenues per DAU were bound to saturate. Using this logistic

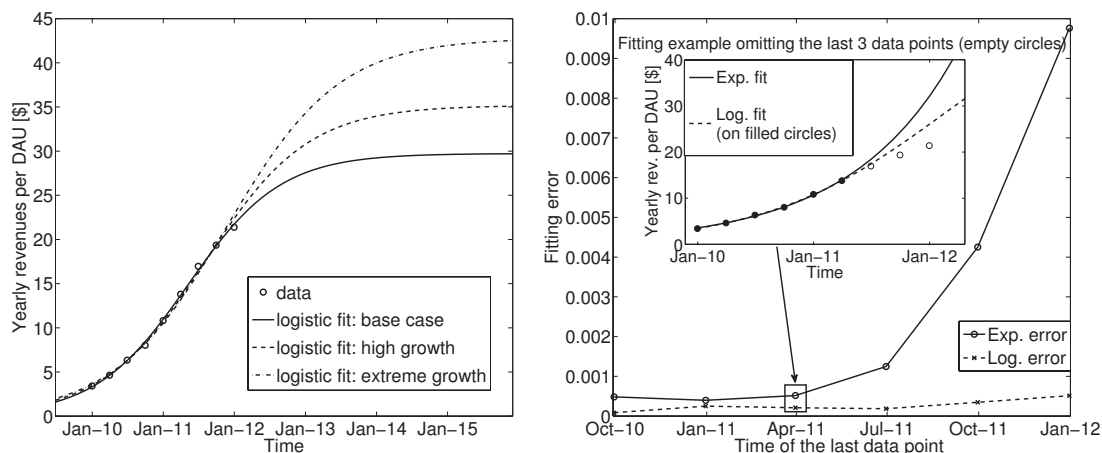


FIGURE 2.7: Left: Yearly revenue per DAU over time. A logistic fit (equation 2.4) was proposed with $K \approx \$30, \35 and $\$43$ for the base case, high growth and extreme growth scenarios. Right: Fitting error of the exponential vs logistic function. Each point was obtained by performing the logistic and exponential regressions on data taking more and more data points into account, starting from a minimum of 4 (as shown in the subpanel where only filled circles were fitted). We can see that the logistic started performing significantly better than the exponential from July 2011. (Source of the data: S1/A and 8-K Forms of the Filings to the SEC of Zynga, 2012.)

description for the revenues per DAU, a valuation of Zynga could be given for each of the 3 growth hypotheses.

2.1.5 Valuation

Combining, through equation 2.2, the hard part of the analysis with the soft part, ie, the number of users over time and the revenues each of them generates per year, the value of the company could be calculated.

We used a profit margin of 15%. This was Zynga's profit margin of fiscal year 2010. As can be seen from table 2.1, this was an optimistic assumption since it was the highest profit margin until 2011 (and even 2015), 2010 being the only profitable year of Zynga so far. We also assumed that all profits would be distributed to the shareholders and used a discount factor of 5% as in Cauwels and Sornette (2012). We computed the company's valuation for all 1000 different scenarios using equation 2.2. The results are shown in figure 2.8 and table 2.2.

Year	2008	2009	2010	2011
Revenue (millions USD)	19.41	121.47	597.46	1065.65
Net income (millions USD)	-22.12	-52.82	90.60	-404.32
Profit margin	-114%	-43%	15%	-38%

TABLE 2.1: Revenue, net income and profit margin of Zynga. (Source of data: S1/A and 8-K forms of the filings to the SEC.)

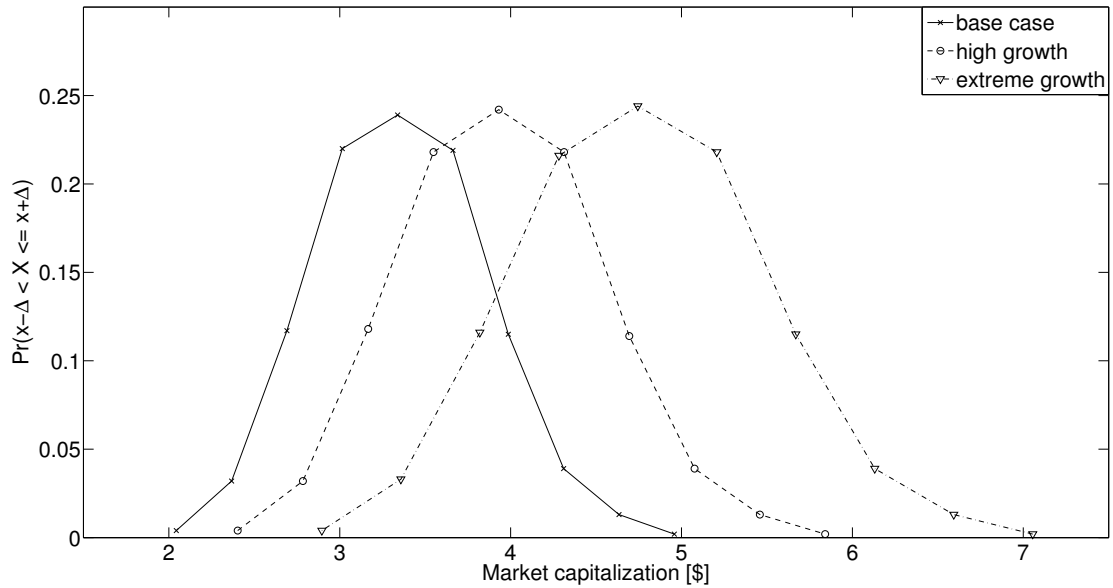


FIGURE 2.8: Distribution of the market capitalization of Zynga according to the base case, high growth and extreme growth scenarios. This shows that the 7 billion USD valuation at IPO or today's 9 billion valuation (April 2012) is not even satisfied in the extreme revenues case.

Scenario	Valuation [\$]	95% conf. interval	Share [\$]	95% conf. interval
Base case	3.4 billion	[2.4 billion; 4.4 billion]	4.8	[3.5;6.2]
High growth	4.0 billion	[2.9 billion; 5.1 billion]	5.7	[4.1;7.3]
Extreme growth	4.8 billion	[3.5 billion; 6.2 billion]	6.8	[4.9;8.9]

TABLE 2.2: Valuation and share value of Zynga in the base case, high growth and extreme growth scenarios.

We obtained a valuation of \$3.4 billion for our base case scenario, well below the \approx \$7 billion value at IPO or the \$9 billion value at the end of March, 2012. Even the unlikely extreme growth case scenario could not justify any of the valuations we have seen in the market so far.

2.1.6 Historic evolution until April 18th 2012

At the time of the IPO, on December 15th, 2011, Zynga was valued at 7 billion USD. Right after the IPO, on December 27th, we published an article on Arxiv (Forró et al., 2011) pointing to an overvaluation of Zynga (which was estimated at 4.2 billion USD in our base case scenario). Since then, Zynga published its earnings for the 4th quarter of 2011. These figures increased the accuracy of our valuation since they contributed to reduce the difference between the three scenarios for the revenues per user. By April 18th 2012, it had traded on the stock market for approximately 4 months. The big question was whether the share price of Zynga moved into the direction of its fundamental value.

As we can see in figure 2.9, it was quite the contrary: after an initial depreciation of the share value reaching a minimum of 7.97 USD on January 9th (still above our extreme case scenario), it was followed by a moderate run-up in price until February 1st 2012, the date of the S1 filing from Facebook. After that, without any solid economic justification, Zynga skyrocketed to a maximum of 14.55 USD/share corresponding to a 10.2 billion USD valuation on February 14th. However, after the release of the 4th quarter results, the company lost more than 15% in a single day, regaining a part of this loss on the following days and peaking again at 14.62 USD on March 2nd, 2012. On March 28th, insiders of Zynga (including its CEO, Mark Pincus) sold 43 million shares in a secondary offering over the counter for 12 USD/share ([424B4 Form of the Filings to the SEC of Zynga, 2012](#)). Zynga's shares subsequently experienced a 6% drop in one day. By the end of April 18th, the publication date of our trading strategy (see section 2.1.7), Zynga closed at 10 USD. More details about Zynga's price trajectory are given in table 2.3 and figure 2.9.

Date	Share price [\$]	Event
2011-12-15	10.00	Zynga goes through its IPO.
2012-01-09	8.00	Zynga closes at its lowest level for the next 4 months.
2012-02-01	10.60	Facebook publishes its S1 Filing. This fuels Zynga's bubble.
2012-02-14	14.35	Zynga unveils its financial results for the 4 th quarter of 2011. The next day, Zynga's share price experiences an 18% drop, its biggest drop until today.
2012-03-01	14.48	Zynga announces that it will launch zynga.com (Takahashi, D, 2012), an independent platform from Facebook. The news is followed by a small increase in share price.
2012-03-21	13.72	Zynga acquires OMGPOP, another social gaming company, for over \$200 million (Cutler, K-M, 2012).
2012-03-27	13.01	Inside investors of Zynga (including its CEO, Mark Pincus) sell 43 million shares at 12\$/piece in a secondary offering. This is followed by a significant 6% drop in share price.
2012-04-18	10.04	We publish our short-term prediction on Arxiv.

TABLE 2.3: Important events in Zynga's price history until April 18th, 2012, the date of our prediction.

The highly volatile, news driven behavior of Zynga's stock price could be quantified using the implied volatility measure. In option pricing, the value of an option depends among others on the volatility of the underlying asset. Knowing the price at which an option is traded, one can reverse engineer the implied volatility, the volatility needed to obtain the market value of the option given a pricing model. This standard measure has the upside of being forward-looking: contrary to the historic volatility, the implied volatility is not computed from past known returns. As such, implied volatility is a good proxy for the mindset of the market. Figure 2.10 compares the implied volatility priced by the market for options written on Zynga with that of Google and Apple in the first

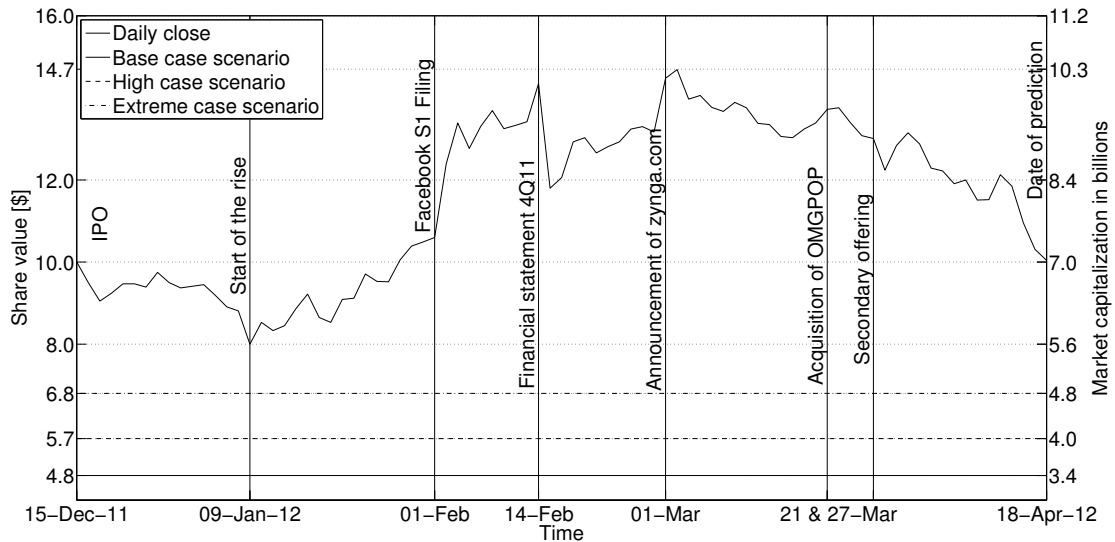


FIGURE 2.9: Share value of Zynga over time. The base case, high and extreme growth scenarios are represented (horizontal lines) as well as important dates for the stock price (vertical black lines). (Source of the data: [Yahoo Finance](#))

half of 2012.

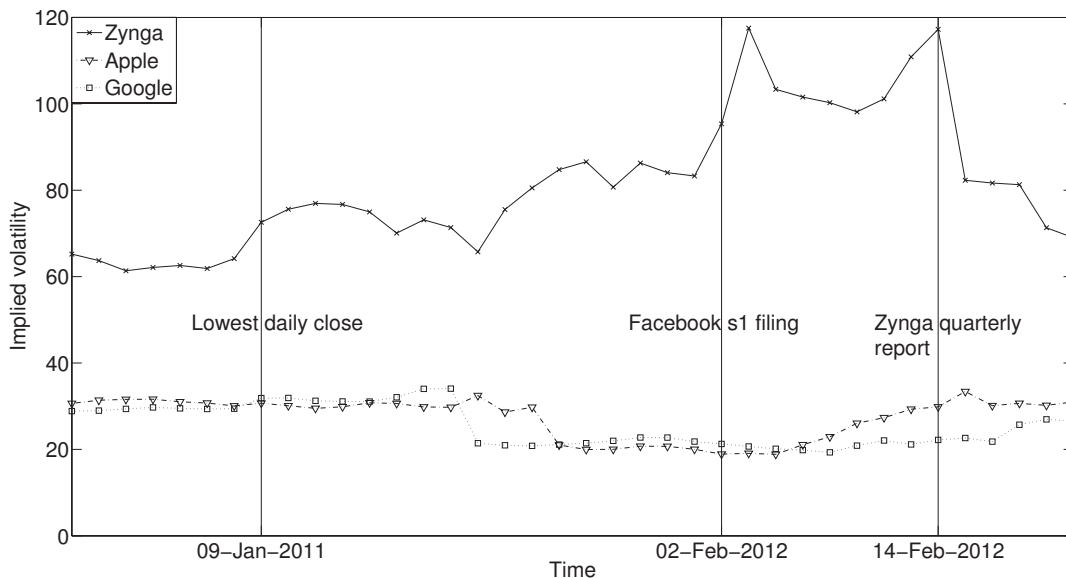


FIGURE 2.10: Implied volatility of Zynga, Apple and Google. Zynga had a much higher and less stable volatility than Apple or Google. (Source of the data: Bloomberg)

We could observe a big difference between the two groups. While Apple and Google had a standard stable implied volatility, there was much more uncertainty surrounding Zynga in the eyes of the market, given its high and unstable volatility. What this told us back then, was that the market players had a hard time putting a value on Zynga. This perception was reinforced by the following event: on February 15th, 2012, the

day following the biggest drop of Zynga since its IPO, most investment banks (with some exceptions) downgraded Zynga's stock rating, readjusting their price target. Some notable examples of actual price targets per share, before the readjustment, were ([Best Stock Watch, 2012](#)):

- Barclays Capital: \$11
- BMO Capital Markets \$10
- Evercore Partners: \$10
- JP Morgan Chase: \$15
- Merrill Lynch: \$13.5
- Sterne Agee \$7

Compared with our analysis (see table 2.2), these price targets seemed to be high. Moreover, even among “experts”, the differences could be significant (the price target of Sterne Agee was less than half of JP Morgan Chase's). One should however keep in mind that the recommendations of most of these companies might not have been independent of their own interest, due to the fact that JP Morgan, Merril Lynch and Barclays Capital were underwriters of the IPO (see [Dechow et al. \(2000\)](#), [Michaely and Womack \(1999\)](#)).

It will be interesting to see the evolution of Zynga on a longer time scale, but so far, our analysis indicated that Zynga was in a bubble, its price not being reflected by its economic fundamentals. Zynga may have been the symptom of a greater bubble, affecting the social networking companies in general, as suggested by the overpricing of Facebook and Groupon ([Cauwels and Sornette, 2012](#)).

2.1.7 Arbitraging Zynga's bubble

While it is commonly assumed that the market price of a stock experiencing a bubble like Zynga should get closer to its fundamental value on the long run, a prediction of its price movements on a shorter time scale is a different problem. It nevertheless possible to develop investment strategies by using a combination of our determination of Zynga's intrinsic value done above with a well known phenomenon, namely the drop of market price when insiders are allowed to sell their shares at the end of their lock-up period.

2.1.7.1 End of lock-up and its implication

When a company goes public, only a fraction of their shares are put on the market (14% in the case of Zynga). The rest of the shares are locked-up for a period of typically 180 days. It is common for IPOs to have a lock-up period in order to prevent insiders from

massively selling their shares after the IPO, and thus driving the market value of the company down. There is a vast amount of literature exploring the effect of the end of the lock-up period on the share value of a company. While there is a broad consensus on the fact that companies, on average, experience abnormal negative returns following the end of the lock-up period, different authors give different explanation of this effect. Some of the relevant results for Zynga are the following: [Field and Hanka \(2001\)](#) found that venture capital backed firms experience the largest price drop at the end of the lock-up period. [Jordan et al. \(2000\)](#) confirmed the finding and added that the “quality” of the IPO underwriters as well as the price increase since the IPO is positively correlated with the drop in share value. [Gao \(2005\)](#) found that firms with the highest forecast bias and the highest forecast dispersion by analysts experience the largest drop. Finally, [Ofek \(2008\)](#) made the argument that a significant increase in share supply could explain the price drop subsequent to the end of the lock-up period. He further argued that the higher the stock price volatility before the end of the lock-up, the bigger the drop in share value.

While most of the above mentioned authors found that it is difficult to develop an arbitrage strategy to take advantage of this effect, we should stress that they all based their works on samples of companies independent of any view regarding their intrinsic value. If they could bias their sample towards the companies whose market value is significantly higher than their fundamental value, we would expect a different outcome. We hypothesized that the overvaluation of the company would be reflected in its market price, as soon as insiders, better informed about the fundamentals of their company, would be allowed to trade freely, i.e., at the end of the lock-up period. We believed that the information asymmetry between outside traders (the only ones who are allowed to trade the shares of Zynga from its IPO) and insiders would be incorporated in the price formation of such a company, and move its market price towards its fundamental value.

2.1.7.2 Prediction for Zynga

On March 23rd, Zynga announced, in an [S1/A Form of the Filings to the SEC of Zynga, 2012](#) that inside investors including CEO Mark Pincus would sell about 43 million shares. This move was surprising since Zynga’s inside investors were subjected to a lock-up period ending on May 29th. However, a secondary offering was authorized by the underwriters of the IPO and was concluded on March 28th with the shares being sold over the counter for \$12/piece (\$0.36 of which went to the underwriters). On that day, the share value of Zynga experienced a large drop, going from \$13.02 to \$12.24 (this corresponded to a 6% decrease). It should be noted that, subsequent to this transaction, these insiders were again subject to a lock-up period and wouldn’t be able to trade until its end. In

practice, the remaining locked shares were to be released in the market in several steps (see [424B4 Form of the Filings to the SEC of Zynga, 2012](#)):

1. Approximately 115 million shares held by non-executive employees around April 30th (or 3 days after Zynga will disclose its financial statement for the first quarter of 2012).
2. Approximately 325 million shares held by non-employee stockholders that have not participated in the secondary offering on May 29th, 2012.
3. Approximately 50 million shares held by directors, executive employees and the stock holders who participated in the secondary offering (such as Mark Pincus) on July 6th, 2012.
4. Approximately 150 million shares held by the same persons as in 3. on August 16th, 2012.

When trying to predict the future price movements of Zynga, one couldn't ignore the fact that on April 26th, 3 days before the end of the first part of the lock-up period, the company was releasing its financial results for the first quarter of 2012. Our basis To understand the implications of this report on Zynga's price movement, was our diagnostic of a bubble. Indeed, during a bubble, phenomena like herding and imitation are dominant among traders ([Sornette, 2003](#)). As such, the market players are very sensitive to new information, giving rise to behaviors inconsistent with its content. We believed that the release of Zynga's financial statement on April 26th could be such an event. According to our model, Zynga's yearly revenue per user was saturating. This can be seen on the left-hand side of figure 2.7, where the yearly revenue per user, computed at each quarter as the running sum of the four previous quarters, were well fitted by a logistic function. The saturation of Zynga's yearly revenues per user was a powerful argument for the diagnostic of a bubble in Zynga's market valuation. Even with a hypothetical 357 million USD of revenues for the first quarter of 2012 (which were published on April 26th) meaning a 15 % increase from the 311 million USD of revenues last quarter, the yearly revenues per user would fall right onto our logistic fit. Hence, even a 15 % increase in quarterly revenues would not have been sufficient to rationally reject the saturating trend of Zynga's revenues per user that we predicted. However, compared with the results of the previous quarter, which saw Zynga's revenue only rise by 1.4% (see figure 2.11), this would have been seen as a very strong performance and would have most likely been followed by an increase in share value, even more so in the bubble environment that we diagnosed.

Up to the end of the lockup period, there were about 150 million shares tradable on the market (100 million from the IPO and about 50 million from the secondary offering). The 115 million shares coming to the market around April 30th represented an important

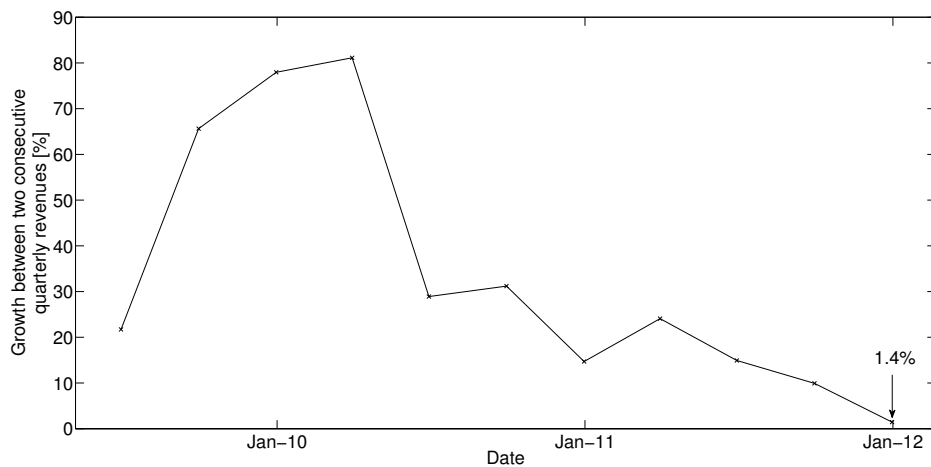


FIGURE 2.11: Percentage difference between the revenues of two consecutive quarters. We can see that Zynga's performance in the last quarter of 2011 was very poor. This suggested that taking the 1.4% figure as a benchmark to evaluate Zynga's performance for the following quarter may have lead investors to be overly optimistic, especially during a bubble period (Source of the data: S1 Filing to the SEC).

increase in the free-floating shares of Zynga. As such, and because Zynga satisfied most of the conditions given in subsection 2.1.7.1 leading to a large price decrease, we predicted a drop of Zynga's market value around that date. We should mention that we thought that what would happen around April 30th would be conditional on what would have happened on April 26th: we predicted this drop to be larger if Zynga's stock price increased on April 26th and smaller if the stock price decreased on the same date. Such a phenomenon could have taken place at each such date.

Given the elements mentioned before, we believed that there was a high probability of strong corrections in Zynga's price after each partial lift of the lock-up period. While we have shown that even an apparently strong performance on April 26th would have been in line with our diagnostic of Zynga's saturating revenues per user, one should not be surprised to see its share value rise in this bubble environment. In the long-run, we predicted Zynga's market value to move towards its intrinsic value of 3.4 billion USD. In summary, the proposed strategy was based on three time periods:

1. From the time of writing (April 16th, 2012) to the announcement of the financial results (around April 26th, 2012): stay out of Zynga or hedge if invested.
2. From the day after the earnings announcement (around April 27th, 2012) to the end of the first lock-up period (around April 30th, 2012): if the financial results are significantly above those of the previous quarter, buy Zynga for a short term holding period. Otherwise short it.

3. From the end of the first lock-up period (after April 30th, 2012): close all open long positions and short. Monitor the subsequent quarterly releases and the successive ends of future lock-up periods to position a strategy in the same spirit as above.

2.1.8 Post-mortem analysis, on May 24th, 2012, of the proposed strategy

This section was added on May 24th after the main paper was accepted for publication on May 17th, 2012. The version of the paper with our ex-ante proposed strategy (section 2.1.7) can be found on Arxiv with the date stamp of April 19th, 2012 (Forró et al., 2012a). In this section, we evaluated our ex-ante prediction in the light of the events that occurred until end of May 2012. Figure 2.12 summarizes the price movements of Zynga from April 19th till May 24th, 2012.

1. *From the time of writing (April 16th, 2012) to the announcement of the financial results (around April 26th, 2012): stay out of Zynga or hedge if invested.* Between April 19th and April 26th, Zynga's share price dropped from \$10.2 to \$8.2 and then rebounded to \$9.42 (the opening price on April 19th). Although the stock went down 7.7% in a week, its behavior was very volatile. As we did not have any strong factual information to support a clear trading strategy before April 27th, not taking a position appears to have been an acceptable advice.
2. *From the day after the earnings announcement (around April 27th, 2012) to the end of the first lock-up period (around April 30th, 2012): if the financial results are significantly above those of the previous quarter, buy Zynga for a short term holding period. Otherwise short it.* This part was undeniably a success. On April 26th, after the markets closed, Zynga revealed its financial results for the first quarter of 2012. Its quarterly revenues were weak, since they only grew 3.1% since the previous quarter, confirming that the company was in its saturation phase. As a result, on April 27th, Zynga experienced a drop of 9.6%, one of its largest daily drops since its IPO.
3. *From the end of the first lock-up period (after April 30th, 2012): close all open long positions and short. Monitor the subsequent quarterly releases and the successive ends of future lock-up periods to position a strategy in the same spirit as above.* As this last part covered a large time-period (from April 30th to May 24th, 2012), we divided it into a short-term part (the first day) and a longer-term part (until the time of writing of the post-mortem analysis, May 24th, 2012).
 - On the first day (April 30th), the prediction was proven successful. Indeed, as a result of the end of the lock-up period, the stock further dropped 2.1% in a

single day. Had someone opened a short position on April 27th (at the opening of the markets) and closed it on April 30th (at the closing of the markets), he would have benefited of a 11.5% drop over 2 trading days.

- On the longer term, the price trajectory, although quite volatile, went significantly down. This was accentuated by Facebook’s IPO on May 18th. Indeed, it was soon clear from the price dynamics after the Facebook’s IPO that it was not a big success. On the other hand, due to the use of the “over-allotment” or “green shoe” option, the price would be kept artificially above the IPO price for a time. Therefore, investors targeted other social networks like Zynga which lost 13%, LinkedIn which lost 6%, Groupon which lost 7% or Renren, the Chinese Facebook which lost 21%. The lack of rebound of Zynga (until May 24th) may have been due to the loss of its status as a “proxy” for Facebook. It is worth noting that, for the first time since it went public, Zynga’s value entered our fundamental valuation bracket, when on May 21st it dropped to 6.5\$/share (intraday), below our extreme case scenario of 6.8\$/share...

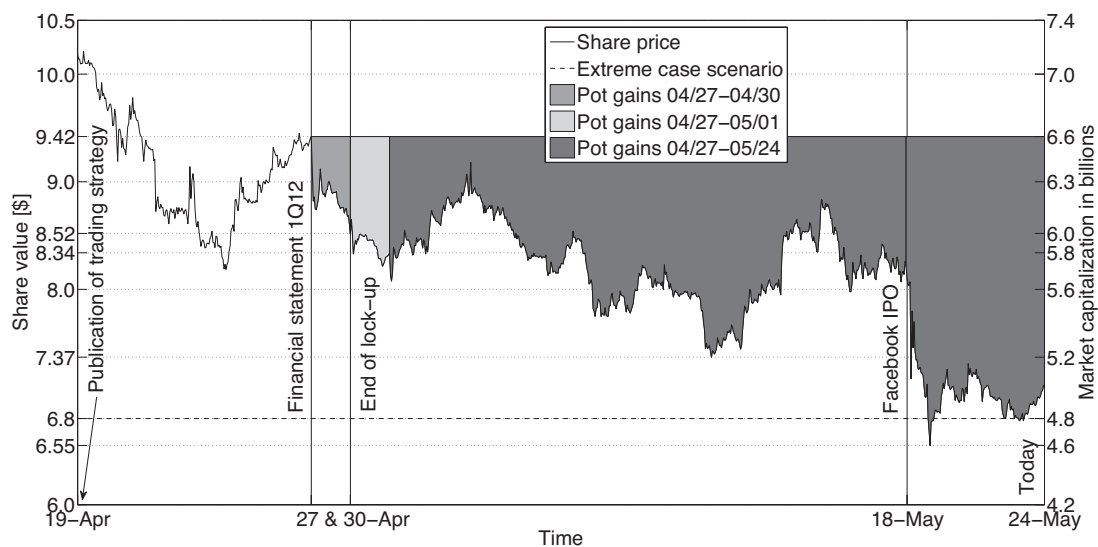


FIGURE 2.12: Price dynamics of Zynga from the publication of our trading strategy on April 19th, 2012 on the arXive until May 24th, 2012 just before going to press. Potential gains (indicated as “Pot gains” in the inset) that could be obtained by opening a short position on April 27th are indicated by the shaded area. The data has a 10 minutes time resolution. (Source of the data: Bloomberg).

2.1.9 Outlook. Added on January 23rd, 2015.

As for every prediction, the question that needs to be asked is how that prediction turned out. In the case of Zynga, more than two and half years have passed since the publication of our results, giving us some perspective. We have thus compared the evolution of Zynga’s market capitalization in comparison to the upper bounds computed

for the base case, high growth and extreme growth scenarios. Results can be seen on figure 2.13.

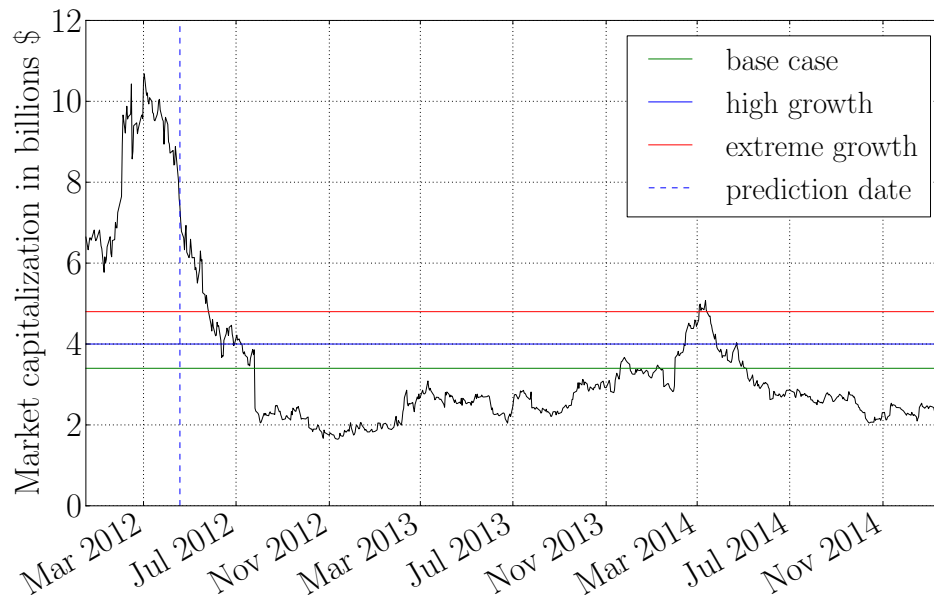


FIGURE 2.13: Evolution of Zynga’s market capitalization until January 2015. One can see that Zynga’s market value has continued to fall after the date of our prediction (April 18th, 2012) and finally stabilized below the upper bound defined by our base case scenario.

Our valuation of Zynga seems to have been very good at forecasting Zynga’s price movements on the short-term (see section 2.1.8) as well as on the long-term. As we can see, after May 18th, 2012 (end of our post-mortem analysis), Zynga continued its correction and crossed the upper-bound of our base case scenario (\$3.4 billion) on July 26th 2012. Zynga’s market capitalization stayed within the confines of our base case scenario ever since, with the exception of a 5-month period between January and May of 2014, when Zynga’s market value hit the limit determined by our extreme case scenario (\$ 4.8 billion) before falling again below our base case valuation.

2.1.10 Conclusion

In this work, we have proposed a new valuation methodology to price Zynga. Our first major result was to model the future evolution of Zynga’s DAU using a semi-bootstrap approach that combines the empirical data (for the available time span) with a functional form for the decay process (for the future time span).

The second major result was that the evolution of the revenues per user in time, r_i , showed a slowing of the growth rate, which we modeled with a logistic function. This

made intuitive sense as these r_i should be bounded due to various constraints, the hard constraint being the economic one, since Zynga's players only have a finite wealth. We studied 3 different cases for this upper bound: the most probable one (the base case scenario), an optimistic one (the high growth scenario) and an extremely optimistic one (the extreme growth scenario).

Combining these hard data and soft data revealed a company value in the range of 3.4 billion to 4.8 billion USD (base case and extreme growth scenarios). Given the optimistic assumptions that were taken in the pricing of the company (power-law decay process, sampled top 20 games with equal probability, took a 15% profit margin and a 5% constant risk premium), our estimates corresponded to upper bounds in Zynga's valuation.

Our valuation allowed us to formulate a short-term trading strategy taking into account the discrepancy of the company's market value and our estimate (around April, 2012) and the end of the lock-up period, i.e. the possibility for insiders to sell shares, on April 30th. Our strategy, that was essentially a bet against Zynga, and proposed ex-ante, would have been very successful Zynga losing about 75% of its value in the subsequent 5 months. On a longer-term horizon, our valuation seems be supported by the fact, that Zynga's price trajectory essentially stayed within the confines of our base case scenario.

Naturally, one successful example such as this one, is not sufficient to prove the forecasting power of our valuation methodology. As such, an interesting topic for future research would be to apply our valuation methodology to many more social networking companies in order to obtain statistically significant results.

2.2 Forecasting future oil production

² While addressing a seemingly different issue than that of Zynga, forecasting future oil production turned out to be a strikingly similar problem as forecasting Zynga's user base, where the individual oil fields of a country played the same role as Zynga's games. Just as Zynga's DAUs, the production of oil fields displays regularity in their decay dynamics and just as Zynga's games, new oil fields are discovered as a function of time. But even more fundamentally, forecasting future production of oil can be seen through the lenses of fundamental valuation. In fact, a difference between the estimated and the real remaining oil reserves can be seen as a bubble component, where this difference can have an impact on long term oil price dynamics. Due to oil's central role in our economy, forecasting its production has been a topic of active interest since the beginning of the past century. Its importance ranges from energy production, through manufacturing to pharmaceuticals industry. Even beyond the impact on long term oil price dynamics, being able to forecast future production is an important problem, since a misestimation of its reserves can have great consequences on our society. In this work, we have developed a Monte-Carlo methodology to forecast the crude oil production of Norway and the U.K. based on the current/past performances of individual oil fields. By extrapolating the future production of these fields and the frequency of new discoveries, we were able to forecast the oil production of these two countries. Our results indicate that standard methodologies tend to underestimate remaining oil reserves. We compared our model to those methodologies by making predictions from various time points in the past. We show that our model gives a better description of the evolution of oil production. This section is based on a paper co-authored with Lucas Fiévet, Peter Cauwels and Didier Sornette (Fiévet, Forró, Cauwels, and Sornette, 2015) titled "A general improved methodology to forecasting future oil production: Application to the UK and Norway".

2.2.1 Introduction

The methodology behind forecasting future oil production has not evolved much since Hubbert who, in 1956, famously predicted that the U.S. oil production would peak around 1965-1970 Hubbert (1956). That prediction proved to be correct. His main argument was based on the finiteness of oil reserves and used the logistic differential equation (equation 2.4) to model oil production. The logistic equation was already used in this thesis to describe Zynga's revenues per users in time. As a reminder,

2. This section is based on the paper co-authored with Lucas Fiévet, Peter Cauwels and Didier Sornette titled "A general improved methodology to forecasting future oil production: Application to the UK and Norway."

the logistic differential equation is characterized by an initial exponential growth, the growth decreasing to zero as the total oil extracted reaches saturation (no more oil is to be found).

$$\frac{dP}{dt} = rP \left(1 - \frac{P}{K} \right) \quad (2.4)$$

r is commonly referred to as the growth rate, and K as the carrying capacity (total quantity of oil extracted). If $P(t)$ is the the amount of oil extracted up to time t , then $\frac{dP}{dt}$ is the oil production rate, the quantity that M. King Hubbert predicted would peak with surprising accuracy. From a methodological point of view, the Hubbert model has enjoyed a longstanding popularity in modeling future oil production given its simplicity. In this work, we introduced a new methodology to forecast future oil production. Instead of taking the aggregate oil production profile and fitting it with the Hubbert curve or its variants (such as the multi cyclic Hubbert curve ([Anderson and Conder, 2011](#), [Nashawi et al., 2010](#))), we used the production profile of each individual oil field. By extending their production into the future and extrapolating the future rate of discovery of new fields, we are able to forecast future oil production by the means of a Monte Carlo simulation. To demonstrate the generality of the methodology presented here, we applied it to two major oil producing countries: Norway and the U.K.

2.2.2 Methodology

The idea behind the methodology was to model the future aggregate oil production of a country by studying the production dynamics of its individual constituents, the oil fields. The main benefit of this approach, compared to working directly with aggregate production data, was the possibility to forecast non-trivial oil production profiles arising from the combination of all the individual fields' dynamics. In order to achieve that, one must be able 1) to extend the oil production of each individual field into the future and 2) to extrapolate the rate of discoveries of new oil fields. As these were the steps taken to model Zynga's future user base, the case for oil is described in a more concise manner.

2.2.2.1 Extending the oil production of individual fields

To be able to forecast the oil production of each individual field, regularity needs to be found in the production's dynamics. Just as in Zynga's case, modeling the whole production profile from the beginning of extraction seemed elusive due to the variety

of the forms it could take. Fortunately, modeling the decay process was sufficient in order to extrapolate future oil production. A preliminary classification was necessary to achieve that goal. Figure 2.14 shows that independent of the country, oil fields can be classified into 2 main categories.

- Regular fields. The fields whose decay show some regularity.
- Irregular fields. Irregular fields are the ones that don't decay in a regular fashion or that don't decay yet. As such, there was no easy way to forecast their future oil production based on past data.

As of January 2013, regular fields made up 85% and 87% of the number of fields and 94% and 71% of the total oil volume in the Norway and the U.K. respectively. As such, being able to model them was crucial. As can be seen on figure 2.15, the stretched exponential (equation 2.15) was a good functional form to fit the decay process of regular fields (Laherrère and Sornette, 1998).

$$P(t) = P_0 e^{\left(\frac{-t}{\tau}\right)^\beta} \quad (2.5)$$

For the minority of irregular fields, we simply modeled their decay as follows. We picked τ to be the average τ over the regular fields. We then fixed β so that the sum of the fields' production over their lifetime be equal to the official ultimate recovery estimates, when such an estimate was available.

2.2.2.2 Discovery rate of new fields

Knowing the future production rate of existing fields was not sufficient; new fields would be discovered in the future. The model describing the discovery rate of new fields needed satisfy two fundamental observations.

1. The rate of new discoveries should tend to zero as time goes to infinity. This is a consequence of the finiteness of the number of oil fields.
2. The rate of new discoveries should depend on the size of the oil fields. As of today, giant oil fields are discovered much less frequently than dwarfs (Höök et al., 2009).

A natural choice for such a model was a non-homogenous poisson process. The Poisson process is a process that generates independent events at a rate λ . It is inhomogeneous if the rate is time-dependent, $\lambda \rightarrow \lambda(t)$. The standard way to measure $\lambda(t)$ is to find a functional form for $N(t)$, the number of events (discoveries) up to time t . Then, $\lambda(t)$ is simply $\frac{dN(t)}{dt}$. Figure 2.16 shows $N(t)$ for Norwegian fields classified according to their

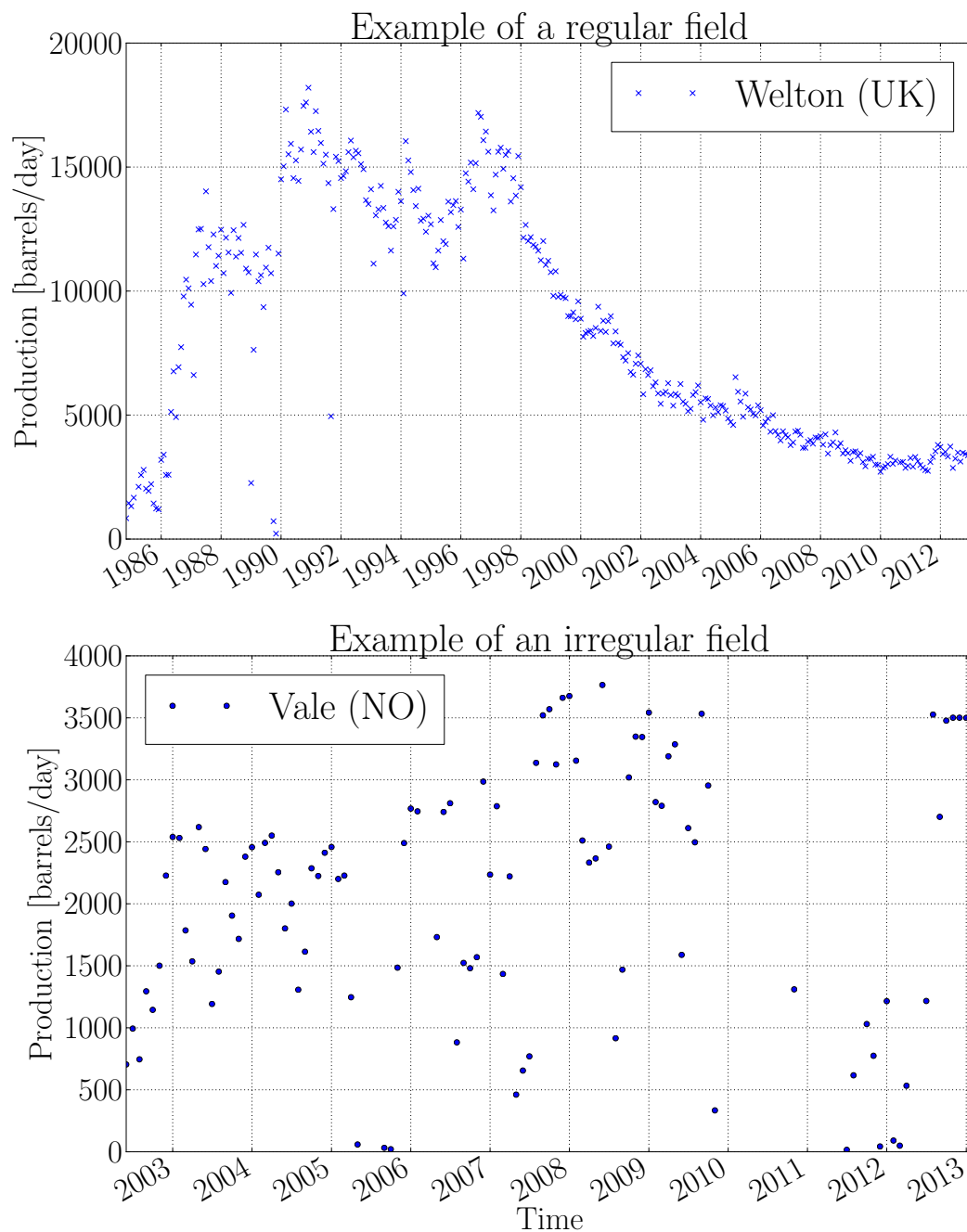


FIGURE 2.14: Example of a regular and an irregular field.

size. As can be seen, the discovery dynamics of new fields is size-dependent; huge fields were expected to have been (nearly) all discovered, while smaller fields were more likely to appear. As such, the fields were divided into two groups; giants and dwarfs as is often found in the literature. In order to choose the size limit below which fields were considered dwarf in the least arbitrary way possible, we turned to the literature to learn that often, dwarf fields that have already been discovered have their exploitation postponed for economic reasons. Consequently, we set the splitting size as small as

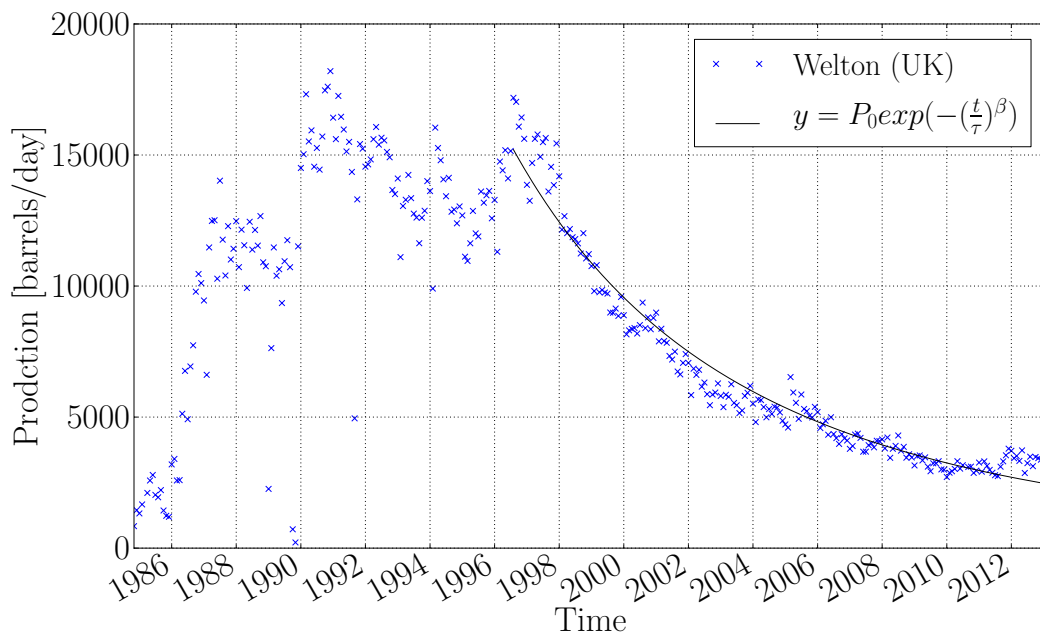


FIGURE 2.15: Stretched exponential fit ($\tau = 144$ days and $\beta = 0.42$) to the production of Welton. This functional form fits the data well.

possible in order to maximize the number of giant fields but large enough to avoid recent discoveries. This resulted in a threshold of $50 \cdot 10^6$ barrels. Functionally, the cumulative rate of new discoveries was described by the logistic curve (integral form of equation 2.4). This implied that after an initial increase, the rate of new discoveries reached a peak followed by a decrease until no more oil fields were to be found, consistent with our fundamental observations. The results were qualitatively similar for the U.K.

2.2.3 Results

Simulating future oil production was straightforward. For each country, new oil fields were generated according to the poisson process, with $\lambda(t)$ chosen according to their size. For each of the generated oil fields, its production profile was chosen randomly from one of the existing fields within the same size category. Figures 2.17 and 2.18 show the average of 1000 simulations. For each country, we could immediately notice the non-symmetric shape of the production's dynamics contrary to what Hubbert would have predicted.

TABLE 2.4: Remaining oil reserves in billion of barrels

Country	Hubbert	our model	Δ
Norway	2.83	5.72	102%
U.K.	1.43	2.93	105%

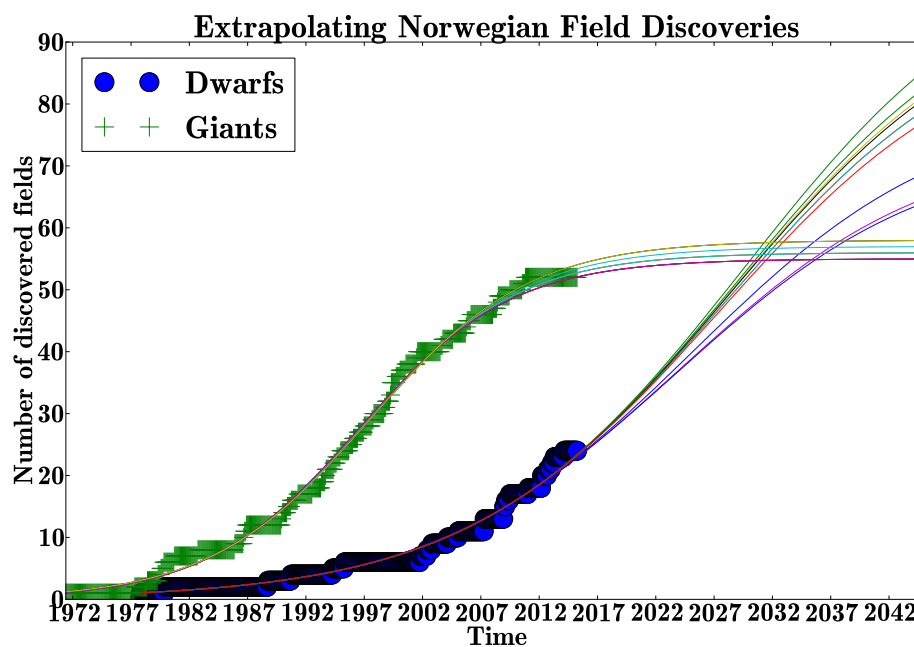


FIGURE 2.16: Logistic fit to the number of discoveries for Norway. The discovery rate of new oil fields is dependent on their size. The logistic model suggests that the majority of small oil fields have not yet been discovered, while the rate of discovery of medium oil fields is in its decreasing phase. Big oil fields are not expected to be discovered anymore. The fitting process developed in [Smith \(1980\)](#) has been used.

The difference between our and the Hubbert-based forecast is striking. In the case of Norway, the standard Hubbert model was used to fit the production data. According to it, Norway's future oil production would decay much faster than in the Monte-Carlo case. As such, the estimated remaining reserves were less than half of the forecasted value of the Monte-Carlo methodology. In the U.K., oil production faced a change of regime during the early nineties due to technological innovation, giving rise to the inverted "w shape" of the oil production profile. The standard procedure to model oil production in those cases is to use a multi-cyclic Hubbert curve. The multi-cyclic Hubbert model is a generalization of the standard one. Conceptually, it is just a superposition of several standard (single-cyclic) Hubbert curves. Two cycles are commonly used to fit the UK's oil production and that is what we did. The difference between the Hubbert-based methodology and the Monte-Carlo one was very similar to the Norwegian case. The former underestimated the remaining oil reserves by about 51% compared to the latter.

Which of the two models is more thrust worthy? Clearly, the implications in adopting one methodology over the other would be significant. The only way to answer this question was to backtest them. In other words we asked: "What would have each of the models predicted had we used them in the past?"

2.2.4 Validation

For each of the two countries, namely Norway and the U.K., we went back in time 6 years (beginning of 2008). The reason for not going back further in time had to do with the number of giants oil fields entering their saturation phase in order to successfully apply the extrapolation algorithm. Figures 2.19 and 2.20 show the outcome of those tests. Comparing the forecast of both models with the oil production of the subsequent 6 years, we found the predictive power of the Monte-Carlo methodology remarkable. While, in 2008, the Hubbert model would have forecasted a significantly lower oil production for Norway than what really happened, the Monte-Carlo simulation perfectly captured the Norwegian oil production dynamics between 2008 and 2013. The U.K. example was not so clear cut: both methodologies perform reasonably well. However, one can notice that the oil production dynamics stay within the confidence of the Monte-Carlo case while they fall outside of the Hubbert case. Table 2.5 quantifies the difference between the Hubbert and the Monte-Carlo forecast for the oil production of both countries between 2008 and 2013. While these results were impressive, one should be aware of the fact that the Monte-Carlo methodology can be sensitive to the definition of the group into which the individual oil fields were divided.

TABLE 2.5: Oil production between 2008 and 2013. One can see that the difference between Monte-Carlo forecast and actual production was smaller in both cases than the difference between the Hubbert extrapolation and actual production.

Country	Actual	Hubbert	our model	$\Delta_{Hubbert}$	$\Delta_{our\ model}$
Norway	4.05	3.29	3.95	-18.8%	-2.5%
U.K.	2.40	1.63	2.28	-18.5%	14.0%

2.2.5 Conclusion

In this work, we have shown that by applying a Monte-Carlo methodology, very much like the one used to forecast Zynga's future user base, we could forecast future oil production with greater level of accuracy than standard methods such as the Hubbert model. It would definitely be interesting to see if these results generalize to the oil production of other countries. But maybe even more importantly, future research should investigate how these results could impact oil price dynamics.

Norwegian oil production and forecast

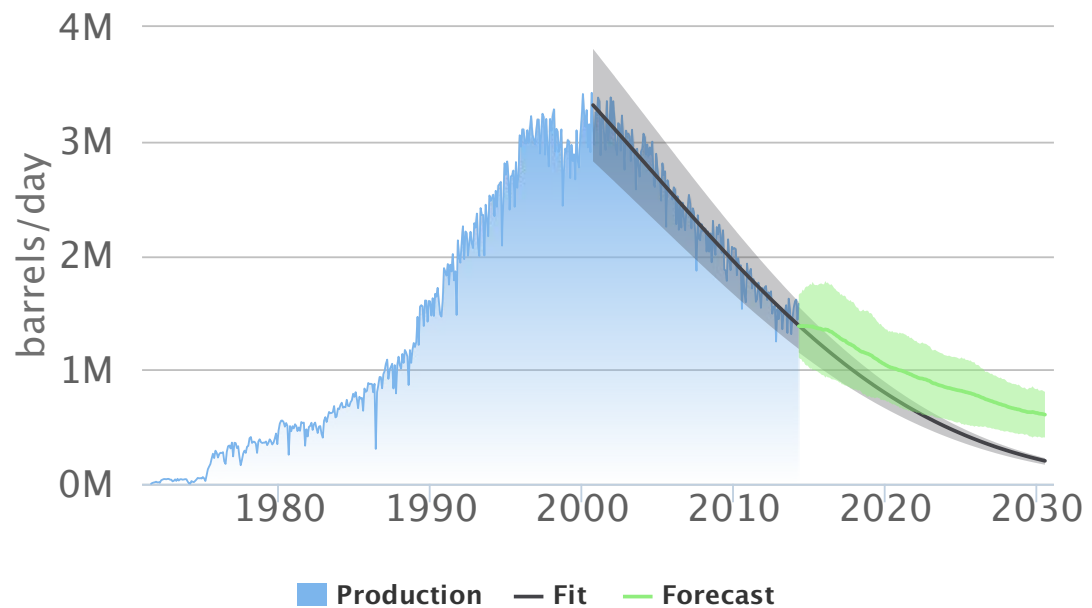


FIGURE 2.17: Monte-Carlo and Hubbert forecast of future oil production for Norway. The Monte-Carlo methodology forecasts a significantly higher oil production than the Hubbert one.

U.K. oil production and forecast

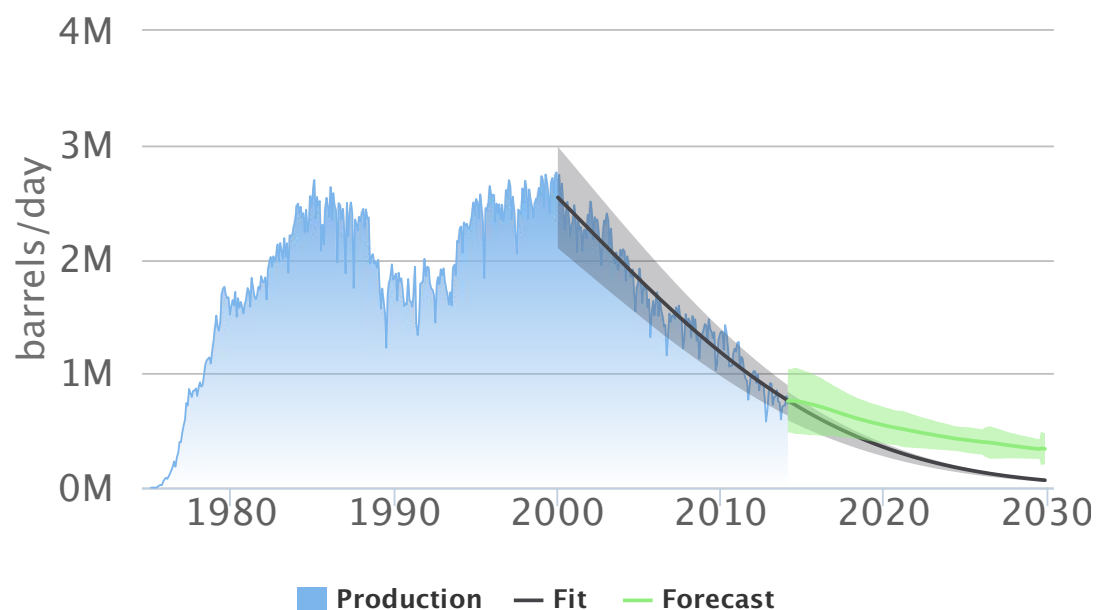


FIGURE 2.18: Monte-Carlo and Hubbert forecast of future oil production for the U.K. The Monte-Carlo methodology forecasts a significantly higher oil production than the Hubbert one.

Norwegian oil production and forecast

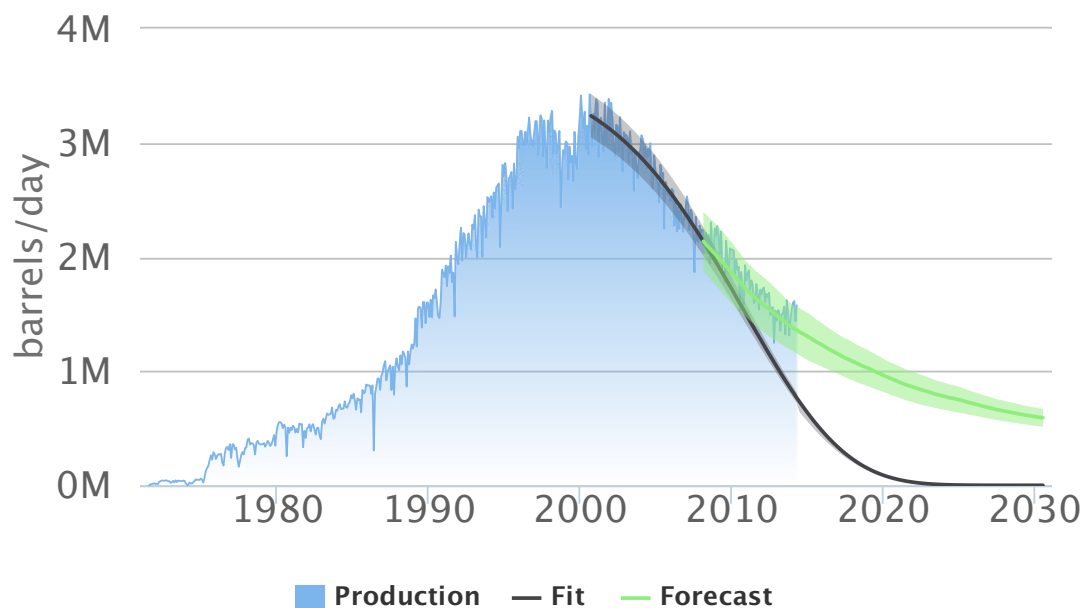


FIGURE 2.19: Backtest of the Monte-Carlo and Hubbert models based on past production data up to 2008 for Norway. The green (black) line corresponds to the Monte-Carlo (Hubbert) forecast where the shaded area represents the uncertainty associated with it. The Monte-Carlo forecast seems to give a much better description of the oil production dynamics between 2008 and 2013.

U.K. oil production and forecast

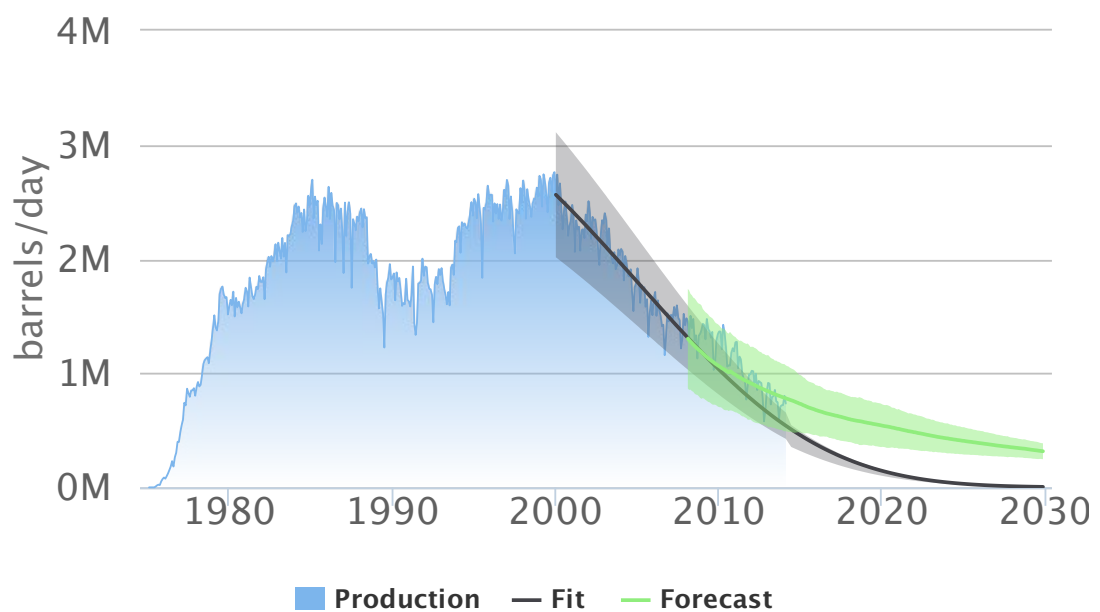


FIGURE 2.20: Backtest of the Monte-Carlo and Hubbert models based on past production data up to 2008 for the U.K. The green (black) line corresponds to the Monte-Carlo (Hubbert) forecast where the shaded area represents the uncertainty associated with it. While the difference is not as striking as in the Norwegian case, The Monte-Carlo forecast has a greater accuracy than the Hubbert model given it's smaller error in the description of the oil production dynamics between 2008 and 2013 (ref table 2.5).

Chapter 3

Bubbles as price dynamical processes

In the previous chapter, we have diagnosed bubbles as the gap between an asset's fundamental and market values. Its implications on long-term price dynamics was investigated through the example of Zynga. Our findings were very encouraging, since Zynga's market value moved towards our fundamental valuation. This hints at the fact that diagnosing bubbles in this manner can have predictive power. However, such a methodology is hard to deploy at large scale since we specifically targeted a social networking company due to its simple business model. This difficulty to scale is amplified by the fact that, even among social networking companies, each firm needs to be studied individually. Another shortcoming of that fundamental value oriented methodology is the absence of proposed mechanisms for bubbles. As such, it is hard to give a time horizon over which the bubble would burst. We managed to do so in the case of Zynga, using the expiration of the lock-up period, but such an event only happens once in the lifetime of a company. These issues motivated us to look at models trying to detect bubbles in the dynamical signatures in the prices of assets. In particular, we turned our attention to the log-periodic power law model (LPPL), a model that looks for super-exponential increases in the timeseries of price due to herding among traders. We asked the question of its predictive power on a large scale. We then extended the notion that non-linear price dynamics are significant in explaining asset returns using a factor regression framework, a standard in economics.

3.1 The predictive power of LPPL

¹ This work was based on a paper co-authored with Ryan Woodard and Didier Sornette (Forró et al., 2015) titled “On The Use of Trading Strategies to Detect Phase Transitions in Financial Markets”. We showed that the log-periodic power law model (LPPL), a mathematical embodiment of positive feedbacks between agents and of their hierarchical dynamical organization, has a significant predictive power in financial markets. We found that LPPL-based strategies significantly outperformed the randomized ones and that they were robust with respect to a large selection of assets and time periods. The dynamics of prices thus markedly deviate from randomness in certain pockets of predictability that can be associated with bubble market regimes. Our hybrid approach, marrying finance with the trading strategies, and critical phenomena with LPPL, has demonstrated that targeting information related to phase transitions enables the forecast of financial bubbles and crashes punctuating the dynamics of prices.

3.1.1 Introduction

Complex systems often exhibit non-linear dynamics due to the interactions among their constituents. In particular, phase transitions are a very common phenomenon that emerge due to coupling leading to positive and negative feedback loops. Examples of such systems can be found in a broad range of fields such as biology with the firing of neurons in the brain (Kinouchi and Copelli, 2006), ecology with the turbidity of lakes (Scheffer et al., 1993), sociology with the Arab Spring (Root, 2013) or, as we propose, the economy with stock market bubbles and crashes (Sornette, 2003).

The stock market is a complex system whose constituents are economic agents. They are heterogeneous in size and preference (financial institutions, individual traders etc. . .) and interact through a complex network topology (Roukny et al., 2013). As such, statistical stationarity and complete unpredictability of price dynamics as postulated by classical economic theory seem unlikely, as exemplified by the emergence and run-up of financial bubbles until the Global Financial Crisis of 2008 (Sornette and Cauwels, 2014). In fact, many models have been put forward to describe bubbles 1.1 or diagnose their occurrence but quantifying their explanatory and predictive power remains an outstanding problem (Kaizoji and Sornette, 2010).

We propose that bubbles and their ensuing crashes can be seen as phase transitions where the behavior of the economic agents becomes synchronized through positive feedback loops, building up the bubble and eventually leading to its collapse. The crash is not the

1. This section was based on the paper co-authored with Ryan Woodard and Didier Sornette titled “On The Use of Trading Strategies To Detect Phase Transitions in Financial Markets”.

result of a new piece of information becoming available to market participants; instead, it is the result of a system close to criticality, where even a tiny perturbation is enough to reveal the large susceptibility associated with the approach to the phase transition. These concepts are captured by the log-periodic power law model (LPPL) (Johansen et al., 2000, Sornette et al., 1996) which has been usefully applied in the description of other physical phenomena such as earthquakes (Johansen et al., 1996) or the rupture of materials (Anifrani et al., 1995). In the case of the stock market, the basic mechanism generating the positive feedback loops is herding, both technical and behavioral: during a bubble regime, the action of buying the asset pushes its price up, which itself leads paradoxically to an increased demand in the asset. This process is unsustainable and inevitably leads to a change of regime, which often results in a financial crash (Sornette et al., 1996).

We have extended previous LPPL-related studies (Johansen et al., 2000, Sornette et al., 1996) in two ways. On a procedural level, we have tested for LPPL's predictive power on a very large scale (by analyzing about 1000 stocks over 20 time periods) in a unified systematic framework to substantiate the significance of our results. But maybe more importantly, on a conceptual level, we can think of this work as developing a hybrid methodology combining physics and finance based on the outcome of trading strategies constructed on the log-periodic power law model. Usually the province of finance, trading strategies are here proposed as genuine nonlinear transformations mapping an input time series (here a price) onto an output profit-and-loss time series that, when coupled with physical mechanism(s), may reveal novel properties of the studied system (Zhou et al., 2011).

3.1.2 The log-periodic power law model

The positive feedback process at the core of LPPL can be described by the simple differential equation 3.1 capturing the market impact of excess demand fueled by growing prices:

$$\frac{dp(t)}{dt} \propto p(t)^\delta, \quad (3.1)$$

with p , the price. When the exponent δ is greater than 1, it represents positive feedback of the price on the instantaneous rate of return. In this case, the solution to equation 3.1 becomes:

$$p(t) \propto (t_c - t)^{-m}, \quad (3.2)$$

where $m = \frac{1}{\delta-1}$. This solution is quite remarkable because of the emergence of the hyperbolic power-law describing a super-exponential regime ending in a finite-time singularity occurring at the movable time t_c determined by initial conditions, beyond which equation 3.2 has no solution. This can be interpreted as a change of regime, where the price dynamics go from a super-exponential to something different, a crash for example. By abandoning the need to describe the full process of the bubble, followed by a crash and the subsequent market recovery and evolution, we gain the insight that the information on the end of the bubble regime, embodied in the critical time t_c , is contained in the price dynamics itself during the bubble. We should stress that this super-exponential increase is one of the important insights of the model; it is precisely the herding mechanism at the origin of the positive feedback process that is at the core of this interesting dynamical signature. One of the best examples of this phenomenon is the Hong-Kong Hang Seng index (figure 3.1).

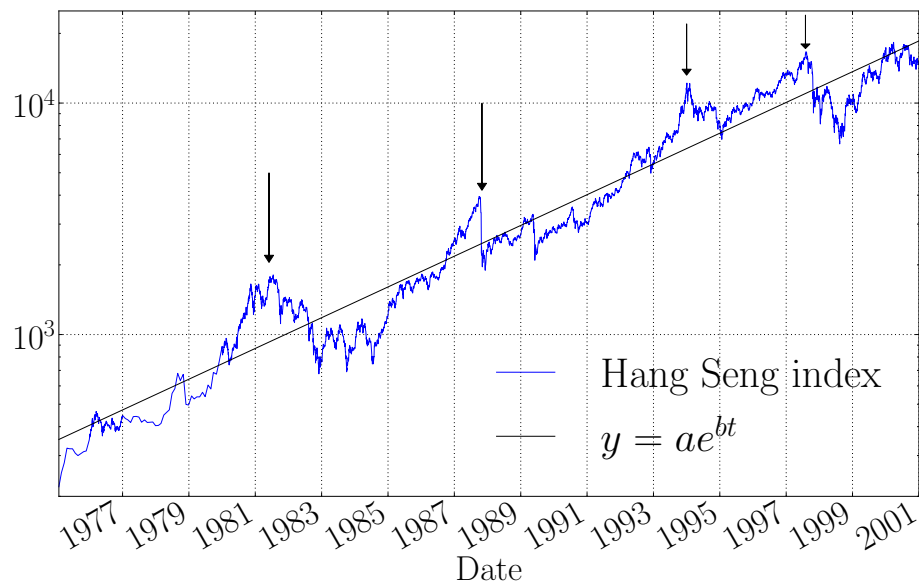


FIGURE 3.1: The Hong-Kong Hang Seng index. The black line represents an exponential fit to the data (straight line on a log-linear plot). One can observe, that while the long-term trajectory of the index can be modeled by an exponential, the price dynamics are constantly overshooting in a super-exponential way until they crash as indicated by the vertical arrows (depicting the biggest crashes of the index).

The Hang Seng index, perfectly exemplifies the fact that, while on the very long-run, price dynamics can be described exponentially (straight line on the log-lin plot), it is clear that this exponential price trajectory is constantly overshoot, corresponding to super-exponential regimes (curvature on log-linear plot), followed by crashes, or at least regime changes.

In addition to the exuberant growth dynamics of the bubble, one can observe the existence of medium-term volatility dynamics decorating the super-exponential price growth. Several mechanisms have been found to be at the origin of these structures (Ide and Sornette, 2002), the most notable one being the hierarchical organization of the network of traders (Sornette and Johansen, 1998) leading to dynamics obeying the symmetry of discrete scale invariance (Sornette, 1998).

Combining these mechanisms with the positive feedback process and imposing that the price should remain finite leads to the so-called log-periodic power law (LPPL) specification for the deterministic component (or expected logarithmic price) of the price dynamics (Sornette et al., 2013):

$$\ln[p(t)] = A + B(t_c - t)^m + C(t_c - t)^m \cos(\omega \ln(t_c - t) - \phi), \quad (3.3)$$

where $A = \ln[p(t_c)]$ is the log price at t_c , B ($B < 0$ for positive bubbles) gives the amplitude of the bubble, $0 < m < 1$ and C determines the amplitude of the oscillations. ω is the angular log frequency determining the scaling ratio of accelerating oscillations and ϕ is a phase embodying a characteristic time scale. An example of an LPPL fit is given in figure 3.2.

3.1.3 Conceptual approach

Because empirical time series such as prices contain many complex unknown features, using trading strategies to extract properties requires a robust null reference, which we took as random strategies, i.e. decisions to buy or sell at random. Because the random strategies are exposed to the same complex patterns as the supposedly LPPL-informed trading strategies, they are exposed to the same bias and same idiosyncratic features. Hence, any significant performance of the LPPL-informed strategies over the random ones signaled a causal relationship between LPPL and the price formation process. This idea can be formulated mathematically as follows: if ϕ_{LPPL} and ϕ_{random} are respectively an LPPL-based and random strategies, then the following statements are true in expectation:

$$\begin{cases} r(\phi_{LPPL}(M)) = r(M) \text{ if } M \text{ is random} & (3.4) \\ r(\phi_{LPPL}(M)) = \omega \neq r(M) \text{ if } M \sim LPPL & (3.5) \\ r(\phi_{random}(M)) = r(M) \forall M & (3.6) \end{cases}$$

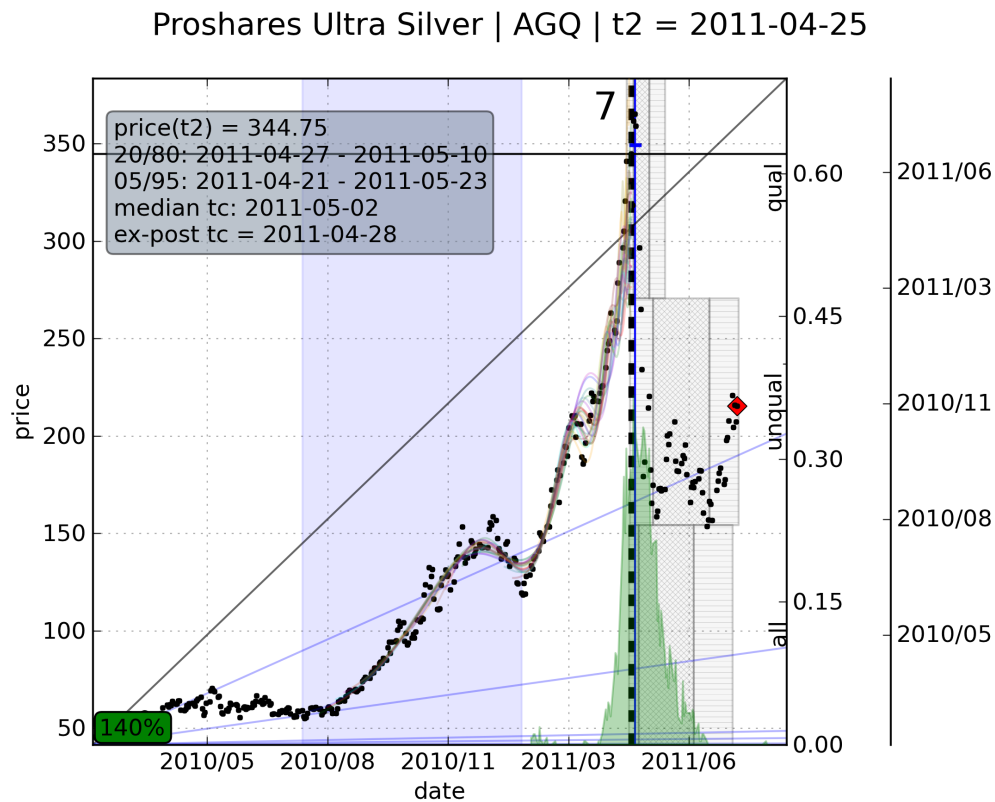


FIGURE 3.2: Example of LPPL fits for AGQ, a silver ETF on the date of observation April 25th, 2011. One can easily see the explosive nature of the timeseries of prices. LPPL fits are shown and their corresponding critical times. Source of the figure: Ryan Woodard.

where M is the time-series of prices of the market and r the annualized return “operator”. Equations 3.4 and 3.5 follow from the martingale condition (no free lunch) and equation 3.6 translates the fact that random strategies have no skill and do not use any pattern that might be existing in the financial market time series. From equations 3.5 and 3.6, it follows that if LPPL-strategies outperform random ones, the market has some structure connected to the LPPL pattern. Trading strategies were thus used as the analogs of “spectrometers” that probe the market, where deviations from the performance of random strategies reveal the existence of information (Zhou et al., 2011). Thus, only the relative performances of the LPPL strategies were relevant within this context; their absolute performance was secondary for our purpose. The strength of this methodology lies in the fact that our tests did not depend on the definition of bubbles, of crash, or of market phase transitions or on any assumption about the underlying process. It should also be noted that the impact of any trading activity on the market dynamics fell outside of the scope of our methodology for the following reason: if a significant number of agents would start applying LPPL-based strategies, they would possibly influence the market dynamics and it is an open question as to whether this

would modify the LPPL patterns that we aim at probing.

3.1.4 The signal

In order to implement LPPL-based strategies, we first needed to create a signal aggregating the information contained in the calibration of the LPPL model to financial prices at different timescales. Not knowing *a priori* which timescales captured a potential bubble dynamic, for any given day t , we fitted equation 3.3 for every interval $[t - \Delta, t]$ with $\Delta \in [20, 21, \dots, 500]$ days, corresponding to one business month to 2 business years. For each day of observation t , we fitted $500 - 20 = 480$ intervals, each of them representing a different time scale. Naturally, not all the fits were relevant, especially outside of a super-exponential regime. In order to distinguish those that were phenomenologically compatible with bubble regimes, we filtered them according to theoretically and empirically motivated criteria. For instance, we wanted $m \in [0, 1]$ (in equation 3.3): $m > 1$ would have lead to a decelerating price, inconsistent with the concept of a super-exponential regime, whereas $m < 0$ would have lead to diverging prices, which is unrealistic. Similarly, we needed $B < 0$ to filter for increasing price and $Bm > C\omega$ to ensure that the probability of a crash remained positive in the rational expectations framework (Graf and Meister, 2003). The remaining criteria were based on empirical observations and were $\omega \in [5, 20]$ and the *number of oscillations until t* should be larger than 2.5 (Johansen and Sornette, 2001). We then defined our signal s_t simply as $s_t = \frac{\text{number of qualified fits}}{\text{number total fits}}$, the fraction of qualified fits, which represented the equally weighted vote among all the timescales of the strength of the super-exponential regime.

Based on the signal, we built simple trading strategies and compared them with two different benchmarks consisting of different ways of randomizing our strategies.

One might wonder why the signal was not simply constructed based on the estimation of t_c , the critical time, of each of the fits since that would have been the most straightforward way of estimating the timing of the crash. The answer lies in a mathematical property of LPPL, namely that t_c is a sloppy parameter (in the mathematical sense) (Brée et al., 2011). This basically means that the quality of the fit (measured by the sum of squares) varies very little even if t_c varies significantly. Figure 3.4) shows the variation of the cost function (sum of squares) as a function of t_c and t_1 . The flatness of the cost function in the t_c direction, i.e. the fact that a change of t_c affects the quality of the fit very little, is evidence of t_c 's sloppiness. Fortunately, there are ways of getting rid of the sloppiness and it is the object of active reasearch.

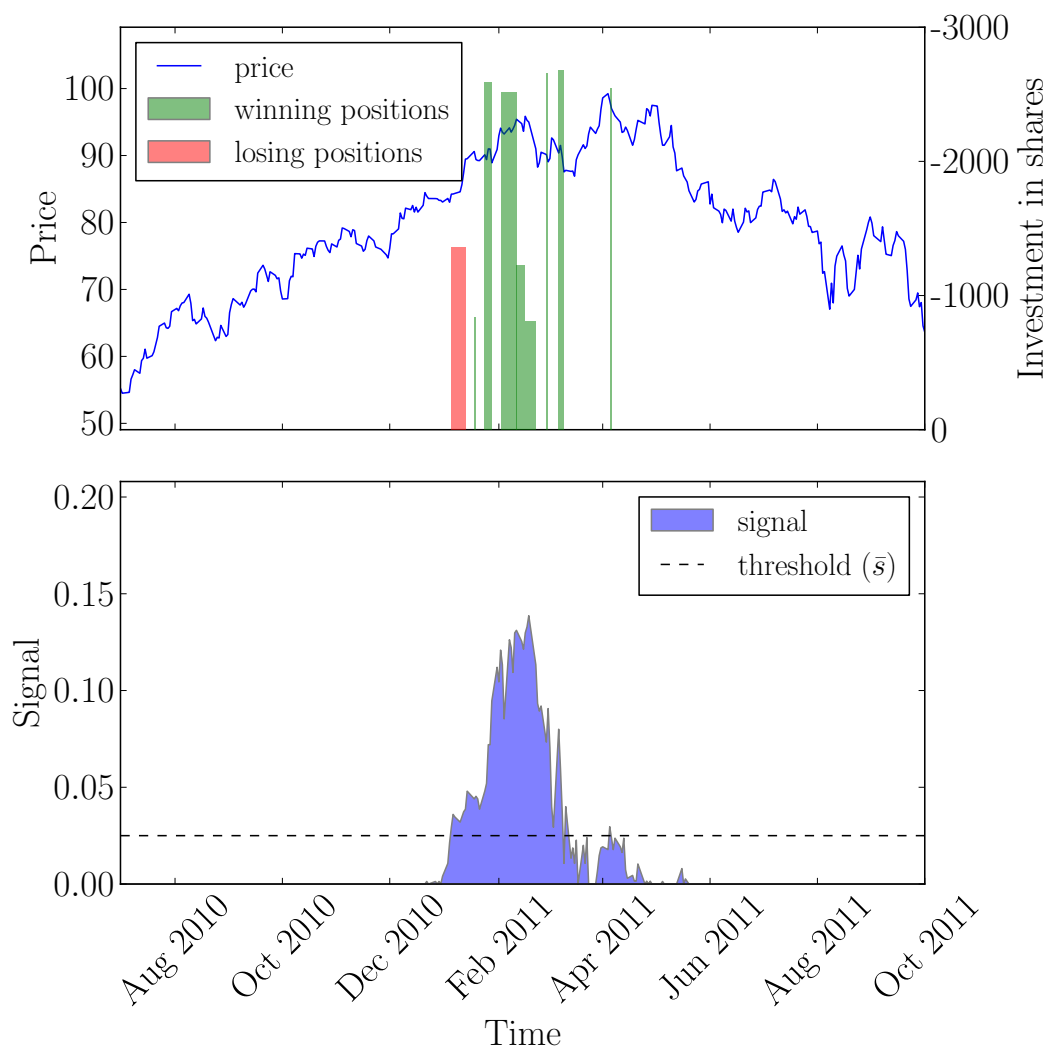


FIGURE 3.3: Example of a strategy applied on Deere and Co (DE). We see a successful example of going short with a signal threshold \bar{s} of 0.025 as manifested by the predominance of green bars on the upper figure. The size of the positions changes due to reallocation of resources on the other assets of the portfolio. Notice that most green bars coincide with the end of a strong price appreciation that the LPPL signal qualifies as a mature bubble ready to burst.

3.1.5 Strategies

The LPPL-based strategies were constructed as follows: when the LPPL signal s_t was above a threshold, \bar{s} , we entered a short position as shown in Figure 3.3. This meant that we expected the price to decrease, as the LPPL signal was detecting a strong super-exponential regime that, due to its unsustainable nature, indicated an imminent correction. We closed the position according to some predefined exit gain / loss (\bar{g} , \bar{l}) and minimum / maximum holding times (h_{min} , h_{max}). We then waited ω_{min} days to open a new position. In practice, we explored all possible combinations of

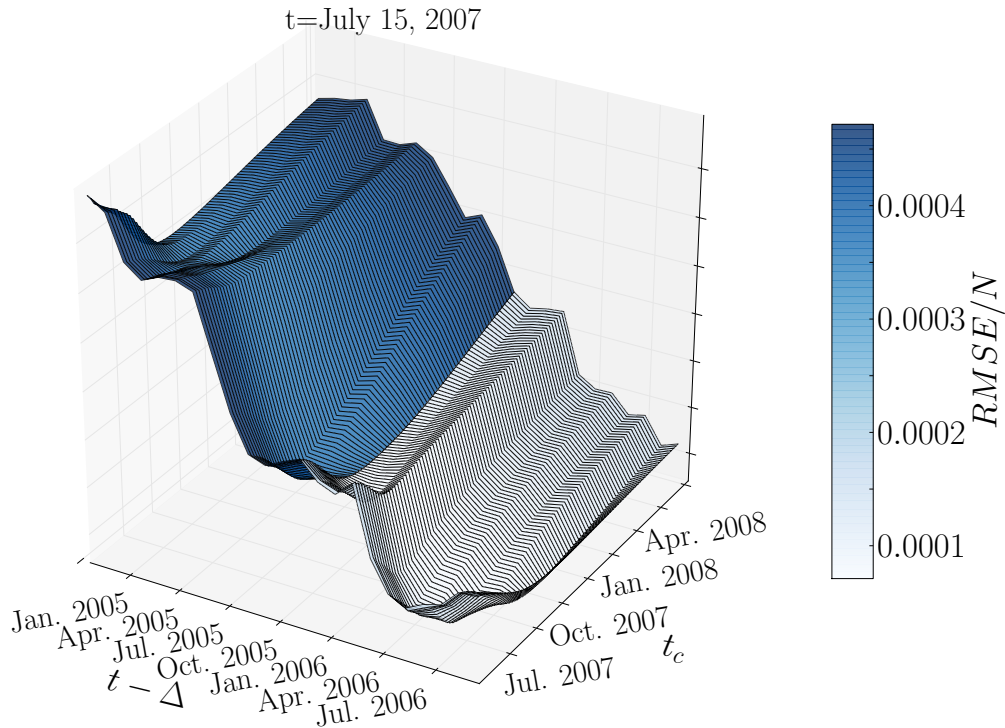


FIGURE 3.4: Normalized cost function for the SP500 as a function of $t - \Delta$, the starting point of the fit, and t_c , the critical time, with the day of observation $t = \text{July 15}^{\text{th}}, 2007$. One can see that the cost function is flat in the t_c direction, indicating that the critical time is a sloppy parameter. Source of the figure: Guilherme Demos.

$\bar{s} \in [0.01, 0.025, 0.05, 0.075, 0.1]$, $\bar{g} \in [1\%, 5\%, 15\%, 50\%]$, $\bar{l} \in [-1\%, -5\%, -15\%, -50\%]$, $h_{min} \in [1, 5, 10]$ trading days, $h_{max} \in [20, 100, \infty]$ trading days and $\omega_{min} \in [1, 5, 10]$ trading days, i.e. 2160 strategies. The parameters were allowed to vary within broad bands to show that the difference in the outcome of the LPPL-based versus random strategies was robust.

The random strategies were constructed in two different ways. In the first one, for each of the 2160 LPPL strategies, the positions taken based on the LPPL signal were shuffled. In other words, for each LPPL position (open-close pair), the entry time of the random position was picked arbitrarily and its duration was set to be the same as the LPPL one. The upside of this method was that the duration and the number of positions was the same in both the LPPL and random cases. Its downside was that the exit strategies were ignored, since the duration of the positions was enforced. The second randomization process consisted in shuffling the LPPL signal (an example of which is shown on the lower panel of Figure 3.3), effectively destroying its structure and applying the strategies as in the LPPL case but on the shuffled signal. While not suffering from the downside of the first method, the number and length of the positions were not

conserved. The two processes were complementary: shuffling the signals ensured that any difference between LPPL and shuffling the positions was not solely due to the exit strategies $(\bar{g}, \bar{l}, h_{min}, h_{max}, \omega_{min})$. While shuffling the positions ensured that any difference between LPPL and shuffling the signal was not solely due to the difference in the number and duration of the positions between the two processes.

3.1.6 Results

Figure 3.5 shows the comparison between LPPL strategies and their two random counterparts in the five-year period starting on January 1st 2008 and ending on December 31st 2012. The strategies were not applied on a single asset, but rather on a portfolio of 50 assets chosen randomly among the SP500 constituents. We see a clear difference between the performances of the LPPL-based strategies and the random ones, in that the annualized returns of the LPPL-based strategies are greater than that of their random counterparts: the vast majority of the points on figure 3.5 lying above the $x = y$ line representing equal performance. Indeed the LPPL-based strategies outperformed the outcome of the shuffled positions and shuffled signal process 92% and 96% of the time respectively. In other words, the portfolio performance metrics of LPPL strategies were significantly and consistently different from those of random strategies in this time period, strongly suggesting that LPPL contains information.

Proving that LPPL-based strategies outperform their random counterpart in a given time period was not sufficient to make a general statement. In fact, asset price dynamics can be very different depending on the time period chosen. For instance, during the 2008 financial crisis, the stock market went down as opposed to the 2003-2006 period. This motivated us to compare the LPPL-based results with the random ones in different time periods. Concretely, we chose the non-overlapping 10 yearly periods, 5 two-year periods, 2 five-year periods and the ten-year period from January 1st 2003 until December 31st 2012. Moreover, in order to show that the deviation from randomness was robust with respect to the choice of the 50 assets portfolio on which the 2160 strategies were applied, for each time period, we ran the strategies on 10 different portfolios of 50 assets chosen randomly among the SP500 constituents.

Figure 3.7 shows the extension of the procedure reported in figure 3.5 to all the time periods and portfolios of 50 assets described previously. The cubes represent the fraction of strategies that performed better in the LPPL case than in the shuffled positions ones in a given time period. The error bars on the yz projection result from repeating the comparison on 10 different portfolios of 50 assets. We can see that LPPL shorting clearly outperformed the outcome of the two randomized processes in most of the time

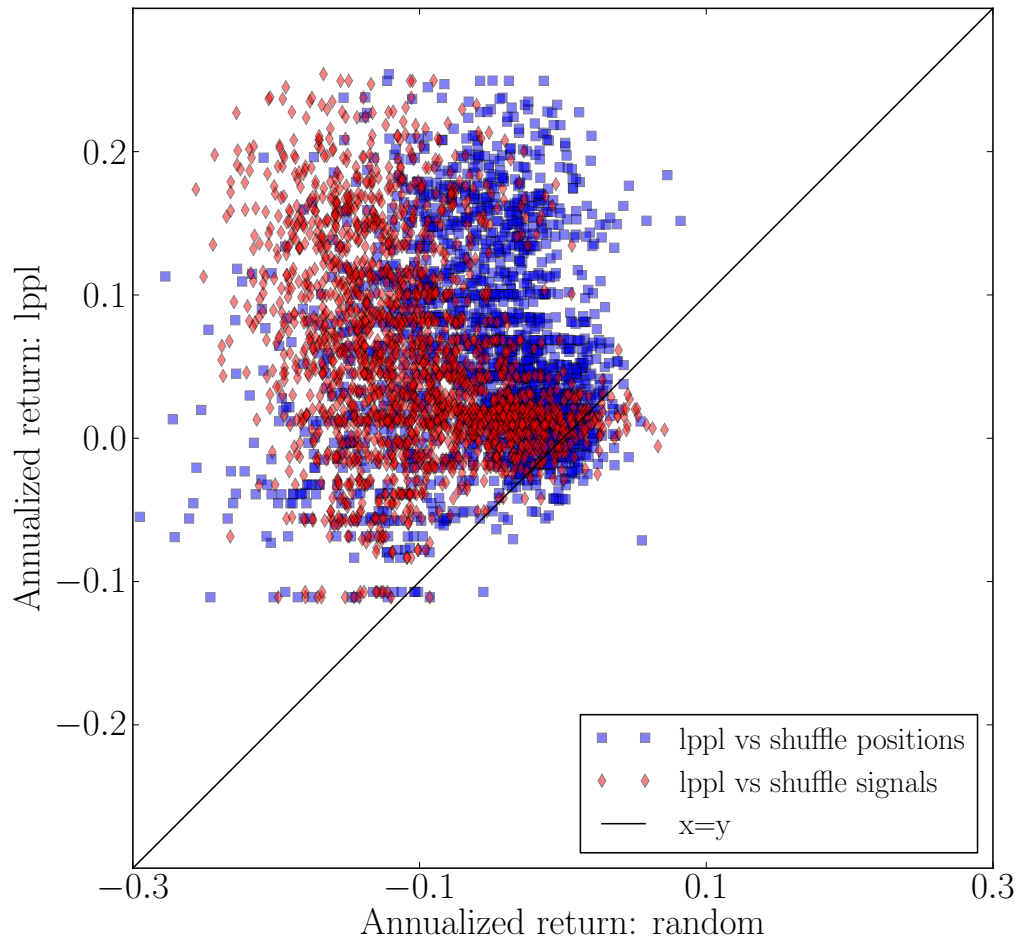


FIGURE 3.5: Annualized returns of the LPPL-based strategies vs the random ones on a basket of 50 assets between January 1st 2008 and December 31st 2012. Each point represents the annualized return of a single strategy.

periods and choices of assets, confirming LPPL’s predictive nature. However, contrary to naive expectations, the few time periods in which the LPPL strategies performed similarly to their random counterparts is between 2008 and 2009, i.e. during the financial crash and the ensuing recession. This seemingly unintuitive behavior has a straightforward explanation: by construction, the signal was built to detect positive bubbles. However, super-exponential regimes of positive price growth are by definition rare during a financial crash as shown by figure 3.6. As such, it is not surprising to see our signal getting very close to “white noise” during such a period. There are ways to take into account the so-called negative bubbles (Yan et al., 2012), but that is beyond the scope of this work.

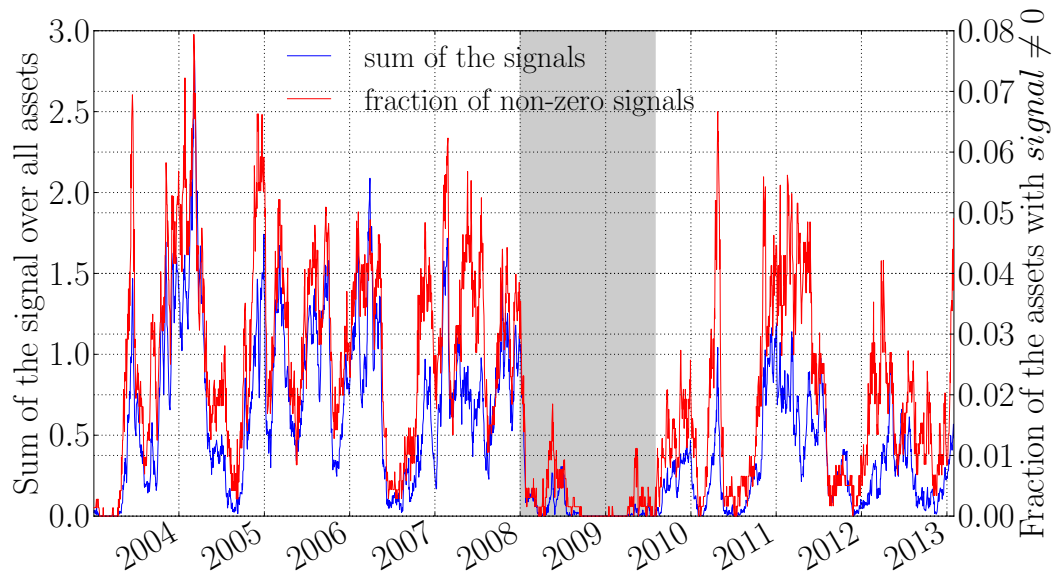


FIGURE 3.6: Two measures of the density of LPPL-signals over time: 1) the sum of the signals over all assets as a function of time and 2) the fraction of assets with non-zero signal over time. The signal indicating a strong super-exponential regimes, the very low density of signals during the years 2008 and 2009, as indicating by the grey rectangle, confirms the absence of positive bubbles during times of crises.

3.1.7 Conclusions

In summary, we have applied a hybrid methodology to financial markets, marrying trading strategies with the log-periodic power-law model which takes its roots in critical phenomena. Our results have clearly demonstrated that the outcome of LPPL strategies persistently outperforms that of the random ones, in other words they have predictive nature. This work supports a view in which financial markets are inherently unstable, out of equilibrium, a view dramatically opposed to the consensus in classical finance and economics.

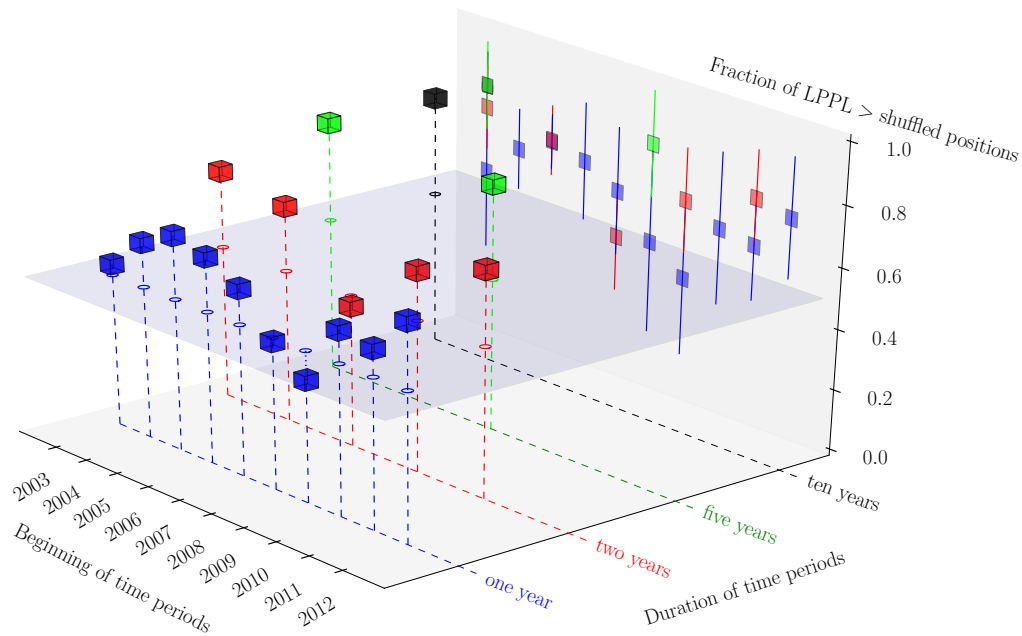


FIGURE 3.7: Each polygon represents the fraction of LPPL > shuffle positions of a cloud plot of figure 3.5. The vertical bars on the yz projections represent the standard deviation of the fraction of LPPL > shuffle positions over 10 different choices of assets. In the vast majority of cases, LPPL is above the 50% plane and thus significantly differs from shuffle positions. Although LPPL vs shuffle signals is not shown here, the results are qualitatively the same.

3.2 The Γ factor

² We have seen in section 3.1 that the log-periodic power law model (LPPL) has predictive power. This was achieved by building strategies on top of LPPL-derived signals and showing that their outcome was persistently different from the one of random strategies. The strategies had several parameters and while the random strategies were implemented in such a way as to leave no doubt about the LPPL-origin of the results, it was hard to identify, based on that framework, in which feature(s) of LPPL did the predictive power reside. We hypothesize that the main feature responsible for LPPL's predictive power is the super-exponential price increase. To test this hypothesis in the simplest way possible, we proceeded to quantify the impact of the first difference of returns of an asset on its future returns in a factor regression framework. Indeed, the basis of a super-exponential price increase is the fact that the growth rate grows itself. One simple way to translate this idea into return space is to look at the first-difference of returns, or in other words the difference between two successive growth rates. After introducing the most popular factor models and explaining the methodology behind them, we proceed to show that the constructed measure on the first difference of returns, which we call the

2. This section is based on work in progress performed with Diego Ardila and Didier Sornette.

Γ factor, can be exploited to create profitable portfolio-based strategies and contains significant information to explain realized returns in the U.S stock market. Consequently, it supports the idea that positive feedbacks play an important role in market price dynamics. This section is based on work in progress performed with Diego Ardila and Didier Sornette.

3.2.1 Introduction

Understanding what factors explain an asset's returns has always been the central question of financial economics. [Markowitz \(1952\)](#) with his portfolio theory laid the foundations of the field. He realized that a portfolio P of N assets with returns being distributed according to a Gaussian with mean μ , volatility σ , and a correlation structure ρ , has a higher ratio $\frac{\mu_P}{\sigma_P}$ than the individual assets that constitute it (as long as $\rho < 1$): in economic terms, a higher return for a lower risk. This is simply due to the fact that the sum of N gaussian random variables scales linearly with N while the standard deviation only scales as \sqrt{N} . Intuitively, the N random variables, not moving in perfect synchronisation, cancel part of each other's noise out. This result can be extended to the case where $\mu \rightarrow \mu_i$, $\sigma \rightarrow \sigma_i$ and $\rho \rightarrow \rho_{i,j}$, i.e. when the assets returns don't come from the same Gaussian distribution. In addition, Markowitz showed that for any given level of risk, σ_P , there is a combination of weights for the stocks composing the portfolio yielding the highest possible return, r_P . This is called the Efficient Frontier. As the name suggests, it would be inefficient for anyone to buy a portfolio with (r_P, σ_P) not on the efficient frontier, since there would be another portfolio which would yield higher returns for the same level of risk (see figure 3.8).

By introducing the possibility of buying or selling the risk-free asset (usually proxied by US treasury bills), the effective efficient frontier becomes a linear function with intercept equal to the risk-free rate, r_f , and tangent to the hyperbolic curve representing the efficient frontier without the possibility of trading the risk-free asset (see figure 3.8). This linear function is also called the Capital Allocation Line. As a consequence, it is inefficient to take a position in anything else than a mixture of the risk-free asset and the tangential portfolio: the left-side of figure 3.8 corresponds to buying the risk-free asset (in other words lending at the risk-free rate) and the tangential portfolio, while the right side corresponds to shorting the risk-free asset (borrowing at the risk-free rate) and buying with leverage the tangential portfolio. Observing that the only portfolio that should be chosen is the tangential portfolio (mixed with risk-free asset) leads to the conclusion that the tangential portfolio *is* the market portfolio and has return r_M and volatility σ_M .

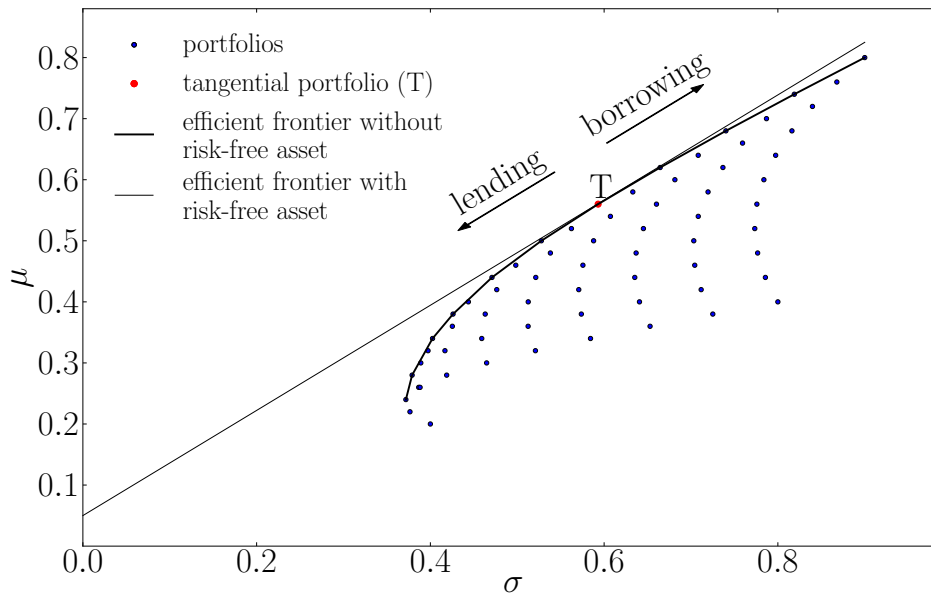


FIGURE 3.8: Markowitz portfolio theory. Each blue circle represents a portfolio of the same 3 assets but with different weighting. The 3 assets have their own μ_i , σ_i and correlation structure. The thick black line represents the efficient frontier without borrowing: it is the line characterizing the highest return one can get for a given level of risk by combining the 3 assets. The thin black line represents the efficient frontier with the possibility of borrowing/lending the risk-free assets. It is tangent to the market portfolio represented by the red circle.

Sharpe (1963) and Lintner (1965), building on Markowitz's results, showed that not only is there a linear relationship between the risk and the return of any combination of the risk-free asset and the market portfolio, but that the linear relationship extends to the risk/return of the individual assets constituting the market (equation 3.7):

$$(r_i - r_f) = \alpha_i + \beta_i(r_M - r_f), \quad (3.7)$$

where r_i is the return of asset i , r_f is the risk-free rate, $\beta = \frac{Cov(r_i, r_M)}{\sigma_m}$ and can be understood as the part of the risk of asset i that cannot be diversified away, and r_M is the return of the market. This model, coined the Capital Asset Pricing Model (CAPM) became a pillar of economic theory, since it was the first time that a quantitative relationship was proposed between the risk and the return of an individual asset as well as an explanation for the source of the risk (the exposure to the market). In essence, the higher the covariance of the stock with the market, the higher should its risk premium be to compensate for the extra risk.

One of the most important challenges to the CAPM came from [Fama and French \(1992\)](#) who showed on empirical grounds that the market factor was not sufficient to explain the cross-sectional returns of assets. They proposed to add two additional factors to the model: the size factor accounting for the fact that smaller firms have historically had higher returns than bigger firms, and the book-to-market factor associated with the observation that firms with a higher book-to-market ratio had higher returns than those with lower book-to-market ratio. This model is commonly known as the 3-factor Fama-French model (equation 3.8):

$$(r_i - r_f) = \alpha_i + \beta_i(r_M - r_f) + s_iSMB + bm_iHML, \quad (3.8)$$

where, in addition to the variables in equation 3.7, SMB and HML are the size and book-to-market and s_i and bm_i are the corresponding factor loadings. While both the CAPM and the 3-factor model are factor regression models, there is an important conceptual difference: CAPM is a theoretical result derived from “first principles”, while the 3-factor model is empirical insofar as the size and book-to-market factors do not have a formal theoretical underlying. In this sense, the 3-factor model belongs to the family of Arbitrage Pricing Theory models ([Ross, 1976](#)).

The Fama-French model has been very successful at explaining stock returns, but has failed to explain the fact that stocks that performed the best (the worst) over three to 12 months periods tend to continue to perform well (poorly) over the subsequent three to 12 months ([Jegadeesh and Titman, 1993](#)). Momentum in stock returns is by now the most studied anomaly based purely on the dynamics of prices. It has been detected in the US, Europe and Asia Pacific ([Fama and French, 2012](#)) as well as in different asset classes ([Asness et al., 2013](#)). Additionally, it has withstood explanations based on industry effects, cross-correlation among assets, and data mining ([Grundy and Martin, 2001](#), [Jegadeesh and Titman, 2001](#)).

Motivated by the momentum anomaly, [Carhart \(1997\)](#) introduced an additional momentum factor (equation 3.9).

$$(r_i - r_f) = \alpha_i + \beta_i(r_M - r_f) + s_iSMB + bm_iHML + \delta_i\Delta, \quad (3.9)$$

with Δ being the momentum factor and δ_i the factor loading.

While [Jegadeesh and Titman \(1993\)](#) and [Carhart \(1997\)](#) showed that momentum (velocity in a physical sense), as a strategy and as a factor has significant explanatory power, we empirically studied the first difference of returns Γ , analogous to the acceleration of prices, as an important source of predictability and of risk premium. To support

our claim, we have conducted a twofold analysis. On one hand, we documented the profitability of portfolio-based strategies based on our measure. On the other hand, we enhanced the Fama-French model with our Γ factor and analyzed its performance explaining portfolio returns.

3.2.2 Methodology

As mentioned above, we wanted to construct a factor that was sensitive to the super-exponential behavior of assets, the distinctive feature of non-sustainable regimes and of bubbles, documented in section 3.1. To that end, we chose to use the simplest measure encapsulating the idea of acceleration or a change in momentum: the first difference of returns.

Formally, we define $\Gamma_t(K)$ as follows:

$$\Gamma_t(K) = r_t(K) - r_{t-K}(K), \quad (3.10)$$

where $r_t(K)$ is the K -months return of the corresponding asset at the end of month t and K is omitted if we refer to Γ based on yearly returns. For example, if $t = \text{March } 31^{\text{st}} 2005$, and $K = 12$, then $t - 12 = \text{March } 31^{\text{st}} 2004$, $\Gamma_t = r_t(12) - r_{t-12}(12)$, $r_t(12) = \frac{p_t}{p_{t-12}} - 1$ and $r_{t-12}(12) = \frac{p_{t-12}}{p_{t-24}-1}$.

The $\Gamma(K)$ strategy is then constructed as follows:

1. At time t , the end of the current month, we sort the stocks according to $\Gamma_t(K)$, the first difference in the K -months returns.
2. Starting at time $t + \epsilon$, i.e. the beginning of next month, we go long the top decile and short the bottom decile of the assets until $t + 1$, the end of next month.
3. The profit of $\Gamma(K)$ at time $t+1$ is then simply the return of the long-short strategy between time $t + \epsilon$ and time $t + 1$.

In nutshell, we create the portfolios based on the first difference of K -months returns, re-balancing monthly. The idea behind the methodology is quite simple: if indeed there is a positive relationship between the first differences of the K -months returns, and the next month monthly return r_{t+1} , longing the top decile and shorting the bottom decile should result in a positive return for a strategy. Furthermore, the long-short strategy *is* the portfolio mimicking $\Gamma(K)$, namely the portfolio that proxies the effect of the first difference in the K -months returns at t on the monthly return at $t + 1$.

3.2.2.1 Data

The data used in this work were all the common stocks on the Center for Research in Security Prices (CRSP) database (share code 10 and 11) between May 1963 and December 2013. We have limited the analysis to this period since the Fama-French factors require accounting data only available after 1963. Stocks with less than two years of existence (by time t) were discarded to control for survival bias. Following [Fama and French \(1992\)](#), the computation of the breakpoints of the different portfolios was done using stocks from the NYSE, the NASDAQ and the AMEX (exchange code 1, 2 and 3). The measures were then computed using stocks from all exchanges. Finally, to study the stock returns by industry, we employed the Fama-French industry classification, which in turn is based on the stock's SIC code.

3.2.3 Results

3.2.3.1 Γ on different time horizons

We first analyzed the performance of $\Gamma(K)$ for $K \in [1, 3, 6, 9, 12, 24, 60]$. [Figure 3.9](#) presents the cumulative returns $\sum_0^t \Gamma_t(K)$ up to time t , corresponding to the return one would get without reinvesting profits. We can see that $\Gamma(K)$ exhibits very positive returns when $K = [9, 12, 24]$, peaking when using yearly returns. In contrast, the strategy yields very negative returns over short formation periods $K = [1, 3, 6]$, bottoming when $K = 6$.

Interestingly, whereas poorly-performing Γ -strategies are (approximately) monotonically decreasing in their cumulative returns, good-performing Γ -strategies exhibit periods of remarkable stagnation, in which cumulative returns neither increase or decrease. For example, the cumulative returns of $\Gamma(K)$ between 1994 and 1999 remain virtually constant. Additionally, we do not observe any trend reversal on the performance of the strategies, i.e. good strategies do not turn badly, while bad strategies continue their negative trend. As these results suggest that accelerating prices over medium term horizons forecast increasing returns, we focused on $K = 12$ for the rest of the study.

3.2.3.2 Γ vs the standard factors

[Figure 3.10](#) presents the cumulative returns up to time t of the Γ -strategy against that of a momentum Δ -strategy using yearly returns, as well as against the cumulative returns of size and B/M strategies. Similar to Γ , all the strategies are based on deciles, which buy the top 10% and sell the bottom 10%.

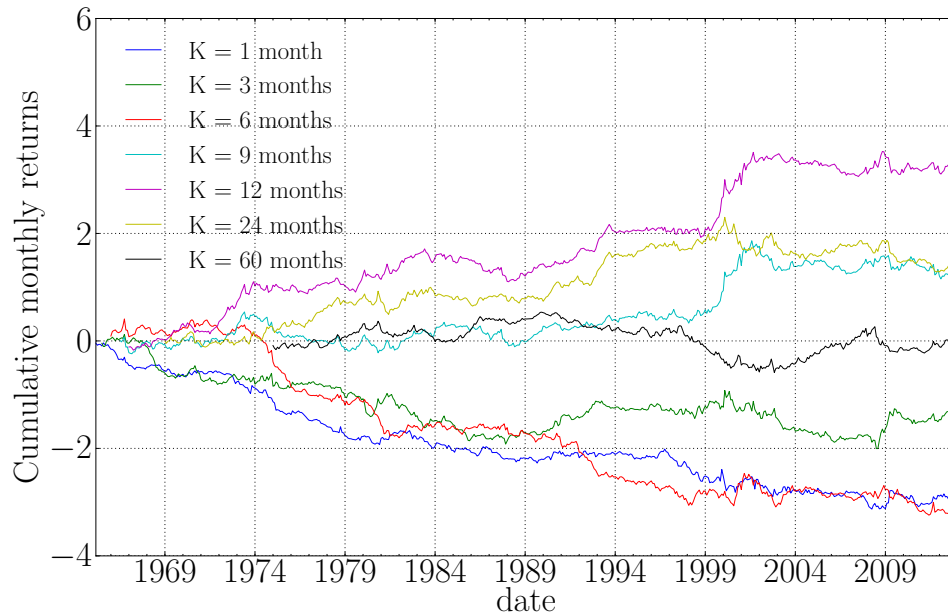


FIGURE 3.9: Cumulative monthly returns of $\Gamma(K)$, i.e. $\sum_0^t \Gamma_t(K)$, for K ranging from 1 month to 5 years. We notice a peak in performance at $K = 12$ months.

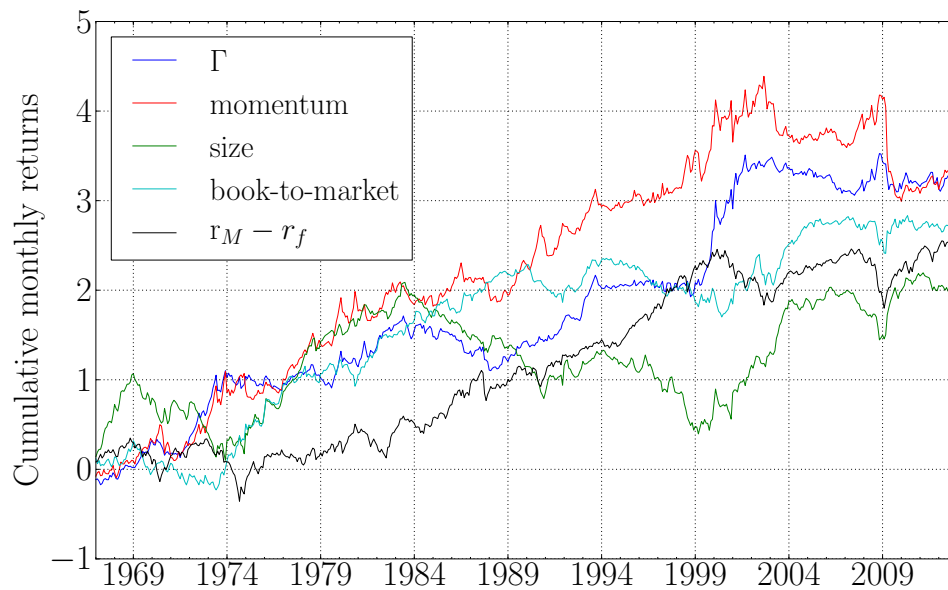


FIGURE 3.10: Cumulative monthly returns, i.e. $\sum_0^t f_t^i$ of the factors. For the factors to be comparable, each of them were constructed on the difference in returns between the top and the bottom decile portfolios, except the market factor. Notice that this is not the standard way to compute the size and the book-to-market factors (Fama and French, 1996). The cumulative monthly returns show the strong and stable performance of Γ .

	$\mu(\Gamma)$	$\mu(\text{momentum})$	$\sigma(\Gamma)$	$\sigma(\text{momentum})$	$\frac{\mu}{\sigma}(\Gamma)$	$\frac{\mu}{\sigma}(\text{momentum})$
utilities	0.0027	0.00046	0.052	0.065	0.052	0.0071
telecom	0.0043	0.0081	0.12	0.17	0.036	0.048
retail	0.0056	0.0068	0.07	0.11	0.08	0.064
non-financial	0.01	0.0085	0.061	0.089	0.17	0.095
manufacturing	0.0031	0.0038	0.07	0.093	0.044	0.041
healthcare	0.0039	0.00037	0.094	0.13	0.041	0.0029
finance	0.0028	0.0086	0.069	0.1	0.041	0.083
energy	0.0065	0.0028	0.088	0.11	0.074	0.026
cons. non-durable	0.0016	0.0056	0.075	0.11	0.022	0.051
cons. durable	0.0051	0.014	0.1	0.13	0.05	0.11
chemical	0.0033	0.0027	0.089	0.11	0.037	0.024
business	0.011	-0.00033	0.082	0.11	0.13	-0.0029
all exchanges	0.0087	0.012	0.06	0.092	0.15	0.13

TABLE 3.1: Performance of the Γ and momentum factors across sectors. Bold figures in the last two columns emphasize which of Γ or momentum had the highest sharpe ratio.

We can observe that Γ 's cumulative returns is higher than the cumulative returns of the size and B/M strategies and even tops the performance of Δ towards the end of the sample. Moreover, Γ 's Sharpe-ratio (the ratio of excess return to volatility) computed over the whole sample is 0.116, which is higher than that of Δ (0.086). This is consistent with the more stable trajectory of Γ 's cumulative returns, already visible from figure 3.10. In addition, momentum strategies have been known to have performed poorly in the post-2005 period in contrast to Γ which didn't experience any major losses. We interpret Γ 's higher sharpe ratio as indication that it is more robust to the various changes of regimes that have happened in the last 40 years than momentum.

3.2.3.3 Performance by industry

To study potential differences in performance among different industries, we analyzed the Γ and Δ strategies for different sectors. The results are presented in the form of summary descriptive statistics that can be found in table 3.1. Not only are the Sharpe ratios of the Γ factors of the different sectors higher 9 times out of 13 than those of momentum (Δ), but their mean over the sectors are higher and dispersion lower. We interpret this as an indication that Γ accounts better for the cross-sectional returns of assets, since its performance depends less than that of Δ on the particular sectors used to compute it.

3.2.3.4 The Fama-French + Γ factor model

Motivated by the performance of the gamma-strategy, we constructed the 4-factor model described in equation 3.11 using the 3-factor Fama-French model plus Γ .

$$(r_i - r_f) = \alpha_i + \beta_i(r_M - r_f) + s_iSMB + bm_iHML + \gamma_i\Gamma, \quad (3.11)$$

Our aim was to assess the performance of Γ as a factor, to study whether Γ is a source of additional risk premium. In order to do so, we compared the explanatory power of our model relative to that of the Fama-French + Δ model via their ability to describe Γ and Δ -sorted portfolios. We focused on these portfolios since it is already known that the Fama-French model by itself is able to explain well portfolios based on alternative rules (i.e. size, B/M, β , etc.). Hence, if Γ captures any discernible source of risk premium, this should be evident from the Γ -portfolios designed to isolate a prevailing anomaly.

Using the GRS statistic, we tested the null hypothesis $\alpha_i = 0, \forall i = 1..N$ at a 5% significance level. Rejecting the null hypothesis implies that the model presents significant pricing errors and is unable to explain the returns of the portfolios. Results can be found in tables 3.2, 3.3, 3.4, where in addition to the regressions of the Fama-French + Γ and Fama-French + Δ models, we have added the results for CAPM as a reference.

The results do not suggest that Γ is an additional source of risk premium. On one hand, the Fama-French + Γ model cannot explain the Δ -sorted portfolios; the p-value of the GRS statistic allows us to reject the null hypothesis at a 0.05 significance level. On the other hand, the Fama-French + Δ model seems to explain the Γ -sorted portfolios, as we are unable to reject the corresponding null hypothesis, as seen in table 3.5. Hence, Γ , at least in its current form, can not be considered as an additional source of risk premium.

3.2.3.5 The correlation between Γ and momentum

To understand why Γ seems to be captured by momentum, we looked at the correlation between the factors (table 3.6). The highest correlation among the factors is precisely that between Γ and Δ (0.69). This suggests that the two factors necessarily have a similar behavior.

The roots of the high correlation between Γ and momentum can probably be found in the correlation between the quantities used to build the factors, $r_t - r_{t-12}$ and r_t respectively. It turns out that contrary to what one might intuitively expect, there is a strong correlation between $r_t - r_{t-12}$ and r_t (assuming r_t to be distributed as a Gaussian):

	α		β		r^2
	μ	σ	μ	σ	
1	-0.00	0.00	1.24	0.03	0.77
2	-0.00	0.00	1.04	0.02	0.83
3	0.00	0.00	0.99	0.02	0.84
4	0.00	0.00	0.90	0.02	0.81
5	0.00	0.00	0.90	0.02	0.83
6	0.00	0.00	0.88	0.02	0.82
7	0.00	0.00	0.92	0.02	0.85
8	0.00	0.00	0.91	0.02	0.81
9	0.00	0.00	0.98	0.02	0.82
10	0.00	0.00	1.08	0.03	0.69
GRS			2.87		
p-value			0.00		

TABLE 3.2: Regression of the 10 Γ decile portfolios on CAPM (equation 3.7). Applying the GRS statistic to test the joint hypothesis that $\alpha_i = 0$ leads to its rejection since $p = 0.0$.

	α		β		s		bm		r^2
	μ	σ	μ	σ	μ	σ	μ	σ	
1	-0.00	0.00	1.14	0.03	0.15	0.04	-0.26	0.05	0.79
2	-0.00	0.00	1.05	0.02	-0.04	0.03	0.00	0.03	0.83
3	0.00	0.00	1.00	0.02	-0.05	0.03	-0.01	0.03	0.85
4	0.00	0.00	0.96	0.02	-0.10	0.03	0.14	0.03	0.83
5	-0.00	0.00	0.96	0.02	-0.09	0.02	0.17	0.03	0.85
6	-0.00	0.00	0.95	0.02	-0.06	0.02	0.23	0.03	0.85
7	0.00	0.00	0.98	0.02	-0.06	0.02	0.19	0.03	0.87
8	0.00	0.00	0.93	0.02	-0.03	0.03	0.09	0.03	0.81
9	0.00	0.00	0.99	0.02	0.06	0.03	0.11	0.03	0.82
10	0.00	0.00	0.96	0.03	0.37	0.04	-0.06	0.05	0.73
GRS					2.21				
p-value					0.02				

TABLE 3.3: Regression of the 10 Γ decile portfolios on the 3-factor model (equation 3.8). Once again the null hypothesis of $\alpha_i = 0$ is rejected with $p = 0.02 < 0.05$.

	α		β		s		bm		δ		r^2
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	
1	0.00	0.00	1.07	0.03	0.11	0.04	-0.33	0.04	-0.21	0.02	0.83
2	0.00	0.00	1.00	0.02	-0.06	0.03	-0.05	0.03	-0.14	0.01	0.86
3	0.00	0.00	0.96	0.02	-0.07	0.02	-0.05	0.03	-0.10	0.01	0.87
4	0.00	0.00	0.94	0.02	-0.11	0.03	0.12	0.03	-0.06	0.01	0.83
5	-0.00	0.00	0.95	0.02	-0.09	0.02	0.15	0.03	-0.04	0.01	0.85
6	-0.00	0.00	0.95	0.02	-0.07	0.02	0.23	0.03	-0.01	0.01	0.85
7	0.00	0.00	1.00	0.02	-0.05	0.02	0.21	0.03	0.05	0.01	0.87
8	0.00	0.00	0.97	0.02	-0.01	0.03	0.13	0.03	0.11	0.01	0.84
9	0.00	0.00	1.04	0.02	0.09	0.03	0.16	0.03	0.13	0.01	0.85
10	0.00	0.00	1.08	0.02	0.44	0.03	0.05	0.04	0.32	0.01	0.87
GRS							1.62				
p-value							0.10				

TABLE 3.4: Regression of the 10 Γ decile portfolios on the 4-factor model (3-factor model + momentum, equation 3.9). This time the null hypothesis of $\alpha_i = 0$ cannot be rejected ($p = 0.1 > 0.05$): the 4-factor model explains the returns of the Γ portfolios.

	α		β		s		bm		γ		r^2
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	
1	-0.00	0.00	1.23	0.04	0.64	0.05	0.25	0.06	-0.60	0.03	0.80
2	-0.00	0.00	1.09	0.03	0.28	0.04	0.25	0.04	-0.44	0.02	0.82
3	0.00	0.00	1.04	0.02	0.11	0.03	0.34	0.04	-0.28	0.02	0.82
4	0.00	0.00	0.95	0.02	0.12	0.03	0.25	0.03	-0.20	0.02	0.83
5	-0.00	0.00	0.93	0.02	-0.02	0.02	0.21	0.03	-0.13	0.01	0.86
6	-0.00	0.00	0.92	0.02	-0.05	0.02	0.15	0.03	-0.05	0.02	0.84
7	-0.00	0.00	0.96	0.02	-0.10	0.02	0.16	0.03	-0.00	0.01	0.86
8	0.00	0.00	0.96	0.02	-0.11	0.03	0.08	0.03	0.09	0.02	0.82
9	0.00	0.00	0.98	0.02	-0.07	0.03	-0.05	0.03	0.19	0.02	0.84
1	0.00	0.00	1.05	0.03	0.21	0.03	-0.31	0.04	0.39	0.02	0.84
GRS					2.10						
p-value					0.02						

TABLE 3.5: Regression of the 10 Δ decile portfolios on the 3-factor + Γ model (equation 3.11). The null hypothesis of $\alpha_i = 0$ is rejected ($p = 0.02 < 0.05$). We conclude that Γ cannot explain the returns of the momentum portfolios.

	market	size	book-to-market	momentum	Γ
market	1.00	0.36	-0.36	-0.21	-0.15
size	0.36	1.00	-0.27	-0.14	0.06
book-to-market	-0.36	-0.27	1.00	-0.04	0.13
momentum	-0.21	-0.14	-0.04	1.00	0.69
Γ	-0.15	0.06	0.13	0.69	1.00

TABLE 3.6: Correlation between factors. One can observe a very large correlation between Γ and momentum.

$$Corr(r_t, r_t - r_{t-12}) = \frac{\mathbb{E}[(r_t - \mathbb{E}[r_t])(r_t - r_{t-12} - \mathbb{E}[r_t - r_{t-12}])]}{\sigma_{r_t} \sigma_{r_t - r_{t-12}}} = \frac{\sigma_{r_t}}{\sigma_{r_t - r_{t-12}}} = \frac{1}{\sqrt{2}} \quad (3.12)$$

Numerically, $\frac{1}{\sqrt{2}} \sim 0.71$. While it is reasonable to assume that a high correlation between r_t and $r_t - r_{t-12}$ leads to a high correlation between the factors, it is not trivial to assess why correlation during the formation period translates into correlation to the holding periods. This is due to the fact that determining the mechanism by which $Corr(r_t, r_t - r_{t-12})$ relates to $Corr(\Delta, \Gamma)$ requires an explicit model for the emergence of momentum which goes beyond the scope of this work.

Nevertheless, one possible way to go forward and to continue exploring a factor in the spirit of Γ but not explained by Δ is to use a quantity similar to $r_t(12) - r_{t-12}(12)$ but less correlated to $r_t(12)$ by construction. The term in $r_t(12) - r_{t-12}(12)$ responsible for the correlation with $r_t(12)$ is $r_t(12)$ itself. Hence, it is straightforward to show that by shifting the first term in $r_t(12) - r_{t-12}(12)$, the correlation with $r_t(12)$ is decreased.

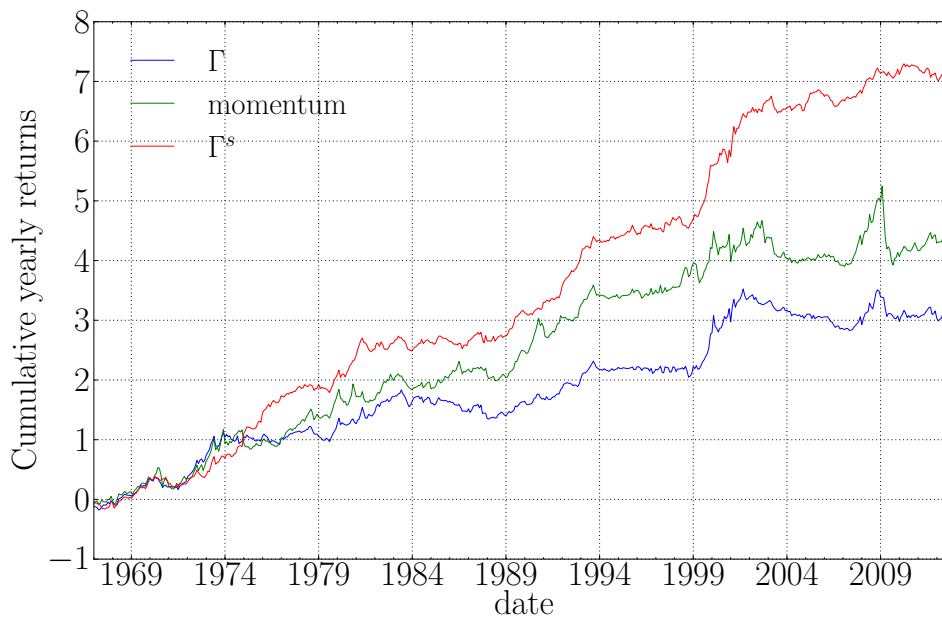


FIGURE 3.11: Cumulative monthly returns of momentum, Γ and Γ^s . Γ^s , the factor constructed based on the decorrelating procedure between Γ and momentum strongly outperforms both of them.

This reasoning led us to study Γ^s , the shifted Γ , constructed on the first difference of half-yearly returns shifted by 6 months: $r_{t-6}(6) - r_{t-12}(6)$. The correlation between Γ^s and momentum decreased to 0.39 (from 0.69 between Γ and momentum), by almost a factor of 2.

3.2.3.6 The Γ^s factor

Figure 3.11 shows the cumulative returns of Γ^s . This factor, designed to reduce its correlation with momentum, went beyond our expectations. Not only does its performance exceeds that of Γ and Δ , but as figure 3.12 shows, when constructing double-sorted portfolios, the average returns seem to vary only along the Γ^s deciles. This suggests that Γ^s explains momentum.

3.2.3.7 The Fama-French + Γ^s factor model

Analogously to Γ , we have regressed the returns of the Γ^s -portfolios on the CAPM, the Fama-French 3-factor model and the Fama-French + Γ^s model. Tables 3.7, 3.8, 3.9 show the results. The null hypothesis of all the regressions intercepts is rejected for all three models, implying that none of the standard factors (market, size, book-to-market and momentum) are enough to explain the returns generated by the Γ^s -portfolios. In

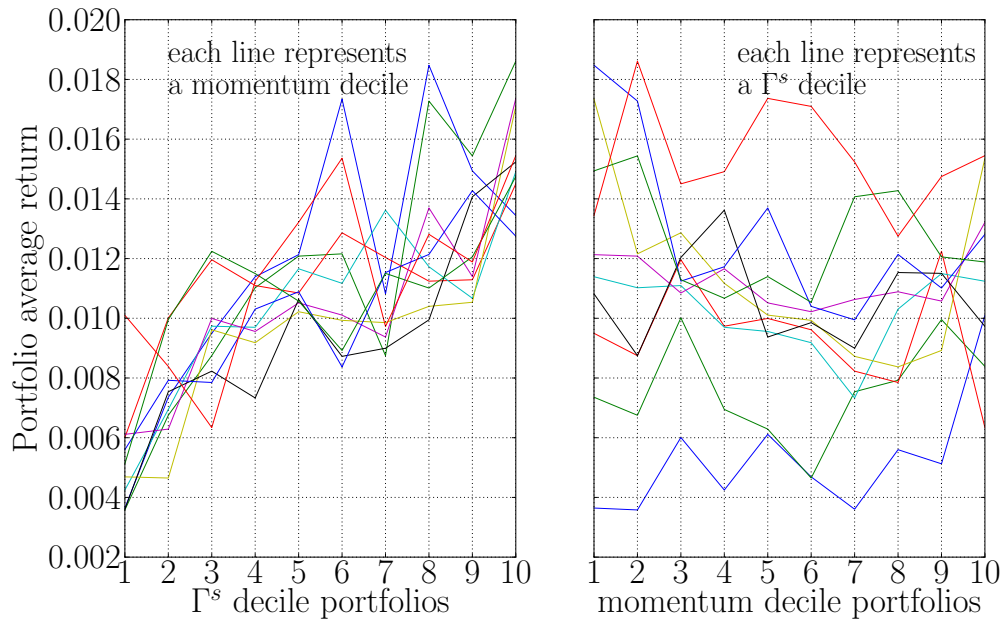


FIGURE 3.12: Left: Average return of Γ^s decile portfolios for a fixed momentum portfolio. It can clearly be seen that average returns increase for higher Γ^s portfolios, for any choice of the momentum decile (any choice of line). Right: Average return of momentum for fixed Γ^s deciles. Average returns don't show any clear trend with respect to increasing momentum decile portfolios. This suggests that the explanatory power on returns lies in Γ^s .

addition, the 3-factor model + Γ^s do explain the returns of the momentum-portfolios as shown in table 3.10. Based on these results, we conclude that Γ^s is a candidate to be a new factor associated to a risk premium.

3.2.4 Discussion

In this work we have studied the behavior of Γ and Γ^s , both measures introduced as proxies for the effect of accelerating prices and positive feedback loops in the stock dynamics. Analysis of these phenomena is important as they constitute a source of predictability, and thus inefficiency. The performance of strategies based on these variables, as well as those incorporating Γ and Γ^s into an APT-like factor model to explain sources of risk premium support our hypothesis. The results are mixed insofar we have observed the strong relationship between our measures and momentum. On one hand, the performance of Γ suggests that increasing returns are already priced into the risk premium explained by momentum. On the other hand, the results for Γ^s suggests that after a simple de-correlation procedure, accelerating prices represent an additional source of risk premium not priced into the classical models. This was underlined by the inability of the standard regression models to account for the returns of the Γ^s -portfolios.

	α		β		r^2
	μ	σ	μ	σ	
1	-0.00	0.00	1.15	0.03	0.77
2	-0.00	0.00	1.02	0.02	0.83
3	-0.00	0.00	0.97	0.02	0.84
4	-0.00	0.00	0.94	0.02	0.87
5	0.00	0.00	0.91	0.02	0.84
6	0.00	0.00	0.91	0.02	0.85
7	0.00	0.00	0.92	0.02	0.85
8	0.00	0.00	0.97	0.02	0.80
9	0.00	0.00	1.02	0.02	0.82
10	0.01	0.00	1.18	0.03	0.77
GRS			4.12		
p-value			0.00		

TABLE 3.7: Regression of the Γ^s decile portfolios on CAPM. The hypothesis of $\alpha_i = 0$ is clearly rejected ($p = 0.0$). CAPM cannot explain the returns of the Γ^s portfolios.

	α		β		s		bm		r^2
	μ	σ	μ	σ	μ	σ	μ	σ	
1	-0.00	0.00	1.05	0.03	0.22	0.04	-0.17	0.04	0.80
2	-0.00	0.00	1.02	0.02	0.01	0.03	0.02	0.03	0.83
3	-0.00	0.00	1.02	0.02	-0.12	0.03	0.09	0.03	0.85
4	-0.00	0.00	0.98	0.02	-0.07	0.02	0.09	0.03	0.88
5	0.00	0.00	0.95	0.02	-0.06	0.02	0.10	0.03	0.85
6	0.00	0.00	0.96	0.02	-0.10	0.02	0.13	0.03	0.87
7	0.00	0.00	0.94	0.02	0.01	0.02	0.13	0.03	0.85
8	0.00	0.00	0.98	0.02	0.01	0.03	0.08	0.03	0.80
9	0.00	0.00	1.01	0.02	0.12	0.03	0.10	0.03	0.82
10	0.00	0.00	1.08	0.03	0.34	0.04	-0.07	0.04	0.81
GRS					3.17				
p-value					0.00				

TABLE 3.8: Regression of the Γ^s decile portfolios on the 3-factor model. Once again, the null hypothesis of $\alpha_i = 0$ is clearly rejected ($p = 0.0$).

	α		β		s		bm		δ		r^2
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	
1	-0.00	0.00	1.0	0.03	0.20	0.04	-0.22	0.04	-0.12	0.02	0.82
2	-0.00	0.00	0.9	0.02	-0.01	0.03	-0.01	0.03	-0.08	0.01	0.84
3	-0.00	0.00	1.0	0.02	-0.13	0.03	0.08	0.03	-0.04	0.01	0.86
4	-0.00	0.00	0.9	0.02	-0.07	0.02	0.09	0.03	-0.00	0.01	0.88
5	0.00	0.00	0.9	0.02	-0.06	0.02	0.10	0.03	0.01	0.01	0.85
6	0.00	0.00	0.9	0.02	-0.10	0.02	0.13	0.03	0.02	0.01	0.87
7	0.00	0.00	0.9	0.02	0.01	0.02	0.13	0.03	0.02	0.01	0.85
8	0.00	0.00	0.9	0.02	0.02	0.03	0.09	0.04	0.03	0.01	0.81
9	0.00	0.00	1.0	0.02	0.13	0.03	0.13	0.03	0.07	0.01	0.83
10	0.00	0.00	1.1	0.03	0.37	0.03	-0.01	0.04	0.16	0.02	0.84
GRS						2.27					
p-value						0.01					

TABLE 3.9: Regression of the Γ^s decile portfolios on the 4-factor model. Even when including momentum, the 4-factor model fails to explain the abnormal returns associated to the Γ^s portfolios ($p = 0.01$).

	α		β		s		bm		γ^s		r^2
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	
1	-0.00	0.00	1.35	0.05	0.55	0.06	0.19	0.07	-0.42	0.04	0.72
2	-0.00	0.00	1.18	0.03	0.22	0.04	0.20	0.05	-0.34	0.03	0.76
3	0.00	0.00	1.10	0.03	0.06	0.04	0.30	0.04	-0.15	0.03	0.77
4	0.00	0.00	0.99	0.02	0.08	0.03	0.22	0.03	-0.12	0.02	0.81
5	-0.00	0.00	0.96	0.02	-0.05	0.02	0.20	0.03	-0.08	0.02	0.84
6	-0.00	0.00	0.94	0.02	-0.06	0.02	0.14	0.03	-0.03	0.02	0.84
7	-0.00	0.00	0.96	0.02	-0.10	0.02	0.15	0.03	0.01	0.02	0.85
8	0.00	0.00	0.95	0.02	-0.10	0.03	0.09	0.03	0.08	0.02	0.82
9	0.00	0.00	0.94	0.02	-0.04	0.03	-0.03	0.03	0.16	0.02	0.82
1	0.00	0.00	0.96	0.03	0.28	0.04	-0.27	0.05	0.27	0.03	0.78
GRS					1.58						
p-value					0.11						

TABLE 3.10: Regression of the 10 Δ decile portfolios on the 3-factor + Γ^s model . The null hypothesis of $\alpha_i = 0$ cannot be rejected ($p = 0.11$). As such, Γ^s explains the momentum anomaly.

More generally, as positive feedback loops dynamics have been theoretically and historically linked to periods of exuberance and panic, one can inquire about the relationship of Γ and Γ^s during periods of instabilities. Figure 3.13 shows the dynamics of the cumulative returns overlaid by the different crisis periods. Table 3.11 presents the average returns of time periods separated by the peaks of the crises for Δ , Γ , and Γ^s . Two observations can be made. First, we can notice that the cumulative returns of Γ^s increase much faster around crisis periods, meaning the long-short strategies based on the acceleration of prices become more successful. This observation is consistent with the idea that bubbles and crashes are characterized by an emergence of positive feedbacks leading to super-exponential price dynamics. Second, the performance of both Γ and Γ^s exhibit an alternating behavior, in which financial crashes marks the beginning of booms and plateaus, a behavior not observed in Δ .

Hence, similarly to Bandarchuk and Hilscher (2013), who maintained that the analysis of momentum should instead focus on the link between momentum profits and extreme past returns, we argue that the analysis of price dynamics needs to be regime dependent. General efforts to validate a theory might fail by not taking into account the market has undergone and adjusted to severe instabilities. Markets are likely to behave efficiently most of the time, while exhibiting limited periods of predictability.

Finally, while our results support the role of the acceleration of prices in the pricing of assets, a question that future research should address is the role of the 6 months shift in the first difference of returns in the construction of Γ^s , a formal model studying the relationship. Namely, the decrease in correlation between Γ^s and momentum is easy

Crisis name	Peak	$\mu(\Delta)$	$\mu(\Gamma)$	$\mu(\Gamma^s)$
Exit of US from the Bretton Woods	1971			
Oil crisis	1973	0.009	0.019	0.003
Energy crisis (second oil crisis)	1979	0.013	0.006	0.019
Savings and loan crisis	1985	0.007	0.007	0.008
Black Monday	1987	0.008	-0.005	0.006
The Mexican peso crisis	1994	0.008	0.006	0.016
Asian and Russian financial crises	1997	0.011	0.003	0.009
Dot-com bubble	2001	0.014	0.038	0.041
Real estate bubble	2006	-0.006	-0.004	-0.002
European sovereign-debt crisis	2009	0.023	0.017	0.014

TABLE 3.11: Chronology of US-related financial and economic crises (adapted from [Kindleberger and Aliber \(2011\)](#) and [Reinhart and Rogoff \(2008\)](#)). The $\mu(\Delta)$, $\mu(\Gamma)$, and $\mu(\Gamma^s)$ correspond to the average return of holding the corresponding portfolio from the end of the previous crisis until the end of the current one.

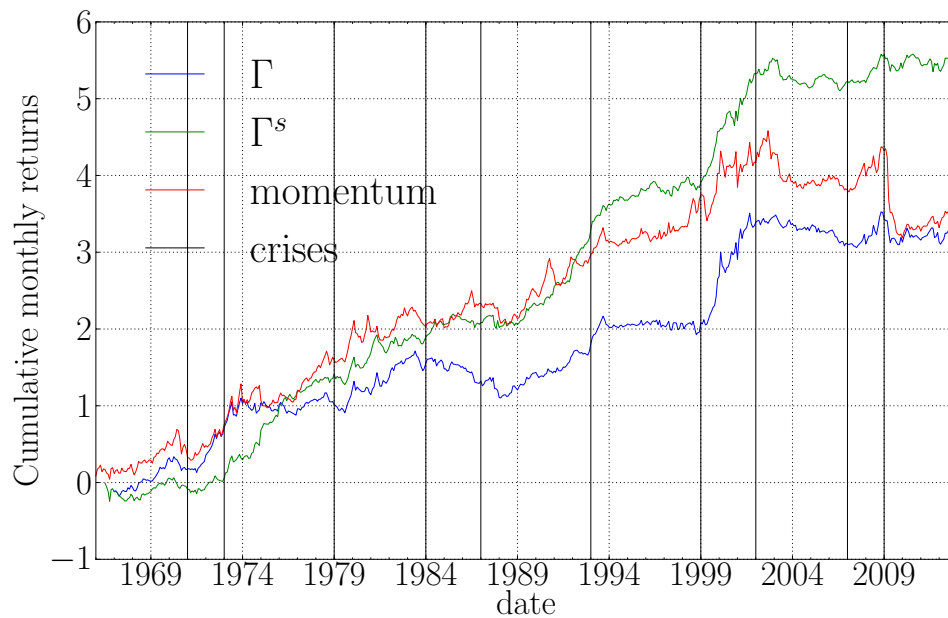


FIGURE 3.13: Cumulative returns of Δ (momentum), Γ and Γ^s over time. The vertical bars represent financial crises (see table 3.11 for details). Interestingly, Γ seems to be alternating between periods of inactivity (flat trend of cumulative returns) and periods of high activity (steep increase of the cumulative returns). The same can be seen for Γ^s albeit not so clearly, given the constant presence of a drift on its cumulative returns. However, momentum doesn't display this alternating behavior. This suggests that mechanisms such as positive feedbacks are at play.

to understand. However, why Γ^s improves its performance with this shift is an open question. It is in fact quite counter-intuitive, since one would expect that a quantity computed with data excluding the last 6 months, and hence containing less information, should have less predictive power.

Chapter 4

Conclusion and outlook

In this thesis, we have addressed the question of how to detect bubbles in financial markets. We have shown, through two fundamentally different approaches, that contrary to the commonly held belief in classical economics, bubbles can be detected and sometimes even their crash can be forecasted. This offers a view in which bubbles and their ensuing crashes are not the result of spontaneous exogenous events such as new pieces of information, but they are rather the product of maturation processes that have distinctive features that can be captured through appropriate models.

In the first part of our work, we have studied bubbles from a fundamental valuation perspective through the example of Zynga. To our surprise, even such a simple approach based on the difference between our fundamental valuation and the market price, agnostic with respect to the mechanisms behind the formation of bubbles, showed the potential of significant predictability.

The power of our methodology lied in the way we valued the company: Zynga, being a social-networking company derived all its revenues from its users. It was by using non-linear decay dynamics coupled with stochastic birth process to forecast its future user base and modeling the revenues each user would generate with a logistic model, that we were able to value the company. All the previously mentioned models take their roots in different fields of science: power-law decay can be observed in many social systems, such as the rate of book sales or the dynamics of video view on YouTube, the Poisson birth process (and generalizations of it) can be found in the field of nuclear physics to model radioactive decay, and the logistic model has been used to describe the dynamical evolution of many biological systems such as the growth of bacteria in a Petri dish. This interdisciplinary approach to model the different components of Zynga relevant to its valuation was at the core of our methodology and probably the reason for its success. This notion was reinforced by the successful application of the very same methodology

to the forecasting of future oil production of Norway and the U.K.

One should nevertheless not lose sight of the inherent limitations of this fundamental valuation approach for the detection of bubbles and especially the prediction of their crashes. By abandoning the need to postulate specific mechanisms for the emergence of bubble, this approach also restricts its ability to foresee crashes. In the case of Zynga, we were able to make a successful short-term prediction of Zynga's crash based on the end of the lock-up period of the company and our diagnosis of a bubble. But such an event only happens once in the lifetime of a company. One can thus usually only make long-term assessments of the asset's price evolution (that we also did successfully in the case of Zynga). The other intrinsic limitation of the fundamental valuation approach lies in the difficulty to scale this analysis to many companies, as the modeling process needs to be done on a case by case basis.

The second part of this thesis took a completely different approach: we turned our attention to the log-periodic power law model, a model proposing a specific mechanism for the origin of bubbles in the form of super-exponential transient price dynamics, taking its roots in the physics of phase transitions and discrete scale invariance. One of the strengths of LPPL is to describe within a single model the emergence of the bubble, i.e. the super-exponential price increase due to the herding behavior of traders, as well as the timing of its crash in the form of a finite-time singularity due to the unsustainable nature of the process.

While the topic is not new, and the LPPL model has been successfully applied to different assets (ex-ante and ex-post), our main contribution has been to show in a systematic way the robustness of LPPL's predictive power on a large scale. This has been done by using an interdisciplinary approach marrying physics and finance: we showed that the outcome of trading strategies based on LPPL persistently deviated from the outcome of random strategies. So as to leave no doubt about the robustness of our results, our study was performed on a universe of 2160 strategies applied to 10 different randomly chosen basket of 50 assets the SP500 in each of the 18 different time periods. The cross-sectional and longitudinal significance of our results support a view in which financial markets are inherently unstable, out-of-equilibrium, permeated by critical phenomena such as phase transitions.

In order to generalize our results and show the importance of super-exponential price dynamics, we studied the impact of the first difference of past returns on future returns in a factor regression framework. The first difference of returns, i.e. the difference between two successive growth rates is the simplest form encapsulating the idea of price acceleration. Our results suggest that a significant risk premium is associated with Γ^s , the factor-mimicking portfolio constructed on the (delayed) first difference of returns.

Moreover, we found that Γ^s could explain the returns of momentum portfolios (embodying the persistence of returns), while the reverse was not true. As such, our results suggest that the momentum factor in the 4-factor model (market, size, book-to-market and momentum), one of the pillars of classical finance, should be replaced by Γ^s .

This indicates that acceleration of prices, as captured by Γ^s , might be the dominant force in the pricing of assets, as opposed to the velocity of prices, as captured by momentum. This interpretation is very much in line with the concept behind the log-periodic power law model, namely that bubbles are a result of super-exponential transient regimes characterized by increasing growth rates.

On an approach specific basis, future research should focus on extending the Zynga-type methodology to many other companies in order to quantify in a statistically significant way this methodology's predictive power. In addition, the whole fundamental valuation approach relying on the discount factor, the modeling of that quantity should be extended to include a dynamical component for its evolution. Indeed, one shouldn't discount with the same rate a company at the early stages of its development and the same company after 10 years of existence.

With respect to the LPPL model, finding a way resolving the sloppiness of the critical time (t_c) would definitely be one of the biggest improvements one could think of. It would allow the use of t_c -related information for the forecast of crashes, something that is dangerous to do in the current state of the model.

From a broader perspective, it would be very interesting to investigate a possible synergy between the two approaches. While both methodologies do not necessarily detect the same bubbles -LPPL doesn't claim to detect all bubbles, but only those with a super-exponential price signature- could the LPPL diagnostic of a bubble be refined by having a fundamental view on the asset? Given an asset exhibiting LPPL dynamics, is the likelihood of a crash higher if its market price is much above its fundamental one? In other words, is there a link between these two approaches? And if so, in what form? This is definifely an exciting question that will be left for another PhD...

Curriculum Vitae

Name: Zalán

Last name: Forró

1987

Born in Zagreb, Croatia

2005 - 2009

Bachelor of Science in Physics, École Polytechnique Fédérale de Lausanne

2009 - 2011

Master of Science in Physics, École Polytechnique Fédérale de Lausanne

2011 - 2015

PhD-Student at the Chair of Entrepreneurial Risks
Department of Management, Technology and Economics
Eidgenössische Technische Hochschule Zürich

Bibliography

- 424B4 Form of the Filings to the SEC of Zynga, 2012. URL <http://goo.gl/7WSro>.
- 8-K Forms of the Filings to the SEC of Zynga, 2012. URL <http://goo.gl/neY30>.
- Dilip Abreu and Markus K. Brunnermeier. Bubbles and Crashes. *Econometrica*, 71(1): 173–204, January 2003. doi: 10.1111/1468-0262.00393. URL <http://dx.doi.org/10.1111/1468-0262.00393>.
- F. Allen, S. Morris, and A. Postlewaite. Finite Bubbles with Short Sale Constraints and Asymmetric Information. *Journal of Economic Theory*, 61(2):206–229, December 1993. ISSN 00220531. doi: 10.1006/jeth.1993.1067. URL <http://dx.doi.org/10.1006/jeth.1993.1067>.
- Ken B. Anderson and James A. Conder. Discussion of Multicyclic Hubbert Modeling as a Method for Forecasting Future Petroleum Production. *Energy Fuels*, 25(4): 1578–1584, March 2011. doi: 10.1021/ef1012648. URL <http://dx.doi.org/10.1021/ef1012648>.
- J. C. Anifrani, C. Le Floc’h, D. Sornette, and B. Souillard. Universal Log-Periodic Correction to Renormalization Group Scaling for Rupture Stress Prediction From Acoustic Emissions. *Journal de Physique I*, 5(6):631–638, June 1995. ISSN 1155-4304. doi: 10.1051/jp1:1995156. URL <http://dx.doi.org/10.1051/jp1:1995156>.
- Clifford S. Asness, Tobias J. Moskowitz, and Lasse H. Pedersen. Value and Momentum Everywhere. *The Journal of Finance*, 68(3):929–985, June 2013. ISSN 00221082. doi: 10.1111/jofi.12021. URL <http://dx.doi.org/10.1111/jofi.12021>.
- Pavel Bandarchuk and Jens Hilscher. Sources of Momentum Profits: Evidence on the Irrelevance of Characteristics. *Review of Finance*, 17(2):809–845, April 2013. ISSN 1573-692X. doi: 10.1093/rof/rfr036. URL <http://dx.doi.org/10.1093/rof/rfr036>.

- Eli Bartov, Partha Mohanram, and Chandrakanth Seethamraju. Valuation of Internet Stocks—An IPO Perspective. *Journal of Accounting Research*, 40(2):321–346, 2002. doi: 10.1111/1475-679x.00050. URL <http://dx.doi.org/10.1111/1475-679x.00050>.
- Best Stock Watch, 2012. URL <http://www.beststockwatch.com/zynga-stock-price-\plunged-as-the-rating-cut.html>.
- Olivier J. Blanchard and Mark W. Watson. Bubbles, Rational Expectations and Financial Markets. *National Bureau of Economic Research Working Paper Series*, pages 945+, June 1983. URL <http://www.nber.org/papers/w0945>.
- David S. Brée, Damien Challet, and Pier P. Peirano. Prediction accuracy and sloppiness of log-periodic functions. *Quantitative Finance*, 13(2):275–280, September 2011. doi: 10.1080/14697688.2011.607467. URL <http://dx.doi.org/10.1080/14697688.2011.607467>.
- Markus K. Brunnermeier and Martin Oehmke. Bubbles, Financial Crises, and Systemic Risk. *National Bureau of Economic Research Working Paper Series*, pages 18398+, September 2012. URL <http://www.nber.org/papers/w18398>.
- Mark M. Carhart. On Persistence in Mutual Fund Performance. *The Journal of Finance*, 52(1):57–82, March 1997. doi: 10.1111/j.1540-6261.1997.tb03808.x. URL <http://dx.doi.org/10.1111/j.1540-6261.1997.tb03808.x>.
- Peter Cauwels and Didier Sornette. Quis Pendit Ipsa Pretia: Facebook Valuation and Diagnostic of a Bubble Based on Nonlinear Demographic Dynamics. *The Journal of Portfolio Management*, 38(2):56–66, January 2012. ISSN 0095-4918. doi: 10.3905/jpm.2012.38.2.056. URL <http://dx.doi.org/10.3905/jpm.2012.38.2.056>.
- Chris Chatfield. *The Analysis of Time Series: An Introduction, Sixth Edition (Chapman & Hall/CRC Texts in Statistical Science)*. Chapman and Hall/CRC, 6 edition, July 2003. ISBN 1584883170. URL <http://www.worldcat.org/isbn/1584883170>.
- Rama Cont and Jean P. Bouchaud. Herd behavior and aggregate fluctuations in financial markets. *Macroeconomic Dynamics*, 4(02):170–196, 2000. doi: 10.1017/s1365100500015029. URL <http://dx.doi.org/10.1017/s1365100500015029>.
- Riley Crane and Didier Sornette. Robust dynamic classes revealed by measuring the response function of a social system. *Proceedings of the National Academy of*

- Sciences*, 105(41):15649–15653, October 2008. ISSN 1091-6490. doi: 10.1073/pnas.0803685105. URL <http://dx.doi.org/10.1073/pnas.0803685105>.
- Riley Crane, Frank Schweitzer, and Didier Sornette. Power law signature of media exposure in human response waiting time distributions. *Physical Review E*, 81(5):056101+, May 2010. ISSN 1539-3755. doi: 10.1103/physreve.81.056101. URL <http://dx.doi.org/10.1103/physreve.81.056101>.
- Cutler, K-M, 2012. URL <http://techcrunch.com/2012/03/21/zynga-omgpop-porter-\sabet-david-ko/>.
- Patricia M. Dechow, Amy P. Hutton, and Richard G. Sloan. The Relation between Analysts' Forecasts of Long-Term Earnings Growth and Stock Price Performance Following Equity Offerings*. *Contemporary Accounting Research*, 17(1):1–32, 2000. doi: 10.1111/j.1911-3846.2000.tb00908.x. URL <http://dx.doi.org/10.1111/j.1911-3846.2000.tb00908.x>.
- Bradford J. DeLong, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. Noise Trader Risk in Financial Markets. *The Journal of Political Economy*, 98(4):703–738, 1990. ISSN 00223808. doi: 10.2307/2937765. URL <http://dx.doi.org/10.2307/2937765>.
- Elizabeth Demers and Baruch Lev. A Rude Awakening: Internet Shakeout in 2000. *Review of Accounting Studies*, 6(2):331–359, June 2001. ISSN 13806653. doi: 10.1023/a:1011675227890. URL <http://dx.doi.org/10.1023/a:1011675227890>.
- F. Deschâtres and D. Sornette. Dynamics of book sales: Endogenous versus exogenous shocks in complex networks. *Physical Review E*, 72(1), July 2005. ISSN 1539-3755. doi: 10.1103/physreve.72.016112. URL <http://dx.doi.org/10.1103/physreve.72.016112>.
- Z. Dezsö, E. Almaas, A. Lukács, B. Rácz, I. Szakadát, and A. L. Barabási. Dynamics of information access on the web. *Physical Review E*, 73(6):066132+, June 2006. ISSN 1539-3755. doi: 10.1103/physreve.73.066132. URL <http://dx.doi.org/10.1103/physreve.73.066132>.
- Eugene F. Fama. Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2):383–417, May 1970. ISSN 00221082. doi: 10.2307/2325486. URL <http://dx.doi.org/10.2307/2325486>.
- Eugene F. Fama and Kenneth R. French. The Cross-Section of Expected Stock Returns. *The Journal of Finance*, 47(2):427–465, 1992. ISSN 00221082. doi: 10.2307/2329112. URL <http://dx.doi.org/10.2307/2329112>.

- Eugene F. Fama and Kenneth R. French. Multifactor Explanations of Asset Pricing Anomalies. *The Journal of Finance*, 51(1):55–84, March 1996. ISSN 00221082. doi: 10.1111/j.1540-6261.1996.tb05202.x. URL <http://dx.doi.org/10.1111/j.1540-6261.1996.tb05202.x>.
- Eugene F. Fama and Kenneth R. French. Size, value, and momentum in international stock returns. *Journal of Financial Economics*, 105(3):457–472, September 2012. ISSN 0304405X. doi: 10.1016/j.jfineco.2012.05.011. URL <http://dx.doi.org/10.1016/j.jfineco.2012.05.011>.
- Laura C. Field and Gordon Hanka. The Expiration of IPO Share Lockups. *The Journal of Finance*, 56(2):471–500, 2001. doi: 10.1111/0022-1082.00334. URL <http://dx.doi.org/10.1111/0022-1082.00334>.
- L. Fiévet, Z. Forró, P. Cauwels, and D. Sornette. A general improved methodology to forecasting future oil production: Application to the UK and Norway. *Energy*, 79: 288–297, January 2015. ISSN 03605442. doi: 10.1016/j.energy.2014.11.014. URL <http://dx.doi.org/10.1016/j.energy.2014.11.014>.
- Zalán Forró, Peter Cauwels, and Didier Sornette. Valuation of Zynga, December 2011. URL <http://arxiv.org/abs/1112.6024>.
- Zalán Forró, Peter Cauwels, and Didier Sornette. When games meet reality: is Zynga overvalued. April 2012a. URL <http://arxiv.org/abs/1204.0350>.
- Zalán Forró, Peter Cauwels, and Didier Sornette. When games meet reality: is Zynga overvalued? *Journal of Investment Strategies*, 1(3), June 2012b. URL http://m.risk.net/digital_assets/5512/jis_sornette_web.pdf.
- Zalán Forró, Ryan Woodard, and Didier Sornette. On the Use of Trading Strategies to Detect Phase Transitions in Financial Markets. *Physical Review Letters E*, in submission, 2015.
- Kenneth A. Froot and Maurice Obstfeld. Intrinsic Bubbles: The Case of Stock Prices. *National Bureau of Economic Research Working Paper Series*, pages 3091+, September 1989. doi: 10.3386/w3091. URL <http://dx.doi.org/10.3386/w3091>.
- John K. Galbraith. *The Great Crash, 1929*. Penguin Group(CA), October 2009. ISBN 014103825X. URL <http://www.worldcat.org/isbn/014103825X>.
- Yuan Gao. Trading and the Information Environment of IPO Stocks Around Lockup Expiration: Evidence from Intraday Data. *Social Science Research Network Working Paper Series*, March 2005. URL <http://ssrn.com/abstract=686566>.

- Peter M. Garber. Tulipmania. *Journal of Political Economy*, 97(3):535–560, June 1989. URL http://econpapers.repec.org/article/ucpjpolec/v_3A97_3Ay_3A1989_3Ai_3A3_3Ap_3A535-60.htm.
- Graf and Christian Meister. Predicting critical crashes? A new restriction for the free variables. *Physica A: Statistical Mechanics and its Applications*, 320:539–547, March 2003. ISSN 03784371. doi: 10.1016/s0378-4371(02)01535-2. URL [http://dx.doi.org/10.1016/s0378-4371\(02\)01535-2](http://dx.doi.org/10.1016/s0378-4371(02)01535-2).
- B. D. Grundy and J. S. Martin. Understanding the nature of the risks and the source of the rewards to momentum investing. *Review of Financial Studies*, 14(1):29–78, January 2001. ISSN 1465-7368. doi: 10.1093/rfs/14.1.29. URL <http://dx.doi.org/10.1093/rfs/14.1.29>.
- Sunil Gupta, Donald R. Lehmann, and Jennifer A. Stuart. Valuing Customers. *Journal of Marketing Research*, 41(1):7–18, February 2004. ISSN 0022-2437. doi: 10.1509/jmkr.41.1.7.25084. URL <http://dx.doi.org/10.1509/jmkr.41.1.7.25084>.
- John R. M. Hand. The Role of Book Income, Web Traffic, and Supply and Demand in the Pricing of U.S. Internet Stocks. *European Finance Review*, 5(3):295–317, January 2001. ISSN 1573-692X. doi: 10.1023/a:1013872530040. URL <http://dx.doi.org/10.1023/a:1013872530040>.
- J. Michael Harrison and David M. Kreps. Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations. *The Quarterly Journal of Economics*, 92(2):323–336, May 1978. ISSN 1531-4650. doi: 10.2307/1884166. URL <http://dx.doi.org/10.2307/1884166>.
- Thorsten Hens and Klaus R. Schenk-Hoppé. Survival of the Fittest on Wall Street. *Institute of Economics University of Copenhagen*, 2004. URL <http://curis.ku.dk/ws/files/32173419/0403.pdf>.
- Mikael Höök, Robert Hirsch, and Kjell Aleklett. Giant oil field decline rates and their influence on world oil production. *Energy Policy*, 37(6):2262–2272, June 2009. ISSN 03014215. doi: 10.1016/j.enpol.2009.02.020. URL <http://dx.doi.org/10.1016/j.enpol.2009.02.020>.
- Marion K. Hubbert. Nuclear Energy and the Fossil Fuel. *Drilling and Production Practice*, January 1956. URL <https://www.onepetro.org/conference-paper/API-56-007>.
- Roger G. Ibbotson and Jay R. Ritter. Chapter 30 Initial public offerings, volume 9 of *Handbooks in Operations Research and Management Science*, pages 993–1016.

- Elsevier, 1995. ISBN 9780444890849. doi: 10.1016/s0927-0507(05)80074-x. URL [http://dx.doi.org/10.1016/s0927-0507\(05\)80074-x](http://dx.doi.org/10.1016/s0927-0507(05)80074-x).
- Kayo Ide and Didier Sornette. Oscillatory finite-time singularities in finance, population and rupture. *Physica A: Statistical Mechanics and its Applications*, 307(1-2):63–106, April 2002. ISSN 03784371. doi: 10.1016/s0378-4371(01)00585-4. URL [http://dx.doi.org/10.1016/s0378-4371\(01\)00585-4](http://dx.doi.org/10.1016/s0378-4371(01)00585-4).
- Narasimhan Jegadeesh and Sheridan Titman. Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance*, 48(1):65–91, March 1993. ISSN 00221082. doi: 10.1111/j.1540-6261.1993.tb04702.x. URL <http://dx.doi.org/10.1111/j.1540-6261.1993.tb04702.x>.
- Narasimhan Jegadeesh and Sheridan Titman. Profitability of Momentum Strategies: An Evaluation of Alternative Explanations. *The Journal of Finance*, 56(2):699–720, April 2001. ISSN 0022-1082. doi: 10.1111/0022-1082.00342. URL <http://dx.doi.org/10.1111/0022-1082.00342>.
- Anders Johansen and Didier Sornette. Significance of Log-Periodic Precursors to Financial Crashes. *Quantitative Finance*, 1(4):452–471, July 2001. URL <http://ssrn.com/abstract=277857>.
- Anders Johansen, Didier Sornette, Hiroshi Wakita, Urumu Tsunogai, William I. Newman, and Hubert Saleur. Discrete Scaling in Earthquake Precursory Phenomena: Evidence in the Kobe Earthquake, Japan. *Journal de Physique I*, 6(10):1391–1402, October 1996. ISSN 1155-4304. doi: 10.1051/jp1:1996143. URL <http://dx.doi.org/10.1051/jp1:1996143>.
- Anders Johansen, Olivier Ledoit, and Didier Sornette. Crashes as critical points. *Int. J. Theor. Appl. Finan.*, 03(02):219–255, April 2000. doi: 10.1142/s0219024900000115. URL <http://dx.doi.org/10.1142/s0219024900000115>.
- Bradford D. Jordan, Daniel J. Bradley, Ivan C. Roten, and Ha-Chin Yi. Venture Capital and IPO Lockup Expiration: An Empirical Analysis. *Social Science Research Network Working Paper Series*, September 2000. URL <http://ssrn.com/abstract=238714>.
- T. Kaizoji and D. Sornette. Market bubbles and crashes. December 2008. URL <http://arxiv.org/abs/0812.2449>.
- Taisei Kaizoji and Didier Sornette. Bubbles and Crashes. 2010. doi: 10.1002/9780470061602.eqf01018. URL <http://dx.doi.org/10.1002/9780470061602.eqf01018>.

- Namwoon Kim, Vijay Mahajan, and Rajendra K. Srivastava. Determining the going market value of a business in an emerging information technology industry: The case of the cellular communications industry. *Technological Forecasting and Social Change*, 49(3):257–279, July 1995. ISSN 00401625. doi: 10.1016/0040-1625(95)00012-y. URL [http://dx.doi.org/10.1016/0040-1625\(95\)00012-y](http://dx.doi.org/10.1016/0040-1625(95)00012-y).
- Charles P. Kindleberger and Robert Aliber. *Manias, Panics, and Crashes: A History of Financial Crises (Wiley Investment Classics)*. Wiley, 5 edition, October 2005. ISBN 0471467146. URL <http://www.worldcat.org/isbn/0471467146>.
- Charles P. Kindleberger and Robert Z. Aliber. *Manias, Panics and Crashes: A History of Financial Crises, Sixth Edition*. Palgrave Macmillan, sixth edition, revised edition, September 2011. ISBN 0230365353. URL <http://www.worldcat.org/isbn/0230365353>.
- Osame Kinouchi and Mauro Copelli. Optimal dynamical range of excitable networks at criticality. *Nature Physics*, 2(5):348–351, April 2006. ISSN 1745-2473. doi: 10.1038/nphys289. URL <http://dx.doi.org/10.1038/nphys289>.
- J. Laherrère and D. Sornette. Stretched exponential distributions in nature and economy: "fat tails" with characteristic scales. *The European Physical Journal B - Condensed Matter and Complex Systems*, 2(4):525–539, April 1998. ISSN 1434-6028. doi: 10.1007/s100510050276. URL <http://dx.doi.org/10.1007/s100510050276>.
- Jure Leskovec, Mary McGlohon, Christos Faloutsos, Natalie Glance, and Matthew Hurst. *Patterns of Cascading Behavior in Large Blog Graphs*, pages 551–556. Society for Industrial and Applied Mathematics, Philadelphia, PA, April 2007. ISBN 978-0-89871-630-6. doi: 10.1137/1.9781611972771.60. URL <http://dx.doi.org/10.1137/1.9781611972771.60>.
- L. Lin and D. Sornette. Diagnostics of rational expectation financial bubbles with stochastic mean-reverting termination times. *The European Journal of Finance*, 19(5):344–365, September 2011. doi: 10.1080/1351847x.2011.607004. URL <http://dx.doi.org/10.1080/1351847x.2011.607004>.
- John Lintner. The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics*, 47(1):13–37, February 1965. ISSN 00346535. doi: 10.2307/1924119. URL <http://dx.doi.org/10.2307/1924119>.

- Thomas Lux. Herd Behaviour, Bubbles and Crashes. *The Economic Journal*, 105(431): 881+, July 1995. ISSN 00130133. doi: 10.2307/2235156. URL <http://dx.doi.org/10.2307/2235156>.
- Harry Markowitz. Portfolio selection*. *The Journal of Finance*, 7(1):77–91, March 1952. ISSN 00221082. doi: 10.1111/j.1540-6261.1952.tb01525.x. URL <http://dx.doi.org/10.1111/j.1540-6261.1952.tb01525.x>.
- R. Michaely and K. L. Womack. Conflict of interest and the credibility of underwriter analyst recommendations. *Review of Financial Studies*, 12(4):653–686, July 1999. ISSN 1465-7368. doi: 10.1093/rfs/12.4.653. URL <http://dx.doi.org/10.1093/rfs/12.4.653>.
- Edward M. Miller. Risk, uncertainty, and divergence of opinion. *The Journal of Finance*, 32(4):1151–1168, September 1977. ISSN 00221082. doi: 10.1111/j.1540-6261.1977.tb03317.x. URL <http://dx.doi.org/10.1111/j.1540-6261.1977.tb03317.x>.
- Ibrahim S. Nashawi, Adel Malallah, and Mohammed Al-Bisharah. Forecasting World Crude Oil Production Using Multicyclic Hubbert Model. *Energy Fuels*, 24(3): 1788–1800, February 2010. doi: 10.1021/ef901240p. URL <http://dx.doi.org/10.1021/ef901240p>.
- Eli Ofek. The IPO Lock-Up Period: Implications for Market Efficiency and Downward Sloping Demand Curves. *Social Science Research Network Working Paper Series*, November 2008. URL <http://ssrn.com/abstract=1298279>.
- Eli Ofek and Matthew Richardson. The Valuation and Market Rationality of Internet Stock Prices. *Oxford Review of Economic Policy*, 18(3):265–287, September 2002. ISSN 1460-2121. doi: 10.1093/oxrep/18.3.265. URL <http://dx.doi.org/10.1093/oxrep/18.3.265>.
- André Orléan. Mimetic contagion and speculative bubbles. *Theory and Decision*, 27 (1-2):63–92, 1989. doi: 10.1007/bf00133988. URL <http://dx.doi.org/10.1007/bf00133988>.
- Carmen M. Reinhart and Kenneth S. Rogoff. This Time is Different: A Panoramic View of Eight Centuries of Financial Crises. *National Bureau of Economic Research Working Paper Series*, pages 13882+, March 2008. doi: 10.3386/w13882. URL <http://dx.doi.org/10.3386/w13882>.
- Jay R. Ritter and Ivo Welch. A Review of IPO Activity, Pricing, and Allocations. *The Journal of Finance*, 57(4):1795–1828, 2002. doi: 10.1111/1540-6261.00478. URL <http://dx.doi.org/10.1111/1540-6261.00478>.

- Hilton L. Root. *Dynamics among Nations: The Evolution of Legitimacy and Development in Modern States*. MIT Press, 2013. URL <https://ideas.repec.org/b/mtp/titles/0262019701.html>.
- Stephen A. Ross. The arbitrage theory of capital asset pricing. *Journal of economic theory*, 13(3):341–360, 1976. URL http://econpapers.repec.org/article/eejetheo/v_3A13_3Ay_3A1976_3Ai_3A3_3Ap_3A341-360.htm.
- Tarik Roukny, Hugues Bersini, Hugues Pirotte, Guido Caldarelli, and Stefano Battiston. Default Cascades in Complex Networks: Topology and Systemic Risk. *Scientific Reports*, 3, September 2013. ISSN 2045-2322. doi: 10.1038/srep02759. URL <http://dx.doi.org/10.1038/srep02759>.
- S1/A Form of the Filings to the SEC of Zynga, 2012. URL <http://goo.gl/h3aZK>.
- M. Scheffer, S. H. Hosper, M. L. Meijer, B. Moss, and E. Jeppesen. Alternative equilibria in shallow lakes. *Trends in Ecology & Evolution*, 8(8):275–279, August 1993. ISSN 01695347. doi: 10.1016/0169-5347(93)90254-m. URL [http://dx.doi.org/10.1016/0169-5347\(93\)90254-m](http://dx.doi.org/10.1016/0169-5347(93)90254-m).
- José A. Scheinkman and Wei Xiong. Overconfidence and Speculative Bubbles. *Journal of Political Economy*, 111(6):1183–1220, December 2003. ISSN 0022-3808. doi: 10.1086/378531. URL <http://dx.doi.org/10.1086/378531>.
- Anna Scherbina. Asset Price Bubbles: A selective survey. *Working Paper*, 2013. URL <http://www.worldcat.org/isbn/9781475515299>.
- Eduardo S. Schwartz and Mark Moon. Rational Pricing of Internet Companies. *Financial Analysts Journal*, 56(3):62–75, May 2000. ISSN 0015-198X. doi: 10.2469/faj.v56.n3.2361. URL <http://dx.doi.org/10.2469/faj.v56.n3.2361>.
- William F. Sharpe. A Simplified Model for Portfolio Analysis. *Management Science*, 9(2):277–293, January 1963. ISSN 0025-1909. doi: 10.1287/mnsc.9.2.277. URL <http://dx.doi.org/10.1287/mnsc.9.2.277>.
- Robert J. Shiller. Bubbles, human judgment, and expert opinion. *Financial Analysts Journal*, pages 18–26, 2002.
- James L. Smith. A Probabilistic Model of Oil Discovery. *The Review of Economics and Statistics*, 62(4):587–594, November 1980.
- D. Sornette, F. Deschâtres, T. Gilbert, and Y. Ageon. Endogenous Versus Exogenous Shocks in Complex Networks: An Empirical Test Using Book Sale Rankings. *Physical Review Letters*, 93(22):228701+, November 2004. ISSN 0031-9007. doi:

- 10.1103/physrevlett.93.228701. URL
<http://dx.doi.org/10.1103/physrevlett.93.228701>.
- Didier Sornette. Discrete-scale invariance and complex dimensions. *Physics Reports*, 297(5):239–270, April 1998. ISSN 03701573. doi: 10.1016/s0370-1573(97)00076-8. URL [http://dx.doi.org/10.1016/s0370-1573\(97\)00076-8](http://dx.doi.org/10.1016/s0370-1573(97)00076-8).
- Didier Sornette. *Why Stock Markets Crash: Critical Events in Complex Financial Systems*. Princeton University Press, 2003.
- Didier Sornette. Endogenous versus Exogenous Origins of Crises. In Sergio Albeverio, Volker Jentsch, and Holger Kantz, editors, *Extreme Events in Nature and Society*, The Frontiers Collection, chapter 5, pages 95–119. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006. ISBN 978-3-540-28610-3. doi: 10.1007/3-540-28611-x_5. URL http://dx.doi.org/10.1007/3-540-28611-x_5.
- Didier Sornette. Physics and financial economics (1776?2014): puzzles, Ising and agent-based models. *Reports on Progress in Physics*, 77(6):062001+, June 2014. ISSN 0034-4885. doi: 10.1088/0034-4885/77/6/062001. URL <http://dx.doi.org/10.1088/0034-4885/77/6/062001>.
- Didier Sornette and Peter Cauwels. 1980–2008: The Illusion of the Perpetual Money Machine and What It Bodes for the Future. *Risks*, 2(2):103–131, April 2014. doi: 10.3390/risks2020103. URL <http://dx.doi.org/10.3390/risks2020103>.
- Didier Sornette and Anders Johansen. A hierarchical model of financial crashes. *Physica A: Statistical Mechanics and its Applications*, 261(3-4):581–598, December 1998. ISSN 03784371. doi: 10.1016/s0378-4371(98)00433-6. URL [http://dx.doi.org/10.1016/s0378-4371\(98\)00433-6](http://dx.doi.org/10.1016/s0378-4371(98)00433-6).
- Didier Sornette, Anders Johansen, and Jean-Philippe Bouchaud. Stock Market Crashes, Precursors and Replicas. *Journal de Physique I*, 6(1):167–175, January 1996. ISSN 1155-4304. doi: 10.1051/jp1:1996135. URL <http://dx.doi.org/10.1051/jp1:1996135>.
- Didier Sornette, Ryan Woodard, Wanfeng Yan, and Wei-Xing Zhou. Clarifications to questions and criticisms on the Johansen–Ledoit–Sornette financial bubble model. *Physica A: Statistical Mechanics and its Applications*, 392(19):4417–4428, October 2013. ISSN 03784371. doi: 10.1016/j.physa.2013.05.011. URL <http://dx.doi.org/10.1016/j.physa.2013.05.011>.
- Takahashi, D, 2012. URL <http://venturebeat.com/2012/03/01/zynga-launches-zynga-com-as-it-expands-beyond-facebook-in-a-big-way/>.

- Earl A Thompson. The tulipmania: Fact or artifact? *Public Choice*, 130(1-2):99–114, January 2007. ISSN 0048-5829. doi: 10.1007/s11127-006-9074-4. URL <http://dx.doi.org/10.1007/s11127-006-9074-4>.
- Yahoo Finance. URL <http://finance.yahoo.com/q?s=ZNGA&q1=1>.
- Wanfeng Yan, Ryan Woodard, and Didier Sornette. Diagnosis and prediction of rebounds in financial markets. *Physica A: Statistical Mechanics and its Applications*, 391(4):1361–1380, February 2012. ISSN 03784371. doi: 10.1016/j.physa.2011.09.019. URL <http://dx.doi.org/10.1016/j.physa.2011.09.019>.
- Wei-Xing Zhou, Guo-Hua Mu, Wei Chen, and Didier Sornette. Investment Strategies Used as Spectroscopy of Financial Markets Reveal New Stylized Facts. *PLoS ONE*, 6(9):e24391+, September 2011. ISSN 1932-6203. doi: 10.1371/journal.pone.0024391. URL <http://dx.doi.org/10.1371/journal.pone.0024391>.