DISS. ETH NO. 25162

FINANCIAL MARKETS IN NATURAL EXPERIMENTS, FIELD EXPERIMENTS, LAB EXPERIMENTS AND REAL LIFE

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

presented by

KE WU
(吴柯)

M.A. in the Mathematics of Finance, Columbia University

born on June 22, 1992

citizen of P.R. China

accepted on the recommendation of
PROF. DR. Didier SORNETTE
PROF. DR. Thorsten HENS
PROF. DR. Zhiyi HE

2018

Acknowledgments

Pursuing a PhD has never been a plan for me before meeting Prof. Didier Sornette. However, just after our first conversation, it became all I wanted to do for the next chapter of my life. Three years later, here I am at the end of this chapter, feeling extremely grateful for making one of the best choices in my life.

First of all, I would like to express my heartfelt thanks to my incredible supervisor and mentor Prof. Didier Sornette. He is exceptionally knowledgeable in a wide range of subjects, and passionately applying his unique expertise in solving down-to-earth yet very challenging problems for the real world. His endless inspiring ideas and valuable experience have taught me so much during my PhD. Moreover, the wide diversity of topics he developed and the large flexibility he allowed in the Chair gives me a freedom to embark on many wonderful adventures. Prof. Sornette has always been open-minded and encouraging for all the ideas I raised and things I wanted to do, including some beyond the scope of this thesis. More importantly, his ambitious, curious, optimistic, adventurous and sustainable attitudes toward life and career have deeply influenced me. Also, I would like to thank Prof. Yougui Wang, who was my supervisor in my bachelor time at Beijing Normal University in China, opened my door to science, supported and advised me for a long time, and introduced me to meet Prof. Sornette at Shanghai in May 2014. I also thank Prof. Thorsten Hens from University of Zurich and Prof. Zhiyi He from Peking University and NewHuadu Business School (Switzerland) for being my co-referees.

Secondly, special thanks go to my collaborators of the papers included in this thesis: Spencer Wheatley, Sandra Andraszewicz, Ryan O. Murphy, Philipp Rindler and Dorsa Sanadgol, from whom I have received countless helps and ideas. In particular, I want to thank Spencer Wheatley for many important statistical expertises he contributed, and for many constructive and interesting discussions we had together. I am also grateful for the supports and idea exchanges from current and former members in the Chair of Entrepreneurial Risks and in ETH Zurich, which together make my journey here remarkable.

Thirdly, I deeply appreciate my funding source: the European Union's Horizon 2020 research and innovation Programme under the Maria Skłodowska-Curie grant agreement No 643073, and the professors and colleagues in the CRITICS network (Critical Transitions in Complex Systems), with whom we had many interesting conferences and discussions.

Lastly but most importantly, I would like to thank my father Hua Wu and mother Xiangting He, who have been constantly providing their unconditional supports, understandings, and useful advises.

Abstract

This thesis includes four submitted research papers and some additional unpublished material. Starting from two major political events in 2016, which served as two natural experiments, the thesis illustrates the market failure in aggregating and reflecting information, and ponders different reasons including political bubbles in the markets. Then the thesis moves on from the natural experiments to a new field experiment design, which is more realistic than the existing set-ups, in order to study the "bubble-and-crash puzzle", the factors that could mitigate the bubble and the impact of information on market prices. The field experiment is then reproduced in a lab environment, to test the impact of different experimental environments. Lastly, we go further to study and quantify a real market case – cryptocurrency markets – which provides fundamental insights to understand this new and young market.

In Chapter 2, we exploit the near-experimental conditions provided by the British Pound and Mexican Peso exchange rate markets during the Brexit vote of June 23rd, 2016, and the Nov. 8th, 2016 U.S. Presidential election, and propose a simple predictive methodology based on re-calibrating the biased ex-ante predictions of pollsters, in real time. In the case of Trump's election, a reliable early prediction of the outcome was possible, and the Peso market partially reflected this. In the case of the Brexit, we document an astonishing delay of the Pound market in reflecting the increasingly high probability of a Brexit outcome, which could have been predicted very early by soundly reconciling the flow of voting information with its prior belief. This failure occurred notwithstanding evidence that the market was actively responding to announcements, showing significant trading activity in reaction to each "news item" associated with the vote announcements. Put together, this constitutes a major breakdown of market efficiency, which we attribute to confirmation biases and social herding, leading to a political bubble and subsequent crash.

In Chapter 3, we test the hypothesis that mispricing is a robust finding that generalises beyond the standard experimental SSW design, which has been characterised by a persistent "bubble-and-crash puzzle". We propose a new paradigm to study strategic decision making and coordination in financial asset markets. This new paradigm is a prediction market with additional features that have been previously shown to mitigate price bubbles. The paradigm accounts for fundamental uncertainty, as there is no "true" underlying asset price. This setup is more realistic than the standard SSW design. Nonetheless, the market's correct pricing can be evaluated. In two experiments, we observed pronounced overpricing, which however does not show the standard "bubble-and-crash" scenario. The overpricing diminishes over time, indicating learning, but does not disappear completely.

Traders' price estimates showed a collective realisation of communal ignorance, pushing the market much closer to its true value.

In Chapter 4, to explore the similarities and differences between human behaviour in the laboratory environment and in a realistic natural setting, with the same type of participants, we translate the field study in Chapter 3 with trading rounds each lasting six full days to a laboratory experiment lasting two hours. The laboratory experiment replicates the key findings from the field study but we observe substantial differences in the market dynamics between the two settings. The replication of the results in the two distinct settings indicates that relaxing some of the laboratory control does not corrupt the main findings, while at the same time it offers several advantages such as the possibility to increase the number of participants interacting with each other at the same time and the number of traded securities.

In Chapter 5, we extend the experiment in Chapter 3, and preliminarily investigate the impact of objective (i.e. reliable and quantifiable) and subjective (i.e. ambiguous) information on market prices. We hypothesize that providing participants with objective information about the "correct" price level of the market, with the clear indication of how to use market inefficiencies, would reduce mispricing by moving the price level to the "rational" range. We present results from two treatments employing the Experimental Asset Markets setup proposed in Chapter 3, each lasting two weeks.

In Chapter 6, we empirically verify that the market capitalisations of coins and tokens in the cryptocurrency universe follow power-law distributions with significantly different values, with the tail exponent falling between 0.5 and 0.7 for coins, and between 1.0 and 1.3 for tokens. We provide a rationale for this, based on a simple proportional growth with birth & death model previously employed to describe the size distribution of firms, cities, webpages, etc. We empirically validate the model and its main predictions, in terms of proportional growth (Gibrat's law) of the coins and tokens. Estimating the main parameters of the model, the theoretical predictions for the power-law exponents of coin and token distributions are in remarkable agreement with the empirical estimations, given the simplicity of the model. Our results clearly characterize coins as being "entrenched incumbents" and tokens as an "explosive immature ecosystem", largely due to massive and exuberant Initial Coin Offering activity in the token space. The theory predicts that the exponent for tokens should converge to 1 in the future, reflecting a more reasonable rate of new entrants associated with genuine technological innovations.

Abstrakt

Die vorliegende Doktorarbeit beinhaltet vier eingereichte Forschungsarbeiten und einige zusätzliche unveröffentlichte Materialien. Ausgehend von zwei wichtigen politischen Ereignissen im Jahr 2016, die als zwei natürliche Experimente dienten, veranschaulicht die Doktorarbeit das Marktversagen bei der Aggregation und Reflexion von Informationen und erwägt verschiedene Gründe, einschließlich politischer Blasen in den Märkten. Dann geht die Arbeit von den natürlichen Experimenten zu einem neuen Feld-Experiment-Design über, das realistischer ist als die bestehenden Anordnungen, um das "Bubbles-and-Crash-Puzzle" zu untersuchen, die Faktoren, die eine Blase abschwächen könnten und die Auswirkungen von Informationen auf Marktpreise. Das Feldexperiment wird dann in einer Laborumgebung reproduziert, um den Einfluss verschiedener experimenteller Umgebungen zu testen. Schließlich gehen wir noch weiter, um einen realen Marktfall - Kryptowährungsmärkte - zu untersuchen und zu quantifizieren, um diesen neuen und jungen Markt zu verstehen.

Im Kapitel 2 nutzen wir die nahezu experimentellen Bedingungen des Wechselkursmärkte des britischen Pfund bei der Brexit-Abstimmung vom 23.Juni 2016 und des mexikanische Peso bei der US-Präsidentschaftswahl vom 8.November 2016 und schlagen eine einfache prädiktive Methodik vor, die auf einer Neukalibrierung der voreingenommenen Ex-ante-Vorhersagen von Meinungsforschern in Echtzeit basiert. Bei der Wahl von Trump war eine zuverlässige frühzeitige Vorhersage des Ergebnisses möglich, und der Peso-Markt spiegelte dies teilweise wider. Im Falle des Brexit dokumentieren wir eine erstaunliche Verzögerung der Reaktion des Pfund-Marktes, da er die zunehmend hohe Wahrscheinlichkeit eines Brexit-Ergebnisses widerspiegelt, das sehr früh vorhergesagt werden konnte, indem der Fluss der Stimmrechtsinformationen mit seiner früheren Überzeugung in Einklang gebracht wurde. Dieses Scheitern ereignete sich trotz des Beweises, dass der Markt aktiv auf Ankündigungen reagierte und als Reaktion auf jede "Nachricht", die mit den Abstimmungsankündigungen verbunden war, signifikante Handelsaktivität aufwies. Beides stellt einen erheblichen Zusammenbruch der Markteffizienz dar, die wir auf die Bestätigungstendenz und „social herding" zurückführen, was zu einer politischen Blase und einem anschließenden Crash führt.

In Kapitel 3 testen wir die Hypothese, dass Fehlbewertung ein robuster Vorkommnis ist, das über das standardmäßige experimentelle SSW-Design hinausgeht, das durch ein hartnäckiges "Bubble-and-Crash-Puzzle" charakterisiert wurde. Wir schlagen ein neues Paradigma vor, um die strategische Entscheidungsfindung und Koordination auf den Märkten für Finanzanlagen zu untersuchen. Dieses neue Paradigma ist ein Prognosemarkt

mit zusätzlichen Funktionen, von denen zuvor gezeigt wurde, dass sie Preisblasen abschwächen. Das Paradigma erklärt die fundamentale Unsicherheit, da es keinen "wahren" zugrunde liegenden Vermögenspreis gibt. Dieses Setup ist realistischer als das Standard-SSW-Design. Dennoch kann die korrekte Preisgestaltung des Marktes bewertet werden. In zwei Experimenten beobachteten wir eine ausgeprägte Überbewertung, die jedoch nicht das Standard-Bubble-and-Crash-Szenario zeigt. Die Überbewertung verringert sich im Laufe der Zeit und zeigt ein Lernpotential an, verschwindet aber nicht vollständig. Die Preiserwartungen der Trader zeigten eine kollektive Realisierung von kommunaler Ignoranz und drängten den Markt viel näher an seinen wahren Wert.

In Kapitel 4 übersetzen wir die Feldstudie in Kapitel 3 mit Handelsrunden, die jeweils sechs volle Tage dauern, zu einem zweistündigen Laborexperiment, um die Ähnlichkeiten und Unterschiede zwischen menschlichem Verhalten in der Laborumgebung und in einer realistischen natürlichen Umgebung mit der gleichen Art von Teilnehmern zu untersuchen. Das Laborexperiment repliziert die wichtigsten Ergebnisse aus der Feldstudie, aber wir beobachten erhebliche Unterschiede in der Marktdynamik zwischen den beiden Umgebungen. Die Replikation der Ergebnisse in den zwei verschiedenen Umgebungen zeigt, dass ein Lockern der Kontrollvariablen im Labor nicht die Hauptergebnisse verfälscht, während sie gleichzeitig mehrere Vorteile bietet, wie etwa die Möglichkeit, die Anzahl der Teilnehmer, die zeitgleich miteinander interagieren, und die Anzahl der gehandelten Wertpapiere zu erhöhen.

Im Kapitel 5 erweitern wir das Experiment im Kapitel 3 und untersuchen vorläufig den Einfluss objektiver (d.h. zuverlässiger und quantifizierbarer) und subjektiver (d.h. mehrdeutiger) Informationen über Marktpreise. Wir stellen die Hypothese auf, dass die Bereitstellung von objektiven Informationen über das "korrekte" Preisniveau des Marktes mit klaren Hinweisen darauf, wie Marktineffizienzen genutzt werden können, die Fehlbewertung reduzieren würde, indem das Preisniveau in den "rationalen" Bereich verschoben wird. Wir präsentieren Ergebnisse aus zwei Experimenten, die den in Kapitel 2 vorgeschlagenen Kapitalmarkt als Versuchsaufbau verwenden und jeweils zwei Wochen dauern.

In Kapitel 6 überprüfen wir empirisch, dass die Marktkapitalisierung von Coins und Token im Kryptowährungsuniversum Potenzgesetzverteilungen mit signifikant unterschiedlichen Werten folgt, wobei der Tail-Exponent zwischen 0,5 und 0,7 für Coins und zwischen 1.0 und 1.3 für Token liegt. Wir liefern eine Begründung dafür, basierend auf einem einfachen proportionalen Wachstum mit dem Geburts- und Todesmodell, das früher verwendet wurde, um die Größenverteilung von Firmen, Städten, Webseiten usw. zu beschreiben. Wir validieren empirisch das Modell und seine Hauptprognosen in Bezug auf proportionales Wachstum (Gibrats Gesetz) der Münzen und Wertmarken. Die theoretischen Voraussagen für die Potenzgesetzexponenten der Coin- und Tokenverteilungen schätzen die wichtigsten Parameter des Modells ein und stimmen angesichts der Einfachheit des Modells in bemerkenswerter Übereinstimmung mit den empirischen Schätzungen überein. Unsere Ergebnisse charakterisieren Münzen eindeutig als "etablierte Incumbents" und Token als "explosives unreifes Ökosystem", hauptsächlich aufgrund massiver und überschwänglicher Initial Coin Offering-Aktivitäten im Token-Bereich. Die Theorie sagt voraus, dass der Exponent für Tokens in Zukunft gegen 1 konvergieren sollte, was eine vernünftigere Rate von Neueinsteigern im Zusammenhang mit echten technologischen Innovationen widerspiegelt.

Contents

Chapter 1

Introduction

1.1 Motivation from Brexit and Trump – two natural experiments

Financial markets have long been regarded by Financial economists as efficient mechanisms for the aggregation of information, and thus effective predictors of the probability of future events [e.g. 1, 2, 3, 4, 5]. Given their intrinsic uncertain nature, real financial markets can be conceived as particular incarnations of prediction markets, where the possible outcomes are known while the underlying probability structure of the outcomes is unknown and fundamentally unknowable. Therefore, the participants of prediction markets make "educated guesses", while the market prices emerging from aggregated traders' beliefs should reflect the probability of future outcomes [6, 7]. In financial markets, traders aggregate their beliefs concerning the future performance of firms, leading to prices that can be interpreted as predictions of the firm value.

In academic thinking, prediction markets, "in which prices are used to predict future events" [7], have been used to successfully predict political elections [8, 9, 10, 11], spells of infectious diseases [12, 13], sports outcomes [14] and new product blockbusters [15, 16, 17], just to name a few examples. Ref. [18] provide a comprehensive review of the use of prediction markets in the laboratory and field studies.

More fundamentally, the belief that markets are a sound and powerful valuation system largely underlies free market systems and economic liberalism. The claim that financial markets are information-efficient was made in the famous Efficient Markets Hypothesis (EMH), formulated by Samuelson[19] and Fama[20], and restated by Fama[21]: "*I take the market efficiency hypothesis to be the simple statement that security prices fully reflect* *all available information*". Three levels of market efficiency have been defined, of which the *semi-strong market efficiency* is tested widely by how quickly security price reflect relevant information that was obviously publicly available [20, 21]. Numerous studies have contributed to this "event studies" research agenda [e.g. 21, 22, 23], and a standard routine of performing such a test has been summarized by [24]. Researchers tested abnormal returns in stock markets, bond markets and currency markets, with respect to macroeconomic news, central bank announcements and other various type of news [see a review by 25].

In the literature supporting or criticizing the EMH, there are on-going controversies over what is meant by "fully reflecting" the "true value" of the price, the boundary of "quickly" (short-term or long-term return horizon), the level and definition of rationality of market participants, the entanglement of different types of event impacts and reactions, the operational market efficiency (liquidity, accessibility, the scale of market participants, transaction costs, etc.), and so on [e.g. 22, 23, 26, 27, 28, 29, 30]. To make a test of the EMH operational, one must specify additional structures like investor's preference, normal return models, etc., which makes the test become a test of several auxiliary hypotheses [31, 32]. This joint hypothesis problem, which is largely discussed in the literature, leads to the claim that market efficiency as such can practically never be rejected. And it is, in fact, very difficult to carry out unambiguous tests outside of an actual laboratory experiment [33].

Two important elections in 2016 – the EU "Brexit" Referendum and the US "Trump" Presidential Election – provided two rare real-life experiments to test the Efficient Market Hypothesis. Crucially, the market – ordinarily robed in complexity – momentarily exposed herself in a simplified state, allowing an exceptionally objective analysis of response to fundamental information. The stream of area vote count announcements provided the stimulus, and the Pound market (British Pound in US Dollar) and the Peso market (Mexican Peso in US Dollars) were taken as the responses.

In Chapter 2, we employ a simple and natural one-factor linear model, to analyze the market efficiency in terms of their response to the stream of voting results. In practical terms for voting events, with high stakes, technological sophistication, and immense resources, one would expect the financial markets to ravenously consume and digest both public and rarefied information streams, and through arbitrage, yield a near-unbeatable market price (i.e., one consistent with the underlying probability). In contrast, we use a humble real-time algorithm, based on a simple linear regression of public voting announcements onto the prevailing expert ex-ante polls[34, 35]. The surprise is that, in the case of the Brexit, our algorithm confidently and robustly predicts a Brexit outcome after only 20-30 out of 382 local voting results had been revealed, hours before the Pound market (in US Dollars) had reflected the outcome. In the case of the Trump election, however, the Peso market (in US Dollars) was efficient, as the growing probability of a Trump presidency was reflected.

It is worthwhile to note that an efficient market is not equivalent to a responsive market. Indeed, we show that the markets were extremely responsive to the announcements, as illustrated by the quick response to Sunderland results in the UK and the strong reactive market activities presented later. However, this responsive market did not give an effective prediction of the result of the UK Referendum in a probabilistic sense. In other words, the available information was not fully reflected in the market, using any reasonable range of models that could interpret the market price.

Based on the limited two elections, we do not claim to fully explain why the Brexit market was inefficient while the Trump market was not. This leaves the full meaning of this event somewhat open to interpretation. However a plausible one is that the US Presidential Election was a more familiar process, in accordance with beliefs about market efficiency, and the ability of our prediction to coincide with the market trajectory validates the strength of the algorithm. And therefore the application of the same algorithm to the Brexit case is a reliable indicator of an anomalous situation, perhaps due to Brexit being a one-off rather than a routine voting event. This dramatic failure of the market to converge to the easily accessible fundamental value could be associated with confirmation bias in the financial markets [36], groupthink and social sentiments, which are evocative of the herding psychology that characterise financial bubbles [37, 38]. In a sense, the markets were in a "Remain bubble", and the Pound crashed only at the very last stage when the outcome became inevitable as in the famous SSW economic experimental design [39].

1.2 From the real markets to field and lab experiments

1.2.1 Two field experiments in the classroom

This SSW design proposed by Smith, Suchanek and Williams[39] is a seminal study with a simple setup where a few persons would trade one risky asset over a period of a few minutes. This setup pioneered the use of experimental asset markets to study how financial markets function and how specific mechanism changes might affect trading behavior and price outcomes, to simplify the complexity and interaction of many variables in the real market.

The SSW design has generated a large experimental literature (see [40] for a review). Price bubbles and crashes are a robust finding in this type of studies, even when the information about the rational price is directly provided to the participants [41]. This phenomenon known as the “bubble-and-crash puzzle” has not been fully understood and “formal theoretical explanation is an area of future work” [42].

Ref. [40] identified a number of factors that have been found to mitigate the price bubbles in an experimental setting. These factors include: expertise of a trader, common expectations of rationality, low cash-to-asset ratio, large accrual dividend, trading teams instead of individual traders, lack of overconfidence, existence of alternatives to trading, short-selling, limit price change, non-tournament type of compensation, and comparison to the best players. Individual factors mitigating mispricing do not make the bubbles disappear completely, even after training of the participants. However, Ref. [43] demonstrated that the declining fundamental value process in the SSW design is confusing for participants and that making this process more intuitive resolves the confusion and reduces mispricing.

This motivates the question of “how well the results [of the SSW experiments] extend to more realistic market settings” [44]. In psychological laboratory studies, the relation between the experimental findings and people’s behaviour “in the wild” is an important point of critique that is addressed by alternating experimental paradigms to test related theories and by conducting field studies. However, in experimental asset markets, the same research method - the SSW design - is repeatedly utilized in studies that derive their theories from previous experiments [45].

Consequently, “many experiments are not aimed at a well-specified real-world target but rather contribute to a ‘library of robust phenomena’, a body of experimental knowledge to be applied case by case” [46]. Repetitive use of the very same experimental design may lead to the “mutual internal validity of theory and experimental test” [47] that creates its own world, where the robust findings from experiments may not be generalisable to the outside world [46]. Artificiality of laboratory experiments and lack of context may reduce their relation to the real trading situations [45, 47].

On the one hand, laboratory experiments offer the possibility to manipulate and measure individual variables in a fully controlled way. On the other hand, it is a crucial question whether the bubbles in the experimental markets are a characteristic artifact of the SSW design or whether it is a general phenomenon resulting from the action of the market players. According to Ref. [44], the structure of the market plays an important role in attenuating or mitigating the bubbles. “What is still missing, however, is a careful analysis of possible new experimental methods that will help increase the external validity [of the experimental asset markets]" [47].

Motivated by these important questions, in Chapter 3, we propose a new experimental design for investigating asset markets. This new design is a prediction market with additional features that have been previously shown to mitigate price bubbles. The paradigm accounts for fundamental uncertainty, as there is no "true" underlying asset price. Our design is more realistic than the SSW design, but simple enough to conduct analysis of controlled variables. In contrast to the SSW design, in which full information about the value of the securities is given to participants, our experiment features true Knightian uncertainty[1] [48]. Nonetheless, the market's correct pricing can be evaluated. In the two experiments, we test the hypothesis that mispricing is a robust finding that generalizes beyond the SSW design. We observed pronounced overpricing, which however does not show the standard "bubble-and-crash" scenario. The overpricing diminishes over time, indicating learning, but does not disappear completely. Traders' price estimates showed a collective realization of communal ignorance, pushing the market much closer to its true value.

1.2.2 Back to the lab

Due to the high complexity of financial markets, most of existing studies are simplified to a highly controlled laboratory setting [49], to disentangle the interaction among individual measured variables from random or not-measured variables. While laboratory studies allow for controlling the variables of interest, they are often open to the criticism that their environment is quite artificial [47]. Ref. [50] argues that lack of realistic conditions is not a problem and that laboratory markets are real markets as long as the general economic principles apply. According to this reasoning, artificiality is not an issue if an experiment allows for testing and comparing particular theories [49]. However, Ref. [45] points out that highly structured markets, such as those implemented in laboratory experiments, are rare in real life. He indicates that "most of the economic transactions [...] are notable for the lack of disciplining mechanisms." Therefore, "laboratory experiments are of limited relevance for predicting field behaviour, unless one wants to insist a priori that those aspects of economic behaviour under study are perfectly general" [51]. Moreover, the control in the laboratory may paradoxically introduce unintended variables that are not present in the wild, such as time pressure, limited time, lack of field-specific knowledge, etc.

One way of investigating the robustness of experimental results is their replicability. Re-

[1] A situation in which we cannot know all the information we need in order to calculate the odds.

peating the experiments is a way to define whether the particular finding is a true stylised fact or rather an artifact generated by inexperience, coincidence or mistake [52]. The issue of replicability in behavioural sciences has been addressed by Ref. [53], who have replicated 100 original studies published in three top journals in psychology. Following this tradition, Ref. [54] replicated 18 studies in the American Economic Review and the Quarterly Journal of Economics. Ref. [53] reported reproducibility of 36%, while Ref. [54] reported that the results were replicated in 61% of the studies.

The need for replicability of results is reflected by the creation of electronic libraries of standard experimental tasks [55]. However, the fact that a particular effect is replicated many times in a very similar setting does not imply that this effect is of any relevance outside of this environment. Following this line of reasoning, one could fall into a trap of testing theories in an isolated environment that hold under the assumptions of this environment[2] [47].

Our approach to replicability of experimental results is different - we aim to evaluate the generalisability of behavioural effects obtained both in more realistic and in artificial experimental environments. Towards this goal, in Chapter 4, we translate our field experimental asset market study in Chapter 3 to the laboratory setting. We use the same experimental material and rules, but we adapt the procedure to the sterile laboratory environment. The point of this exercise is to challenge a frequent misconception about field studies that field experiments are the "uncontrolled variants of laboratory experiments"[56]. On the contrary, we propose that the domain of experimental asset markets conducted in the laboratory resulted in such a large literature investigating interactions among individual variables (see [40, 49] for a review) that the next direction in this experimental domain could be to relax some of the control restrictions to obtain additional insights into how people behave in more realistic settings and to use advantages that such non-laboratory experiments offer.

As a stepping stone between transferring from the highly controlled laboratory experiment to only loosely controlled field or natural experiments, it is necessary to investigate the replicability of the main effects in these experiments in the field and in laboratory settings[51]. For example, Ref. [56] replicated in the field a standard experimental design used in the environmental policy experiments. He found contrasting results to the previous laboratory studies by Ref. [57] and Ref. [58]. Ref. [59] postulated that "one highly important question about the external validity of experiments is whether the same individuals

[2] In the "classical" economic research, theories would be proven by mathematical derivation, ignoring anomalies in the data and variables not considered by the model. Analogically, experimental economists may fall prey to making the same mistake by ignoring important experimental methodological issues related to artificiality of the experimental setting.

act in experiments as they would in the field." Ref. [60] investigated the differences in behaviour in computerised matrix games between student, professional card game players and professional football players, conducted in the laboratory and in the professionals' natural environment. They found both professionals and students fall prey to cognitive biases when in the laboratory. They surmised that professionals come to the laboratory with the pre-learned skills and knowledge and, when exposed to the same role as in real life, they transfer this knowledge to the laboratory task. In contrast, when exposed to a novel task or novel environment, the professionals fall prey to the same biases as students, indicating that the environment in which one preforms a task may have a crucial role on the performance.

Levitt and List [61, 62, 63] advocated the use of field studies for economic experimentation as opposed to laboratory experiments that according to the authors lack generalisability to the real life behaviour. Their work was heavily criticized by Ref. [64], who reviewed a number of studies that directly compared field studies with their laboratory counterparts. According to Ref. [64], by 2011, there were only 6 studies designed for direct field-lab comparison. None of these studies used experimental asset markets but there was a high correlation between the lab and field results.

Experimentally studying complex systems such as asset markets poses a number of challenges. In our opinion, the biggest challenge is that real asset markets offer a large number of securities to a large number of market participants who can interact with each other at various times during trading hours. The interaction can happen over buy and sell orders and through interpersonal communications. The laboratory setting reproduces these features in a very limited and reductionistic way while, on the other hand, reducing possible noise contributions (effects resulting from these uncontrolled variables).

In the lab experiment in Chapter 4, we seek to answer the question whether moving to a less controlled setting can open opportunities for experimental investigations without distorting the relations between individual variables clearly observed in the laboratory. First, we test whether we find the same behavioural effects in the field and in the laboratory. Also, we aim to investigate the dynamics of the two types of experimental markets populated by the same type of participants, in order to assess the impact of the environment on their behaviour. Finally, we aim to close the gap on the field-lab comparison for experimental asset markets with multiple securities. For this purpose, we replicate in the laboratory the field study in Chapter 3, using the same experimental material and the same type of participants. The design in Chapter 3 is sufficiently engaging as a field study conducted over a few days, while being simple enough to be run within one experimental round. This property allows for testing the impact of experimentation in the artificial laboratory environment on the experimental results and behavioural dynamics of the study participants.

1.2.3 Two small extensions

Numerous studies, employing the experimental asset markets, including the seminal study by [39], report substantial mispricing followed by a market crash despite explicitly providing market participants with the information about the “rational” price [see 40, 44, for a review]. [65] challenged this “bubble-and-crash puzzle” by showing that several features, such as presentation of the fundamental value (i.e. the “rational” price of the asset) process in form of a graph rather than a table can diminish or completely eliminate bubbles from the market. They also showed that making participants better understand the task structure by asking participants about the fundamental value at the beginning of each trading round and clear instruction writing reduce mis-pricing.

Several studies investigated the dependency between the information presentation and trading behaviour. [66] demonstrated that framing of the information about particular assets impacts trading behaviour. They showed that telling participants that the purchase price of the stock increased resulted in lower purchasing frequency than in the situation, where participants were informed that the stock’s purchase price decreased. Similarly, when participants were informed that the price of the stock that they own increased, they would sell it quicker than when they were informed that the price decreased, despite the fact that the selling price in both conditions was the same. In contrast, [67] found no framing effect [68] when adding a positive, negative or no description of a traded asset. None of these descriptions influenced the market price, volume and trading strategies. This was because resolving the uncertainty about the underlying outcome probability by providing sufficient explanation reduced the impact of information on the trading decisions.

Information that reduces or enhances uncertainty about the future state of the market may impact the trading behaviour and the market prices differently. In a study about the effect of salience and statistical reliability of information about the traded assets on trading behaviour, [69] found that participants who received salient information with low reliability systematically transferred wealth to participants who received less salient but more reliable information. This was motivated by the fact that, in a situation of high uncertainty, people tend to form more extreme estimates of the price.

Possessing more information about the market does not always yield higher returns [70], while markets with asymmetric information distribution among traders may result in lower mis-pricing [71]. While ambiguous information destabilises the market, reliable signals should help rebalance it. However, [72] argue that judging the quality of the information is the most difficult task and requires the highest skill level. When the information is difficult to judge, investors treat the information as ambiguous, for which they expect higher future premiums for trading based on ambiguous information. Also, their beliefs about the future outcomes are spread among multiple likelihoods.

The way with which the information is distributed plays an important role in pricing particular assets. In a field experiment, [73] demonstrated that markets can be manipulated by putting and canceling bets. The study showed that the market players infer information, from the orders put by other players. [74] found that providing only 50% of participants with information about the profitability of firms whose securities are traded, leads the less-informed players to buying high and selling low, indicating that the less-informed participants indirectly infer the information from the asset prices. This finding confirmed in the study by [75], in which experimental asset market participants without insider information inferred the information from the prices where the insiders partially exploited the information. However, when the less-informed participants were provided with the guidance about the statistical reliability about the provided information, the systematic wealth-transfer from the less-informed to more-informed participants disappeared. Both [69] and [74] point out that, the less-informed participants are over-confident about their investment decisions.

In Chapter 5, we extend the experiment in Chapter 3, and investigate the impact of objective (i.e. reliable and quantifiable) and subjective (i.e. ambiguous) information on market prices. We hypothesize that providing participants with objective information about the "correct" price level of the market, with the clear indication of how to exploit market inefficiencies, would reduce mispricing by moving the price level to the "rational" range. In line with [65], we expect that well-explained reliable information about the correct price should make the market more efficient. Also, we hypothesize that presenting participants with subjective and ambiguous information would make participants form more extreme opinions about prices and should therefore inflate prices of particular assets. However, it is an open question, if and to what degree one piece of this subjective information would "override" the objective information, continuously provided as a quantitative measure of market mispricing. To our knowledge, the latter has not been investigated before.

We present results from two treatments employing Experimental Asset Markets, each lasting two weeks. In this experiment, participants trade assets whose outcomes are uncertain, all participants have the same information and the only source of information in the market is the one distributed by the experimenters. In Treatment 1, we provide traders with an objective quantitative measures of what the correct price of the market would be, accompanied with instructions on how to use market mispricing to generate profit. In Treatment 2, in the market that already includes these objective metrics, we distribute subjective and uncertain information (c.f. an opinion) to half of the participants, while informing all participants that a subjective expert opinion will be distributed to randomly selected 50% of the traders.

1.3 Back to the real markets again: Cryptocurrencies

The markets, which we examined in the field and lab experiments from Chapter 2 to 5, are all relatively simplified. The market prediction power, the reactions to news and information, bubbles mitigation factors, and other issues we studied could be broken down in such simplified markets, and the rationales are intuitive and easy to be understood. However, there is a new market – cryptocurrency market – rising all over the world, which people do not understand the reason of its rise, the fundamental value of the market, and even the nature of the asset.

In 2008, under the pseudonym Satoshi Nakamoto, the decentralized cryptocurrency, Bitcoin [76], and its innovative and disruptive blockchain technology[3] was introduced. From its techno-libertarian beginnings, Bitcoin, and a host of other cryptocurrencies have turbulently erupted into the mainstream. In an overall story of tremendous growth, by Feb 2018 around 1500 cryptocurrencies exist with their total market capitalization hitting an all-time high of $830 billion on Jan 7, 2018, and then crashing to $280 billion in the following month – a sensational drop, but only partially undoing gains made in Q4 2017. Growth potential and market action have therefore attracted huge attention among retail and institutional investors, who are rushing into the new "crypto-world"[4], whose hype is based on the key promise that cryptocurrency technology can deliver decentralized systems that avoid trust and reliance upon centralized authorities, and keep power in the hands of the users. A range of disruptive use cases, some more speculative than others, are foreseen[5]. At the same time, well-known figures from central banks, governments, financial institutions, and other status quo agents, have censured the cryptocurrency space – calling it a "scam" with zero fundamental value. Unsurprisingly, regulators are watching

[3]Powered by a public decentralized ledger that records and validates all transactions chronologically, called the blockchain. These transactions are secured and verified by encryption techniques, and shared between network participants in the absence of a central authority.

[4]According to Fintech research house Autonomous NEXT, the number of crypto hedge funds more than doubled in the four months to Feb 15, 2018 [77].

[5] Such as being a global decentralized currency, avoiding central banks, and "banking the un-banked"; a secure digital asset, within the class of safe haven assets, such as gold or perceived stable currencies; and even a fully decentralized internet, whose protocol hosts a wide range of distributed applications.

the space, and their early statements about potential regulation send shock waves through the market.

Regarding academic studies of cryptocurrency, aside from some comprehensive surveys [78, 79], studies have mostly focused on Bitcoin. This includes: economics [80, 81, 82, 83]; network properties [84, 85, 86, 87]; social signals [88, 89, 90, 91, 92] and price dynamics [93, 94, 95, 96, 97, 98, 99, 100, 101]. Focusing on overall market dynamics and growth mechanisms, some models have been proposed [102, 103], but failed to reliably explain the market dynamics. For instance, in [104], their ecological model predicted a gradual drop of Bitcoin to 50 percent of the total market capitalization in a decade from now, but that same drop then happened within months of the paper being published.

In Chapter 6, we empirically verify that the market capitalisations of coins and tokens in the cryptocurrency universe follow power law distributions with significantly different values, with the tail exponent falling between 0.5 and 0.7 for coins, and between 1.0 and 1.3 for tokens. With a simple birth-proportional growth-death model previously introduced to describe firms, cities, webpages, etc., we validate the proportional growth (Gibrat's law) of the coins and tokens, and find remarkable agreement between the theoretical and empirical tail exponent of the market cap distributions for coins and tokens respectively. Our results clearly characterizes coins as being "entrenched incumbents" and tokens as an "explosive immature ecosystem", largely due to massive and exuberant ICO activity in the token space. The theory predicts that the exponent for tokens should converge to Zipf's law in the future, reflecting a more reasonable rate of new entrants associated with genuine technological innovations. With our analysis of this relatively new market, we argue that the market has the same nature as many other systems and markets. The cryptocurrencies are evolving towards being an alternative investment asset instead of a currency, with massive bubbles and crashes due to its endogenous instabilities[105].

1.4 Contribution

Chapter 2 is based on the submitted paper "Brexit vs Trump: quantifying Pound and Peso market efficiency with a natural experiment"[106], coauthored by Ke Wu, Spencer Wheatley and Didier Sornette. Ke Wu initiated this project and designed the research with Spencer Wheatley and Didier Sornette. Ke Wu retrieved and analyzed the data. Ke Wu, Spencer Wheatley and Didier Sornette wrote the paper.

Chapter 3 is based on the submitted paper "Overpricing persistence in experimental asset markets with intrinsic uncertainty"[107], coauthored by Didier Sornette, Sandra Andraszewicz, Ke Wu, Ryan O. Murphy, Philipp Rindler, Dorsa Sanadgol. Chapter 4 is based on the submitted paper "Behavioural Effects and Market Dynamics in Field and Laboratory Experimental Asset Markets"[108], coauthored by Sandra Andraszewicz, Ke Wu and Didier Sornette. Chapter 5 is based on a unpublished work, jointly done by Ke Wu, Sandra Andraszewicz and Didier Sornette. For papers in Chapter 3 to 5, Didier Sornette conceived the research. Ke Wu, Sandra Andraszewicz and Didier Sornette designed the experiments. Ke Wu and Sandra Andraszewicz conducted the experiments. Ke Wu analyzed the data. All coauthors wrote the papers.

Chapter 6 is based on the submitted paper "Classification of crypto-coins and tokens from the dynamics of their power law capitalisation distributions"[109], coauthored by Ke Wu, Spencer Wheatley and Didier Sornette. Didier Sornette conceived the research. Ke Wu retrieved the data and performed major part of data analysis. Ke Wu, Spencer Wheatley and Didier Sornette wrote the paper.

All of these papers are reprinted with permissions from all of the co-authors.

Chapter 2

Brexit vs Trump: quantifying Pound and Peso market efficiency with a natural experiment

In 2016, two important elections – the EU "Brexit" Referendum and the US "Trump" Presidential Election – took place, injecting turmoil into financial markets, both before, during voting, and after the outcome. In particular, in advance of voting, the Pound (British Pound in US Dollar) and the Peso (Mexican Peso in US Dollars) were taken as barometers for the outcome of the Brexit and Trump votes, respectively. As the votes were announced throughout the night, the two currencies actively responded to this information, eventually settling at a level much lower than before the vote, once market participants had accepted the outcomes. But the financial markets were not the only casualty: the Brexit and Trump outcomes were contrary to the majority of polls, pundits, betting markets, and expert predictions, who very zealously predicted a Remain and Clinton victory, respectively – calling their objectivity, competence and authority into question.

Ex-ante prediction of elections is difficult and plagued by biases in sampling, survey design, and response. Sophisticated methods exist to attempt to deal with this, in these two cases with limited success. In a recent comprehensive study [110], it was shown that bias in ex-ante poll-based predictions can be predicted based on multiple macro-level variables, and therefore corrected to provide a powerful predictor. Here, our focus is different as we investigate to what extent fatally biased conventional ex-ante estimates can be used to predict the outcome of an election in real time, by re-calibrating these estimates based on the stream of voting results. This re-calibration can then repair myriad systematic errors in sampling, response, and – crucially – turnout. This also allows the computation of something of fundamental value: an evolving probability of the outcome is given, starting with the initial ex-ante prediction, and converging to a certain outcome as the voting information becomes sufficiently complete. This sounds like a Bayesian setting, where the posterior reflects the updated prior belief, however we do not use Bayesian methods and hence use the terms ex-ante, real time, and ex-post to indicate quantities before, during, and after the event.

In this chapter, we provide a natural and simple real-time predictive methodology that successfully adjusts for myriad systematic errors in ex-ante polls/predictions. And we use the case of the US Presidential Election to test our predictive methodology, finding that our modest approach was competitive with the trajectory of the Peso market. Using this method, and exploiting the near-experimental properties of the market during the Brexit event, we document the remarkable delay that the Pound market exhibited in reflecting reality, i.e. in converging to the fundamental value by soundly reconciling the flow of voting information with its prior belief. Together with the fact that the market was highly active and fully operational, this provides an example of a major breakdown of market efficiency, as well as market-based predictions [111]. We discuss the impact of the social and political environments in both cases, which apparently had an unusual influence in shaping both the predictions of pundits and the decisions of voters, in particular in the form of an overarching categorization of political correctness – and a political bubble – associated with the Remain/Clinton position. This social climate may have led to the failure of financial markets that were influenced and biased by group-think conviction among the media, pundits and politicians (e.g. [112, 113, 114, 115, 116, 117, 118]).

2.1 Brexit Referendum as a Natural Experiment

On Jun 23rd 2016, the EU Referendum took place from 0700 BST until 2200 BST, and by a majority vote of 51.9% – with about 17.4 Million voting Leave – the people of the United Kingdom (UK) decided to leave the EU. The referendum was held across all four countries of the UK, as well as in Gibraltar, as a single majority vote. Under the provisions of the European Union Referendum Act 2015, there were a total of 382 voting areas across twelve regions using the same boundaries as used in European Parliamentary elections since 1999 under the provisions of the European Parliamentary Elections Act 2002 with votes counted at local authority level. Counting began as soon as the polls closed on June 23, from 22:00 British Summer Time (BST) onwards, making it the first UK-wide referendum to be counted overnight, and took nine hours and twenty minutes to complete.

Over this roughly nine hour period, 382 local voting results were announced, one by one.

This night offered a rare and high quality natural experiment [119], where the market – ordinarily robed in complexity – momentarily exposed herself in a simplified state, allowing an exceptionally objective analysis of response to fundamental information. Further, our modest predictive algorithm – although done with the benefit of hindsight – effectively provides a lower bound for what could be considered feasible and sound for real-time use.

Characterizing this high quality natural experiment in detail, the experimental phenomenon is the response of the market to fundamental information, and the hypothesis is that the market is efficient in semi-strong form. The Pound market in US Dollars during the Brexit voting event provides the response, with the nine hour stream of vote announcements from 382 voting areas being the stimulus. The structure and outcome of the experiment could not be simpler: Leave or Remain in the EU, as decided by a simple majority vote. As in previous event studies, expectations of outcomes of future events are usually taken from survey forecasts of professionals; see examples in [120, 121].

The two outcomes, relating to distinct market regimes, can also be reliably mapped to Pound levels. There was a general consensus amongst major market participants about the market values of the two end states: Having a value of around \$1.4 prior to the vote, Ref. [122] suggested that the Pound would fall to between \$1.10 and \$1.30 if voters chose to leave. In February 2016, HSBC forecasted the Pound to drop $15 - 20\%$ if Brexit would happen [123], and J.P. Morgan forecasted the Pound bottoming at \$1.32 (the exact realized outcome) in a forecast note in January 2016 [124]. Besides these end-state market values of the final result, the public expectations of each local voting result were also formed based on broadly circulated studies. Thus the corresponding "surprise" or impact of each voting result, in the sense of defining the direction of market moves, is clear in principle. For example, the Pound dropped 3% in one minute after the unexpected result of Sunderland was revealed. As will be seen, after our algorithm predicted a Brexit outcome (after vote announcement 20-30), the Pound market remained around \$1.45, pricing in a high probability for the Remain decision. Only after having observed over 300 out of the 382 voting results did it appear to accept the Brexit outcome, with the Pound reaching \$1.32 – the lowest value since 1985. This discrepancy questions the deep-rooted notion that the market was efficient [20, 21, 115].

The experiment was also effectively controlled, avoiding various difficulties present in traditional event studies. Indeed, during the Brexit vote, the local announcements were the only relevant news (the experimental stimulus) affecting the value of the Pound (the experimental response). This relevant information stream was simple, and closely followed by the public via Twitter. Worth mentioning is that, atypically, there were no exit-polls done to bias or possibly instantaneously disclose the result of the election. These controlled conditions naturally avoid the complexity in studying price responses in ordinary market regimes, or on longer time frames, which entangle market impacts of different news, interventions, and other factors.

The Pound market is one of the largest and “most efficient” markets in the world [125, 126], and unsurprisingly the Brexit market was even more active than usual, even reacting with bursts of trading activity to early “private information”, identified by early local results tweeted by non-influential individuals. The operational market efficiency issues are naturally resolved in such a well-functioning market, therefore skirting the potential to attribute findings to standard causes or limits of arbitrage [127]. Moreover, it goes without saying that the market participants were fully incentivized, and in this case well primed and prepared – with all relevant ex-ante information being simple, widely analyzed and discussed, and publicly disseminated.

The Trump case is similar, however with more experimental limitations/defects. Firstly, the structure of the vote is different, where the Electoral College system is not based on the popular vote (which Hillary Clinton received), but instead where the winner needs the votes of at least 270 of the 538 electors, allocated at state level, who in practice give their vote to the winner of the popular vote within their state. In this case, Trump won with 306 electoral votes, Clinton collecting only 232. This makes prediction more difficult, by making the result highly sensitive to the outcome of states near the 50-50 mark – so-called swing states. Next, the Peso market is smaller than the Pound market, and although there was consensus on the effect of the election on the direction of the Peso, there was less consensus on the effect size [128]. Further, the US elections have widely published exit polls. However, the US exit polls – with about 20 questions – often perform poorly [129] and suffer relatively low response rates relative to the simple exit polls of other countries. In this case, they predicted a Clinton victory, with key swing states (North Carolina, Pennsylvania, Wisconsin, and Florida) in favour, all of which ended up voting for Trump. As will be seen from the Peso market, the exit poll was not taken seriously.

2.2 Pre-Election Expectations

In both of 2016 EU Referendum and 2016 US Election, the whole election prediction industry was predicting a substantial win for the wrong side, i.e. Remain in the EU Referendum and Clinton in the US Election. This was reflected in the opinion polls, betting markets and financial markets, each discussed below, with the relevant quantities for beliefs about the Brexit vote shown in Fig 2.1, which includes the evolution of the Pound in 2016, up to a month after the Brexit announcement of Jun 24.

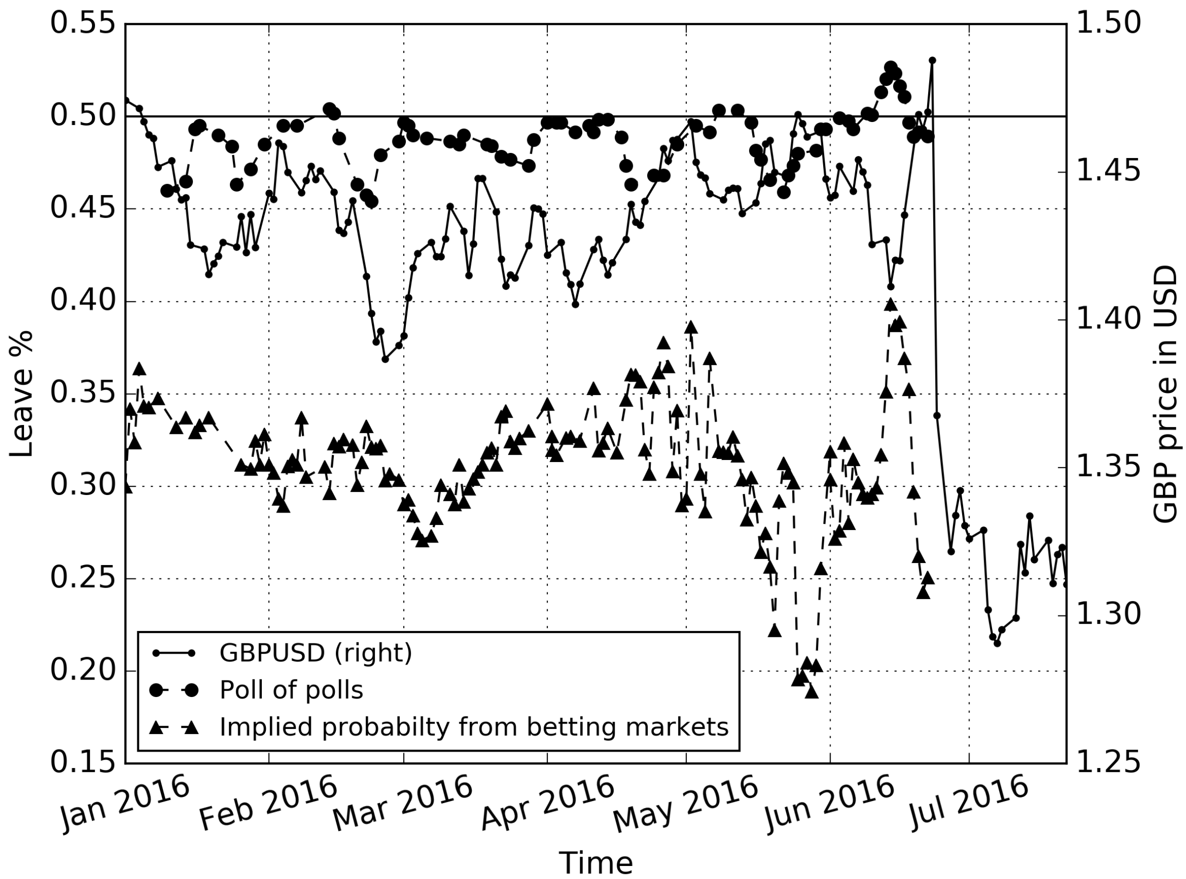

Figure 2.1: Pound market price, betting market odds implied probability of Brexit and poll of polls results of Leave voting tendency. GBP price in USD is plotted with the continuous line with dots (right y-axis). The probability of Leave implied by betting markets is plotted as the dashed line with triangles (left y-axis); the poll of polls showing the Leave voting tendency is plotted with circles (left y-axis). It is calculated by the rolling average of 6 polling results listed by *the Financial Times (https://ig.ft.com/sites/brexit-polling/)*. The betting markets odds data is compiled by Bell [130], from University of Stirling, Management School and Centre on Constitutional Change, based on *www.oddschecker.com*.

2.2.1 Opinion polls

Opinion polls are the most popular indicators for the future outcome of political elections, and are widely reported by the media. Aggregation of opinion polls is considered as an effective prediction methods, and is usually the basis of many sophisticated statistical prediction models [110]. Looking at the “poll of polls” of the EU Referendum, i.e. the average result of the 6 most recent polls, the Remain campaign was always leading with a brief exception one week before the referendum. *YouGov* [131] published an opinion poll just after voting closed, showing that 52% of decided respondents backed remaining in the EU, and another *YouGov* [132] poll, published a day earlier, showed 51% of decided respondents supporting remaining. The margin of error in both polls was quoted as being 3%. Interpreting 52% as the mean, the margin of error of 3% as the standard deviation, and assuming a normal distribution of the probability of Leave, the *YouGov* poll [131] just after voting closed predicts a 32.7% probability of Leave.

In the US Election, Clinton was always leading Trump in the "poll of polls", according to FiveThirtyEight[133] which had collected 1,106 national polls before the Election Day. In the 22 polls published in November 2016, only 2 predicted a victory for Trump, and most were heavily in favour of Clinton.

2.2.2 Betting markets

In betting markets (or prediction markets), participants directly bet on the outcome of a future event. Betting markets constitute another kind of financial market, directly implying beliefs of the odds of the outcome. Due to their simple and direct quantification of odds, betting markets are an increasingly popular method for predicting future events, which have been relatively successful in predicting a large variety of outcomes (e.g. [111, 130, 134, 135]), including the future presidency of the United States (e.g. [8, 136, 137]), sport winners (e.g. [138, 139]), success of businesses (e.g. [15, 140]), movie box office sales (e.g. [137, 141]), etc.

The world's largest Internet betting exchange, *Betfair*[142], initiated a UK referendum betting on February 26, 2016, and received more than 100 million GBP worth of bets for the UK referendum on EU membership. On Betfair, one day before the referendum, Remain was given a strong advantage with odds-implied-probability 77% [143]. On the day of the referendum, the odds-implied-probability of Remain increased to 89% [144]. The betting-market implied probability of Leave, averaged across 20 of the largest betting markets [130] shown in Fig 2.1, consistently favored a Remain result, giving the Remain outcome a probability between 60% and 80%.

In the 2016 US Election, on the day of the election, Betfair showed the odds-implied-probability was 93% for Clinton to win the presidency [145]. PredictWise[146] aggregated several prediction markets results and gave an 89% probability for a Clinton win. The famous prediction market Iowa Electronic Markets also predicted a Clinton win with a 83% probability. All of these prediction or betting markets, which are believed to be a better predictor than polls, were strongly in favour of a Clinton presidency.

2.2.3 Financial markets

Financial markets hinge on (anticipated) probabilities. Financial markets have long been regarded by Financial economists as efficient engines for the aggregation of information, and therefore as effective predictors of the probability of future events (e.g. [1, 2, 3, 4, 5].

Before the EU Referendum, a study of GBPUSD options implied that the market expected the Pound to fall to between \$1.10 and \$1.30 if voters chose to leave, and to rise to around \$1.47 if the UK voted to remain in the EU [122]. On June 22, 2016, the risk of a Brexit implied by currency options was only 19.6% according to Bloomberg [147]. Before the Brexit result was known, the Pound was comfortably above its \$1.38 low for the year, and generally over the \$1.40 level, which had been a floor for the currency since the mid-1980s, as can be seen in Fig 2.1. Ironically, just before the voting results started to roll in on the night of Jun 23rd, the Pound was at \$1.50, the highest point of 2016, leaning toward pricing in a Remain victory.

Days before the 2016 US Election, currencies from emerging markets had mostly priced-in a Clinton win, which would be neutral or positive for all countries except Russia, according to a report from Société Générale SA, and other media reports [148, 149]. Mexico's peso would benefit from a Clinton win more than any other emerging-market currency, which had become a barometer of market-based election expectations. In other terms, the Peso strengthened with signs of weakness for Trump's candidacy, due tho his nationalistic rhetoric, including criticism of NAFTA, and the vow to build a US-Mexico border wall. On Nov 8th, the day of the election, iShares MSCI Mexico ETF touched its highest level since mid-August and closed up 1.8 percent [150].

2.3 Ex-post Analysis of Predictive Factors in the Brexit Case

The two fundamental variables of a vote are (i) preference, which is the fraction of eligible voters wishing for a Leave or Remain outcome, and (ii) turnout defined as the fraction of those voters who actually vote. Regarding voting preference, the simplest ex-ante estimate is from raw opinion polls. One simple version in the EU Referendum case was the Euro-skepticism ranking polls [151, 152]. A more sophisticated one, using demographic information (age, education, ethnicity, etc.) was constructed by Hanretty [34], and was well publicized and widely quoted in the media. We use the one from Hanretty in this paper, and present the result based on Euro-skepticism in the Appendix A. Common systematic errors relating to such polling estimates include a non-representative sample frame (e.g., surveying those accessible by land-line telephone, or people who happen to be around a university), and bias in disclosure (e.g., pro Brexit voters being less willing to participate in a survey, or more likely to respond as undecided). Similarly, polls about turnout need not be reliable. And, importantly, there can be significant relationships between preference and turnout. For instance, in the Brexit vote, overall turnout was 72.2%, being 6% higher than in the 2015 UK General Election, and turnout in pro-Leave areas was on average nearly 4% higher than pro-Remain ones. A well known Brexit pollster Matt Singh [153] attributed this to the re-engagement of formerly discouraged voters, enticed by the prospect of regime-change, calling them the “2.8 Million non-voters who delivered Brexit”. This has been cited as a primary factor missed by pollsters in their ex-ante predictions. In the supplement, we analyze this further. See [154, 155] for rationalization of voting turnout and the consequences of protest voting.

To simplify the analysis, we propose that systematic bias in estimation of preference and turnout can be corrected together, by simply mapping the ex-ante estimates for voter preference onto the actual voting outcome. This will control for turnout insofar as it is predictable on the basis of preference, but neglects additional predictability e.g., due to demographic information or previous turnout history as discussed in the supplement. A linear mapping is natural for capturing this turnout effect, as well as bias in estimation of preference, given the systematic nature of errors. For the Brexit vote, there were $N = 382$ voting areas, with L_i and R_i defined as the number of votes in the i^{th} voting area, corresponding to Leave and Remain respectively, and l_i and $r_i = 1 - l_i$ are the corresponding voting percentage. To estimate the Leave percentage l_i at each local area i, we assume the ex-ante predicted Leave percentage called PL_i $(PR_i), i = 1, \ldots, 382$. For this, we use the prediction made by Hanretty [34]. The mapping is then done by the simple linear regression of observed Leave vote percentages, l_i, against the factor PL_i,

$$l_i = \alpha PL_i + \beta + \varepsilon_i \ , i = 1, 2, \cdots , 382 \ , \tag{2.1}$$

for all voting areas. The result, summarized in the left side of Fig 2.2, is a stable positive linear relationship with a low residual variance of 3.9% relative to the unconditional variance of the response l_i of 10.7%. Thus, on average, the ex-ante estimate can be linearly scaled and shifted to the actual result, with the scale parameter 1.19 and shift parameter -7.95%. In principle, nonlinear mappings could also be considered, however the linear one is perhaps most natural, given that we are dealing with high quality ex-ante estimates of large groups with systematic errors. Same analysis based on crude Euro-skepticism polls is presented in the Appendix A.

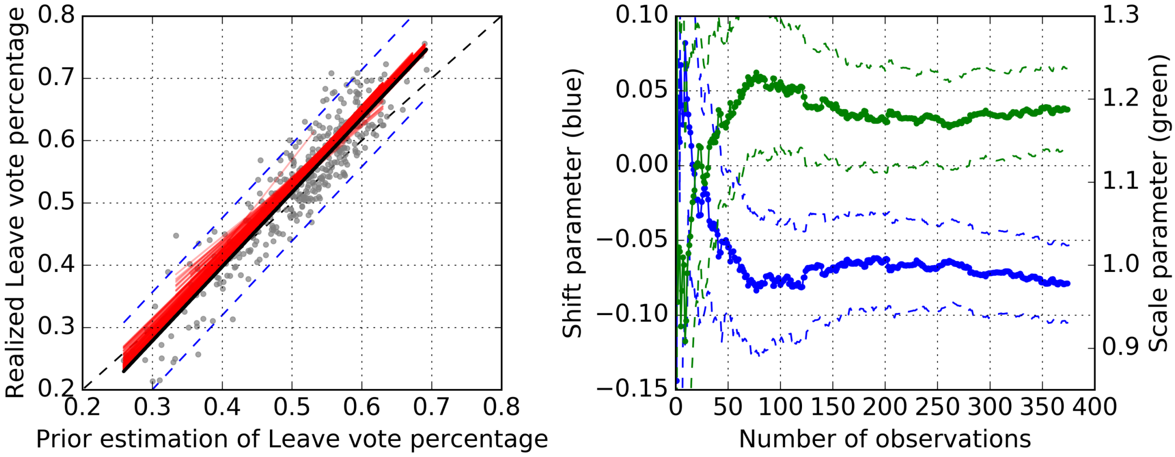

Figure 2.2: Regression of the actual Leave vote percentage against the ex-ante prediction by Hanretty, and evolution/stability of parameter estimates for rolling regressions. On the left side, data is plotted in gray dots, while the black thick line is the regression line for all of the data, together with its 95% confidence intervals (blue dashed lines); red lines are the regression lines for the 378 rolling regressions. On the right side, the evolution of shift and scale parameters are plotted in blue (left y-axis) and green (right y-axis) respectively, together with their 95% confidence intervals in dashed lines.

But what about estimating this one factor model in real-time with only a subset of areas announced? To test this, we perform rolling regressions – fitting only the first 3 observations (l_i for areas 1-3), then 4, 5, and up to all 382 – and compare the parameter estimates. As shown in Fig 2.2, the estimates are remarkably stable, after about 50 areas had been announced. This indicates a consistent positive relationship between the prior Leave factor and the outcomes, and that the areas that were announced early were representative of the later areas. Such a diagnostic would inspire confidence in a real-time prediction setting, which we will present in next section.

2.4 Mock Real-Time Probabilistic Prediction of Brexit versus the Market

2.4.1 The Real-Time Prediction & the Market

In line with the prior beliefs about the outcome implied by the markets, the market did not foresee the result of the referendum. Instead, before the results came out, it had almost priced in a Remain result, reaching to the 2016 highest price at \$1.50. In addition to the surprise, it appears that the market was in denial of the outcome as it was being progressively revealed, while our model predicted it hours earlier. In Fig 2.3, we plot the Pound market, and the evolution of the voting result in the form of the cumulative percentage difference between Leave and Remain votes, where positive values correspond to Brexit. In this figure, the inverse relationship between Leave percentage and the Pound is apparent. The first local result was announced at around 23:36 BST from Gibraltar where 95.91% people voted Remain. Next, the Pound tumbled by 3% due to a surprising result from Sunderland with a 61.34% Leave vote percentage, which had been forecasted by pollsters to be only 56% [34]. After this, the Leave advantage grew, with the Pound falling, then reversed briefly as a series of pro-Remain areas were announced. Around 3am, there were already 111 areas (7,516,379 out of 33,577,342 valid votes) announced, and the current vote was roughly tied. For about two hours, the Pound responded only in proportion (linearly) as the Leave position continued to strengthen, to finally settle at a lead of about 4%. Only once Brexit was a mathematical certainty did the Pound settle roughly to its final level, reaching to a 30-year low value.

We now construct an algorithm to test the efficiency of the market in this case. Following the idea from the previous section, we implement a mock real-time voting result prediction scheme that neither benefits from hindsight nor from privileged data. For this, we use the minimal model that distinguishes between the 382 voting areas using a single essential factor: voter preference to leave or to remain. We use only simple data that was publicly available before the election: The tick data of GBPUSD is from [156]; The electorate (i.e,. number of registered voters by area), as well as turnout for the 2015 UK General Election, was published by the Electoral Commission by June 21st and is available on-line [157]; We take the BBC announcements (via Twitter) as the time stamp of record for the 382 local counting areas.

Formally, we say that the districts, ordered in time by announcement from BBC, are indexed by $i = 1, 2, \cdots, 382$, with announcement time t_i ($t_{i-1} < t_i$), and resulting Leave fractions l_i, Remain fractions $r_i = 1 - l_i$, turnouts w_i, and electorate size S_i. Defining

$$D_{total} := \sum_{i=1}^{382} (L_i - R_i) = \sum_{i=1}^{382} S_i \times w_i(2l_i - 1) \tag{2.2}$$

as the final difference of number of Leave votes over Remain, our objective is to estimate the real-time probability of Brexit,

$$B(t) = \Pr(D_{total} \geq 0 \mid I_t) \tag{2.3}$$

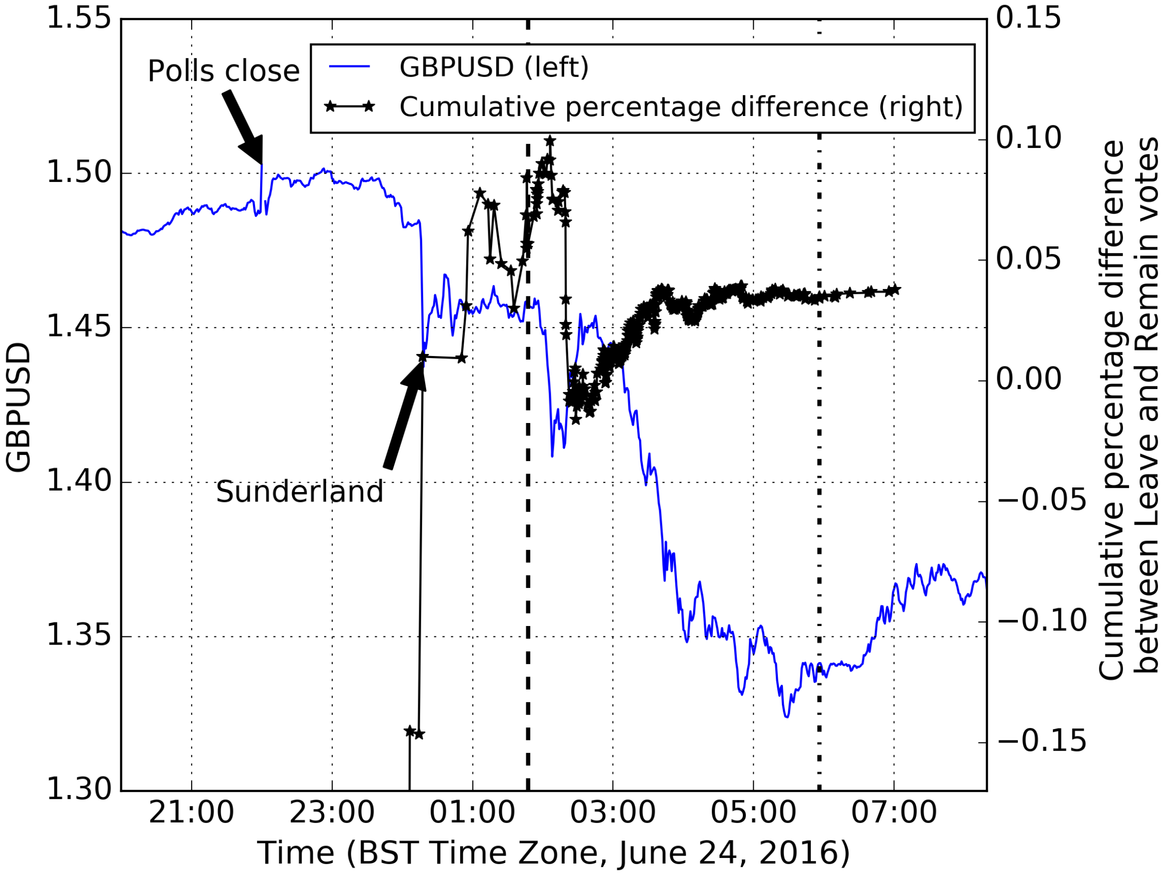

Figure 2.3: Pound value against USD (continuous blue line and left y-axis) and the cumulative percentage difference between Leave and Remain votes (black line with stars and right y-axis). The dashed vertical line is the time when our model presented in section 5 predicted a sure Brexit, while the dashed dot vertical line is the time when the outstanding votes were unable to reverse the result.

where I_t is the voting result information available at time t (i.e., in real time), which consists of the already observed Leave fractions $\{l_1, \dots l_k \; : t_k < t\}$, and turnouts $\{w_1, \dots w_k \; : t_k < t\}$.

Given that the electorate size is known, the model for $B(t)$ needs to account for two random variables: *Leave fraction* l_i and *turnout fraction* w_i. For this, we consider a one-factor model M_1 that only uses the prediction of the Leave fraction, where we regress the observed leave fractions onto the ex-ante prediction as expressed by Eq (2.1) and shown in Fig 2.2. Given this regression, one can then predict the leave fraction for outstanding areas, considering also the randomness inherent in the regression model, as well as the uncertainty of the estimated parameters. For the turnout factor w_i, we take the naive/modest approach of simply sampling it uniformly from the already announced turnout ratios, instead of correlating it with its historically known turnout tendencies. One can then sample leave fractions and turnouts for outstanding areas and compute the total voting difference D_{total} for each sample. The percentage of these samples that result in a Brexit provide the Monte-Carlo estimate of the Brexit probability $B(t)$ (Eq (2.2) and (2.3)). At each new announcement of the results for another area, one additional point is made available, thus enabling an updated calibration and prediction. The detailed prediction procedure can be found in the Appendix A.

We compare our model-estimated probabilities with a market-implied probability, assuming that the market price implies a time-varying expected probability of Brexit $b(t)$, which is the linear combination of fundamental value of Brexit BV and Remain RV, i.e., the mathematical expected value over the two possible outcomes,

$$GBPUSD(t) = (1 - b(t)) \times RV + b(t) \times BV \ . \tag{2.4}$$

If we denote s_0 and s_1 as the price of GBP in USD at the beginning and end of the announcement period, then we can infer from Eq (2.4) that $BV = s_1$ and $RV = \frac{s_0 - b_0 \times BV}{1 - b_0}$, where b_0 is the initial probability of Brexit implied by the market (which we take equal to 50% as a non-informative prior). Thus, the market-implied probability of a Brexit $b(t)$ is,

$$b(t) = \frac{s_0 - (1 - b_0)GBPUSD(t) - b_0 s_1}{s_0 - s_1} \ . \tag{2.5}$$

In Fig 2.4, the model-estimated probabilities, together with the market-implied probabilities, are shown. The simulated probability of Brexit is essentially undistinguisble from 100% after just 20 areas had announced their results and, at the same time, the standard deviation converges to 0 at around 2am BST. At that time, the Pound was still around 1.45, which is still far from the closing level of $1.34 - 1.36$. This result shows that with just a few data points, we can see that the polls were significantly wrong, but the bias could have been adjusted. With such a one-factor linear model that only takes into account the polls, together with the real-time re-calibration and rigorous accounting of uncertainties, we are able to predict the Brexit result very early. In fact, even omitting the factor (the ex-ante predicted Leave percentage) from the model – e.g., in the US case like treating California and Texas as the same – providing a lower bound for a probabilistically valid approach, we find that our prediction still beats the market. In other terms, at some point, it becomes sufficiently unlikely that the outstanding votes, if consistent with already cast votes (i.e., drawn from that population), will reverse the outcome. In this case the market was therefore extremely slow, such that this result is highly significant and robust.

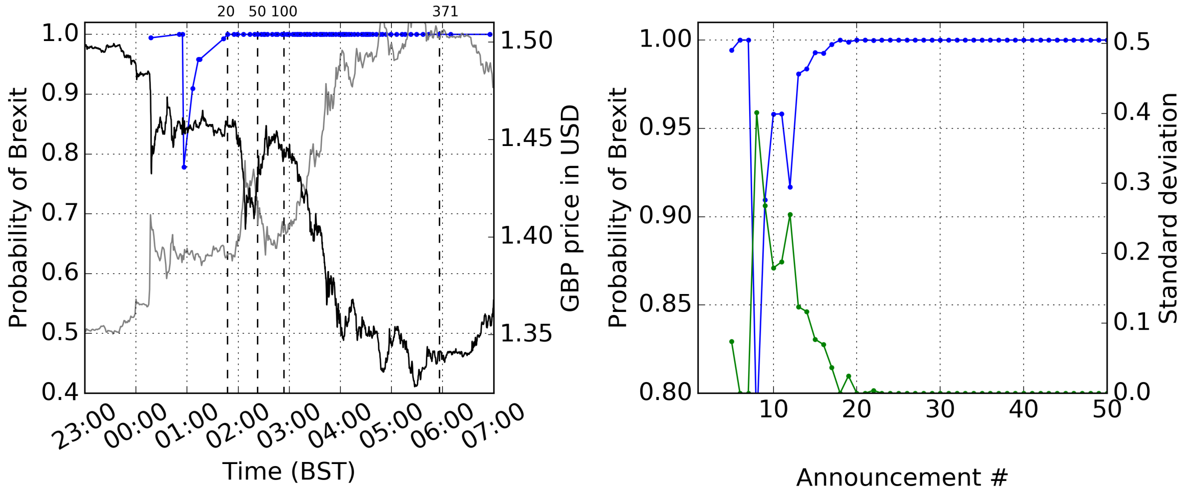

Figure 2.4: Evolution of the Monte-Carlo estimated probability of Brexit based on Hanretty's ex-ante prediction. Left panel: as a function of time, the blue line is the simulated probability, the gray line is the market implied probability, and the black line is the GBP price in USD. All of these probabilities are plotted once for every 5 areas. The vertical dashed lines indicate the time of the 20th, 50th, 100th, and 371st announcements. Right panel: higher resolution of the two probabilities in the range of the first 50 announcements. Blue line shows the probability of Leave, and green line is the corresponding standard deviation of the simulated probabilities. The Monte Carlo estimation is performed with $N = 1000, M = 1000$ simulations, and the initial probability of Brexit b_0 used in calculating market implied probability is 50%.

Such a simple and effective model should have been within the capability of the major market participants. However, the convergence of the price – and effectively the market-implied probability, represented by any reasonable transformations – was embarrassingly slow, implying a market that was some combination of unsophisticated and highly biased. The following subsection shows how one could validate the results of the model in real time.

2.4.2 Real-time Diagnosis of Prediction Performance

To address this question, we define an error measure to assess systematic bias in real time. In detail, at each time step t_k, the *one-step prediction error* is the difference between the observed number of Leave votes $L_k = S_k w_k l_k$ for the recently announced k^{th} area, and the prediction $\widehat{L}_k = E[S_k w_k l_k \mid I_{k-1}]$ made in the previous time step. Therefore, the cumulative error at time t_k is the sum of the observed one-step errors,

$$E_k = \frac{\sum_{j=1}^{k} \left(\widehat{L}_j - L_j\right)}{\sum_{j=1}^{k} S_j w_j} \,. \tag{2.6}$$

Here, we have normalized by the cumulative turnout $\sum_{j=1}^{k} S_j w_j$ to make the error a percentage.

To measure whether this cumulative error is large enough to flip the prediction of the Brexit result, at each time step, we also calculate the predicted final voting difference,

$$E\left[\frac{D_{total}}{TB} \mid I_{t_k}\right] = \frac{\sum_{j=1}^{k} (L_j - R_j) + \sum_{j=k+1}^{382} \left(\widehat{L}_j - \widehat{R}_j\right)}{\sum_{j=1}^{k} S_j w_j + \sum_{j=k+1}^{382} S_j \hat{w}_j} \,, \tag{2.7}$$

where TB is the total turnout.

The error (Eq (2.6)) thus provides our best estimate of the bias of the predicted final percentage difference (Eq (2.7)). In Fig 2.5, we plot the evolution of these two quantities as the announcements out. The model systematically overestimates the number of Leave votes by typically less than just 1%, while the predicted end-state percentage voting difference is always larger than 4%. Thus our real-time estimate of the error does not put our prediction in jeopardy. It is also remarkable that the error lines are always within the 95% null confidence intervals, which indicates that our simple model cannot be rejected at a 5% level.

2.5 Brexit Market Response to Announcements

The failure of the market in predicting the Brexit cannot be attributed to a lack of interest or depth: the trading activity, represented by the number of mid-price changes, was 3 times that of the previous week, and remained high through the next day. And the volatility was 10 times the average value over the previous month. Moreover, the market response to specific announcements could also be observed. For instance, the minute after the announcement from Sunderland witnessed a negative return worse than any other minute in the past 15 years.

To study such responses, it is important to know the relevant set of announcements and their timing. Here we consider two announcement times for each area: the time of the earliest tweet T^1 and the time of the de-facto official source, the BBC tweet announcement T^2, such that $T^1 \leqslant T^2$. In other words, sometimes there was a tweet before the BBC one.

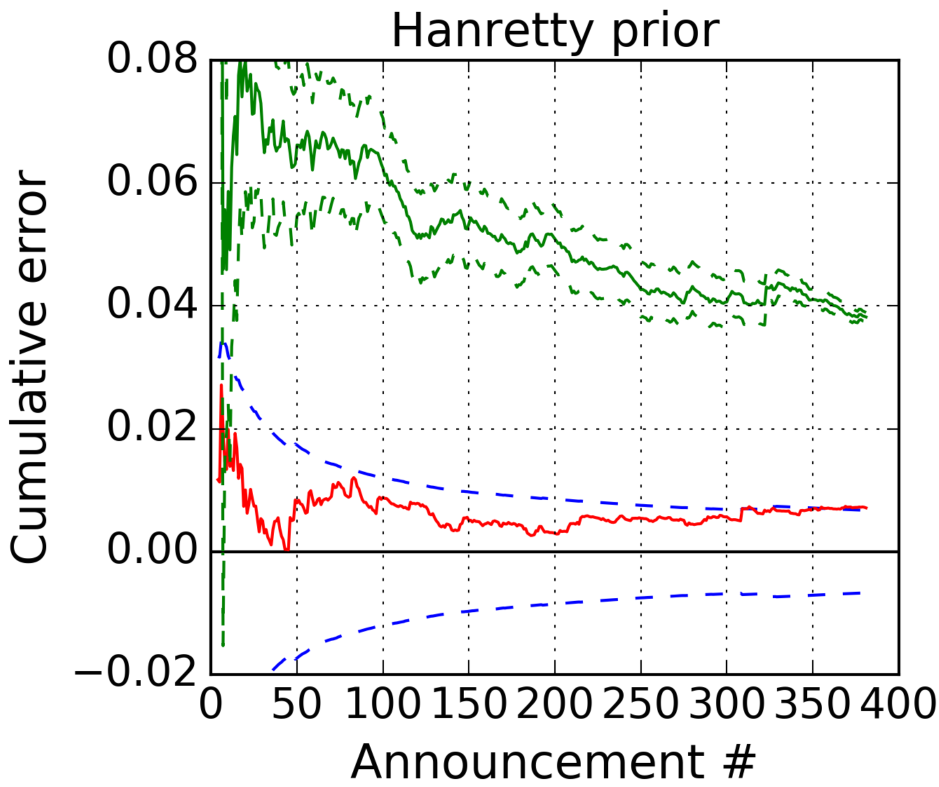

Figure 2.5: Cumulative errors (red) given by Eq (2.6) based on one-step–forward prediction error, together with the evolution of predicted end-state percentage voting difference (green) given by Eq (2.7). The dashed green and red lines correspond to one standard deviation intervals respectively. The dashed blue lines represent the 95% confidence interval under the null model.

There was no official stream from the election authority. Rather, as the count of each area was confirmed, it was announced locally, and then to the media, who then released the announcement on Twitter. The de-facto official stream was thus the Twitter feed of the BBC, which announced all local results in a timely fashion, drawing a wide audience and dominating Twitter activity about the referendum. Hence, we use the time-stamps of BBC tweets as the announcement time for each area. In addition to BBC tweets, we crawled all tweets containing voting result information, to identify the earliest tweet announcing each local result. Among all the tweets, the BBC tweet was consistently among the earliest tweets, and always received the most re-tweets. Early tweets received less attention, as they were mostly from staff in the counting office, local government, and apparently unaffiliated individuals. For each local result announcement, there were on average 6 tweets announcing the results before BBC announcements, and these tweets got

36 re-tweets in total, while BBC announcements got 193 re-tweets on average.

Disappointingly, even in this near-experimental setting, the price response to each announcement was neither clear nor consistent for all but the most pivotal announcements. This was not helped by the high intensity of announcements, leading to a flurry of overlapping and asynchronous responses by different market participants. Further, the assessment of the consistency of price-response is somewhat ill-posed as it requires characterizing a single and steady mind of the market, concerning both the importance and expectations for area announcements. However, the response in trading activity, represented by the number of mid-price changes, was clear and consistent. This measure is less demanding, not requiring a consistent direction of price reaction, but instead simply aggregating activity of different participants, with potentially diverse understandings and strategies. We thus define the trading activity $A(t)$ as the number of mid-price changes from $t-60$ (seconds) to t. To demonstrate the information flow and price response, the Sunderland and Swindon area announcements are presented as two examples. These pro-Leave areas (61.3% and 54.7%) were in the early phase of announcements (5th and 8th), where the announcement frequency was quite low. Fig 2.6 shows that the activity level roughly doubled and tripled for Sunderland and Swindon respectively following the first announcement, being about 30 seconds before the BBC one. For Swindon, one can notice a coincidence between the earliest tweet and a drop of price, but this loss falls within the small 10bps range of subsequent fluctuations – far less conspicuous than the obvious excursion seen in the volume. More importantly, the increase of trading activity before and after the earliest tweet, which received few re-tweets and drew much less attention, shows clearly that the market was functioning well in aggregating and responding to diverse and early information.

For a more comprehensive comparison of responses, we consider the first 100 announcements, which were broadcasted before 2:50am. We restrict ourselves to this earlier announcement period because it had much longer time intervals between announcements, avoiding overlapping responses. To see the evolution of trading activity after the announcements, we stack and average the trading activity $A(t)$ by centering and scaling $A(t)$ at T^1 and T^2 respectively. Specifically, the trading activities centered at T^1 and at T^2 are defined by

$$\overline{A^j}\,(\tau) = \frac{1}{N}\sum_{k=1}^{N}\frac{A(T_k^j+\tau)}{A(T_k^j)}, \tau = -m, -m+1, \cdots, m, \tag{2.8}$$

by setting $j = 1$ or $j = 2$ respectively, where N is the number of areas announced, and τ is the lag or lead time to the announcement time. Here, we use $m = 300$ seconds to study the average evolution of trading activity 5 minutes before and after each announcement.

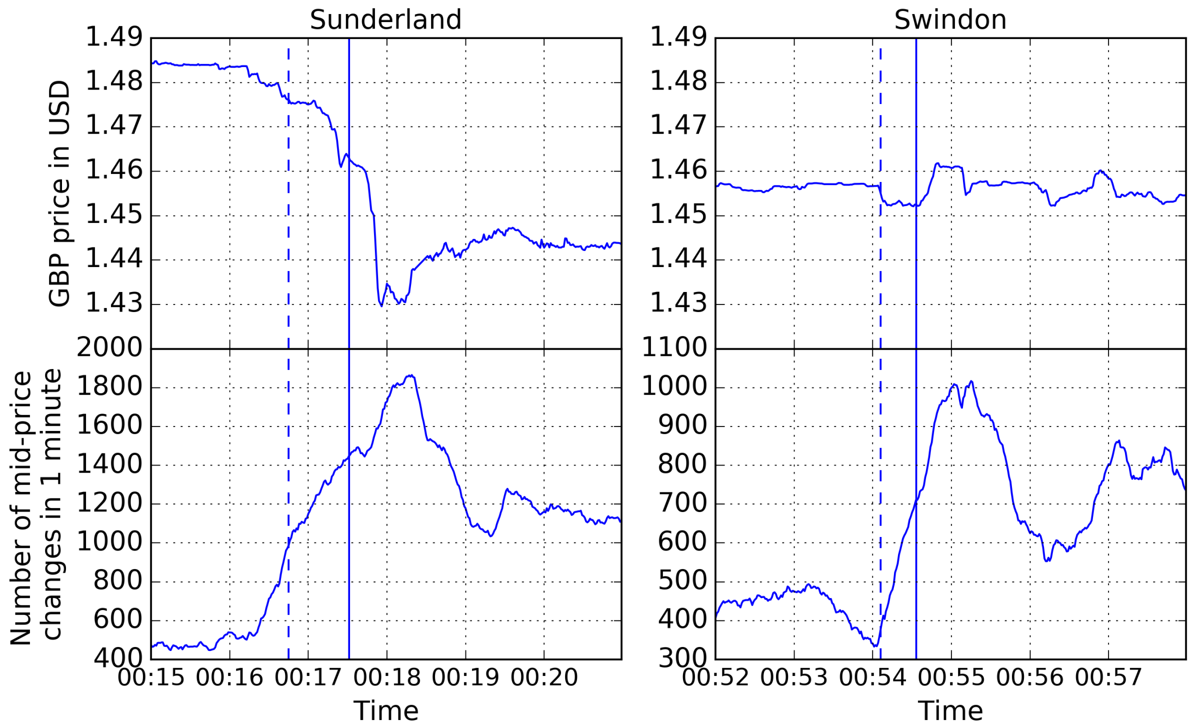

Figure 2.6: Price evolution (upper panel) and trading activity (lower panel) around the Sunderland (left panel) and Swindon (right panel) announcements. The vertical dashed lines indicate the timestamps of the earliest tweets, and the vertical solid lines are the timestamps of the BBC tweets.

As shown in Fig 2.7, for cases with different T^1 and T^2, the activity tends to increase by 7% in 1 minute after the first tweet (despite few re-tweets), and 5% more after the BBC tweet (with many re-tweets). For the case where BBC was the first tweet (27 instances), the completely novel information came all at once and created a much stronger reaction: the number of price changes increased by 15% in one minute after the announcement, and reached 30% in the next minute. It is also possible that the BBC was more timely in its reporting of highly anticipated announcements. Thus, the market quickly captured and aggregated information, even before the official announcement when early unofficial tweets were present.

This evidence confirms that the market was fully functioning in terms of actions, reactions, liquidity, and so on. Further it was successful in aggregating information in a timely way. This also includes remarkably significant reactions to early “private” information identified on Twitter by diverse early tweets.

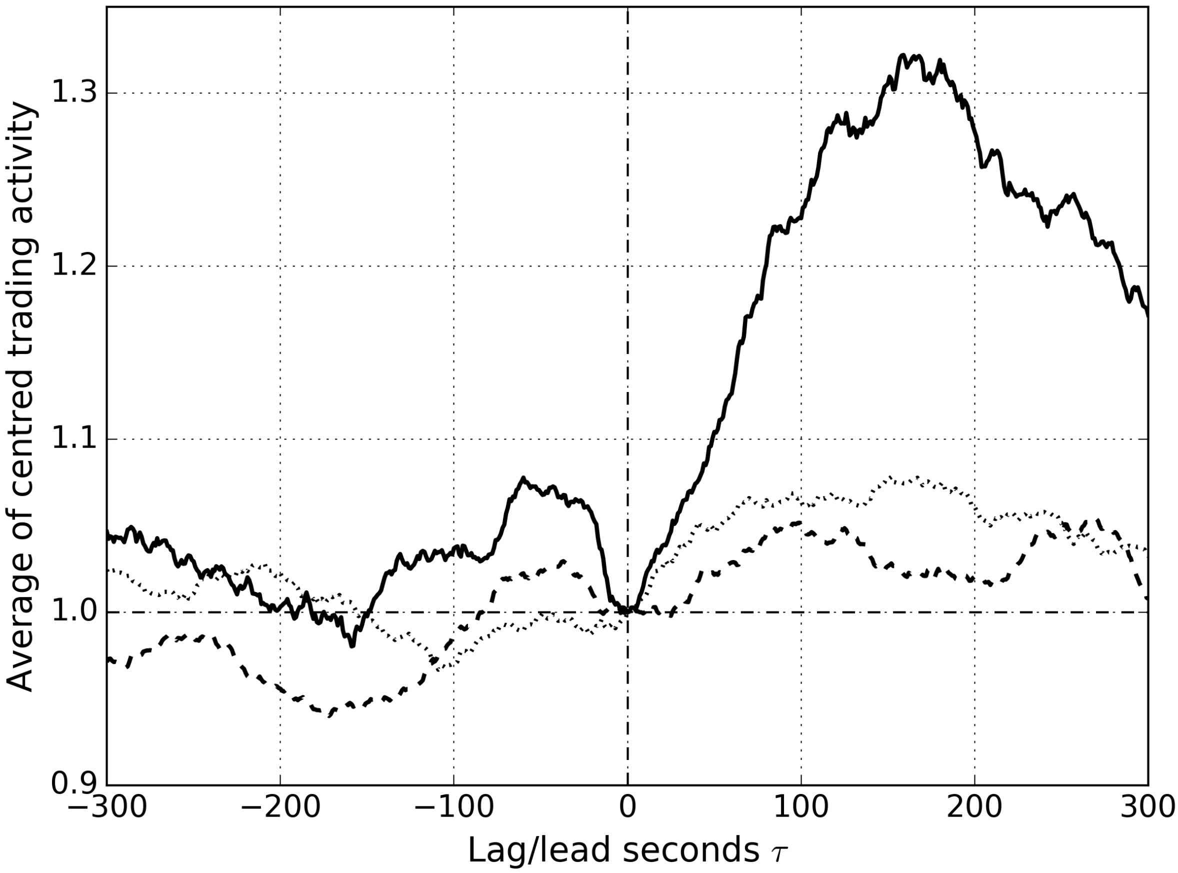

Figure 2.7: Average trading activity (number of mid-price changes), centered at the announcement time, for the first 100 announcements. The dotted line indicates the trading activity centered at T^1 for areas that satisfy $T^1 < T^2$ (73 instances); the dashed line is the trading activity centered at T^2 for these 73 areas; the solid line plots the trading activity for areas where $T^1 = T^2$ (27 instances).

2.6 2016 US Presidential Election

On November 8, 2016, the voting took place to select between Donald Trump as the Republican candidate and Hillary Clinton as the Democrat candidate, to be President of the United States. The so-called Electoral College system stipulates that the winner needs at least 270 of the 538 electors, allocated at state level, which in practice give their vote to the winner of the popular vote within their state. Tallies were continually reported for each state, and by 3:00 AM Eastern Time on November 9, 2016, Trump had secured over 270 electoral votes. In the end, Trump won 30 states yielding 306 electoral votes, and Clinton 20 states collecting only 232 votes despite winning the popular vote by a couple of percent. About 60 percent of the eligible population voted. Importantly, Trump won the hotly contested "swing states" of Florida, North Carolina, Ohio, and Iowa, and also took Michigan, Pennsylvania, and Wisconsin – three states that were expected to vote Clinton with a track record of voting for the Democrats.

We test another market and our predictive methodology on the "Trump" 2016 US Presidential Election, adapted to the Electoral College voting structure. There are 51 local areas (50 states and 1 special district) in the US, which award electoral votes. We take the ex-ante state-by-state US Presidential Election forecasts from the well known pollsters at FiveThirtyEight[133], run by Nate Silver. This ex-ante prediction is based on combinations of polling data only, and was widely circulated in the media well before the election due to their success in correctly predicting the outcome of the presidential contest in all 50 states in 2012. The explanation of their methods can be found in [35].

Regarding the information flow, the voting development was reported live throughout the counting process at state level. We take CNN Election Night Live TV as the data source for this information flow, which broadcast the evolving vote counts[158]. Note that we start to calculate our model when there are at least 5 states having been updated 5 times each. Similarly to the Brexit analysis, the state level results were compared to their ex-ante predictions, adjusting for bias, and enabling the prediction of the outcome – see the Appendix A for details. We don't have the data at the more detailed county level. It is likely that a better model and prediction would be possible at county level – where voting and counting takes place, and where counties are more homogeneous than the large states that contain them.

The Mexican Peso - US Dollar currency pair was regarded as a proxy of the market expectation towards the election result, which became effectively a prediction market for the presidential election during the election night. In Fig 2.8, we plot the estimated probability of a Trump win, together with the US Dollar price in Mexican Peso. The estimated electoral votes for Trump and Clinton are also plotted according to the right axis. The evolving probability follows a similar path to the currency pair from 20:30 until 24:00, when the Peso crashed as the Trump presidency became almost sure. Actually, traders reacted quickly after the intermediate results of a few states were reported. The earliest signal results came from Florida, which completed 90% counting by 20:30, and gave 48.8% votes to Trump, and 48.1% votes to Clinton, while the polls showed 47.5% votes to Trump and 48.1% to Clinton. By 21:00, five more states (Georgia, North Carolina, Ohio, Texas, Virginia) had finished 20%-60% counting, and all leaned to Trump more than predicted by the polls. Unlike in the Brexit case, the Peso market responded quickly on this early information, driving the Peso to crash more than 5%.

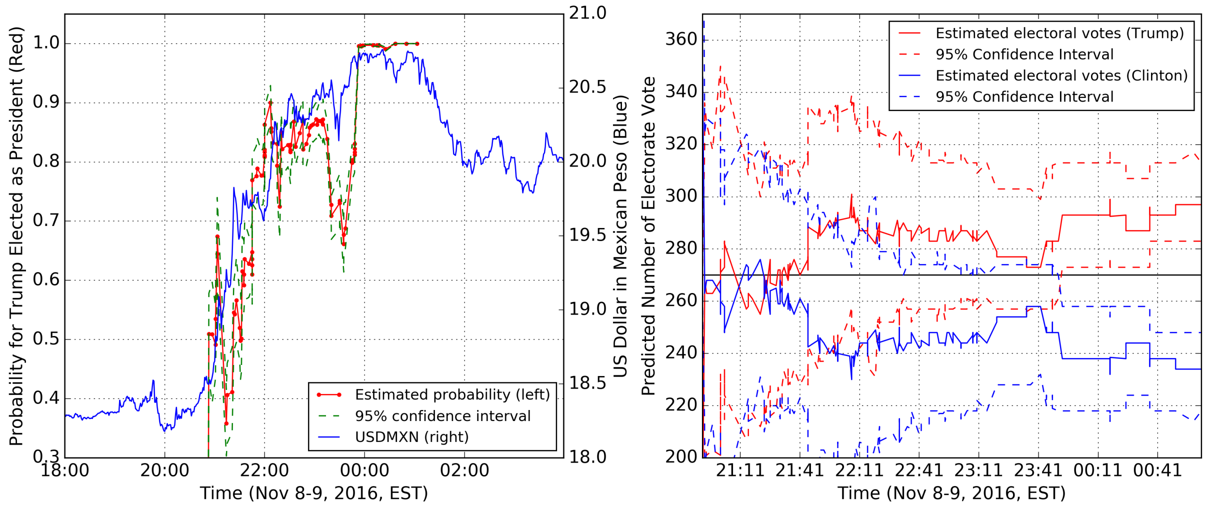

Figure 2.8: Estimated probability for a Trump win and estimated electoral votes for Trump and Clinton. Left: red line is the estimated probability of a Trump win, together with 95% confidence interval in green dashed lines. Blue line on the right y-axis is the one-minute price of US Dollar in Mexican Peso, from `www.histdata.com`. Right: red and blue solid lines are the estimated electoral votes for Trump and Clinton respectively, together with their 95% standard deviation confidence intervals in dashed lines.

Therefore, our simple algorithm with limited state level data managed to compete with the market, in the sense that the most of the market movements are consistent with developments of our estimated probability. Further, as in the Brexit case, a stable linear mapping from ex-ante predictions to actual voting results was uncovered (Fig 2.9). In this case, to adjust for bias and turnout effects, on average, the ex-ante estimate of Trump can be linearly scaled and shifted to the actual result, with the scale parameter 1.31 and shift parameter -12.2%.

2.7 Discussion

Both the EU Referendum and US President Election provided an exceptional natural experiment to reveal how markets endogenize information and to study market response to fundamental information. Crucially, the market – ordinarily robed in complexity – was momentarily exposed in a simplified state, having a single "experimental response" measure being the Pound and the Peso in USD, respectively, a single "stimulus" being the vote announcements, and most importantly being targeted on a binary outcome, whose market values were roughly known to market participants (e.g. [122, 123, 159]).

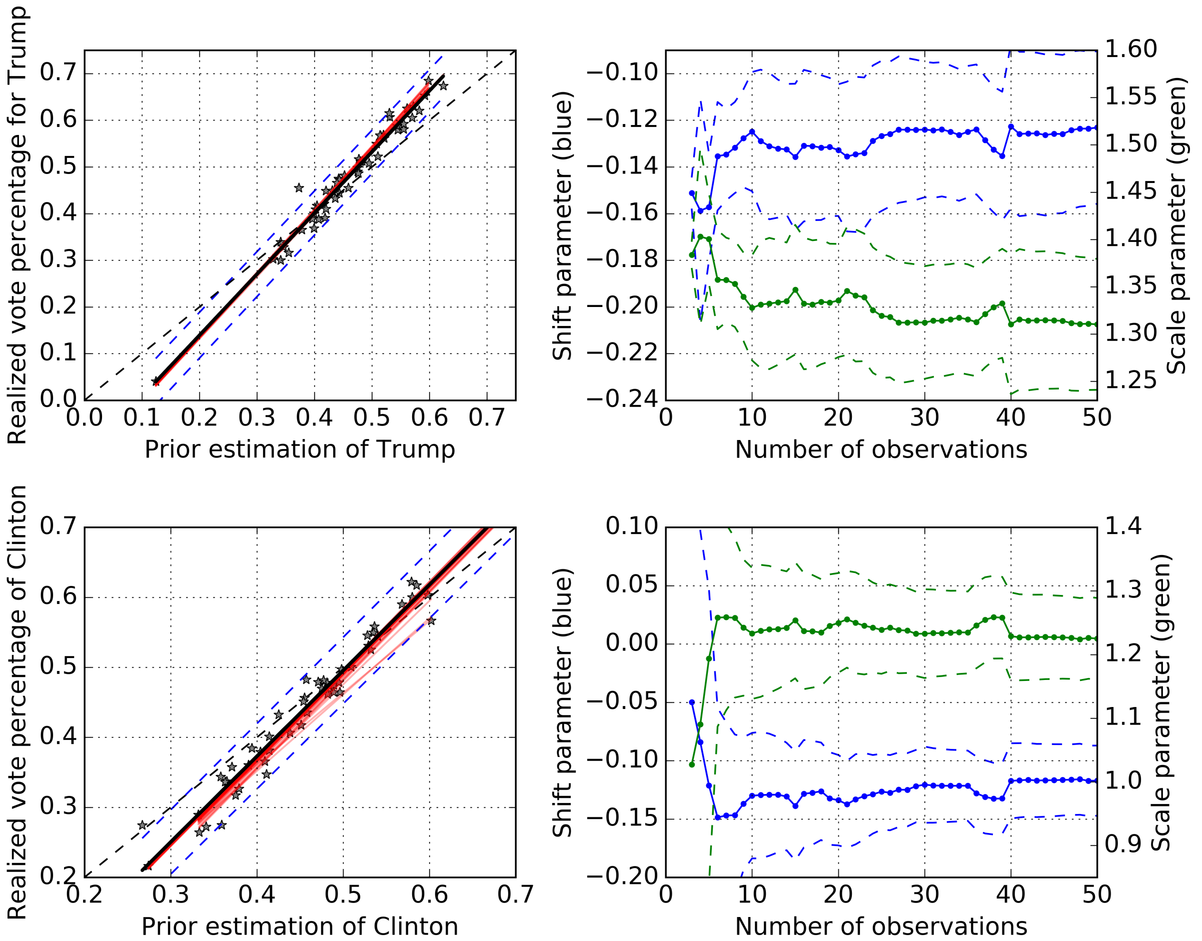

Figure 2.9: Regression of the actual vote percentage against the prior for Trump (upper panel) and Clinton (lower panel), and evolution/stability of parameter estimates for rolling regressions. On the left side, data is plotted in gray dots, while the black thick line is the regression line for all of the data, together with its 95% confidence intervals (blue dashed lines); red lines are the regression lines for the 48 rolling regressions. On the right side, the evolution of shift and scale parameters are plotted in blue (left y-axis) and green (right y-axis) respectively, together with their 95% confidence intervals in dashed lines.

To demonstrate the predictability of the outcome in such an ideal setting, we mimicked the real-time situation of a market participant. With the simplest reasonable model and basic publicly available data, we were able to predict the Brexit result with high confidence after only 20 out of 382 local results had been revealed, and predict the Trump victory in the same pace as the market, with very limited information. This provides evidence that such an approach is useful. The high probabilities early on also demonstrate that the Brexit and the Trump victory was not coming down to aleatory "luck of the day" – in statistical terms, the outcome was significantly different from a tie. Further, it shows the power of our real-time prediction approach based on the calibration of rich but biased ex-ante polls/predictions, by comparing them with early voting information. In the one-factor prediction model, this is captured by the scale and shift coefficients of Eq (2.1) where a scaling (amplification) of 1.2-1.3 was necessary to map ex-ante estimates onto actual voting results, perhaps integrating the effect of so-called undecided voters and skewed turnout. Thus this bias and turnout effect can be quantified and adjusted for in real time.

The contrast between these results and the market responses quantifies a departure from market efficiency, in the form of a strong mispricing of the Pound market during the Brexit vote. While a clean probability was accessible indicating a Brexit early on, the market was more than an hour delayed in reflecting it – apparently waiting until the outcome was a mathematically certainty rather than a virtual certainty. As the market was active and operating well, and comparing with the more efficient market in the Trump election, the most plausible explanation of the Brexit Pound market inefficiency is behavioral. In particular confirmation bias, "the seeking or interpreting of evidence in ways that are partial to existing beliefs" [36, 160] seems highly relevant.

Understanding the formation of biased a priori beliefs requires characterizing the public perception of the Brexit and Trump votes, social norms and pressures ("political correctness"), and underlying identity factors within the population. It is uncontroversial that the both Clinton and Remain votes were characterized generally as being related to personality characteristics such as openness, agreeableness [161, 162], and to "politically correct" [163] ideologies, multi-culturalism, and globalism. Quite naturally then, groups endorsing these position included: the "left" media (and even The Economist[164]), younger people, celebrities (e.g., as measured on social media [165]), and academia [166]. In such a context, Euro-skeptics and Trump supporters – often dismissed by opponents as being racially motivated – may tend to not disclose their preference, allowing their latent right-wing opinion to be counted as undecided, or even left (pro-Bremain or pro-Clinton), in the polling numbers. Thus for ex-ante prediction, it is not surprising that such social and psychological factors are now being exploited (see [162, 167, 168]) and can perhaps be corrected for in the sense of Kennedy[110]. The hidden preferences of the right-wing are revealed, we could say "calibrated", by our one-factor model that was necessary to map prior polls onto actual voting results, as discussed above.

Further, the potential for group-think/herding psychology, leading to ex-ante collective blindness and both real-time and ex-post denial, should not be ignored here. It seems likely that, in addition to generic market inefficiency, there was a herding of market participants away from a "fundamental value" that could have been determined quite early and precisely, as we have demonstrated. Such herding psychology is reminiscent of financial bubbles when the market is dominantly driven by sentiment and no longer reflects a sound indicator of the real underlying value [38]. In the past, financial bubbles have been characterized as the most blatant market failures [37]. In a sense, the market was in a 'Remain bubble', and the British Pound crashed only at the very last stage when the outcome became inevitable, as in the famous SSW design [39]. We propose that the Brexit experiment reported here exhibits a market failure of comparable importance and significance, suggesting that markets can become massively inefficient only when there is a "collective bubble spirit", in a general sense. Here we hypothesize that this spirit was formed by a large-scale group-think based on the political and social attractiveness of the left-wing vote, inflated and galvanized by the intense atmosphere of the election debate. In contrast, in the US Election case, this bubble bursts quickly and the market adjusted to the reality much faster than Brexit night.

It has become clear that there is a "Dragon" [169, 170] hidden in the dis-enfranchised traditionally non-voting population. And this dragon has been ignored by pollsters and pundits that said a Brexit or a Trump victory was out of the question [153, 171]. In other words, the occurrence of a vote for a perceived legitimate regime change may re-engage formerly disaffected voters with a synchronized vote. The larger the population of eligible non-voters, the stronger the degree of frustration, and the greater the perceived potential for regime change – or even revenge – the greater the hazard. In the case of the Brexit, this was clearly a decisive factor. The good news presented here is that this can be quantified rapidly in real time with simple models that combine priors with early revealed information.

Chapter 3

Overpricing persistence in experimental asset markets with intrinsic uncertainty

In this chapter, we present the results of two experiments in which we explored pricing behaviour in the new design. In these experiments, we address the research question of how robust the mispricing effect is, when using a design different from the SSW-design and implementing several features that previously have been shown to mitigate bubbles. In the first experiment, we test the basic experimental setup. In the second experiment, we replicate the design with a few small improvements, test for robustness of the effects found in the first experiment and conduct analysis of traders' strategies. For simplicity, we present the method of Experiment 2 in the main text, while we provide the experimental details of Experiment 1 in Appendix B.2. In Section 3.3, we present results from both experiments. Due to the fact that the core results from both experiments are the same, we present figures from Experiment 2 in the main text, while figures corresponding to Experiment 1 are attached in Appendix B. Appendix B.3 outlines additional analyses from both experiments.

3.1 From SSW Markets to the New Paradigm

3.1.1 The Gap between Real and SSW Markets

The SSW studies share one property - there is a well-defined fundamental value of the traded security, which allows for precise calculations of deviations of the market prices from the fundamental value. The actual "true" value for each period and the probability distribution is either given directly to participants or can be calculated precisely with relative ease.

The fundamental value of a security is a theoretical construct, while some authors (e.g. [172]) claim that it is a "convention" that is extremely difficult to estimate in real markets. The same problem has been faced by experimental asset markets, where various measures of mispricing may lead to inconsistent results [41]. In academic finance thinking, the logic is often turned around by taking for granted that the market price is (almost) always right and any difference from a theoretical value may be due to incorrect choices of the dividend growth and discount rates.[1] The difficulty in quantifying what is the "true" value of a security is often at the source of failures in diagnosing financial bubbles in real time [174].

The information available to agents in standard experiments departs strongly from the situation in real financial markets where the probabilities of possible future outcomes are generally unknown. Ref. [175] point out important differences in decision making in situations under risk (i.e. when the probabilities of events are known) and uncertainty (i.e. when the probabilities of events are unknown) leading to a "description-experience gap" in decision making. This gap is analogous to the gap between real and experimental markets, which can have an important impact on studying mispricing.

3.1.2 Alternatives to the SSW Design

Despite the extensive use of the SSW design in experimental asset markets, other designs have been developed to investigate the dynamics of complex financial markets. For example, a review by Ref. [176] outlines how experimental environmental markets – markets on which one trades tickets/permits for pollution limits or use of natural resources (i.e. fishing quota) – are used to investigate the impact of regulation on individual behaviour of traders in this complex trading environment. Depending on the set of trading rules,

[1] In this logic, Fischer Black once famously observed that "we might define an efficient market as one in which price is within a factor of 2 of value, i.e., the price is more than half of value and less than twice value" [173].

speculative bubbles occur (i.e. allowing for permit banking[2]) or can be diminished (i.e. when "permanent transfers are allowed only after traders have had some experience with temporary lease transfers" [176]).

Another class of studies, uses pari-mutuel betting games, where the market players can purchase tickets for particular state of an event such that tickets are purchased at fixed prices (see [177] for a discussion on this type of markets). An example of a pari-mutuel betting market is a betting market for horse races, where event can have multiple states (i.e. a given horse can end up on place, 1, 2, 3, etc.). Herding (i.e. "betting in disagreement with one's private signal but in favour of the consensus based on prior bets" [177]) is a commonly observed behaviour in this type of markets, but eliciting bettors' beliefs directs their attention more to the probability of each state. These designs, however, are not applicable for studying asset markets.

Some authors introduced changes to the SSW design. For example, Refs. [178], [179] and [180] use a double auction market implemented in Vecon Lab[3], which allows for various dividend generating mechanisms, payoff schemes, transaction costs and taxes etc. All of these studies featured trading sessions lasting 1-2-minute and were repeating 10-25 times by the same group of students. Ref. [178] introduced another change by conducting their study online, with a number of students enrolled in a finance class that could participate in the experiment at a designated time from any place they wanted as long as they had access to the Internet. In all three studies, the dividend was paid out to the stock holders at the end of the trading period and there was one or two assets available for trading. The major change to the SSW design related to the various structures of dividends and fundamental values, including flat fundamental values, random dividends, etc. In a SSW-like design with multiple short trading periods with a single asset or assets with a complete number of states with known probabilities, Ref. [181] implemented dividends that were dependent on the state of events at the end of the trading period.

Given their intrinsic uncertain nature, real financial markets can be conceived as particular incarnations of prediction markets, where the possible outcomes are known while the underlying probability structure of the outcomes is unknown and fundamentally unknowable. Therefore, the participants of prediction markets make "educated guesses", while the market prices emerging from aggregated traders' beliefs should reflect the probability of future outcomes [6, 7]. In financial markets, traders aggregate their beliefs concerning the future performance of firms, leading to prices that can be interpreted as predictions

[2]Permit banking refers to treating permits for the use of environmental resources as assets that can be bought, held or leased.

[3]`http://veconlab.econ.virginia.edu/da/da.php`

of the firm value. Indeed, in the efficient market hypothesis, the present price is equal to the discounted expectation of all future prices. The present price is thus supposed to be informed by all possible future scenarios that impact the value of the firm. It is thus fundamentally determined from the aggregate forecasts of investors on future performance.

In academic thinking, prediction markets, "in which prices are used to predict future events" [7], have been used to successfully predict political elections [8, 9, 10, 11], outcry of infectious diseases [12, 13], sports outcomes [14] and new product blockbusters [15, 16, 17], just to name a few examples. Ref. [18] provide a comprehensive review of the use of prediction markets in the laboratory and field studies. The key focus of these studies is the predictive power of these markets rather than the dynamics of the market or behaviours of the market participants.

One of the mechanisms underlying the predictive performance of prediction markets is the "wisdom of crowds" [182, 183, 184] – a phenomenon in which the weak existing information diluted over many individuals may emerge above the large noise by aggregation over the group. Another mechanism is that experts, and even insiders who have special private information, may reveal their knowledge by trading [185].

All these predictive mechanisms are at play in real financial markets, where the investors exchange their opinion over the bid and ask offers to identify more or less promising trades. In standard SSW experiments, this is not possible, as usually these experiments have just one risky asset and one safe asset (i.e. cash), with known values so that the predictive power of the market in the sense discussed previously becomes irrelevant.

3.1.3 New Paradigm

The new proposed experimental paradigm is a prediction market offering a large number of securities. Over a period of six full days, students of a Financial Market Risks course trade financial assets that correspond to slides of a professor, to predict the page number of the final lecture slide, on which the professor will end in the next week's lecture. The professor always prepares more slides than he can cover, he does not know exactly how many he will cover and he uploads slides a week in advance to a student portal. After the market closes, only one security - the one on which the professor finishes the lecture, pays out the dividend equal to 100 units of experimental currency, while other securities are priced at 0. Therefore, to perform well in that task, one has to trade to make a lot of cash and/or correctly predict the finishing slide by holding the corresponding securities (i.e. the promising venture).

The new paradigm is derived from two distinct experimental approaches - the classical asset market experiments and prediction markets. In this setup, dozens of securities form a complete market associated with all possible outcomes, while the underlying probability structure of the outcomes is unknown and fundamentally unknowable. While all outcomes are known and can be priced by their corresponding assets, there is no objective way by which participants can fully learn or estimate the probabilities of the possible outcomes. This allows us to study how well a trading market aggregates opinions into a consensus and how stable this consensus is with respect to changes in the environment.

The new setup offers various improvements to the SSW design. First, the market participants trade various numbers of assets. In real markets, the number of securities varies depending on the market and traders have to adjust. This set-up thus captures the heterogeneous structure of offered securities inherent in real financial markets. In contrast to the standard experiments with trading periods lasting a few minutes, in our setup one trading period lasts six full days.

Second, participants do not know the expected value of the securities and they (should) know that no one knows them (i.e., there should be common knowledge of ignorance). Our experimental setup focuses on the dynamics of pricing in an environment with limited information.

Further, the trading occurs on a realistic trading platform[4] that replicates the trading rules of the Swiss Stock Exchange Market.[5] The software is equipped with multiple tabs, with the possibility to monitor the price evolution and volume of every asset available on the market, list of the last traded securities and hot stocks, and the possibility to monitor the price distribution of all the securities at any time. This information is supported with charts and numerical values.

Similar to the SSW studies, in our setup, the rational value of each security is determined by the dividend. However, one does not know which security will pay out the dividend, which is analogous to composing one's financial portfolio in a highly uncertain economic and political context. The securities in our setup are a type of an "all-or-nothing" option (binary option). In sum, the market structure of this new design is characterised by persistent uncertainty about the state of the event at the end of the trading period, trading restrictions by not allowing short selling, open communication among the market players and a rank-based incentive schemes.

[4]`https://xyotta.com` : This platform has been internally developed at the ETH Zurich for teaching and research purposes. Researchers who would like to use the platform for their studies should contact the corresponding authors.

[5]`http://www.six-swiss-exchange.com/rule_book/01-RB_en.pdf`

3.1.4 Features Mitigating Bubbles

Our experimental paradigm features several mechanisms that have been shown to mitigate bubbles in previous studies. First, we use equal endowment and a fixed and deferred dividend. Results in [42] and [186] show that a deferred dividend payment, and a single possible dividend, reduces the incidence of bubbles by concentrating common endogenous expectations. With a single bullet dividend, participants focus more on long-term value than on short-term gains through intermittent dividend payments.

Second, our experimental setup features a constant and relatively small cash-to-asset ratio where bidding at high prices is not possible, thus curtailing bubble formation [186]. Ref. [43] reports that increasing the cash-to-asset ratio due to intermittent dividend payments significantly increases the likelihood and magnitude of bubbles.

Third, since the experimental market is open throughout the week, participants will not be required to continuously monitor the market. According to the *active market hypothesis* [187], irrational trading is reinforced when participants do not have any alternative to trading actively (as is the case in a standard laboratory study).

Fourth, we allow participants to openly communicate among each other and the market features an open book. Among others, Refs. [186] and [188] show that this reduces the incidence and magnitudes of bubbles. One possible explanation is that when traders receive information about the motivations, strategies and dispositions of other players (revealed by price, bids, and order evolution), they integrate the optimisation strategy of others in their own strategy. In the game-theoretical reasoning, a trader who has access to the strategies of others will account for the reasoning of others [186].

Finally, our market has a large number of securities. Despite the mixed evidence with only two assets (see [189, 190, 191, 192, 193], for example), the overall direction of these previous results indicates that multiple assets tend to reduce overly exuberant pricing especially if the assets differ.

3.1.5 Goals of Experiments

The main goal of Experiment 1 was to test whether mispricing would occur in a design more realistic than the SSW design, which includes features that should mitigate overpricing. The second goal of the experiment was to investigate the emergence of opinion in a situation of inherent uncertainty and the development of that opinion in the course of coordination among traders over the bid and ask offers. An additional aim of this experiment was to explore the trading behaviour over the six full days of trading in the new setup based on prediction market methodology.

Experiment 2 had three main goals. First, we aimed to measure the robustness of key results from Experiment 1 by replicating with a different group of participants: a) quick price emergence, b) approximately constant price across the week, and c) mispricing as indicated by the market index. The second goal was to explain the emergence of consensus about the price of securities observed in the first experiment. Towards this aim, we introduced a belief elicitation mechanism before and after each trading round. Third, based on the results from Experiment 1, we improved the experimental procedure to close even more the gaps between the real markets and experimental markets.

3.2 Method

3.2.1 Participants

In the Fall semester 2015, 221 students were enrolled in the course of Financial Market Risks for Master students at the ETH Zurich (Swiss Federal Institute of Technology in Zurich). 122 students participated in the experiment and 95 of them took the exam at the end of the course. The experiment was voluntary and offered the possibility to obtain bonus credit points added to the exam grade of the course. The maximum grade for the course cannot be exceeded in the case when the sum of the exam grade and the bonus would be larger than 6.0. 80% of the participating students were male. Due to the personal data protection of students attending the course, we did not collect any demographic data.

3.2.2 Procedure

As outlined in panel B of Figure 3.1, the class was held on Mondays at 10:15am to noon each week. Every Monday afternoon, the professor uploaded slides for the next lecture and these were accessible by all participants. All possible outcomes were known to participants before trading for the week began, so that the market was a complete contingent market. Based on the content of the slides, the participants could form an educated guess about the likelihoods of different outcomes. Their task was to make money by translating their expectations into prices and trading accordingly. On the next Monday, at the end of the class, the realised security was recorded, announced to all students and publicly confirmed by the lecturer.[6] The securities that were not realised did not pay any dividend and were

[6] Professors with more structured lecturing style could ask students to predict the slide that will be uploaded during the lecture at some predefined time, say 45 minutes into the lecture. Even the most

priced at 0.

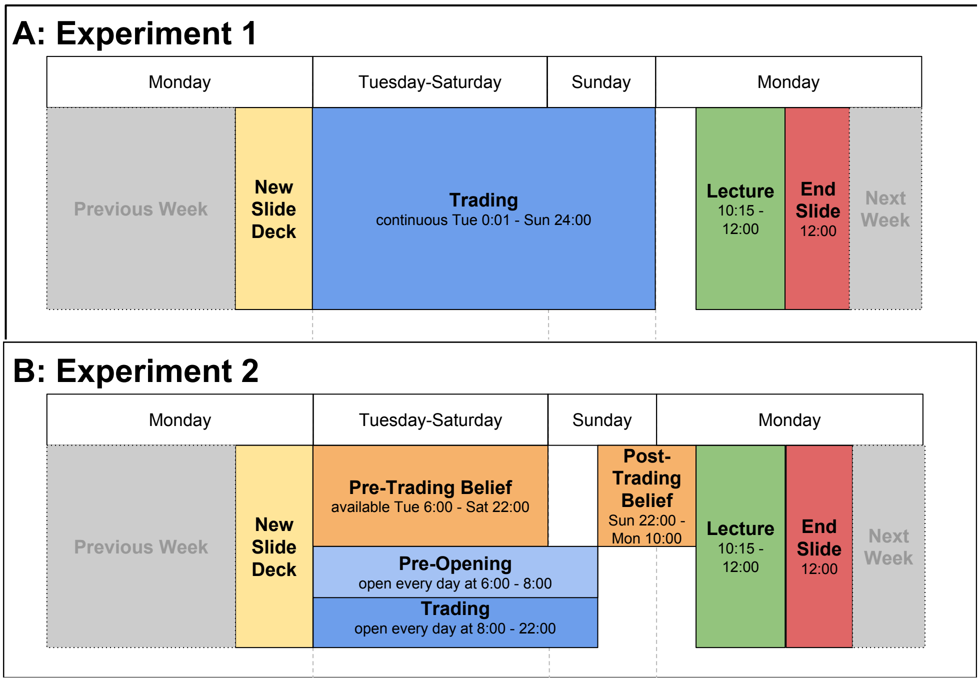

Figure 3.1: A: Timeline of one week of the procedure in Experiment 1. The market was continuously open from Tuesday morning until Sunday night; B: Timeline of one week of the procedure in Experiment 2. In Experiment 2, entering of the market was only possible after submitting pre-trading belief. Participants who did not enter their pre-trading belief could not see the prices on the market. Similarly, accruing the results of one's trading and banking on the payment of the dividend were only possible after submitting one's post-trading belief.

The market was open every day from Tuesday to Sunday between 8am and 10pm. Every day, there was a market pre-opening at 6-8am. We selected the market opening times based on the trading activity in Experiment 1.[7] All buy orders had to be covered by sufficient cash in their account and sell orders were only allowed if the participant had the necessary quantity of securities in their portfolio. No short sells and no buying on margin was allowed. The trading rule follows the standard continuous double auction mechanism.[8] For Experiment 1, the sessions started after the 9th lecture allowing the participants to familiarize with the professor's teaching style and lasted 4 weeks. During lecture 6, the trading task was announced and explained in detail. For experiment 2, the

disciplined and well-prepared lecturer exhibits some variability in her pace, which can be used as the stimuli source.

[7]Results regarding trading activity are provided in Appendix B.3.

[8]A trade was successful only if there was a buyer that wanted to buy one or more units of a security for a price at least as high as a seller was offering.

trading experiment was announced in the second week of the semester. In week 4, the participants could take part in the practice period, which did not count for their final rank. The periods that counted for the final rank lasted for four weeks.

At the beginning of each week (trading period), participants were endowed with 300 units of experimental currency and 3 units of each security, which was equal to 600 units of loan that they had to repay to the experimenters after the market closed. Results in Ref. [194] indicate that unequal endowments and the associated portfolio rebalancing motives are not necessary to induce participants to trade actively. Because the rational price of a complete set of assets is 100 (as further explained in Section 3.3.2), the market had a cash-to-asset ratio of 1. The market was completely reset every week, so no asset or cash was carried over to the following rounds. Before the experiment, the participants could take part in one practice period lasting one week. The purpose of the practice period was to let the participants familiarize with the trading software and the task. Performance during the practice period was not included in the participants' final earnings.

At the end of each week after the realised state had been announced, the cash holdings and any earned dividends were added together to form a ranking of participants. The ranking was based on the following formula for the earnings of participant i in week t:

$$\text{Earnings}_{i,t} = \max\left\{\text{Cash}_{i,t} + \text{Dividends}_{i,t} - 600, 0\right\}. \tag{3.1}$$

The total ranking was not published but the online platform allowed participants to see their weekly ranking. The instructions of the experiment are outlined in Appendix B.1.

In order to activate their trading accounts, every week, the participants had to submit their belief about the success of each security in the market. We used the modified *roulette prior belief elicitation* method [195, 196, 197]. Participants were asked to allocate 100% of their belief among all available securities before and after each week's period. They were presented with a dynamic bar diagram with all securities listed on the x-axis and the belief expressed in percentages on the y-axis. The participants could allocate any natural number between 0-100 to any security and to any number of securities, as long as the sum of the allocated belief was equal 100. By default, the participants were presented with uniform allocation of the probability and could drag and block the bar for each security, while the bars of the remaining securities would automatically adapt to ensure normalisation. After the professor uploaded the new stack of slides on Monday afternoon, he submitted his own belief in the same way as the participants did.

The belief elicitation was not incentivised separately from the trading task. The participants were asked to give their honest belief and were told that their submitted belief is anonymous and will have no influence on their final grade from the course. However, submission of beliefs was the necessary requirement for opening and closing of the portfolio. Therefore, the participants had no incentive to provide false or misleading information during the belief submission, but they were aware that the experimenters had access to their submitted beliefs during the course of the experiment. However, we cannot rule out the fact that some students provided dishonest belief. While submitting the second belief, the participants were presented with their pre-trading belief for reference. The participants could enter the market and activate their account by submitting the pre-trading belief at any time between Tuesday 6am and Saturday 10pm. In order to have one's portfolio in a given week included in the final ranking, the participants had to submit their second belief between 10pm on Sunday (when the market closes) and 10am on the following Monday.

For experiment 2, after completion of the four week trading sessions, the participants were asked to fill out a questionnaire that included questions regarding the cues that they used to predict the end-slide, whether they realised that the sum of prices should not exceed 100, whether they used the opportunity when the prices exceeded 100 to apply arbitrage and what other trading strategies they used. 114 participants responded to the questionnaire.

3.2.3 Materials and Apparatus

Due to the large number of slides used by the professor during each lecture, we grouped them into sets of 3. In other words, one security corresponded to 3 consecutive slides, such that security 1 covered slides 1, 2 and 3, security 2 covered slides 4, 5 and 6, etc. For instance, if the uploaded presentation contained 69 slides, this would give 23 securities, the first security for slides 1-3, the second one for slides 4-6, ..., and the 23rd security for slides 67-69. In each of the four weeks for experiment 2, the lecture slide decks had 168, 157, 144 and 83 slides, which corresponds to 69, 54, 49 and 29 securities. Within each set of securities, there was one security that corresponded to the class not taking place due to unexpected events (e.g. when the professor being sick). To avoid ambiguity, the number of slides referred to the actual number of the pdf page of the slide shown in the lecture.

The set-up was "double blind," in the sense that the participants did not know on which slide the professor would end, and the professor did not know how participants were betting. Moreover the professor himself did not know precisely on which slide he would finish the class. The professor covered 44 (22%), 67 (43%), 61 (42%) and 60 (72%) of all the slides, which corresponds to 15, 23, 21 and 21 securities. Slides that were not covered in one lecture, were added to the new stack of slides for the next lecture.[9]

[9]For a taste of the professor's teaching style, see the video lectures at `http://www.er.ethz.ch/`

The schedule and the course content are the same every year, but the professor adapts his slides depending on the important financial and political events in the world, new research or other changes that should be implemented in the course.

3.2.4 Compensation

At the end of the four trading periods (weeks) of experiment 2, participants were rewarded with bonus grade points that was added to their final grade from the course. According to the ranking, the top quartile (best 25%) of participants with the highest earnings received 0.5 point bonus grade to their final exam grade. The next quartile (second best 25%) received a bonus of 0.25 grade point. The rest of the participants did not receive any bonus point. The grade system is based on a 6-point system, such that 6 is the higher grade that is usually obtained in case of exceptionally good student performance.[10] The minimum grade required to pass a course is 4 and the final grade is awarded with 0.25 steps (i.e. if a student receives 4.15 from an exam, the final grade will be rounded up to 4.25). Therefore, half a point grade bonus can help a student with an exam grade of 3.5 pass the course and is highly valued by the students. Note that students could also receive the maximum grade from the course without participating in the experiment and by performing well in the exam.

During the experiments, we observed that participants had the intrinsic motivation to experience a realistic trading experience, which often is part of the curriculum of many finance courses. Ref. [45] outlines that monetary incentives in experiments do not fully relate to the real monetary incentives and can be powerfully influenced by other motives, such as social aspect or a desire to appear smart, etc. We direct a curious reader to [108], where we provide a discussion on the incentive compatibility in different experimental settings.

media/presentations/Videos.html or the TED Global talk at http://www.ted.com/talks/didier_sornette_how_we_can_predict_the_next_financial_crisis?language=en that was scheduled to last 15 minutes, took 18 minutes and the producers reduced it to 17 minutes.

[10] See explanation of the Swiss grading system here: https://www.swissuniversities.ch/en/higher-education-area/swiss-education-system/grading-system/

3.3 Results

3.3.1 Market Prices and Beliefs

To examine the distribution of prices over each week, we use the median price of all transactions within an epoch (e.g., a 4-hour block) as the price per security. As presented in Figure 3.2 where the securities on the x-axis are sorted consecutively, relative prices reflect the market assessment of the likelihood ratio of two states (the dividend paying out or not). Hence, the distribution of prices provides a direct representation of the market assessments for each week of the likelihood of the lecture stopping at a particular block of slides.

In both experiments, each week, there are a number of securities with essentially zero prices and almost no price fluctuations. For these securities, the participants seem to agree that the corresponding block of slides is very unlikely to be realised. In Experiment 1, we observed a double-peaked distribution of prices, whereas all price distributions in Experiment 2 has a single peak. In three out of four weeks in Experiment 1 and in Experiment 2, the peak of the distribution falls close, but not exactly on the realised security indicating high predictive power of the market. In all weeks in Experiment 2, the market price distribution was more spiky than both belief distributions (as indicated by the smaller entropy of the distributions listed Table 3.1), but this discrepancy between the market and the beliefs disappears progressively from week 1 to week 4, as indicated by decreasing kurtosis differences between the market and the two beliefs across the four weeks (see Table 3.1).

The prices emerged early during each week and stayed relatively constant until the end of the trading period. This pattern occurs in all trading periods (see Figure 3.3 in the main text and Figure B.3 in Appendix B.3). We computed Jensen-Shannon Divergence (JSD) for the end-of-day prices from all six trading days. JSD is a measure of similarity between two distributions (see Table 3.2 in Appendix B.3). In both experiments, JSD values across the week are close to 0 indicating very high similarity. It appears as if participants use existing prices to inform their probability assessments, which can be interpreted as a behaviour consistent with the status quo bias that prevents deviations from initial prices even in the presence of persistent uncertainty [198, 199, 200].

Further, according to Figure 3.2, the distribution of average pre- and post-trading beliefs were very strongly aligned with the price distribution in each week. The post-trading distribution was more strongly correlated with the price distribution ($r = .98, .98, .97, .98, p < .001$ for weeks 1-4) than the pre-trading belief ($r = .91, .97, .93, .98, p < .001$ for weeks 1-4)

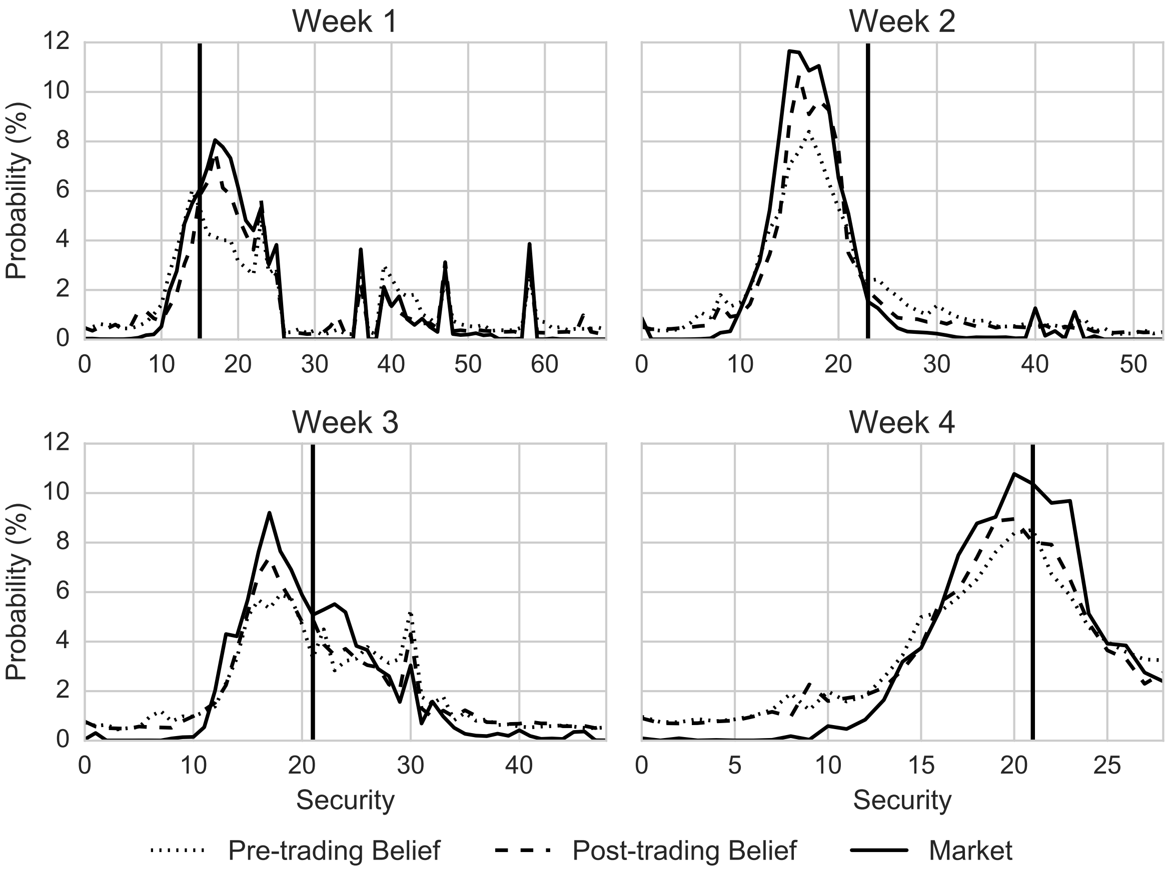

Figure 3.2: Experiment 2: Average distribution of prices based on the median of all transactions in each epoch and pre-trading and post-trading beliefs averaged across all participants in the market. The price is normalised to ensure that the sum of all security prices is 100. The security whose state was realised at the end of the week (payout of 100 units) is marked with the vertical line.

and both belief distributions were slightly less correlated with each other than with the market ($r = .88, .97, .96, .98, p < .001$ for weeks 1-4). We conducted a multiple correlation analysis with the pre-trading and the market price distribution as two independent variables correlated with the post-trading belief (the dependent variable in the regression). According to Table 3.3, the sum of the two coefficients of the two factors is close to one in all 4 weeks, with high R-squared values, showing that the two factors explain the post-trading belief very well. More importantly, the market impact on the post-trading belief decreased across the four weeks, while the impact of the pre-trading belief increased. Additionally, to account for multicollinearity in the multiple correlation analysis, we conducted a linear regression analysis, in which we use the difference between the pre-training belief and the market as a predictor of the pos-trading belief (see Table 3.3). This analysis indicates an

Table 3.1: Experiment 2: standard deviation, skewness, kurtosis and entropy of the market price distribution, pre-trading belief distribution and post-trading distribution in each of the four weeks of the experiment.

	SD	Skewness	Kurtosis	Entropy
		Week 1		
Pre-trading belief	16.22	0.86	2.62	3.76
Post-trading belief	14.96	1.18	3.51	3.62
Market	11.98	.1.52	4.37	3.16
		Week 2		
Pre-trading belief	9.67	1.30	4.86	3.46
Post-trading belief	8.99	6.30	3.51	3.25
Market	6.06	2.75	12.95	2.78
		Week 3		
Pre-trading belief	9.32	0.50	2.63	3.55
Post-trading belief	9.45	0.71	3.60	3.51
Market	6.20	1.09	4.81	3.11
		Week 4		
Pre-trading belief	5.99	-0.72	3.29	3.11
Post-trading belief	5.75	-0.81	3.56	3.07
Market	3.83	-0.28	3.37	2.73

Table 3.2: Jansen-Shannon Divergence of end-of-day market price distributions in Experiments 1 and 2.

Week	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
			Experiment 1			
1	0.29	0.16	0.13	0.21	0.13	0.18
2	0.29	0.16	0.13	0.21	0.13	0.18
3	0.29	0.16	0.13	0.21	0.13	0.18
4	0.29	0.16	0.13	0.21	0.13	0.18
			Experiment 2			
1	0.46	0.07	0.05	0.03	0.04	0.07
2	0.42	0.03	0.02	0.12	0.14	0.07
3	0.37	0.12	0.09	0.03	0.09	0.08
4	0.22	0.03	0.01	0.02	0.03	0.03

increasing impact of the difference between the pre-trading belief and the market on the post-trading belief. Overall, these results indicate that the participants learned that there is no new information in the market and learned to ignore the opinion of the others. In

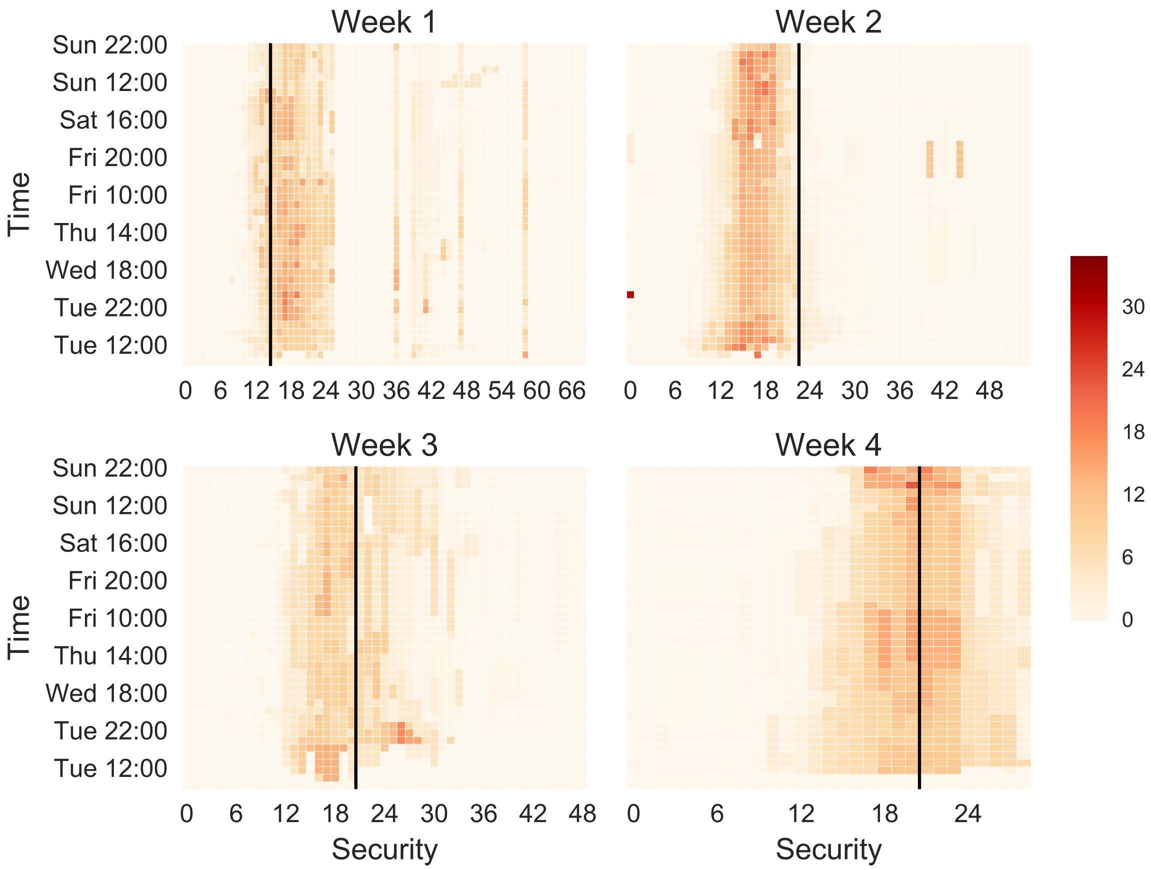

Figure 3.3: Experiment 2: The evolution of security prices over time. Within a trading period (one week) the prices are generally stable; price levels are established rather early in the week and remain relatively constant. In week one, the distribution of prices is multi-peaked, with four securities emerging as favorites. The tendency disappears in week 2, which has a single-peaked distribution and further the distribution becomes wider and less-peaked, indicating larger uncertainty about the success of each security.

other words, we observe the emergence of a communal ignorance across the four trading periods. These strong correlations support the notion that the price distribution was a result of the aggregated a priori beliefs of the market participants, which was then further consolidated in the post-trading beliefs.

Figure 3.4 shows that the individual beliefs of each participant were close to the security that paid the dividend. Overall, all participants assigned belief weights to the same group of securities. However, there is a large variability in the number of securities that each participant assigned weights to - some participants diversified their beliefs among many securities, while some participants indicated that only 2-3 securities as likely to pay out the dividend. Participants were consistent in applying the same strategy of assigning their

Table 3.3: Experiment 2: Multiple linear correlation coefficients showing the dependence of the post-trading belief as a function of the pre-trading belief and the market prices, according to the formula $Belief_{post} = \beta_0 + \beta_1 Belief_{pre} + \beta_2 Market$ (Model 1), and coefficients of the linear regression according to the formula $Belief_{post} = \beta_0 + \beta_1(Belief_{pre} - Market)$ (Model 2).

	Model 1		
Week	β_1	β_2	R^2
1	0.152*	0.737***	0.9628
2	0.503***	0.485***	0.9721
3	0.580***	0.381***	0.9835
4	0.716***	0.268***	0.9755
	Model 2		
Week	β_1		
1	0.345***		0.3650
2	0.522***		0.6180
3	0.646***		0.8010
4	0.756***		0.8330

* $p < 0.05$
*** $p < 0.001$

belief to very few or many securities in pre- and pos-trading belief elicitation. However, they adjusted their beliefs after experiencing the market, such that their post-trading beliefs became less divergent according to the average JSD measures for the pre-trading and post-trading beliefs (Pre-trading: 1.12, 0.88, 0.91, 0.77; Post-trading: 1.05, 0.75, 0.92, 0.72 for weeks 1-4). These adjustments resulted in shifts of the belief by a few securities only.

The belief of the professor was also elicited, following the same procedure as for the participants. The professor's belief was divided among 4 to 5 securities, each having 15%, 20% or 25% of the assigned weight. In weeks 1, 3 and 4, the security that paid dividend turned out to be either the first or the second security to which he assigned any weight. As the securities are ordered by increasing slides, this corresponds to the security paying the dividend being the lowest or second lowest guess of the professor on which slide he will end up the class with. In week 2, the security that paid dividend was before any of the securities indicated by the professor. Overall, the professor was over-confident about the number of slides that he would be able to cover during the lecture. The average distance between the expectation of the distribution in each week and the security that paid dividend was higher for the professor ($M = 5.14$, averaged across all four weeks) than for the participant pre-trading belief ($M = 4.64$), participant post-trading belief

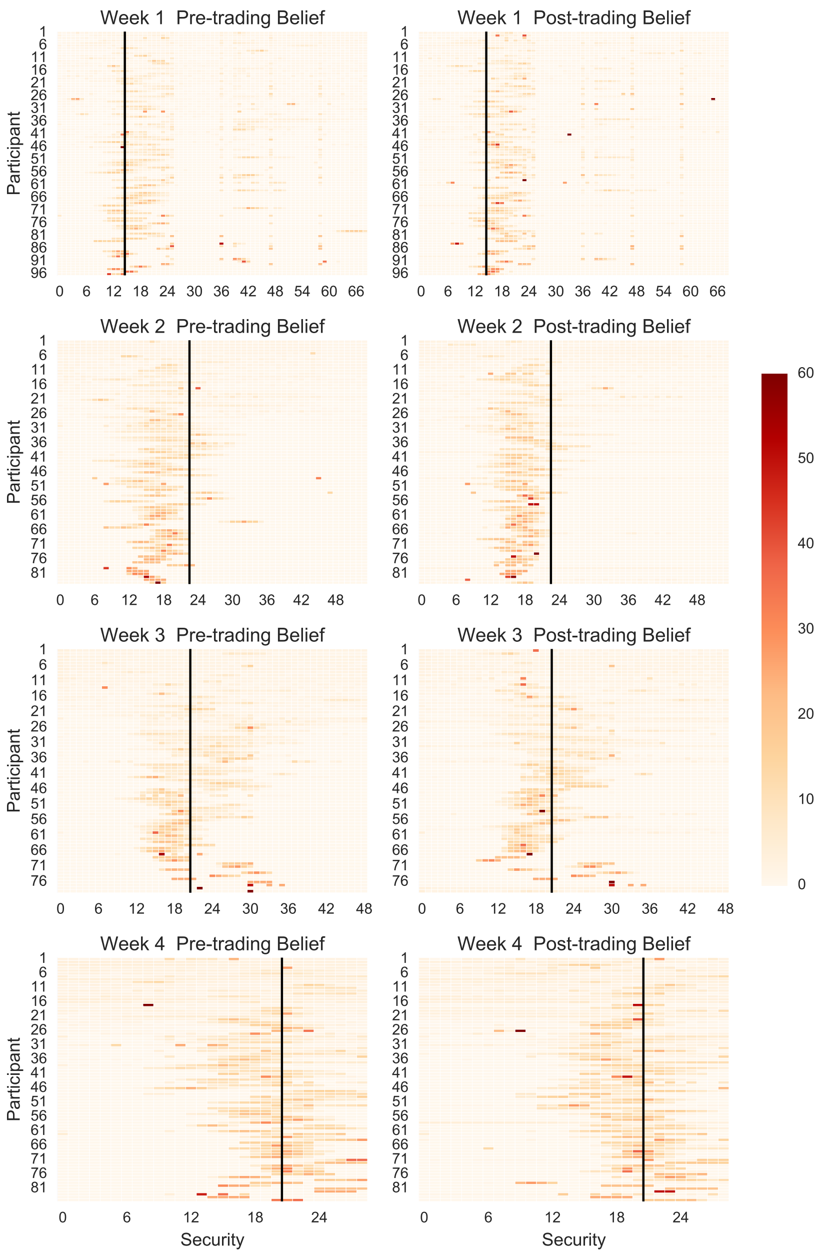

Figure 3.4: Left: Individual pre-trading beliefs; Right: individual post-trading beliefs. The black line indicates the security that paid dividend.

($M = 4.31$) and the market ($M = 3.56$). This confirms the fact that the experiment was double-blind.

3.3.2 Mispricing and Market Rationality

It is important to note that it cannot be said whether the prices at which securities are traded are rational, as there is no knowable fundamental value. However, in our setup, we are still able to make normative statements about the total price level of the market. Since one and only one security pays off 100 units of currency, the sum of the security prices – *the market index* – should equal 100 at all times. If the sum was above 100, it would be profitable to sell one unit of every security and vice versa (assuming that the mispricing would eventually to zero).

Figure 3.5 presents the progression of the sum of all prices for each week in real time. The two other indices provide information on whether the market was a so-called "buyer's market" or "seller's market". Most of the time, one can observe that the sum of highest bid prices is much closer to the smoothed sum of prices than is the sum of lowest ask prices. The latter often tends to be much larger, suggesting that this market was mostly a buyer's market, i.e. supply for securities exceeded demand so that buyers can buy at low prices.

The sums of highest bids and lowest asks can also help identify periods of blatant arbitrage opportunities – if index 2 (the sum of highest bids) is larger than 100, the arbitrage opportunity can be exploited by selling one share of each security (with the sum of the sale values being above 100) and thus obtain a certain profit at maturity. Such strategy would tend to push down the overall price level. If index 3 (the sum of lowest asks) is lower than 100, an arbitrage opportunity would also occur, which could be implemented by buying one share of each security, with the sum of the paid prices being smaller than 100. There is a clear but short-lived arbitrage opportunity in week 1, while arbitrage was possible most of the time in weeks 2-4. Also, every week starts with very high ask prices, which stabilise across the week, apart from week 1, which is characterised by high price fluctuation. In weeks 2-4, the sum of highest bids is almost always above 100 revealing the presence of overpricing.

As shown in Figure 3.5, the market is persistently over-priced in all trading periods. However, in week 4, the index is much closer to the 100-level than in earlier weeks, indicating a learning effect across periods. The common pattern is that the over-pricing is larger at the beginning of each trading period and it decreases towards the end of the period. As in Experiment 1, each week starts with very high ask prices. The overpricing decreases throughout the week but remains until the end of the trading. This mispricing resulted in two arbitrage opportunities in week 1 (see discussion associated with Experiment 1).

To quantify the mispricing of the market, we computed the Relative Deviation (RD)

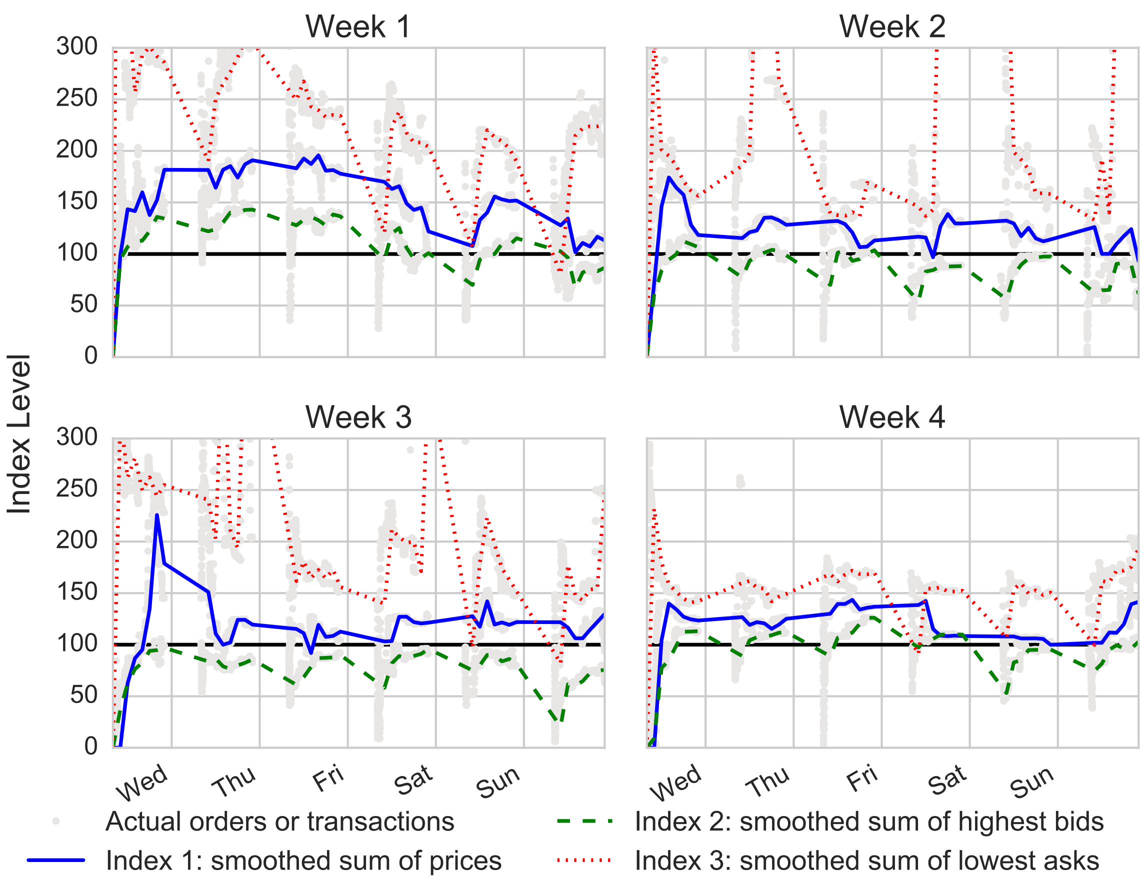

Figure 3.5: Experiment 2: Evolution of the sum of all security prices for the four weeks of trading. The sum of all security prices should be 100 but there are several pronounced deviations from this normative prediction. The indices are smoothed using a 2-hour moving average.

[201]) of the three indices. This measure, outlined in Table 3.4, indicates that in both experiments, the over-pricing decreases from week 1 to week 4 (according to index 1). The Relative Deviation values for indices 2 and 3 indicate that the arbitrage opportunities in all four weeks in the two experiments were very limited.

3.3.3 Grades and Performance

There was a strong correlation between the number of submitted orders and total earnings by participants from all four weeks, Experiment 1: $r = .48, p < .001$, Experiment 2: $r = .51, p < .001$. Also, the participants with the bonus points had higher grades in the exam (Kruskal-Wallis, $p < .001$ in both experiments), such that the median grade of the

Table 3.4: A market mis-pricing measure - Relative Deviation for the three market indices: the sum of prices (index 1), the sum of highest bid prices (index 2) and the sum of lowest ask prices (index 3) in each trading period (week) in Experiments 1 and 2.

Index	Week 1	Week 2	Week 3	Week 4
1	0.3	0.42	0.27	0.21
2	-0.5	0.28	0.06	0.16
3	1.78	0.87	0.82	1.07

participants with 0-bonus was 4.25 (Experiment 1 and 2), of participants with .25 bonus it was 4.50 and 4.75 in Experiments 1 and 2, while the participants with .5 bonus points from the trading would obtain a median grade of 5.63 and 5.5 in Experiments 1 and 2 (without counting the bonus). This means that performance was related to the traders' activity, knowledge and involvement in the course, involving a cumulative effect reminiscent of the Matthew effect [202, 203], which states that "the rich get richer and the poor get poorer", or here in the context of grades, those were the "richest" in their grade obtained the bonus, the "poorest' grades did not get anything. Additional analyses for Experiments 1 and 2 are provided in Appendix B.3.

3.4 Discussion

In two experiments, we tested a new design for investigating experimental asset markets. This design substantially differs from the well-established SSW design [39]. The aim of implementing the new design was twofold. First, we investigated whether the "bubble-and-crash" pattern is a typical phenomenon only found in this type of experimental markets, or it is a general bias that is reflected in other artificial and real markets. Second, we aimed at testing the coordination of opinions in a situation of intrinsic uncertainty.

Our new experimental design is more realistic than the classical SSW paradigm. In the new design, the market players not only have to agree on the "right" price of each of dozens securities, but also have to predict a real future uncertain event whose outcome affects the market. This experimental setup employs a prediction market approach to study stylised results observed in real financial markets and classical asset market experiments. In contrast to standard prediction market studies, the result of our market directly impacts the traders and is experienced by them after the market closes.

To do well in such a market, one has to position his/her portfolio by anticipating the outcome of this event, while using the opportunities in the market to obtain cash. Such situations are very common in real markets. As recent examples, let us mention Brexit in June 2016, the US elections in November 2016 and the Italian referendum in December 2016, whose anticipations influenced the asset allocations of investors according to their beliefs and how they would impact the security prices. Evidence for this can be found in the large price impact that these events triggered on financial markets as a result of the reassessment of investment opportunities following each event, leading to significant changes in portfolio allocations [106].

In the two reported experiments, we observed a few robust effects. First, market players quickly approached a consensus price in spite of the intrinsic Knightian uncertainty, i.e. the lack of well-defined fundamental value resulting from the impossibility to know the probabilities of different outcomes. The market emerged despite the lack of initial price and showed higher predictive accuracy than the professor himself - who was the underlying stochastic process.

The price consensus occurred already in the pre-opening phase as a result of the convergent beliefs of individual players and the price was merely fine-tuned during the actual trading. This is an important finding that reaches beyond standard prediction market experiments, in which researchers traditionally focus on the change of the price over the trading period. Instead, our report of an early agreement on the price reveals the effect of coordination facilitated by the information flow provided by the order book of bid and ask quotes. This transcends the “wisdom of crowd” phenomenon, whose mechanism is based on the averaging of distributed noise to make a small systematic signal emerge by aggregation.

Second, the price emergence was strongly influenced by the prior beliefs of individual market participants, whose initial beliefs were remarkably convergent, despite the intrinsic uncertainty. Participants based their beliefs on vague information in the historical data to extrapolate the future events.

Third, substantial overpricing occurred, despite the features of the new design mitigating bubbles. This shows a general bias that persistently occurs in the markets. However, the overpricing pattern shows much more variability and departs from the typical “bubble-and-crash” scenario found in previous experiments. We found that roughly half of the participants were aware of the mispricing but this situation could not have been always arbitraged due to insufficient liquidity in the market.[11] This could be one possible explanation for why mispricing is so persistent in experimental settings, despite the large number of market participants. Also, not all mispricing observed in the markets may be irrational but instead, it may partly result from various constraints as discussed in a vast literature (i.e. [40, 44]).

[11] See Appendix B.3 for more details.

By replicating the experiment a year later with a different group of participants, we conclude that all observed behavioural effects are independent of the content of the lecture slides. Our two experiments used different decks of slides and started at two different time-points of the semester. Nevertheless, the mispricing is the most pronounced in week 1 of both experiments, the mispricing diminishes over time and the price distribution over time stays approximately constant across the week. The participants were able to adjust their expectations and strategies, given the new securities and differing number of securities. Most of the participants spent relatively little time on analysing the content of the slides, but rather developed technical analysis tools.[12]

In our setup, paying a high price for securities that one likes would lead to underperformance, because the final value that counts for the grade bonus is the sum of cash plus the dividends of only one security (reflecting the real-world "all-or-nothing" security) and does not take into account the value of the portfolio at the last trading price. In other words, the value of the portfolio of each participant was just determined by the payoff of the held securities at maturation, i.e. when the winning slide / security was revealed: namely, 0 value for all securities except the one containing the slide ending the lecture. Therefore, inflated prices that increase the instantaneous values of participants' portfolios were irrelevant for the final valuation of the portfolio of each participant.

The current setup suggests numerous extensions. Most of the design changes we introduced that can be compared to traditional experimental asset markets are such that excessive speculation and bubble formation are decreased. Recall that our market features delayed, bullet dividends, a low cash-asset-ratio, a large number of assets (30-60), a long time horizon (one week), equal endowments among participants and reward unrelated to (or even deterring) high prices. Varying these features would be interesting to understand the conditions under which even more bubbly at one extreme or rational markets at the other extreme can emerge. Also, making the market fully multi-periods where capital can accumulate could result in different price distributions and evolutions during each trading period. Another improvement to the design would be to add news (i.e. subjective information, expert opinion, information about events or recent analyses), which could potentially change the price stationarity observed during the week of trading.

Also, imposing a time constraint on the trading would bring more insight on the active market hypothesis and on the impact that measurements done in controlled laboratory settings have on the market dynamics. Our two experiments were conducted in a semi-controlled environment, where the trading environment was fully controlled by the experimenters and no non-course related events had an impact on the market. However, the

[12]See additional analyses in Appendix B.3.

conditions in which the participants traded were not controlled. This approach mimics well the real-life situation of individual investors.

Further, the inventor of the design is also the lecturer and the first author of the current paper. Prof. Sornette is a lecturer with a particularly unpredictable teaching style, which we tested during Experiment 2 by taking notes and conducting a quantitative analysis on the time devoted to each slide, number of slides covered during one lecture, number of skipped slides etc. (see[108] for a more elaborate discussion on the analysis of the professor's lecturing style). This analysis showed no predictive power of any of these measures. This may not be the case for other lecturers.

Therefore, extensions that would allow for replication of the market design by other experimenters could be necessary, such as predicting the slide or subject reached at a given time within the lecture rather than at the end, as mentioned above. Thus, such a design change would be to make the dividend be dependent on the slide, which the lecturer discusses at, for example, 30^{th} minute in the lecture. Even for very controlled lecturers, it may be difficult to always discuss the same number of slides within the designated time. Another extension would involve predicting the price change of real financial assets. For example, one could ask participants to predict the percentage price change of S&P500 from Monday end-of-day to the end-of-day on the following Monday. This design would be particularly interesting because there would be new information available to the traders during the trading period. While the information would not be controlled by the experimenters, it would be easily accessible. Also, this extension would allow participants to apply technical and fundamental analysis of the real assets underlying the assets traded in the experimental market. [13]

Finally, it is important to note that despite the fact that our market participants had a quite accurate a priori belief about the success of each security, there were persistent errors, suggesting that prediction markets involving real Knightian uncertainty, and financial markets in particular, are useful in the face of intrinsic uncertainty but are not panacean oracles. This could help explain why, in real markets, sometimes resources are allocated to losing ventures and biases are persistent across a substantial number of the market players and over long trading periods.

[13] We implemented this design in Fall Semester 2016 but, due to technical difficulties, our data was not collected correctly and is therefore inconclusive.

Chapter 4

Behavioural Effects and Market Dynamics in Field and Laboratory Experimental Asset Markets

In this chapter, we seek to answer the question whether moving to a less controlled setting can open opportunities for experimental investigations without distorting the relations between individual variables clearly observed in the laboratory. First, we test whether we find the same behavioural effects in the field and in the laboratory. Also, we aim to investigate the dynamics of the two types of experimental markets populated by the same type of participants, in order to assess the impact of the environment on their behaviour. Finally, we aim to close the gap on the field-lab comparison for experimental asset markets with multiple securities. For this purpose, we replicate in the laboratory the field study in Chapter 3, using the same experimental material and the same type of participants. The design in Chapter 3 is sufficiently engaging as a field study conducted over a few days, while being simple enough to be run within one experimental round. This property allows for testing the impact of experimentation in the artificial laboratory environment on the experimental results and behavioural dynamics of the study participants.

4.1 Preliminary considerations

4.1.1 Between the laboratory and the field

Borrowing from Ref. [51], we now discuss five factors that can be used to define the taxonomy of field versus laboratory studies.

First, in the laboratory, usually the participants are students, while field studies would seek to recruit participants within a particular target group. Second, in a field experiment, participants (e.g. finance professionals) can bring specific knowledge about trading, which could affect the experimental market. Third, in the laboratory, participants usually trade abstract assets, while many field studies and natural experiments[1] may use naturally occurring goods. Fourth, Ref. [51] point out that the stakes in experimental asset markets are usually not comparable to the real traders' payments. Fifth, the nature of the task defines whether an experiment is a field study or a laboratory experiment. For example, implementing the SSW design [39] in the field (i.e., on a trading floor) would result in an artificial task, even if conducted at a professional site instead of at the university, and would remain a kind of laboratory experiment.

Our laboratory-field comparison focuses on evaluating whether the strictly controlled experimental environment is necessary for obtaining reliable results. We test whether implementing the experimental task in the participants' natural environment could potentially yield richer data on people's behaviour concerning stock markets. For this purpose, we recruit the same type of participants (students with uniform educational background), who are in general naive with respect to trading with no or little experience, in both laboratory and field study. To equalise the level of information for the field and the laboratory participants, we descriptively present the information that the participant in the field study could experience over a longer period of time. This "story-telling" is often exercised in laboratory studies. This procedure emphasises the direct difference between experiencing a particular process rather than being presented with its description. This difference can influence people's decision making [175]. However, due to time constraints, providing descriptive information about the task at hand is a standard procedure in laboratory experiments. Therefore, our study could potentially reveal the impact of the natural environment experienced over a long period of time, on the market dynamics.

Further, in our study, students trade the same goods in both settings. The assets correspond to the lecture slides of the professor (see below). Therefore, for the participants in

[1]Studies that collect naturally occurring data

the field study, the assets should be similar to naturally occurring goods, while for the participants in the laboratory the securities are an abstract part of the story of the experiment. In both settings, our participants are rewarded competitively and appropriately to the environment in which they act – bonus grades that could help one pass a course (classroom-based field study) and monetary compensation that is substantially higher than student hourly wages (laboratory experiment). Students enrolled in the class may find it natural to receive grades for the task completed along their coursework, while laboratory participants should be used to receive money for performing tasks. In both settings, we strictly enforce the same payoff function.

One could argue that the field study we describe is nothing but a classroom experiment. However, the main goal of classroom experiments in economics is to demonstrate to the students the law of finance and economics for pedagogical purposes. Our aim is different – we aim to test whether introducing an engaging, entertaining and partially educational task to student groups can result in valuable data that could be difficult to collect in the laboratory. In a second step, we adapt the field experiment conducted in a classroom to the controlled laboratory conditions, while preserving the goal and the procedure of the task but in very different conditions, field versus laboratory. Therefore, two groups of participants perform the same task. One group works in a controlled environment within a short time frame. The second groups acts "in the wild" where the task can be performed at their time of convenience and with engagement in the trading environment.

4.1.2 Incentive compatibility

An additional aim of the present study is to investigate whether different types of incentives proposed to subjects to perform experimental tasks lead to compatible results. This topic has gained a lot of attention in experimental economics and resulted in a large literature (see [40, 44] for reference). In economic thinking, the true behaviour can only be elicited if the appropriate monetary incentive is applied. However, Ref. [204] find that different incentive structures can lead to the same results regarding belief elicitation. Ref. [205] claim that the intrinsic motivation of participants can be so high that incentives do not matter or even can be harmful for the task, resulting in over-learning and putting "too much effort".

In order to resolve the debate between psychologists and economists about whether monetary rewards have positive (economic view) or negative (psychological view) impact on performance, Ref. [206] conducted a set of economic experiments in which they found non-monotonic relationship between monetary payment and performance. Their results indicate that high payments increase performance while small payments yielded poorer performance than no rewards. Ref. [207] demonstrated that the brain's reaction to reward is context-sensitive and scales the reward with respect to the possible range of outcomes. Ref. [208] showed that higher hypothetical monetary rewards (i.e. the rewards presented as experimental money rather than small values of real money) result in higher activation of the brain regions responsible for processing rewards. In an fMRI-based study, Ref. [209] found that the same region of the brain – the medial orbitofrontal cortex (mOFC) – is activated when people receive tangible monetary rewards and when they imagine rewards that are important for them. Along the same lines, Ref. [210] showed that the same brain regions (i.e. the ventromedial prefrontal cortex) are involved in computation of monetary and social rewards.

These findings indicate that, on the neurobiological level, monetary or non-monetary rewards have to be well-suited to the context of the task and the scale of possible outcomes, while real tangible money is not necessary to elicit good performance in a task. Along these lines, Ref. [45] criticises monetary compensation by not accounting for other motives, such as the need of performing well in the group.

In their review on the neural underpinnings of intrinsic motivation, Ref. [211] propose a new scientific direction – the neuroscience of intrinsic motivation – which highlights personality, biological and physiological differences in how individuals exhibit intrinsic motivation (i.e. motivated by one's intrinsic motives such as curiosity) as opposed to extrinsic motivation (i.e. motivated by external stimuli such as money). This proposition is in particular motivated by the observations that intrinsic motivation tends to elicit performance in a more persistent way than extrinsic motivation.

Another important aspect of incentives is the way the final compensation is computed. Ref. [212] recall that experiments with multiple trials can implement a variety of payment by the experimenter to the participants: (i) payment based on a single randomly selected round, (ii) payment based on the cumulative performance over all rounds, (iii) payment of only a subset of selected participants or to all of them. Overall, their investigation shows that paying either for a subset of trials or to a subset of participants is the most effective to motivate participants to perform.

Here, we propose that the compensation scheme should be appropriate for a particular setting and group of participants to be compatible with their intrinsic motivation to perform in the task. We use the same compensation function in two experimental settings, where in both experiments only the subset of the best-performing participants receives very significant compensations. In the field experiment conducted with students enrolled in a financial market risks class, the compensation function is converted to bonus credit points. In the laboratory experiment, the same compensation scheme is converted to a monetary payment.

According to Ref. [213], grades work like monetary rewards. In our study, 0.5 of a grade point is valuable and can be decisive of passing a course. The Swiss academic grading system has 6-point grades, with 6 being the maximum grade, 4 being passing grade and 1 being the lowest.[2]. In the laboratory experiment, we offer monetary payment that, for the best performing students, is over 1.5 as much as a standard hourly payment for a student job (27 Swiss francs per hour, in year 2016). In this study, we can directly compare the behavioural effects from experiments that have the same compensation function that is converted to different assets (i.e. money vs. grades), such that each of these compensation schemes is compatible with the experimental setting at hand.

4.2 Method

4.2.1 Field Study

The field study described here corresponds to Experiment 2 in Ref. [107], which provides in depth details of this experiment[3].

In a trading experiment, students of the Financial Market Risks course in Fall semester 2015 in the Department of Management, Technology and Economics at ETH Zurich were trading the lecturer's slides and had to predict the slide on which the professor will finish the next lecture. The professor always prepares more slides than he needs and he does not know precisely himself on which slide he will finish the lecture. The number of slides per lecture varied between 78 and 168. Each security on the market corresponded to three consecutive slides. For the purpose of the experiment, every week, the professor uploaded the slides to a student portal a week in advance. 122 (55% of the enrolled students) students participated. Participation was voluntary and had no negative impact on the students' final grade. At the end of the semester, the best 25% of the students received 0.5 bonus credit point, the second best quartile would receive 0.25 bonus credit point, while the worst half of the students would receive no bonus.

Each experiment had four experimental rounds, each round lasting 6 days (Tuesday – Sunday) preceding the class. The class would take place at 10:15am - 12:00pm on Mon-

[2] See explanation of the Swiss grading system here: `https://www.swissuniversities.ch/en/higher-education-area/swiss-education-system/grading-system/`

[3] We decided to replicate Experiment 2 in the laboratory, because it included important improvements in comparison to Experiment 1.

day. At the end of the lecture, the professor announced the ending slide. The security corresponding to this slide would pay out a dividend of 100 units of experimental currency, while all other securities would be priced at 0. Therefore, to perform well in the task, one would have to trade to either obtain a lot of cash and/or correctly predict the ending slide by buying as much as possible of the corresponding security.

The design is characterised by a few features that should mitigate mispricing: 1) equal endowment and a fixed deferred dividend, 2) small cash-to-asset ratio, 3) trading time lasting six full days, and 4) possibility to communicate among the players and an open order-book. Despite these features, Ref. [107] found substantial mispricing of the market. This mispricing pattern departs from the typical "bubble-crash-scenario" often found in the SSW experimental asset markets [39]. Also, the prices reflected the traders' ex-ante belief about the success of each of the securities. The initial distribution of the price demonstrated a communal agreement about which securities are "good" and "bad" despite the Knightian[4] uncertainty and lack of fundamental value.

4.2.2 Laboratory experiment

Participants

Thirty six students of a Swiss University enrolled in either a natural science or social science programme were recruited over the UAST database[5] to participate in a trading competition experiment. From the UAST participant pool, we selected students with majors (engineering, natural sciences and social sciences such as management and economics) that matched the background of the participants in Ref. [107]. In the invitation e-mail, we informed participants that, in the study, they would compete against other participants and that the compensation will be competitive. The point of providing this information was twofold: to obtain self-selection of participants in similar ways as it occurred in the field study and to comply with ethical guidelines of conducting behavioural experiments (i.e. informing participants about the purpose of the study). Seventeen (47%) of the participants were female, which reflects the standard recruitment procedure in laboratory experiments. The age range was 18 to 32 years (mean age = 24 years). The number of participants corresponded to the full capacity of the laboratory. None of the participants attended the course of Professor Sornette and all were unfamiliar with his lecturing style. This assured that all participants had the same base knowledge about the task, which

[4]Knightian uncertainty refers to a situation in which outcomes of events are known but probabilities of their occurrence are not known and/or cannot be computed.

[5]`https://www.uast.uzh.ch`

is usually the case in laboratory experiments. The maximum capacity of the Decision Science Laboratory[6] of ETH Zurich determined the number of participants.

Procedure

In the invitation e-mail, the participants were informed that they will participate in a competitive trading experiment in which they will play against each other. The e-mail included information about the possible minimum and maximum payment. Participants arrived at the laboratory and were promptly seated at 2pm to randomly assigned seats in the laboratory room. After reading the instructions,[7] the participants watched a movie describing the professor's lecturing style and video instructions on how to use the trading platform, both lasting about 15 minutes in total. Next, a trading task consisting of one practice round and three experimental rounds with the trading time of 10 minutes each followed. The practice round did not count to the final rank and participants were informed about that. After each trading round when the ending slide and the corresponding security were announced, the winning security was priced at 100 while other securities were priced at 0. There were 117, 168, 157 and 144 slides in the practice round and rounds 1-3 respectively, which corresponded to 39, 57, 54 and 49 securities (3 slides per security). The winning securities were 15, 15, 23 and 21.

Before and after trading in every round, the participants were asked to submit their belief about the success of each slide, using the roulette belief elicitation method [195, 196, 197]. To submit their belief, participants were asked to allocate 100% of their belief among all available securities, in any fashion that they wanted, as long as the sum of the allocated beliefs summed to 100. For that purpose, they were presented with a bar graph with all securities listed on an x-axis with uniformly assigned weights to each security. The participants could freely adapt these weights according to their true beliefs. After the trading task, the participants completed a short questionnaire including demographics, trading strategies and the illusion of control [214].

The experiment followed a fixed time schedule that had to be obeyed by all participants. The exact timing of the schedule is provided in Figure 4.1. Each of the steps of the schedule were announced to the participants in writing on a black screen of their computer. We presented the information to all participants at the same time. Participants had access to the previous rounds and their account balance at any time during the trading task. During the experiment, the participants were allowed to take notes[8] on a blank sheet of

[6] `https://www.descil.ethz.ch`
[7] The instructions are included in C.1.2.
[8] The scanned notes can be downloaded from `https://polybox.ethz.ch/index.php/s/`

paper. The notes were collected by experimenters and were anonymous such that they were not assigned either to the real name or the experimental ID of any person attending the experiment.

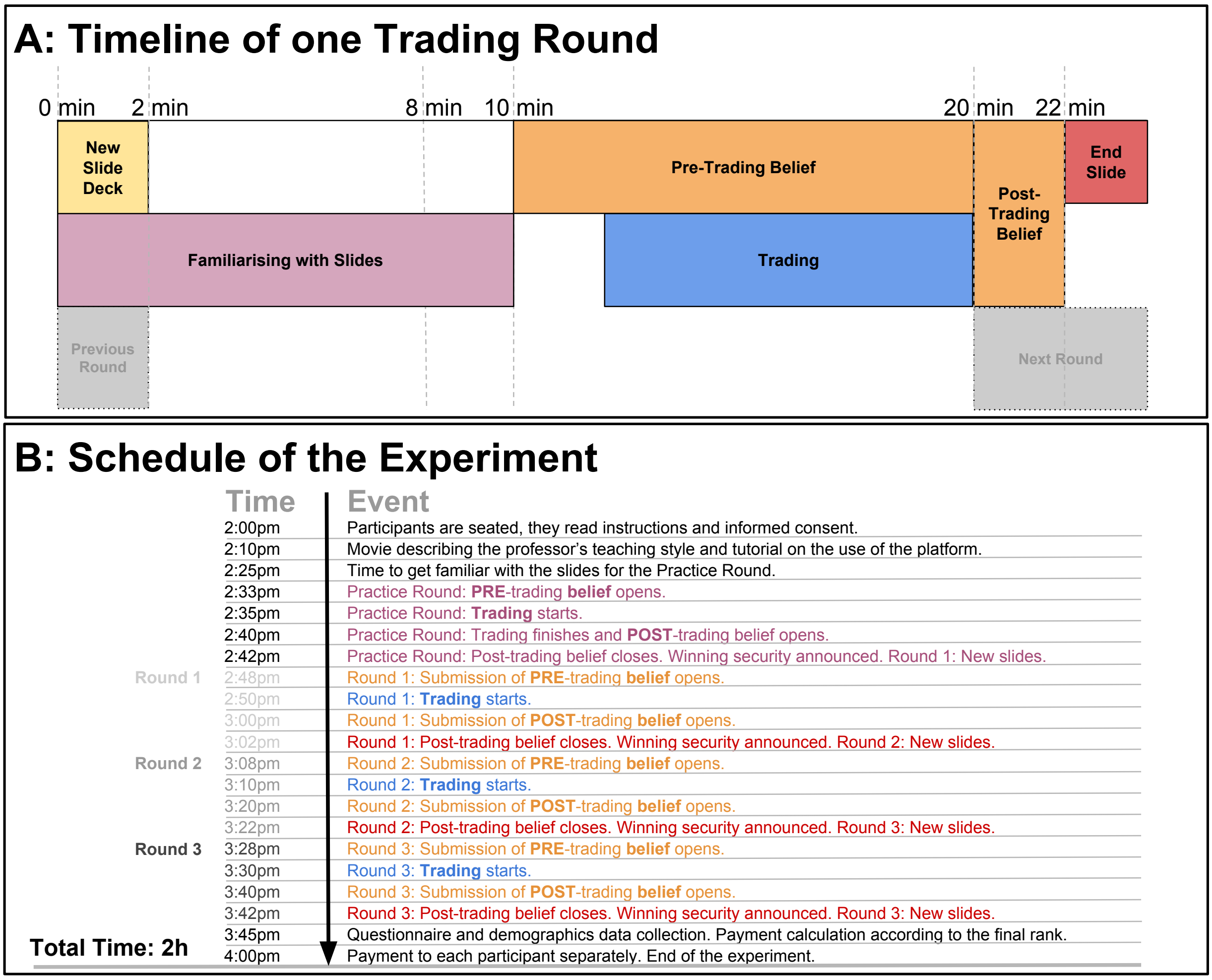

Figure 4.1: A: Timeline of one trading round of the procedure in the laboratory experiment; B: Schedule of the whole experiment including the exact timing that was the same for all participants. The particular elements of the experiment are colour-coded, such that blue corresponds to the trading time, red to belief elicitation and purple correspond to the practice round.

Before conducting the main experiment, we conducted three pilot studies with 6-12 student traders in the room.[9] During these pilot experiments, we calibrated the length of

H5LLGucyK0Ynn89

[9]We do not report the results of these pilot studies because markets in these studies were not liquid enough with such a low number of participants. The purpose of the pilot studies was to set technical issues of the experiment, such as timing.

the individual trading rounds and the length of the whole experiment so that the whole experiment took not longer than two hours.[10]

Compensation

As in Ref. [107], for each round, the trading platform provided a ranking. The market was reset after every trading round and no assets were carried over to the consecutive round. The final rank was calculated based on the sum of earnings in each of the three trading rounds.

The best 25% (thus 9) of the participants in the final rank received a bonus of 60 Swiss francs (worth approximately 60 US dollars), the second best 9 participants received a bonus of 30 Swiss francs and the worst 18 participants did not receive any bonus. This bonus scheme was intended to correspond to the payment of 0.5 and 0.25 of the grade credit points awarded in the field study as described above. All participants received a show-up fee of 30 Swiss francs, which was compliant with the rules of the Decision Science Laboratory of ETH Zurich. Therefore, the top performing students received 90 Swiss francs for a 2-hour experiment, which is very attractive compared to the standard participant payment in Zurich of about 27 CHF/h.

4.3 Presentation of the main results

For the purpose of direct comparison of the laboratory and field studies, we provide results from the laboratory and contrast them with the findings from the field study presented in Ref. [107]. Each comparison comes with a discussion about the similarities and differences between the two experimental settings. Please note that, while we expected differences between the laboratory and field settings, we did not have clear expectations on the nature of these differences because our investigation provides the first such direct comparison for a study with a large number of participants and of traded securities. We further summarise this analysis in Section 4.4.

4.3.1 Trading activity

We observe an increase of participants' activity from the first to the third round. The total number of orders increased from 676 in Round 1, through 844 to 922 in the final

[10] The experimental procedure has been approved by the ETH Zurich Ethics Committee.

trading round. As shown in Figure 4.2, the number of transactions increased within the first 1-5 minutes (5 minutes equals half of the trading time), when it reached the peak and then fluctuated at around 30 transactions per minute. This indicates that participants learned the task and started to react quicker in later rounds. This pattern of trading activity in the laboratory is in contrast to the trading activity in the field experiment, where the number of orders in each round decreased across rounds and the activity within each trading round had a clear cyclical pattern, with the daily peaks of activity in the morning and in the evening and the weekly peaks of activity just after the market opened and just before it closed (similarly to real financial markets). We do not observe such patterns in the laboratory.

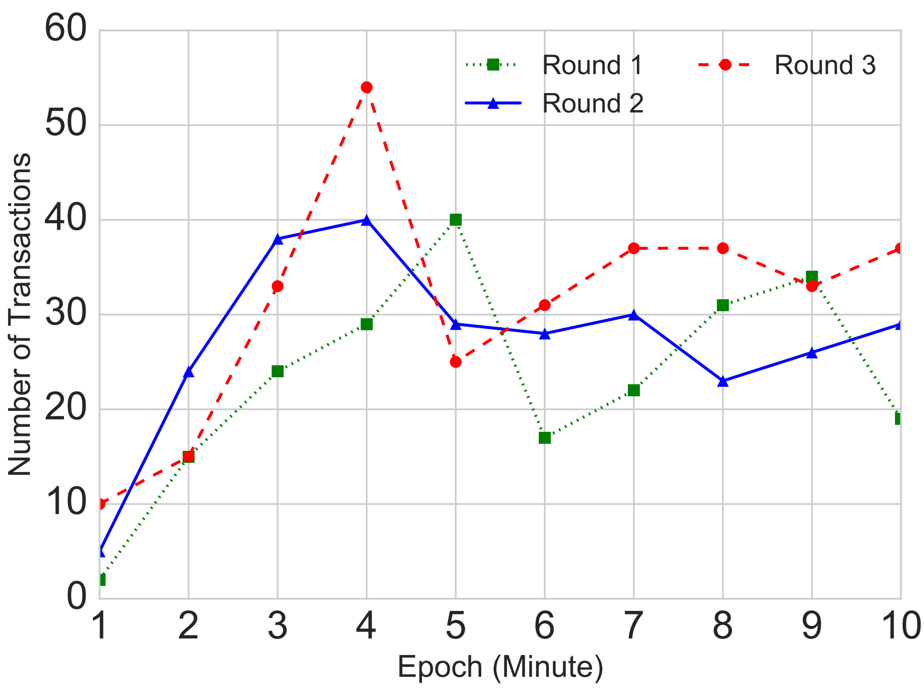

Figure 4.2: Number of transactions per minute in three rounds. The figure shows how the number of transactions changes during the trading time.

On average, each student submitted 18.8, 23.4 and 25.6 orders in rounds 1-3. This indicates very high activity during the short trading periods lasting 10 minutes, compared to the classroom setting with the average number of orders per students within the 6-day period would equal 34.1, 26.5 and 19.7. We surmise that this increase in activity in the laboratory was related to learning and improving at the task. In contrast, the decreasing activity in the field could have resulted from the lack of interest in the task or improvement of trading strategies such that one would become more efficient with fewer trades. To correctly disentangle these effects, it would be necessary to conduct an experiment in which experienced students trade within a limited time-frame, e.g. 2 hours.

Figure 4.3 shows the trading volume of each security in the laboratory and in the field experiment. Similarly to the classroom setting, the prices in the laboratory market were strongly correlated with the trading volume ($r = .73, .81$ and $.91$, $p < .001$ for Rounds 1-3). While, in both settings, the security listed as the first one (i.e., left-most) has a relatively high volume, in the laboratory the volume of that security was higher relative to the volume of other securities. This is especially pronounced in Round 1, where the first security on the list (i.e. Security 1) was traded twice as much as the next most traded security. Also, in all three rounds, securities with larger numbers (on the right tail of the probability distribution) exhibit little or no activity. This is due to the limited time of laboratory trading rounds that restricted exploration and exploitation of all available securities.

4.3.2 Market prices and participants' beliefs

Figure 4.4 shows that the price distribution emerged in the first 30% of the total trading time, compared to the 6% of the available trading time in the field experiment. However, in absolute terms, the price emergence in the laboratory was very quick as it took only 3 minutes, likely forced by the fact that all participants had a strictly designated limited trading time.

Further, in Round 1, only Security 1 had a price in the first minute of the trading round. Also, the securities that were priced early during the trading were much more expensive than the securities for which the price is established later in the trading round. The prices of these first securities diminished after minute 3 of the trading. We did not observe a similar pattern in the field experiment. In the laboratory experiment, 23, 17 and 10 (40%, 31% and 20% of available securities) securities remained without price in Rounds 1-3, in comparison to none in the field experiment. The fact that fewer securities remained without price across rounds shows that participants learned to explore all available securities and traded them.

Once the prices of the securities were established, the "expensive" (i.e. good and possibly paying out the dividend) securities remained expensive and the "cheap" (i.e. bad and possibly not paying out the dividend) securities remained cheap till the end of each trading round,[11] which replicates the effect observed in the field experiment. This conclusion is

[11]The distinction between the good and bad securities is based on the median split of the final price at the end of a trading round.

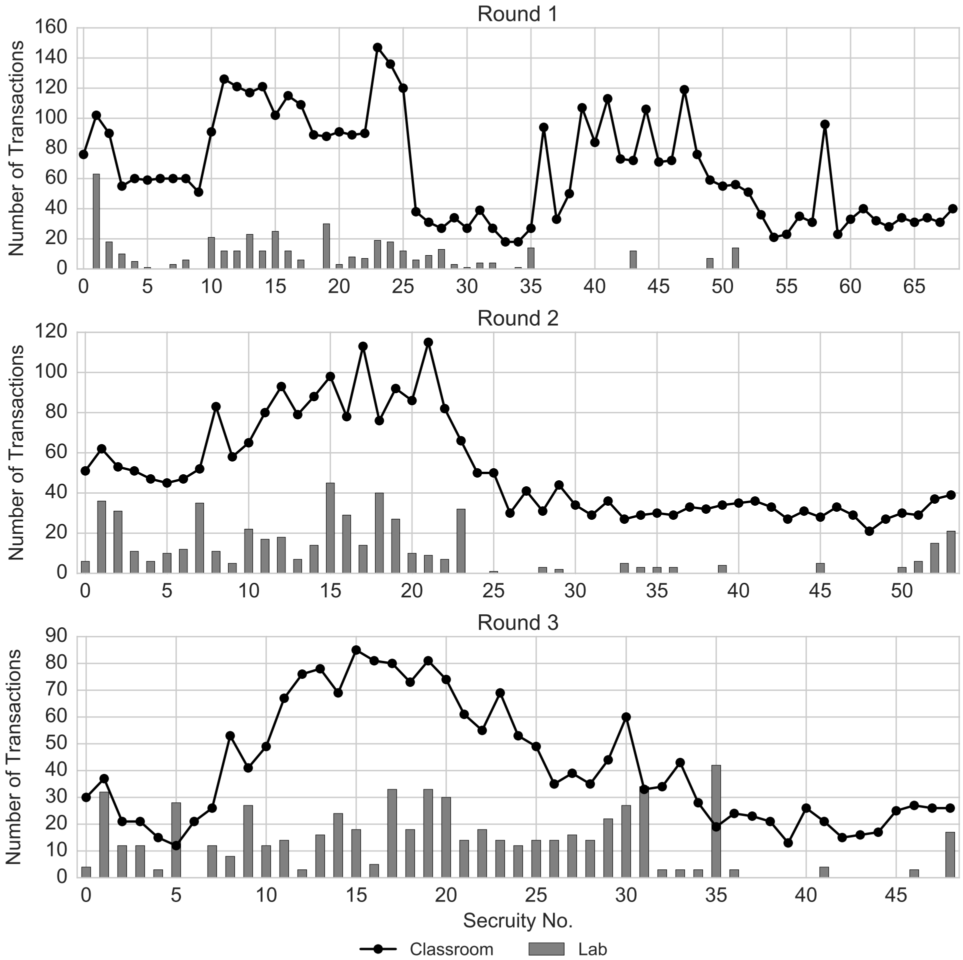

Figure 4.3: Trading volume of each security on the market in the laboratory (bars) and in the field experiment (line).

confirmed by the Jensen-Shannon Divergence (c.f. JSD, see Table 4.1) ranging between 0 and 1, such that values close to 0 indicate almost identical distributions and values close to 1 indicate substantially different distributions.

Table 4.1: Jansen-Shannon Divergence of end of each minute in the laboratory study. For minute 1, the values correspond to the divergence between the price distribution and a uniform distribution.

	Minute									
Round	1	2	3	4	5	6	7	8	9	10
1	0.89	0.61	0.28	0.23	0.08	0.11	0.12	0.09	0.07	0.08
2	0.86	0.60	0.13	0.04	0.15	0.09	0.10	0.06	0.11	0.03
3	0.73	0.34	0.29	0.08	0.02	0.06	0.03	0.04	0.07	0.04

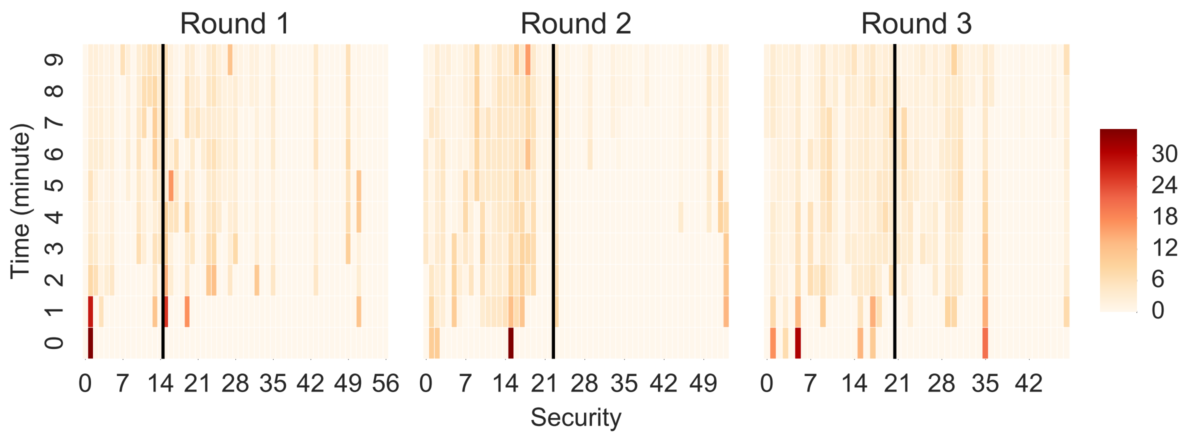

Figure 4.4: The evolution of security prices over time for the three trading rounds. The price distribution emerges within the first 3 minutes.

This price emergence resulted from the aggregated initial beliefs of the market participants. According to Figure 4.5, the average pre- and post-trading beliefs were very strongly aligned with the price distribution in each week. The post-trading distribution was more strongly correlated with the price distribution than the pre-trading belief (Pearson correlations of the price with the post-trading belief: $r = .62, .71, .51$; Pearson correlations of the price with the pre-trading belief: $r = .56, .71, .38$, $p < .001$ for all correlations), while the beliefs were more correlated with each other than with the price ($r = .79, .83, .83$, $p < .001$ for all correlations). This replicates the corresponding finding from the field experiment. As outlined in Equation 4.1, for each round, we implemented a regression analysis demonstrating that the difference between the post-trading belief and the market can be predicted by the difference between the pre-trading belief and the market:[12]

$$Belief_{post-trading} - Price = \beta_0 + \beta_1 \times (Belief_{pre-trading} - Price) \tag{4.1}$$

In all three rounds β_1 (β_1 equaled 0.70, 0.84, 0.80 in Rounds 1-3) was significant at $p < 0.001$ and the percentage of explained variance was medium and high (R^2: 0.57, 0.38, 0.64).

Further, the peaks of the distribution for each week were always the lowest for the price distribution, second highest for the pre-trading belief and the highest for the post-trading distribution. This is in contrast to the field experiment, where the peak of the price distribution was always higher than the peaks of the belief distributions. This means that, in the laboratory, the beliefs of the market players were directed towards particular

[12]The dependent and independent variables in this regression are expressed as differences between beliefs and the marked distributions to avoid the multicollinearity problem.

securities more than the market (showing a coordinated opinion of the players), while it is the opposite in the field experiment.

Figure 4.6 shows that the beliefs of individual participants were convergent on which of the securities would pay out a dividend. The securities that were assigned with more weight are close to the realised securities. Overall, the beliefs in Round 1 were more dispersed than beliefs in Rounds 2-3 and most of the belief were assigned close to the executed securities. This finding is consistent for the two experimental settings.

4.3.3 Mispricing and market rationality

To analyse the pricing rationality of the market, we calculated three market indices: index 1 – sum of security prices in the market, index 2 – sum of highest bid offers and index 3 – sum of lowest ask offers. Due to the fact that the dividend pays 100 units of currency, index 1 should not exceed the value of 100 and for the market to be rationally priced, index 1 should equal to 100. If index 2 exceeds 100, or index 3 is lower than 100, there would be a straightforward arbitrage opportunity against positive and negative bubble on the market respectively.

Figure 4.7 shows the evolution of the three indices across 10 minutes of each trading round. First, the market was overpriced in all three experimental rounds, which is confirmed by the Relative Deviation (c.f. RD;[201]) presented in Table 4.2. However, this overpricing was not as pronounced as in the field experiment. In Round 1, index 1 exceeded 100 only after 4 minutes of trading (i.e. 40% of the trading time) and stayed at the level of about 150. In Rounds 2 and 3, index 1 exceeded 100 after about 2 minutes. In the laboratory setting, we did not observe decrease of this mispricing across rounds.

Table 4.2: A market mis-pricing measure - Relative Deviation for the three market indices: the sum of prices (index 1), the sum of highest bid prices (index 2) and the sum of lowest ask prices (index 3) in each trading period (week) in Experiments 1 and 2.

Index	Round 1	Round 2	Round 3
1	0.45	0.19	0.30
2	-0.07	-0.51	-0.25
3	1.01	1.50	1.39

Second, the over-pricing was particularly well characterised by the time intervals during which index 2 becomes larger than 100: in Round 1 briefly at the end of the sixth minute and during the eight and ninth minutes, in Round 2 during the second, third and fourth minutes, and in Round 3 from the second to the fifth minute. The fact that the best bid was larger than 100 means that any transaction had to be concluded at a price that would result in an aggregate price significantly above 100, in clear violation of the rationality and fair value argument.

As the average bid prices were smaller than 100 almost all the times, there were no obvious arbitrage opportunities in the laboratory setting *on average*. However, given that the prices were at times very large for some securities, in the self-reported questionnaire, seven participants reported that they applied an arbitrage strategy, selling the securities with high prices. In the field experiment, we observed one strong arbitrage opportunity in Round 1 and one in Round 4, in the sense that the best bid price became transiently larger than the best ask price. The development of all three indices is very similar in all three trading rounds.

In the field experiment, the overpricing was the highest during the first half of the day when the market opened (i.e. 5% of the trading time) and it decreased towards the end of each trading round. Also, the mispricing diminished across rounds. We attribute these differences to the time constraint and late formation of the price distribution in the laboratory.

4.3.4 Trading performance and the Illusion of Control

In the final questionnaire that followed the trading task, 18 participants (50%) responded that they realised that the market index should equal 100, while 7 persons claimed to have applied an arbitrage strategy based on this normative fact. Three of these persons were in the top quartile, two were in the second best quartile and only the remaining two did not receive any bonus, but were in the third quartile. This supports the observation that there were some arbitrage opportunities only based on recognising that the market (and a number of securities) were overpriced.

In the post-trading questionnaire, one participant reported to have had a few years of experience in trading, two people reported having 3-6 months experience (an equivalent of an internship) with trading, while others had no experience. The person with a few years of experience was fifth on the final rank.

In contrast to the field setting, we found no correlation between the number of submitted orders and participants' earnings. There was only one person (an outlier), who not only submitted substantially more orders ($N_{orders} = 156$) than other participants (Range: 16-119, $M = 68$), but also, this person had a substantially higher total earnings ($Earnings = 3529$) than the rest of the participants (Range: 2471-1041, $M = 1800$). Therefore, this participant had rank 1. This suggests that the laboratory setup promotes more of a gambling atmosphere with insufficient time to ponder and evaluate the options as well as keep or recover a cool trading mind.

Further, for each participant, we calculated the primary illusion of control [214], which relates to the belief that one has a control over the outcome of the stochastic process, the secondary illusion of control, which defines that a person aligns themselves with having extraordinary skills such as "feeling lucky moments". For each participant, we computed the average of responses from the questions corresponding to each subscale (primary and secondary), where each question was measured on the scale 1-10. The total score of the illusion of control is the average from all questions in the survey. Overall, all participants had a low primary ($M = 3$, Range: $.5 - 5.67$) and secondary ($M = 1.44$, Range: $0 - 5.33$) illusion of control, as well as the total score ($M = 2.58$, Range: $.8 - 4.8$) of the illusion of control. The last question of the illusion of control questionnaire asks on a scale 1-10 whether "It was all chance". Six people replied 0 on this question meaning that they believed that their performance was completely attributed to their actions. Only three people responded 10 (maximum value) indicating that they believed that they had no influence on their performance. The distribution of responses was slightly positively skewed, with the median of 4 and mean equal 4.06.

We found a moderate correlation between the final earnings at the end of the three trading rounds and the total illusion of control ($r = .44, p < .01$). This correlation was driven by the strong correlation between the secondary illusion of control and the final earnings ($r = .57, p < .001$), while there was no correlation of the final earnings and the primary illusion of control. Given the fact that the survey of the illusion of control was preceded by the trading task and that the participants generally had no trading experience, we interpret that those participants, who received better scores in the trading task, attributed their success to their skills such as "feeling the market". This relation was also reflected in the negative correlation between the final rank and the total illusion of control ($r = -.57, p < .001$), the negative correlation between the final rank and the secondary illusion of control ($r = -.71, p < .001$) and no correlation between the final rank and the primary illusion of control. There was no correlation between the trading volume or number of orders and any measure of the illusion of control, which means that the illusion of performing well was attributed only to the final results of the trading.

Ref. [214] found that higher illusion of control was correlated with people's prior beliefs about the outcome of a gambling task that their participants performed. In our experiment, we find that participants in the laboratory condensed their beliefs to fewer securities than the participants in the field experiment. The distribution of the prior beliefs had a larger peak and thinner tails than in the field. We surmise that the laboratory participants formed more extreme beliefs while being less confident about these beliefs and their actions.

4.4 Discussion

4.4.1 Common findings in the laboratory and field experiments

In this study, we adapted a complex field experiment involving an experimental asset market to laboratory conditions. We replicated the procedure of the field experiment in the highly controlled experimental setting for the purpose of testing the relation between the laboratory results and complex trading environment.

In the laboratory experiment, we replicated a number of key effects found in the field experiment. First, we observe that the initial price emerges early during the trading round and the price distribution stays relatively constant until the end of the trading time. Second, this price emergence is a result of the initial belief of the market participants. The post-trading belief was more correlated with the price distribution of the securities than the pre-trading beliefs, but the two beliefs correlated more strongly with each other than with the price distribution. Third, we observe significant mispricing, despite the fact that half of the participants realised that the market was overpriced.

The fact that we replicated these behavioural effects in a highly controlled setting with time constraints highlights the robustness of the findings. This speaks in favour of the reliability of these results independently of the environment, in which the experiment was conducted. The most robust effect found across many studies is the market mispricing. It is worth noting that, in our laboratory study, despite the fact that 20-40% of the securities were not priced, the market was overpriced over half of the trading time. Surprisingly, the mispricing in the laboratory occurred at a relatively later point during trading (in percentage of total trading time) than in the field experiment, which at prima facie seems to contradict the Active Market Hypothesis [187] but may also be associated with the incompressible time for participants to make up their mind within the few minutes available in the laboratory.

Also, we showed that, in the laboratory, the market forms even when the securities are abstract and the participants have minimum knowledge about the traded assets. Our participants formed an opinion (i.e. belief) about a stochastic process, with minimum prior knowledge about it. This questions the validity of the experimental findings, because the laboratory participants formed a more extreme opinion about securities and they had less knowledge about the underlying securities than the field participants who indicated more uncertainty in their beliefs.

On the other hand, the robustness of the main effects between the laboratory and the field study demonstrates that it is possible to relax some of the controlled measures in the laboratory in favor of additional advantages of field experimentation. For example, in the field study, there was no limit in the number of participants, while in the laboratory, we were restricted by the capacity of the laboratory. Additionally, allowing participants to complete the task from any place that they find convenient offer the possibility to record their behaviour in their “natural” environment and to run the study for a much longer time (i.e. four weeks instead of two hours). Thanks to the larger number of participants, the market in the field study was more liquid, which demonstrates that increasing the complexity of the task may require increasing the number of participants in each particular round and extending the duration of one round.

Our results confirm the hypothesis that the participation in an economic experiment should be endowed with the compensation scheme that is relevant for the particular experimental setting. We obtained the same key effects when compensating students enrolled in a class with bonus grade points and endowing laboratory participants with competitive amount of money.

Here, we purposefully used the same experimental materials (i.e. the professor’s slides) as in the field experiment in order to directly compare the two settings. However, this design allows for several extensions. For example, in order to investigate how important is familiarity with a particular stock, one could conduct an experiment in which students trade abstract stocks such that the one paying out the dividend would be chosen according to a stochastic process. Another extension could test the predictive power of the market by asking students to predict a real life event such as outcomes of sports events. Imagine that each security corresponds to one athlete in the 400-meter run competition at the Olympics. Experiment participants could trade these securities before the run. Another variation would involve environment that is rich in information about the traded securities - students could trade securities that correspond to the percentage change from one day to another of an index from a market that is closed in the timezone of the experiment.[13] This variation would also allow for investigating the impact of the time allowed for trading on the trading activity and performance on the market.

[13]We implemented this variation in a setting where students traded derivatives of FTSE from home, within strictly designated 2-hour time. Unfortunately, we experienced technical problems during data collection, which made the data non-reliable.

4.4.2 Observed differences between the laboratory and field setting

Despite replicating the main findings from the field experiment, we observed a few differences between the field experiment and the laboratory experiment. First, we did not replicate the effect of the decrease of mispricing across trading rounds. We surmise that this was due to the short trading time in the laboratory setting. Also, there are substantial differences in the market dynamics between the laboratory and the field setting.

Second, the price distribution became stable at a later stage during the trading round in the laboratory compared to the field. It is important to note that the definition of “late” is a relative concept, measured as percentage of the total available trading time. In absolute terms, exceeding the rational price level after 2-4 minutes after the market opens is comparable to the time needed for bubble development in the experiments using the SSW design [40, 44].

Third, many securities remained without price, which is not the case in the field setting. Also, the price distribution is less “smooth” in the laboratory, which makes it difficult to judge the predictive power of the market. The differences in the price distribution are related to the short trading time and complexity of the task. Our results demonstrate that the time allowed for trading is a very important component that not only makes the market more liquid, but also gives the market players more opportunity to explore the complexity of the market.

Fourth, in the field experiment, we observed a characteristic daily and weekly cycles of trading activity. These fluctuations show that the market liquidity differs at different time points. For example, some orders were executed immediately when many traders were logged to the trading platform, while other orders had a longer waiting time or could be canceled by the issuer, at times when few traders were active. In that logic, there were times at which participants could “think twice” and times at which they had to react fast. In the laboratory, it was impossible for the participants to thoroughly think about their strategies and they had to react fast at all times. This was reflected by more diversified self-reported trading strategies and higher trading volume of the first security listed in the platform.

Further, while transferring the field experimental design to the laboratory, we experienced a few challenges. First, given the rather large number of traded securities, the market was not as liquid as in the field setting, despite the high trading activity and our use of the maximum capacity of the laboratory. This points to the limitations of the laboratory experiments – implementing a large number of securities requires a large number of participants as well as long trading rounds. Implementing an online experiment would provide a solution to this problem. However, the experimenters would not be able to control what the participants really do. We propose that this high degree of control of the experimental setting introduces artificiality. In real life, traders constantly face distractions, check e-mail, browse the Internet, and are continuously subjected to a flow of news through various channels. Forcing participants to focus on one task only does not resemble the real markets. In contrast, the field experiment captures well this condition.

Next, using a realistic, complex trading platform requires teaching the participants on how to use it. The trading platform that we used in both experiments is a multi-tab software, which mimics some functionality of professional trading software. In order to make it possible for our naïve participants to use it, we created a video with instructions that worked as a 7-minute crash course to the software. We cannot eliminate the possibility that some participants underperformed because they had to learn how to use the software "on the go". This is another demonstration that a realistic trading task may be too challenging for a short laboratory experiment. Participants need time to learn how to use the software and how to perform well in the task [215].

4.4.3 Motivation for the changes between the field and laboratory setting

In order to adapt the field study to the laboratory conditions, we had to make a few changes to the design. First, the main change was the number of participants reduced from over 100 to exactly 36. On the one hand, the laboratory setting allows for the control of an exact number of participants (In the field setting, the number of participants fluctuated across experimental rounds). On the other hand, the number of participants was strictly limited by the laboratory capacity, which in settings with low market liquidity caused by a large number of securities can pose an important problem. Indeed, our market had lower liquidity in the laboratory than in the field. In that sense, the field experiment has the advantage of measuring price emergence and development of complex markets with multiple securities. Also, in real life markets, the number of traders is not controlled. Despite the standard criticism of non-laboratory experiment in which the experimenter "cannot control what participants are really doing", the less controlled setting can shed more light on how people really behave.

Second, in order to make the participants learn to use the trading software with several tabs and to explain the relatively complex task for a short experiment, we had to present a manual on how to use the software in a form of a concise and comprehensible movie, while in the field setting we were able to provide a presentation of the software in the classroom. This is a general limitation of implementing realistic complex tasks in short laboratory experiments with participants that are not familiar with the task and software.

Third, due to time restrictions of 2 hours that was partially dictated by the Decision Science Laboratory, all information had to be presented in a very coherent way and we had to reduce the number of trading rounds from 4 to 3. While this change reduced the number of obtained data and statistical power, the main effects held.

Fourth, in the laboratory, we presented a movie summarising the professor's lecturing style while participants in the field setting could experience first-hand the professor lecturing. This change raises two types of criticism. The first arises from the description-experience gap [175], which states that people tend to under-sample the outcomes of events and make their decisions accordingly. In a similar fashion, it is likely that each participant experienced the professor's teaching style differently, which could have impacted their trading strategies. In the laboratory setting, all participants received the same information about the professor's teaching style. On the one hand, presenting the same information gives more control over the flow of the experiment. On the other hand, presenting information descriptively results in a standard criticism of artificiality of laboratory experiments in all behavioural sciences.

4.5 Conclusion

The goal of this study was to compare an experimental asset market in the field and laboratory experiments, while using the same experimental design in two settings. We did not aim to find new behavioural effects in the laboratory experiment. On the contrary, this experiment was an exercise whose goal was to test whether the key findings found in the field study would be replicated in the time-constrained more controlled laboratory setting. The laboratory results replicate the three main findings from the field experiment, which demonstrates their robustness. The most robust finding is mispricing of the market, which has been widely reported in experimental asset market experiments.

Despite the replication of the key results, we found the existence of substantial differences in the market dynamics in the two experimental settings. The key reason for these differences was the time pressure that limited the learning to trade and to use the software by the participants in the laboratory. In spite of this time limitation, and in the presence of an intrinsic uncertainty about the market fundamentals and very limited knowledge of the market process and of the securities, we observed a very high market activity and a rapid price formation dynamics in the laboratory conditions. This poses the question of what information can reliably be extracted from trading experiments in the laboratory, where this is the only task performed by the participants in very unrealistic conditions.

The confirmation that the key effects of the field experiments were reproduced by the laboratory version, together with the fact that the field conditions did not suffer from the many unrealistic constraints, while presenting other findings better in accordance with empirical observations in the real world, suggests that these new class of field experiments, as introduced by Ref. [107] can have a promising future. Nevertheless, the goal of this paper has been to raise researchers' awareness to the fact that standard laboratory experiments may not mimic the behaviour of real complex financial systems. Alternative setups can be developed with intermediate levels of control and complexity that may help close the gap between the maximally controlled laboratory conditions and the real financial markets.

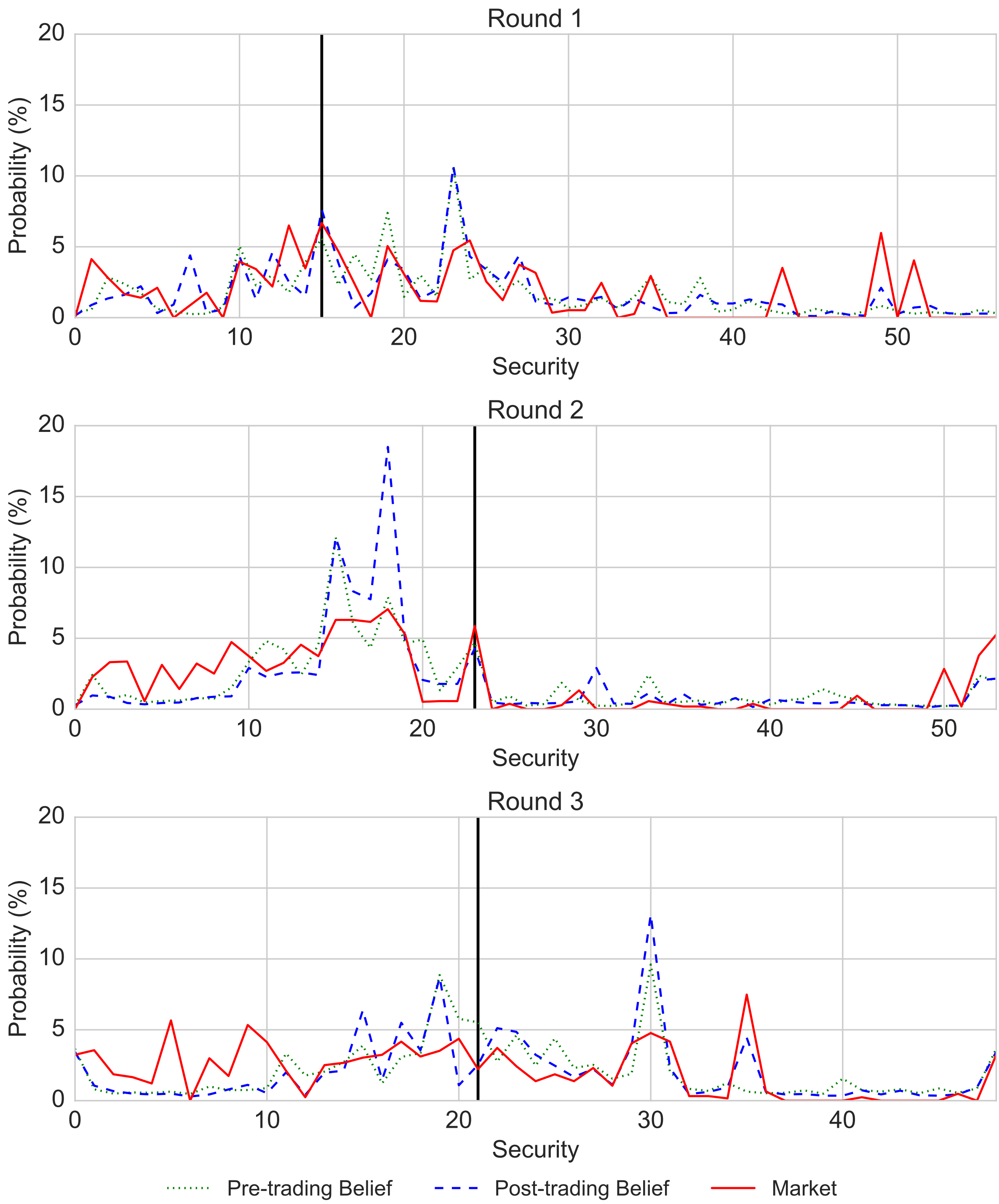

Figure 4.5: Distribution of securities' prices, pre-trading and post-trading beliefs across 10 1-minute epochs. For each epoch, we calculated the median price for each security. These median prices were then averaged across all 10 epochs. These prices were normalised so that their sum is 100. The pre- and post-trading beliefs were obtained by averaging the submitted beliefs across all 36 participants.

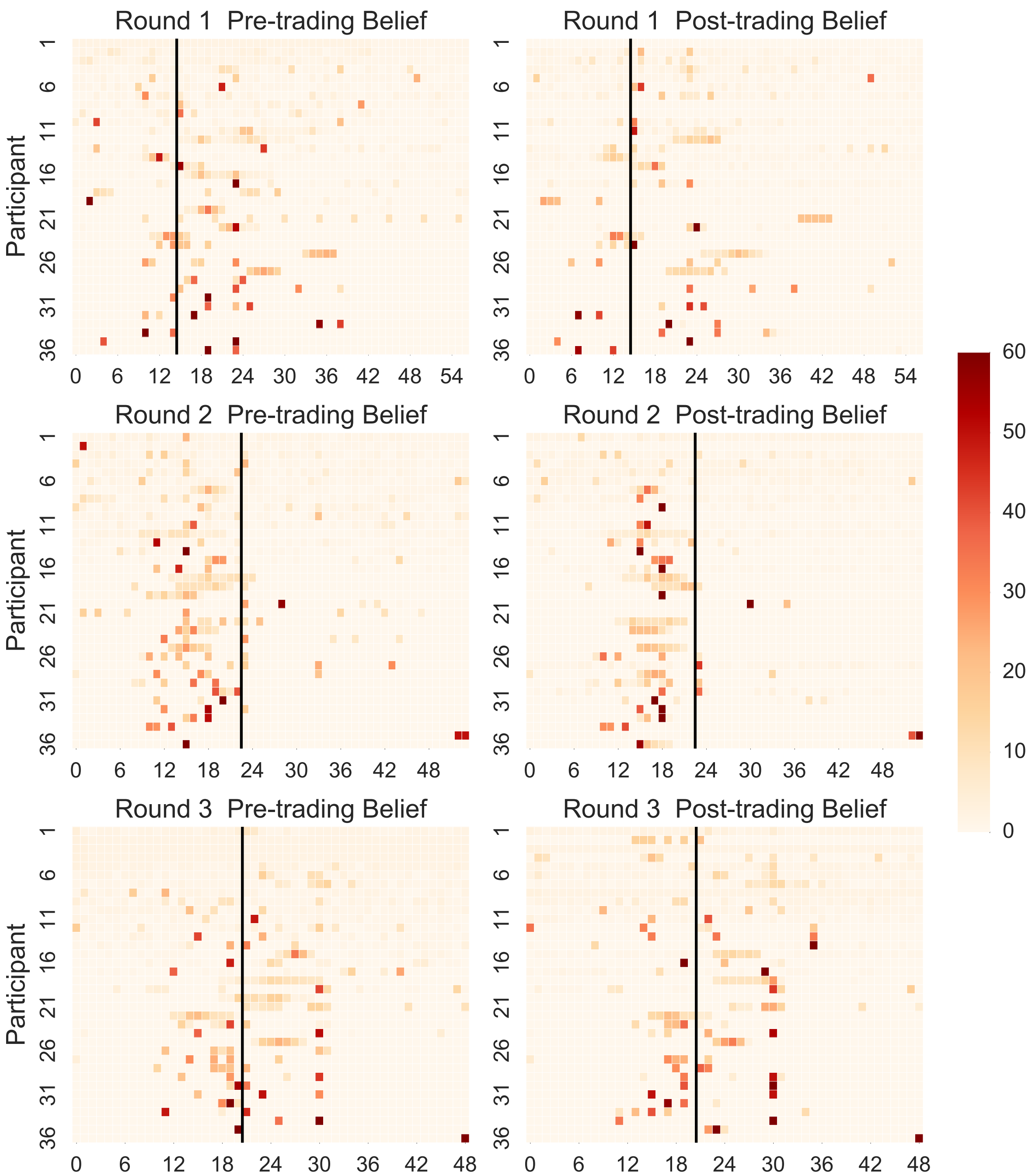

Figure 4.6: Heat maps of pre- (left) and post-trading (right) beliefs, such that each cell corresponds to the belief assigned by one participant to one security. The individual belief distributions are sorted in a decreasing fashion, according to the number of securities with non-zero weights.

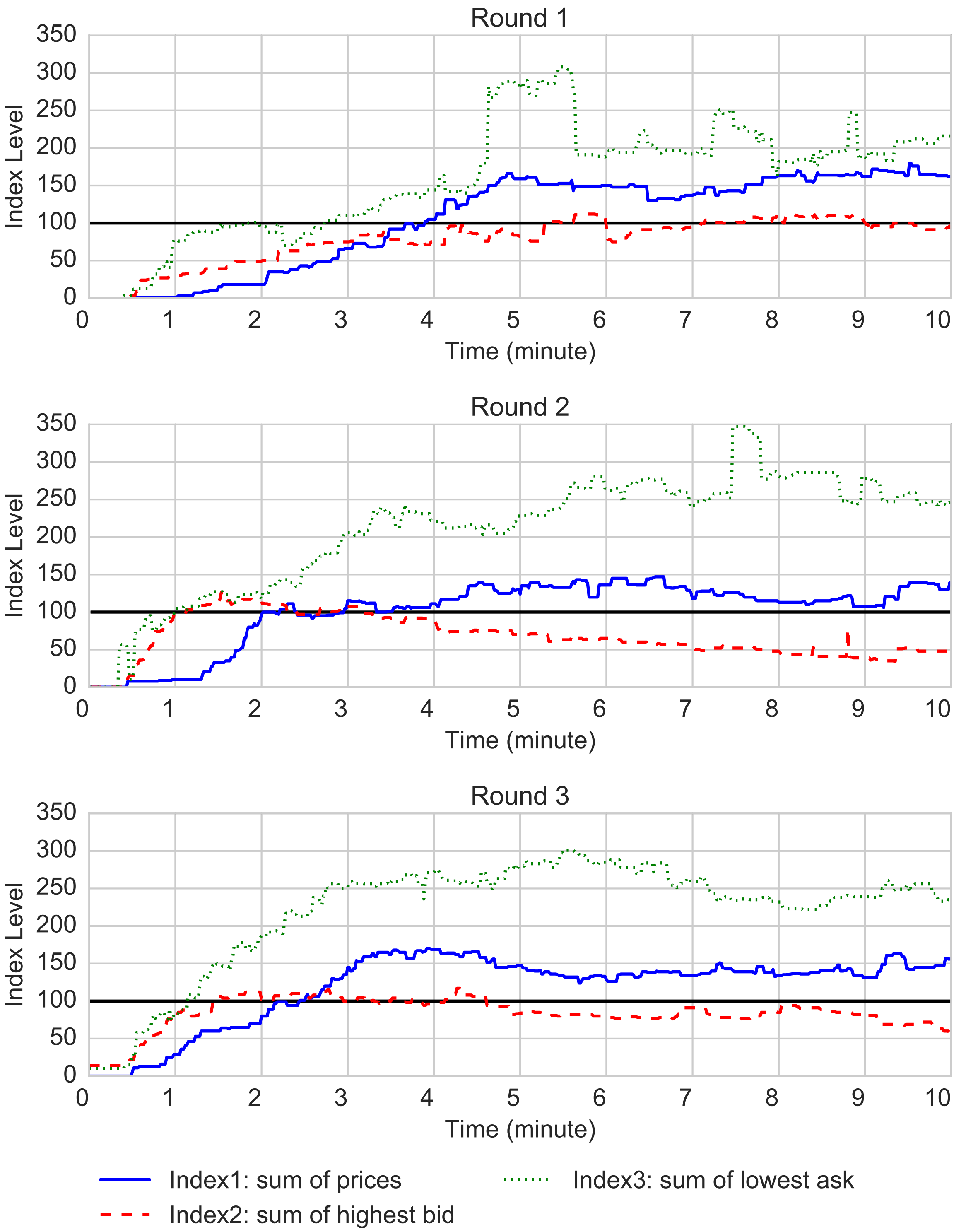

Figure 4.7: Evolution of three market indices corresponding to the sum of all securities' prices, sum of the highest bid prices, and sum of the lower ask prices. The sum of all security prices should be 100 but there are several pronounced deviations from this normative prediction.

Chapter 5

Impact of objective and subjective information on market prices

In this Chapter, we extend the experiment in Chapter 3, and investigate the impact of objective (i.e. reliable and quantifiable) and subjective (i.e. ambiguous) information on market prices. We hypothesize that providing participants with objective information about the "correct" price level of the market with the clear indication of how to use market inefficiencies, would reduce mispricing by moving the price level to the "rational" range. In line with [65], we expect that well-explained reliable information about the correct price should make the market more efficient. Also, we hypothesize that presenting participants with subjective and ambiguous information would make participants form more extreme opinions about prices and should therefore inflate prices of particular assets. However, it is an open question, if and to what degree one piece of this subjective information would "override" the objective information continuously provided as a quantitative measure of market mispricing. To our knowledge, the latter has not been investigated before.

We present results from two treatments employing Experimental Asset Markets, each lasting two weeks. In this experiment, participants trade assets whose outcomes are uncertain, all participants have the same information and the only source of information in the market is the information distributed by the experimenters. In Treatment 1, we provide traders with an objective quantitative measures of what the correct price of the market would be, accompanied with instructions on how to use market mispricing to generate profit. In Treatment 2, in the market that already includes these objective metrics, we distribute subjective and uncertain information (c.f. an opinion) to half of the participants while informing all participants that a subjective expert opinion will be distributed to randomly selected 50% of the traders.

It is worth to note that the extensions of the experiment and the results are very preliminary here, and it is not enough to draw sound conclusion at this stage. Due to the time limits and other constraints, we did not run further experiments in this setup. However, we believe that these two treatments here can shed a light on studying the influence of subjective and objective information in more realistic experimental markets.

5.1 Method

5.1.1 Experimental Design

Each experiment employing the Chapter 3 Experimental Asset Market design consisted of two experimental rounds, each lasting two weeks. In Fall 2015, students enrolled in a Financial Market Risks Class at the Swiss Federal Institute of Technology in Zurich could voluntarily participate in a trading experiment. Their good performance was rewarded with bonus credit points that were added to their exam grade such that the maximum grade from the course could be obtained without participation in the experiment and the total grade from the exam and trading could not exceed the maximum grade in the Swiss grading system. The best 25% of the participants in the trading experiment were awarded with 0.5 bonus credit point, the second best quartile received 0.25 bonus credit point, while the worst half of the participants did not receive any bonus. The rank was cumulative throughout the whole semester that was equivalent to eight trading rounds, but the participants could monitor their relative performance for a given experimental round. The first four trading rounds constituted Experiment 2 are reported in Chapter 3, while the latter four trading rounds constitute two experiments reported here. All of the market setups and experiment procedure is the same as the one in Chapter 3.

5.1.2 Participants

Out of 209 students enrolled in the course, 97, 80, 79 and 73 participated in the four consecutive experimental rounds. All participants had a science and engineering or management and technology background. 70, 63, 56 and 47 participants submitted their post-trading belief to include their portfolio in the ranking. 6, 8, 9 and 10 participants were classified as very active traders (i.e. traders whose number of submitted orders exceeded 1.5 times interquartile range). Before starting the experiment reported here, the participants had five weeks of experience with the task (one practice and four experimental rounds) and they had eight weeks of experience of the professor's lecturing style. About 80% of the participants were male.

5.1.3 Treatment 1

Due to the fact that one security paid out the dividend worth 100 units of experimental currency, the rational price of the market was when the market index (i.e. the sum of prices of all securities on the market) equaled 100. If the index exceeded 100, the market was overpriced (i.e. positive bubble). If the index was below 100, the market was underpriced (i.e. negative bubble). Chapter 3 found substantial overpricing in all four trading rounds when the arbitrage was possible. Also, they found that 66% of the participants realized the market index should equal 100, but 44% of the participants did not know how to arbitrage this opportunity.

In this experiment, we implemented three indices to the trading platform: *index 1* - sum of prices of all securities available on the market, which indicates the rationality of the market prices; *index 2* - sum of the highest bid prices, which indicates an arbitrage opportunity when the index exceeds 100. The arbitrage opportunity can be exploited by selling one unit of each security and thus obtain a certain profit at maturity. Such strategy would tend to push down the overall price level; *index 3* - the sum of the lowest ask prices, which indicates an arbitrage opportunity when its value is below 100. One could make profit by buying one unit of each security. During the lecture preceding the experiment, a teaching assistant of the professor provided an explanation of the indices and a tutorial on how to implement an arbitrage strategy. All three indices could be continuously monitored by all participants during a six-day trading round. The participants were explicitly instructed to use the indices to monitor for the possible arbitrage strategies.

5.1.4 Treatment 2

We informed all participants that on Thursday at 6pm, by e-mail, we would distribute the information about which slide could possibly be the end slide of the upcoming lecture to 50% of the randomly selected students enrolled in the class. The remaining 50% of the students received this additional information in the second round of the experiment. The estimation was made purely based on teaching assistants' knowledge without any insider information, and this was clearly stated to the students. We sent the following e-mail:

You are receiving this email because you are randomly picked in the group who shall receive the additional information for this week's trading. It is up to you how to use this *information and whether it helps your trading. 50% of all students enrolled in the class received this information.*

Ms. Sandra Andraszewicz and me agreed on an estimation saying that the most possible security for the next Monday's lecture is L7.P031-033. This is our prediction based on our knowledge without any insider information, which may or may not be correct.

Actually, both of the recommendations the teaching assistants made in the two weeks turned out to be wrong in the end. The choice of the time at which the information was distributes was selected such that, in the previous six trading rounds, roughly half of the trades occurred before the time when the information was distributed and the other half would occur after this time. The reason for this time selection was to have a balanced trading activity before and after the distribution of the additional information.

The three market indices were visible during Treatment 2. In order to fully understand the effect of Treatments 1 and 2, some of the results will be compared to the results from before the treatments, which are presented in detail in Chapter 3.

5.2 Results

According to Table 5.1, the activity of the market was stable across the four weeks of the experiment. The Relative Absolute Deviation, Relative Deviation and the percentage of the time when the index 1 (sum of prices of all securities) was under-priced [201] indicate that the mispricing decreased during the two treatments, compared to the time before the treatments were applied. We cannot rule out that the mispricing decreased as a result of learning experience of the traders, which is a common finding in experimental asset markets [see 40, for a review].

Further, we analyzed the percentage of the trading time during which there was an arbitrage opportunity, based on the three indices. The arbitrage was possible for 65%, 23%, 0% and 51% of the time in weeks 1-4, it was the case for only 0%, 0%, 21%, 0% in weeks 5-8, which shows again the effectiveness of Treatment 1.

Next, we compared the end of day price distributions for each day from Tuesday to Sunday. We calculate Jensen-Shannon Divergence (JSD) to measure the similarity between the end of day price distributions for all of two consecutive days within the week, and plot the result in Figure 5.1. A low JSD value means the two ends of day price distributions are similar, i.e., the price distribution does not change much after one day's trading. For week 1–6, the price distribution stabilize soon on Tuesday with low JSD afterwards (there is

Table 5.1: Descriptive measures of the market in each experimental week: Relative Deviation, number of submitted orders and average number of orders submitted per trader.

Treatment	Week	RD	Number of Orders	Average Orders per Trader
0	1	50%	4443	34.71
	2	21%	2700	26.47
	3	15%	2006	19.67
	4	14%	1493	15.39
1	5	7%	1163	11.99
	6	-8%	1173	14.66
2	7	-3%	928	11.75
	8	3%	906	12.41

one exception in week 6). However, the price distribution only stabilize after Thursday for week 7 and 8.

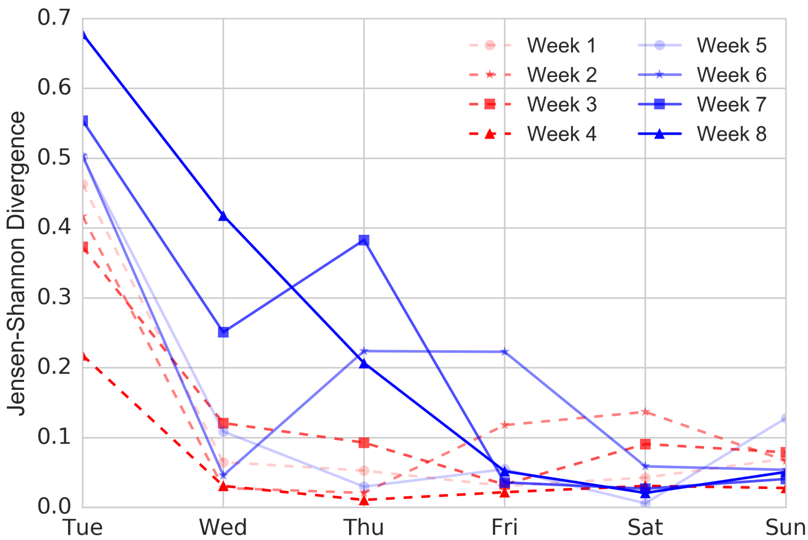

Figure 5.1: Jansen-Shannon Divergence of end-of-day market price distributions of 8 weeks. Each data point is the JSD between today's and yesterday's end-of-day price distribution. The first 4 weeks (Treatment 0) are plotted in red dashed lines, with earlier weeks marked in lighter red color. The last 4 weeks (Treatment 1 and 2) are plotted in blue solid lines, with earlier weeks marked in lighter blue color.

We investigate the impact of the "news", which is the subjective stock recommendation made by the teaching assistants, on the price of the securities in Figure 5.2. We plot the price evolution of the stock recommended by the teaching assistants, and the sum of prices of the 6 securities closest to the recommended stock. These 6 securities are corresponding to the slide pages around the most possible end-slide predicted by the teaching assistants. In the first week of Treatment 2 (week 7), the price of the recommended stock shot up immediately after the recommendation sent out to the randomly selected students. There were 3 and 29 trades on the recommended stock respectively before and after the stock recommendation announcement. The average transaction price of these trades were 8.9 and 22.1, with standard deviation 0.8 and 4.2 respectively. As for the 6 stocks closest to the recommended stock, their prices increased together as well, showing that the participants were buying a group of stocks closed to the recommended stock to diversify. The recommended stock in the first week turned out to be wrong, but only one stock away from the realized stock.

In the second week of Treatment 2 (week 8), the price of the recommended stock did not shoot up intensively as the week before: there are 11 and 13 trades on the recommended stock respectively before and after the announcement, and the average price of these trades are 9.9 and 15.8, with standard deviation 3.2 and 5.2 respectively. Nevertheless, the participants seemed to buy potential stocks before the announcement, which pushed up the prices of stocks closest to the recommended stock, as we can see from the right panel in Figure 5.2. This might be because the participants learned to act in advance, and diversify their bets on the group of stocks around the stocks which might be included in the recommendation information, given that the previous recommendation was closed to the realized one.

To better understand how the participants received and used the stock recommendation information sent by us, as the experiment in Chapter 3, we sent out a questionnaire to all of the participants after the experiment. The question was the following:

Were you informed directly (by e-mail) or indirectly (by your friends, or any other indirect case) about the "expert" information? If Yes, did you do any trading on that?

A. I was informed directly and I traded on it

B. I was informed indirectly and I traded on it.

C. I was informed directly but I did not trade on it

D. I was informed indirectly but I did not trade on it

E. I was not informed at all.

56 and 47 participants voluntarily responded to this post-trading questionnaire in weeks 7 and 8 respectively. 25% and 15% of the participants directly received and used the

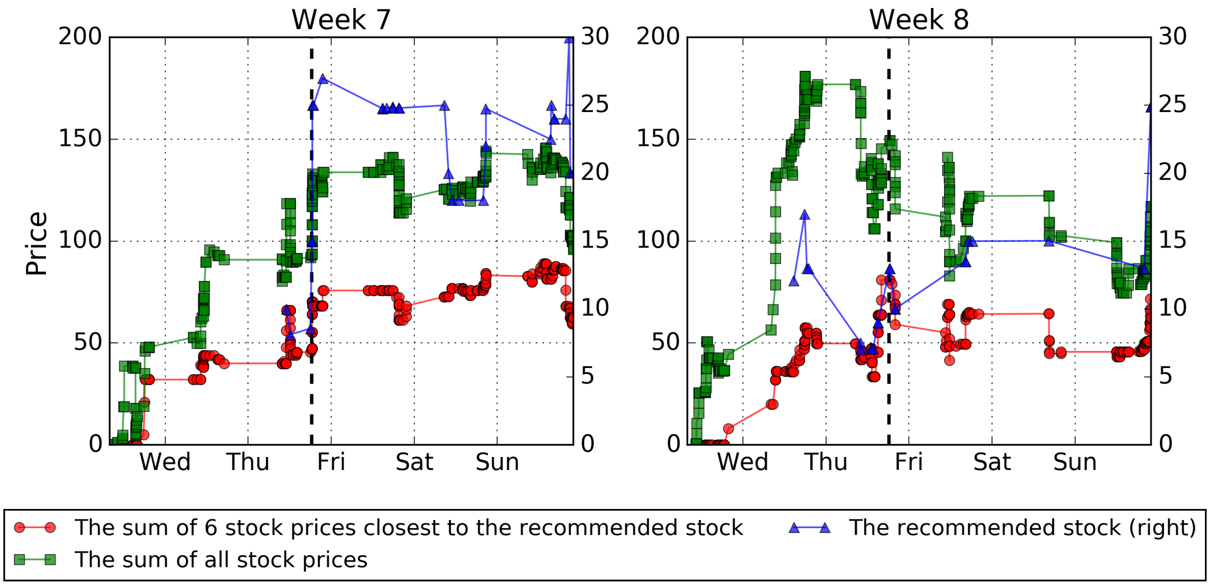

Figure 5.2: The price evolution of the stock recommended by teaching assistants to the students (right axis, in blue triangles), the sum prices of 5 stocks near the estimated stock (red circles), and the sum prices of all stocks (green squares). Each data point indicates a trade. The vertical black dashed line indicates the time when the information of estimated stock was sent.

information in week 7 and 8 respectively, while 29% and 28% of the respondents were informed directly but did not use the information for trading. We can see that 46% of the students, who received the information, acted on the recommended stock in the first week. However, this fraction decreased to 35% in the second week, which shows again that they might have less confidences in the recommended stock, due to the wrong recommendation in the previous week. There were only 1 and 3 students received the information indirectly (from their friends or other ways), showing that there was little incentive for students to share this subjective recommendation to their friends, even all of the them enrolled in the same class. Also, we did not find any difference in earnings between the participants that received and did not receive the stock recommendation information.

5.3 Discussion

In this trading experiment lasting four full weeks, we showed that including the objective information of the market mispricing level available to all traders reduces arbitrage opportunities, and make the market more efficient. However, it is not clear whether this is due to the learning effect of the participants, which shall be tested in a controlled experiment. Due to our time limits, we keep this question for future studies. Also, we showed that subjective information distributed to only half of the participants has a significant impact on the prices of the securities linked to the news, but not on the market in general. Interestingly, traders did not use the information contained in the news for trading and used the prices of these securities instead.

Our results resonate with a few of the previous findings. Possessing more subjective information than other market players does not result in higher returns[70]. Similarly as in [65, 69], objective quantitative information, accompanied with clear explanation on how to interpret it, reduced over-pricing and kept the price level close to the rational level. Also, providing ambiguous information (i.e. the recommendation subjectively made by TAs) resulted in more extreme price levels [72]. However, this was only the case for the securities to which this ambiguous information was linked, while the market in general was not impacted.

Our findings provide a few new contributions to the current knowledge. In a non-laboratory experimental environment with more realistic conditions (multiple securities and substantial long trading time), we made the first attempt to demonstrate the impact of different type of information on the market prices. Our experiment demonstrates how the impact of objective and reliable information interacts with ambiguous information. The key insight from the current study is that quantitative measures of mispricing have a moderating effect on the market prices, even when prices of individual securities are affected by ambiguous information that drives the prices of these securities to extreme values.

Also, our study shows that, it is not the content of the ambiguous message that plays a role, but the expectation that such message will arrive and affect the prices for only particular assets. In the second week of Treatment 2, we observed that, participants bought the stocks which are likely to be included in the news before the announcement, and sold these stocks when their prices increased afterwards. In contrast, in the first week of Treatment 2, participants only bought the related stocks after the news was distributed. This indicates learning to use the news. Also, the impact of the news could be high due to the fact that, in the fist week of Treatment 2, the recommended stock was only one stock away from the realised stock. However, the recommendation was less accurate in the second week of Treatment 2. This proves that the experimental design was double-blind and occurred under true uncertainty. Participants were not only predicting which security will be executed, but also which security will be recommended by the "experts" to be executed.

Despite its unconventional design and procedure, this experiment had some limitations. First of all, in this within-subject design, we were unable to disentangle the impact of the three indices and the ambiguous information on the price levels. We learned that a price bubble can be developed for some securities, while keeping the overall market price level within reasonable limits. However, given the fact that learning would likely to play an important role in this complicated task, we cannot be sure whether the three indices, which indicate the level of bubble, or participants' experience had the main bubble moderating influence. A few more controlled experiments are opened for future work, such as introducing the indices from the beginning, excluding the indices and including the news, etc.

Chapter 6

Classification of cryptocurrency coins and tokens by the dynamics of their market capitalisations

In this Chapter, we empirically verify that the market capitalisations of coins and tokens in the cryptocurrency universe follow power law distributions with significantly different values, with the tail exponent falling between 0.5 and 0.7 for coins, and between 1.0 and 1.3 for tokens. With a simple birth-proportional growth-death model previously introduced to describe firms, cities, webpages, etc., we validate the proportional growth (Gibrat's law) of the coins and tokens, and find remarkable agreement between the theoretical and empirical tail exponent of the market cap distributions for coins and tokens respectively. Our results clearly characterizes coins as being "entrenched incumbents" and tokens as an "explosive immature ecosystem", largely due to massive and exuberant ICO activity in the token space. The theory predicts that the exponent for tokens should converge to Zipf's law in the future, reflecting a more reasonable rate of new entrants associated with genuine technological innovations.

For guidance, we look to fundamental work on the nature of growth of firms and other entities. In particular, Zipf's Law has been identified as an ubiquitous empirical regularity for firm sizes [216], city sizes [217], connections between Web pages [218], connections between open source software packages [219], etc. – manifesting as a power law distribution of sizes with a unit parameter, such that $Pr\{Size > x\} \propto x^{-1}$ for sufficiently large size level, x. Since Simon's pioneering work [220], the primary generating mechanism of Zipf's law is understood to be proportional growth ("Gibrat's law"), also popularized as "preferential attachment" when recast in the context of networks [221]. Malevergne et al. [222]

extended the proportional growth framework to feature realistic birth and death, which again yields Zipf's law, but not necessarily with unit parameter – depending on a balance between the growth of new firms versus old ones. We employ this framework to study the growth process of cryptocurrencies, according to their market capitalization, from April 2013 to Feb 2018 [1]. We make an essential distinction that some cryptos are "coins" – which operate on their own independent network – and others "tokens" – which operate on top of a coin network as a platform. Notably, the coin market capitalization distribution is heavier tailed than Zipf's Law, and that of the token market somewhat lighter. The framework of Malevergne et al. [222] allows this to be explained, and identifies that the coins and tokens currently exist in distinct market regimes. This requires confirming Gibrat's law, estimating the birth and death parameters, and comparing the predicted exponent of the market capitalization distribution with its empirical counterpart, both for coins and tokens. Despite the clear limitations of the model, and a highly non-stationary market, we argue that this provides a reliable and meaningful result which may be refined with extended methods.

6.1 Evolution of Crypto-Currencies and Token Market Capitalization

After going through about a two year bear market, the cryptocurrency market started to grow again at the beginning of 2016 (Figure 6.1). The total market capitalization of all cryptocurrencies achieved a 250% return in 2016, and 3170% return in 2017. Although the first token, "Maid Safe Coin", appeared in April 2014, not until 2017 did the number of tokens explode, from less than 50, to more than 400 by the end of the year. On the other hand, the birth of coins has been relatively stable and the market therefore more mature. The evolution of the number of token deaths is more noisy, in some cases due to external events such as the bankruptcy of a large exchange in 2016 [225].

Although we are interested in market capitalization, to briefly isolate the relative size of different cryptocurrencies, we examine the distribution of market shares (the fraction of each coin or token to the total market capitalization of all coins or tokens, respectively) in Figure 6.2. For coins, the distribution is well described by a Pareto ("power law")

[1] This study uses the daily data of 2499 cryptocurrencies from April 28, 2013 until February 7, 2018; 1497 of which are still alive on Feb 7. The data is taken from Coin Market Cap [223], including daily closing price, market capitalization (the product of the price and the circulating supply), and the type of the cryptocurrency (coin or token) [224].

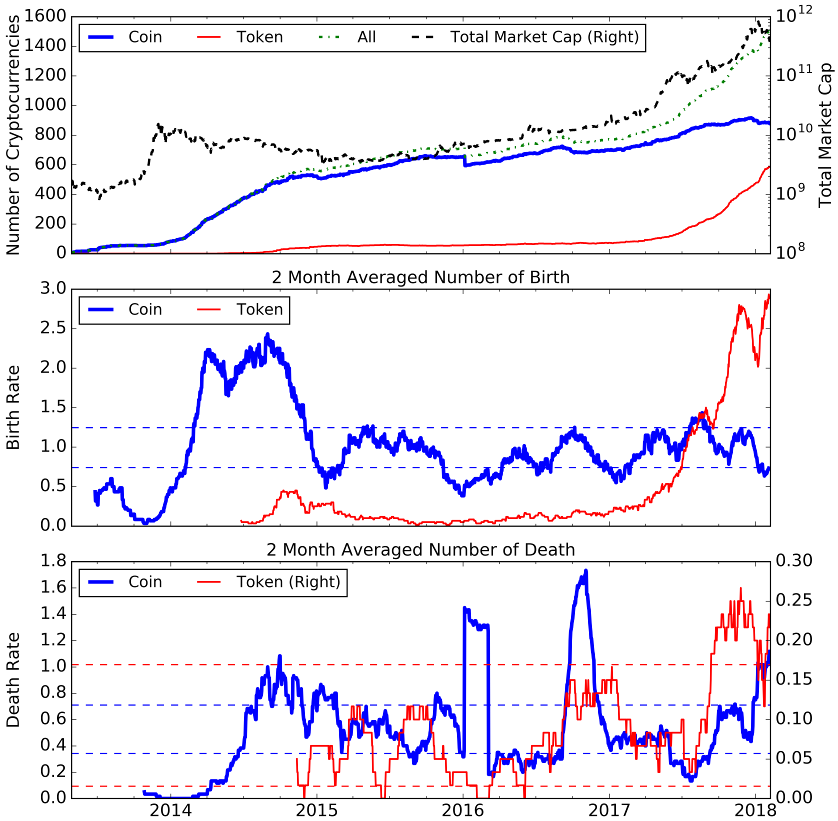

Figure 6.1: The upper panel depicts the evolution of the number of all cryptocurrencies (green dash dotted line), including both coins (blue thick line), and tokens (red thin line). The total market capitalization of all cryptocurrencies is plotted with a black dashed line against the right y-axis (log scale). The middle panel plots the birth rate for coins (blue thick line), and tokens (thin line) respectively. The lower panel is the corresponding death rate. Birth and death rates are the averaged number of births (deaths) in 2 month moving window. The horizontal blue (resp. red) dashed lines in the middle and lower panels are the 95% confidence intervals, assuming a Poisson process for coins (resp. tokens), whose mean is estimated over the whole period. The number of births is calculated as the number of new coins/tokens that appeared on CoinMarketCap[223] each day, and the number of deaths is calculated as the number coins/tokens removed from CoinMarketCap (i.e. marked as "inactive" by CoinMarketCap)[223]

distribution,

$$Pr\{X > x\} = (x/u)^{-\mu}, \; x > u > 0. \tag{6.1}$$

with the tail exponent, $\mu > 0$, fluctuating between 0.5 and 0.7 over time[2], and not being significantly worse than the more flexible 2-parameter Lognormal distribution for the top

[2]The tail exponent is estimated by Maximum Likelihood

275, out of more than 500 coins, in the most recent snapshot of February 2018 [3]. The market share distribution of tokens was closer to Lognormal instead of power-law at the earlier stage of 2017, but has been recently evolving towards a power-law in the past months, with parameter around 1.1, and not being significantly worse than the Lognormal for the top 50, out of more than 400 tokens, at a 0.05 test level. This confirms that the tail of coin and token market capitalization distributions are now well described by power laws with different exponents. Recall that the Pareto distribution with $\mu = 1$ is a border case called Zipf's law [227] where all moments of order larger than or equal to 1 are infinite. In the next section, we will consider a model to explain why coin and token distributions fall on different sides of this border case $\mu = 1$.

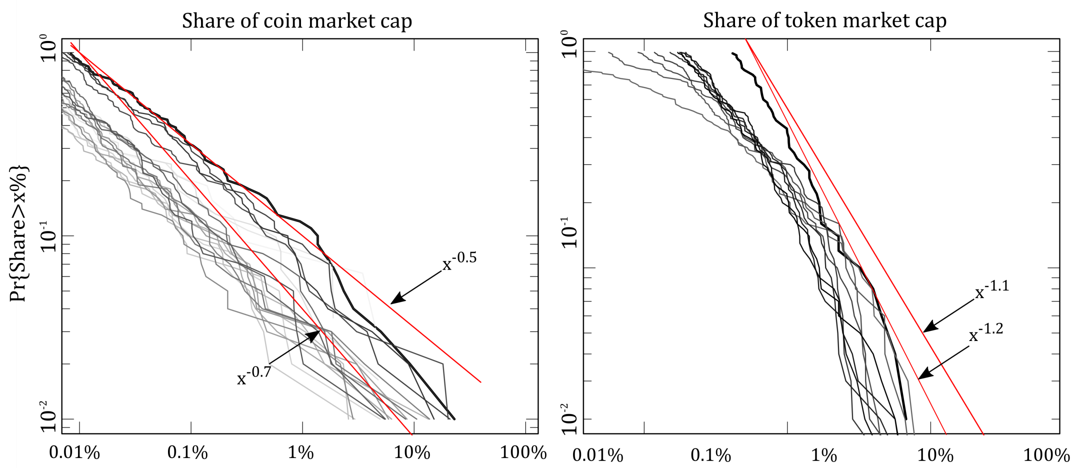

Figure 6.2: Left: Empirical complementary cumulative distribution function (CCDF) of top 100 coin market capitalizations at snapshots from 2014 (grey) to February 2018 (black) with the range of fitted Pareto tails indicated by the two red lines. Right: The same for tokens, but for the top 50 tokens, starting in early 2017.

Before this, we address the "Bitcoin maximalism" belief that Bitcoin would be the one and only winner, and all alternative coins ("altcoins") are destined to fail. This degenerate scenario would preclude our growth framework. However, time has largely settled the debate on this, which we briefly address in Figure 6.3, where Bitcoin dominance (its coin market share) has dropped from above 80 percent to at times well below 50 percent. The follow-up question is then, if Bitcoin is, or has been, an outlier. Bitcoin dominance alone may be misleading since the market capitalization distribution has changed over time. However, the market size of Bitcoin can be compared to the other top 100 market capitalization coins via a transformation [228]. As shown in Figure 6.3, in 2016, Bitcoin

[3]Lognormal not superior at p=0.05 level, using uniformly most powerful unbiased test [226].

was at its relative largest, at times exceeding the 0.9 quantile of the null distribution. Since then, it has descended to its relative smallest size, near the 0.1 null quantile. This indicates a change of fortune for Bitcoin, and effectively rejects it as an outlier[4].

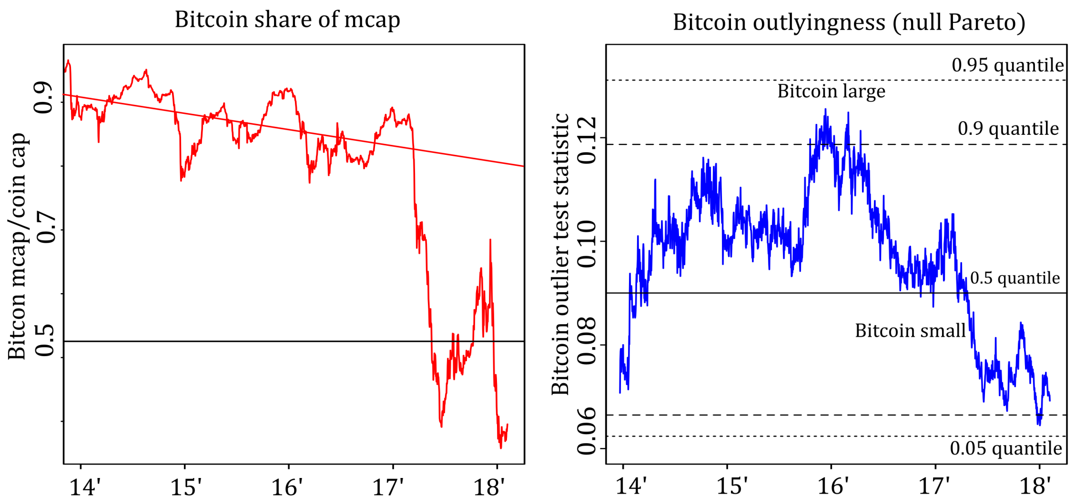

Figure 6.3: Left: Bitcoin dominance (share of total coin market capitalization) over time. Right: Test statistic and null (under Pareto) quantiles for testing Bitcoin as an outlier. In particular, the test statistic is $E_1/(E_2 + \cdots + E_{100})$ where $E_i = log(X_i/X_{100}), i = 1, \ldots, 100$ is a transformation of the top 100 market capitalizations $X_1 > \cdots > X_{100}$ that transforms X with a Pareto distribution to E with an exponential distribution [228]. With this transformation, the test statistic is independent of the parameter.

6.2 Proportional Growth with Stochastic Birth and Death

6.2.1 Definition of the Model and Main Properties

Proportional growth is a general and ubiquitous mechanism, as discussed in the introduction, and is quite natural for cryptocurrencies given the pervasiveness of proportional growth in complex networks. Within the cryptocurrency community, "network effects" have often been attested as a reason for the sustained dominance of Bitcoin. Further,

[4]Note that here Bitcoin forks are treated as separate independent cryptocurrencies, however including them all together within the Bitcoin value provides similar results

allowing for birth and death, we employ the framework of Malevergne et al. [222], which is based on the following assumptions:

1. **Gibrat's rule of proportional growth holds**. This implies that, in the continuous time limit, the market capitalization $MC_i(t)$ of the i_{th} cryptocurrency at time t, conditional on its initial market cap, is the solution to the stochastic differential equation (i.e. geometric Brownian motion)

$$dMC_i(t) = MC_i(t)\left[rdt + \sigma dW_i(t)\right], \tag{6.2}$$

   where r is the drift and σ is the standard deviation, and $W(t)$ is a standard Wiener process. Parameters r and σ are assumed to be the same for all cryptocurrencies, but the Wiener process $W_i(t)$ is specific to each.

2. **Independent random birth time and size**. The birth flow of each crypto, at time $t_i, i \in N$, follows a Poisson process with exponentially growing intensity $v(t) = v_0 e^{d \cdot t}$, and initial size $s_0^i = s_{0,i} \cdot e^{c_0 t_i}$, where $\{s_{0,i}\}_{i \in N}$ are independent draws from a common random variable.[5]

3. **Cryptocurrencies exit (die) at random with a constant hazard rate, $h \geq 0$, independent of size**.

Under these assumptions and mild conditions, asymptotically, the process generates a power-law distribution with tail index μ_{TH}

$$\mu_{TH} := \frac{1}{2}\left[(1 - 2\frac{r - c_0}{\sigma^2}) + \sqrt{(1 - 2\frac{r - c_0}{\sigma^2})^2 + 8\frac{d + h}{\sigma^2}}\right]. \tag{6.3}$$

It is important to stress that this is a very simple model with a number of limitations, listed below, and therefore the objective is only to capture the rough fundamentals of the dynamics of the crypto ecology.

- It does not capture the strong non-stationarities (e.g., bubbles and crashes) of the market;
- It treats crypto-currencies as independent despite the overall market being highly correlated, including some pairs being more correlated than others;

[5]Exponential growth is a standard feature of economic systems and financial markets. However, extension of the birth time process to a vast class of non-Poisson processes does not alter key results.

- It neglects difference in "fitness" (i.e., quality) of different crypto-currencies, which has been shown to be important in complex networks [229], and is clearly present as newer technology is introduced in newer coins;
- It neglects "forking", which is similar to spin-off/divestiture of a company
- The process only applies above a sufficiently high threshold, acknowledging that an entire complex ecosystem cannot be described by such a simple model.

6.2.2 Direct Empirical Quantitative Confirmation of Gibrat's Law of Proportional Growth

Gibrat's law of proportional growth embodied in Equation 6.2 implies that, for sufficiently small time intervals Δt, the mean change in market capitalization $\langle \Delta MC \rangle$ and the standard deviation of the change, $\sigma(\Delta MC)$, are both proportional to MC for large coins and tokens. Figure 6.4 shows the mean and standard deviation of ΔMC as a function of MC, setting Δt as one day, within a one year time window, confirming proportional growth.

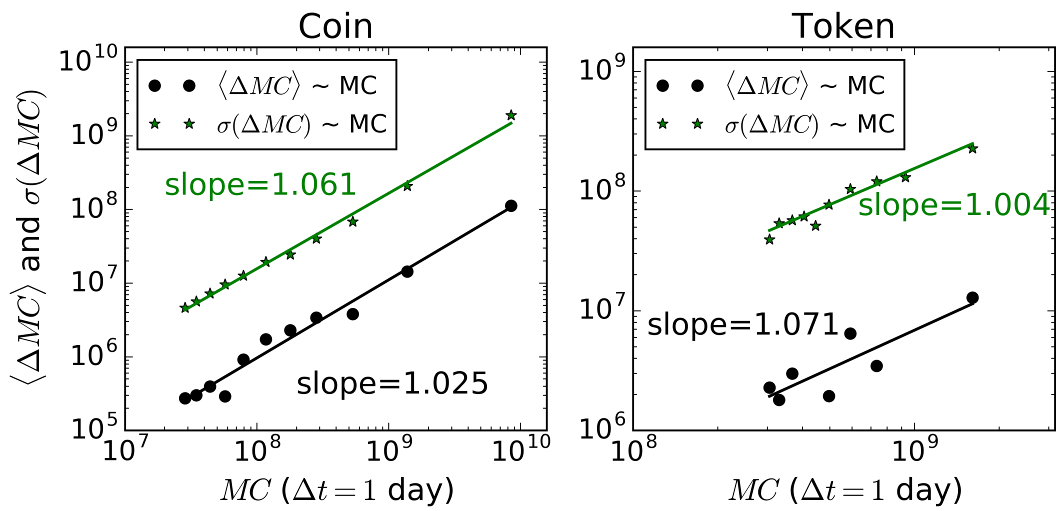

Figure 6.4: Test of Gibrat's law of proportional growth for market capitalization of large coins (left panel) and tokens (right panel) within a one-year window, up to Feb 7, 2018. The black circles are the mean of the increments (i.e., $\langle \Delta MC \rangle$) versus its current market capitalization MC. The green stars are the standard deviation of the increments (i.e., $\sigma(\Delta MC)$) versus its current market capitalization. Every 2000 increments (i.e. ΔMC_t) for coins (resp. 250 for tokens) are grouped into a subset for calculating these means and standard deviations. Only positive points are shown. In both panels, the lines show the least squares fit to the data points.

Moreover, Equation 6.2 implies that, over a small time interval Δt, the average growth rate $\langle\frac{\Delta MC}{MC}\rangle$ and its standard deviation should be given by,

$$\langle\frac{\Delta MC}{MC}\rangle = r \times \Delta t, \quad \sigma(\frac{\Delta MC}{MC}) = \sigma \times \sqrt{\Delta t}, \tag{6.4}$$

where the later square-root dependence reflects the property of the Wiener (random walk) process

This is verified via Figure 6.5, where we estimate the drift r (resp. the standard deviation σ) as the slope of the linear regression of the average growth rate (resp. standard deviation of the growth rate) as a function of Δt (resp. $\sqrt{\Delta t}$), with $\Delta t \leq 10$ days. We can see that the growth rate of coins is roughly two times that of tokens, while their volatilities are similar – the relatively large growth in the token market capitalization is therefore a result of the high birth rate, not an exceptional growth of individual tokens.

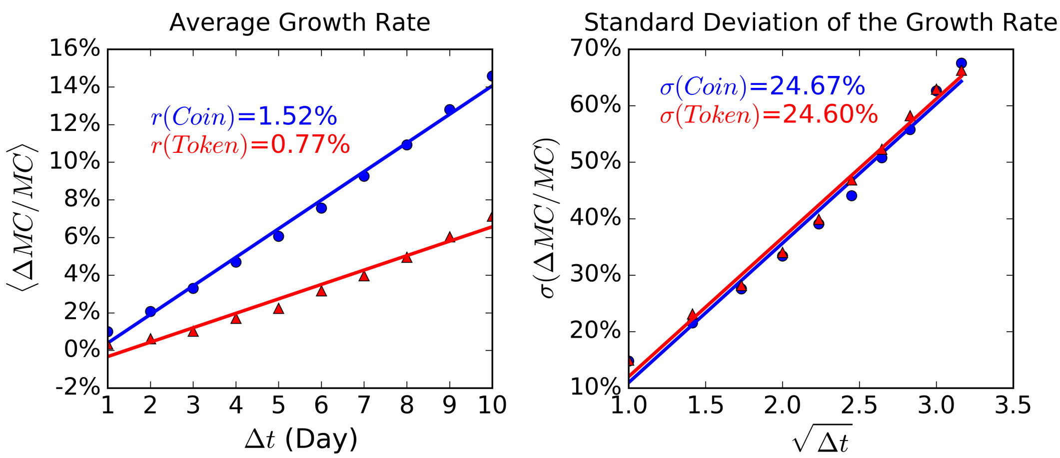

Figure 6.5: Left: The relationship between the average growth rate $\langle\frac{\Delta MC}{MC}\rangle$ versus the time interval Δt, for coins (blue circle) and tokens (red triangles) respectively. Right: The standard deviation of the growth rate $\sigma(\frac{\Delta MC}{MC})$ versus $\sqrt{\Delta t}$, for coins and tokens respectively. Data values were taken in the one year window ending Feb 7, 2018.

6.2.3 Estimation of the Birth and Death Parameters

For both coins and tokens, the distribution of birth market capitalization[6] (see Figure 6.6) has a substantially thinner tail (0.59 for coins and 1.48 for tokens) than the current distri-

[6]Note that the birth market capitalization we get from CoinMarketCap may have some delays, because: 1) sometimes cryptocurrencies are listed to CoinMarketCap after they have been traded for a while; 2) for some cryptocurrencies, the information of the circulating supply is not available to calculate the market

bution (0.48 for coins and 1.29 for tokens), whose exponents are estimated for the largest 100 coins and 30 tokens respectively, based on Maximum Likelihood. This confirms that the observed market capitalization distribution is not simply a consequence of the distribution of initial market capitalizations. Rather, the distributions becomes heavier-tailed due to proportional growth.

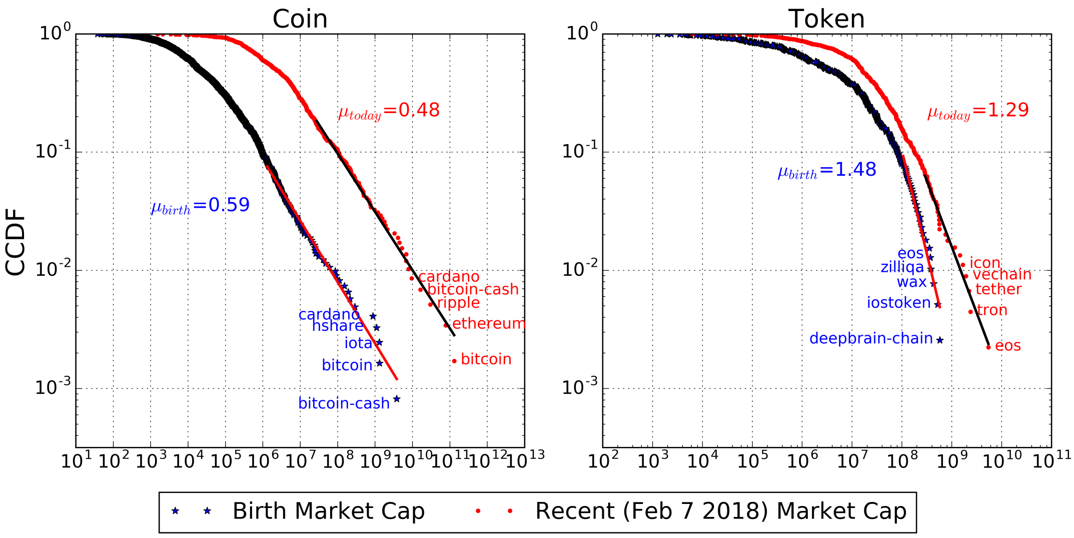

Figure 6.6: Comparison between the distributions of birth market capitalization and the recent market cap, for coins (left panel) and tokens (right panel) respectively. The black stars are the market capitalization at birth (taking the market capitalization 1 week after the birth as a proxy where supply and price are both known). The red dots are the market capitalization on Feb 7, 2018. The tail exponents are estimated by the largest 100 coins and 30 tokens respectively. The largest five coins/tokens are labeled in the upper panel.

Moving to the extended Gibrat's Law framework of Malevergne et al. [222], the growth rate of birth size c_0, the growth of the birth intensity d, and the exit hazard rate h are estimated, as summarized in Figure 6.7. Importantly, as a threshold is necessary, only coins (resp. tokens) having an average market capitalization over their lifetime larger than US Dollars $10^{7.3}$ (resp. $10^{8.1}$) are considered, which correspond to roughly the top 10% of coins and tokens. Further, due to non-stationarities not permitted by the framework, we focus on estimating the parameters in the most recent relatively stable window. In particular, the birth size of (high market cap) coins does not have a significant trend (p-value ¿ 10%), but rather shifted from one level to another around May 2017. Therefore, we fix the growth rate of birth size to be $c_0 = 0$. For tokens, however the birth size is significantly growing

capitalization at the birth time. Therefore, the distribution of market capitalization at birth ("initial market capitalization") we show in Figure 6.6 is later than the actual birth time.

with time (p-value ¡ 0.1%), especially after July 2017. Thus, c_0 of tokens is estimated to be 1.19%. The birth intensity of coins is relatively stable, giving the growth rate of birth intensity $d = 0$. However, the number of tokens has been growing significantly since May 2017 due to a large amount of ICOs (Initial Coin Offerings – like an initial public offering of equity). A linear approximation of the growing birth intensity implies a Token birth intensity growth rate of d to be 0.59%. In terms of the death process, there have been less than 3 dead large coins and tokens, so we consider the exit hazard rate h to be 0 for both coins and tokens.[7]

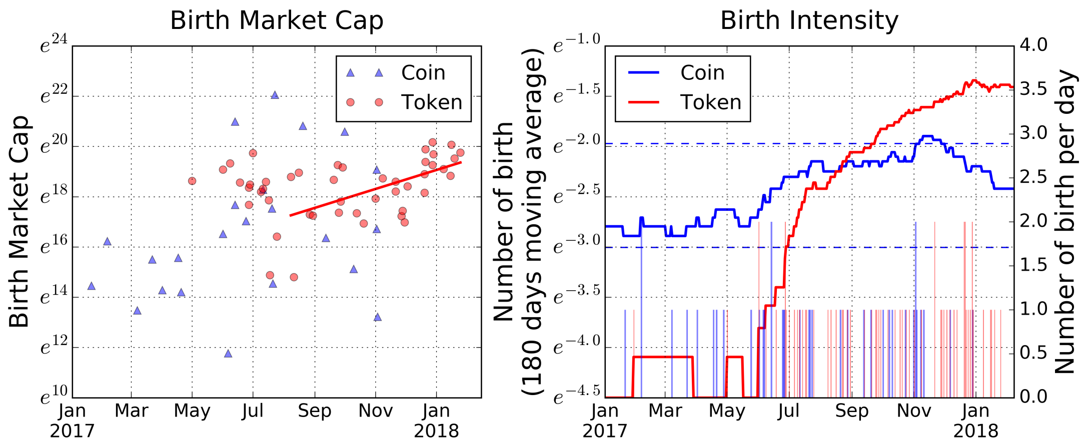

Figure 6.7: Birth market capitalization and birth intensity (frequency) of large coins (having average market capitalization over life larger than $10^{7.3}$) and tokens (average market capitalization over life time larger than $10^{8.1}$) since 2017. The left panel is the birth market capitalization of coins (blue triangles) and tokens (red circles). The red solid line is the best linear fit of the tokens' birth market capitalization. The right panel plots the number of birth of coins (blue) and tokens (red), smoothed by a 180 day moving average. The horizontal blue dashed lines in the middle and lower panels are the 95% confidence intervals assuming a Poisson process for coins, whose mean is estimated on the full window (since 2017). The number of births per day is shown in blue (coins) and red (tokens) bars, against the right y-axis.

6.2.4 Comparing Empirical and Theoretical Predicted Distributions

Given estimates of the five parameters in Equation 6.3, for coins and tokens separately, the theoretically predicted power law exponents are computed and compared with their

[7]However, note that the empirical distributions of the lifetimes of all coins and tokens suggest similar Exponential distributions, suggesting a similar death hazard rate for coins and tokens.

empirical counterparts[8], and summarized in Table 6.1.

Despite admitted model limitations, the empirically and theoretically predicted tail exponents of the market capitalization distributions are consistent, with the theoretical coin and token tail exponents, being less than and greater than 1 respectively. Comparing the empirical and theoretical tail exponents on time windows different from the one presented here is complicated by non-stationarities in the birth and death parameters, but still produces consistent results, with the theoretical and empirical tail exponents falling within similar ranges. This, in combination with the confirmation of Gibrat's law, effectively verifies the proposed model, delivering a robust insight into the underlying nature of the two fundamentally different coin and token ecosystems.

For coins, we have $r - h > d + c_0$ which means that the capitalisation growth (corrected for death) of existing "entrenched incumbents", such as Bitcoin, Ethereum, and so on, exceeds the growth of capitalisation of recent market entrants, of which there are relatively few. This inequality also theoretically implies a tail exponent μ less than 1 under the framework of [222]. In contrast, the token market has the opposite features, with $r - h \leq d + c_0$, implying a thinner tail with $\mu \leq 1$. Indeed in the token market, the high rate of birth of tokens is the dominant feature driving the market, and the limited growth r in excess of death h restricts growth of older tokens in relative terms, leading to a market capitalization distribution that is slightly lighter tailed than Zipf's law and reflecting an immature system.

6.3 Discussion

Having looked at the market capitalization of all cryptocurrencies, and treating coins and tokens separately, we aimed to understand the basic growth mechanism with a simple model. We have empirically verified that, for large coins and tokens, their market capitalizations follow power law distributions with significantly different values – with the tail exponent falling between 0.5 and 0.7 for coins, and between 1.0 and 1.3 for tokens.

Despite recognized limitations, the simple stochastic proportional growth model of Malevergne et al. [222] successfully recovers these tail exponents based on statistically estimated birth, death, and proportional growth parameters. This clearly characterizes coins as being "entrenched incumbents" and tokens as an "explosive immature ecosystem", largely due to massive and exuberant ICO activity in the token space.

[8]Empirical exponents are based on the market capitalization distribution on Feb 7, 2018, as shown in Figure 6.6.

Table 6.1: Comparison between the theoretically predicted power law exponents μ_{TH} and the empirical exponents μ_{MLE}, estimated by maximum likelihood, for coins and tokens respectively. The theoretical values μ_{TH} are given by Equation 6.3 with estimated birth and death parameters plugged in – see the previous subsections for their estimation. Numbers in brackets are the 95% confidence interval estimated by bootstrap.

	Coin	**Token**
Growth rate of market capitalization r	1.52% [1.45%, 1.59%]	0.77% [0.68%, 0.86%]
Growth volatility σ	24.67% [22.64%, 26.58%]	24.6% [22.74%, 26.45%]
Exit hazard rate h	0	0
Growth rate of birth size c_0	0	1.19% [0.48%, 1.90%]
Growth of the birth intensity d	0	0.59% [0.57%, 0.61%]
Empirical tail exponent μ_{MLE}	0.48 [0.39, 0.57]	1.29 [0.83, 1.75]
Theoretical tail exponent μ_{TH}	0.50 [0.41, 0.57]	1.29 [1.12, 1.48]

With Zipf's Law having a unit tail exponent and being a statistical signature of an optimal economy [222], it is perhaps unsurprising that the coin and token markets have different tail exponents. Undoubtedly, if more productive regulation is introduced [230, 231, 232, 233], and institutional investors flood the market and adoption grows, the markets will become more mature. One can then expect a better balance between the growth of incumbents and a healthy rate of new entrants associated with technological innovations. However, as the cryptocurrencies are evolving towards being an alternative investment asset, one should remain extremely cautious, where massive endogenous instabilities exist [105] and risks are poorly understood.

Looking forward, the methodology presented here could be productively extended to allow for varying quality (i.e., fitness) of cryptocurrencies. This would be realistic, as improved technology enters the market in new coins. In particular, such a framework could more adequately address the question of if and when Bitcoin will be overtaken, as pure proportional growth frameworks perhaps overly emphasize the strength of the so-called "first mover advantage".

Chapter 7

Conclusions

This thesis discusses financial markets' information efficiency, bubbles, different factors that might mitigate bubbles, fundamental values and other related issues, under the framework of different setups: two natural experiments, a newly proposed field experiment design, a lab environment, and a new and young market.

Starting from the two natural experiments of the UK Referendum and the US President Election, we employ a simple and natural one-factor linear model, to analyze the market efficiency in terms of their response to the stream of voting results. In practical terms for voting events, with high stakes, technological sophistication, and immense resources, one would expect the financial markets to ravenously consume and digest both public and rarefied information streams. However, in the case of the Brexit, our algorithm confidently and robustly predicts a Brexit outcome after only 20-30 out of 382 local voting results had been revealed, hours before the Pound market had reflected the outcome. In the case of the Trump election, however, the Peso market was efficient, as the growing probability of a Trump presidency was reflected.

The high probabilities early on demonstrate that the Brexit and the Trump victory were not coming down to aleatory "luck of the day" – in statistical terms, the outcome was significantly different from a tie. The strong delay of the market in the case of Brexit exhibits a market failure of comparable importance and significance, suggesting that markets can become massively inefficient when there is a "collective bubble spirit", in a general sense. We hypothesize that this spirit was formed by a large-scale group-think based on the political and social attractiveness of the left-wing vote, inflated and galvanized by the intense atmosphere of the election debate. In contrast, in the US Election case, this bubble bursts quickly and the market adjusted to the reality much faster than Brexit night.

The second part of this thesis proposes a new design for investigating experimental asset markets. This design is more realistic than the well-established SSW design. This experimental setup employs a prediction market approach to study stylised results observed in real financial markets and classical asset market experiments. The aim of implementing the new design was twofold. First, we investigated whether the "bubble-and-crash" pattern is a typical phenomenon only found in this type of experimental markets, or it is a general bias that is reflected in other artificial and real markets. Second, we aimed at testing the coordination of opinions in a situation of intrinsic uncertainty. In the two experiments, we show that mispricing is a robust finding that generalizes beyond the SSW design. We observed, however, the overpricing diminishes over time, indicating learning effect. Traders' price estimates showed a collective realization of communal ignorance, pushing the market much closer to its true value. We extend this experiment and investigate the impact of different type of information on market prices, by adding two additional treatments. We test that the information of the "correct" price level of the market with the clear indication reduces mispricing. We demonstrate that it is not the content of the ambiguous message that plays a role, but the expectation that such message will arrive and affect the prices for only particular assets.

Furthermore, we compare this experiment in the field and laboratory experiments, while using the same experimental design in two settings. We reproduce the three main findings from the field experiment, which demonstrates their robustness, especially the mispricing of the market, which has been widely reported in experimental asset market experiments. Through our comparison between the field and the lab, we demonstrate that standard laboratory experiments may not mimic the behaviour of real complex financial markets. Alternative setups can be developed with intermediate levels of control and complexity that may help close the gap between the maximally controlled laboratory conditions and the real financial markets.

After examining the natural, field and lab experiments carefully, we turn to a rising market, the cryptocurrency market. Modeling the market cap of cryptocurrency coins and tokens respectively, we managed to understand basic growth mechanism with a simple model. We have empirically verified that, for large coins and tokens, their market cap follow power law distributions with significantly different values – with the tail exponent falling between 0.5 and 0.7 for coins, and between 1.0 and 1.3 for tokens. Despite recognized limitations, the simple stochastic proportional growth model of Malevergne et al. [222] successfully recovers these tail exponents based on statistically estimated birth, death, and proportional growth parameters. This clearly characterizes coins as being "entrenched incumbents" and tokens as an "explosive immature ecosystem", largely due to massive and exuberant ICO activity in the token space. We argue that the market is not fundamentally different from many other systems and markets, where massive bubbles and crashes exist due to its endogenous instabilities, and risks are usually poorly understood.

This thesis studies the financial markets empirically and experimentally, which together give useful insights to understand the universal nature of all the markets. If we take into accounts of many behaviors of human beings, which are the basic components of a market, it is not difficult to understand ubiquitous bubbles, crashes and other "anomalies" in the markets. However, there are many factors that are not included in the models and the setup we proposed in this thesis. As mentioned, the turnout factor, demographic and psychology factors are very important in predicting the election, and thus in understanding the market. In the experimental asset market, short selling, leverage, and other factors are still excluded in the current setup and could be further tested. In the model of cryptocurrency market capitalization, a number of limitations, such as non-stationarity, correlations, qualities of different coins/tokens, have been discussed in Chapter 6. Nevertheless, the simple and intuitive models and setup in this thesis already provide useful and direct insights of financial markets, and have a great potential for future developments and studies.

Appendix A

Appendix to Chapter 2

A.1 2016 UK Referendum analysis based on crude Euro-skepticism polls

In this section we present the results of our prediction method based on a normalized average of the crude Euro-skepticism polls from Sky News [151], and YovGov [152] [1].

Similar in the main text, we show the rolling regressions of actual Leave vote percentage against the ex-ante prediction in Fig A.1, the simulated probability of Brexit in Fig A.2, and the cumulative errors in Fig A.3. With this crude prior of Euro-skepticism, we still find very stable linear relationship between the ex-ante prediction and the realized results. The simulated probability for Brexit was typically above 0.95 since the 20th announcement, with the convergence to 100% happening at around the 100th announcements, which is still quite early compared to the market. It is worth to note that this crude prior is nothing but ranking of the local voting areas by Euro-skepticism. Even with this information, one has a strong indication of a Brexit result early on, and a confident result after less than 100 announcements.

[1] We divide both of the rankings by the maximum rank, and take the average of the divided ranking as the prior estimation. This normalization combines the results of the two rankings from two sources, and constrains them into the same interval from 0 to 1. Regarding the 6 areas that are not in the rank, which are Gibraltar, Orkney Islands, Isles of Scilly, Shetland Islands, Comhairle Nan Eilean Siar and Northern Ireland, we give them a conservative value of 40%, which means that they are at the bottom 40% in terms of Euro-skepticism ranks.

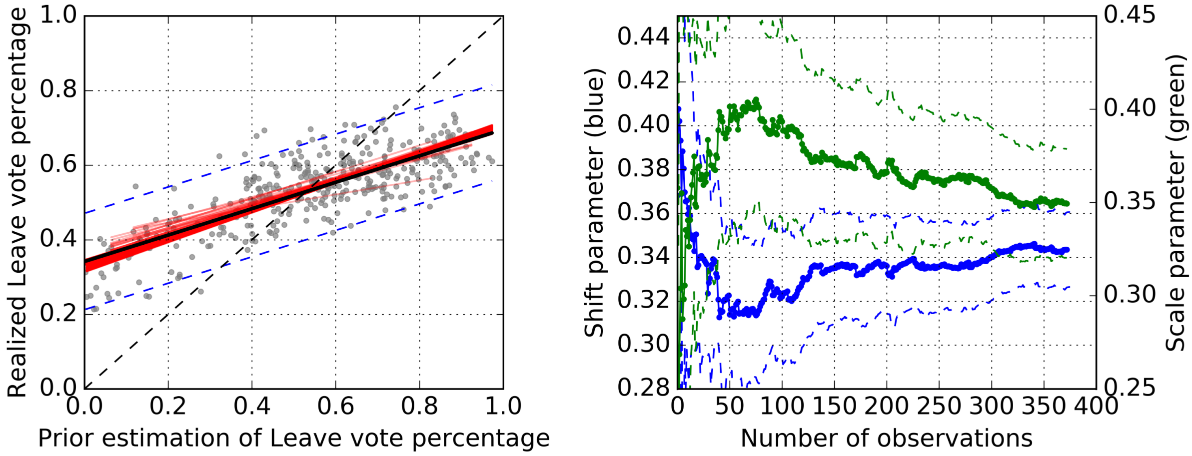

Figure A.1: Regression of the actual Leave vote percentage against the ex-ante prediction based on Euro-skepticism polls, and evolution/stability of parameter estimates for rolling regressions. On the left side, data is plotted in gray dots, while the black thick line is the regression line for all of the data, together with its 95% confidence intervals (blue dashed lines); red lines are the regression lines for the 378 rolling regressions. On the right side, the evolution of shift and scale parameters are plotted in blue (left y-axis) and green (right y-axis) respectively, together with their 95% confidence intervals in dashed lines.

A.2 The Monte-Carlo Estimation of Brexit Probability

Note that for the ex-ante prediction of Hanretty, there are 4 areas without prior estimation: Isles of Scilly, Gibraltar, Anglesey, and Northern Ireland. According to the poll conducted by Mcbride[234] on June 17th, 2016, 40% of people in Northern Ireland intending to vote would choose to Leave, excluding undecided people. Hence, we give Northern Ireland a conservative prior estimate of Leave of 40%. For Isles of Scilly and Anglesey, we give both of them a conservative prior estimate of 40% Leave, while the actual Leave vote percentages are 43.61% and 50.94% respectively. These priors are used for predictions for these areas, but the regressions for estimating parameters in Eq (A.1) excludes them to avoid the potential for “tuning” results.

a) At each time $t_k > t_3$, when the result of area k is announced, we do the regression with the equation

$$l_i = \alpha PL_i + \beta + \varepsilon_i \ , i = 1, 2, \cdots, k. \tag{A.1}$$

Hence, the parameters α, β follow a Normal distribution $N\left(\hat{\alpha}, \hat{\Sigma}_\alpha\right)$, $N\left(\hat{\beta}, \hat{\Sigma}_\beta\right)$,

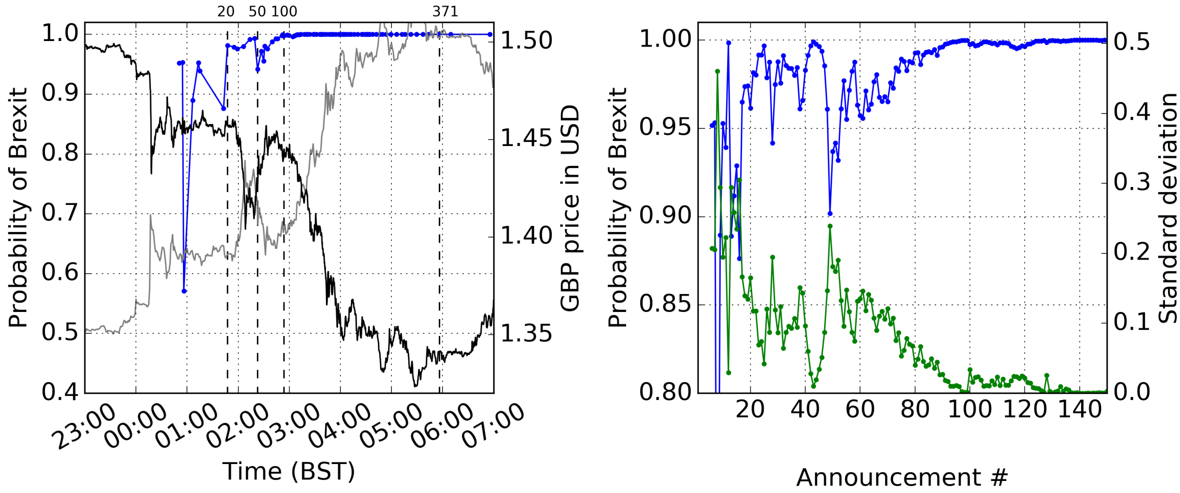

Figure A.2: Evolution of the Monte-Carlo estimated probability of Brexit based on Euro-Skepticism. Left panel: as a function of time, the blue line is the simulated probability, the gray line is the market implied probability, and the black line is the GBP price in USD. All of these probabilities are plotted once for every 5 areas. The vertical dashed lines indicate the time of the 20th, 50th, 100th, and 371st announcements. Right panel: higher resolution of the two probabilities in the range of the first 50 announcements. Blue line shows the probability of Leave, and green line is the corresponding standard deviation of the simulated probabilities. The Monte Carlo estimation is performed with $N = 1000, M = 1000$ simulations, and the initial probability of Brexit b_0 used in calculating market implied probability is 50%.

where α, β, $\hat{\Sigma}_\alpha$, $\hat{\Sigma}_\beta$ are the standard linear regression estimates. If the standard deviation of the residuals is denoted by σ, then the sum of squares of residuals divided by σ^2 has a chi-squared distribution with $k-2$ degrees of freedom, where $\hat{\sigma}$ is the standard deviation of the residuals from the above regression.

b) Simulate α and β from their Normal distributions for M times, and simulate σ based on the chi-squared distribution. Then, we have M groups of parameters $(\alpha_m, \beta_m, \sigma_m), m = 1, 2, \cdots, M$.

c) For each group of parameters $(\alpha_m, \beta_m, \sigma_m)$, simulate the Leave vote percentage l_i for each unknown local area $i \geqslant k+1$ for N times, from the distribution $N\left(\hat{\alpha}_m PL_i + \hat{\beta}_m, \hat{\sigma}_m\right)$.

d) For each unknown local area $i \geqslant k+1$, simulate the turnout ratio w_i for N times, by uniformly sampling from the turnout ratios of the k already announced areas.

e) With the simulated Leave vote percentage l_i and turnout ratio w_i for each unknown area i, we have N series of full voting results, which generate N final total vote

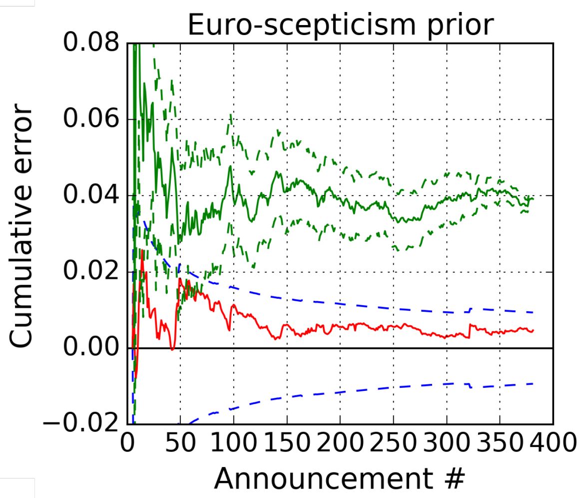

Figure A.3: Cumulative errors (red) based on one-step-forward prediction error, together with the evolution of predicted end-state percentage voting difference (green). The dashed green and red lines correspond to one standard deviation intervals respectively. The dashed blue lines represent the 95% confidence interval under the null model.

difference D_{total} in favor of Leave. The simulated frequency of Brexit at time t_k is simply the number of $D_{total|t_k}$ that is larger than 0 divided by N.

f) At each time t_k, we have M simulated frequencies, and the simulated Brexit probability $\widehat{B}(t_k)$ is simply the arithmetic average of these frequencies, i.e.

$$\widehat{B}(t_k) = \frac{1}{M}\sum_{M}\frac{1}{N}\sum_{N}\mathbb{1}_{D_{total|t_k}>0} \tag{A.2}$$

A.3 The Monte-Carlo Estimation of Trump Victory Probability

In US President Election case, the result of the election is determined by the number of Electorate Votes, and there are more than two candidates of the election. Therefore, we regress the realized intermediate voting results against prior estimates for Trump and Clinton respectively, and then calculate the number of electorate votes. Note that in the end CNN only reported 23 states' intermediate results and they are used in our regressions. These 23 states include all the swing states. For the other states, we take the voting results without the voting percentages into our model, when they are called by the media.

Let's say that there are K announcements with announcement time t_k $(k = 1, 2, \cdots, K)$, updating voting percentage for Trump r_i, Clinton l_i and completed voting percentage w_i for state i $(i \in S_k)$, where S_k is the set of states that have completed counting at least 15% votes by t_k. The corresponding prior estimate of voting percentage for Trump and Clinton are PR_i and PL_i. Therefore, at each time t_k, if $|S_k| \geq 5$, we do the regression for Trump and Clinton respectively with the equations

$$\begin{aligned} l_i &= \alpha_1 PL_i + \beta_1 + \varepsilon_{1i} \ , i \in S_k \ , \\ r_i &= \alpha_2 PR_i + \beta_2 + \varepsilon_{2i} \ , i \in S_k \ . \end{aligned} \tag{A.3}$$

Then we sample results for outstanding states with the process a) to c) in the previous section, and can thus estimate the real-time probability of a Trump win,

$$\begin{aligned} PR(t_k) &= \Pr(EV_R(t_k) - EV_L(t_k) > 0 \mid I_{t_k}) \\ &= \Pr(\sum_{i=1}^{51} EV_i \mathbb{1}_{r_i > l_i} - \sum_{i=1}^{51} EV_i \mathbb{1}_{r_i < l_i} > 0 \mid I_{t_k}) \ , \end{aligned} \tag{A.4}$$

where I_{t_k} is the voting result information available at time t_k (i.e., in real time), $EV_R(t_k)$ and $EV_L(t_k)$ are the final electorate votes for Trump and Clinton estimated at time t_k respectively, and EV_i is the number of electorate votes in the i^{th} state.

A.4 Turnout

It is always difficult to untangle the dependence between the turnout factor and voting preference, which partly accounts for the linear relationship between the ex-ante prediction of local voting percentage and the observed voting percentage.

For 2016 EU Referendum, turnout factor has been regarded as one of the major reasons for the surprising results [153]. Comparing the 2016 EU Referendum turnout with the *Turnout factor*, PT_i, $i = 1, \ldots, 382$, being turnout fractions for each voting area in the 2015 General Election (Fig A.4, left), a consistently positive relationship exists (including for rolling regressions), implying that this "prior information" would have been useful for prediction.

Across the voting areas, there was a range of -10% to 15% in the difference in turnout percentage relative to the 2015 election, as shown in the right plot of Fig A.4. Taking this difference as an explanatory variable, there is a clear positive relationship with the

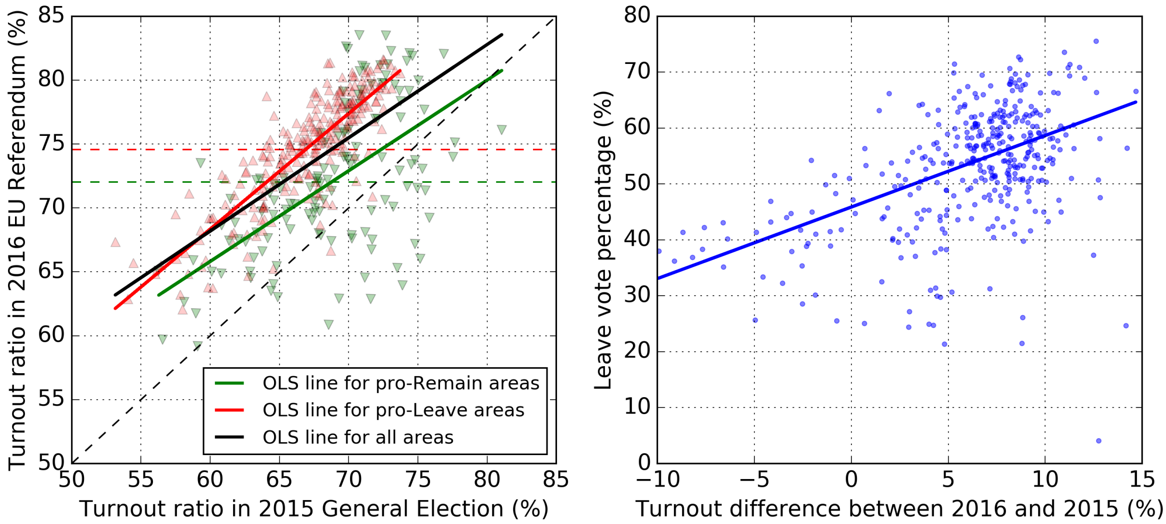

Figure A.4: Left panel: turnout ratio in the EU Referendum versus the turnout ratio in the 2015 General Election. The solid red line is the Linear Regression line for pro-Leave areas (red triangles), the solid green line is for the pro-Remain areas (green triangles), and the solid black line is the result of all areas. The horizontal lines correspond to the average turnout ratios for pro-Leave areas (red dashed) and pro-Remain areas (green dashed) respectively. Here, the pro-Leave areas are those whose voting preference for Leave is larger than 50% in Hanretty's prior. The black dashed line is the diagonal $y = x$. Right panel: the actual Leave vote percentage versus the increase of turnout percentage from 2015 to 2016, with a linear regression line.

observed voting area Leave fractions, l_i, in the Brexit vote, having intercept 44%, and slope about 1.4 when modelled by simple linear regression[2],

$$l(\Delta w) = \frac{L}{L+R} \sim 0.44 + 1.4\Delta w + \varepsilon, \tag{A.5}$$

where, for each voting area L and R are the actual number of Leave and Remain votes, l and r their percentage versions, and Δw the difference in turnout percentage between the Brexit vote and 2015 UK Election. The scale parameter in the regression is estimated to be 1.4, meaning that a 1% increase in the turnout percentage corresponds to a 1.4% increase in the Leave vote percentage. The causal inference here is that pro-Brexit voters were more likely to turn out, and there must be a significant amount of voters shifted from pro-Remain to pro-Leave, whereas a slope of 0 would have indicated no relationship between turnout and vote preference. Importantly, had Brexit turnout remained equal to

[2]The regression method is the standard Weighted Linear Regression with the weights set as electorate size. The fit and coefficients are highly significant (F-test and parameter p-values $\leq 10^{-16}$), but the data is noisy (R^2 measure about 0.3).

the one in 2015, such that $\overline{\Delta w} = 0$, then an average Leave fraction of $l(0) = 44\%$ would be expected! However, with the actual average turnout being 6% higher than in 2015, the average Leave fraction, $l(6) = 0.44 + 1.4 \times 6\% = 52.4\%$, was large enough to deliver a Brexit result.

For 2016 US President Election, as turnout data is only available at state level rather than local county level, it is difficult to draw a conclusion about the dependence between turnout and voting preference. However, there have been evidences and reports showing that low turnouts in a few key states, especially poor Democratic turnouts contributed significantly to the Trump's victory [e.g. 235, 236].

Appendix B

Appendix to Chapter 3

B.1 Experimental Instructions

Dear Students, This document will provide you with the necessary information for the trading market for this class. Should you have any questions that are not answered in here, please contact Philipp Rindler at prindler@ethz.ch. In the interest of fairness, please address all questions by e-mail. Answers will be sent to all students so that everybody has the same basic information.

Goals of the Experiment

There are three primary goals that we pursue with this experiment. First, we are scientifically interested in the results of the market and how you will trade. Second, we want to offer you a pedagogical experience in a real trading environment so that you can apply some of the concepts learned in class. Finally, it is an opportunity for you to earn extra points in addition to the final exam.

Your Compensation

Each week, your final earnings are recorded. At the end of the experiment (end of the semester), your total earnings over all weekly sessions will be used to compile a ranking of all students. The top 25% students with the highest earnings will receive a bonus of 0.5 grade points. The next 25% will receive a bonus of 0.25 grade points. These grades will be added to your grade on the final exam (whereby 6.0 cannot be exceeded).

General Information

Every week, you will be asked to predict the page number of the final lecture slide that Prof. Sornette will talk about in next week's lecture. To do so, you will trade securities that pay out 100 FMRF (FMR francs) if and only if a particular number is realised. Your goal is to accumulate as many FMRF as possible in order to gain bonus points to increase your grade in the class. You can gain FMRFs by trading during the week and by the final pay off at the end of each week.

To be concrete, page number refers to the actual page number in the pdf file. At the end of the class, Prof. Sornette will publicly announce the realised state. Tradable securities are available for possible results only. Securities for pages that have already been covered are not available in any given week.

Prof. Sornette will not be aware of the trading results during any week. Each week, he will be completely ignorant of the trading results and your portfolio holdings. The flow of the lecture will not be, in any way, affected by your trading.

Similarly, market manipulation, in the form of excessive questioning or interruption of the lecture, will not be tolerated. Prof. Sornette will resist any such attempts during the lecture.

Market Description

You will trade on a market platform on Innovwiki. In order to gain grade points, you have to participate in the market. To do that, you will need to create an Innovwiki account. You need to have an account by November 9^{th}.

The market consists of all students in this class and you each trade individually from your own account. Every week, you will receive a 300 FMRF and 3 units of every security. You may buy or sell securities at any time during the week. You are allowed to input market orders or limit orders. When you issue a market order, the trade (buy or sell) is executed at the current price (if you can afford it). A limit order is entered into the trading book until someone else agrees to the trade. Note that you cannot enter into a trade that you cannot afford: your cash balance and asset balance cannot drop below zero.

During the week, you can make as many trades as you like, subject to the limit that you have to be able to honor your commitments: if you offer to sell a certain number of securities, you have to own them at the time that you submit that offer. If you want to buy a certain number of securities at a certain price, you have to have the cash balance available for that trade.

Each trading week, the market opens after class at midnight between Monday and Tuesday and remains open until the end of the week until midnight between Sunday and Monday. During this time, you are free to use the trading mechanism to obtain your optimal portfolio. You may rebalance your current portfolio should market situations change or your opinion changes.

At the end of each class, the realised number of slides is announced and your account will be credited with your payoff for that week. Since the securities pay off 100 FMRF if their respective page number is realised, your final payoff is equal to 100 FMRF times the number of units of the correct state security that you have in your portfolio at the closing of the market.

The Innovwiki trading platform provides information on the time series of the prices of all securities as well as the order book, which presents all the standing orders by you and other students to buy or to sell that are waiting to be fulfilled. On the basis of this information and your own analysis, you will trade along the week to position your portfolio in your assessed optimal way.

Each week, the market is reset completely. No money or asset positions are carried over. Your earnings for the week are recorded by the system. You can access information about your own earnings but not of others.

Earnings

Each week, you get an endowment of 300 FMRF and 3 units of each security. This loan has to be repaid at the end of the week at 600 FMRF. This is the payout you would receive if you do nothing during the week. Hence, the total earnings at the end of each week is your cash balance at the end of the week, plus the pay offs from your securities minus 600 FMRF. Therefore, you can be ahead of the rest of the class by buying and selling intelligently or by predicting the outcome correctly, or both.

Definition of Securities

In the market, securities are available for each outcome that can happen every week. Each security pays off if class ends on one of three consecutive slides. The last security may cover fewer depending on how many slides there are in total. Securities in the market are named in the following manner: Li.Pa-b. The letter i refers to the lecture, the numbers a-b to the page number that needs to be realised for the security to pay off. For example, L3.P4-6 is the security that pays off if class ends on page 4, 5, or 6 of lecture 3.

Each week, the following securities are available for trading. Starting from the ending slide of the last lecture, there is a security for each following page in the same lecture. In addition, if a new lecture is uploaded that week (which will always happen on Monday after class), securities that cover all pages of that lecture will also be available for trading. At the end of each lecture, Professor Sornette will announce whether the next lecture will continue exactly from he left off or whether he will start from a slide further ahead, for instance at the first slide of the next lecture notes. The securities available for trading will always reflect this information.

For the number of slides, the page number in the pdf document is relevant, not the page number shown on the slides! Please note that on display, some slides appear to be "animated" but are in fact a set of different pdf pages. Each pdf page counts as a different slide!

Example

As an example, let's consider the following made up situation: class on a Monday ends on page 12 of lecture 3. No additional lectures are uploaded that week. The total number of pages in lecture 3 is 104. So for the next class the following Monday, there are 92 possible pages on which class could end: page 13 to page 104 of lecture 3. Therefore, there would be 31 different securities available for trading: L3.P13-15, L3.P16-18, and so forth until L3.P100-102, L3.P103-104 that each pay off if and only if the final page number is among their respective pages indicated by their names.

At the beginning of the week, you would receive 300 FMRF and 3 units of each of the 31 securities. You are free to trade any of these at any price. Of course, for trades to occur, someone else has to take the other side of the bargain.

As an alternative scenario, assume again that class on Monday ends on page 12 of lecture 3 but now lecture 4 is uploaded on that Monday as well. In this case, you will be able to trade on each page in lecture 3 and 4. Lecture 4 contains 90 pages. Therefore, in this market there would be 61 different securities available for trading:

- 31 Securities that pay off if and only if the final page number is among their respective pages of lecture 3: L3.P13-15, L3.P16-18, and so forth until L3.P100-102, L3.P103-104.
- 30 Securities that pay off if and only if the final page number is among their respective pages of lecture 4: L4.P1-3, L4.P4-6, and so forth until L4.P85-87, L4.88-90.

In this case, you would receive the same 300 FMRF and 3 units of each of the 61 securities.

Your Task

Your goal is to hold a portfolio that you find optimal. To do so, you will have to estimate the probabilities for the different states. You can use the market prices unfolding over the week pushed by buys and sells by yourself and your fellow students to infer the markets assessment of these probabilities in a kind of a wisdom of crowds mechanism. Use the market to buy state securities that you find undervalued and sell state securities that you find overvalued.

To help you with your task, the online platform provides you with a number of tools. For example, for each security you can see the price history and transaction volume over time, both graphically and numerically.

Literature

In order to arrive at your optimal decisions, you will have to compare the market behavior to your own expectations and adjust your portfolio accordingly. You can find further information on designing an optimal portfolio in a setup such as in this experiment (state-preference approach) in the following articles:

- Hens, T. & M. Rieger (2010). Two Period Model: State-Preference Approach In T. Hens, T. & M. Rieger (Eds.), Financial Economics (pp. 141-209). Heidelberg, Springer.

- Hirshleifer, J. (1966). Investment Decision Under Uncertainty: Applications of the State-Preference Approach. Quarterly Journal of Economics, 80(2), pp. 252-277.

- Kerruish, A. (1983). Can we use State Preference Theory? Managerial Finance, 9(3/4), pp. 52-57.

- Sauer, R. (1998). The Economics of Wagering Markets. Journal of Economic Literature, 36(4), pp. 2021-2064.

B.2 Details of Experimental Method

B.2.1 Participants

In a class of 234 students with different majors at the MSc level at ETH Zurich (the Swiss Federal Institute of Technology in Zurich, Switzerland), enrolled in the course "Financial Markets Risks" in Fall 2014, the students were asked to take part in a trading experiment. Participation in the experiment was voluntary and 102 (44 %) of the total students actively participated. The ratio of men to women was approximately 4:1.

B.2.2 Materials

The specific numbers of securities for each of these four weeks were 46, 26, 48 and 53, associated with respectively 138, 78, 144 and 159 slides. The professor finished on slide 60 (43% coverage), 68 (87%), 47 (33%) and 57 (36%), in four weeks.

B.2.3 Procedure

As outlined in panel A of Figure 3.1, each week, the market was continuously open from Tuesday at 00:01 until Sunday midnight before the class on Monday. Orders could be put at any point during this period. All buy orders had to be covered by sufficient cash in their account and sell orders were only allowed if the participant had the necessary quantity of securities in their portfolio. No short sells and no buying on margin was allowed. The trading rule follows the standard continuous double auction mechanism.[1] The experiment started after the 9^{th} lecture allowing the participants to familiarize with the professor's teaching style and lasted 4 weeks. During lecture 6, the trading task was announced and explained in detail.

B.3 Additional Analyses

B.3.1 Experiment 1: Trading activity

The market was very active with the largest numbers of trades occurring just after the market opened and just before it closed, similarly to real financial markets. According to Figure B.1, there is substantial trading activity during each day, although the number of transactions decreased towards the end of the week. During each week, the largest number of transactions occurred in the afternoon after the new set of slides was uploaded by the professor and in late evening hours just before the closing of the market. When we look at intra-day pattern, there were always more trades in the evening (i.e., after 8 p.m.), and almost no transactions occurred between 2.00am and 6.00am. After we exclude these times with very few trades, we observe an average of 19.4 transactions per hour with a median value of 6.

For the further analysis, we bin the trading data in four-hour intervals. Excluding the time from 2am to 6am with virtually no transactions, there are no four-hour intervals during which no transaction occurred. Every trading period (week) is therefore subdivided into 31 epochs of four hours each, with only the first epoch of each week lasting for 2 hours, from midnight to 2.00am on Tuesday night. On average, there were 75.1 transactions in any four-hour interval with a median of 56.

The activity of participants decreased from week 1 to week 4 – 34%, 32%, 41% and 53% of the participants did not make any trade in weeks 1 – 4. The number of participants

[1] A trade was successful only if there was a buyer that wanted to buy one or more units of a security for a price at least as high as a seller was offering.

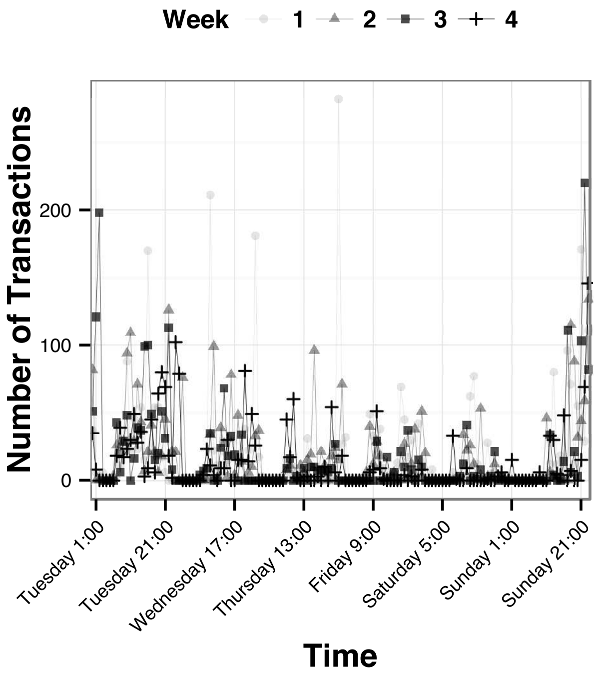

Figure B.1: Experiment 1: Trading Activity during the four repetitions over the four weeks.

that were very active traders (i.e. the traders whose number of submitted orders was 1.5 times the distance between the 75% and 25% quantile) decreased from the first week and stabilised over the three following weeks (15, 18, 15, 10 for weeks 1 – 4). The number of traders in each week decreased from week 1 to week 4 and was 72, 72, 68 and 55.

The number of trades decreased from 2696, through 2378, 2088 to 1618 across weeks. This decrease was not only due to the decrease in the number of market participants, but the activity of the participants also decreased – The average number of trades per participant in each week was 37, 33, 31, 29.

B.3.2 Experiment 1: Order Book Summary

To provide more insights into how the opinion about the prices was formed, we analysed the order book. For each week, we split the securities into good (indicating participants' high expectation to pay the dividend) and bad (indicating participants' low expectation to pay the dividend), according to the median prices at the closing of the market. Therefore, the good securities were the securities around the peak of the price distributions in Figure

B.2 in the main text, whereas the bad securities were in the tails of the distributions. There are three main results that we can highlight from these data.

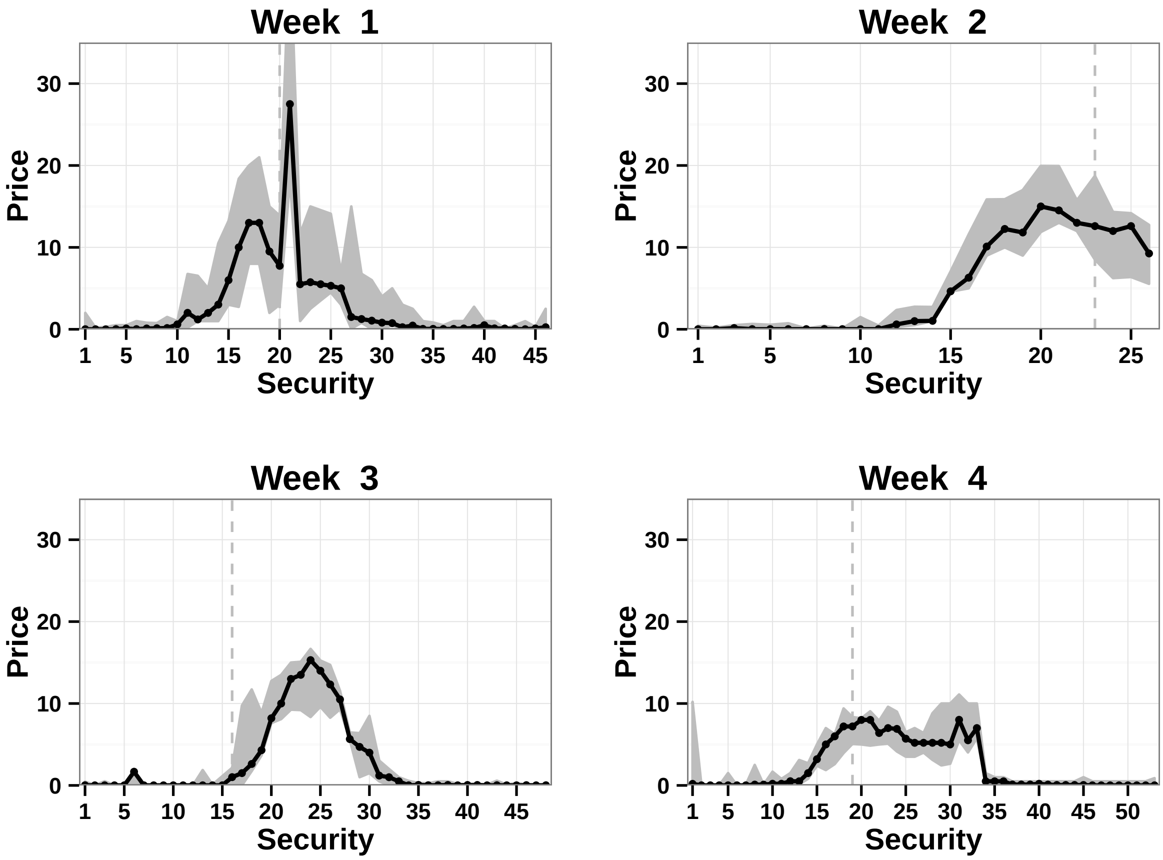

Figure B.2: Experiment 1: For a given week and a given epoch of four hours, we take the median price of all transactions within that epoch. Then we construct the average price over the 31 median prices of the 31 epochs of each week with 95% confidence intervals. The four panels show the resulting distribution of prices over all securities in each week. The security whose state was realised at the end of the week (payout of 100 units) is marked with the dashed line.

First, the number of orders was very strongly correlated with the final prices of the securities ($r = .64$, $r = .78$, $r = .73$ and $r = .85$, $p < .001$ for weeks 1 – 4 consecutively), such that the distribution of prices had the same shape as the distribution of the number of transactions. This implies that people traded the good securities more ($N_{\text{orders}} = 1987$, 1949, 1721, 1454 for weeks 1 – 4) than the bad ones ($N_{\text{orders}} = 1118$, 853, 795, 579 for weeks 1 – 4) and this correlation became more pronounced from week 1 to week 4, indicating a learning effect.

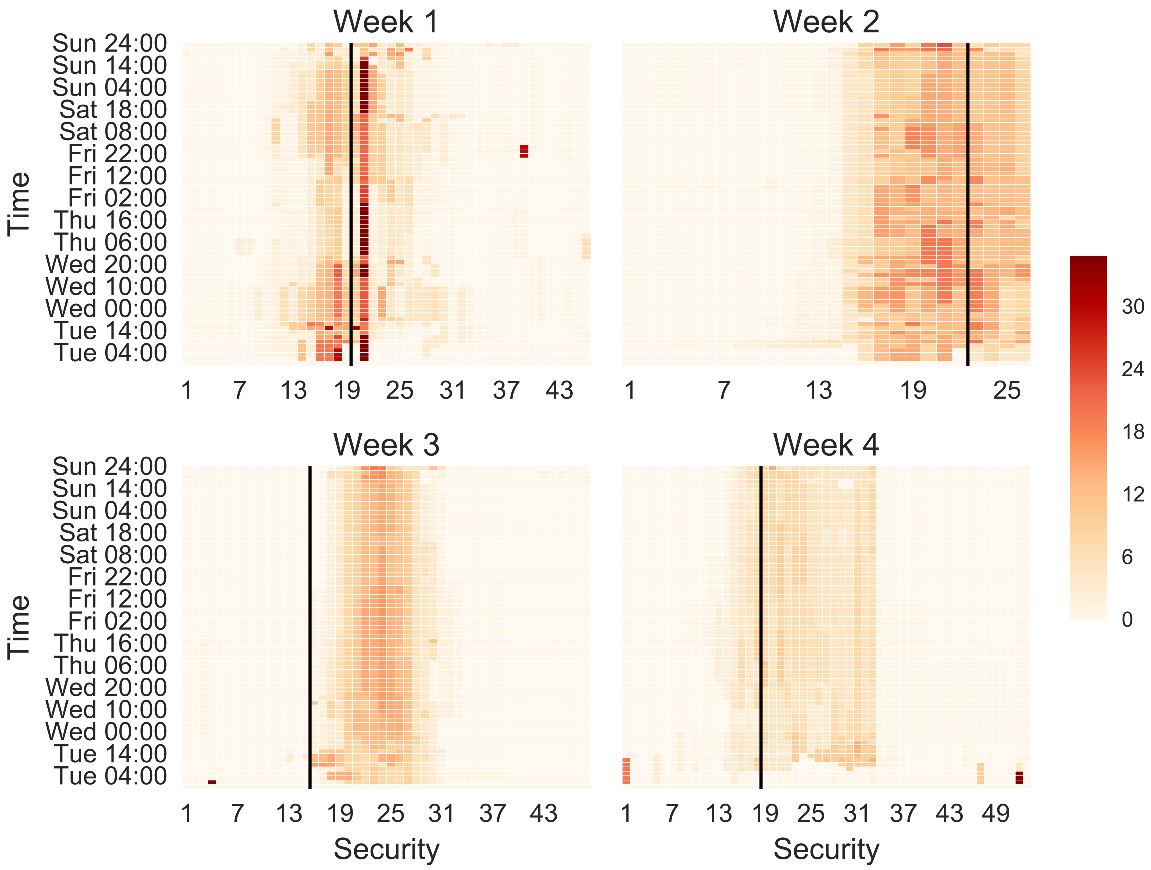

Figure B.3: Experiment 1: The evolution of security prices over time. Within a trading period (one week) the prices are generally stable; price levels are established rather early in the week and remain relatively constant. In week one, the distribution of prices is very peaked, with one security emerging as a clear favorite. The tendency however diminishes over periods, and by the last week of the experiment, aggregate price levels are much flatter.

Second, on average, we did not observe a significant difference between the spreads between asks and bids for the good securities compared to the bad ones (Mean difference between spreads of good and bad securities is Δ_M = .32, .18, -2.72 and .81 for weeks $1-4$,), apart from week 3, in which the spreads were relatively high for the first 20 securities. Overall, we did not observe differences in variance of spreads between good and bad securities. Median prices for both asks and bids were higher for the good securities (Ask: $Md = 6$, 5.01, 1, 1.5; Bid: Md =3.07, 4.7, 0.5, 0.41) than for the bad securities (Ask: $Md = 0.03, 0.15, 0.01, 0.1;$, Bid: Md =0.04, 0.02, 0.01, 0.01 for weeks $1-4$).[2] Median bid-ask spreads during the first hours after the market opened were much smaller for the bad securities than for the good

[2]We calculate the median prices from the whole trading period.

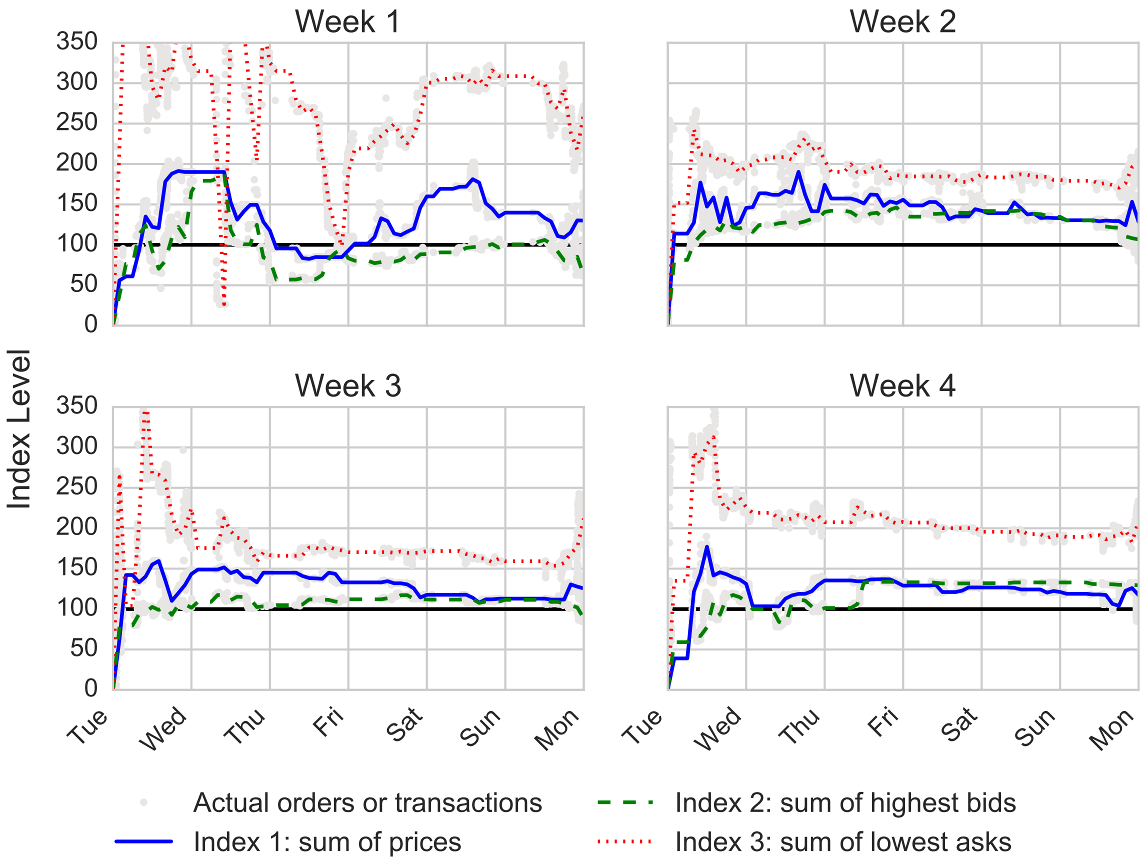

Figure B.4: Experiment 1: Evolution of the sum of all security prices for the four weeks of trading. The sum of all security prices should be 100 but there are several pronounced deviations from this normative prediction. The indices are smoothed using a 2-hour moving average.

securities. During the first few hours, the spreads were negative for the bad securities and positive for the good securities. Spreads between asks and bids were approaching 0 at the end of each week. These results reflect the agreement among the participants about which securities are the most valuable. In the very first orders, the good securities had higher prices and these securities were traded more frequently until the end of each week.

B.3.3 Experiment 2: Trading Activity

The market was very active – 128, 102, 102 and 97 participants submitted pre-trading beliefs and 99 (77%), 86 (84%), 82 (80%) and 87 (90%) post-trading beliefs. This means that some traders had access to the market by submitting their initial beliefs but did not want to have their portfolios included in the final ranking, or some simply forgot to submit their second belief, as reported by some of them. We did not exclude any of these from the analysis because this structure reflects the real-life imperfections of complex social systems. Also, all students that had access to the market formed it. Therefore, excluding activity of the students that did not submit their final belief would not reflect the market the way the participants experienced it.

As in experiment 1, the activity of the market decreased from week 1 to week 4. First, the number of submitted orders decreased across weeks, from 4443, through 2700 and 2006 to 1493 in the last week. Also, the number of active traders (i.e. the traders whose number of submitted orders was 1.5 times the distance between the 75% and 25% quantile) was 12, 15, 12 and 9 in the consecutive weeks, indicating a stability in the activity, except for the last week.

As in Experiment 1, the market was the most active on Tuesday, when it opened and on Sunday, just before it closed. Different from Experiment 1, the highest daily activity was in the morning.

B.3.4 Experiment 2: Self-reported Measures

The post-trading questionnaire revealed that 83% of the participants used the number of slides covered in the previous lectures to predict the next end slide.[3] However, only 50% of the participants studied the professor's slides before trading and the majority (60%) would spend less than 30 minutes on studying the slides. The second most popular cue (45%) was participant's own belief submitted before the trading. Other strategies were used by 24-37% of the participants. These included the average time spent by the professor on each slide, the time spent on presenting slides, the number of topics usually covered by the professor, bid-ask prices offered by the traders and security prices in the market. Two participants indicated using a model based on the previous lectures. It is important to note that most of the participants claimed to use more than one strategy.

66% of the participants realised that the sum of the prices should be equal to 100 at any time. However, only 44% of the participants realised that this mispricing can be used to arbitrage the market. Out of these, 45% applied the arbitrage strategy explained in the description of the results of experiment 1. The main reason for not applying the arbitrage strategy was insufficient market liquidity, while only less than one third of those who did not apply the arbitrage strategy did not know how to do this. Participants used multiple trading strategies at a time, where the most popular were "buy-and-hold" (62%) and

[3]Please, note that more than strategy was allowed in the questionnaire.

"mean-reverting" (50%).

The main motivation to participate in the voluntary trading experiment was to gain additional credit points (76% of the participants), 53% of the respondents participated to gain trading experience, while 61% found the task interesting. This provides an additional argument that grades can be very motivating to participate in experiments and also that many participants have motivations different from purely monetary to participate in financial experiments.

B.3.5 Experiment 2: Trading Strategies

Depending on the percentage of dividend in their total earnings, we distinguish three types of traders: a) *fundamental traders* whose earnings come more from the dividends than from the cash accumulated from trading, b) *technical traders* whose earnings come from cash more than from the dividends and c) *silent traders* who do not apply any of these strategies and whose ratio of earnings from dividends to earnings from cash equals 1. According to this measure, we identified 63 fundamental traders, 56 technical traders and 17 silent traders.

Technical traders put the largest median number of orders out of the three groups: $Me_{Technical} = 72$, $Me_{Fundamental} = 21$, $Me_{Siletnt} = 0$. According to a Kruskal-Wallis test, these group differences were significant between all types of traders: technical vs. fundamental, $p = .021$, fundamental vs. silent, $p < .001$, technical vs. silent, $p < .001$

Also, the technical traders had the highest grade from the exam (Median grades: $Me_{Technical} = 5.7$, $Me_{Fundamental} = 5.2$, $Me_{Siletnt} = 5.0$). However, this difference was statistically significant only between technical and silent traders, and fundamental and silent ($p < .001$, according to a Kruskal-Wallis test). Further, technical traders had the highest earnings ($Me_{Technical} = 3660.67$, $Me_{Fundamental} = 2366.56$, $Me_{Siletnt} = 600$, where the difference was significant between technical and silent traders, and fundamental and silent traders $p < .001$ according to Kruskal-Wallis test).

From these findings, we derive that the more active traders gained more money and those with the highest predictive skills were the most successful. Better financial knowledge was related to "profiting from the market", but the causality in this relation is not clear.

Appendix C

Appendix to Chapter 4

C.1 Details of the experimental method

C.1.1 Materials and apparatus

The movie describing the professor's lecturing style was based on the two lectures (Lectures 1 and 2) professionally recorded by the university services in Fall 2015 (Experiment 2 of Ref. [107]). The movie included the most characteristic features of Prof. Sornette's lecturing style and a written summary of these features[1].

The selection of the features was based on the notes and observations of two Teaching Assistants (cf. TAs) that were present during the professor's lectures. Based on the notes systematically taken by one of the TAs, the number of the slides, time spent per one slide and time spent for one topic were not related to how many slides the professor would cover. This qualitative and quantitative analysis supported the hypothesis that the lecturing style of the professor is a truly stochastic process with a number of characteristic features.

To define securities in the market, we used the same lecture slides as in Experiment 2 in Ref. [107]. These slides would correspond to lectures 5-7 in the Fall semester 2015. The number of slides were 168, 157 and 144, which corresponded to 57, 54 and 49 securities (the slides were grouped by 3 to define one security). The practice round had 117 slides, which corresponds to 39 securities.[2] The numbers of the 'executed' securities (i.e. which paid

[1] The movie can be viewed under this link: `https://polybox.ethz.ch/index.php/s/jNdUVCXHnz4qu43`. The features of the professor's lecturing style, as displayed in the movie, are listed in C.1.1.

[2] The decks of slides can be downloaded here: `https://polybox.ethz.ch/index.php/s/z1fol6od9IoWX4N`

out 100 monetary units, corresponding to the ending slide of the lecture) in the Practice and three experimental rounds were 15, 15, 23 and 21, respectively. The slide decks were printed in color such that each security would have 3 slides on one sheet and the number of that security would be marked on each slide with large font.

In contrast to Ref. [107], we implemented only three instead of four trading rounds and we reduced the number of slides in Round 1 from 201 to 168, such that the slide deck ends when a certain topic ends. The ending slide was the same as in Ref. [107]. We chose to reduce the number of slides in Round 1 in order to adapt it to the available 10-minute period and the number of available securities in other rounds. Therefore, the final number of slides in Round 1 was similar to those in other rounds. We used the same trading platform as Ref. [107].

C.1.2 Experimental instructions

Trading Competition (2016)

Instructions to the Experiment

Experimenters:

Prof. Didier Sornette, Dr. Sandra Andraszewicz, MSc Ke Wu, Prof. Ryan Murphy, Dr. Dorsa Sanadgol

Please read the instructions and follow the schedule.

If you have any questions, please raise your hand and ask the assistant in the room.

Dear Participant,

In this experiment, you will trade financial assets with all participants in the experimental room. The financial assets represent slides of Professor Didier Sornette and your task will be to predict on which slide he will finish his lecture because only this slide will pay out a dividend.

You will stay anonymous to other traders for the whole duration of the experiment. No persons from outside of the experimental room have rights to trade on this market and all traders participate in this market for the first time.

There will be three trading sessions. After each session, there will be a short break. Your final performance at the end of the experiment will be calculated as cumulative from all three sessions. Before the three sessions, you will participate in a practice session, which will not count to your final compensation.

Your Compensation

Your performance is based on the cumulative earnings of the three trading sessions. The **top 25%** traders with the highest earnings will receive a bonus of **60 Swiss francs**. The **next 25%** will receive a bonus of **30 Swiss francs**. The remaining 50% of the traders in the rank will not receive any bonus. The bonus will be added to your **base payment of 30 Swiss francs**.

How to Earn the Bonus

In each session, you get an **endowment of 300 Experimental Francs (EFR) and 3 units** of each security. This **loan** has to be **repaid** at the end of the session at the **value of 600 EFR**. At the end of the trading session, only one security pays out a **dividend of 100 EFR**. The stock balance at the end of the session does not contribute to your final earnings, only the cash balance and number of dividends matters. If you generate losses, your balance will be turned to 0. Your balance at the end of each session equals:

$$\boldsymbol{max\{cash\ balance + n * 100\ (dividend)\ -\ 600\ (initial\ endowment), 0\}}$$

Therefore, doing nothing will also earn you nothing. To earn the bonus you have to trade intelligently and/or correctly predict, which security will pay the dividend. Every session will start a new market and earnings or losses are not carried over to the next session.

You can make as many trades as you want, as long as you have enough cash to buy shares and you have enough shares to sell. Short-selling is not allowed.

Which Security Pays a Dividend

Securities correspond to slides of Professor Didier Sornette that he presented in his Financial Markets Risks class in Fall 2015. The professor's teaching style is non-typical in a sense that he prepares more slides than needed. **The security that pays out the dividend corresponds to the final lecture slide that Prof. Sornette presents in his lecture. To receive the dividend you have to correctly predict the final slide of the professor’s lecture.** You will receive the stack of slides before the trading session. The final slides of each lecture were recorded in Fall 2015 and they will be announced after every trading session.

Timeline of the Experiment

<u>Step 1 – Professor's Lecturing Style</u>

First, you will see a short movie describing the lecturing style of Prof. Sornette. It will take about 8.5 minutes.

<u>Step 2 – Trading Software Tutorial</u>

Next, you will see a short movie-tutorial on how to use the trading software. It will take about 7.5 minutes.

<u>Step 3 – Practice Session: 5 min trading</u>

In the practice session, you will trade for 5 minutes. Before the trading session starts, you will have 10 minutes to familiarize yourself with the professor's slides and submit your belief. Use the slides that are provided on your desk. This is the time to ask any questions to the experimenters. The experimenters will be present in the room during the practice session. Please, make sure that you understand the software and the procedure before the proceeding to Step 4.

<u>Step 4 – Three Experimental Trading Sessions: 10 min trading</u>

There are three trading sessions that count to your final rank. Each trading session will last 10 minutes and will be proceeded by 10 minutes time to familiarize yourself with the new stack of slides and enter your belief about which slide could be the end slide of the lecture

<u>Step 5 – Debriefing Questionnaire</u>

After the last trading session, you will be automatically re-directed to the website with the questionnaire. Please, enter your trader ID that is provided on your desk and that you used while trading. While you fill out the questionnaire, we will compute the final score on the trading floor.

Practice Session

Each trading session opens at the same time for all participants and will last exactly 5 minutes (10 minutes in the experimental session. The time until the end of the trading will be displayed on your screen.

Login to your account using the login data provided on the table.

At the end of each trading session, the realized number of slides is announced and your account will be credited with your payoff for that session.

In each session, you can monitor your rank.

Before you start trading, you should submit your HONEST and SERIOUS assessment of the probability distribution that the particular security will pay out the dividend.

The Securities

Due to the large number of slides, each security corresponds to three consecutive slides. For your convenience, on each slide, there is a security number and the slides are printed 3 on one page, such that one page corresponds to one security.

There is one "NO-SHOW" security, which would pay 100 EFR in the case where the lecture would be cancelled and the lecture wouldn't happen.

Practice session will start at 14:35.

Good luck!

C.1.3 Summary of the lecturing style of Prof. Sornette, as displayed in the movie

- This movie presents Prof. Didier Sornette's lecturing style. He prepares more slides than needed and he doesn't know how much material he will cover.
- Slides that were prepared for a given lecture but were not presented, are presented in a consecutive lecture.
- You will predict the final slide of the lectures that took place in weeks 5-7 of the semester.
- The slides prepared for week 4 will be used in the practice round.
- In weeks 1-3, the professor covered 45, 35 and 30 slides consecutively which corresponds to 69%, 39% and 54% of the available slides.
- There are a few characteristics of the professor's lecturing style.
- The professor sometimes stops the flow of the lecture to provide a more detailed mathematical derivation of a problem on the blackboard.
- He jumps to a different slide or a topic that either has been shown previously or has not been shown at all.
- Professor Sornette has two lecture ending styles:
- 1) He finishes a topic and ends on the last slide of that topic.
- 2) He finishes a lecture by showing the first slide of the next topic to give an overview on what he will be talking about in the next lecture.
- Some lectures might start a few minutes later due to organisational issues, important announcements or presentation of an assignment.
- Now, you will see samples of the professor's lecturing style, based on material recording during two consecutive lectures in the Fall 2015.

C.2 Summary of the self-reported measures

17 participants claimed to have applied a buy-and-hold strategy, 14 classified themselves as using a mean-reverting strategy, 9 reported as trend followers, and 7 people reported

“other” strategies. The “other” category included buying cheaply and inflating the prices of the purchased securities to sell at a higher price, buying cheaply and trying to sell expensive and selling securities that were unlikely to pay out the dividend.

The majority of the participants (N = 21, 58% of the participants) used the number of slides as the cue for estimating the end slide, while the second most frequent cue was their initial probability estimate (N = 18, 50% of the participants). The participants used the bid and ask prices of other traders and the number of topics covered by the professor equally likely (N = 15 and 14 consecutively). In the field experiment, substantially more participants would anchor their prediction on the number of the slides covered in the previous rounds but also only twelve participants claimed to be using only one strategy, compared to seven participants (20%) in the laboratory.

Bibliography

[1] Erik Snowberg, Justin Wolfers, and Eric Zitzewitz. Information (in)efficiency in prediction markets. In *Information Efficiency in Financial and Betting Markets*, pages 366–386. Cambridge University Press, 2005.

[2] Terrill R Keasler and Delbert C Goff. Using Fed Funds Futures to Predict a Federal Reserve Rate Hike. *Journal of Economics and Finance*, 6(2):9–14, 2007.

[3] Øyvind Eitrheim, Espen Frøyland, and Øistein Røisland. Can the price of currency options provide an indication of market perceptions of the uncertainty attached to the krone exchange rate? *Economic Bulletin*, pages 266–277, 1999.

[4] Jens Carsten Jackwerth and Mark Rubinstein. Recovering Probability Distributions from Option Prices. *The Journal of Finance*, 51(5):1611–1631, 1996.

[5] Justin Wolfers and Eric Zitzewitz. Prediction Markets. *Journal of Economic Perspectives*, 18(2):107–126, 2004.

[6] Joyce E. Berg and Thomas A. Rietz. Prediction markets as decision support systems. *Information Systems Frontiers*, 5:79–93, 2003.

[7] Charles F. Manski. Interpreting the predictions of prediction markets. *Economics Letters*, 91:425–429, 2006.

[8] Joyce Berg, Robert Forsythe, Forrest Nelson, and Thomas Rietz. Results from a Dozen Years of Election Futures Markets Research. *Handbook of Experimental Economics Results*, 1(C):742–751, 2008.

[9] Robert Forsythe, Forrest Nelson, George R. Neumann, and Jack Wright. Anatomy of an experimental political stock market. *The American Economic Review*, 82:1142–1161, 1992.

[10] Robert Forsythe, Thomas A. Rietz, and Thomas W. Ross. Wishes, expectations and actions: a survey on price formation in election stock markets. *Journal of Economic Behavior and Organization*, 39:83–110, 1999.

[11] Jan Hansen, Carsten Schmidt, and Martin Strobel. Manipulation in political stock markets - preconditions and evidence. *Applied Economics Letters*, 11:459–463, 2004.

[12] Philip M. Polgreen, Forrest D. Nelson, and George R. Neumann. Use of prediction markets to forecast infections disease activity. *Clinical Infectious Diseases*, 44:272–279, 2007.

[13] Chen-Yuan Tung, Tzu-Chuan Chou, and Jih-Wen Lin. Using prediction markets of market scoring rule to forecast infectious diseases: a case study in Taiwan. *BMC Public Health*, 15:766–777, 2015.

[14] Kyle J. Kain and Trevor D. Logan. Are sports betting markets prediction markets?: Evidence from a new test. *Journal of Sports Economics*, 1:45–63, 2014.

[15] Bo Cowgill, Justin Wolfers, and Eric Zitzewitz. Using prediction markets to track information flows: Evidence from Google. In *Auctions, Market Mechanisms and Their Applications*, page 3, 2009.

[16] Anita Elberse and Jehoshua Eliashberg. Demand and supply dynamics for sequentially released products in international markets: The case of motion pictures. *Marketing Science*, 22:329–354, 2003.

[17] Teck H. Ho and Kay-Yut Chen. New product blockbusters: The magic and science of prediction markets. *California Management Review*, 50:144–158, 2007.

[18] Cary Deck and David Porter. Prediction markets in the laboratory. *Journal of Economic Surveys*, 27(3):589 – 603, 2013.

[19] Paul A Samuelson. Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6(2):41–48, 1965.

[20] Eugene F Fama. Efficient Capital Markets : A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2):383–417, 1970.

[21] Eugene F Fama. Efficient Capital Markets : II. *The Journal of Finance*, 46(5):1575–1617, 1991.

[22] Burton G Malkiel. The Efficient Market Hypothesis and Its Critics. *Journal of Economic Perspectives*, 17(1):59–82, 2003.

[23] Burton G Malkiel. Reflections on the Efficient Market Hypothesis: 30 Years Later. *The Financial Review*, 40:1–9, 2005.

[24] Stephen J Brown and Jerold B Warner. Using daily stock returns: The case of event studies. *Journal of financial economics*, 14(1):3–31, 1985.

[25] A Craig MacKinlay. Event studies in economics and finance. *Journal of economic literature*, 35(1):13–39, 1997.

[26] Victor L. Bernard and Jacob K. Thomas. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting and Economics*, 13(4):305–340, 1990.

[27] Mark L. Mitchell and J. Harold Mulherin. The Impact of Public Information on the Stock Market. *Journal of Finance*, 49(3):923, 1994.

[28] Eugene F. Fama. Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics*, 49(3):283–306, 1998.

[29] Leighton Vaughan Williams. Semi-strong and strong form information efficiency in betting markets. In *Information Efficiency in Financial and Betting Markets*, pages 123–156. Cambridge University Press, 2005.

[30] Camelia Oprean. Testing the financial market informational efficiency in emerging states. *Review of Applied Socio- Economic Research*, 4(2):181–190, 2012.

[31] Keith Cuthbertson and Dirk Nitzsche. *Quantitative financial economics: stocks, bonds and foreign exchange*. John Wiley & Sons, 2005.

[32] J Doyne Farmer and Andrew W Lo. Frontiers of finance: Evolution and efficient markets. *Proceedings of the National Academy of Sciences*, 96(18):9991–9992, 1999.

[33] John Y Campbell, Andrew Wen-Chuan Lo, and Archie Craig MacKinlay. *The econometrics of financial markets*. princeton University press, 1997.

[34] Chris Hanretty. The EU Referendum: what to expect on the night. `https://medium.com/@chrishanretty/the-eu-referendum-what-to-expect-on-the-night-521792dd3eef`, jun 2016. Jun 7.

[35] Nate Silver. A User's Guide To FiveThirtyEight's 2016 General Election Forecast. `https://fivethirtyeight.com/features/a-users-guide-to-fivethirtyeights-2016-general-election-forecast/`, 2016. Jun 29.

[36] Raymond S Nickerson. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, 2:175, 1998.

[37] Robert J. Shiller. Speculative asset prices. *American Economic Review*, 104(6):1486–1517, 2014.

[38] Didier Sornette and Peter. Cauwels. Financial bubbles: mechanisms and diagnostics. *Review of Behavioral Economics*, 2(3):279–305, 2015.

[39] Vernon L Smith, Gerry L Suchanek, and Arlington W Williams. Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica*, 65:1119–1151, 1988.

[40] Stefan Palan. A review of bubbles and crashes in experimental asset markets. *Journal of Economic Surveys*, 27:570–588, 2013.

[41] Owen Powell. Numeraire independence and the measurement of mispricing in experimental asset markets. *Journal of Behavioral and Experimenatl Finance*, 9:56–62, 2016.

[42] V. L. Smith, M. V. van Boening, and C. P. Wellford. Dividend timing and behavior in laboratory asset markets. *Economic Theory*, 16:567–583, 2000.

[43] M. Kirchler, J. Huber, and T. Stöckl. Thar she bursts, reducing confusion reduces bubbles. *American Economic Review*, 102:865–883, 2012.

[44] Owen Powell and Natalia Shestakova. Experimental asset markets: A survey of recent developments. *Journal of Behavioral and Experimenatl Finance*, 12:14–22, 2016.

[45] G Loewenstein. Experimental economics from the vantage-point of behavioural economics. *The Economic Journal*, 1999.

[46] Francesco Guala and Luigi Mittoe. Experiments in economics: External validity and the robustness of phenomena. *Journal of Economic Methodology*, 12:495–515, 2005.

[47] Arthur Schram. Artificiality: The tension between internal and external validity in economic experiments. *Journal of Economic Methodology*, 12:225–237, 2005.

[48] Frank H. Knight. *Risk, uncertainty and profit.* Houghton Mifflin Company, Boston, New York, 1921.

[49] Simone Nuzzo and Andrea Morone. Asset markets in the lab: A literature overview. *Journal of Behavioural and Experimental Finance*, In press, 2017.

[50] Charles R. Plott. Industrial organization theory and experimental economics. *Journal of Economic Literature*, 20:1485 – 1527, 1982.

[51] Glenn W. Harrison and John A. List. Field experiments. *Journal of Economic Literature*, XLII:1009 – 1055, 2004.

[52] J. O. Ledyard. *Handbook of Experimental Economics*, chapter Public goods: A survey of experimental research, pages 111 – 194. Princeton University Press, 1995.

[53] B. A. Nosek, G. Alter, G. C. Banks, D. Borsboom, S. D. Bowman, S. J. Breckler, S. Buck, C. D. Chambers, G. Chin, G. Christensen, M. Contestabile, A. Dafoe, E. Eich, J. Freese, R. Glennerster, D. Goroff, D. P. Green, B. Hesse, M. Humphreys, J. Ishiyama, D. Karlan, A. Kraut, A. Lupia, P. Mabry, T. Madon, N. Malhotra, E. Mayo-Wilson, M. McNutt, E. Miguel, E. Levy Paluck, U. Simonsohn, C. Soderberg, B. A. Spellman, J. Turitto, G. VandenBos, S. Vazire, E. J. Wagenmakers, R. Wilson, and T. Yarkoni. Promoting and open research culture. *Science*, 348:1422 – 1425, 2015.

[54] Colin F. Camerer, Anna Dreber, Eskil Forsell, Teck-Hua Ho, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Johan Almenberg, Adam Altmejd, Taizan Chan, Emma Heikensten, Felix Holzmeister, Taisuke Imai, Siri Isaksson, Gideon Nave, Thomas Pfeiffer, Michael Razen, and Hang Wu. Evaluating replicability of laboratory experiments in economics. *Science*, 351(6280):1433–1436, 2016.

[55] Dniel Chen, Martin Schonger, and Chris Wickens. otree - an open-source platform for laboratory, online, and field experiments. *Journal of Behavioural and Experimental Finance*, 9:88 – 97, 2016.

[56] John A. List. Do explicit warnings eliminate the hypothetical bias in elicitation procedures? evidence from field auctions for sportcards. *The Merican Economic Review*, 91(5):1498 – 1507, 2001.

[57] Ronald G. Cummings, Glenn W. Harrison, and E. Elisabet Rutström. Homegrown values and hypothetical surveys: Is the dichotomous choice approach incentive compatible? *American Economic Review*, 95(1):260 – 266, 1995.

[58] Ronald G. Cummings and Laura O. Taylor. Unbiased value estimates for environmental goods: A cheap talk design for the contingent valuation method. *American Economic Review*, 89(3):649 – 665, 1999.

[59] Matthias Benz and Stephan Meier. Do people behave in experiments as in the field? – evidence from domains. *Journal of Experimental Economics*, 11:268 – 281, 2008.

[60] Steven Levitt, John A. List, and David Railey. What happens in the field stays in the field: Professionals do not play minimax in laboratory experiments. *Econometrica*, 78(4):1413 – 1434, 2010.

[61] Steven Levitt and John A. List. Viewpoint: On the generalizability of lab behaviouur to the field. *Canadain Journal of Economics*, 40(2):347 – 370, 2007.

[62] Steven Levitt and John A. List. What do laboratory experiments measuring social preferences reveal about the real world. *Journal of Economic Perspectives*, 21(2):153 – 174, 2007.

[63] Steven Levitt and John A. List. Homo oeconomicus evolves. *Science*, 319(5865):909 – 910, 2008.

[64] Colin F. Camerer. The promise and success of lab-field generalizability in experimental economics: A ritical reply to Levitt and List*. In *Handbook of Experimental Economic Methodology*, pages 1 – 64.

[65] Jürgen Huber and Michael Kirchler. The impact of instructions and procedure on reducing confusion and bubbles in experimental asset markets. *Experimental Economics*, 15:89 – 105, 2011.

[66] Erich Kirchler, Boris Maciejovsky, and Martin Weber. Framing effetcs, selective information, and market behavior: An experimental analysis. *Journal of Behavioral Finance*, 6(3):90 – 100, 2005.

[67] Matthias Stefan. On the impact of semantic framing in experimental asset markets. *Journal of Behavioral and Experiemntal Finance*, 9:81 – 87, 2016.

[68] Amos Tversky and Daniel Kahneman. The framing of decisoins and the psychology of the choice. *Scinece*, 4481:453 – 458, 1981.

[69] Mark W. Nelson, Robert Bloomfield, Jeffrey W. Hales, and Robert Libby. The effetc of information strength and weight on behavior in financial markets. *Organizational Behaviour and Human Decision Process*, 86:168 – 196, 2001.

[70] Jürgen Huber, Michale Kirchler, and Matthias Sutter. Is moreinformation always better? Experimental financial markets with cumulative information. *Journal of Economic Bheaviour and Organization*, 65:86 – 104, 2008.

[71] Matthias Sutter, Jürgen Huber, and Mchael Kirchler. Bubbles and informatoion: An experiment. *Management Science*, 58(2):384 – 393, 2012.

[72] Larry G. Epstein and Martin Schneider. Ambiguity, information quality and asset pricing. *Journal of Finance*, 63:197 – 228, 2008.

[73] Colin Camerer. Can asset markets be manipulated? A field experiment with racetrack betting. *Journal of Political Economy*, 106(3):457 – 482, 1998.

[74] Robert Bloomfield, Robert Libby, and Mark W. Nelson. Confidence and welfare of less-informed investors. *Accounting, Organizations and Society*, 24:623 – 647, 1999.

[75] Werner Grüth, Jan P. Krahnen, and Christian Rieck. Financial markets with asymmetric information: A pilot study focusing on insider advantages. *Journal of Economic Psychology*, 18:235 – 257, 1997.

[76] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. 2008.

[77] Maiya Keidan and Jemima Kelly. Number of crypto hedge funds surges amid bitcoin volatility. `https://www.reuters.com/article/uk-hedgefunds-bitcoin/number-of-crypto-hedge-funds-surges-amid-bitcoin-volatility-idUSKCN1FZ189`, 2018. Feb 15.

[78] Carey Caginalp and Gunduz Caginalp. Opinion: Valuation, liquidity price, and stability of cryptocurrencies. *Proceedings of the National Academy of Sciences*, 115(6):1131–1134, 2018.

[79] Dean Fantazzini, Erik Nigmatullin, Vera Sukhanovskaya, and Sergey Ivliev. Everything You Always Wanted to Know About Bitcoin Modelling But Were Afraid to Ask. *Applied Econometrics, Forthcoming.*, 2016. Available at SSRN: https://ssrn.com/abstract=2794622.

[80] Robleh Ali, John Barrdear, Roger Clews, and James Southgate. The economics of digital currencies. 2014. Available at SSRN: https://ssrn.com/abstract=2499418.

[81] Paul Vigna and Casey Michael J. *The age of cryptocurrency: How bitcoin and digital money are challenging the global economic order*. St. Martin's Press, 2015.

[82] Jonathan Chiu and Thorsten V Koeppl. The economics of cryptocurrencies–bitcoin and beyond. 2017. Available at SSRN: https://ssrn.com/abstract=3048124.

[83] Pavel Ciaian, Miroslava Rajcaniova, and d'Artis Kancs. The economics of bitcoin price formation. *Applied Economics*, 48(19):1799–1815, 2016.

[84] Ernie GS Teo. Emergence, growth, and sustainability of bitcoin: The network economics perspective. In *Handbook of Digital Currency*, pages 191–200. Elsevier, 2015.

[85] Dániel Kondor, Márton Pósfai, István Csabai, and Gábor Vattay. Do the rich get richer? an empirical analysis of the bitcoin transaction network. *PloS one*, 9(2):e86197, 2014.

[86] Micha Ober, Stefan Katzenbeisser, and Kay Hamacher. Structure and anonymity of the bitcoin transaction graph. *Future internet*, 5(2):237–250, 2013.

[87] Paolo Tasca, Adam Hayes, and Shaowen Liu. The evolution of the bitcoin economy: extracting and analyzing the network of payment relationships. *The Journal of Risk Finance*, pages https://doi.org/10.1108/JRF–03–2017–0059, 2016.

[88] Ladislav Kristoufek. Bitcoin meets google trends and wikipedia: Quantifying the relationship between phenomena of the internet era. *Scientific reports*, 3:3415, 2013.

[89] David Garcia, Claudio J Tessone, Pavlin Mavrodiev, and Nicolas Perony. The digital traces of bubbles: feedback cycles between socio-economic signals in the bitcoin economy. *Journal of the Royal Society Interface*, 11(99):20140623, 2014.

[90] David Garcia and Frank Schweitzer. Social signals and algorithmic trading of bitcoin. *Open Science*, 2(9), 2015. DOI: https://doi.org/10.1098/rsos.150288.

[91] Florian Glaser, Kai Zimmermann, Martin Haferkorn, Moritz Weber, and M Siering. Bitcoin - Asset or currency? Revealing users' hidden intentions. *ECIS 2014 Proceedings - 22nd European Conference on Information Systems*, 01 2014.

[92] Young Bin Kim, Jurim Lee, Nuri Park, Jaegul Choo, Jong Hyun Kim, and Chang Hun Kim. When Bitcoin encounters information in an online forum: Using text mining to analyse user opinions and predict value fluctuation. *PLoS ONE*, 12(5):1–14, 2017.

[93] Svetlana Sapuric and Angelika Kokkinaki. Bitcoin is volatile! isn't that right? In *International Conference on Business Information Systems*, pages 255–265. Springer, 2014.

[94] Lukáš Pichl and Taisei Kaizoji. Volatility Analysis of Bitcoin Price Time Series. *Quantitative Finance and Economics*, 1(4):474–485, 2017.

[95] Andrew Urquhart and Hanxiong Zhang. Is Bitcoin a Hedge or Safe-Haven for Currencies? An Intraday Analysis. pages 1–25, 2018. Available at SSRN: https://ssrn.com/abstract=3114108.

[96] Elie Bouri, Peter Molnár, Georges Azzi, David Roubaud, and Lars Ivar Hagfors. On the hedge and safe haven properties of Bitcoin: Is it really more than a diversifier? *Finance Research Letters*, 20:192–198, 2017.

[97] Aurelio F. Bariviera. The inefficiency of Bitcoin revisited: A dynamic approach. *Economics Letters*, 161:1–4, 2017.

[98] Joerg Osterrieder and Julian Lorenz. a Statistical Risk Assessment of Bitcoin and Its Extreme Tail Behavior. *Annals of Financial Economics*, 12(01):1750003, 2017.

[99] Jeffrey Chu, Saralees Nadarajah, and Stephen Chan. Statistical analysis of the exchange rate of bitcoin. *PloS one*, 10(7):e0133678, 2015.

[100] Ladislav Kristoufek. What are the main drivers of the bitcoin price? evidence from wavelet coherence analysis. *PloS one*, 10(4):e0123923, 2015.

[101] Anne Haubo Dyhrberg. Bitcoin, gold and the dollar–a GARCH volatility analysis. *Finance Research Letters*, 16:85–92, 2016.

[102] Stefan Bornholdt and Kim Sneppen. Do bitcoins make the world go round? on the dynamics of competing crypto-currencies. *arXiv preprint arXiv:1403.6378*, 2014.

[103] Garrick Hileman and Michel Rauchs. Global cryptocurrency benchmarking study. *Cambridge Centre for Alternative Finance*, 2017.

[104] Abeer ElBahrawy, Laura Alessandretti, Anne Kandler, Romualdo Pastor-Satorras, and Andrea Baronchelli. Evolutionary dynamics of the cryptocurrency market. *Royal Society Open Science*, 4(11):170623, 2017.

[105] Didier Sornette. *Why stock markets crash: critical events in complex financial systems*. Princeton University Press, 2003.

[106] Ke Wu, Spencer Wheatley, and Didier Sornette. The british pound on brexit night: a natural experiment of market efficiency and real-time predictability. 2017.

[107] Didier Sornette, Sandra Andraszewicz, Ke Wu, Ryan O. Murphy, Philipp Rindler, and Dorsa Sanagdol. Mispricing robustness in experimental asset arkets with persistent uncertainty. *PloS one*, under review.

[108] Sandra Andraszewicz, Ke Wu, and Didier Sornette. Behavioural effects and market dynamics in field and laboratory experimental asset markets. *PloS one*, under review.

[109] Ke Wu, Spencer Wheatley, and Didier Sornette. Classification of crypto-coins and tokens from the dynamics of their power law capitalisation distributions. *arXiv preprint arXiv:1803.03088*, 2018.

[110] Ryan Kennedy, Stefan Wojcik, and David Lazer. Improving election prediction internationally. *Science*, 355(6324):515–520, 2017.

[111] Kenneth J Arrow, Robert Forsythe, Michael Gorham, Robert Hahn, Robin Hanson, John O Ledyard, Saul Levmore, Robert Litan, Paul Milgrom, Forrest D Nelson, George R Neumann, Marco Ottaviani, Thomas C Schelling, Robert J Shiller, Vernon L Smith, Erik Snowberg, Cass R Sunstein, Paul C Tetlock, Philip E Tetlock, Hal R Varian, Justin Wolfers, and Eric Zitzewitz. The promise of prediction markets. *Science*, 320(5878):877–878, 2008.

[112] Janis Irving. *Victims of groupthink: a psychological study of foreign-policy decisions and fiascoes.* Houghton Mifflin, Boston, 1972.

[113] Paul Hart. *Groupthink in government: A study of small groups and policy failure.* Swets & Zeitlinger Publishers, 1990.

[114] Philip E Tetlock, Randall S Peterson, Charles McGuire, Shi-jie Chang, and Peter Feld. Assessing political group dynamics: a test of the groupthink model. *Journal of Personality and Social Psychology*, 63(3):403–425, 1992.

[115] James K. Esser. Alive and well after 25 years: a review of groupthink research. *Organizational Behavior and Human Decision Processes*, 73(2-3):116–141, 1998.

[116] Paul Kowert. *Groupthink or deadlock: when do leaders learn from their advisors?* SUNY Press, 2012.

[117] Robert S. Baron. So right it's wrong: groupthink and the ubiquitous nature of polarized group decision making. *Advances in Experimental Social Psychology*, 37:219–253, 2005.

[118] James D Rose. Diverse perspectives on the groupthink theory–a literary review. *Emerging Leadership Journeys*, 4(1):37–57, 2011.

[119] John DiNardo. Natural experiments and quasi-natural experiments. In *Microeconometrics*, pages 139–153. Springer, 2010.

[120] Torben G Andersen, Tim Bollerslev, Francis X Diebold, and Clara Vega. Micro effects of macro announcements: Real-time price discovery in foreign exchange. *The American economic review*, 93(1):38–62, 2003.

[121] Gikas A Hardouvelis. Economic news, exchange rates and interest rates. *Journal of International Money and Finance*, 7(1):23–35, 1988.

[122] Iain Clark and Saeed Amen. Implied Distributions from GBPUSD Risk-Reversals and Implication for Brexit Scenarios. 2016. Available at SSRN: https://ssrn.com/abstract=2794888.

[123] Gary Howes. Pound to Suffer 15-20% Slump on Brexit, Surge to 1.60 if Brexit Avoided. `https://www.poundsterlinglive.com/gbp-live-today/3686-hsbc-forecast-for-the-pound-on-brexit`, 2016. Feb 12.

[124] Will Peters. Forecasts 2016: JPMorgan See GBP/USD Bottoming at 1.32. `https://www.poundsterlinglive.com/exchange-rate-forecasts/3404-jpmorgan-chase-exchange-rate-forecast-summary`, 2016. Apr 25.

[125] Rudiger Dornbusch. Exchange Rate Economics: Where Do We Stand? *Brookings Papers on Economic Activity*, 11(1, Tenth Anniversary Issue):143–206, 1980.

[126] Louis H Ederington and Jae Ha Lee. How markets process information: News releases and volatility. *The Journal of Finance*, 48(4):1161–1191, 1993.

[127] Andrei Shleifer and Robert W Vishny. The limits of arbitrage. *The Journal of Finance*, 52(1):35–55, 1997.

[128] Silvio Cascione. Mexico peso seen plunging 10 percent if Trump wins U.S. election - Reuters poll. `https://www.reuters.com/article/us-forex-poll-latam/mexico-peso-seen-plunging-10-percent-if-trump-wins-u-s-election-reuters-poll-idUSKBN12W4DN`, 2016. Nov 1.

[129] Nate Cohn. Exit Polls: Why They So Often Mislead. `https://www.nytimes.com/2014/11/05/upshot/exit-polls-why-they-so-often-mislead.html?_r=0`, 2014. Nov 4.

[130] David Bell. Brexit betting odds: lesson from Scotland is not promising for Leave. `https://theconversation.com/brexit-betting-odds-lesson-from-scotland-is-not-promising-for-leave-60494`, 2016. Jun 6.

[131] YouGov. YouGov on the day poll: Remain 52%, Leave 48%. `https://yougov.co.uk/news/2016/06/23/yougov-day-poll/`, 2016. Jun 23.

[132] YouGov. YouGov's eve-of-vote poll: Remain leads by two. `https://yougov.co.uk/news/2016/06/22/final-eve-poll-poll/`, 2016. Jun 22.

[133] FiveThirtyEight. https://projects.fivethirtyeight.com/2016-election-forecast/.

[134] David M Pennock, Steve Lawrence, C Lee Giles, Finn Arup Nielsen, et al. The real power of artificial markets. *Science (New York, N.Y.)*, 291:987–988, 2001.

[135] Joyce E. Berg, Forrest D. Nelson, and Thomas A. Rietz. Prediction market accuracy in the long run. *International Journal of Forecasting*, 24(2):283–298, 2008.

[136] Joyce E Berg and Thomas a. Rietz. Prediction Markets As Decision Support Systems. *Information Systems Frontiers*, 5(1):79–93, 2003.

[137] Saul Levmore. Simply Efficient Markets and the Role of Regulation: Lessons from the Iowa Electronic Markets and the Hollywood Stock Exchange. *Journal of Corporation Law*, 28(4):589–606, 2003.

[138] Yiling Chen, Chao-Hsien Chu, Tracy Mullen, and David M. Pennock. Information markets vs. opinion pools. *Proceedings of the 6th ACM conference on Electronic commerce*, pages 58–67, 2005.

[139] Earl S Rosenbloom and William Notz. Statistical Tests of Real-Money versus Play-Money Prediction Markets. *Electronic Markets*, 16(1):63–96, 2006.

[140] Bo Cowgill and Eric Zitzewitz. Corporate prediction markets: Evidence from google, ford, and firm x. *The Review of Economic Studies*, 82(4):1309–1341, 2015.

[141] Lyric Doshi, Jonas Krauss, Stefan Nann, Peter Gloor, Anita Elberse, Bharat N Anand, Ekaterina V. Karniouchina, Saul Levmore, Koleman Strumpf, Anita Elberse, Jehoshua Eliashberg, and Ekaterina V. Karniouchina. Are Virtual Markets Efficient Predictors of New Product Success? The Case of the Hollywood Stock Exchange. *The Journal of Product Innovation Management*, 28(4):470–484, 2011.

[142] Betfair. http://www.betfair.com.

[143] Max Liu. NoBrexit: Leave and Remain make final pitches to voters. https://betting.betfair.com/politics/brexit/eu-referendum-betting-latest-polling-and-odds-june-22-2016-220616-204.html, jun 2016. Jun 22.

[144] Max Liu. Brexit: Bettors back high turnout and victory for Remain. https://betting.betfair.com/politics/brexit/eu-referendum-betting-latest-polling-and-odds-june-23-2016-230616-204.html, jun 2016. Jun 23.

[145] Naomi Totten. US President Betting: £200m traded as market smashes Betfair record. https://betting.betfair.com/politics/us-politics/200m-us-president-market-breaks-all-time-betfair-record-091116-745.html, Nov 2016. Nov 9.

[146] PredictWise. http://predictwise.com/politics/2016-president-winner.

[147] Christopher Langner. Brexit's Probability Gambit. https://www.bloomberg.com/gadfly/articles/2016-06-23/brexit-risks-skewed-by-gamblers-pray-they-re-right, jun 2016. Jun 22.

[148] Michelle Davis and Netty I.l Ismail. Mexico Peso Trades in Line With Peers After Presidential Debate. https://www.bloomberg.com/news/articles/2016-10-20/mexico-peso-is-little-changed-as-u-s-presidential-debate-starts, 2016. Oct 20.

[149] Phil Kuntz. U.S. Election Guide to Markets: What to Watch as Results Come In. https://www.bloomberg.com/news/articles/2016-10-30/u-s-election-guide-to-markets-what-to-watch-as-results-come-in, 2016. Oct 31.

[150] Chuck Mikolajczak. Stocks, Mexican peso climb ahead of U.S. election results. http://www.reuters.com/article/us-global-markets/stocks-mexican-peso-climb-ahead-of-u-s-election-results-idUSKBN133030?il=0, 2016. Nov 8.

[151] Harry Carr. How Eurosceptic Is Your Home Town? http://news.sky.com/story/how-eurosceptic-is-your-home-town-10206150, 2016. May 15.

[152] Will Dahlgreen. The Eurosceptic map of Britain. https://yougov.co.uk/news/2016/02/28/eurosceptic-map-britain/, 2016. Feb 28.

[153] Matt Singh. The 2.8 Million Non-Voters Who Delivered Brexit. https://www.bloomberg.com/view/articles/2016-07-04/the-2-8-million-non-voters-who-delivered-brexit, jul 2016. Jul 4.

[154] David P. Myatt. A theory of protest voting. *London Business School, August*, 2015.

[155] David P. Myatt. A theory of voter turnout. *London Business School, September*, 2015.

[156] histdata. http://www.histdata.com.

[157] UK Electoral Commission. Provisional electorate figures published ahead of the EU referendum. http://www.electoralcommission.org.uk/i-am-a/journalist/electoral-commission-media-centre/news-releases-referendums/provisional-electorate-figures-published-ahead-of-the-eu-referendum, 2016. Jun 24.

[158] CNN Election Night TV Program. `https://youtu.be/w31B49W03io`.

[159] Stephanie Flanders. Brexit: What investors should consider. *J.P. Morgan Asset management*, 2016. Feb.

[160] Sebastien Pouget, Julien Sauvagnat, and Stephane Villeneuve. A mind is a terrible thing to change: confirmatory bias in financial markets. *The Review of Financial Studies*, 30(6):2066–2109, 2017.

[161] Jacob B Hirsh, Colin G DeYoung, Xiaowen Xu, and Jordan B Peterson. Compassionate liberals and polite conservatives: associations of agreeableness with political ideology and moral values. *Personality and social psychology bulletin*, 36(5):655–664, 2010.

[162] Joachim I Krueger. The Personality of Brexit Voters. `https://www.psychologytoday.com/blog/one-among-many/201606/the-personality-brexit-voters`, 2016. Jun 29.

[163] Scott Barry Kaufman. The Personality of Political Correctness. `https://blogs.scientificamerican.com/beautiful-minds/the-personality-of-political-correctness/`, 2016. Nov 20.

[164] How our Brexit model fared on election night. `http://www.economist.com/blogs/graphicdetail/2016/06/daily-chart-18`, 2016. Jun 27.

[165] Kosha Gada. What Social Media Told Us About BREXIT That Polls And Betting Markets Missed. `https://www.forbes.com/sites/koshagada/2016/06/24/the-united-kingdom-makes-history-why-social-media-like-bookmakers-and-investors-got-brexit-wrong`, 2016. Jun 24.

[166] Nature Editorial. Better together: The European Union has its issues, but a Brexit could spell problems for science. *Nature*, 530(7588):6–6, 2016.

[167] Vyacheslav Polonski. Social Media Voices in the UK's EU Referendum. `https://medium.com/@slavacm/social-media-voices-in-the-uks-eu-referendum-brexit-or-bremain-what-does-the-internet-say-about-ebbd7b27cf0f`, 2016. May 5.

[168] Paul Wood. The British data-crunchers who say they helped Donald Trump to win. `http://www.spectator.co.uk/2016/12/the-british-data-crunchers-who-say-they-helped-donald-trump-to-win/`, 2016. Dec 3.

[169] Didier Sornette. Dragon-Kings, Black Swans and the Prediction of Crises. *International Journal of Terraspace Science and Engineering*, 2(1):1–18, 2009.

[170] Didier Sornette and G. Ouillon. Dragon-kings: Mechanisms, statistical methods and empirical evidence. *The European Physical Journal Special Topics*, 205(1):1–26, 2012.

[171] Nate Silver. How I Acted Like A Pundit And Screwed Up On Donald Trump. `https://fivethirtyeight.com/features/how-i-acted-like-a-pundit-and-screwed-up-on-donald-trump/`, 2016. May 18.

[172] A. Orléan. Bayesian interactions and collective dynamics of opinion: Herd behavior and mimetic contagion. *Journal of Economic Behavior and Organization*, 28:257–274, 1995.

[173] F. Black. Noise. *The Journal of Finance*, 41:529–543, 1986.

[174] R. S. Gurkaynak. Econometric tests of asset price bubbles: taking stock. *Journal of Economic Survey*, 22:166–186, 2008.

[175] R. Hertwig and I. Erev. The description-experience gap in risky choice. *Trends in Cognitive Sciences*, 13:517–523, 2009.

[176] Lana Friesen and Lata Gangadharan. Enviromental markets: Wwhat do we learn from the lab? *Journal of Economic Surveys*, 27(3):515 – 535, 2013.

[177] Charles N. Noussair and Steven Tucker. Experimental research on asset pricing. In *Discussion Paper Center*, pages 1 – 26. Tilburg University, 2013.

[178] AJ A. Bostian and Charles A. Holt. Price bubbles with discounting: A web-based classroom experiment. *Journal of Economic Education*, 40(1):27 – 37, 2009.

[179] Charles A. Holt, Megan Porzio, and Michelle Yingze Song. Price bubbles, gender, and expectations in experimental asset markets. *European Economic Review*, 100:72 – 94, 2017.

[180] Alec Smith, Terry Lohrenz, Justin King, P. Read Montague, and Colin F. Camerer. Irrational exuberance and neural crasg warning signals during endogenous experiemntal asset market bubbles. *Proceedings in National Academy of Sciences*, 111(29):10503 – 10508, 2014.

[181] Charles R. Plott and Shyam Sunder. Rational expectations and the aggregation of diverse information in laboratory security markets. *Econometrica*, (5):1085 – 1118, 1988.

[182] A. E. Mannes. Are we wise about the wisdom of crowds? the use of group judgments in belief revision. *Management Science*, 55:1267–1279, 2009.

[183] R. Ray. Prediction markets and the financial "wisdom of crowds". *Journal of Behavioral Finance*, 7:2–4, 2006.

[184] J. Surowiecki. *The Wisdom of Crowds*. Anchor Books, 2005.

[185] Marc Chesney, Remo Crameri, and Loriano Mancini. Detecting abnormal trading activities in option markets. *Swiss Finance Institute Research Paper No. 11-42*, 2015. Available at SSRN: http://ssrn.com/abstract=1522157.

[186] G. Caginalp, D. P. Porter, and V. L. Smith. Financial bubbles: excess cash, momentum, and incomplete information. *Journal of Psychology and Financial Markets*, 2:80–99, 2001.

[187] V. Lei, C. Noussair, and C. R. Plott. Non-speculative bubbles in experimental asset markets: Lack of common knowledge of rationality vs. actual irrationality. *Econometrica*, 69:831–859, 2001.

[188] J. Oechssler, C. Schmidt, and W. Schnedler. On the ingredients for bubble formation: Informed traders and communication. *Journal of Economic Dynamics and Control*, 35:1831–18951, 2011.

[189] L. F. Ackert, N. Charupat, B. K. Church, and R. Deaves. Margin, short selling, and lotteries in experimental asset markets. *Southerns Economic Journal*, 73:419–436, 2006.

[190] L. F. Ackert, N. Charupat, R. Deaves, and B. D. Kluger. Probability judgment error and speculation in laboratory asset market bubbles. *Journal of Financial and Quantitative Analysis*, 3:719–744, 2009.

[191] G. Caginalp, V. Ilieva, D. P. Porter, and V. L. Smith. Do speculative stocks lower prices and increase volatility of value stocks? *Journal of Psychology and Financial Markets*, 3:118–132, 2002.

[192] K. S. Chan, V. Lei, and F. Veseley. Differentiated assets: An experimental study on bubbles. *Economic Enquiry*, 51:1731–1749, 2013.

[193] E. O. Fisher and F. S. Kelly. Experimental foreign exchange markets. *Pacific Economic Review*, 5:365–387, 2000.

[194] R. R. King, V. L. Smith, A. W. Williams, and M. V. van Boening. *Nonlinear Dynamics and Evolutionary Economics*, chapter The robustness of bubbles and crashes in experimental stock markets, pages 183–200. Oxford University Press, New York, 1993.

[195] S. Gore. Biostatistics and the medical research council. *Medical Reseacrh Council News*, 35:19–20, 1987.

[196] S. R. Johnson, G. A. Tomlinson, G. A. Hawker, J.T. Granton, H. A. Grosbein, and B.M. Feldman. A valid and reliable belief elicitation method for Bayesian priors. *Journal of Clinical Epidemiology*, 63:370–383, 2010.

[197] E. Morris, David, E. Oakley, Jeremy, and A. Crowe, John. A web-based tool for eliciting probability distributions from experts. *Environmental and Modelling Software*, 52:1–4, 2014.

[198] S.M. Fleming, C.L. Thomas, and R. J. Dolan. Overcoming status quo in the human brain. *Proceedings of National Academy of Sciences, USA*, 107:6005–6009, 2010.

[199] D. Kahneman, J.L. Knetsch, and R. H. Thaler. Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives*, 5:193–206, 1991.

[200] W. Samuelson and R. Zeckhauser. Status quo bias in decision making. *Journal of Risk and Uncertainty*, 1:7–59, 1988.

[201] Thomas Stöckl, Jürgen Huber, and Michael Kirchler. Bubble measures in experimental asset markets. *Experimental Economics*, 13:284 – 298, 2010.

[202] Robert K Merton. The matthew effect in science: The reward and communication systems of science are considered. *Science*, 159(3810):56–63, 1968.

[203] Robert K Merton. The matthew effect in science, ii: Cumulative advantage and the symbolism of intellectual property. *ISIS*, 79(4):606–623, 1988.

[204] Stefan. T. Trautman and Gijs van de Kuilen. Belief elicitation: A horse race among truth serums. *The Economic Journal*, 125:2116 – 2135, 2015.

[205] Colin Camerer and Robin Hogarth. The effets of financial incentives in experiments: A review and capital-labor-production framework. *journal of Risk and Uncertainty*, 19(3):7 – 42, 1999.

[206] Uri Gneezy and Aldo Rustichini. Pay enough or don't pay at all. *Quarterly Journal of Economics*, 115(3):791 – 810, 2000.

[207] Sander Nieuwenhuis, Dirk J. Heslenfeld, Niels J. Alting von Geusau, Rogier B. Mars, Clay B. Holroyd, and Nick Yeung. Activity in uman reward-sensitive brain areas is strongly contect dependent. *NeuroImage*, 25:1302 – 1309, 2005.

[208] Krishna P. Miyapuram, Philippe N. Tobler, Lucy Gregorios-Pippas, and Wolfram Schulz. BOLD responses in reward regions to hypothetical and imaginary monetary rewards. *NeuroImage*, 59:1692 – 1699, 2012.

[209] Signe Bray, Shinsuke Shimojo, and John P. O'Doherty. Human medial orbitofronal cortex is recruited during experience of imagined and real rewards. *Journal of Neurophysiology*, 103:2506 – 2512, 2010.

[210] Alice Lin, Ralph Adoplhs, and Antonio Rangel. Social and monetary reward learning engage overlapping neural substrates. *SCAN*, 7:274 – 281, 2012.

[211] Stefano I. Di Domenico and Richard M. Ryan. The emerging neuroscience of intrinsic motivation: A new frontier in self-determination research. *Frontiers in Human Neuroscience*, 11(145):1 – 14, 2017.

[212] Gary Charness, Uri Gneezy, and Brianna Halladay. Experimental methods: Pay one or pay all. *Journal of Economic Behavior and Organization*, 131:141 – 150, 2016.

[213] J.E. Beatty. Grades as money and the role of the market metaphor in management education. *Academy of Management Learning and Education*, 3:187–196, 2004.

[214] Anastasia Ejova, Paul H. Delfabbro, and Daniel J. Navarro. The illusion of control: Structure, measurement and dependence on reinforcement frequency in the context of a laboratory gambling. In *ASCS09: Proceeding of the 9th Conference of the Australasian Society for Cognitive Science*, pages 84–92, 2009.

[215] K. Binmore. *Does Game Theory Work? The Bargaining Challenge*. MIT Press., 2007.

[216] Robert L Axtell. Zipf distribution of US firm sizes. *science*, 293(5536):1818–1820, 2001.

[217] Xavier Gabaix. Zipf's law for cities: an explanation. *The Quarterly journal of economics*, 114(3):739–767, 1999.

[218] Joseph S Kong, Nima Sarshar, and Vwani P Roychowdhury. Experience versus talent shapes the structure of the web. *Proceedings of the National Academy of Sciences*, 105(37):13724–13729, 2008.

[219] Thomas Maillart, Didier Sornette, Sebastian Spaeth, and G Von Krogh. Empirical tests of Zipf's law mechanism in open source linux distribution. *Physical Review Letters*, 101(21):218701, 2008.

[220] Herbert A Simon. On a class of skew distribution functions. *Biometrika*, 42(3/4):425–440, 1955.

[221] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[222] Yannick Malevergne, Alexander Saichev, and Didier Sornette. Zipf's law and maximum sustainable growth. *Journal of Economic Dynamics and Control*, 37(6):1195–1212, 2013.

[223] CoinMarketCap. `https://coinmarketcap.com/`.

[224] Ke Wu, Spencer Wheatley, and Didier Sornette. Classification of crypto-coins and tokens from the dynamics of their power law capitalisation distributions. `http://datadryad.org/review?doi=doi:10.5061/dryad.t577f19`.

[225] Stan Higgins. Cryptsy Threatens Bankruptcy, Claims Millions Lost in Bitcoin Heist. `https://www.coindesk.com/cryptsy-bankruptcy-millions-bitcoin-stolen/`, 2016. Jan 15.

[226] Yannick Malevergne, Vladilen Pisarenko, and Didier Sornette. Testing the Pareto against the lognormal distributions with the uniformly most powerful unbiased test applied to the distribution of cities. *Physical Review E*, 83(3):036111, 2011.

[227] Alexander Saichev, Yannick Malevergne, and Didier Sornette. *Theory of Zipf's Law and Beyond*. Springer, 2010.

[228] Spencer Wheatley and Didier Sornette. Multiple outlier detection in samples with exponential & Pareto tails: Redeeming the inward approach & detecting dragon kings. *Swiss Finance Institute Research Paper Series*, pages 15–28, 2015. Available at SSRN: https://ssrn.com/abstract=2645709.

[229] Ginestra Bianconi and A-L Barabási. Competition and multiscaling in evolving networks. *EPL (Europhysics Letters)*, 54(4):436, 2001.

[230] FINMA. FINMA publishes ICO guidelines. https://www.finma.ch/en/news/2018/02/20180216-mm-ico-wegleitung/, 2018. Feb 16.

[231] Tian Chuan and Rachel-Rose OĹeary. China Outlaws ICOs: Financial Regulators Order Halt on Token Trading. https://www.coindesk.com/china-outlaws-icos-financial-regulators-order-halt-token-trading/, 2017. Sep 4.

[232] Stan Higgins. Not a Surprise: Blockchain Industry Saw SEC ICO Action Coming. https://www.coindesk.com/not-surprise-blockchain-industry-saw-sec-ico-action-coming/, 2017. Jul 26.

[233] Cynthia Kim. South Korea bans raising money through initial coin offerings. https://www.reuters.com/article/us-southkorea-bitcoin/south-korea-bans-raising-money-through-initial-coin-offerings-idUSKCN1C408N, 2017. Sep 29.

[234] Sam Mcbride. New poll: Catholic voters in Northern Ireland moving from Remain to Leave. http://www.newsletter.co.uk/news/new-poll-catholic-voters-in-northern-ireland-moving-from-remain-to-leave-1-7435854, 2016. Jun 17.

[235] Paul Waldman. Why did Trump win? In part because voter turnout plunged. https://www.washingtonpost.com/blogs/plum-line/wp/2016/11/10/why-did-trump-win-in-part-because-voter-turnout-plunged/?utm_term=.bddb6f6e6aaf, 2016. Nov 10.

[236] Brad Plumer. Low voter turnout in a few key states may have helped Donald Trump. https://www.vox.com/policy-and-politics/2016/11/9/13573904/voter-turnout-2016-donald-trump, 2016. Nov 9.