UNIVERSITY OF SOUTHAMPTON
FACULTY OF BUSINESS, LAW AND ART
SOUTHAMPTON BUSINESS SCHOOL

IT TAKES ALL SORTS:

THE COMPLEXITY OF PREDICTION MARKETS

VALERIO RESTOCCHI

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
JUNE 2018

Supervisors: Prof. Frank MC.GROARTY
Dr. Enrico GERDING
Prof. Johnnie E.V. JOHNSON

Examiners: Prof. Claudio J. TESSONE
Dr. Tiejun MA

UNIVERSITY OF SOUTHAMPTON
FACULTY OF BUSINESS, LAW AND ART
SOUTHAMPTON BUSINESS SCHOOL

IT TAKES ALL SORTS:
THE COMPLEXITY OF PREDICTION MARKETS

BY:
VALERIO RESTOCCHI

ABSTRACT

Prediction markets represent a great tool to harness the wisdom of the crowd and, for this reason, they are used to provide accurate forecasts on great variety of events. However, current models of prediction markets do not capture their full complexity, and fail to give satisfactory explanations of the price formation process and mispricing anomalies. This thesis consists of six separate, yet interconnected papers that address these gaps.

The first three papers analyse the favourite-longshot bias, a well known empirical regularity whereby contracts (or bets) on likely events are underpriced, whereas contracts on unlikely events are overpriced. The favourite-longshot bias has been widely observed especially in sports betting markets but, in contrast with other pricing anomalies, it did not disappear over time. In the first paper, we propose the first model that can explain the favourite-longshot bias and other related phenomena in different contexts. To achieve this, we introduce an agent-model in which market participants possess heterogeneous beliefs and risk attitudes, and find that such a model can accurately explain betting markets mispricing. Moreover, we shed new light on the role bookmakers have in generating mispricing, by considering two different strategies bookmakers can adopt to set prices and show that, in contrast to previous results, bookmakers are more likely to be risk minimisers (i.e., balancing the books only depending on demand) than profit maximisers. The second paper builds on the heterogeneous agents model to investigate the impact of transaction costs on mispricing. Our results suggest that transaction costs alone cannot create mispricing, as suggested by previous work, but significantly amplify its magnitude if mispricing exists already. In the third paper, we provide an analysis of the favourite-longshot bias in political prediction market exchanges, and characterise its temporal behaviour. We find that, on average, mispricing is negatively correlated

with duration, i.e., the longer the market, the smaller the favourite-lonsghot bias, but, surprisingly, we find that duration is strongly, and positively correlated to the magnitude of the favourite-longshot bias in the last days of trading, and argue that this is caused by herding dynamics.

The second part of the thesis continues the analysis of prediction market exchanges. Specifically, the fourth and fifth paper provide a comprehensive list of empirical regularities (or *stylised facts*) that we find in prediction market. This list comprises stylised facts on price changes, volume, and calendar effects. Overall, we find that prediction markets behave differently than financial markets, but share some common characteristics, especially regarding price changes, with emerging financial markets. In the sixth and last paper, we build on this work to introduce a model that can replicate the statistical properties of prediction markets. To achieve this, we propose a model in which agents belong to a social network, and can interact with each others by exchanging their opinions about the probability of a specific event to occur. We find that such a model is particularly suitable to explain prediction markets dynamics, and that it qualitatively reproduces the empirical properties of price changes even in the worst case scenario, suggesting strong robustness.

Contents

Declaration of Authorship

I, Valerio Restocchi, declare that the thesis entitled *It takes all sorts: the complexity of prediction markets* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- parts of this work have been published as:
  Valerio Restocchi, Frank McGroarty, Enrico Gerding, and Johnnie E. V. Johnson. It takes all sorts: a heterogeneous agent explanation for prediction market mispricing. *European Journal of Operational Research* (2018). doi: 10.1016/j.ejor.2018.04.011.
  Valerio Restocchi, Frank McGroarty, Enrico Gerding, and Johnnie E. V. Johnson. The impact of transaction costs on state-contingent claims mispricing. *Finance Research Letters*, November (23) (2017). doi: 10.1016/j.frl.2017.02.006.
  itemize

Signed:...

Date:...

List of Figures

List of Tables

"***It takes all sorts (to make a world)***"

UK SAYING

said to emphasize that people have different characters, opinions, and abilities, and that you should accept this.

– Cambridge English Dictionary

Chapter 1

Introduction

Perfectly rational agents and representative agent models have been for decades the tool of choice for economists willing to model financial markets. This approach, although having the merit of often being the only way to simplify economic models to the point of analytical tractability, have repeatedly been critizised. In financial economics, one of the first and most notable examples is the no-trade theorem (Milgrom and Stokey, 1982). This theorem states that, if all agents in a market are perfectly rational, and prices reflect the information available to traders, then none of them would be able to make any profit, and then no trade would exist. Clearly, this clever yet simple example, contradicts the idea that all agents can be assumed to have the same, rational behavior. Ten years after the no-trade theorem, Kirman (1992) publishes one of the first open cirtiques to the representative agent approach. He argues that, by grouping all the behaviors in a single one representing the average, all the complexity generated by the interaction of different behaviors would be missed, losing an important component of the final resulting dynamics. For these reasons, heterogeneous agent models have been experiencing a surge in their popularity in recent years (Hommes, 2006). However, the introduction of heterogeneous agents also increases the complexity of a model, which often becomes analytically intractable. The solution to this problem is to simulate the interaction between those agents constituting the heterogeneous agent models to find the numerical solutions to the problems.

In this work, we introduce new heterogeneous agent models, and simulate them when appropriate, to study prediction markets. Prediction markets are one of the most recent incarnation of future contracts, as they allow to trade on future events, whose outcomes are revealed at a predetermined point in time. Since their creation, prediction markets have been heralded as effective mechanisms for harnessing the *wisdom of the crowd* to make accurate forecasts (e.g., Berg et al. (2008); Arnesen and Bergfjord (2014)). The most common type of prediction markets allows trading on the outcome of political events (e.g., Iowa Electronic Market, Intrade, PredictIt), but this is not the only application. In fact, prediction markets can be divided into public, i.e., open to anyone who wants to bet on political elections, sports events, etc., and private. This latter type of market is used by large companies, such as Google, Intel, and General Electric, to provide accurate evaluate internal activities such as sales forecasts, or the likelihood for a team to meet certain performance goals (Plott and Chen, 2002; Cowgill et al., 2009). Despite their obvious importance, prediction markets have not been yet studied as a complex system, approach that has been widely employed in financial markets, and has led to surprising breakthroughs (Sornette, 2003; Hommes, 2006; Lux and Marchesi, 2000).

This thesis represents a collection of papers which aims at addressing this gap, each of them analysing a different aspect of the complexity of prediction markets. Broadly, this corpus of work can be divided in two parts. The first three papers tackle the issue of mispricing in prediction markets, and investigate its causes. Specifically, we focus on the favorite-longshot bias (FLB), an empirical regularity observed in many state-contingent-claim markets in the past seven decades. This anomaly has been extensively studied in sports betting markets, but to date there is till debate on why this pricing anomaly still exists and what drives it. In the first three papers, we introduce the first heterogeneous agent model to study the FLB, and show that such a model it is far better at explaining historical data. Then, we analyse the impact of transaction costs and find that market frictions cannot cause the FLB, as previously suggested (Hurley and McDonough, 1996), but only amplify its effects. Lastly, we conduct a thorough analysis of the FLB in political prediction markets, focussing on its temporal dynamics.

The second three papers represent a comprehensive analysis of political prediction markets. Specifically, we extensively analyse data from PredictIt, a political prediction market exchange, and compile a list of empirical properties that characterise prediciton market time series, including price changes, volume, and calendar effects. In the last paper, we employ these findings to build the first agent-based model of prediction markets that matches the stylised facts of historical data.

Below, a more detailed overview of each paper is presented.

The paper, presented in Chapter 2, consists of a paper published in the *European Journal of Operational Research*, under the title **It takes all sorts: a heterogeneous agent explanation for prediction market mispricing**. In this paper, we introduce a heterogeneous agent model that can capture the complexity of prediction market anomalies. Pricing anomalies threaten the value of prediction markets as a means of harnessing the 'wisdom of the crowd' to make accurate forecasts. The most persistent and puzzling pricing anomaly associated with price-implied prediction probabilities is the favourite-longshot bias (FLB). We demonstrate that existing models of the FLB fail to capture its full complexity, thereby preventing appropriate adjustments to market forecasts to improve their accuracy. We develop an agent-based model with heterogeneous agents in a fixed-odds market. Our agent-based simulations and comprehensive analysis using market data demonstrate that our model explains real market behaviour, including that of market makers, better than existing theories. Importantly, our results suggest that adequately complex models are necessary to describe complex phenomena such as pricing anomalies. We discuss how our model can be used to better understand the relation between market ecology and mispricing in contexts such as options and prediction markets, consequently enhancing their predictive power.

In the paper constituting Chapter 3, titled **The impact of transaction costs on state-contingent claims mispricing** and published in *Financial Research Letters*, we build on the model presented in Chapter 2 and analyse the impact that transaction costs have on asset mispricing in state-contingent claims markets. In particular, we examine betting markets, in which, it has been argued, transaction costs cause the FLB. By using a heterogeneous agents model, we prove that transaction costs alone cannot cause mispricing. Also, we run agent-based simulations to characterise the response of market prices to increments in transaction costs. We find that transaction costs have a significant impact on market inefficiency, by amplifying existing mispricing both directly, influencing market prices, and indirectly, inducing a non-linear response from the agents.

Chapter 4 consists of a paper titled **The temporal evolution of mispricing in prediction markets**, and is currently under review by *Finance Research Letters*. In this paper, we analyse mispricing in prediction markets, a powerful forecasting tool that harnesses the wisdom of the crowd. Specifically, we focus on the FLB. We show that prediction market prices exhibit the FLB, and we quantify its temporal evolution. Our results suggest that the average level of the FLB decreases with market duration but always grows towards the end of the market. Also, we find that the level of mispricing the day before the end of the market is negatively correlated with market duration. We propose an explanation to this results based on the temporal evolution of trading volume in prediction markets, and argue that the high level of mispricing in the last trading days may be caused by herding behaviour caused by media coverage of prediction markets.

Chapters 5 and 6 represent the beginning of the second part of this thesis, and are constituted by two papers with title **The stylised facts of predicition markets: analysis of price changes** and **Statistical properties of volume and calendar effects in prediction markets**, both currently under review by *Physica A*. In these two papers, we argue that prediction markets are a powerful tool to make accurate predictions about the outcome of an event and, for this reason, they attract the interest of researchers and practitioners alike. To date, there exist no robust means of validation for quantitative models of prediction markets. To address this shortcoming, we analyse the statistical properties of volume as well as the seasonal regularities (i.e., calendar effects) shown by volume and price. Specifically, we compile a list of empirical regularities (stylised facts) of price changes we find by analysing daily price changes from 3385 prediction markets on political events, a data set provided by PredictIt. Furthermore, we find that volume, with the exception of its seasonal regularities, possesses different properties than what is observed in financial markets. Finally, our results show that price does not exhibit any calendar effect. These findings suggest a significant difference between prediction and most of financial markets (although they possess similarities with emerging markets), and offer evidence for the need of studying prediction markets in more detail.

The final chapter of this thesis is constituted by a paper that builds on chapters 5 and 6 to introduce a heterogeneous agent model that accurately describes prediction markets. In this paper, titled **Opinion dynamics and price formation in prediction markets**, we argue that models to describe prediction markets often fail to capture the full complexity of the underlying mechanisms that drive price dynamics. To achieve this, we propose a model in which agents belong in a social network, and have an opinion about the probability of a particular event to occur and bet on the prediction market accordingly. Agents update their opinions about the event by interacting with their neighbors in the network, following the Deffuant model of opinion dynamics. Our results suggest that a simple model that takes into account opinion formation dynamics is capable to replicate the empirical properties of historical prediction market time series, such as volatility clustering and fat-tailed distribution of returns. Moreover, this is the first paper that addresses the problem of validating opinion dynamics models against real data by using historical data obtained from PredictIt, an exchange platform that has never been used before to validate models of opinion diffusion.

Chapter 2

It Takes All Sorts: A Heterogeneous Agent Explanation for Prediction Market Mispricing

2.1 Introduction

Financial instruments such as forward and futures contracts have long been used to reveal people's collective expectations about future outcomes by tapping into demand and supply in the market today to set a price for delivery (settlement) of a commodity (asset value) at some specified future date. The forward contract is one of the most ancient financial instruments in existence. The most recent incarnation of the futures contract, so-called 'political futures' related to the US Presidential election outcomes, have been traded on the Iowa Electronic Markets since 1988. The Iowa market and its emulators, such as Intrade and PredictIt, together with event-based betting (e.g. election outcomes, interest or tax rate changes, sports events etc.) offered by bookmakers and betting exchanges, are known collectively as *prediction markets*. These markets are heralded as effective mechanisms for harnessing the *wisdom of the crowd* to make accurate forecasts (e.g., Berg et al. (2008); Arnesen and Bergfjord (2014)).

The ability of the forecasts drawn from prediction markets to fully reflect all relevant information is predicated on the Efficient Market Hypothesis (EMH) (Fama, 1970), which states that market prices always incorporate all relevant information. In particular, this would suggest that those who hold information would continue to trade in prediction markets until they believe that their information is fully assimilated into the prevailing market prices. However, pricing anomalies, which result in market prices failing to fully reflect all relevant information, present a potential threat to the forecasting accuracy of prediction markets. If the mechanisms which lead to a particular pricing anomaly are well understood, then it may be possible to identify when and how to adjust final market prices to annul the impact of the anomaly. However, if overly simplistic models of these anomalies are employed, adjustments based on these models may be wholly inadequate, or in some cases harmful, to the accuracy of resulting forecasts.

We argue that in prediction markets, as much as in other fields, it is important that a model describe a phenomenon with the sufficient degree of complexity. Otherwise, it risks missing those meaningful dynamics that are a consequence of complex interactions between the components of the system (e.g., market participants). Models that do not describe these phenomena precisely are unlikely to accurately forecast prices (e.g., fail to spot early signals of herding behaviour) or, in the case of state-contingent claims markets, such as options and prediction markets, to estimate the true probabilities associated to an event.

For more than a century, it was argued that models with only a representative agent could sufficiently describe systems such as financial markets, in which participants are supposed to be rational. However, with the growth of computational power and the introduction of behavioral economics, many criticise this approach. Indeed, although the representative agent approach has the important merit of making models analytically tractable, it has been argued that the representative agent is an unjustified, incorrect assumption that leads to fundamentally wrong conclusions (Kirman, 1992). Furthermore, Heckman (2001) argues that it is vital that models account for the obvious differences among individuals, since heterogeneity plays a considerable role in economic behaviour. More recently, heterogeneous agents models (HAM) have been employed to study a range of complex phenomena across a number of disciplines. In finance, HAMs have been used to explain phenomena such as heavy tails in the distribution of returns, bubbles, and other pricing anomalies that would have been impossible to describe with representative agents (see e.g. Brock and Hommes (1997), Lux and Marchesi (1999a), Buckley et al. (2012) and Joëts (2015)).

Although a considerable amount of effort has been expended in finding the roots of pricing anomalies in financial markets, some of these anomalies are left without a fully convincing explanation, mainly due to the use of overly simplistic models. One of these mispricing phenomena is the favourite-longshot bias (FLB).

The FLB is an empirical regularity found in state-contingent claims markets, whereby the average return on likely outcomes is greater than the average return on less likely outcomes. That is, to use sports betting terminology, favourites are underbet and longshots overbet; i.e. the chance of high/low probability events are under-/over-estimated. This market anomaly has been extensively studied in sports betting markets, which have the essential characteristic of interest to us, i.e. being able to extract forecast probabilities about future outcomes from the spending behavior of the crowd (Ma et al. (2016)). However, the importance of the FLB goes well beyond sports betting markets. In particular, given the similarities between sports betting markets and traditional state-contingent claims markets it is not surprising that the FLB has been observed in a variety of such markets (Hodges et al. (2003) for a study of the bias in S&P500 and FTSE100 index futures options, and Wolfers and Zitzewitz (2004) for a discussion on the FLB in prediction markets). Thus, a model that adequately explains the FLB could be used to estimate the magnitude of mispricing in prediction markets under given conditions, allowing suitable adjustments to be made; thus significantly improving the accuracy of the predictions.

In this paper, we make an important first step towards this by showing a way to accurately reconstruct price curves that account for mispricing caused by the FLB. This model can be reverse engineered to derive the true probability of an event given the biased market prices, consequently improving the forecasting power of prediction markets.

The existing literature proposes a variety of theories to explain the FLB. All these indicate that the bias arises given some assumptions.[1] These theories all lead to qualitatively similar results, but they exhibit two common drawbacks. First, they lack empirical support across different markets. That is, the models are usually only tested on a single market. Second, these theories usually employ a representative agent to model bettors (or, at most, a representative agent beside noise traders). Consequently, these models are insufficiently flexible to explain related phenomena such as the reverse FLB (Woodland and Woodland, 1994, 2003) or the FLB in markets beyond the one studied.

Our goal is to build an appropriately complex model capable of providing a comprehensive explanation of the FLB in its various forms. We achieve this by generalizing the most important theories that try to explain this phenomenon into a HAM with five agent types. The potential for a HAM to explain the FLB is highlighted by the observation that betting markets feature a variety of traders who display significantly diverse behaviours (Rhoda et al., 1999). Furthermore, Crawford and Pendakur (2013) find that a representative agent can only explain two thirds of the variation in consumption behaviour, while using four or five classes of agents can completely rationalise all consumer choices in their data.

[1] See Ottaviani and Sørensen (2008) for a survey on the FLB in sports betting markets.

We model a fixed-odds market using an agent-based model. We focus on developing a model to explain the FLB in a betting market with a market maker, since these markets have been shown to be those most prone to the FLB. In the model, a bookmaker faces five different types of bettors, each associated with a different behaviour suggested in the literature and modelled using prospect theory. That is, we derive the maximum price (i.e. minimum odds) bettors are willing to accept, depending on the behaviour class to which they belong. This allows us to build a model in which agents with different behaviours, borrowed from literature explaining the FLB with risk preference (e.g. Ali (1977), Golec and Tamarkin (1998)) and misbelief (e.g. e.g. Snowberg and Wolfers (2010), Gandhi and Serrano-Padial (2015)), participate simultaneously in the market as separate entities, hence enabling us to measure their relative contribution to price formation.

Most similar to our approach are the works by Chiappori et al. (2012) and Gandhi and Serrano-Padial (2015), who have recently added to the debate by introducing heterogeneity in their models, which focus only on either heterogeneous preferences (Chiappori et al., 2012) or heterogeneous beliefs (Gandhi and Serrano-Padial, 2015). In contrast, the HAM we propose in this paper considers heterogeneity on both beliefs and risk preferences simultaneously. Our results, similarly to theirs, suggest that heterogeneous agents are fundamental to capture the full complexity of the FLB and, more generally, of prediction markets. However, we identify three important contributions that distinguish our paper from previous literature. First, we show that our HAM, unlike any of the existing representative agent models, can explain different degrees of the FLB and the reverse FLB. Indeed, we prove that, in the presence of heterogeneous bettors, the FLB and its negative counterpart can occur regardless of whether the market maker adopts a profit maximisation or a risk minimisation pricing strategy.

Second, we analyse two different pricing strategies the market maker can adopt. Knowing the market maker's pricing strategy is fundamental to being able to reconstruct the price curve, and a necessary first step for more accurate forecasts. We show that using our HAM, the best fit to FLB data from three sports with very diverse degrees and types of FLB is achieved if bookmakers act as risk minimisers, whereas the common assumption is that bookmakers are profit maximisers. The only exception is by Fingleton and Waldron (1999), as they analyse the odds from 1696 races in Ireland in 1993 and reject the hypothesis that bookmakers seek to maximise their expected profit, but also fail to reject the hypothesis that bookmakers are risk minimisers, suggesting that this might be their behaviour. Also, our results agree with recent empirical work by Kopriva (2009) and Feess et al. (2016), who found that, on data from Betfair and the New Zealand Racing Board respectively, bet sizes are significantly decreasing in odds (i.e., volumes on longshots are systematically lower). Our finding is important since this knowledge enables appropriate construction of price curves which can enable us to adjust prediction market forecasts to better reflect the true underlying event probabilities.

Third, we determine the distribution of market prices for different market compositions (i.e. agent proportions) using agent-based simulations. Consequently, by observing the FLB for the three sports betting markets we examine (i.e. tennis win market, baseball win market, and under-over 2.5 goals in football), we are able to find their market compositions. This enables us to reconstruct the price curves given the probability of an outcome and adjust for the FLB, hence improving the accuracy of forecasts derived from this prediction market.

Our results suggest that our HAM is better than representative agent models at describing prices in betting markets. More generally, these findings suggest that a HAM is necessary to fully understand the FLB, supporting the idea that representative agent models cannot describe complex systems in a general way with sufficient accuracy.

The paper is organised as follows. In Section 7.2 we define the agents and introduce the formal model. In Section 2.3 we derive and discuss the theoretical results, proving that a HAM can explain the FLB under different market conditions. Section 2.4 focuses on empirical findings, such as the estimation of equilibrium price curves and market compositions, and the comparison with existing theories. In this section we also perform a sensitivity analysis of agent classes' contributions to equilibrium prices. We draw conclusions in Section 5.

2.2 The Model

To model the betting market, we use an agent-based model, in which one of the agents act as a price-setter (the bookmaker) and the others represent traders (bettors). The bookmaker faces many bettors, and sets the odds in the market depending on the bettors' actions, according to one of two possible pricing strategies. Bettors have heterogeneous beliefs and attitudes towards risk depending on one of the five classes to which they are assumed to belong. As a result, the value they assign to gambles differs depending upon the *behaviour class* to which they belong. We derive the value they assign to gambles from prospect theory by computing the maximum price they are willing to pay to bet on an event.

We consider a time-limited market in which there is only one event with two possible outcomes A and B, occurring with probabilities p_A and $p_B = 1 - p_A$ respectively. Let π_A and π_B be the prices of two contracts that pay one pound if the corresponding outcome occurs, and zero otherwise (Arrow-Debreu securities). Hence, prices are bounded by $0 < \pi_i \leqslant 1$, where $i \in \{A, B\}$. The boundary value $\pi_i = 0^+$ implies a possible infinite return on the bet amount and $\pi_i = 1$ implies no possible returns or, from the bookmaker's point of view, that bets are not accepted.

2.2.1 Heterogeneous Bettors and Subjective Fair Prices

We model six different types of agents, namely the bookmaker and five behaviour classes of bettors. Each behaviour class is associated with a set of beliefs and risk attitude such that, given a probability p_i, each agent considers $\pi_i^s(p_i)$ to be the fair price for buying the *i-th* ticket. That is, they are willing to bet on the *i-th* outcome if and only if $\pi_i < \pi_i^s(p_i)$, where $\pi_i^s(p_i)$ is referred as the *subjective fair price*. We assume that agents do not possess any information regarding the market, the bookmaker, or other participants. Specifically, they do not know the market composition, other agents' utility functions, or the bookmaker's strategy. Therefore, their decision making process (i.e., whether to bet or not) is solely based on their utility function, which is derived below for each bettor type.

To derive bettors' subjective fair price functions we use prospect theory, because it allows us to derive all the agents' subjective fair price functions from the same equation. Also, prospect theory has shown to be a good estimator of bettors' preferences in sports betting (Jullien and Salanié, 2000).

In more detail, we define functions for both attitude towards risk and probability weighting. Following Tversky and Kahneman (1992), we represent the value function $v(x)$ as:

$$v(x) = \begin{cases} x^\alpha & \text{if } x \geqslant 0 \\ -(-x)^\alpha & \text{if } x < 0 \end{cases} \tag{2.1}$$

Then, we use Prelec's function (Prelec, 1998), which is a probability weighting function that accounts for people oversestimating low probabilities, as the probability weighting function:

$$w(p) = e^{-[-ln(p)]^\beta} \tag{2.2}$$

This means that each agent associates a (subjective) probability $w(p_i)$ to the *i-th* outcome, with a corresponding potential profit of $1 - \pi_i$, and therefore expects to lose π_i with probability $w(1 - p_i)$. Combining the two, we can write the utility function as follows:

$$u(\pi_i, p_i) = w(p_i)v(1 - \pi_i) + w(1 - p_i)v(-\pi_i) \tag{2.3}$$

At equilibrium, every agent must be indifferent between trading and not trading, so we find the subjective fair price by setting $u(\pi_i, p_i) = 0$. There is no generic closed-form solution, but specific solutions for each class of agents can be found after setting α and β. Therefore, we define five different agent classes by assigning them specific combinations of values for α and β, deriving the subjective fair price functions from Eq. (2.3). We do so for all the agents except the random bettors, who do not seek utility maximisation, and are detailed below. Next, we describe the 5 specific trader classes in more detail (derivations of the fair price functions used are given in the Appendix A).

The first class, random bettors, represents noise traders, and is modelled in a similar manner to that conducted by Shin (1992). Noise traders are the only bettors who do not have subjective fair prices: they attach probability 1 to an outcome randomly picked from the uniform distribution Unif$\{1, i, ..., N\}$ (where N is the number of outcomes). In our case, they bet on either A or B with probability $p = 0.5$ each. These agents can be seen as bettors who do not possess good information or bet without taking it into account. For example, one can think of these traders as betting on their favourite team in sports betting or investors hedging their position on physical commodities in the options market (Buckley and Long, 2015).

The second class we introduce is the informed bettor class. The literature often defines informed traders as those who have perfect information about the future state of nature. This characterisation, introduced in sports betting by Shin (1991), assumes that insider traders know the outcome of an event with probability 1. However, it could be argued that this is very unlikely to be the case in practice and the assumption has been criticised. Consequently, we assume that informed agents are expert traders, and they are informed in the sense that they have access to all the public information on the event and are capable of processing it without bias, as assumed by Sauer (1998). Under these premises, they can be considered *professionals*. Thus, they are risk-neutral agents who know the true probabilities associated with the event, implying $\alpha = \beta = 1$. As a result, their subjective fair prices correspond exactly to the true probabilities:

$$\pi^s(p) = p \tag{2.4}$$

The third behaviour class models misperceptions of probability, and has been extensively described in the literature. Most notably, Snowberg and Wolfers (2010) show that, under the representative-agent assumption, misperceptions are more likely to explain the FLB in horse racing than a risk-loving behaviour. We follow them in modelling these agents as risk-neutral bettors ($\alpha = 1$) who weight probabilities with Prelec's function with a coefficient $\beta = 0.928$. Although β might seem too close to 1, this is sufficient to distinguish misperceiving bettors from informed bettors. Then, the obtained subjective fair price function is:

$$\pi^s(p) = -\frac{e^{-(-ln(p))^{0.928}}}{-e^{-(-ln(1-p))^{0.928}} - e^{-(-ln(p))^{0.928}}} \tag{2.5}$$

The last two behaviour classes represent agents who are not risk neutral.

Risk aversion is found to be one of the behaviours of bettors by Rhoda et al. (1999). Furthermore, risk aversion is a standard assumption in most economic and financial models. Consequently, considering this behaviour class is essential when developing a comprehensive, general model. To determine the parameters for their utility function we follow Gonzalez and Wu (1999). They conduct an experiment in which participants are asked to take gambles; finding that players' median risk attitude value is $\alpha = 0.49$ (with a standard deviation of $\sigma = 0.04$). For sake of tractability, we round this number to 0.5, without making any significant change to the function. The agents who belong to this behaviour class are not affected by misperceptions of probability, hence $\beta = 1$

Thus, choosing $\alpha = 0.5$ to model their risk attitude, their subjective fair price function is:

$$\pi^s(p) = \frac{p^2}{1 - 2p + 2p^2} \tag{2.6}$$

The last behaviour class we model represents risk-loving agents. Risk-seeking behaviour is frequently employed to model representative agents in sports betting markets because this behaviour alone can cause the FLB. We model this behaviour class with a risk-seeking value of $\alpha = 2$ but assume no misperceptions. Since α is the exponent of the power function, setting $\alpha = 2$ gives us two possible solutions. One of them is negative for $0 \leqslant p < 0.5$ and discontinuous in $p = 0.5$, with limits of minus and plus infinity as p approaches 0.5 from left and right, respectively. This is unrealistic and consequently we employ the other solution to model the subjective fair price function for risk-loving agents:

$$\pi^s(p) = \frac{p - \sqrt{p - p^2}}{2p - 1} \tag{2.7}$$

which is defined for $p \in [0, 0.5) \cup (0.5, 1.0]$ and has a removable discontinuity at $p = 0.5$. However, as shown in Section 2.4, this is not an issue in computing subjective fair prices.

2.2.2 The Betting Process

We model the betting mechanism as an iterative process, in which the bookmaker can adjust prices depending on its utility function and on bettors' actions. In order to reproduce the time-limited nature of sports betting markets, this process is repeated until equilibrium is reached or an arbitrary number T of rounds are played. The betting mechanism can be described by three stages:

STAGE 1. Nature chooses the true probabilities $\{p_A, p_B\}$ for the outcomes $\{A, B\}$ and chooses the market composition distribution. That is, Nature fixes the proportion of each behaviour class participating in the market.

STAGE 2. The bookmaker sets a pair of prices $\{\pi_A, \pi_B\}$ and allows the agents to bet on either A or B.

STAGE 3. An arbitrarily large number N of bettors is chosen to participate in the market. The number of agents belonging to each class is determined by the market composition distribution selected in Stage 1. All the bettors see the prices and compute their corresponding utility. This process has three steps. First, bettors compute their subjective fair price on all the outcomes according to their subjective fair price function. Second, they compute their utility for betting on each outcome and the utility for not betting. Last, they perform the action that maximises their utility (i.e. betting on A, on B, or not betting).

Note that the market composition (i.e., the distribution of bettors) represents only potential bettors. That is, since agents have the possibility to choose not to bet, the initial market composition chosen at Stage 3 does not necessarily reflect the distribution of classes that actually bet throughout the market. This allows the model to replicate the nature of betting markets, in which prices fluctuate depending on bettors choices. Also, since price is path dependent, the initial market composition (i.e., the potential distribution of bettors) is more indicative than the final one, because it allows to understand what is the necessary starting condition to have given prices at the end.

Finally, modelling the betting process in this way allows our HAM to replicate two essential features of betting markets, without making unnecessary assumptions about boomaker's and agents' information and behaviour. First, prices are allowed to change at each betting round, and consequently bettors get different odds depending on when they trade. Second, volumes are not evenly distributed over the two outcomes (see Section 4), and bet sizes on longshots are consistently lower, as observed in recent work by Kopriva (2009) and Feess et al. (2016).

2.2.3 Price Setting

Depending on the bookmaker's selected strategy, the process has different equilibria (i.e. prices converge to different values). In what follows, we first introduce the general setting in which the bookmaker fixes the odds. We then analyse two possible pricing strategies s/he can adopt, namely profit maximisation and risk minimisation. Profit maximisation has often been assumed to be the only strategy available to the bookmaker (e.g. Shin (1991, 1992, 1993), Levitt (2004)). However, even if a small subset of bettors can outperform bookmakers in terms of predicting outcomes, it is possible for the bookmaker to lose money in a given game if s/he sets prices to maximise the expected profit.

Let π_i^f be the fair price on the *i-th* outcome. We define the *fair* price on the *i-th* outcome as the price of a bet such that $\pi_i = \pi_i^f = p_i$. We assume that the bookmaker sets the fair prices, and that a total amount of money V (volume) is bet on the event. We also denote by V_A and V_B the volumes on A and B, respectively, hence $V = V_A + V_B$. The bookmaker's expected profit is given by:

$$\mathbb{E}(P) = V - \frac{p_A}{\pi_A^f} V_A - \frac{p_B}{\pi_B^f} V_B \tag{2.8}$$

therefore:

$$\mathbb{E}(P) = 0 \qquad \forall\, V_A, V_B \tag{2.9}$$

However, in any single game the bookmaker's actual profit may vary depending upon whether A or B occurs, as follows:

$$P = \begin{cases} V - \frac{V_A}{\pi_A^f} & \text{if A occurs} \\ V - \frac{V_B}{\pi_B^f} & \text{if B occurs} \end{cases} \tag{2.10}$$

Without loss of generality, we normalise volumes by setting $V = 1$, so that V_A and V_B become the percentage of money, or relative volume, bet on A and B. Now, it is clear that the bookmaker has a possibility to incur losses every time that $V_A \neq \pi_A^f$, which also implies $V_B \neq \pi_B^f$. That is, if $V_A > p_A$ (consequently $V_B < p_B$) the bookmaker will experience a loss with probability p_A, while in case that $V_A < p_A$ ($V_B > p_B$), s/he will experience a loss with probability p_B.

Next we study the price-setting strategy the bookmaker uses to maximise their utility. We consider two different utility functions, one assuming the bookmaker wants to maximise their expected profit and the other assuming s/he is only concerned to make the same profit regardless of the outcome, as long as the profit is non-negative. Since the bettors are price sensitive, it is not possible to find the prices explicitly as they depend on the volumes, which in turn depend on the choices made by the agents, which in turn are a function of prices. However we show under both these assumptions, how a model with bettors with heterogeneous behaviours can account for both the FLB and its negative counterpart.

ASSUMPTION 1. The bookmaker is a profit maximiser (PMB). That is, s/he prefers to maximise their expected profit function rather than minimise profit volatility. Therefore, their utility is $u(p, V_A, V_B, \pi_A, \pi_B) = \mathbb{E}(P) = V - \frac{p_A}{\pi_A} V_A - \frac{p_B}{\pi_B} V_B$.

ASSUMPTION 2. The bookmaker is a risk minimiser (RMB). That is, s/he prefers to minimise profit volatility rather than maximise the expected profit function. Therefore, their utility is $u(V_A, V_B, \pi_A, \pi_B) = -|\frac{V_A}{\pi_A} - \frac{V_B}{\pi_B}|$.

The latter utility function has a global maximum for $\frac{V_A}{\pi_A} = \frac{V_B}{\pi_B}$. That is, risk is minimised if the bookmaker's profit is the same regardless of the outcome. By assuming that the bookmaker is willing to make zero profit, we can substitute $V - \frac{V_A}{\pi_A} = 0$ and $V - \frac{V_B}{\pi_B} = 0$ in Eq. (2.10). Then, by solving it for π_A, π_B, and recalling that $V = 1$, we obtain:

$$\begin{cases} \hat{\pi}_A = V_A \\ \hat{\pi}_B = V_B \end{cases} \tag{2.11}$$

where $\hat{\pi}_A$ and $\hat{\pi}_B$ are those prices that allow the bookmaker to have a profit of zero whatever the outcome, and we refer to them as risk-neutral prices.

To keep the model as simple as possible, we also assume that bookmakers do not have any knowledge of bettor utilities. Consequently, bookmakers do not take into account how changing prices may affect the absolute betting volumes, and only optimise their profits per bet. This is a possible limitation to our description of a bookmaker, which however allows us to make the assumption that bookmakers do not possess full information about the market. This latter assumption has two advantages. First, it enables us to keep the model simple, which is essential to accurately quantify the contribution of each agent class to market prices. Second, it does not force us to make any unrealistic assumption about the degree of information that bookmakers possess about their (existing and new) customers, which is not known and, more generally, may vary among different bookmakers.

2.3 Theoretical Results

In this section we analyse the implications of Assumptions 1 and 2, outlined in Section 2.3. We show how, under the assumption that the market is composed of heterogeneous agents, a profit maximisation or a risk minimisation strategy can result in the FLB. Moreover, we show how both strategies can also account for the reverse FLB, a market anomaly for which the favourite is overbet and the longshot is underbet. This empirical anomaly has been observed by Busche and Hall (1988) in Hong Kong horse racing and by Woodland and Woodland (1994) in US baseball.

There is no closed-form equation for prices set by a profit-maximising bookmaker under the assumption that agents are price-sensitive. Consequently, we prove that both the FLB and the reverse FLB can be obtained in at least one case by using fixed agent proportions. That is, we make use of two special cases of the profit-maximising bookmaker model which make it analytically tractable, and prove that in these cases either the FLB or the reverse FLB exist.

2.3.1 The favourite-longshot Bias

In a two-outcome market, the FLB exists if the following holds:

$$\frac{p_A}{p_B} > \frac{\pi_A}{\pi_B} \Leftrightarrow p_A > p_B \tag{2.12}$$

Therefore, by assuming A and B are the most and least likely outcomes (favourite vs. longshot, or $p_A > p_B$), we can formulate three propositions, whose rigorous proofs can be found in the Appendix B. The propositions share the assumptions that bettors are price-sensitive, belong to one of the five classes described in Section 2.2.1 and behave accordingly. In addition, we assume that the market has the structure described in Section 2.2.2.

PROPOSITION 1. With a profit-maximising bookmaker (PMB), in a market populated only by risk-loving traders with subjective fair prices defined by Eq. (2.7), there is always a FLB.

PROPOSITION 2. With a profit-maximising bookmaker (PMB), in a market populated only by risk-averse traders with subjective fair prices defined by Eq. (2.6), there is always a reverse FLB.

The reasoning behind the propositions is as follows: a bookmaker will make a maximum profit if s/he persuades as many people as possible to bet on the outcome that gives the bookmaker the highest expected profit. In a market in which all the traders share the same beliefs, the bookmaker's best strategy is to set the maximum price the agents would pay for betting on the outcome that guarantees the bookmaker the best expected return. For example, say there is a tennis match in which player A has a true probability of winning of $p_A = 0.8$. For risk-loving bettors, subjective fair prices for this event are $\pi^s_A = 0.\overline{6}$ and $\pi^s_B = 0.\overline{3}$. This is the equilibrium for the agents, meaning that they are indifferent between betting on either A or B, or not betting at all. Since all the traders bet on the same outcome we can split the expected profit in two parts. Hence, the expected profit for the bookmaker is $P(A) = 1 - \frac{p_A}{\pi^s_A} \simeq -0.2$ and $P(B) = 1 - \frac{p_B}{\pi^s_B} \simeq 0.4$ if all the bettors bet on A, or B, respectively. This implies that it is sensible for the bookmaker to set prices $\pi_B = \pi^s_B$ and $\pi_A > \pi^s_A$ to ensure all the agents bet on B.[2] A similar (but opposite) reasoning, can be employed to prove proposition 2.

PROPOSITION 3. With a risk-minimising bookmaker (RMB), in a market in which relative volumes are not equal to true probabilities, there is always either the FLB or the reverse FLB.

The proof for proposition 3 is straightforward. Let k be the percentage deviation of V_A from the true probability on the same outcome p_A, so that $V_A = p_A(1 + k)$. For example, if $p_A = 0.6$ and the percentage V_A of money bet on A is $V_A = 0.8$, then V_A is 33% greater than p_A, or $k = 0.33$. Therefore, since $V_A + V_B = 1$, we obtain the following system of equations for the relative volumes:

$$\begin{cases} V_A = p_A(1 + k) \\ V_B = p_B - p_A k \end{cases} \tag{2.13}$$

By substituting these in Eq. (3.3) we obtain the new risk-neutral prices. Then,

[2] For this example we assume, without loss of generality, that traders are willing to bet even if $\pi^s_B = \pi_B$ rather than the stricter $\pi^s_B < \pi_B$.

we can check when inequality (4.1) holds, i.e. when FLB exists:

$$\frac{p_A}{p_B} > \frac{p_A + p_A k}{p_B - p_A k} \implies k < 0 \tag{2.14}$$

Recalling Eq. (3.3) it follows that, if $k \neq 0$, we have either the FLB or the reverse FLB. More specifically, if $k < 0$ ($k > 0$), so that more bets are placed on the favourite A (longshot B), the (reverse) FLB occurs. That is, whenever more money than is fair (i.e. the percentage of total money bet on an outcome is greater than the probability of such outcome to happen, that is $V_i > p_i$) is bet on the longshot, there is a FLB, and whenever more money than is fair is bet on the favourite, there is a negative FLB.

2.4 Empirical Analysis

In this section we find those market compositions that best fit historical FLB data from three diverse betting markets. We start by describing the three data sets and the agent-based simulations we use to find equilibria of the betting process described in Section 2.2.2 using different combinations of agent proportions (i.e., populations of potential bettors). We then present the results obtained from the agent-based simulations and discuss these by comparing them with data from real betting markets. We find strong evidence that bookmakers are more likely to follow a risk-maximisation (cf. profit-maximisation) strategy. In particular, the former fits historical data better, returns more realistic market compositions, and gives a more accurate representation of the FLB.

To decide which agent composition is the closest to reality for each market, we run a simulation for each possible market composition combination. We then compute the mean squared error (MSE) between the equilibrium prices generated by the simulations and historical prices for all the tested market compositions and choose the market composition that minimises the MSE. Furthermore, we compare our models with two of the most widely cited models, namely those of Shin (1993) and Snowberg and Wolfers (2010), and show that our RMB model displays significantly smaller MSEs. We also show that our RMB model exhibits higher adaptability, since the errors do not vary significantly between markets. We argue that these results suggest that our HAM can be regarded as a generalisation of existing models of the FLB.

2.4.1 Data

We use three data sets in our analysis, involving historical prices on tennis and baseball win bets and under-over 2.5 goals bets in football. The data on tennis and football are publicly available from specialised websites[3], while the baseball data are supplied in Woodland and Woodland (2003). These data have been employed in several studies (e.g. Forrest and Simmons (2008), Hvattum and Arntzen (2010), Forrest and Mchale (2007), Malueg and Yates (2010)). One of the reasons we choose these data is that they relate to two outcome events, since in tennis and baseball a draw is not possible and in a football match the total number of goals scored is always greater or less than 2.5. Most importantly, the data sets present three diverse levels of the FLB. That is, the tennis and baseball win markets are very efficient compared to the under-over 2.5 (which displays a large FLB). The reason why we choose under-over 2.5 is that this is the score bet category that attracts the most bets and for which the most data is available. [4] Baseball odds, instead, have been shown to display a negative FLB. Hence, these data enable us to test our model under three different scenarios (small, large, and negative FLB).

The tennis data incorporate the mean value of the odds provided by five different bookmakers on both the competitors in 27,069 matches in the ATP circuit over ten seasons, from January 2005 to November 2014. The football data incorporate a total of 13,013 matches in the seasons between 2005/2006 and 2013/2014, considering mean odds across 30 bookmakers on the four major European football leagues (Italian Serie A, English Premier League, Spanish Liga, and German Bundesliga). The baseball data incorporate odds data on 44,675 Major League Baseball matches between 1978 and 1999. The baseball data differs to that of the football and tennis in that they use odds lines rather than decimal odds, and odds across bookmakers are already aggregated.

The data sets we employ are sufficiently large to test the goodness-of-fit of our model, and are one to two orders of magnitude larger than most of the earlier studies on the topic (other than a few studies which employ data for horse racing, e.g. Snowberg and Wolfers (2010)).

[3] *www.tennis-data.co.uk* and *www.football-data.co.uk* for tennis and football, respectively.

[4] Cain et al. (2003) examine the FLB in nine win markets from eight different sports, but only the odds on Cricket (although fwith only 264 observations) seem to be affected by the bias to a greater extent than that found in our under-over 2.5 data set.

Table 1
Best Fitting Market Compositions and Respective MSEs Between the Prices Generated and Historical Data

	Tennis		Under-over 2.5		Baseball	
	RMB	PMB	RMB	PMB	RMB	PMB
Random Agents	10%	0%	45%	20%	5%	0%
Risk-loving Agents	0%	5%	5%	20%	5%	0%
Risk-averse Agents	0%	0%	35%	0%	55%	0%
Misperceiving Agents	60%	40%	15%	5%	5%	40%
Informed Agents	30%	55%	0%	55%	30%	60%
MSE	0.00003	0.00034	0.00020	0.00127	0.00017	0.00022

2.4.2 The Tennis Market

We define the *best market composition* as that distribution of agents (bettors) which fits the data best (i.e. the market composition whose generated prices minimise the MSE between these prices and those displayed in the real market). Table 1 displays the best market composition for the PMB and the RMB models for all the data sets. In tennis, both the PMB and the RMB models give similar results, namely that most of the betting volume comes from informed and misperceiving bettors.

Although at a first sight it may seem strange that 90% to 95% of the total money wagered on tennis is placed by these classes, this is consistent with Forrest and McHale (2005)'s argument . In fact, this and a subsequent paper (Forrest and Mchale, 2007) agree that most tennis bettors can be considered sophisticated for the following reasons: first, it is easier to assess the relative quality of tennis players, and thus their win probabilities, than it is to assess the relative merits of competitors in other sports (e.g. horse and greyhound racing). This arises because there are only two possible outcomes for a game and because available information such as the ATP rankings already incorporate a strong predictive power (del Corral and Prieto-Rodríguez, 2010). Hence it is easier to become an informed bettor.

Second, in contrast to team sports, there is no significant sentiment betting in tennis. Only a few players can boast a large number of fans, and even these numbers are not comparable to the millions of supporters of major football or basketball teams. Since sentiment bettors place bets regardless of the probability of a team winning, they can be included in the random bettors class. Consequently, in tennis we expect fewer random bettors.

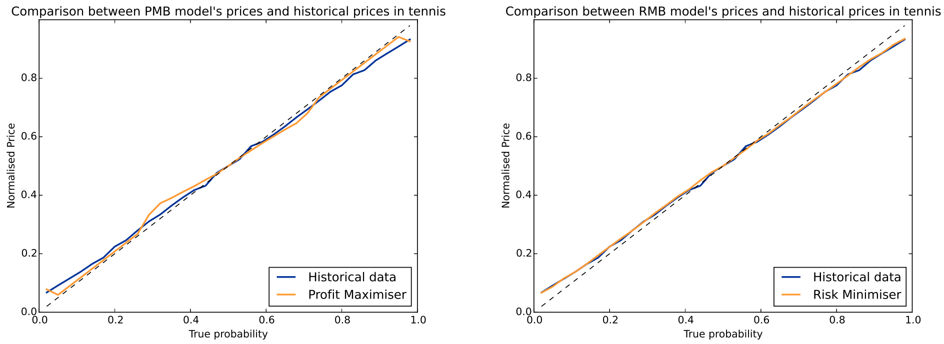


Figure 2.1: Comparison between equilibrium prices generated by the best market compositions (i.e. those market compositions for which the prices generated are closer to historical prices) in tennis and historical prices. The figure on the left represents prices fixed by a profit maximising bookmaker, while the figure on the right represents prices fixed by a risk minimising bookmaker. The dashed line represents the fair prices.

In the tennis betting market, even though both the PMB and the RMB models support similar market compositions, equilibrium prices tell a different story. The equilibrium prices fixed by the RMB give an MSE one order of magnitude smaller than that in the PMB model (i.e. 0.00003 for the RMB and 0.00034 for the PMB). Moreover, by looking at figure 2.1(a) we note that the prices generated by the PMB are closer to the diagonal, which represents the fair prices, than to the historical prices. By contrast, the equilibrium prices generated by the risk minimising boomaker simulations overlap almost perfectly with the historical data. Therefore, we conclude that the RMB model better explains the prevailing FLB in the tennis win market.

2.4.3 The Football Market

The under-over 2.5 goals market shows a very different equilibrium price curve than that related to tennis, with a more pronounced FLB. The MSEs are an order of magnitude greater than those obtained using the tennis data.[5]

[5] Part of the increment of the MSE can be attributed to the higher noise present in the under-over 2.5 data. The smaller number of observations and the shorter range of odds are likely to be the main causes of this noise. In figures 2.2(a) and 2.2(b) the effect of the noise in the historical prices curve can be seen clearly.

However, the RMB model still returns a smaller MSE than the PMB. In addition, the prices obtained from the simulation based on a PMB fail to trace the historical price curve properly. In particular, the curve fails to reproduce its peculiar S shape (see figure 2.2(a)).

The most interesting observation is that, with this data set, the simulation results suggest very different market compositions when assuming a PMB or a RMB. Previous literature and some of the features of the goal scoring process in football suggest that the RMB model is likely to give the most reasonable market composition. This is likely because, amongst other things, the number of *points* is rather limited (cf. other sports such as tennis or basketball). This means that random events (e.g. players' or referee's errors) can readily make the difference between more or less than 2.5 goals being scored. Furthermore, there is strong evidence of momentum in football (Bittner et al., 2007) implying that, after a team score, they are more likely to score again. Moreover, Armatas et al. (2007) suggest that in-play-only observable variables such as physical condition, coaches' choices, and lapses in concentration are key features in determining the number of goals scored by each team. The most significant evidence for the randomness of the goal scoring process comes from the odds on offer in these markets. In our under-over 2.5 market data the average normalised odds value is 2.03, with a standard deviation $\sigma = 0.3$, suggesting that, on average, bookmakers rate both the under and the over 2.5 goals to be almost equally likely. Therefore, we can conclude that goal scoring in football is a highly unpredictable stochastic process. As a result, it would be reasonable to expect many random bettors and few informed individuals to bet on such events.

As shown in Table 1, the risk minimisation model meets these expectations. Assuming that bookmakers follow the RMB model, the simulation suggests a market composition in which 45% of the money is bet by random bettors and no informed bettors play in the market. The PMB model, on the other hand, suggests a market composition including 55% informed bettors and no random bettors. Similarly, the incidence of risk-loving and risk-averse bettors suggested by the RMB appears more realistic than those suggested by the PMB model. In particular, the raw average odds value (i.e. not normalised) is 1.9 ($\sigma = 0.28$), which is rather low compared to other sports, especially those with more than two outcomes. This would suggest that only a few risk-loving agents would participate, while many risk-averse bettors should be attracted to bet. This is reflected in the RMB model's market composition, in which 35% of the total money is bet by risk-averse agents and only 5% comes from risk lovers. By contrast, the PMB model suggests a market with no risk-averse bet-

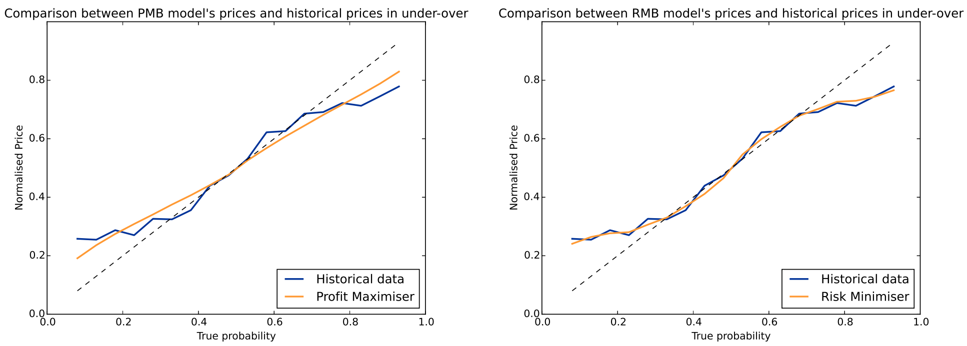


Figure 2.2: Comparison between equilibrium prices generated by the best market compositions (i.e. those market compositions for which the prices generated are closer to historical prices) in under-over 2.5 and historical prices. The figure on the left represents prices fixed by the profit maximising bookmaker, while the figure on the right represents prices fixed by the risk minimising bookmaker. The dashed line represents the fair prices.

tors and where 20% of the volume derives from the bets of risk lovers. Therefore, we conclude that the risk minimisation model is the one that gives the results that are most plausible for the under-over 2.5 market.

2.4.4 The Baseball Market

The RMB model again gives more plausible results when applied to baseball data (see Table 1). Although the MSE is similar for both the models (0.00017 for the RMB and 0.00022 for the PMB), the equilibrium prices derived by the PMB model do not reflect the odds derived from historical data. In fact, the simulation results indicate that the equilibrium prices generated by the PMB model display a positive FLB, which is unsupported by historical prices. Even though its level is rather low, this implies that the PMB model fails to give qualitatively good results. Instead, the price curive obtained by the simulation assuming bookmakers are risk minimisers, supports a negative FLB, as displayed in historical prices. The market composition derived by the simulation suggests that 35% and 55% of the money is bet by informed and risk-averse agents, respectively. The proportion of informed bettors may appear low given that baseball bettors are generally considered sophisticated (Woodland and Woodland, 1994). However, this can be explained by the odds on

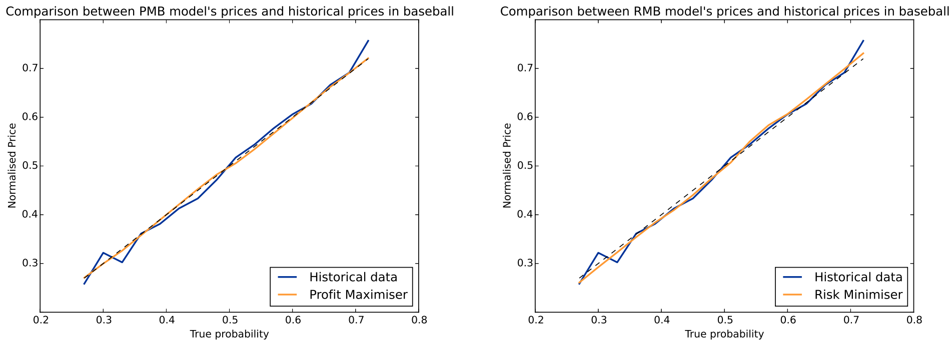


Figure 2.3: Comparison between equilibrium prices generated by the best market compositions (i.e. those market compositions for which the prices generated are closer to historical prices) in baseball and historical prices. The figure on the left represents prices fixed by the profit maximising, while the figure on the right represents prices fixed by the risk minimising. The dashed line represents the fair prices.

offer in these markets. Indeed, across the 44,675 games, the odds range between 1.3 and 4.1, and 90% of them lie between 1.59 and 2.7. With such low odds, it is no surprise that risk-averse bettors are attracted to the market.

2.4.5 Sensitivity Analysis

In Section 2.4.2 we presented empirical evidence that suggests that the average bookmaker is more likely to adopt a risk minimising (cf. profit maximising) pricing strategy. In this section we show, through a sensitivity analysis, that every agent behaviour class plays a significant role in shaping equilibrium prices. Consequently, we examine the robustness of employing representative agents, and suggest that a HAM is necessary to explain the FLB.

The above analysis indicates that certain types of bettors are attracted to particular markets. All the classes contribute to shaping price curves (as can be seen in figures 2.4, 2.5, and 2.6), but they simply bet in the markets which are best suited to maximising their utility. This is especially true for risk-loving agents, who are not present in the best fitting market composition in tennis and baseball, and only account for the 5% of the money bet on under-over 2.5. That can be intuitively explained by the fact that odds in those markets are rarely high, mainly because there are only two possible outcomes. Thus, it seems reasonable that risk lovers prefer to bet on other sports. For instance,

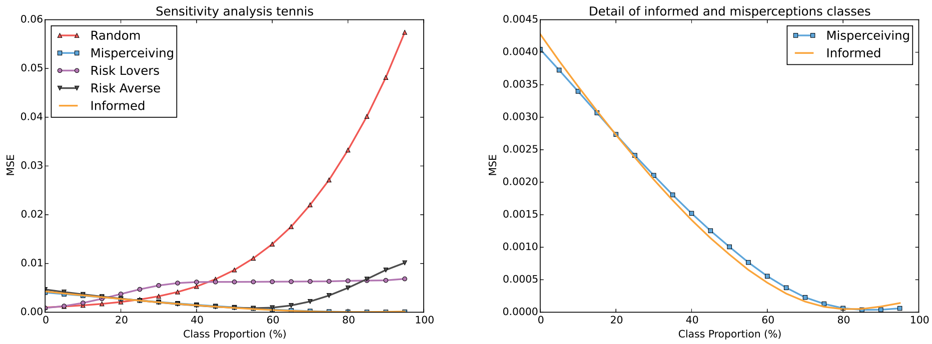


Figure 2.4: Average MSE in tennis betting markets associated with different proportions of agents from the five behaviour classes. The right-hand side figure shows in detail the curves for misperceiving and informed agents.

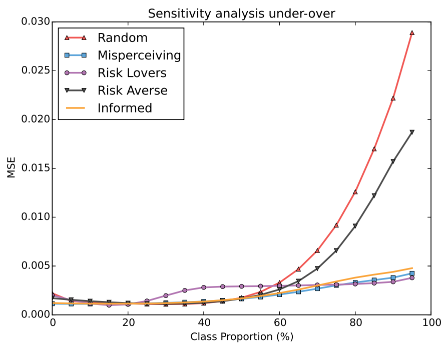


Figure 2.5: Average MSE in under-over 2.5 betting markets with different proportions of agents from the five behaviour classes.

in horse racing, strong favourites have odds similar to the average odds in the sports we examine. In addition, in horse racing it is possible to find bets paying off up to 1000 times the amount bet. In particular, risk-loving bettors are, by definition, inclined to bet on high odds, since this would return a higher utility (as described in Eq. (2.7)).

We analyse the specific contribution of each agent behaviour class under the RMB assumption. This is computed as follows: We fix the proportion of a specific agent class in the market and then we compute the average MSE of all the combinations that include this behaviour class proportion. By doing so, we can measure the contribution of each behaviour class. Then, by changing the proportion of that behaviour class we see how the MSE changes. For example, in tennis, the average MSE of all the market compositions with 0% random bettors is 0.00099, and it increases exponentially, reaching an average MSE of 0.0619 when only random bettors participate in the market. As can be seen in figure 2.4, random bettors are those that affect tennis equilibrium prices the most, increasing the MSE exponentially as they grow in number. Also, we note that misperceiving and informed bettors contribute in a similar degree to tennis equilibrium prices, but their sensitivity is significantly different. This confirms that choosing $\beta = 0.928$ as the weighting function's parameter is sufficient to model misperceptions. Figure 2.4 provides another important observation, showing that the misperceptions behaviour class would be the best to describe the FLB in a representative agent model in a fixed-odds market with a RMB. This confirms the findings of Snowberg and Wolfers (2010).

Figure 2.5 demonstrates that the random agents' MSE curve in the under-over 2.5 market is similar to that for the tennis market, but this time random agents are more likely to play a significant role in price shaping. In fact, when random agents account for 30%-45% of the money bet in the market this produces the lowest MSEs. Other low MSEs can be found when misperceiving or informed bettors are almost or completely absent, suggesting that these categories are not attracted to under-over 2.5 betting markets.

Finally, we find that risk lovers are the class of bettors that make the best representative agent in under-over 2.5 betting markets. It is interesting to note that misperceptions and risk-love, which define the two most widely employed bettor types in the literature, are the most plausible explanations with a representative agent approach in both tennis and under-over 2.5. In this sense, our work supports the claims of Snowberg and Wolfers (2010). They compare misperceptions and risk-love explanations for the FLB in parimutuel horse racing markets and conclude that misperceptions better describe the bias.

However, as a result of our sensitivity analysis, we find that the best type of representative agent changes depending on the market. A further confirmation of the necessity of heterogeneous agents is provided by bettors' sensitivity in baseball. While risk-averse bettors make up the majority of the market, accounting for 55% of the money bet, the best representative agent would be an informed bettor, with an average MSE roughly thirty times smaller than a risk-averse agent (figure 2.6). However, choosing informed agents to represent all the bettors would make the negative FLB impossible to explain.

To conclude, our sensitivity analysis provides strong evidence that HAMs are

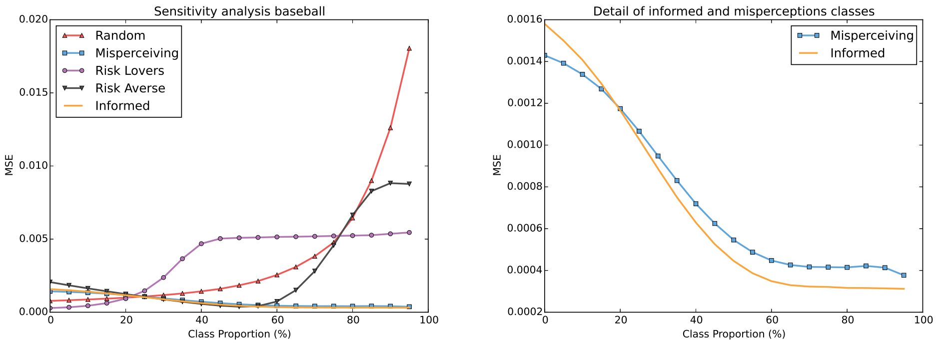


Figure 2.6: Average MSE in tennis betting markets with different proportions of agents from the five behaviour classes. Figure (b) shows in detail the curves for misperceiving and informed agents.

necessary to fully rationalise the FLB, thus enabling one model to explain the bias in different markets.

2.4.6 Comparison with the models by Shin and Snowberg and Wolfers

In Section 2.4.2 we showed that the RMB model outperforms the PMB in all the three data sets we use, generating smaller MSE and suggesting more plausible market compositions. In this section, we compare it with the models developed by Shin (1993) and Snowberg and Wolfers (2010). Snowberg and Wolfers (2010) build two representative agent models based on previous literature related to risk attitude and misperceptions of probabilities. They conclude that the latter better models the FLB. However, they do not account for information asymmetry as a possible cause of the FLB. Consequently, we also compare our results with those derived from the FLB model developed by Shin (1993). In fact, his is one of the most cited works in the information asymmetry stream and a foundation for many other papers studying market efficiency in sports betting (e.g. see Bruce et al. (2012) and Smith et al. (2006)).

We compute the MSE between our data and the prices generated by the models of Shin and Snowberg and Wolfers and compare these MSEs with those obtained by using the RMB model (see table 2). We note that all three models give a similar MSE on baseball, but the RMB model performs significantly bet-

Table 2
Comparison Between Models to explain FLB in Literature and Our RMB Model

	Mean Squared Error		
	Tennis	Under-Over 2.5	Baseball
RMB Model	0.00003	0.00024	0.00017
Snowberg and Wolfers (2010)	0.00020	0.00276	0.00046
Shin (1993)	0.00015	0.00158	0.00028

ter on the data from tennis and under-over 2.5. This suggests that our HAM is highly adaptable and can describe different markets without significantly decreasing its accuracy. By contrast, the representative agent models are tailored to specific markets, so they often only work well in conditions analogous to those assumed in the original model. This, of course, should not be surprising. Indeed, in the last decades, many economists have criticised the representative agent approach more generally because of its lack of flexibility. Another common criticism is that a representative agent approach inevitably misses some key features of complex systems, since it completely neglects interactions among different types of agents.

The MSEs of the models displayed in Table 2 show that the RMB model explains historical prices far better than the other two models. This clearly shows the advantages of adopting a heterogenous agents model to describe the FLB. In fact, we note that by using Snowberg and Wolfers (2010) we implicitly assume that all bettors misestimate probabilities in the same way, and that they are perfectly neutral towards risk. However, this is not the behaviour that it is found in many studies exploring how people gamble. In fact, after Kahneman and Tversky (1979) introduced prospect theory, many papers focused on measuring people's preferences for risk, seldom finding them to be neutral.

Similarly, applying Shin (1993) to different markets results in predictions of high incidence of insider trading. This is very surprising given that Shin's insiders are assumed to know exactly the outcome of an event. Analysis of our data using the Shin model suggests that 5% and 10% of the money bet on baseball and tennis respectively are from bettors who know the exact outcome of the game. This percentage rises dramatically to 46% in the under-over 2.5 market. These values seem unrealistically high, especially considering the highly stochastic nature of the goal scoring process, as explained in Section 2.4.2. Interestingly, we find that the percentage of money bet by Shin's insiders is inversely proportional to the percentage of money bet by informed bettors in the RMB model. This reflects the different views of the market represented in these models. Shin's insiders make the market more inefficient, while in our model informed bettors contribute to increase market efficiency.

The comparison between the results derived from our HAM and those derived from the Shin and Snowberg and Wolfers models provide a clear picture of the important difference between these representative and heterogeneous agent models. Snowberg and Wolfers (2010) build their model to explain the FLB in the horse race pari-mutuel market, which has a different structure than that in fixed-odds markets. Thus, the parameters chosen will inevitably no longer be optimal for the markets we consider. In addition, Shin makes the assumption that bookmakers set odds such that they do not make any profit or loss, thus making it difficult to compare with our model.

To overcome this problem, we run new simulations by calibrating both Shin's and Snowberg and Wolfers' models for each data set[6]. We find that errors remain of the same magnitude[7]. However, we notice interesting changes in the models' key parameters. Shin's measure of the incidence of insiders does not change significantly and remains high especially in the under-over market; his calibrated model suggesting that 45% of the money is bet by insider bettors. His calibrated model also suggests that 5% and 10% of betting in baseball and tennis, respectively, is undertaken by insiders. However, we find important changes in Snowberg and Wolfers' model. Prelec function's β values in this model are considerably lower (0.87 and 0.64 for tennis and under-over 2.5, respectively) than the value they proposed in their paper, that is $\beta = 0.928$. This implies greater misperceptions of probability in these markets compared to horse racing. This result is in line with what we expected for the under-over 2.5 market. However, it is at odds with our expectations for tennis, where the FLB is lower. In addition, this is out of line with the behaviour of tennis bettors, as discussed in Section 2.4.2.

For baseball, we find an optimal value of $\beta = 1.06$ using Snowberg and Wolfers' model. A value of $\beta > 1$ is necessary to lead to the reverse FLB in a representative agent model where the representative bettor is one who misperceives probabilities (since small probabilities are underweighted and large probabilities are overweighted). However, a value of $\beta > 1$ is out of line with theoretical work on misperceptions of probabilities, including the work of Pr-

[6] For the model by Snowberg and Wolfers we optimise the free parameter β in Prelec's probability weighting function. For Shin's model, we allow the bookmakers to seek positive profit, setting the margin they are willing to achieve to the average margin found in the historical data sets

[7] The greatest relative change in MSE values is for Snowberg and Wolfers' model in tennis (decreasing from 0.00020 to 0.00012), while the greatest absolute change occurs in Snowberg and Wolfers' model for under-over 2.5 (decreasing from 0.00276 to 0.00240)

elec himself, which suggests that $0 < \beta < 1$ must hold (Prelec, 1998). This provides further evidence that a representative agent model cannot explain all the aspects of the FLB.

2.5 Concluding Remarks and Future Work

In this paper, we seek to understand the factors that lead to a persistent pricing anomaly (FLB), with the overall aim of demonstrating the value of HAMs for understanding the true causes of such phenomena. We argue that such an approach offers the prospect of significantly improving the accuracy of forecasts derived from prediction markets. Specifically, we generalise existing theories designed to establish the causes of the FLB, and use these to develop a HAM. Our results show that a HAM is necessary to fully understand this important pricing anomaly, supporting the idea that representative agent models cannot describe complex systems in a general way with sufficient accuracy. Indeed, our comparisons with established models show that those using a representative agent cannot adapt to different markets. Furthermore, we demonstrate that our bookmaker risk minimisation model outperforms two of the most influential FLB models (i.e. Shin (1993) and Snowberg and Wolfers (2010)) in terms of reproducing real prices in three different sports betting markets and, most importantly, in explaining the regular and reverse form of the FLB.

Our results help to confirm the important role which analysing behaviour in sports betting markets can perform for understanding behaviour in other state-contingent claims markets. The importance of understanding the behaviours that cause pricing anomalies lies in that knowing the causes enables us to account for such pricing anomalies beforehand. This is fundamental in a number of contexts way beyond sports betting, such as prediction and financial markets.

Our results also suggest that our HAM is better than models based on representative agents at describing betting markets, and, most importantly, our findings provide new evidence that HAMs provide a robust way of effectively modelling pricing anomalies that are caused by heterogeneous behaviours of agents. Our results, therefore, offer the prospect of being able to appropriately adjust the prices that emerge in prediction markets to correct for pricing anomalies, thus leading to more accurate forecasts. However, a shortcoming of this model is that it is difficult to validate statistically. We advocate that future works should address this issue by identifying a procedure that can statistically validate these models.

We identify three main lines for future work that will enable us to have a more powerful forecasting tool. First, we believe it is particularly important to extend this model beyond two-outcome events, and fixed-odds markets. This would be especially useful for many internal prediction markets, which consider multi-outcome events, and political elections. Second, in future work, we would like to extend our HAM to better forecast the true probabilities of events given the corresponding market prices. For instance, knowing the circumstances of a particular prediction market (e.g., the range of odds associated with the various potential outcomes) will enable us to estimate the proportions of different agent classes likely to be attracted to the market, and consequently allow us to estimate (and thus adjust for) the degree of FLB displayed in the market prices. Finally, we plan to consider additional market-maker behaviours beyond what has been proposed in literature. Specifically, we believe it will be useful to address more complex scenarios in which bookmakers employ different, mixed strategies such as seeking maximised risk adjusted profit, and compete with each other to attract bettors.

Chapter 3

The Impact of Transaction Costs on State-Contingent Claims Mispricing

3.1 Introduction

Sports betting markets are a good proxy for financial markets and have been extensively used to analyze market efficiency. In particular, many number of studies have examined the favorite-longshot bias (FLB), an empirical regularity whereby betting on favorites yields higher expected returns than betting on longshots. This anomaly is of noteworthy importance, as it has been found in most state-contingent claims markets, such as prediction markets (Wolfers and Zitzewitz, 2004) and options markets (Hodges et al., 2003). It has also been observed in other contexts (Vaughan Williams et al., 2016).

One of the several theories proposed to explain the FLB argues that transaction costs cause the bias. Hurley and McDonough (1996) suggest that, in a parimutuel market with short-selling constraints, positive transaction costs cause the FLB. However, recent findings demonstrate the need for significant adjustments to their model. In fact, despite the possibility of short-selling in exchange markets, the FLB still exists (Smith et al., 2006). However, both transaction costs and the level of the FLB are significantly lower in such markets, suggesting correlation between the two.

We show how a model in which traders have heterogeneous behaviors can explain all these observations. To achieve this, we extend the model developed by Restocchi et al. (2018c) by incorporating transaction costs. We prove that, in a market with a book-balancing bookmaker playing against heterogeneous agents, transaction costs cannot cause the FLB. Then, we investigate price formation by using agent-based simulations. We reproduce three different levels of FLB, including its negative form, and analyze prices under different values of transaction costs. Our results suggest that transaction costs amplify the mispricing generated by the agents' sub-optimal behavior, and that they are positively correlated with the absolute level of the FLB (i.e. with market inefficiency). This finding suggests that, in contrast to the arguments of Hurley and McDonough (1996) and Vaughan Williams and Paton (1998), transaction costs are not correlated with the direction of the FLB.

3.2 Model and analysis

We consider the model introduced by Restocchi et al. (2018c), in which agents belonging to five different behavioral classes bet on a two-outcome event in a fixed-odds market. Four of the five types of agents bet according to their utility functions, which are derived from prospect theory, and represent risk loving, risk averse, misperceiving, and informed bettors. The fifth class represents noise bettors, who randomly bet on either outcome A or B. Given the outcome $i \in \{A, B\}$, we denote with p_i its probability to happen, with π_i the associated price of an Arrow-Debreu contract set by the bookmaker, and with V_i the fraction of money bet on such outcome (i.e. we assume that the total money bet, V, is $V = V_A + V_B = 1$). Then, the expected profit for the bookmaker is:

$$\mathbb{E}(P) = V - \frac{p_A}{\pi_A}V_A - \frac{p_B}{\pi_B}V_B \tag{3.1}$$

Then, the profit in the single game is defined in the following way:

$$P = \begin{cases} 1 - \frac{V_A}{\pi_A} & \text{if A occurs} \\ 1 - \frac{V_B}{\pi_B} & \text{if B occurs} \end{cases} \tag{3.2}$$

Restocchi et al. (2018c) take into account two possible pricing strategies that the bookmaker can employ, namely a profit maximization and a risk minimization strategy. Their results, obtained by using both historical data and simulations, suggest that the average bookmaker is more likely to minimize the risk (cf. maximize the profit). This behavior is similar to that of market makers in option pricing, as suggested in the relevant literature. Given the assumption that bookmakers seeks risk minimization, their optimal pricing strategy is to set both potential payoffs to be equal, i.e. $1 - \frac{V_A}{\pi_A} = 1 - \frac{V_B}{\pi_B} = \lambda$. By substituting this expression in Eq. (3.2), we obtain the expression of the prices:

$$\begin{cases} \pi_A = \frac{V_A}{1-\lambda} \\ \pi_B = \frac{V_B}{1-\lambda} \end{cases} \tag{3.3}$$

The profit λ is linked to transaction costs by the following relation:

$$TC = \pi_A + \pi_B - 1 = \frac{1}{1-\lambda} - 1 \tag{3.4}$$

Now, we formulate the following proposition: Transaction costs are not a sufficient condition to cause the FLB. In a two-outcome market the FLB exists if and only if the expected return of a bet on A is higher than that of a bet on B. Hence:

$$\frac{p_A}{p_B} > \frac{\pi_A}{\pi_B} \Leftrightarrow p_A > p_B \tag{3.5}$$

By substituting Eq. (3.3) in Eq. (3.5), we obtain the following condition for the FLB to exist:

$$\frac{p_A}{p_B} > \frac{V_A}{1-\lambda}\frac{1-\lambda}{V_B} \tag{3.6}$$

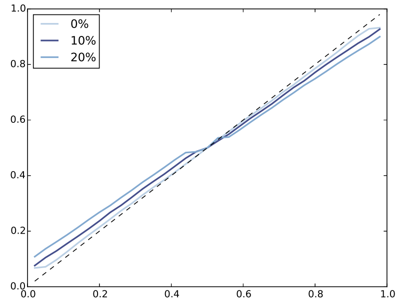
The terms including the bookmaker's profit λ cancel each other out. Hence, assuming a risk-minimizing bookmaker, the FLB (or its reverse counterpart) cannot be caused by transaction costs. Transaction costs amplify the level of the FLB and, as a consequence, market inefficiency. To test this hypothesis, we run agent-based simulations using the market compositions (i.e. distribution of agents) obtained by Restocchi et al. (2018c) for three markets with very diverse levels of the FLB, and compute market prices for transaction costs in the range 0%-50% by steps of 1%. The resulting price curves are presented in Figure 3.1 for the market compositions found for the tennis, the under-over 2.5 goals in football, and the baseball market, which display a weak, strong, and weak negative FLB, respectively.

To analyze in more detail the relation between transaction costs and mispricing, we estimate the parameters of the following equation by using an OLS regression for transaction costs, with the level of the FLB as the dependent variable:

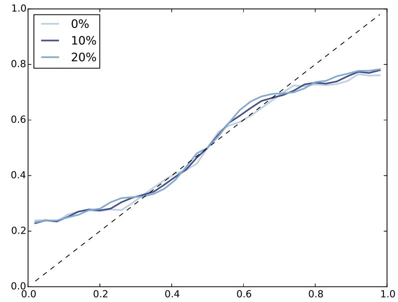
$$LFLB_i = \alpha + \beta\, TC_i + \epsilon_i \tag{3.7}$$

The level of the FLB is measured as the sum of the differences between each price and the corresponding true probability value, as follows:

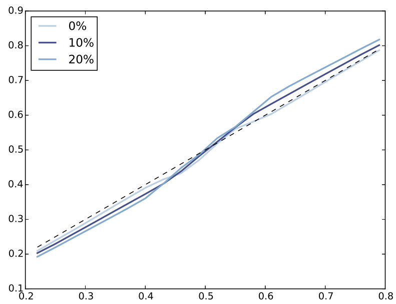
$$LFLB = \int_0^1 |\pi(p, \lambda) - p| dp \tag{3.8}$$



(a) Tennis.



(b) Under-over 2.5



(c) Baseball

Figure 3.1: Normalized market price for a given outcome given its probability, generated for the three different markets described by Restocchi et al. (2018c). They represent a weak positive (fig. 3.1a), a strong positive (fig. 3.1b), and a weak negative FLB form (fig. 3.1c). Market prices are displayed for transaction costs of 0%, 10%, and 20%.

Our results, displayed in Table 3.1, suggest that transaction costs have a direct impact on mispricing. In these regressions, the intercept value indicates the level of FLB caused by agent behavior in a market without transaction costs. It is possible to see that regression statistics, especially the R^2, are surprisingly good. However, the reason of these results can be easily explained. Transaction costs directly increase prices, and consequently increase market inefficiency, which, in our model, is captured by the FLB metric. Thus, the large goodness-of-fit was expected, and not much informative about the non-trivial impact that transaction costs have on prices. Also, prices are generated by

Table 3.1: OLS linear regressions for transaction costs with the level of FLB as dependent variable

	Weak FLB		Strong FLB		Weak negative FLB	
Variable	Coefficient (std. err.)	t	Coefficient (std. err.)	t	Coefficient (std. err.)	t
Intercept	0.8488 (0.034)	24.88*	6.4252 (0.064)	100.23*	0.7327 (0.029)	25.67*
TC	18.6904 (0.120)	155.76*	7.7950 (0.225)	34.576*	18.7383 (0.100)	186.66*
Model Statistics						
Adj. R-squared	0.998		0.961		0.999	
F-statistic	24,260		1,196		34,840	

* Significant at 0.1%

Table 3.2: OLS quadratic regressions for transaction costs with the level of FLB after price normalization as dependent variable

	Weak FLB		Strong FLB		Weak negative FLB	
Variable	Coefficient (std. err.)	t	Coefficient (std. err.)	t	Coefficient (std. err.)	t
Intercept	0.2528 (0.015)	17.26*	2.2995 (0.014)	162.61*	0.3122 (0.018)	37.81*
TC	8.1445 (0.138)	58.92*	1.1144 (0.133)	8.35*	4.1292 (0.078)	52.99*
TC^2	-4.652 (0.273)	-16.96*	-2.1902 (0.263)	-8.31*	-4.6798 (0.154)	-30.43*
Model Statistics						
Adj. R-squared	0.998		0.583		0.995	
F-statistic	14,100		35		4,745	

* Significant at 0.1%

an agent-based model, which by definition is a stylized representation of the market, and might lack of the same sources of noise. However, by examining the regression coefficients for all the markets, we can infer something else.

As shown in Table 3.1, we find that, for both the positive and negative weak forms of FLB, the coefficients of both TC and the intercepts are similar, and are all largely significant, which suggests that transaction costs are not correlated with the direction of the FLB, but only amplifies its size. Otherwise, the negative FLB would become positive in presence of sufficiently high transaction costs. This result is in contrast to what argued by Hurley and McDonough (1996) and Vaughan Williams and Paton (1998), who suggested that high transaction costs can turn a negative FLB into a positive FLB.

From the results in Table 3.1 it is also possible to see that, in the market displaying a strong form FLB, the base level of inefficiency is much higher, but the impact of transaction costs on mispricing is lower than in the other two cases. Both these observations are consistent with the fact that, in this market, 45% of the money is wagered by noise traders (Restocchi et al., 2018c). That is, noise traders generate a higher level of mispricing compared with more rational agents, but do not respond to increments in transaction costs, since they continue to trade randomly.

This substantial difference between the coefficients of the variables for the strong form FLB and the two weak ones, which possess a similar market composition, suggests that diverse market compositions respond in considerably different ways to increments of transaction cost values. To investigate this further we develop a quadratic regression to find the relation between transaction costs and normalized market prices. That is, by discounting for transaction costs and then computing the level of FLB, the direct contribution of transaction costs to prices is neglected. Therefore, it is possible to find the mispricing generated by the agents due to a variation of transaction costs:

$$LFLB_i = \alpha + \beta\, TC_i + \gamma\, TC_i^2 + \epsilon_i \tag{3.9}$$

As shown in Table 3.2, a quadratic equation fits very well the response of the agents to changes in transaction costs. Once again, we find that the R^2 are very close to 1, except from that of the strong form FLB regression. This provides additional evidence for our argument. That is, that, in evaluating the goodness-of-fit of these regressions, one should consider that they are performed on simulated data. Indeed, for the strong-form FLB market, which is significantly populated with noise traders, compared with the other markets, the R^2 is considerably lower. However, once again, the regression coefficients are informative.

In these regressions, the coefficients retain the same sign across all the data sets (positive for the intercept and the first-order term, negative for the second-order term). This implies that, regardless of the market composition, the supplementary inefficiency introduced by the agents always follows a downward concave curve. One possible explanation is that increasing transaction costs progressively deter the most informed bettors. Then, when transaction costs are high and only the least informed bettors are left in the market, the extent of mispricing provides informed bettors with profitable trading opportunities. Consequently, they push prices towards their fundamental values, thus reducing inefficiency.

3.3 Conclusion and future work

We used agent-based modeling to study the effect of transaction costs on asset mispricing. We find that, in a market in which the bookmaker minimizes risk, transaction costs alone cannot generate mispricing. However, our empirical results show that transaction costs significantly affect market prices, amplifying the effect of any existing FLB, but not affecting its direction. This is in contrast with previous literature (Hurley and McDonough, 1996; Vaughan Williams and Paton, 1998), in which it was argued that transaction costs both cause the FLB and make it positive. Furthermore, we find that transaction costs also augment market inefficiency indirectly by affecting the response of agents, increasing mispricing to an even greater extent.

We believe that the agents' response to changes in transaction costs, and the price inefficiency it generates, is worth further analysis. For instance, such a study could serve to design optimal taxation policies tailored to specific markets, by finding the best tax value for a given market composition.

Chapter 4

The Temporal Evolution of Mispricing in Prediction Markets

4.1 Introduction

In this paper we analyze mispricing in prediction markets. Prediction markets are effective tools that gauge the *wisdom of the crowd*, thus greatly improving prediction accuracy on a wide range of events (Berg et al., 2008). In fact, although prediction markets are most famous for election forecasts, often outperforming polls and experts (Wolfers and Zitzewitz, 2006), they have been recently employed by prominent companies such as Google, Microsoft, Intel and many others to improve predictions of a number of key variables, e.g., revenues, sales volume of specific products, company share price, etc. (Plott and Chen, 2002; Cowgill et al., 2009; O'Leary, 2011). Also, because they possess a definite end-point at which the outcome of an event is observed, prediction markets represent an ideal test bed to study decision making under uncertainty and investor behavior in financial markets (Croxson and Reade, 2014).

In our analysis, we focus on the favorite-longshot bias (FLB), an empirical regularity whereby bets on likely (unlikely) outcomes are underpriced (overpriced). The FLB represent an important topic of study, which has attracted researchers for decades (Griffith, 1949; Ottaviani and Sørensen, 2003; Snowberg and Wolfers, 2010), mainly because, unlikely other price anomalies, it has not disappeared in time, but it is still exhibited by betting market prices Mclean and Pontiff (2016). Previous work has shown that the FLB also exists in prediction markets (Rhode and Strumpf, 2012; Cowgill et al., 2009), and provided analysis of the correlation between this type of mispricing and time left to expiration (Page and Clemen, 2012). However, to date, there is no comprehensive analysis of how mispricing evolves over time in prediction markets depending on their duration.

In this paper, we fill this gap by showing that political prediction markets exhibit the FLB in most cases, and we characterize its temporal evolution. Our analysis complements and adds to the one by Page and Clemen (Page and Clemen, 2012). In their paper, they find that mispricing becomes significant only when the time left to expiration is sufficiently long. We differentiate our analysis by providing a complete temporal profile of mispricing (i.e, evaluating the FLB at each trading day) depending on market duration (i.e., the number of days between the start and the end of a market), which allows us to find different dynamics of mispricing over time. Specifically, in contrast to what has been observed by Page and Clemen (Page and Clemen, 2012), we find that, during the last days of trading, the FLB is significantly positive and also positively correlated with duration (i.e, the longer the market, the greater the mispricing). Also, our results suggest that, on average, markets are the more efficient the longer they last. That is, the level of the FLB averaged over the entire period is lower the longer the duration.

The remainder of this paper is organized as follows. In Section 4.2 we outline the data and methods we use to perform our analysis, and in Section 4.3 we show and describe our findings. In Section 4.4 we discuss our results and argue that the FLB in the last days is caused by herding.

4.2 Data and Method

To perform our analysis, we use end-of-day prices from 3363 markets, from October 2014 to November 2016, and constitutes all markets traded on PredicIt[1], a platform that allows betting on the outcomes of political events, during this period.. Each market represents a possible outcome on a political event, and it is possible to buy or sell a contract at a price $0 < \pi < 1$. Such a contract pays 1 dollar if the selected outcome occurs, and 0 dollars otherwise. For example, a user who buys a "Clinton will lead" contract on "Who will lead in Trump vs. Clinton polling on September 14?", will pay π dollars to the seller, who, in turn, at the end of the market (in this case, 14 September 2016), will pay the buyer 1 dollar if Clinton was leading the polls, and nothing otherwise.

By definition, the FLB is observed if prices on the favorite (longshot) are lower (higher) than the corresponding true probability of the outcome to occur. However, it is fair to assume that it is impossible to know the *fair price* of a single contract (i.e., the true probability of the associated outcome to oc-

[1] www.predictit.org

cur). To address this issue, we analyze all markets by comparing the price on a given day with the final outcome. By aggregating these results, we can see the discrepancy between the realized frequencies of positive outcomes and corresponding prices.

The FLB exists if, for low-probability outcomes, prices are systematically higher than the true probability of the associated outcome to occur, and lower otherwise. Formally, prices exhibit the FLB if:

$$\frac{\pi_A}{\pi_B} < \frac{p_A}{p_B} \iff p_A > p_B \tag{4.1}$$

where π_A and π_B represent the price to buy a contract on outcomes A and B, and p_A and p_B represent the true probabilities of such outcomes to occur. In our analysis, since we cannot know a priori the true probability of an event, p_i is the observed frequency of positive outcomes for contracts with prices $\pi = i$, with $0 < i < 1$. For a sufficiently large sample, p_i tends to the average true probability associated with price π_i, thus it is possible to represent the FLB with the following measure:

$$\Phi = \Phi_- + \Phi_+ \tag{4.2}$$

where

$$\Phi_- = \sum_{i=0}^{i=0.49} \pi_i - p_i \tag{4.3}$$

$$\Phi_+ = \sum_{i=0.51}^{i=1} p_i - \pi_i \tag{4.4}$$

where Φ_+ and Φ_- represent the cumulative difference between prices and true probabilities, for high and low probability events respectively. To reduce noise in the estimation of the FLB, we allow increments of i to vary according to the number of observation we possess at a given point in time (i.e., a given number of days before the end of the market). Such increments are always between 0.02 and 0.1.

4.3 Empirical Analysis

In this section we examine the temporal dynamics of the FLB. We start by showing that there exists an FLB in political prediction markets when we average over the entire time period. We then examine this in more detail by

Duration (τ, in days)	**Markets**	**Average Φ**	**St.Dev. Φ**	Φ_{last}	**t-test**
$\tau > 0$	3363	0.025	0.007	0.061	15.92*
$0 < \tau \leqslant 10$	1784	0.038	0.009	0.058	12.30*
$10 < \tau \leqslant 25$	600	0.068	0.015	0.089	13.87*
$25 < \tau \leqslant 50$	435	0.039	0.042	0.136	4.51*
$\tau > 50$	544	0.008	0.030	0.151	1.92**

* = significant at the 0.01%.
** = not significant.

Table 4.1: Summary statistics for the level of FLB Φ, average value of Φ the day before the end of the market, and t-test statistic for the temporal evolution of Φ.

grouping the markets depending on their duration. Following this, we examine prices at different points in time (i.e., different number of days to expiration), which enables analysis of the temporal evolution of the FLB (see Fig. 4.1). In these analyses, we consider all markets with at least one day of trading activity in our data set, resulting in a total of 3363 markets.

Specifically, we examine prices depending on the number of days to the end of the market, which enables analysis on the temporal evolution of the FLB (see Fig. 4.1). We find that the average level of the FLB throughout the last 50 days of trading is small but positive, at $\Phi = 0.025 \pm 0.007$. Hence, we perform a t-test to check whether the time series of Φ is significantly different than random zero. These results, shown in Table 4.1, suggest that there exists a positive FLB in political prediction markets, and that, on average, its level grows significantly during the last 15 days of trading.

To gain more insights, we also group markets depending on their duration and analyze the FLB separately for each group. Results are shown in Table 1 and Fig. 2. We find that prices exhibit the FLB with a high level of confidence for all market groups except those longer than 50 days, for which a low score in the t-test (0.224) and a high p-value (0.06) do not allow us to reject the null hypothesis that mispricing is just noise with a 5% significance. However, we find that, if we measure the FLB within the 24 hours before expiration[2] (denoted with Φ_{last}), this market group shows the highest level of the FLB among all the groups considered. In fact, our results suggest that Φ_{last} increases with the

[2]To perform our analysis, we use end-of-day prices. Since each market has a definite and different end point during the last day, to measure the FLB in the last trading day we use the last available end-of-day price, which is the close price of the day before the event occurs. Consequently, the FLB measured in this way is the mispricing that exists 0-24 hours before the end of the market

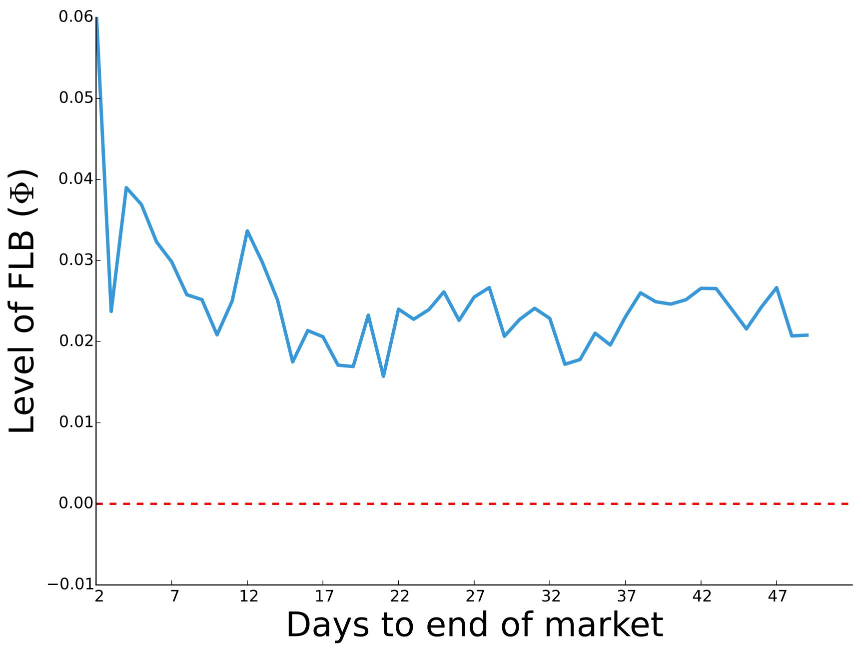
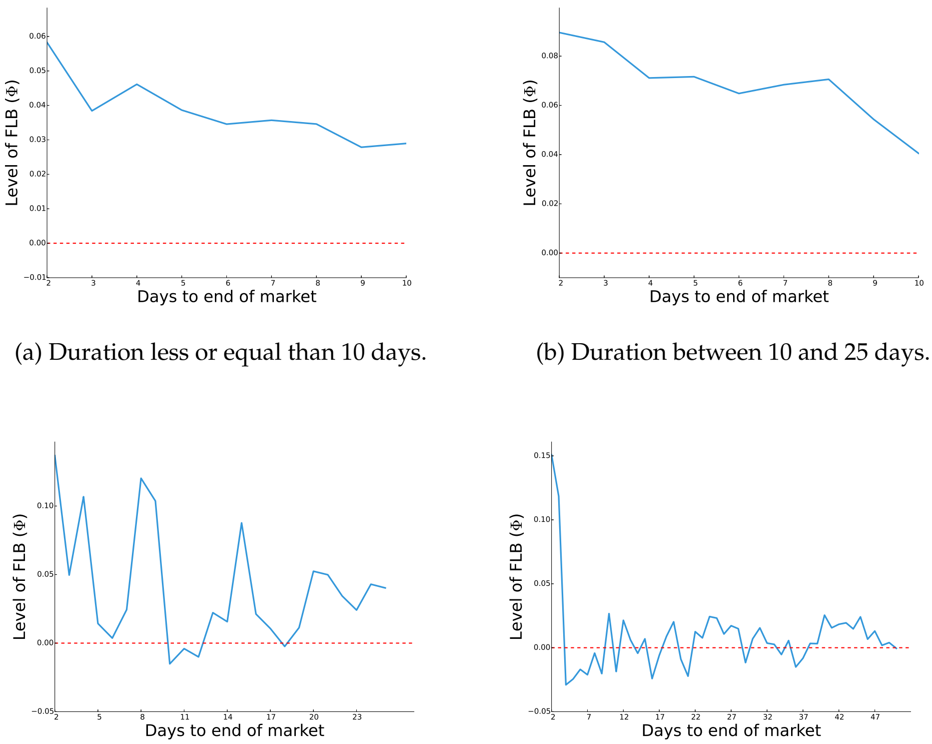


Figure 4.1: The level of the FLB depending on the number of trading days left to the end of the market, computed considering all markets.

duration of the market, growing from $\Phi_{last} = 0.058$ for markets that last no more than 10 days, to $\Phi_{last} = 0.151$ for markets longer than 50 days. From Table 4.1, it is also possible to see that long markets display, on average, lower levels of the FLB, and thus they exhibit little mispricing. Specifically, except from very short markets, our results suggest a negative correlation between the average mispricing and that of the last days of trading.

These observations are consistent with the finding that, in prediction markets, volumes grow exponentially towards the end of the market. In the data we analyze, short markets are usually based on events of lesser importance, that relate to main events for which it is possible to trade for months. For example, markets such as *"Will Democrats lead by 2.9% or less, or tie or trail Republicans, in generic congressional polling on November 7?"* are open only for a few days and exhibit low liquidity, whereas those on more significsant events such as *"Will Hillary Clinton win the 2016 Democratic presidential nomination?"* can last



(a) Duration less or equal than 10 days.

(b) Duration between 10 and 25 days.

(c) Duration between 25 and 50 days.

(d) Duration longer than 50 days.

Figure 4.2: Level of the FLB for different market durations.

for a considerably longer time (in this case, two years) and attract a significant amount of bets. Given this, it is possible that the liquidity of longer-term markets is sufficient to reduce or eliminate the FLB. Also, towards the end of the market, important events are extensively covered by both traditional and social media, and it is not rare that for such events, prediction markets prices are also reported. This may trigger herding behavior, and attract people who are not rational and attach a higher probability of realization to the most probable outcome. In this scenario, rational and informed traders would not be able to push the price back to its fundamental value, due to time and liquidity constraints (Lux and Marchesi, 1999a), while this is possible during the rest of the market, when liquidity is lower and time is not a constraint to move the price.

4.4 Conclusions

We analyzed political prediction markets and showed that their prices exhibit a regular mispricing pattern, known as the favorite-longshot bias (FLB). To perform a more detailed analysis of the mispricing, we group them by duration, which allows us to find that the FLB is both related to the time left to expiration and the market duration. In particular, we observe two important trends. First, we find that, on average, the level of the FLB, when taken over the entire period, is higher the shorter the market duration. Second, we find that this dynamic changes during the last 24 hours of trading, when the FLB becomes larger the longer is the market duration, i.e., there is positive correlation between the two quantities. We argue that these observations are consistent with herding behavior and the fact that market efficiency increases the higher the trading volumes. Indeed, on the one hand, prediction markets that allow trading on the outcomes of important events typically have a longer duration, and, on average, important events attract more money (i.e., a higher daily volume). Consequently, the extent of the mispricing in long markets is expected to be lower, as markets which exhibit high volumes attract more informed bettors, who use private information to improve price efficiency. On the other hand, important events receive far more attention than unimportant ones, and gain increasingly more media exposure the closer they get to the end. Moreover, in the recent past, the media use prediction markets to help their forecasts of the probability of specific events to occur. This attracts more bettors to prediction markets, potentially causing herding behavior in the last 24 hours of trading, making it difficult for informed bettors to push the price back to its fundamental value before the end of the market, due to shortage of time. This explanation is in agreement with our observation that, during the last trading day, the FLB is positively correlated to market duration, which is in stark contrast to the findings of Page and Clemen (Page and Clemen, 2012) that FLB decreases overtime. Since Page and Clemen use an older data set in their analysis, and consider markets from 10-15 years before ours, we argue that one potential explanation of these conflicting results may be that, nowadays, media make more conspicuous use of prediction markets to support their forecasts, thereby attracting more late bettors which engenders herding. We believe this is an important and interesting issue that requires further investigation in future work, to better understand the mechanisms that generate mispricing in prediction markets, with the ultimate goal of improving predictive power.

Chapter 5

The stylized facts of prediciton markets: analysis of price changes

5.1 Introduction

Prediction markets are markets that enable trading on the outcomes of events. These markets are heralded as effective tools for making accurate forecasts, by harnessing the *wisdom of the crowd* (Berg et al., 2008), and are used to predict a number of diverse events. For instance, there exist public prediction markets that allow everyone to trade on political events (e.g. elections, referendum outcomes) or sports events (e.g. football, horse racing, tennis), and private prediction markets that some companies such as Google, Intel, and General Electric use to forecast a variety of business activities such as product sales or the likelihood of meeting project deadlines (Plott and Chen, 2002; Cowgill et al., 2009). Moreover, prediction markets constitute an excellent laboratory to test theories on decision making and financial markets, as they have characteristics which facilitate analysis. Importantly, they have a definite end-point at which all uncertainty is resolved. However, whereas in financial markets there has been extensive data-driven analysis of stylized facts (Mantegna and Stanley, 2000), i.e., empirical regularities observed in most financial time series, there is no such work for prediction markets, mainly due to historical insufficiency of data. Consequently, many of the models of prediction markets lack robust empirical validation.

To this end, in this paper we compile a set of stylized facts for predictions markets using a dataset from PredictIt[1], containing the trades and outcomes of 3385 prediction markets on political events. Our analysis results in three main contributions. First, we show that percentage and logarithmic returns, which are commonly used to analyze price dynamics in financial markets, are inadequate to describe prediction markets, and demonstrate that raw returns

[1]https://www.predictit.org/

possess characteristics that best suit time-series analysis. Second, we characterize the statistical properties of the distribution of price changes, focusing on the power-law behavior of the tails, the long-range memory of returns and the gain/loss asymmetry. Third, we analyze linear and non-linear time-dependence properties of price changes. Specifically, we examine the autocorrelation of returns, the volatility clustering regularity, and the leverage effect.

Overall, we find that prediction markets behave similarly to emerging financial markets, despite their differences in structure and participants. However, to account for the differences between prediction markets and established financial markets, both from an empirical and a structural point of view (e.g. the fixed time horizon and the binary nature of the payoff), we suggest that new models, particularly agent-based ones, should be designed to examine prediction markets in more depth, with the goal of improving their predictive power. We believe that the use of this set of stylized facts will provide an additional, more robust layer of validation for such quantitative models.

Moreover, the analysis we provide in this paper extends the boundaries of the Econophyiscs literature, which has historically revolved around financial markets and financial economics (see Jovanovic and Schinckus (2017), and Richmond et al. (2013)), and shows the potential of applying tools from this discipline to the new types of financial markets that have been growing in the recent past, such as prediction markets and cryptocurrencies (Bariviera et al., 2017).

This paper is organized as follows. In Section 5.2 we present our data set and outline the methods we employed to conduct analysis on it. In Section 5.3 we show in detail why raw returns, and not log returns, should be used to examine price time series in prediction markets. Section 5.4 and 5.5 present our analysis of the distribution and time dependence properties of returns, respectively. We discuss results and suggest directions of future work in Section 5.6.

5.2 Methods and Data

As mentioned, we analyze 3385 betting markets on political events from PredictIt, for which we have OLHC prices for each day of the market. To perform our analysis we use the changes of the close price, for a total of 109173 observations of daily returns. To examine the properties of the distribution of returns, we aggregate returns from all the markets in our data, because most of them provide too few daily observations to guarantee a statistically significant distribution if considered alone. In fact, even compared with the time series of other state-contingent claim financial instruments, those of prediction markets are much shorter. Half of these markets last less than 15 days, and 75% less than 43 days. Fig. 5.1 displays the distribution of the market duration (in days). For all other statistical properties we analyze (i.e., volatility clustering, autocorrelation of returns, etc.), we first compute them market by market, and then take the average across all of them.

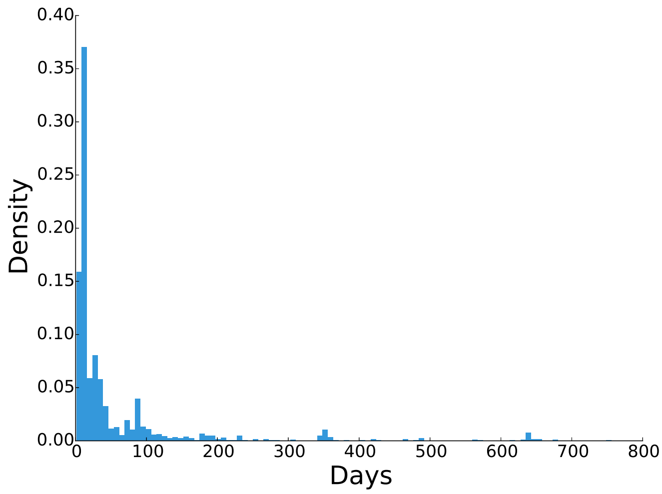


Figure 5.1: Distribution of market durations in days.

5.3 Analysis of returns

In this section we discuss the procedure for analyzing returns in prediction markets and, in particular, we highlight the differences with the analysis of financial time series.

On the PredictIt platform, contracts are linked to the realization of a particular outcome of a future polotical event. For instance, for the market *"Who will lead in Trump vs. Clinton polling on September 14?"*, one could buy a contract on the outcome *Trump will lead* or on the outcome *Clinton will lead* to occur at the market price, which always lies between 0 and 1 dollars. These contracts are

Arrow-Debreu securities, i.e., they pay 1 dollar if a particular state is realized at a given time and 0 otherwise, and are tradable at market price until the expiry date, which always coincide with the end of the market itself.

The price of such a contract at time t is denoted by P_t, where t always coincides with the end of a period Δt. Throughout our analysis, we only consider a daily time scale, i.e., $\Delta t = 1$day. As a consequence, all prices P_t refer to the close price of day t. We also define $P_{t-\tau}$ as the price of the security τ days before t. Then, the daily price change (or *raw return*) at time t is

$$r_t = P_t - P_{t-1} \tag{5.1}$$

Similarly, the log return at the same time is:

$$r_t^l = ln(P_t) - ln(P_{t-1}) \tag{5.2}$$

and, finally, the percentage (or relative) return is:

$$r_t^{\%} = \frac{P_t - P_{t-1}}{P_t} \tag{5.3}$$

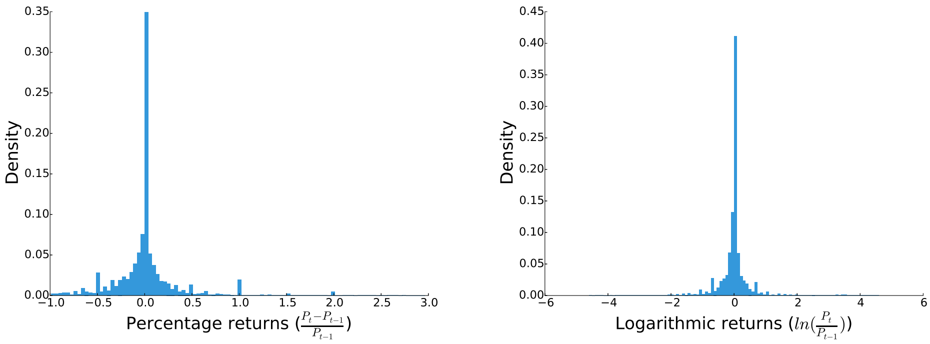
5.3.1 Why raw returns

Log returns are used to analyze financial time series, since this has a number of advantages over using raw or percentage returns. However, in this section we explain why we advocate the use of raw returns in prediction markets analysis, by showing that, for Arrow-Debreu securities, the use of raw returns suits the analysis of price time series far better than log returns.

To ensure the correctness of financial time series analysis, there are three requirements that returns must meet. First, price changes must be comparable across different time series, because otherwise it would not be possible to compare two stocks with a significantly different price magnitude (e.g., two stocks priced at 100$ and 1$). Second, the distribution of returns needs to be symmetric around zero, to allow the comparison of negative and positive returns. Third, returns need to aggregate over time. That is, if $r_{t,T}$ is the return between time t and time T, and $0 < t < T$, the following must hold to ensure that analysis can be comparable at different time scales:

$$r_{0,T} = r_{0,t} + r_{t,T} \tag{5.4}$$

In financial time series, log returns, which provide an approximation for percentage returns, satisfy all these three conditions, whereas percentage returns



(a) (b)

Figure 5.2: Figure (a) displays the distribution of relative returns, for $-1 \leqslant r_t^{\%} \leqslant 3$. Figure (b) displays the distribution of log returns r_t^l.

and raw returns do not. Conversely, in prediction markets, since contracts are Arrow-Debreu securities, raw returns possess properties that guarantee that all requirements are met. In more detail, Arrow-Debreu securities (also known as state-price securities) are contracts that pay one unit of a currency or a commodity if a particular state occurs at a specific point-in-time, and pay zero otherwise. Consequently, Arrow-Debreu securities are always priced between 0 and 1 and, therefore, the price change boundaries are symmetric, i.e., $1 < r_t < 1$. Furthermore, since all contracts are priced according to the same boundaries, prices and their changes can be easily compared across different markets. Finally, the linearity of raw returns assures that they do aggregate over time.

Although these properties are also common to log returns, raw returns possess two important advantages in the analysis of time series of Arrow-Debreu securities. First, there is no need for approximations. For example, the approximation from percentage returns to log returns holds almost always in the stock market, because percentage returns often fall in the neighborhood of 0, but this is not the case for prediction markets. For instance, for $|r_t^{\%}| < 0.1$, we have that $0.95 < \frac{r_t^l}{r_t^{\%}} < 1.05$ (see Fig. 5.4). If $|r^{\%}| > 0.1$, the approximation by which $r_t^l \approx r^{\%}$ becomes weak, since log returns differ more than the 5% from percentage returns. However, in our data, we find that 42.8% of percentage price changes fall outside the optimal range for the approximation to hold (i.e., $|r_t| > 0.1$). Second, percentage and log returns can take only a limited

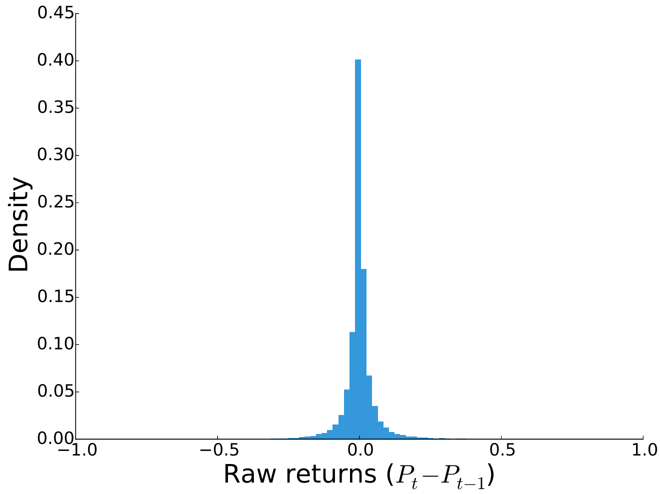


Figure 5.3: Distribution of raw returns.

range of values, and some returns occur too often. This is a consequence of the market structure, as all contracts are priced between 0 and 1, and the minimum value for price changes is 0.01. The implications are twofold: first, some percentage return values cannot occur. For example, if at time $t-1$ the price is $P_{t-1} = 0.02$, next day's percentage return $r_t^{\%}$ must be $r_t^{\%} \leqslant 0.5 \vee r_t^{\%} \geqslant 0.5$, unless $P_{t-1} = P_t$. For the same reason, some values occur more often than others. In the example above, $r_t^{\%} = 0.5$ and $r_t^{\%} = -0.5$ would be the most common returns, because generated by the smallest possible price change. Second, percentage returns values are bounded between $-1 \leqslant r_t^{\%} \leqslant 99$, generating a strongly asymmetric and irregular distribution as shown in Fig. 5.2(a). In this case, using log returns only solves the latter problem (see Fig.5.2(b)).

For these reasons, it is not possible to use percentage returns to analyze the statistical properties of prediction markets' time series. Similarly, log returns are asymptotic to relative returns in the neighbourhood of zero, and, as displayed in Fig. 5.2(b), some values occur too often. On the contrary, in Arrow-Debreu securities, raw returns do not possess any of these issues (see Fig. 5.3). Specifically, price changes are bounded by $-1 < r_t < 1$, which provides strict, symmetric boundaries, thus facilitating analysis. Also, no return value is impossible (that is, considering that the minimum price change is $|r_t| = 0.01$), which results in more informative distributions of returns. For these reasons,

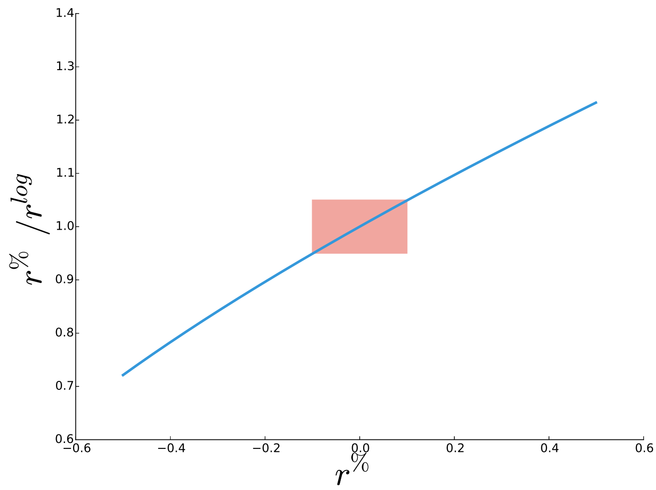


Figure 5.4: Function of the ratio $\frac{r^{\%}}{r^{l}}$ depending on the value of percentage returns. The shaded area represents the range in which $r_t^l = r_t^{\%} \pm 5\%$ holds, and corresponds to $|r^{\%}| \leqslant 0.1$.

in this paper we study the properties of price changes by analyzing the raw returns, and we advocate the usage of raw returns for all quantitative studies of prediction markets and, more generally, of Arrow-Debreu securities.

5.4 Return distribution and heavy tails

In this section we analyze the unconditional distribution of returns which, in the stock market, has been a topic of major interest for more than a century (Bachelier, 1900). Although the probability density function (PDF) of returns has been commonly modeled as Gaussian, evidence of heavy tails can at least date back to the 1960s (Mandelbrot, 1963) and, nowadays, heavy tails in the PDF are one of the most important properties a model has to reproduce for validation (Lux and Marchesi, 1999b). Since a PDF with heavy tails is defined as a PDF whose decay is slower than exponential, by definition a heavy-tailed distribution cannot be Gaussian. However, identifying the distribution of financial returns is a difficult task. In fact, return distributions vary across assets and time scales, and many statistical distributions have common properties,

Table 5.1: Summary statistics for the distribution of raw returns r_t from 3385 prediction market contracts from the website PredictIt.

N.Observations	Mean	c_v (μ/σ)	Skewness	Kurtosis
109173	0.001	0.013	1.593	39.31

such as negative skewness (positive price changes occur more, but negative returns are, on average, larger) and high, positive values of kurtosis (the distribution displays heavy tails and its center is taller than that of the Gaussian distribution) (Peiro, 1999). For this reason, many distributions whose characteristics satisfy such statistical properties have been suggested (Mandelbrot, 1963; Fama, 1965; Praetz, 1972; Blattberg and Gonedes, 1974; Cont et al., 1997; Bucsa et al., 2011). In our data, we find that the descriptive statistics (displayed in Table 6.1) of the return distribution in prediction markets show some differences to those of financial time series. Specifically, the values of mean, kurtosis, and the variation coefficient $c_v = \frac{\mu}{\sigma}$ are in line with what is found in financial times series (Campbell et al., 1997; Cont, 2001). For the skewness we find that, although it has an absolute value similar to those found in the literature, it is opposite in sign (Peiro, 1999). This means that, opposite to financial time series, price changes in prediction markets are more likely to be negative, but upward movements of the market tend to be stronger. We further discuss this observation in Section 5.4.3.

The high value of the kurtosis suggests that the distribution of returns might be heavy tailed. We analyze the distributions of positive and negative returns separately, and we use the maximum likelihood estimation (MLE) to estimate the tail exponent for both our distributions, and find that the power law accurately fits the tails.

5.4.1 Fitting procedure and goodness-of-fit

Among the proposed distributions, the power law is the most commonly used to describe the behavior of the tails of the distribution of returns in financial time series, although the question of which distribution is the correct one and why is still open (Malevergne et al., 2005). A random variable x is said to follow a power law distribution if:

$$p(x) \propto x^{-\alpha} \tag{5.5}$$

where α is the power-law *exponent* (or *scaling parameter*), which is found to be between two and five in most financial returns distributions (Plerou et al., 2000; Mantegna and Stanley, 2000; Cont, 2001). Graphical methods are often used to fit the data to a power-law distribution, thanks to their practicality. These methods consist in fitting the logarithm of the empirical probability distribution with a straight line by using the linear least squares method to determine its slope, which represents the power-law exponent $-\alpha$. However, these methods produce a poor estimate and exhibit large errors, even when improved by using logarithmic binning to estimate the density of the distribution (Bauke, 2007; Clauset et al., 2009). Although more complicated to implement, the maximum likelihood estimation (MLE) is found to be an exceptionally reliable and accurate method to fit empirical data to a power-law distribution (Bauke, 2007; Deluca and Corral, 2013).

In our case, we deal with values that lie in the range $|r_t| \in [0.01, 0.02, ..., 0.99]$, which means that our distributions has discrete values that do not belong to the set of natural numbers $\mathbb{N}$. This raises a number of issues with the fitting procedure. First, for discrete distributions, $\hat{\alpha}$ cannot be found analytically (Bauke, 2007). Second, the normalization constant for the discrete power law PDF contains the Hurwitz zeta function, defined as

$$\zeta(\alpha, x_{min}) = \sum_{i=0}^{\infty} \frac{1}{(i + x_{min})^{\alpha}} \tag{5.6}$$

whose value can only be found numerically. Third, our return distributions are characterized by another constraint. That is, since $0.01 \leqslant |r_t| \leqslant 0.99$, we need to add an upper bound $x_{max} = 0.99$, which complicates the PDF of the corresponding power law (Bauke, 2007). Last, although the values of r_t are discrete, they do not belong to the set of natural numbers $\mathbb{N}$, which is a requisite for fitting discrete power law distributions.

Although there is a number of methods to fit power law distributions to empirical data (e.g., Clauset et al. (2009), Ausloos (2014)), to overcome these problems, we propose a solution mostly based on the work of Bauke (Bauke, 2007), whose proposed method has been shown to work well for discrete distributions with natural upper bounds. To this end, we first need to convert raw returns from decimal values to integer, which we achieve by multiplying our returns by a constant $c = 10^2$. Then, if we adopt the notation $x_i = c \cdot r_t$ we have that $x_i \in [1, 99]$. Note that the exponent of the distribution, α, does not change its value after this transformation. The second step of the procedure is to find the PDF for a discrete power law distribution in which $x_i \in \mathbb{N}$ and

$x_i \in [x_{min}, x_{max}]$. Such distribution is given by:

$$p(x) = \frac{x^{-\alpha}}{\Delta\zeta} \tag{5.7}$$

where

$$\Delta\zeta \equiv \zeta(\alpha, x_{min}) - \zeta(\alpha, x_{max}) \tag{5.8}$$

Then, it is possible to compute the likelihood function for $p(x)$, which is given by

$$L(\alpha) = -\alpha\left(\sum_{i=0}^{N} ln(x_i)\right) - Nln\big(\Delta\zeta\big)\big) \tag{5.9}$$

Then, the maximum likelihood estimator, $\hat{\alpha}$ is given by:

$$\hat{\alpha} = \underset{\alpha}{\text{argmax}}[L(\alpha)] \tag{5.10}$$

Since, in this case, there exists no closed-form solution for $\hat{\alpha}$, we find the value that maximizes Eq. (6.5) numerically.

5.4.2 Estimation of the lower bound

The value $x_{max} = 99$ is the natural upper bound of our distributions, but we need to find which value x_{min} is optimal to for the fitting procedure (i.e., at which value of x our return distributions start behaving like a power law), since this is paramount to achieve a good fit, as found by Clauset et al. (Clauset et al., 2009). To estimate x_{min}, we follow the method they propose. Specifically, we perform a two-sample Kolmogorov-Smirnov (KS) test on our dataset, restricted to values $x_{min} < x < x_{max}$, and a sample drawn from a power law distribution with exponent $\alpha = \hat{\alpha}$, and evaluated in the same range. We repeat the same procedure for all possible values of x_{min}. Clauset et al. suggest that the optimal value of x_{min} is the one that minimizes D, the value of the KS statistic.

Fig. 5.5 shows the values of $\hat{\alpha}$ and D depending on x_{min}. D has a minimum in $x_{min} = 9$ for positive returns and in $x_{min} = 14$ for negative returns, which implies that the distributions mostly behave as power laws. Now, by using these values of x_{min}, we find that the distribution of positive returns is best fit by a power law with exponent $\alpha = 2.33$ whereas the distribution of negative

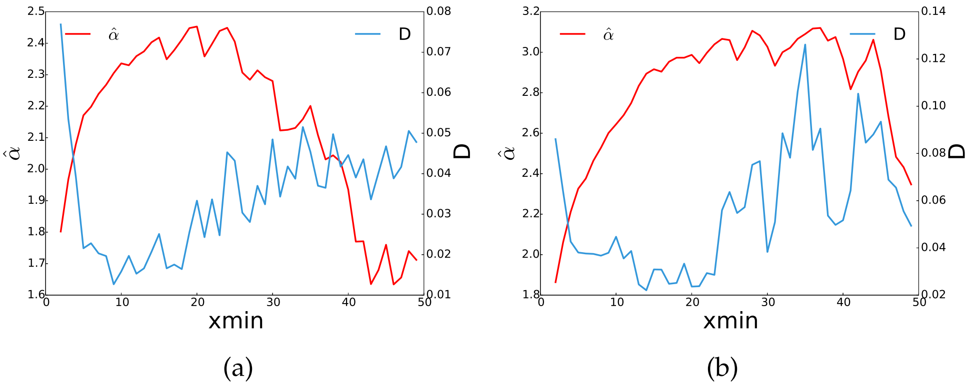


Figure 5.5: Estimated values of α and corresponding results of the KS test for a given x_{min}. Note that the KS statistic D has a minimum in $x_{min} = 9$ for positive returns **(a)** and a minimum in $x_{min} = 14$ for negative returns **(b)**.

returns is best fit by a power law with exponent $\alpha = 2.904$. Table 5.2 summarizes these results, and also displays the asymptotic variance of the estimations $\Delta\hat{\alpha}^2$, which gives us a measure of the error, and is given by

$$\Delta\hat{\alpha}^2 = \frac{1}{n}\frac{\Delta\zeta^2}{\Delta\zeta''\Delta\zeta^2 - \Delta\zeta'^2} \tag{5.11}$$

The values we found for α are in line to those of financial asset returns. In fact, for financial time series, this value usually lies in the range $2 < \alpha < 5$ (Lux, 1996; Plerou et al., 2000; Peiro, 1999).

5.4.3 Gain/loss asymmetry

We find two significantly different values $\hat{\alpha}$ for positive and negative returns. As we showed in Table 6.1, the number of observations and the value of the skewness (Table 6.1) are also different. This suggests that the distributions of positive and negative returns might be significantly different. To check this, we perform a two sample KS test between the two distributions. The null hypothesis, i.e., that the negative and positive returns are distributed in the same way, is rejected if either the p-value is small enough or if the statistics D is greater than the critical value $D_c = c(\alpha)\sqrt{\frac{n_1+n_2}{n_1 n_2}}$, where n_1 and n_2 are the numbers of observations in our distributions, and $c(\alpha)$ is a constant that represents the inverse of the Kolmogorov distribution at a level α of significance.

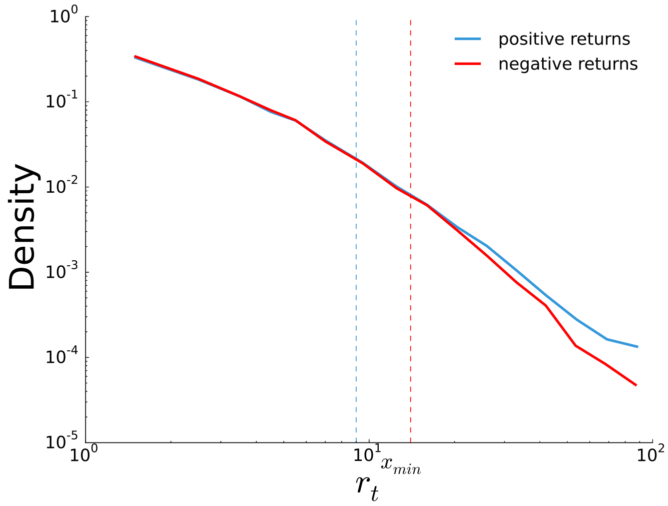


Figure 5.6: Comparison of the PDFs between positive and negative returns.

We compute those values and find that $D = 0.017$, $D_c = 0.011$ and $p = 10^{-5}$. These results confirm that the two return distributions are significantly different and, as it is possible to see in Fig. 5.6, the difference is mainly found in the tails ($x > 10$).

This result is non-trivial. In fact, it confirms that the distribution of returns is right skewed. This implies that negative price changes are more frequent, but positive returns are bigger. On the contrary, in financial time series the opposite is usually true (Peiro, 1999; Cont, 2001). This phenomenon, which is referred to as the *gain/loss asymmetry* and has been observed many times in financial markets (Jensen et al., 2003, 2004), and is one of the few properties of prediction markets time series that considerably differ from their financial markets counterpart. One possible explanation might be due to low liquidity. Indeed, this property heavily depends on the time scale considered, which is, among other things, indicative of liquidity levels. Also, Karpio et al. (2007) found that emerging markets, which possess low liquidity, display an asymmetry which is opposite to that of established markets. Prediction markets, due to low liquidity and short life span, can be compared with emerging markets, which we believe is the reason why they show an inverse asymmetry.

Table 5.2: Power law fitting of the return distributions.

	N.Observations	x_{min}	x_{max}	$\hat{\alpha}$	$\Delta\hat{\alpha}$
Positive returns	38316	9	99	2.330	0.001
Negative returns	40905	14	99	2.904	0.012

5.4.4 Analysis of the Hurst exponent

In this section we examine the Hurst exponent H, a measure of long-term memory of time series introduced by Hurst (1951), which represents a synthetic measure that encapsulates all the information regarding the long-range memory of a time series. Our analysis of H provides further confirmation of the presence of long-range memory in prediction markets. That is, $0 < H < 0.5$ indicates that the time series is mean-reverting, $0.5 > H > 1$ that the time series has long-range memory (trending), and $H = 0.5$ corresponds to a perfect Brownian motion (i.e., in the case of financial time series, returns are random and the market is perfectly efficient). To ensure an accurate and significant estimation of H, we decide to include in our analysis only the 800 markets in our data set that last for more than 30 days.

There are several methods to compute the Hurst exponent. The most common method is perhaps the rescaled range analysis (R/S analysis), which is based on the R/S statistic (Mandelbrot, 1972). However, in the past this test has been criticized. Notably, Lo (1991) examines the results by Mandelbrot (1971) and Greene and Fielitz (1977) and concludes that this test is sometimes unable to discriminate between short and long memory, proposing a modified version of the R/S statistic. However, Lo's work itself has been subject to criticism. For instance, Teverovsky et al. (1999) showed with synthetic data that Lo's modified R/S test is too strict, and sometimes fails at rejecting the null hypothesis that a time series does not display long-range memory. Another method to compute the Hurst exponent is the Detrended Fluctuation Analysis (DFA), originally introduced by Peng (1994) and employed in several financial applications (e.g. Vandewalle and Ausloos (1997), Ausloos et al. (1999), Lillo and Farmer (2004)).

Given the nature of our data, we decide to follow the standard R/S method. In fact, the duration of prediction markets is orders of magnitude shorter than that of financial markets (see Fig. 5.1), which implies that it is unlikely to mistake short-range memory for long-range effect, and Lo's modified R/S statistic

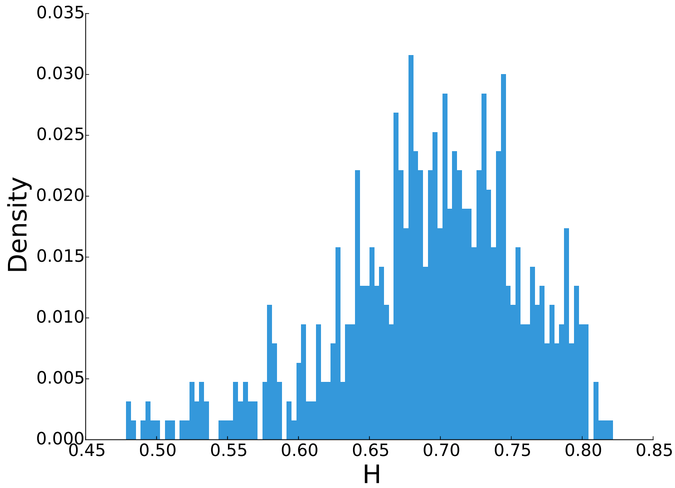


Figure 5.7: Distribution of H for the markets which last more than 30 days.

might be unnecessarily too strict for our data. Also, prediction markets' time series are relatively stable (i.e., there is no strong trend), but are occasionally shocked by sudden (and large) price changes when major news regarding the corresponding event are released. Thus, detrending the time series would not be an easy task, and might introduce significant biases.

To compute the R/S statistic, we need to calculate the ratio of R, the rescaled range, and S, the standard deviation of the time series. R is given by:

$$R_\tau = max(z_1, z_2, ..., z_\tau) - min(z_1, z_2, ..., z_\tau) \tag{5.12}$$

where

$$z_\tau = \sum_{t=1}^{\tau}(r_t - \bar{r}_\tau) \tag{5.13}$$

and $\bar{r}_\tau$ is the average return of a given market which lasts τ days. The standard deviation S_τ is given by the standard estimator

$$S_\tau = \sqrt{\frac{1}{\tau}\sum_t (r_t - \bar{r}_\tau)^2} \tag{5.14}$$

Table 5.3: Summary statistics for the distribution of H.

Mean	St.Dev.	Median	Min.	Max.
0.69	0.07	0.69	0.48	0.82

The Hurst exponent H is related to $(R/S)_\tau$ by

$$(R/S)_\tau = \left(\frac{\tau}{2}\right)^H \tag{5.15}$$

Fig. 5.7 shows the distribution of H, and Table 5.3 displays its summary statistics. The mean and the median have the same value ($H = 0.69$), and the minimum is slightly smaller than 0.5. Indeed, only a few markets display Hurst exponent values smaller than 0.5. These results are in line with the observation that emerging markets (i.e., less efficient markets with lower liquidity) show higher values of H than developed markets (Di Matteo et al., 2003), which have Hurst exponents slightly smaller than $H = 0.5$, the value corresponding to Brownian motion. The fact that markets displaying a high value of the Hurst exponent is consistent with the fact that long-memory markets cannot be efficient, because information on past prices affect future prices. Therefore, when high Hurst exponents are found, like in this case, this implies that markets exhibit long memory and, therefore, are not perfectly efficient.

5.5 Time dependence properties

In this section we discuss the time dependence properties of prediction market returns. These statistical properties have been of great interest (Mantegna and Stanley, 2000), since they have been central in the discussions about market efficiency. Moreover, these properties have direct implications in both understanding the market microstructure and in trading, hence being attractive to academics and practitioners alike. Indeed, one of the most important properties of stock markets is the absence of linear correlation in the returns.

5.5.1 Autocorrelation of returns

We first consider autocorrelation of returns, which is defined as:

$$C(\tau) = corr(r(t, \Delta t), r(t + \tau, \Delta t)) \tag{5.16}$$

In financial time series, this is found to be insignificant in most cases, with very few exceptions (Chakraborti et al., 2011; Sewell, 2011). A significantly non-zero correlation is only found at very short time-scales, and it disappears for lags greater than 1-15 minutes, depending on the market (Cont, 2001; Chakraborti et al., 2011). Cont (Cont, 2001) suggests an easy and compelling reasoning to explain this phenomenon: *"if price changes exhibit significant correlation, this correlation may be used to conceive a simple strategy with positive expected earnings; such strategies, termed statistical arbitrage, will therefore tend to reduce correlations except for very short time scales, which represent the time the market takes to react to new information"*. We find that this observation also holds for our data.

Indeed, we find that raw returns in our data are not correlated at any lag except $\tau = 1$ (see Fig. 5.8), for which they display negative autocorrelation. Although this phenomenon, called the bid/ask bounce, has been found to be a regularity in financial markets (Goodhart and Hara, 1997), we believe that in this case, it is more likely that the negative correlation value is a statistical artifact caused by the large amount of null returns in the time series due to low volumes, which are observed frequently, especially in the first days of long-term markets (i.e, longer than six months).

5.5.2 Volatility clustering

Volatility clustering, that is, the positive autocorrelation of several non-linear measures of returns, is perhaps one of the most famous stylized facts in financial markets, and is commonly cited as evidence that price changes are not independent (Mandelbrot, 1963; Chakraborti et al., 2011). Although volatility can be defined in several ways, usually it is computed by using absolute or squared returns. Then, their autocorrelation function at a given time scale ΔT is defined by:

$$C(\tau)^{sq} = corr(r(t)^2, r(t + \tau)^2) \tag{5.17}$$

or:

$$C(\tau)^{abs} = corr(|r(t)|, |r(t + \tau)|) \tag{5.18}$$

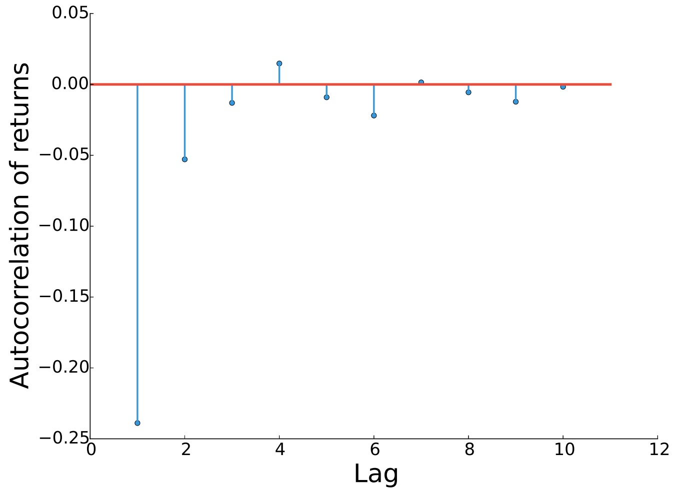


Figure 5.8: Autocorrelation of raw returns for lags in the range $0 > \tau \geqslant 10$ days.

which yield similar results. In fact, both functions are observed to decay with a power law behavior

$$C(\tau) \propto \frac{K}{\tau^{\alpha}} \tag{5.19}$$

where K is the normalization constant and α the power law exponent, empirically found to be lying in the range $\alpha \in [0.1, 0.4]$ by many authors (Liu et al., 1997; Cont et al., 1997; Cizeau et al., 1997; Mantegna and Stanley, 2000; Chakraborti et al., 2011). For our data, the autocorrelation functions of these two volatility measures are shown in Fig. 5.9.

We analyze the decay behavior of their autocorrelation functions by fitting such functions with a power law. In prediction markets, analyzing volatility clustering presents a significant issue. To fit the data with a power law, we need a large number of observations, i.e., we need as many values of $C(\tau)$ as possible. However, as shown by Fig. 5.1, most markets last less than 50 days (approximately 84%). This implies that, if we want to have 50 lags to fit our autocorrelation function, the number of markets we consider at each lag would vary. To avoid this problem, we decide to keep only those markets with more than 50 trading days, which however reduces the number of markets examined to 544. Using this approach, we find that the power-law index for

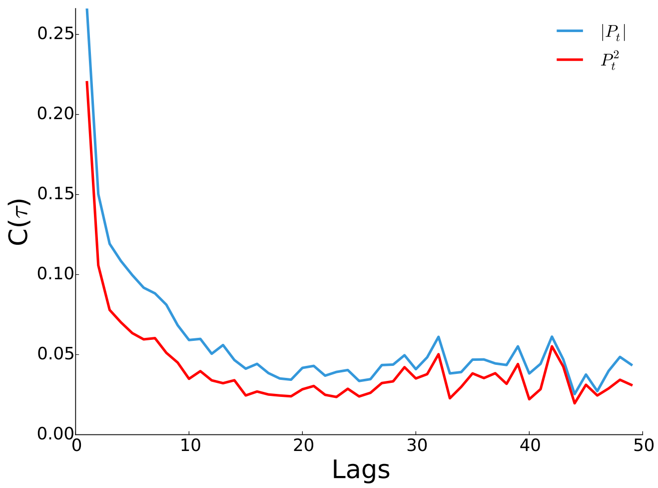


Figure 5.9: Autocorrelation of raw returns for lags in the range $0 > \tau \geqslant 50$ days.

$C(\tau)^{abs}$ and $C(\tau)^{sq}$ are similar, more precisely $\alpha^{abs} = 0.54$ and $\alpha^{sq} = 0.58$. Therefore, compared to that of stock markets, the decay of the autocorrelation of the volatility in prediction markets is slightly faster.

5.5.3 Leverage effect

Another measure of nonlinear dependence of returns is the so-called leverage effect. This effect shows that, in most financial markets, volatility is negatively correlated with past returns, which implies that negative returns increase price volatility (Bouchaud, 2001). In our data, we do not find such correlation (see Fig. 5.10), similar to what is observed in other state-contingent claims markets (Pagan, 1996). Instead, we find a small, positive correlation (between 2% and 4%) that rapidly decays to zero. One possible explanation is that prediction markets, compared with the stock market, attract fewer speculators. In fact, low volumes combined with *appealing* markets such as those on political elections we examine in this paper, provide a much lower entry barrier, both in terms of capital and knowledge needed by traders to participate in such markets.

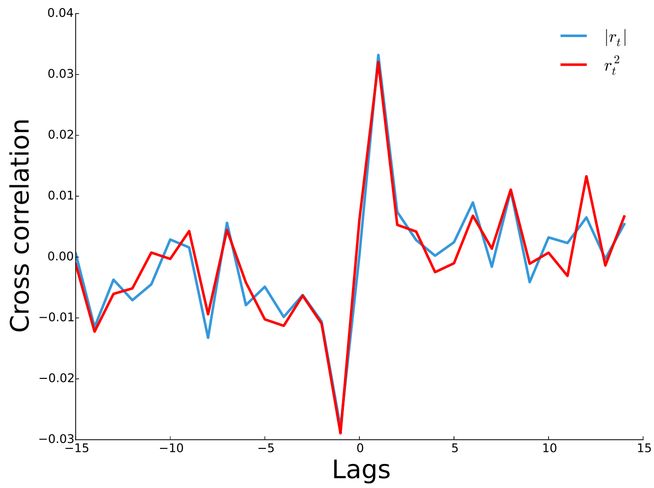


Figure 5.10: Correlation between two measures of volatility and past returns.

5.6 Conlcusions

In this paper, we analyzed several statistical properties of price changes in prediction markets by using a dataset comprising 3385 time series of security prices on political events. Our analysis is motivated by the need of new, robust means of validation for models of prediction markets, which has been fulfilled in finance by the study of empirical regularities, or stylized facts, of price and volume time series, but has been absent in prediction markets due to historical insufficiency of data.

We find that the behavior of prediction markets' price changes is remarkably similar to that of emerging financial markets, although these two types of markets possess a different structure. Specifically, we find that prediction markets' price time series exhibit high Hurst exponent values, and that their positive returns are on average larger, but less frequent, than negative returns. These two properties are often observed in emerging financial markets, but not in established ones (Chakraborti et al., 2011). Given the structural differences of prediction markets and emerging financial markets, we argue that these similarities may be caused by low volumes, which the most important characteristics these markets have in common.

In conclusion, examining the statistical properties of prediction markets allows us to gain greater insight into how they work, which is essential to both improve their accuracy. Moreover, the parallel between emerging financial markets and prediction markets can shed new light on decision making processes under uncertainty and strengthen the role of prediction markets as a testbed for financial theories.

Future work will focus on studying the statistical properties of other aspects of these markets, such as the order book, by using order-level data. This will provide an important contribution towards having a complete picture of the behavior of prediction markets. Future work will focus on studying the statistical properties of other aspects of these markets, such as price impact and mispricing, by using order-level data. This will provide an important contribution towards having a complete picture of the behavior of prediction markets.

Chapter 6

Statistical properties of volume and calendar effects in prediction markets

6.1 Introduction

Prediction markets are effective tools that harness the *wisdom of the crowd* to make accurate forecasts on a number of events (Berg et al., 2008). Although prediction markets are most famous for allowing anyone to bet on political events, often resulting in better predictions on political election outcomes than polls and experts (Wolfers and Zitzewitz, 2006), they are also used in many other contexts, e.g., to forecast business output by companies such as Google, Intel, and General Electrics, to predict the likelihood of natural disasters, or the future value of macroeconomic parameters (Plott and Chen, 2002; Cowgill et al., 2009). Moreover, due to features such as possessing a definite end-point, prediction markets represent an ideal test bed to study decision making under uncertainty. This allows, opposite to financial markets, to observe the outcome of an event, and all uncertainty is resolved at a fixed point-in-time.

However, historical insufficiency of data has limited the number of empirical studies of prediction markets. Notably, there is no comprehensive work on the empirical regularities observed in prediction markets (or *stylized facts*), whereas in financial markets data-driven analysis has always represented a prominent, valuable field of study (Mantegna and Stanley, 2000; Cont, 2001). One of the main consequences is that quantitative models of prediction markets lack an important means of validation.

In this paper, we focus on the analysis of daily volumes (measured as the number of shares traded on a given day), and calendar effects, i.e., regularities that occur during a trading period, such as a week, or a year. We find that volume in political prediction markets shares only few of the characteristics typical of stock market time series. Specifically, we find that volume properties, including calendar effects, seem to be similar to those observed in the stock market, whereas we find no evidence of any price seasonalities.

The paper is organized as follows. In Section 2 we present the data set and explain how prediction markets work. In Section 3, we perform a statistical analysis of volume, and Section 4 depicts our findings on volume and price calendar effects. Finally, in Section 5 we summarize and discuss our results.

6.2 Data and Methods

Our data set comprises the daily volumes and the OHLC contract prices of 3385 betting markets on political events, provided by PredictIt[1], for a total of 112761 valid observations (i.e., after removing all days in which there was no trading activity). Contracts on the PredictIt exchange market are Arrow-Debreu securities, i.e., contracts which are priced between 0 and 1 dollars, and whose payoff is either 0 or 1 dollars and solely depends on the the outcome of a future event. For instance, one could buy a contract on either *"Trump will lead"* or *"Clinton will lead"* in the market *"Who will lead in Trump vs. Clinton polling on September 14?"* (or sell a contract on *"Clinton will lead"* or *"Trump will lead"*, respectively). Then, one contract *"Clinton will lead"* pays 1 dollar if Clinton will be leading in Trump vs. Clinton polling on September 14, and 0 otherwise. As a consequence, rational traders are willing to buy a contract on a given outcome only if the current price of such a contract is lower than the probability they attach to the respective outcome to occur.

To perform our analysis, we use this data in two ways. To examine the distribution of daily traded shares, we aggregate volumes across all markets, which allows us to have sufficient observations to reconstruct a significant distribution. Conversely, to examine other properties such as calendar effects, we analyze each market separately and then take both the average and the median results among all markets, which allows us to have a more detailed statistical description of these phenomena.

In the next sections, we present our findings and describe in more detail how the results are obtained.

Table 6.1: Summary statistics for the distribution of traded volume.

N.Observations	Mean	St.Dev.	Minimum	$q_{25\%}$	$q_{50\%}$	$q_{75\%}$	Maximum
112761	3515.68	18950.04	1	43	306	1761	1388889

6.3 Statistical analysis of traded volume

In this section, we analyze the statistical properties of volume, which is measured as the number of daily traded shares, from the PredictIt data set. Specifically, we examine its distribution, its temporal evolution, and its long-term memory.

6.3.1 Volume distribution

To analyze the distribution of the number of contracts traded each day for each market, we exclude those days in which no contract has been traded, which leaves 3363 markets and a total of 112761 observations (i.e., trading days with positive volume). The summary statistics of the distribution of volumes (shown in Table 6.1) indicate that most of the markets examined display a small number of daily trades. Specifically, we find that only in half of the days with trading activity the number of transactions is greater than 306, and only during 25% of the active days 1761 or more contracts are purchased. Also, we observe that the mean is one order of magnitude larger than the median, and the kurtosis and skewness values are high. This may indicate that the distribution of volumes is characterized by heavy tails, i.e., most of the trading activity is concentrated in few trading days. Many probability distributions that characterize natural and social phenomena display such heavy tails. More specifically, most of these distribution have a power-law like asymptotical behavior Newman (2004); Sornette (2006). In financial markets, the tails of distribution of price changes have been shown to be heavy for most stocks and indexes (Campbell et al., 1997; Cont et al., 1997) and, although the exact asymptotic behavior of such tails is still under debate (Malevergne et al., 2005), the power-law decay, given by:

$$p(x) \sim x^{-\alpha} \tag{6.1}$$

is the most widely used (Gopikrishnan et al., 2000; Plerou et al., 2004) to fit the decay of the tails.

[1] www.predictit.org

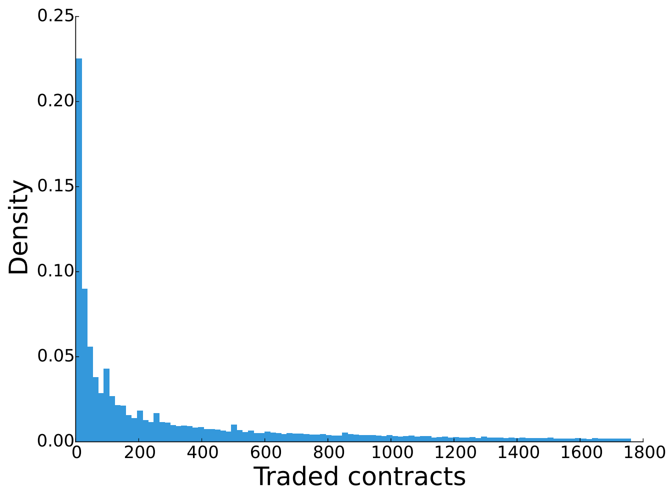


Figure 6.1: Distribution of the number of daily traded contracts. The distribution is shown only for $v < 1761$, corresponding to the 75% of the observations.

Both the summary statistics and Fig. 6.1 suggest that this might also be the case of our distribution. We check this by fitting the tail of our distribution by following a procedure which enables us to estimate the power-law exponent for discrete data (Bauke, 2007), and relies on a maximum likelihood estimation. This procedure, in contrast to other methods such as graphical methods and linear regression, is found to be more robust and reliable (Bauke, 2007; Deluca and Corral, 2013)).

In more detail, this method, which is essential to fit a PDF to a discrete power-law form (Bauke, 2007; Clauset et al., 2009), consists in finding the value α, such that:

$$p(x) = \frac{x^{-\alpha}}{\Delta\zeta} \tag{6.2}$$

where x represents the daily volumes, and $\Delta\zeta$ is the difference:

$$\Delta\zeta \equiv \zeta(\alpha, x_{min}) - \zeta(\alpha, x_{max}) \tag{6.3}$$

where ζ is the Hurwitz zeta function, defined as:

$$\zeta(\alpha, x_{min}) = \sum_{i=0}^{\infty} \frac{1}{(i + x_{min})^{\alpha}} \tag{6.4}$$

Here, x_{min} is the number of traded shares after which the distribution of volume starts behaving like a power law. The theoretical limit of the distribution, i.e., the largest possible value of x, is denoted by x_{max}. However, for volumes, there is no such a constraint. Indeed, in theory, any number of shares can be exchanged during a single trading day. Therefore, we can assume that $x_{max} = \infty$ and, consequently, $\zeta(\alpha, x_{max}) = 0$.

Given this, it is possible to compute the likelihood function for $p(x)$, which is given by

$$L(\alpha) = -\alpha\left(\sum_{i=0}^{N} ln(x_i)\right) - Nln\big(\Delta\zeta\big)\big) \tag{6.5}$$

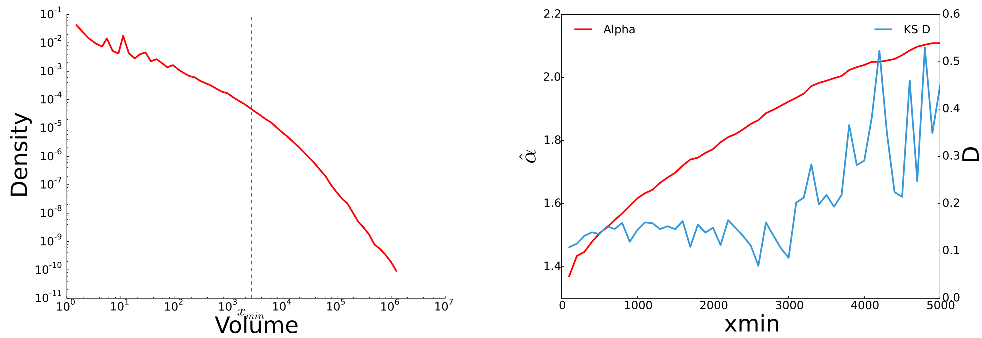
Then, the maximum likelihood estimator, $\hat{\alpha}$ is given by:

$$\hat{\alpha} = \underset{\alpha}{\operatorname{argmax}}[L(\alpha)] \tag{6.6}$$

Since, in this case, there exists no closed-form solution for $\hat{\alpha}$, we find the value that maximizes Eq. (6.5) numerically.

Finally, the last step required in order to accurately estimate α, is to find the numerical value of x_{min}. To achieve this, we perform a two-sample Kolgomorov-Smirnov test (KS), as suggested by Clauset et al. (Clauset et al., 2009). The procedure they introduce is as follows: first, we fix the value of x_{min}, starting from the smallest possible, and remove from our data all values of x such that $x < x_{min}$, if any. Second, we fit a power-law distribution to these values, and find $\hat{\alpha}$. Third, we perform the KS test between our data and a sample drawn from a power law distribution with exponent $\hat{\alpha}$, hence computing the KS statistic (D). Finally, we increase by the smallest possible increment the value of x_{min}, and we repeat the procedure until all possible values of x_{min} have been considered.

Then, we choose the x_{min} that minimizes the value of D, and take the corresponding $\hat{\alpha}$ as the power-law exponent for our distribution. By following this procedure, we find that the distribution of traded shares follows a power-law with exponent $\hat{\alpha} = 1.865 \pm 0.002$ for values greater than 2600, corresponding to the 20% of the total observations. This value is not distant from the power-law exponent $\gamma_q = 1.53 \pm 0.07$ estimated for financial markets (Gopikrishnan et al., 2000; Gabaix et al., 2007), from which we can conclude that, although in prediction markets volumes are lower than in the stock market, the decay of the number of traded shares is similar.



(a) (b)

Figure 6.2: Figure (a) displays the PDF of volumes in a logarithmic scale. Figure (b) shows the KS statistic and the corresponding values of $\hat{\alpha}$ for different x_{min}

6.3.2 Autocorrelation of volumes

Next, we examine the long-memory properties of volumes. To achieve this, compute the autocorrelation function of the number of traded shares, and fit it to a power-law distribution. To obtain an accurate estimation, we computed the autocorrelation function for lags in the range $1 < \tau < 100$, i.e., we used all markets longer than 100 days, for a total of 236 markets. We find that the volume autocorrelation function can be described as:

$$\langle V(t),\ V(t+\tau)\rangle \sim \tau^{-\lambda} \tag{6.7}$$

where we estimate the exponent to be $\lambda = 0.094 \pm 0.003$ (see Fig. 6.3). This result suggests that trading activity behaves in the same way in both prediction and stock markets, in which the power-law exponent is observed to be of the same order of magnitude. More specifically, its value is estimated to be $\lambda = 0.30$ for US stocks (Plerou et al., 2001), and $\lambda = 0.21$ for the Chinese stock market (Qiu et al., 2009), which also suggests that the decay of the volume autocorrelation function is faster the more liquid the market is.

6.3.3 Temporal evolution of traded volume

An interesting aspect of prediction markets time series (and, more generally, state-contingent claims) is that, in contrast to those of the stock market, they have a fixed end-point. In this section, we examine this aspect of prediction

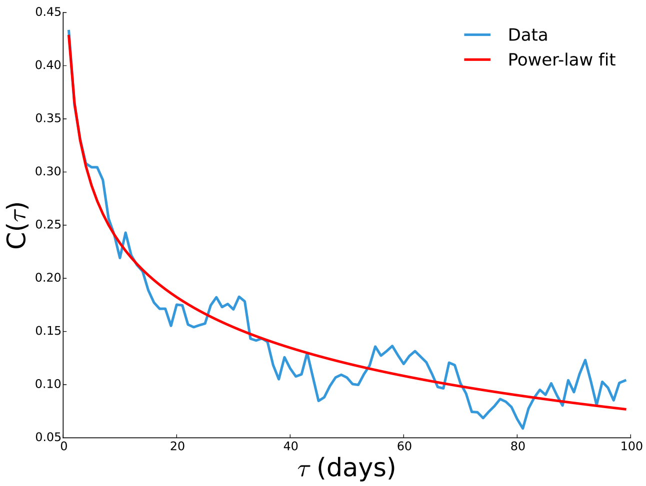


Figure 6.3: Autocorrelation function of traded volume and the fitted power law with exponent $\lambda = 0.094$.

markets, i.e., the temporal distribution of volume, and find that, towards the end of the market, the average daily volume grows significantly. Specifically, the number of traded shares depends on the number of remaining days τ until the end of the market, and, as shown in Fig. 6.4, this relation follows power-law decay:

$$V(T - \tau) \sim \tau^{-\zeta} \tag{6.8}$$

where T denotes the final day of the market. We fit this function with a power law, and we estimate the exponent to be $\zeta = 2.44 \pm 0.06$, which suggests that the during the last days of trading, volumes are higher than during all the rest of trading days combined. This result can be explained in several ways. For example, those who invest in prediction markets, may be waiting for a lower uncertainty on the outcome (i.e., waiting for new information to be revealed), or they simply have a higher utility to bet in the days right before the end of the market, hence reducing the time between the investment and the (potential) gain. Another explanation is provided by Alfi et al. (2009), that analyses how people respond to deadlines by examining the time distribution of conference registrations. Either way, we believe this is a crucial result for building realistic models of prediction markets, because this phenomenon may generate non-

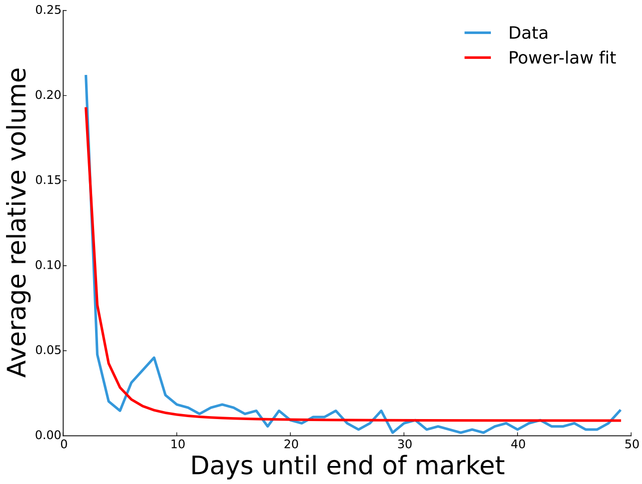


Figure 6.4: Relative volume depending on the number of days τ until the end of the market.

trivial price dynamics during the last days of trading.

6.3.4 Volume-volatility correlation

In this section we examine the correlation between volume and price in prediction markets. In the stock market, it has been observed in a number of contexts that volume changes and the volatility of returns are correlated (Chordia et al., 2001; Podobnik et al., 2009). For instance, it is shown that volatility grows proportionally to the total number of trades in a market (Podobnik et al., 2009). Unfortunately, for the prediction markets, we do not possess order-level data, and hence we show that volume and volatility are correlated on a daily time scale. That is, we compute the correlation coefficient

$$C(\tau)^{sq} = \langle r_t^2,\ v(t+\tau)\rangle \tag{6.9}$$

and find that correlation is significant only for $\tau = 0$. Fig. 6.5(a) shows the cross correlation function between traded volume and volatility and also between traded volume and raw returns, defined as:

$$C(\tau) = \langle r_t,\ v(t+\tau)\rangle \tag{6.10}$$

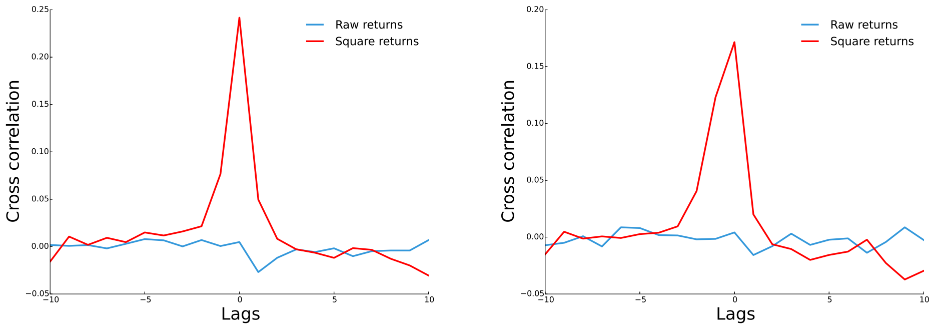


Figure 6.5: Figure (a) shows the cross-correlation between traded volume and both volatility (square returns) and price changes (raw returns). Figure (b) shows the cross-correlation between changes in traded volume and both volatility (square returns) and price changes (raw returns)

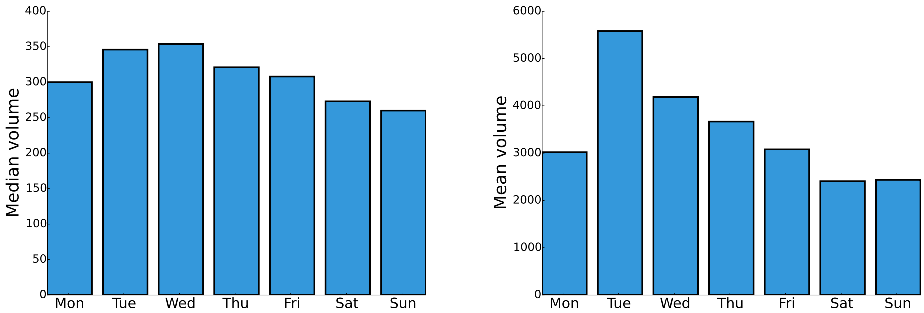
for which the correlation coefficient is insignificant at all lags τ. This implies that volume is only correlated with volatility (at lag 0) but not with price changes, which is a well known fact in financial markets (Podobnik et al., 2009). Interestingly, we find similar results when computing the cross correlation between returns and volume changes (Fig. 6.5(b)). This is in contrast to what is found in the stock market, for which it has been observed that the correlation between volume changes and volatility decays with a power law (Podobnik et al., 2009). Conversely, in our data set we find that volatility is correlated with volume changes only at lag 0.

6.4 Calendar Effects

Calendar effects, or *seasonalities*, are cyclical regularities that occur throughout a trading period, be it a year, a week, or a day, and have been observed in both returns and volume by a number of authors who examined international stock markets (Sewell, 2011). In this section we examine some well-known effects that are present in financial markets (Dzhabarov and Ziemba, 2010), and we find that only some of them can be observed in prediction markets. Specifically, we first describe cyclical regularities exhibited by trading activity and then focus on price changes, for which we examine the Weekend and the January effects in detail.

6.4.1 Trading activity calendar effects

There is evidence that, in financial markets, trading activity significantly varies depending on the time of the day and the day of the week. The first comprehensive study of volume calendar effects (Jain and Joh, 1988) examines several years of NYSE-listed stock data and find that liquidity is lowest on Monday, peaks on Wednesday, and drops until Friday. A similar, more recent study (Chordia et al., 2001), which analyzes U.S. stocks between 1988 and 1998, find that the volume peak has shifted to Tuesdays, whereas Fridays have become the days with the lowest liquidity. In this section we analyze trading activity in our data, and find that it significantly varies across days of the week and across months of the year. Although this behavior is similar to that of the U.S. stock market, this is a non-trivial result, since prediction markets possess two main differences compared with stock markets. Specifically, in prediction markets, it is possible to trade during weekends. Also, since liquidity in prediction markets is much lower than in financial markets, we find that the average number of traded shares is significantly affected by those markets in which volumes are largest. Specifically, to overcome this issue, we present our results using both the average and the median volumes.



(a) (b)

Figure 6.6: Figure (a) and Figure (b) display the median and the mean, respectively, number of contracts traded by day of the week.

Despite these differences, we find that most of our results are comparable with those of U.S. stocks. In fact, we conclude that, in our data, trading activity is lowest during weekends, but otherwise shows a trend similar to that found in the U.S. stock market (see Fig. 6.6). Table 6.2 shows that the average vol-

Table 6.2: This table displays summary statistics of the trading activity (expressed as the number of contracts traded) across the days of the week. The t statistic is used to either accept or reject the null hypothesis that the mean volume value of a given day of the week is the same as the mean value for the other days.

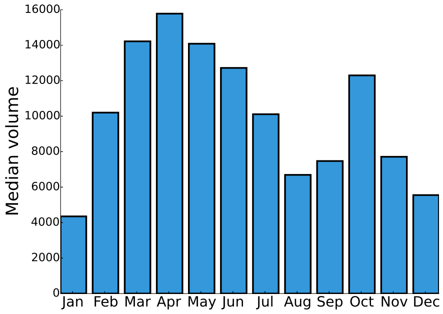
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Mean	3019.45	5580.16	4187.65	3667.96	3079.47	2406.12	2436.10
St. Dev.	10550.43	36248.83	18319.36	19204.85	11540.58	10730.70	8266.70
Median	300	346	354	321	308	273	260
t-stat.	-5.50*	8.55*	5.09*	1.11**	-4.62*	-11.85*	-13.34*

* corresponds to a significance level of 0.01%.
** indicates that the result is not significant.

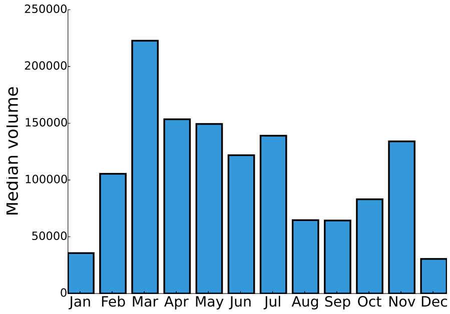
ume is low on Mondays, peaks on Tuesdays, and then decreases gradually for the rest of the week, and it reaches its lowest value during weekends, which agrees with the analysis by Chordia et al. (2001). The analysis of the median number of traded shares (Fig. 6.6, and Table 6.2) shows a similar pattern, although the volume differences across the days of the week become less pronounced compared to the average value, and the number of traded contracts has a high on Wednesdays instead. We repeat the analysis for the months of the year, and we find that, although the differences between mean and median are more pronounced than in the weekly analysis, both measures show similar trends (see Fig. 6.7). First, January and December are the months with the least trades in both cases. Second, both the mean and the median volumes increase from January to Spring (April and March for the median and the mean value, respectively), then have a local low in August, and then a new high in Autumn (October for the median volume, November for the mean volume). These findings suggest that, despite the structural differences, volume temporal regularities in prediction markets are similar to those found in stock markets.

6.4.2 Price calendar effects

In this section, we examine price changes across days of the week and months of the year. We first introduce these regularities, also presenting the results found in financial markets, and then show that these two patterns are not exhibited by our data. Indeed, we find that, opposite to volume, price in prediction markets does not follow the same behavior as in the stock market and, more generally, does not seem to exhibit any regularity. Conversely, in numer-



(a)



(b)

Figure 6.7: Figure (a) and Figure (b) display the median and the mean number of contracts traded by month of the year, respectively.

Table 6.3: This table displays summary statistics of the trading activity (expressed as the number of contracts traded) across the months of the year. The t statistic is used to either accept or reject the null hypothesis that the mean volume value of a given day of the week is the same as the mean value for the other days.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Mean	35621.59	105487.06	222737.46	153514.76	149391.7	121854.35	139041.75	64646.92	64314.74	83033.61	134040.53	30493.27
St. Dev.	18903.99	66802.39	156852.56	140245.76	183499.35	112634.55	118178.35	42586.11	45412.31	49532.94	151787.61	13093.59
Median	4350.0	10200.0	14220.0	15780.0	14085.0	12720.0	10110.0	6690.0	7470.0	12300.0	7710.0	5550.0
t-stat.	-35*	0.02**	12.39*	5.83*	4.46*	2.79*	5.65*	-15.35 *	-15.37*	-8.63*	3.2*	-40.36

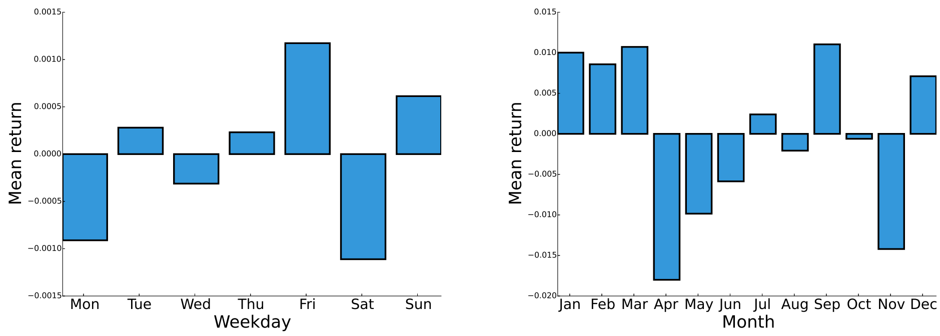
* corresponds to a significance level of 0.01%.
** indicates that the result is not significant.

ous stock markets, it has been observed that prices display more calendar regularities than volume, and the study of this topic has generated a large body of literature (Thaler, 1992; Constantinides et al., 2003). After their discovery, many of these anomalies have reduced or even disappeared (Mclean and Pontiff, 2016), but some of the most important calendar effects, among which the *January effect* and the *Weekend effect* are the most documented (Sewell, 2011), are still present in many stock markets (Dzhabarov and Ziemba, 2010).

6.4.3 The Weekend and the January effects

The weekend effect (sometimes referred to as *Monday effect*) is an empirical regularity by which average returns on Mondays are significantly lower than those of the rest of the week, and is often regarded as the strongest of calendar effects (Rubinstein, 2001). This anomaly was firstly observed in the 1930s (Fields, 1931), but the first comprehensive discussion was provided by Kenneth French (French, 1980), who analyzed more than twenty years of stock returns in the U.S. market to test two hypotheses. The first, called calendar time hypothesis, states that the expected returns on Mondays should be three times those for the other days of the week, since the the risk accumulated during weekends should be reflected in Monday's returns. The second, named trading time hypothesis, states that, if only trading time matters to generate returns, there should be no distinction between Mondays and other days. However, French found that neither of these hypotheses were true. In fact, he found that, on average, Mondays display lower returns than all of other days of the week and, more specifically, Monday is the only day of the week during which average returns are negative.

Lakonishok and Maberly (Lakonishok and Maberly, 1990) provide an explanation of the weekend effect based on the analysis of trading patterns of individual and institutional investors. First, they find that, on Mondays individual investors tend to trade more compared with the rest of the week, and also that the number of sell transactions relative to buy transactions increase significantly. Second, they observe that, in their data, the traded volume by institutional investors was the lowest on Mondays. They claim that these two regularities combined provide a partial explanation for the weekend effect.



(a) (b)

Figure 6.8: Figure (a) and Figure (b) display the mean return across days of the week and months of the year, respectively.

The January effect is another important calendar regularity, whereby returns on January are significantly higher than in other months. It has been first observed in the U.S. and Australia stock markets (Wachtel, 1942; Praetz, 1972; Officer, 1975; Rozeff and Kinney, 1976), and in several international stock markets afterwards (Gultekin and Gultekin, 1983; Agrawal and Tandon, 1994). Similarly to the weekend effect, the January effect has proven to be a regularity whose causes are puzzling (Haugen and Lakonishok, 1988). There are many competing explanation attempts, but most of these theories revolve around small firms. Indeed, there is evidence that this phenomenon is related with the capitalization of firms, and then that it is likely to be a consequence of a small-firm effect (Reinganum, 1983), low share prices (Bhardwaj and Brooks, 1992), or tax-motivated trading (Sias and Starks, 1997; Poterba and Weisbenner, 2001).

6.4.4 Analysis of returns

In this section we examine the seasonality of returns, to find whether the Weekend and the January effects exist in prediction markets. To achieve this, we follow the same procedure employed to analyze calendar effects on volume, and take into account both the mean and the median return. However, in contrast to traded volume, returns do not seem to possess any significant differences across days of the week (see Fig. 6.8). Mean daily returns, as it is shown in Table 6.4, lie between -0.001 and 0.001 for all days of the week, i.e., they are one order of magnitude smaller than the minimum possible raw return $|r_t| = 0.01$, and these small differences disappear completely when considering the median returns. Accordingly, we find that all the p-values from the t-test are greater than 0.7, and hence the null hypothesis that average returns are the same across the days of the week cannot be rejected. Similarly, we find that monthly returns do not display any significant difference (see Table 6.5). These findings are consistent with the hypothesis that the January effect is due to smaller-capitalization stocks and tax-loss selling (Roll, 1983). Indeed, in prediction markets, there is no equivalent of capitalization since contract prices purely reflect the likelihood of a given event to occur as perceived by market participants. Also, losses from these markets do not impact on fiscal contribution, since prediction markets fall under the gambling legislation in most countries and, importantly, volumes are too low to affect fiscal contribution whatsoever.

Table 6.4: This table displays summary statistics of the returns for each day of the week. The t statistic is used to either accept or reject the null hypothesis that the mean return of a given day of the week is the same as the mean return for the other days.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Mean	-0.0009	0.0003	-0.0003	0.0002	0.0012	-0.0011	0.0006
St. Dev.	0.11	0.11	0.11	0.11	0.14	0.08	0.07
Median	0	0	0	0	0	0	0
t-stat.	-1.11*	0.34*	-0.41*	0.3*	1.21*	-1.75*	1.1*

* indicates that the result is not significant.

Table 6.5: This table displays summary statistics of the returns for each month of the year. The t statistic is used to either accept or reject the null hypothesis that the mean return of a given month of the year is the same as the mean return for the other months.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Mean	0.01	0.0086	0.0107	-0.018	-0.0098	-0.0059	0.0024	-0.0021	0.011	-0.0006	-0.0142	0.0071
St. Dev.	0.498	0.579	0.699	0.545	0.563	0.579	0.593	0.614	0.558	0.582	0.655	0.549
Median	0	0	0	0	0	0	0	0	0	0	0	0
t-stat.	0.34*	0.25*	0.25*	-0.55*	-0.32*	-0.2*	0.08*	-0.07*	0.4*	-0.03*	-0.37*	0.21*

* indicates that the result is not significant.

6.5 Conclusions

We analyzed calendar effects and several statistical properties of volumes in prediction markets, by using a data set comprising 3385 time series of security prices and trading volumes on political events. We find volume properties and price seasonalities mostly differ from those observed in financial markets, with the exception of volume calendar effects. We argue that this may be caused by the low liquidity and short duration that characterize prediction markets, especially compared with those of the stock market. These results suggest that studying prediction markets could provide additional insights on people's individual and collective behavior when investing under uncertainty, and we advocate the use of our results to build and validate new models of prediction markets.

Chapter 7

Opinion dynamics and price formation in prediction markets

7.1 Introduction

Futures contracts have long been used in finance to harness the *wisdom of the crowd* and make predictions about the future value of an asset by exploiting people's aggregate expectations. Prediction markets are one of the most recent forms of futures and, although they were originally only meant to forecast the outcomes of important political events, nowadays they are used in a number of different contexts. For instance, alongside public markets that allow betting on political or sports events, there exist private prediction markets that are used by companies such as Google, Intel, and General Electric to gather people's beliefs about business activities such as sales forecasts or the likelihood of a team meeting certain performance goals (Plott and Chen, 2002; Cowgill et al., 2009). Although prediction markets are often heralded as effective mechanisms to make highly accurate predictions (Berg et al., 2008; Arnesen and Bergfjord, 2014), it has been shown that they are prone to bias (Restocchi et al., 2018c) and manipulation (Goodell et al., 2015), and that these phenomena can spread to financial markets with dreadful consequences (Goodell and Bodey, 2012; Goodell et al., 2015).

To better understand prediction markets and consequently account for these adverse events, thereby limiting their impact outside the prediction market itself, there is the need for models that realistically reproduce the underlying processes that drive price and opinion formation. This can be achieved by building adequately complex frameworks that can be validated against real-world data. That is, models that are simple enough to be understood and controlled, but complex enough to allow the emergence of realistic and complex dynamics generated by simple interactions between agents. Despite the vast amount of existing models of prediction markets (Shin, 1993; Ottaviani and Sørensen, 2008; Snowberg and Wolfers, 2010; Restocchi et al., 2018c), there is little work on modelling prediction market exchange platforms, arguably the most important type of prediction market.

To address this gap, in this paper we propose a model that matches the empirical properties (often referred to as *stylized facts*) of historical price and volume time series of political prediction markets. To achieve this, we consider a social network where agents possess an opinion about the probability of a given event occurring, and either buy or sell contracts on a prediction market exchange based on their opinion. To model the opinion dynamics process, we use the Deffuant model (Deffuant et al., 2000). Opinion dynamics has long been a topic on which a number of physicists applied statistical physics tools to better understand human interactions and the complex phenomena that emerge from them (see, for example, Deffuant et al. (2000); Sznajd-Weron and Sznajd (2000); Clifford and Sudbury (2002); Tessone et al. (2004); Baronchelli et al. (2006) for seminal opinion dynamics models and Castellano et al. (2009) for a thorough review of the topic). Among the many models proposed to describe opinion diffusion in social networks, we choose to follow the Deffuant model, which has three features that make it especially suitable to represent the underlying opinion propagation process that determines prices in prediction markets. First, the Deffuant model considers continuous opinions bounded by arbitrary values. This makes it perfect to describe the diffusion of opinions about the probability of an event to occur, as the opinion can be bounded between 0 and 1 and take any value in this range. Second, since it has only two free parameters, the Deffuant model has the merit of being extremely simple, which allows us to gain deeper insights on what drives price properties in prediction markets. This also guarantees a good degree of realism without having to make assumptions on other parameters in the model, which have otherwise to be fine-tuned to guarantee a good degree of realism, which is a common (and often necessary) practice for models of financial markets (Lux and Marchesi, 1999a; Zhou and Sornette). Third, similar to other bounded confidence models of opinion diffusion, in the Deffuant model only people with similar opinions update their beliefs after interacting, which allows the possibility of not reaching consensus at equilibrium. This property enable us to analyze the similarity of our results to historical data depending on the number of opinion clusters that coexist at equilibrium.

We identify two important contributions that this paper provides. First, we introduce the first model of prediction markets to use opinion dynamics in social networks as its underlying mechanism. This model has the merit of being particularly simple, since it possesses only two free parameters, but highly robust at the same time, it is capable of generating price time series that match the stylized facts of historical data for any combination of the two free parameters. Second, we provide empirical validation for the Deffuant model. This contribution helps tackling an important challenge that Sobkowitz posed by arguing that opinion dynamics models are often disconnected from the real world and lack of empirical validation(Sobkowicz, 2009). Since his paper, the explosion of popularity social media experienced has certainly offered means of compelling validation (Del Vicario et al., 2016, 2017), but at the same time the availability of such vast data sets and the ease with which information can be collected from these platforms has polarized the source of data, with the consequence that rarely other data sets are considered. In this paper, for the first time, we use empirical market data to validate continuous opinion diffusion models, and specifically the Deffuant model. To achieve this, we use data from PredictIt, a political prediction market exchange platform, and show that the Deffuant model provides an excellent representation of the opinion diffusion process in social networks when opinions are scalar.

The remainder of the paper is organized as follows. In Section 7.2 we define the model of opinion formation and market exchange. In Section 7.3 we explain the experimental setting we used to run agent-based simulations of our model, and discuss our findings, showing that our model provides a qualitatively good description of historical price and volume time series even in the worst case scenario. Finally, we conclude with a short discussion and outline future work in Section 7.4

7.2 Model

In this section we describe the price formation model and argue why it is appropriate to represent prediction markets. We start by describing how prediction markets work, also defining normalized prices and true probabilities, and then we describe the model dynamics in detail.

Prediction markets are time-limited markets in which contracts are traded on the outcomes of a given event, $\mathbb{E}_i \in \{0, 1\}$, where i denotes the *i-th* event, and can take only two values: $\mathbb{E}_i = 1$ if i occurs, and 0 otherwise. Similarly, the payoff of a contract on the i-th event is 1 if the corresponding event occurs and 0 otherwise. Let us denote with $\tilde{\pi}_j \in (0, 1)$ the price of the contract on the event $\mathbb{E}_i$. Since, in prediction markets, $\sum_j \tilde{\pi}_j = 1 + \delta$, where δ is the turnaround (e.g., spread, bookmaker fees, etc.), for our analysis we consider normalized prices $\pi_i = \frac{\tilde{\pi}_i}{\sum_j \tilde{\pi}_j}$, which give a better representation of the corresponding realization probabilities. Also, if a prediction market is completely efficient, the normalized price of a security reflects exactly the probability of the corresponding outcome to happen, i.e., $\pi_i = p_i = P(\mathbb{E}_i = 1)$, where p_i is referred to as *true probability*.

To model the opinion diffusion process we follow the Deffuant model (Deffuant et al., 2000). We start by considering a population of N agents, who belong in an undirected, unweighted social network $\mathbb{G}$. Without loss of generality, we assume that there is only one event with two possible outcomes $\mathbb{E} = 0$ and $\mathbb{E} = 1$. Then, agent j possesses an opinion $o_j(t) \in [0, 1]$, which can take any real value and corresponds to its subjective probability the agent attaches to the outcome $\mathbb{E} = 1$. The opinion update process is iterative: at each time step, agents may discuss the event with their neighbors, and update their opinion. To model this process, every round an agent i is randomly chosen from the network to discuss with agent j, which is, in turn, randomly chosen among agent i's neighbors. If their opinions are too different, they refuse to update their beliefs. More precisely, the agent pair (i, j) interacts only if $|o_i(t) - o_j(t)| < \varepsilon$, where ε is the threshold of this process and can take any real value between 0 and 1. If the agent pair interacts, they update their opinions as following:

$$\begin{aligned} o_i(t+1) &= o_i(t) + \mu\left[o_j(t) - o_i(t)\right] \\ o_j(t+1) &= o_j(t) + \mu\left[o_i(t) - o_j(t)\right] \end{aligned} \tag{7.1}$$

where μ is the *convergence parameter*, and $\mu \in [0, \frac{1}{2}]$.

In this model, ε represents the *open-mindedness* of agents which would discuss with, or listen to other agents, only if their opinions are sufficiently close. In this paper we follow the basic Deffuant model and consider ε constant among all agents, but there exist other versions of this model in which agents have heterogeneous open-mindedness (Weisbuch et al., 2002; Lorenz). In general, ε affects the number of clusters at equilibrium (i.e., the number of opinions that coexist), while μ drives the convergence time (Deffuant et al., 2000).

To reflect the temporal volume dynamics observed by Restocchi et al. (Restocchi et al., 2018b), each agent has a probability $p = (T - \tau)^{-\gamma}$ to participate in the market, where T is the duration of the market (in days), τ is the number of days until the end of the market, and $\gamma = 2.44$ is a scaling parameter, estimated by Restocchi et al. (2018b). If agent i is chosen to participate in the market, they can either buy or short sell a contract, whose payoff is 1 if $\mathbb{E} = 1$ and 0 otherwise. For sake of simplicity, we assume that there exists only one event, with two possible outcomes. Since agent i believes that $P(\mathbb{E} = 1) = o_i$, they will buy a contract only if the current price $\pi < o_i$, and sell (or short sell) it if $\pi > o_i$. They will neither buy or sell if $\pi = o_i$. The demand of agent i, D_i, is proportional to the distance between their opinion and the price, and is described by

$$D_i(t) = o_i(t) - \pi(t) \tag{7.2}$$

That is, the more mispriced the agent believes the contract is, the more they will trade. The excess demand (ED) is simply the sum of each agent's demand, multiplied by a noise term $\nu \sim N(0, \sigma)$, as follows:

$$ED(t) = \nu \sum_i D_i(t) \tag{7.3}$$

where, following Lux and Marchesi (2000), $\sigma^2 = 0.05$. The reason we choose to model the process with a multiplicative (cf. additive) noise term is that, to ensure that the price does not go beyond the boundaries too often, we need a quenching term. By running extensive simulations where the equation for the excess demand is $ED(t) = \beta \sum_i D_i(t) + \nu$ we found that our model gives qualitatively similar results for $0 < \beta < 0.25$. However, adding such a quenching term would require us to run additional calibration and optimization, eventually risking to overfit our model. By adding a multiplicative noise, we avoid this issue. Importantly, below we show that the empirical properties of the time series obtained by running our simulations heavily depend on μ and ε, suggesting that the noise term has little impact on the emerging properties of the model. At each trading round, the price gets updated depending only on $ED(t)$. Since prices are bounded between 0 and 1, they are set to 0 or 1 if they become less than 0 or greater than 1, respectively, following an update. Therefore, $\pi(t+1)$ takes the following values:

$$\pi(t+1) = \begin{cases} \min\{1, \pi(t) + ED(t)\} & ED(t) \geqslant 0 \\ \max\{0, \pi(t) + ED(t)\} & ED(t) < 0 \end{cases} \tag{7.4}$$

where, without loss of generality, we have included the equation for the price update for $ED = 0$ in the first case (i.e., $ED \geqslant 0$).

7.3 Results

In this section we describe the experimental setting we use to run agent-based simulations of the model, and the results we obtained from such experiments. The simulations are run with all possible combinations of μ and ε, which are the only two free parameters in our model, within the ranges $0 \leqslant \mu \leqslant 0.5$ and $0 \leqslant \varepsilon \leqslant 1$, which represent the whole possible space for the Deffuant model, with a precision of 0.01 units. Our results show that this model provides a particularly accurate description of prediction markets, because even under the worst-performing conditions, the synthetic time series produced by our simulations capture (at least qualitatively) the emerging properties of prediction markets, such as volatility clustering and absence of autocorrelation of returns (Restocchi et al., 2018a).

Social networks $\mathbb{G}$ are generated following the Barabasi-Albert preferential attachment algorithm (Barabási and Albert, 1999), since the opinion diffusion process happens across a social network, which are well described by networks generated with this algorithm, and consist of 1000 nodes. Although it has been demonstrated that, at least under some circumstances, diffusion processes display finite-size effects (Tessone et al., 2004; Toral and Tessone, 2006), we decide to run experiments with 1000 agents to reflect the scarce liquidity prediction markets usually exhibit (Restocchi et al., 2018b). In fact, if any finite-size effect exists, this must be displayed by real prediction markets and, therefore, by keeping the number of participants low, we can capture such an effect in our simulations. We leave a detailed analysis of the relation between network size and price formation in prediction markets to future work.

To specify the other parameters of the simulations, we follow Restocchi et al. (2018a,b), who provide a thorough quantitative analysis of the emprical properties of prediction markets and also use the same data set from PredictIt. Specifically, for each combination of μ, ε, we run $M = 3385$ simulations, which is the number of markets used by Restocchi et al. (Restocchi et al., 2018a,b) in their analysis, and the duration T of each market is randomly drawn from the empirical distribution of durations observed among these markets. In this way, our simulations will produce time series which are directly comparable with historical data on prediction markets. Also, for sake of simplicity, we assume that the true probability is equal to $p = P(\mathbb{E} = 1) = \frac{1}{2}$ and is constant throughout the market. This assumption reflects the fact that, on PredictIt and other prediction markets exchange platforms, it is possible to bet on an event and on its opposite (i.e., there is the possibility to buy and sell contracts on the

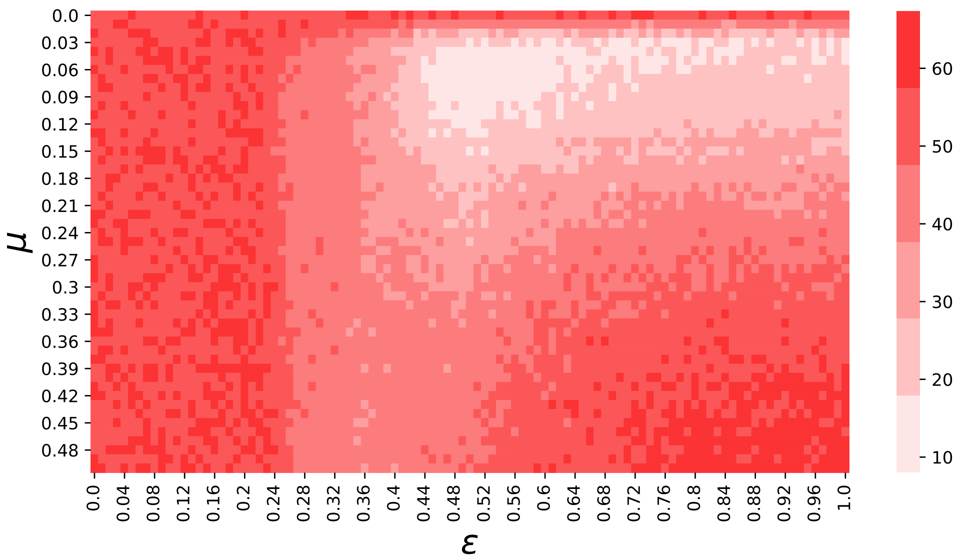


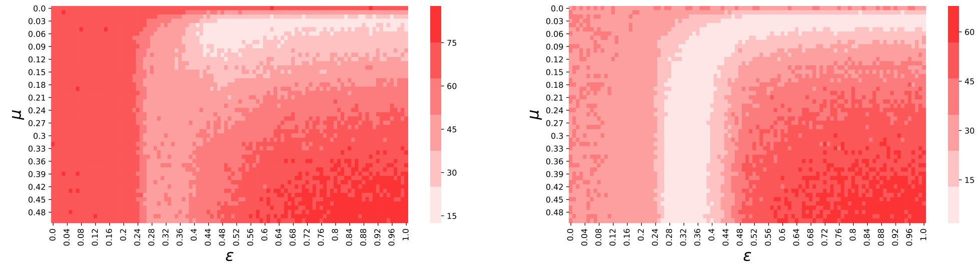
Figure 7.1: Objective function values (see Eq. 7.5) depending on μ and ε. These results show that there is a region, approximately delimited by the area $(\mu, \varepsilon) \in [0.03, 0.11] \times [0.45, 0.64]$, where the objective function f reaches its minima. Each color used in this figure represents an interval of 10 for f, starting with $7.99 < f < 17.99$. We chose to discretize the colors to smooth our results over noise, and make the regions easily recognizable.

event *Will* $\mathbb{E}$ *happen?* but also on the event *Will* $\mathbb{E}$ *not happen?*).

To find the optimal values for the pair μ, ε, we follow Gilli and Winker (2003) and define the following objective function:

$$f = |k_{sim} - k_{emp}| + \lambda|\alpha_{sim} - \alpha_{emp}| \tag{7.5}$$

where k_{sim} and k_{emp} represent the kurtosis value of the distribution of the returns of all 3385 markets, for simulated and historical data, respectively. Gilli and Winker suggest the use of kurtosis to capture the characteristic heavy tail properties of the distribution of returns. For the second term, rather than using the value of ARCH(1), i.e., the first autoregressive term of the time series, as suggested by Gilli and Winker, we use the value of the scaling parameter α that describes the power-law decay of the autocorrelation function of absolute returns. Since the two terms can significantly differ in magnitude, to ensure that



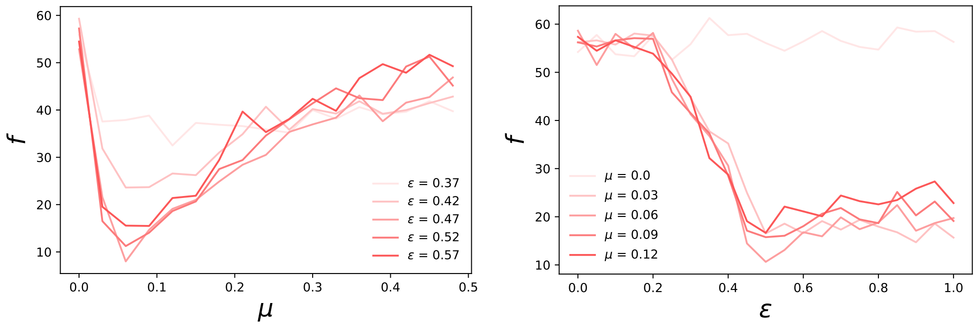
(a) (b)

Figure 7.2: Heatmaps for the objective functions $f_2 = |k_{sim} - k_{emp}| + \lambda|\alpha_{sim} - \alpha_{emp}| + \lambda^a|a_{sim} - a_{emp}|$ and $f_3 = |k_{sim} - k_{emp}| + \lambda^a|a_{sim} - a_{emp}|$, respectively. These figures show that including a in the objective function f does not add information (Fig. (a)), and that removing α reduces the granularity of f (Fig. (b)).

no component in the objective function outweighs the other, Gilli and Winker suggest that the term $|\alpha_{emp} - \alpha_{sim}|$ is multiplied by a constant $\lambda = \frac{k_{emp}}{\alpha_{emp}}$ that rescales the magnitude of the second term. For our data set, Restocchi et al. Restocchi et al. (2018a) found that $\alpha_{emp} = 0.54$ and $k_{emp} = 39.31$, from which we derive $\lambda = 72.78$. We choose to use α, instead of the first-order autocorrelation term a, because we believe this gives the calibration a better accuracy. Specifically, we also tried to calibrate the model by using two different functions, namely $f_2 = |k_{sim} - k_{emp}| + \lambda|\alpha_{sim} - \alpha_{emp}| + \lambda^a|a_{sim} - a_{emp}|$, where $\lambda^a = \frac{k_{emp}}{a_{emp}}$, and $f_3 = |k_{sim} - k_{emp}| + \lambda^a|a_{sim} - a_{emp}|$, and found that f_2 exhibits a qualitatively similar behavior to f, without adding information. Also, we found that by using f_3, as suggested by Gilli and Winker, reduces the sensitivity of the objective function to μ and ε, and generates two regions of local minima in which values are not significantly different. These results are shown by Fig. 7.2.

The objective function values computed from our simulations for each pair μ, ε are displayed by the heatmap in Fig. 7.1. It is clear to see from this figure that there exists one region, approximately within the interval $(\mu, \varepsilon) \in [0.03, 0.11] \times [0.45, 0.64]$ where f is significantly lower than in the rest of the space, and also includes the global minimum, which is found in $\mu = 0.06$ and $\varepsilon = 0.47$.

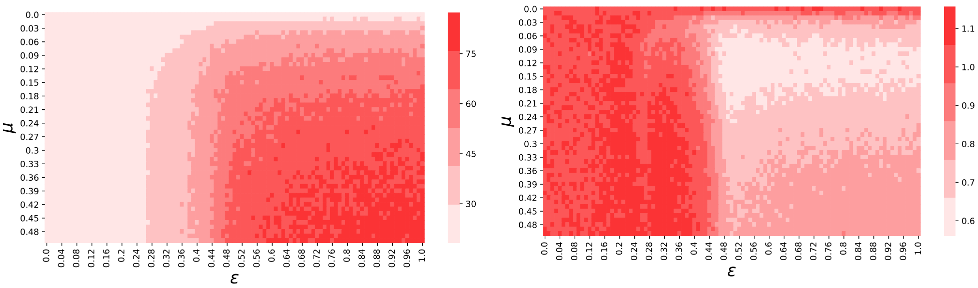
It is also interesting to note that f displays a regular behavior when varying μ and ε. To better visualize this, in Fig. 7.3 we cut the objective function space in slices and show the behavior of f separately depending on μ (ε), for a few ε (μ) around its optimal value. By observing these two figures, it is easy to



(a) (b)

Figure 7.3: Detailed view of the objective function f values, depending on μ (ε), for five values of ε (μ) in the neighborhood of the optimal values. The objective function shows strong dependence on both μ and ε, but these figures suggest that ε has a greater impact on f.

see two regularities. First, from the right-hand side plot one can note that, in the range $0.27 \lesssim \varepsilon \lesssim 0.5$, there is a sharp transition from high values of f to low ones, and that this behavior does not depend on μ (apart from $\mu \approx 0$). Second, the left-hand side plot shows that f has always a minimum around $\mu \approx 0.05$, but, depending on ε, this minimum is more or less pronounced, and it almost disappears for $\varepsilon \lesssim 0.4$. In fact, in this region, f has a sharp drop as soon as μ moves away from 0, but does not change significantly afterwards. These results suggest that both μ and ε have a significant impact on the objective function, i.e., both parameters contribute in shaping the statistical properties of the time series generated by the model. This is in a way expected, since, within short time horizons, they both contribute towards the emergence of consensus, and the speed at which this happens. For instance, for high values of μ and ε, consensus is reached too soon, and the generated time series become less accurate, as suggested by the high value of f in the region $\mu \gtrsim 0.4 \cup \varepsilon \gtrsim 0.7$. However, our results suggest that ε has a greater impact on f than μ. Specifically, we observe that there is one region, delimited by $\varepsilon \lesssim 0.26$, for which the objective function value becomes particularly high. Remarkably, this is the same value of ε below which consensus on a single opinion is not reached, and two or more opinions coexist at equilibrium (Deffuant et al., 2000). These results imply that the quantitative accuracy at reproducing k_{sim} and α_{sim} heavily depends on the number of coexisting opinions at equilibrium, and on the time it takes for opinions to converge. Further ev-



(a)

Figure 7.4: Values of k (Fig. (a)) and α (Fig. (b)) depending on μ and ε. These figures show that the kurtosis of the return distribution depends on both μ and ε equally, but that the value of α is heavily affected by ε, suggesting a link with the number of coexisting opinions at equilibrium.

idence is represented by the results shown in Fig. 7.4, in which we show the dependence of k and α on μ and ε.

These results suggest that the kurtosis of the distribution of raw returns is affected both by μ and ε, but loses its dependence on μ for low values of ε, approximately for $\varepsilon < 0.27$, that is, in the region where multiple opinions coexist at equilibrium. However, the dependence on μ when $\varepsilon > 0.5$ suggests that, when only one opinion exists at equilibrium, the kurtosis highly depends on the time it takes to reach consensus. Not surprisingly, the shorter the time to reach consensus, the higher the kurtosis, as once consensus is reached only few agents trade, and those who trade have low absolute values of demand, since their opinion is closer to the mean. Similarly, α seems to depend mostly on ε, but exhibits a far sharper phase transition around $\varepsilon = 0.5$, significantly decreasing its value for $\varepsilon > 0.5$. Also, Fig. 7.4a shows that, similarly to k, in the region $\varepsilon > 0.5$ α depends mainly on μ, and its value becomes the larger the faster the convergence rate.

These results are further evidence that the accuracy of our model depends on the number of coexisting opinions at equilibrium, except for the region approximately delimited by $(\mu, \varepsilon) \in [0.35, 0.5] \times [0.60, 1]$, where consensus is reached early in the market and little or no trading exists after a certain point. Finally, Fig. 7.4 shows that, for any pair μ, ε, the time series generated by our simulations exhibit excess kurtosis and volatility clustering, which are the two main features of prediction markets time series we want to replicate. This is important because the optimal values of μ and ε can significantly change depending on the data set used for calibration, but our results suggest that our model is capable to reproduce, at least qualitatively, the empirical properties of prediction markets regardless the value of its parameters. Fig. 7.5 shows two comparisons between historical data and simulations results, obtained with both the best and worst configurations, found for the pairs found at $\mu = 0.06, \varepsilon = 0.47$ and $\mu = 0.46, \varepsilon = 0.84$, respectively. The comparisons are based on the autocorrelation of absolute return and the probability density function of absolute returns. From this figure it is possible to note that the both the best and worst configurations generate distributions of returns which are similar to the historical one, only with slightly thicker tails. Results displayed in Fig. 7.5 suggest that, whereas time series generated by the best configuration reproduce almost perfectly the profile of the absolute return autocorrelation function, those generated by the worst configuration do not match historical data accurately. Indeed, although even in this case the decay of the autocorrelation function can be modeled by a power law, its values become negative when the lag considered is greater than 25 days. This is because, for the point $\mu = 0.46, \varepsilon = 0.84$, the underlying opinion dynamics process converges to one opinion cluster far earlier than the end of the market for most market durations. This causes price changes to be little or zero, in contrast to the beginning of the market, when price changes are larger due to the high heterogeneity of opinions.

7.4 Conclusions

In this paper we propose a model that accurately describes price formation in prediction markets. To achieve this, we propose, for the first time, an exchange market model in which participants are part of a social network, and exchange opinions about the realization of a particular event following the Deffuant model. Depending on their opinion, agents buy or sell contracts on a prediction market exchange. By running agent-based simulations, we show that our model generates price time series whose statistical properties closely mimic those shown by historical data on prediction markets. Interestingly, our findings show that even in the worst-case scenario our model reproduces prediction markets qualitatively, generating price time series that display volatility clustering and fat-tailed return distributions. These results suggest that a model of prediction markets in which agents interact with each other by sharing and updating their beliefs about an event is particularly suitable to represent prediction markets. At the same time, using historical data,

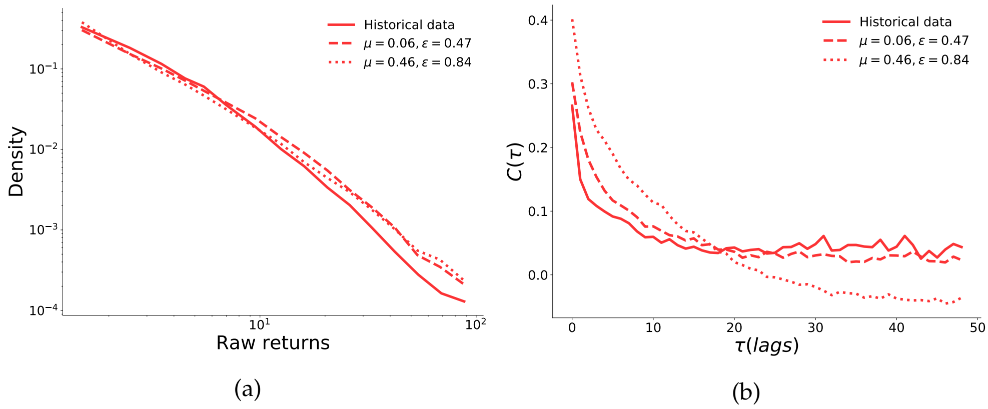


Figure 7.5: Comparison between historical data and simulation data generated by the best and worst pairs of μ, ε ($\mu = 0.06, \varepsilon = 0.47$ and $\mu = 0.46, \varepsilon = 0.84$, respectively). Fig. (a) shows the comparison between the probability density distributions of the distributions of raw returns. Fig. (b) shows the decay of the autocorrelation functions.

our findings corroborate the validity of the Deffuant model as a representation of real-world phenomena such as opinion formation with continuous opinions. In future work we intend to include endogenous and exogenous shocks that can affect both the value of the true probability of an outcome, as well as opinion shifts.

Chapter 8

Conclusions

This thesis comprises six distinct, yet interconnected papers that constitute novel work on prediction markets, and address the need for adequately complex models to study these tools and, consequently, decision making under uncertainty. Specifically, these six papers can be broadly categorised in two parts: the study of prediction market mispricing and the characterisation of the phenomena that drive price in prediction markets.

More in detail, in the first part we focus on the favourite-longshot bias (FLB), a pricing anomaly whereby bets on probable events are underpriced whereas bets on unlikely events are overpriced. In the first chapter, we show that models to explain such an anomaly fail to capture its full complexity, often being able to explain single aspects of the FLB, and often under specific conditions. Also, we introduce a model in which agents with heterogeneous beliefs and betting habits interact with each other and a bookmaker, and show that such a model can account for all forms of the FLB without need for limiting assumptions. Furthermore, we investigate two possible versions of such a model, that are different from each other in the bookmaker's behaviour. We find that, despite both models are capable of replicate historical data qualitatively, the model in which the bookmaker aims at minimise their risk by balancing the books is far more accurate at reproducing historical prices than the model in which the bookmaker seeks profit maximisation. In Chapter 3 we extend this model to consider the role of transaction costs and bookmaker fees in generating market mispricing. Our results suggest that it is the bettors' behaviour, and not transaction costs as previously suggested in the literature, that cause the FLB. In fact, we find that transaction costs only amplify the extent of the mispricing, but cannot generate it if it is not pre-existent. Chapter 4 extends our analysis of prediction market mispricing by providing an analysis of the temporal evolution of the FLB in political prediction market exchanges, contrary to the first two chapters in which only sports betting data is considered. This analysis shows that mispricing is highly correlated with the duration of the markets. Specifically, we find that the average mispricing throughout the market is lower the longer the duration, but that the extent of mispricing during the last days of trading is highly correlated with duration, and argue that this is an effect of herding.

Since in Chapter 3 we focus on prediction market exchanges, this work forms a clear link with the second part of the thesis, in which exchanges are analysed in greater detail. Specifically, we use an extensive data set of 3385 markets from PredictIt, a political prediction market platform, to study the empirical properties of prediction market time series. In these two papers, we compile a comprehensive list of stylised facts that include analysis of price changes, volume, and calendar effects. In the last chapter of the thesis, findings from these papers are used to build a agent-based model of prediction markets exchanges that matches such stylised facts. To achieve this, we consider a social network where agents have an opinion of the probability of a specific event to happen, but change this opinion after interacting with other agents.

This thesis provides several theoretical and empirical contribution to the prediction market literature. We divide these contributions in two categories, empirical and theoretical.

Empirical contributions are provided by the papers constituting chapters 4, 5 and 6. The contributions of these papers are many, but are all interconnected. Specifically, we greatly added to the literature by analysing price and volume time series in prediction markets, thereby compiling a comprehensive list of empirical properties these markets exhibit. By doing so, not only we add insight on how prediction markets work, but we also provide new and complete means of validation for models of prediction markets, which have historically failed to capture the full complexity of reality mainly for insufficiency of historical data with which compare the results.

The remaining papers offer instead theoretical contributions. In chapters 2 and 7 we introduce two agent-based models that allow the study of prediction market exchanges and mispricing in fixed-odds markets. The paper constituting chapter 2 provides several contribution to the literature. First, ours is the first model capable of reproducing both the FLB and its reverse counterpart, suggesting that a heterogeneous agent model is necessary to describe such complex phenomena as the FLB. Second, we are the first to show that bookmakers are more likely to be risk averse (i.e., they balance the books only depending on trading volume) than profit maximisers, which has been the common assumption so far in literature. Third, we show that the model here introduced can be use backwards to estimate the market composition, and argue that this can have important consequences for regulators and policy makers alike. In the paper constituting Chapter 7, we show that a simple model in which agents update their beliefs by interacting on a social network can accurately describe prediction markets, suggesting that this may be indeed the dynamics people follow when betting on prediction markets. Finally, Chapter 3 shows that transaction costs are not responsible of causing mispricing, but only to amplify so, contrary to what has been argued so far in the literature.

To conclude, this thesis represents an important step forward in understanding prediction markets, social dynamics, and decision making under uncertainty. As a whole, this collection of papers shows how important it is to appropriately model socio-economic complex systems and that, without any doubt, it takes all sorts (to make a world). And we should accept it.

Appendix A

Subjective Fair Prices Derivation

In this section we derive the subjective fair prices function shown in Section 2.2.1. Recalling the utility function for the agents

$$u(\pi, p) = w(p)v(1 - \pi) + w(1 - p)v(-\pi) \tag{A.1}$$

where the value function is

$$v(x) = \begin{cases} x^{\alpha} & \text{if } x \geqslant 0 \\ -(-x)^{\alpha} & \text{if } x < 0 \end{cases} \tag{A.2}$$

and the probability weighting function is

$$w(p) = e^{-[-ln(p)]^{\beta}} \tag{A.3}$$

we derive each subjective fair price function by varying the values of α and β.

A.0.1 Informed Bettors

The informed bettors are defined by having no misperceptions of probabilities and being risk-neutral, hence $\alpha = 1$ and $\beta = 1$. By substituting such values in equation (A.1), we obtain:

$$u(\pi, p) = (1 - \pi)e^{-[-ln(p)]} - (\pi)e^{-[-ln(1-p)]} = (1 - \pi)p - \pi(1 - p) \tag{A.4}$$

Since at equilibrium $u(\pi, p) = 0$ must hold, then $(1 - \pi)p - \pi(1 - p) = 0$ which, solved by π, gives $\pi = p$.

A.0.2 Misperceiving Bettors

Misperceiving bettors are characterised by being risk-neutral and having a coefficient $\beta = 0.928$. Hence their utility function is:

$$u(\pi, p) = (1 - \pi)e^{-[-ln(p)]^{0.928}} - (\pi)e^{-[-ln(1-p)]^{0.928}} = 0 \tag{A.5}$$

By expanding and collecting in terms of π the equation becomes:

$$e^{-[-ln(p)]^{0.928}} + \pi(-e^{-[-ln(p)]^{0.928}} - e^{-[-ln(1-p)]^{0.928}}) = 0 \tag{A.6}$$

By subtracting $e^{-[-ln(p)]^{0.928}}$ from both sides and dividing both sides by $-e^{-[-ln(p)]^{0.928}} - e^{-[-ln(1-p)]^{0.928}}$ the subjective fair price function can be rewritten as:

$$\pi = \frac{e^{-[-ln(p)]^{0.928}}}{-e^{-[-ln(p)]^{0.928}} - e^{-[-ln(1-p)]^{0.928}}} \tag{A.7}$$

A.0.3 Risk-loving Bettors

Risk-loving agents have no misperceptions of probabilities, and a risk coefficient of $\alpha = 2$. Hence their utility function is:

$$u(\pi, p) = (1-\pi)^2 e^{-[-ln(p)]} - (\pi)^2 e^{-[-ln(1-p)]} = (1-\pi)^2 p - \pi^2(1-p) = 0 \tag{A.8}$$

By expanding and collecting in terms of π we can rewrite the equation as $p - 2\pi p + \pi^2(2p-1) = 0$. Then, it is possible to divide both sides by $2p-1$ and subtract $\frac{p}{2p-1}$ from both sides, obtaining $\pi^2 - \frac{2\pi p}{2p-1} = -\frac{p}{2p-1}$. By adding $\frac{p^2}{2p-1}$ to both sides it is possible to rewrite the left-hand side of the equation as a square, and consequently to take the square root of both sides, which gives the two following solutions:

$$\pi - \frac{p}{2p-1} = \pm\sqrt{\frac{p^2}{(2p-1)^2} - \frac{p}{2p-1}} \tag{A.9}$$

This equation can be reduced to:

$$\pi = \frac{p \pm \sqrt{p-p^2}}{2p-1} \tag{A.10}$$

The solution $\pi = \frac{p+\sqrt{p-p^2}}{2p-1}$ goes to minus and plus infinity for p approaching $\frac{1}{2}$ from left and right, respectively. Therefore, this is not a realistic solution and cannot be considered as the subjective fair price function. Hence, the subjective fair price function for risk-loving agents is $\pi = \frac{p-\sqrt{p-p^2}}{2p-1}$, which is defined for $p \in [0, 0.5) \cup (0.5, 1.0]$ and has a removable discontinuity in $p = \frac{1}{2}$.

A.0.4 Risk-averse Bettors

Risk-averse bettors are characterised by the values $\alpha = 0.5$ and $\beta = 1$. Consequently, their utility function is

$$u(\pi, p) = (1-\pi)^{\frac{1}{2}} e^{-[-ln(p)]} - (\pi)^{\frac{1}{2}} e^{-[-ln(1-p)]} = (1-\pi)^2 p - \pi^2 (1-p) = 0 \quad \text{(A.11)}$$

By expanding and rearranging the terms it is possible to rewrite the equation as $2\sqrt{-\pi(\pi-1)p(p-1)} = -\pi(p-1)^2 - p^2(1-\pi)$. Then, it is possible to square both sides to obtain $-4\pi(\pi-1)p^2(p-1)^2 = (-\pi(p-1)^2 - p^2(1-\pi))^2$. By expanding and collecting again in terms of π, we obtain the following:

$$\pi^2 + \frac{\pi(4p^4 - 4p^3 + 2p^2)}{-4p^4 + 8p^3 - 8p^2 + 4p - 1} = \frac{p^2}{-4p^4 + 8p^3 - 8p^2 + 4p - 1} \quad \text{(A.12)}$$

Then, we add

$$\frac{(4p^4 - 4p^3 + 2p^2)^2}{4(-4p^4 + 8p^3 - 8p^2 + 4p - 1)^2} \quad \text{(A.13)}$$

to both sides, and rewrite the left-hand side as a square, as follows:

$$(\pi + \frac{(4p^4 - 4p^3 + 2p^2)^2}{2(-4p^4 + 8p^3 - 8p^2 + 4p - 1)^2})^2 = \\ \frac{p^2}{-4p^4 + 8p^3 - 8p^2 + 4p - 1} + \frac{(4p^4 - 4p^3 + 2p^2)^2}{4(-4p^4 + 8p^3 - 8p^2 + 4p - 1)^2} \quad \text{(A.14)}$$

The right-hand side is 0. Therefore, by taking the square root of the left-hand side, the equation becomes:

$$\pi = -\frac{(4p^4 - 4p^3 + 2p^2)^2}{2(-4p^4 + 8p^3 - 8p^2 + 4p - 1)^2} \quad \text{(A.15)}$$

which can be written as follows:

$$\pi = \frac{p^2}{2p^2 - 2p + 1} \quad \text{(A.16)}$$

Appendix B

Proof of Propositions 1-3

B.0.1 Proof of Proposition 1

Let us assume that a bookmaker sets prices π_A and π_B for the two possible outcomes of a competitive event, which have probabilities p_A and p_B, respectively. Let us assume that V is the total amount of money bet on the event, and that V_A and V_B represent the amount of money bet on the outcomes A and B, respectively.

PROPOSITION 1. With a profit-maximising bookmaker (PMB), in a market populated only by risk-loving traders with subjective fair prices defined by Eq. (2.7), there is always a FLB.

PROOF To maximise their profit, the bookmaker needs all the agents to bet on the outcome that will assure the bookmaker the maximum profit. Since the market is populated only by bettors who share the same utility function and beliefs, all the bets will be wagered on a single outcome, either A or B. Therefore, either $V_A = V$ or $V_B = V$ is true. Recalling the derivation in Appendix A, the risk-loving bettors' subjective fair price is:

$$\pi = \frac{p - \sqrt{p - p^2}}{2p - 1} \tag{B.1}$$

By substituting this equation in Eq. (3.1), we can rewrite the bookmaker's expected profit as follows:

$$\mathbb{E}(P) = V - \frac{p_A(2p_A - 1)}{p_A - \sqrt{p_A - p_A^2}} V_A - \frac{p_B(2p_B - 1)}{p_B - \sqrt{p_B - p_B^2}} V_B \tag{B.2}$$

Since there are only two possible outcomes, $p_B = 1 - p_A$, Eq. (B.2) can be rewritten as follows:

$$\mathbb{E}(P) = V - \frac{p_A(2p_A - 1)}{p_A - \sqrt{p_A - p_A^2}} V_A - \frac{(1-p_A)(2(1-p_A)-1)}{(1-p_A) - \sqrt{(1-p_A)-(1-p_A)^2}} V_B \quad \text{(B.3)}$$

Then, at equilibrium, the bookmaker should set risk-lovers' subjective fair price for A if and only if:

$$\frac{p_A(2p_A - 1)}{p_A - \sqrt{p_A - p_A^2}} < \frac{(1-p_A)(2(1-p_A)-1)}{(1-p_A) - \sqrt{(1-p_A)-(1-p_A)^2}} \quad \text{(B.4)}$$

This can be reduced to $2p_A < 1$, so that the solution is $p_A < \frac{1}{2}$. However, since p_A is the probability associated with the most likely outcome (i.e. $p_A \in (0.5, 1]$), there are no valid solutions to the inequality. Therefore, the bookmaker would lose more money should all the bets be wagered on A instead of B. This suggests that a profit-maximising bookmaker should set the maximum price risk-loving bettors are willing to pay (i.e. the subjective fair price, or $\pi_B = \pi^(s)_B$) on B, and a price $\pi_A > \pi^(s)_A$ on A. For sake of simplicity, we can assume that the bookmaker sets $\pi_A = \pi^(s)_A + \epsilon$, where ϵ has an arbitrarily small positive value. We can then substitute the equations for these prices in inequality (4.1), which must hold in order to produce the FLB. Then, it is possible to rewrite inequality (4.1) as:

$$\frac{p_A}{1-p_A} > \frac{(2(1-p_A)-1)(p_A - \sqrt{p_A - p_A^2})}{(2p_A - 1)(1 - p_A - \sqrt{(1-p_A)-(1-p_A)^2})} \quad \text{(B.5)}$$

By subtracting $\frac{p_A}{1-p_A}$ from both sides and simplifying the expression, the inequality can be rewritten as:

$$\frac{2p_A - \sqrt{p_A - p_A^2}}{p_A - 1} < 0 \quad \text{(B.6)}$$

The denominator is positive if $p_A > 1$ and negative otherwise, while the numerator is positive only for $p_A > \frac{1}{5}$. In the domain given by $p_A \in [0.5, 1)$, the denominator is always negative and the numerator always positive. Hence inequality (B.6) always holds. Therefore, if the bookmaker is a PMB and all the bettors are risk-lovers, the FLB always exists.

B.0.2 Proof of Proposition 2

PROPOSITION 2. With a profit-maximising bookmaker (PMB), in a market populated only by risk-averse traders with subjective fair prices defined by Eq. (2.6), there is always a reverse FLB.

To prove proposition 2, we will follow the same reasoning we followed for proposition 1. Let assume all bettors are risk-averse. That is, their subjective fair price function is:

$$\pi(p) = \frac{p^2}{2p^2 - 2p + 1} \tag{B.7}$$

Recalling Appendix B.0.1, the bookmaker should set risk-averse bettors' subjective fair price for A if and only if the following holds:

$$\frac{p_A(2p_A^2 - 2p_A + 1)}{p_A^2} < \frac{(1 - p_A)(2(1 - p_A)^2 - 2(1 - p_A) + 1}{(1 - p_A)^2} \tag{B.8}$$

By expanding and rearranging the terms the inequality can be rewritten as:

$$\frac{(p_A - \frac{1}{2})(p_A^3 - \frac{5}{2}p_A^2 + \frac{3}{2}p_A - \frac{1}{2})}{(p_A - 1)^2 p_A} < 0 \tag{B.9}$$

By analysing the sign of the denominator, it is possible to note that, since $p_A \in [0.5, 1)$, the denominator is always positive. Similarly, the term $p_A - \frac{1}{2}$ is always positive. To find the solutions to the inequality, we only need to find the solutions to $(p_A^3 - \frac{5}{2}p_A^2 + \frac{3}{2}p_A - \frac{1}{2}) < 0$. It is possible to rewrite this as $p_A(p_A - 1)(2p_A - 3) < 1$. Clearly, for $p_A \in [0.5, 1)$, the function $f(p_A) = p_A(p_A - 1)(2p_A - 3)$ has a maximum in $f(1) = 0.5$. Hence, it is always less than one, and the inequality always holds. Therefore, when all bettors are risk-averse, for a PMB, it is optimal that the agents bet on A. Following again the procedure employed to prove proposition 1, we rewrite the condition for the FLB to exist in terms of risk-averse bettors' subjective fair prices. However, this time we want to prove that the reverse FLB exists, so we want that $\frac{p_A}{p_B} < \frac{\pi_A}{\pi_B}$:

$$\frac{p_A}{1 - p_A} < \frac{p_A(2(1 - p_A)^2 - 2(1 - p_A) + 2)}{(1 - p_A)(2p_A^2 - 2p_A + 1)} \tag{B.10}$$

By expanding, collecting, and rearranging the terms, it is possible to simplify the expression to $(p_A - 1)^2 p_A(2p_A - 1) > 0$. Recalling that $p_A \in (0.5, 1)$, it is clear that all the terms are always positive in this domain. Therefore, we can conclude that in presence of only risk-averse bettors, if the bookmaker is a PMB, there is always a negative FLB.

B.0.3 Proof of Proposition 3

PROPOSITION 3. With a risk-minimising bookmaker (RMB), in a market in which relative volumes are not equal to true probabilities, there is always either the FLB or the reverse FLB.

PROOF Let k be the percentage deviation of V_A from the true probability p_A on the outcome A, such that $V_A = p_A(1 + k)$. Therefore, since $V_A + V_B = 1$, we obtain the following system of equations for the relative volumes V_A and V_B:

$$\begin{cases} V_A = p_A(1 + k) \\ V_B = p_B - p_A k \end{cases} \tag{B.11}$$

Recalling that a RMB sets the prices $\hat{\pi}_A = V_A$ and $\hat{\pi}_B = V_B$, we can substitute these equations of the volumes in $\frac{p_A}{p_B} > \frac{\pi_A}{\pi_B}$ (the necessary condition for the FLB to happen, assuming $p_A > p_B$). Then, we can rewrite it as:

$$\frac{p_A}{p_B} > \frac{p_A + p_A k}{p_B - p_A k} \tag{B.12}$$

To solve the inequality for k, we first rewrite the inequality in terms of p_A, as follows:

$$\frac{p_A}{1 - p_A} > \frac{p_A(1 + k)}{1 - p_A - p_A k} \tag{B.13}$$

Since $p_A(1 + k)$ represents the relative volume wagered on outcome A, the bounds are $0 \leqslant p_A(1 + k) \leqslant 1$, which implies $-1 \leqslant k \leqslant \frac{1}{p_A} - 1$.

By rearranging the terms, inequality (B.13) becomes $\frac{k p_A}{(p_A - 1)(k p_A + p_A - 1)} < 0$. Considering the boundaries for p_A and k, we can write a system of inequalities as follows:

$$\begin{cases} \frac{kp_A}{(p_A-1)(kp_A+p_A-1)} < 0 \\ -1 \leqslant k \leqslant \frac{1}{p_A} - 1 \\ \frac{1}{2} < p_A \leqslant 1 \end{cases} \tag{B.14}$$

Inequality (B.13) has many solutions, but the system has only one, that is $-1 < k < 0$. Therefore, with a RMB, the FLB exists only if the relative volume of money bet on A is greater than the probability of A happening, i.e. $V_A > p_A$. However, following the same reasoning, one can prove that this implies that, if $k > 0$, a negative FLB exists. Furthermore, this implies that a market structured in this way can only be efficient if $k = 0$, i.e. if $V_A = p_A$ and $V_B = p_B$.

Appendix C

Results on Prediction Markets

In this section, we perform an additional analysis on political prediction markets using the RMB model presented in Chapter 2. Our data comprises 3363 observations from the exchange platform PredictIt (www.predictit.org), which are used to reconstruct a price curve for prediction markets. Although our model could be extended to reproduce exchange markets (i.e., by removing bookmakers and allowing agents to short sell), this is beyond the scope of this paper. Therefore, we present our analysis solely with the intention of showing that the results obtained using sports betting data are robust and not the result of overfitting. Our results, displayed in Figure .7, suggest that the RMB can capture the positive FLB exhibited by prediction markets. However, the large value of the MSE (0.0051) indicates that our model is not capable to fully explain the data. Indeed, the MSE computed between historical prices and those results of the simulations, is an order of magnitude greater than those found on the Under-over 2.5 and Baseball markets, and 170 times larger than the error on Tennis. Since our model is not suitable to analyse exchange markets, such results were expected, and provide additional evidence that the price curves reconstructed by the RMB for sports betting markets are not a consequence of overfitting, rather, they are a genuine display of the our model's accuracy. Similarly, the market composition identified by the RMB suggests that, for prediction markets, participants are mostly risk-averse (75%), while only some of them are represented by informed bettors (20%), and the rest are accounted as noise traders (5%). This market composition, however, is difficult to be validated, due to the lack of information about the trader population of prediction markets.

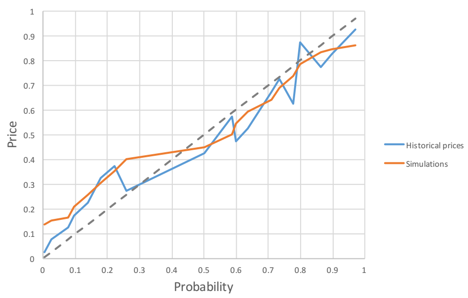


Figure C.1: Comparison between equilibrium prices generated by the best market compositions (i.e. those market compositions for which the prices generated are closer to historical prices) in political prediction markets and historical prices. The dashed line represents the fair prices.

Bibliography

Agrawal, A. and Tandon, K. Anomalies or illusions? Evidence from stock markets in eighteen countries. *Journal of International Money and Finance*, 13: 083–106, 1994.

Alfi, V., Gabrielli, A., and Pietronero, L. How people react to deadlines: time distribution of conference registrations and fee payments. *Central European Journal of Physics*, 7:483–489, 2009.

Ali, M. M. Probability and Utility Estimates for Racetrack Bettors. *The Journal of Political Economy*, 85(4):803, 1977.

Armatas, V., Yiannakos, A., and Sileloglou, P. Relationship between time and goal scoring in soccer games: Analysis of three World Cups. *International Journal of Performance and Analysis in Sport*, 7(2):48–58, 2007.

Arnesen, S. and Bergfjord, O. Prediction markets vs polls–an examination of accuracy for the 2008 and 2012 elections. *Journal of Prediction Markets*, 8(3): 24–33, 2014.

Ausloos, M., Vandewalle, N., and Boveroux, P. Applications of statistical physics to economic and financial topics. *Physica A: Statistical Mechanics and its Applications*, 274:229–240, 1999.

Ausloos, M. Toward fits to scaling-like data, but with inflection points and generalized Lavalette function. *Journal of Applied Quantitative Methods*, 9: 1–21, 2014.

Bachelier, L. Théorie de la spéculation. *Annales scientifiques de l'E.N.S.*, 17: 21–86, 1900.

Barabási, A.-L. and Albert, R. Emergence of scaling in random networks. *Science*, 5439(286):509–512, 1999.

Bariviera, A. F., Basgalla, M. J., Hasperué, W., and Naiouf, M. Some stylized facts of the Bitcoin market. *Physica A: Statistical Mechanics and its Applications*, 484:82–90, 2017. doi: https://doi.org/10.1016/j.physa.2017.04.159.

Baronchelli, A., Dall'Asta, L., Barrat, A., and Loreto, V. Topology Induced Coarsening in Language Games. *Physical Review E*, 73(015102(R)), 2006.

Bauke, H. Parameter estimation for power-law distributions by maximum likelihood methods. *The European Physical Journal B*, 58(2):167–173, 2007.

Berg, J., Nelson, F., and Rietz, T. Prediction market accuracy in the long run. *International Journal of Forecasting*, 24(285-300), 2008.

Bhardwaj, R. and Brooks, L. The January Anomaly: Effects of Low Share Price, Transaction Costs, and Bid-Ask Bias. *The Journal of Finance*, 47(2):553–575, 1992.

Bittner, E., Nussbaumer, A., Janke, W., and Weigel, M. Self-affirmation model for football goal distributions. *Europhysics Letters*, 78(5):1–5, 2007.

Blattberg, R. and Gonedes, N. A Comparison of the Stable and Student Distributions as Statistical Models for Stock Prices. *The Journal of Business*, 47(2): 244–280, 1974.

Bouchaud, J.-P. Power laws in economics and finance: some ideas from physics. *Quantitative Finance*, 1(1):105–112, 2001.

Brock, W. and Hommes, C. A Rational Route to Randomness. *Econometrica: Journal of the Econometric Society*, 65(5):1059–1095, 1997.

Bruce, A., Johnson, J. E., and Peirson, J. Recreational versus professional bettors: performance differences and efficiency implications. *Economics Letters*, 114(2):172–174, 2012.

Buckley, W. S. and Long, H. A discontinuous mispricing model under asymmetric information. *European Journal of Operational Research*, 243(3):944–955, 2015.

Buckley, W. S., Brown, G. O., and Marshall, M. A mispricing model of stocks under asymmetric information. *European Journal of Operational Research*, 221 (3):584–592, 2012.

Bucsa, G., Jovanovic, F., and Schinckus, C. A unified model for price return distributions used in econophysics. *Physica A*, 390(20):3435–3443, 2011.

Busche, K. and Hall, C. D. An exception to the risk preference anomaly. *The Journal of Business*, 61(3):337–346, 1988.

Cain, M., Law, D., and Peel, D. The Favourite-Longshot Bias, Bookmaker Margins and Insider Trading in a Variety of Betting Markets. *Bulletin of Economic Research*, 55:3(1998):263–274, 2003.

Campbell, J., Lo, A., and MacKinlay, A. *The econometrics of financial markets*. Princeton University Press, Princeton, New Jersey, 1997. ISBN 0691043019.

Castellano, C., Fortunato, S., and Loreto, V. Statistical physics of social dynamics. *Review of Modern Physics*, 81:591–646, 2009.

Chakraborti, A., Toke, I. M., Patriarca, M., and Abergel, F. Econophysics review: I. Empirical facts. *Quantitative Finance*, 11(7):991–1012, 2011.

Chiappori, P., Gandhi, A., Salanié, B., and Salanié, F. From Aggregate Betting Data to Individual Risk Preferences. *Columbia University Working Paper, n. 13-453*, 2012.

Chordia, T., Roll, R., and Subrahmanyam, A. Market Liquidity and Trading Activity. *The Journal of Finance*, LVI(2), 2001.

Cizeau, P., Liu, Y., Meyer, M., Peng, C.-K., and Stanley, H. E. Volatility distribution in the S&P500 stock index. *Physica A: Statistical Mechanics and its Applications*, 245:441–445, 1997.

Clauset, A., Shalizi, C., and Newman, M. Power-law distributions in empirical data. *SIAM review*, 51(4):661–703, 2009.

Clifford, P. and Sudbury, A. Minority Opinion Spreading in Random Geometry. *European Physical Journal B*, 25:403–406, 2002.

Constantinides, G., Harris, M., and Stulz, R. *Handbook of the Economics of Finance*. Elsevier North-Holland, Amsterdam, 2003.

Cont, R. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2):223–236, 2001.

Cont, R., Potters, M., and Bouchaud, J.-P. Scaling in stock market data: stable laws and beyond. In *Scale invariance and beyond*, pages 75–85. 1997.

Cowgill, B., Wolfers, J., and Zitzewitz, E. Using prediction markets to track information flows evidence from google. In Das, S., Ostrovsky, M., Pennock, D., and Szymanski, B., editors, *Auctions, Market Mechanisms and Their Applications.* Springer, Berlin, lecture no edition, 2009.

Crawford, I. and Pendakur, K. How many types are there? *The Economic Journal*, 123:77–95, 2013.

Croxson, K. and Reade, J. J. Information and Efficiency: Goal Arrival in Soccer Betting. *The Economic Journal*, 124(575):62–91, 2014.

Deffuant, G., Neau, D., Amblard, F., and Weisbuch, G. Mixing beliefs among interacting agents. *Advances in Complex Systems*, 87(03), 2000.

del Corral, J. and Prieto-Rodríguez, J. Are differences in ranks good predictors for Grand Slam tennis matches? *International Journal of Forecasting*, 26(3): 551–563, 2010.

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., and Quattrociocchi, W. The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113(3):554–559, 2016.

Del Vicario, M., Zollo, F., Caldarelli, G., Scala, A., and Quattrociocchi, W. Mapping social dynamics on Facebook: The Brexit debate. *Social Networks*, 50: 6–16, 2017.

Deluca, A. and Corral, A. Fitting and goodness-of-fit test of non-truncated and truncated power-law distributions. *Acta Geophysica*, 61(6):1351–1394, 2013.

Di Matteo, T., Aste, T., and Dacorogna, M. Scaling behaviors in differently developed markets. *Physica A: Statistical Mechanics and its Applications*, 324 (1-2):183–188, 2003.

Dzhabarov, C. and Ziemba, W. T. Do seasonal anomalies still work? *The Journal of Portfolio Management*, 36(3):93–104, 2010.

Fama, E. The Behavior of Stock-Market Prices. *The Journal of Business*, 38(1): 34–105, 1965.

Fama, E. Efficient Capital Markets: A Review of Theory and Empirical Work. *The journal of Finance*, 25(2):28–30, 1970.

Feess, E., Muller, H., and Schumacher, C. Estimating risk preferences of bettors with different bet sizes. *European Journal of Operational Research*, 249(3):1102–1112, 2016.

Fields, M. Stock prices: A problem in verification. *The Journal of Business of the University of Chicago*, 4(4):415–418, 1931.

Fingleton, J. and Waldron, P. Optimal Determination of Bookmakers' Betting Odds: Theory and Tests. *Trinity Economic Paper Series*, 9(96), 1999.

Forrest, D. and McHale, I. Longshot bias: insights from the betting market on men's professional tennis. In Vaughan Williams, L., editor, *Information Efficiency in Financial and Betting Markets*, chapter 8, pages 215–230. Cambridge University Press, 2005. ISBN 9780521816038.

Forrest, D. and Mchale, I. Anyone for Tennis (Betting)? *The European Journal of Finance*, 13(8):751–768, 2007.

Forrest, D. and Simmons, R. Sentiment in the betting market on Spanish football. *Applied Economics*, 40(1):119–126, 2008.

French, K. R. Stock Returns and the Weekend Effect. *Journal of Financial Economics*, 8:55–69, 1980.

Gabaix, X., Gopikrishnan, P., Plerou, V., and Stanley, E. A unified econophysics explanation for the power-law exponents of stock market activity. *Physica A: Statistical Mechanics and its Applications*, 382:81–88, 2007.

Gandhi, A. and Serrano-Padial, R. Does Belief Heterogeneity Explain Asset Prices: The Case of the Longshot Bias. *The Review of Economic Studies*, 82(1):156–186, 2015.

Gilli, M. and Winker, P. A global optimization heuristic for estimating agent based models. *Computational Statistics and Data Analysis*, 43(3):299–312, 2003.

Golec, J. and Tamarkin, M. Bettors love skewness, not risk, at the horse track. *Journal of Political Economy*, 106(1):205–255, 1998.

Gonzalez, R. and Wu, G. On the shape of the probability weighting function. *Cognitive psychology*, 166:129–166, 1999.

Goodell, J. W. and Bodey, R. A. Price-earnings changes during US presidential election cycles: Voter uncertainty and other determinants. *Public Choice*, 150 (3-4):633–650, 2012.

Goodell, J., McGroarty, F., and Urquhart, A. Political uncertainty and the 2012 US presidential election : A cointegration study of prediction markets, polls and a stand-out expert. *International Review of Financial Analysis*, (February), 2015.

Goodhart, C. A. E. and Hara, M. O. High frequency data in financial markets : Issues and applications. *Journal of Empirical Finance*, 4:73–114, 1997.

Gopikrishnan, P., Plerou, V., Gabaix, X., and Stanley, H. E. Statistical Properties of Share Volume Traded in Financial Markets. *Physical Review E*, 62:R4493–R4496, 2000.

Greene, M. T. and Fielitz, B. D. Long-term dependence in common stock returns. *Journal of Financial Economics*, 4:339–349, 1977.

Griffith, R. Odds adjustements by american horse-race bettors. *The American Journal of Psychology*, pages 9–13, 1949.

Gultekin, M. and Gultekin, N. Stock Market Seasonality: International Evidence. *Journal of Financial Economics*, 12(December):469–481, 1983.

Haugen, R. and Lakonishok, J. *The Incredible January Effect*. Dow-Jones-Irwin, Homewood, IL, 1988.

Heckman, J. J. Micro Data , Heterogeneity , and the Evaluation of Public Policy: Nobel Lecture. *Journal of Political Economy*, 109(4):673–748, 2001.

Hodges, S. D., Tompkins, R., and Ziemba, W. T. The Favorite Longshot Bias in S & P 500 and FTSE 100 Index Futures Options: The Return to Bets and the Cost of Insurance. *EFA 2003 Annual Conference Paper No. 135; Sauder School of Business Working Paper*, 2003.

Hommes, C. H. Heterogeneous agent models in economics and finance. In Tesfatsion, L. and Judd, K. L., editors, *Handbook of computational economics, Volume 2*, volume 2, chapter 23, pages 1110–1186. Elsevier B.V., 2006.

Hurley, W. and McDonough, L. The Favourite-Longshot Bias in Parimutuel Betting: A Clarification of the Explanation That Bettors Like to Bet on Longshots. *Economics Letters*, 52:275–278, 1996.

Hurst, H. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116:770–799, 1951.

Hvattum, L. M. and Arntzen, H. Using ELO ratings for match result prediction in association football. *International Journal of Forecasting*, 26(3):460–470, 2010.

Jain, P. and Joh, G.-H. The Dependence between HoIGNOREdoiy Prices and Trading Volume. *The Journal of Financial and Quantitative Analysis*, 23(3):269–283, 1988.

Jensen, M., Johansen, a., Petroni, F., and Simonsen, I. Inverse statistics in the foreign exchange market. *Physica A: Statistical Mechanics and its Applications*, 340(4):678–684, 2004.

Jensen, M. H., Johansen, A., and Simonsen, I. Inverse Statistics in Economics : The gain-loss asymmetry. *Physica A: Statistical Mechanics and its Applications*, 324(1-2):338–343, 2003.

Joëts, M. Heterogeneous beliefs, regret, and uncertainty: The role of speculation in energy price dynamics. *European Journal of Operational Research*, 247 (1):204–215, 2015.

Jovanovic, F. and Schinckus, C. *Econophysics and Financial Economics: An Emerging Dialogue*. Oxford University Press, 2017.

Jullien, B. and Salanié, B. Estimating Preferences under Risk : The Case of Racetrack Bettors. *Journal of Political Economy*, 108(3):503–530, 2000.

Kahneman, D. and Tversky, A. Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2):263–292, 1979.

Karpio, K., Zaluska-Kotur, M. A., and Orlowski, A. Gain–loss asymmetry for emerging stock markets. *Physica A: Statistical Mechanics and its Applications*, 375:599–604, 2007.

Kirman, A. P. Whom or What Does the Representative Individual Represent? *Journal of Economic Perspectives*, 6(2):117–136, 1992.

Kopriva, F. Constant Bet Size? Don't Bet on It! Testing Expected Utility Theory on Betfair Data. *CERGE-EI Working Paper series*, 2009.

Lakonishok, J. and Maberly, E. The Weekend Effect: Trading Patterns of Individual and Institutional Investors. *The Journal of Finance*, 45(1):231–243, 1990.

Levitt, S. D. Why are gambling markets organised so differently from financial markets? *The Economic Journal*, 114:223–246, 2004.

Lillo, F. and Farmer, J. D. The Long Memory of the Effcient Market. *Studies in Nonlinear Dynamics and Econometrics*, 8(3), 2004.

Liu, Y., Cizeau, P., Meyer, M., Peng, C.-K., and Stanley, H. E. Correlations in economic time series. *Physica A: Statistical Mechanics and its Applications*, 245: 437–440, 1997.

Lo, A. Long-Term Memory in Stock Market Prices. *Econometrica*, 59(5):1279–1313, 1991.

Lorenz, J. *Managing Complexity: Insights, Concepts, Applications*. Springer, Berlin.

Lux, T. The stable Paretian hypothesis and the frequency of large returns: an examination of major German stocks. *Applied Financial Economics*, 6(6):463–475, 1996.

Lux, T. and Marchesi, M. Scaling and criticality in a stochastic multi-agent model of a Financial market. *Nature*, 397(February), 1999a.

Lux, T. and Marchesi, M. Scaling and criticality in a stochastic multi-agent model of a Financial market. *Nature*, 397(February):498–500, 1999b.

Lux, T. and Marchesi, M. Volatility Clustering in Financial Markets: a Microsimulation of Interacting Agents. *International Journal of Theoretical and Applied Finance*, 03(04):675–702, 2000.

Ma, T., Tang, L., McGroarty, F., Sung, M.-C., and Johnson, J. E. V. Time is money: Costing the impact of duration misperception in market prices. *European Journal of Operational Research*, 255(2):397–410, 2016.

Malevergne, Y., Pisarenko, V., and Sornette, D. Empirical distributions of stock returns: between the stretched exponential and the power law? *Quantitative Finance*, 5(4):379–401, 2005.

Malueg, D. A. and Yates, A. J. Testing Contest Theory: Evidence from Best-of-Three Tennis Matches. *Review of Economics and Statistics*, 92:689–692, 2010.

Mandelbrot, B. The Variation of Certain Speculative Prices. *The Journal of Business*, 36(4):394–419, 1963.

Mandelbrot, B. When Can Price be Arbitraged Efficiently? A Limit to the Validity of the Random Walk and Martingale Models. *The Review of Economics and Statistics*, 53(3):225–236, 1971.

Mandelbrot, B. Statistical Methodology for Nonperiodic Cycles: From the Covariance to R/S Analysis. *Annals of Economic and Social Measurement*, 1(3): 259–290, 1972.

Mantegna, R. and Stanley, H. E. *An introduction to econophysics*. Cambridge University Press, 2000.

Mclean, R. D. and Pontiff, J. Does Academic Research Destroy Stock Return Predictability? *The Journal of Finance*, 71(1):5–32, 2016.

Milgrom, P. and Stokey, N. Information, trade and common knowledge. *Journal of Economic Theory*, 1982.

Newman, M. E. J. Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, (46):323–351, 2004.

Officer, R. Seasonality in Australian capital markets: Market efficiency and empirical issues. *Journal of Financial Economics*, 2(1):29–51, 1975.

O'Leary, D. E. Prediction Markets as a Forecasting Tool. In Lawrence, K., editor, *Advances in Business and Management Forecasting*, volume 8, chapter 8, pages 169–184. Emerald Group Publishing Ltd, London, 2011.

Ottaviani, M. and Sørensen, P. N. Late Informed Betting and the Favorite-Longshot Bias. (July), 2003.

Ottaviani, M. and Sørensen, P. N. The Favorite-Longshot Bias : An Overview of the Main Explanations. In Hausch, D. B. and Ziemba, W. T., editors, *Handbook of Sports and Lottery markets*, chapter 6, pages 83–101. North-Holland, 2008.

Pagan, A. The Econometrics of Financial Markets. *Journal of Empirical Finance*, 3:15–102, 1996.

Page, L. and Clemen, R. T. Do prediction markets produce well-calibrated probability forecast? *The Economic Journal*, 123(May):491–513, 2012.

Peiro, A. Skewness in Financial Returns. *Journal of Banking & Finance*, 23, 1999.

Peng, C.-K. Mosaic organization of DNA nucleotides. *Physical Review E*, 49(2): 1685–1689, 1994.

Plerou, V., Gopikrishnan, P., Amaral, L. a. N., Gabaix, X., and Stanley, H. E. Economic Fluctuations and Anomalous Diffusion. *Physical Review E*, 62: R3023–R3026, 2000.

Plerou, V., Gopikrishnan, P., Gabaix, X., Amaral, L. A. N., and Stanley, H. E. Price fluctuations, market activity and trading volume. *Quantitative Finance*, 1:2:262–269, 2001.

Plerou, V., Gopikrishnan, P., Gabaix, X., and Stanley, H. E. On the Origin of Power-Law Fluctuations in Stock Prices. *Quantitative Finance*, 4:C11–C15, 2004.

Plott, C. and Chen, K. Information aggregation mechanism: concept, design, and implementation for a sales forecasting problem. 2002.

Podobnik, B., Horvatic, D., Petersen, A. M., and Stanley, H. E. Cross-correlations between volume change and price change. *PNAS*, 106(52): 22079–22084, 2009.

Poterba, J. and Weisbenner, S. Capital gains tax rules, tax-loss trading, and turn-of-the-year returns. *The Journal of Finance*, 51(1):353–368, 2001.

Praetz, P. The Distribution of Share Price Changes. *Journal of business*, 45(1): 49–55, 1972.

Prelec, D. The Probability Weighting Function. *Econometrica*, 66(3):497–527, 1998.

Qiu, T., Zhong, L. X., Chen, G., and Wu, X. R. Statistical properties of trading volume of Chinese stocks. *Physica A*, 388(12):2427–2434, 2009.

Reinganum. The anomalous stock market behavior of small firms in January: Empirical tests for tax-loss selling effects. *Journal of Financial Economics*, 12 (1):89–104, 1983.

Restocchi, V., McGroarty, F., and Gerding, E. The stylized facts of prediction markets: analysis of price changes. *Under review: Physica A*, 2018a.

Restocchi, V., McGroarty, F., and Gerding, E. Statistical properties of volume and calendar effects in prediction markets. *Under review: Physica A*, 2018b.

Restocchi, V., McGroarty, F., Gerding, E., and Johnson, J. E. It Takes All Sorts: A Heterogeneous Agent Explanation for Prediction Market Mispricing. *European Journal of Operational Research - doi: 10.1016/j.ejor.2018.04.011*, 2018c.

Rhoda, K. L., Olson, G. T., and Rappaport, J. M. Risk preferences and information flows in racetrack betting markets. *The Journal of Financial Research*, XXII(3):265–285, 1999.

Rhode, P. W. and Strumpf, K. S. Historical Presidential Betting Markets. *Journal of Economic Perspectives*, 18(2):127–142, 2012.

Richmond, P., J., M., , and S., H. *Econophysics and Physical Economics*. Oxford University Press, 2013.

Roll, R. Vas ist das? The turn-of-the-year effect and the return premia of small firms. *The Journal of Portfolio Management*, 9(2):18–28, 1983.

Rozeff, M. and Kinney, W. Capital Market Seasonality: the Case of Stock Returns. *Journal of Financial Economics*, 3(4):379–402, 1976.

Rubinstein, M. Rational Markets: Yes or No? The Affirmative Case. *Financial Analysts Journal*, 57(3):15–29, 2001.

Sauer, R. D. The Economics of Wagering Markets. *Journal of Economic Literature*, 36(4):2021–2064, 1998.

Sewell, M. Characterization of financial time series. *Research Note RN/11/ 01*, UCL, London, 2011.

Shin, H. S. Optimal Betting Odds Against Insider Traders. *The Economic Journal*, 191(408):1179–1185, 1991.

Shin, H. S. Prices of State Contingent Claims with Insider Traders, and the Favourite-Longshot Bias. *The Economic Journal*, 102(411):426–435, 1992.

Shin, H. S. Measuring the Incidence of Insider Trading in a Market for State-Contingent Claims. *The Economic Journal*, 103(420):1141–1153, 1993.

Sias, R. and Starks, L. Institutions and individuals at the turn-of-the-year. *The Journal of Finance*, 52(4):1543–1562, 1997.

Smith, M. A., Paton, D., and Williams, L. V. Market Efficiency in Person-to-Person Betting. *Economica*, 73(292):673–689, 2006.

Snowberg, E. and Wolfers, J. Explaining the Favorite–Long Shot Bias: Is it Risk-Love or Misperceptions? *Journal of Political Economy*, 118(4):723–746, 2010.

Sobkowicz, P. Modelling opinion formation with physics tools: Call for closer link with reality. *Journal of Artificial Societies and Social Simulations*, 12(1), 2009.

Sornette, D. *Why stock markets crash?* Princeton University Press, 2003.

Sornette, D. *Critical phenomena in natural sciences: chaos, fractals, selforganization and disorder: concepts and tools*. Springer-Verlag, Berlin, 2nd edition, 2006.

Sznajd-Weron, K. and Sznajd, J. Opinion evolution in closed community. *International Journal of Modern Physics C*, 11(06):1157, 2000.

Tessone, C. J., Toral, R., Amengual, P., Wio, H. S., and Miguel, M. S. Neighborhood models of minority opinion spreading. *European Physical Journal B*, 39 (4):535–544, 2004.

Teverovsky, V., Taqqu, M. S., and Willinger, W. A critical look at Lo's modified R/S statistic. *Journal of Statistical Planning and Inference*, 80:211–227, 1999.

Thaler, R. H. *The Winner's Curse: Paradoxes and Anomalies of Economic Life*. Princeton, New Jersey, 1992.

Toral, R. and Tessone, C. J. Finite size effects in the dynamics of opinion formation. *Communications in Computational Physics*, 2(2):177–195, 2006.

Tversky, A. and Kahneman, D. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323, 1992.

Vandewalle, N. and Ausloos, M. Coherent and random sequences in financial fluctuations. *Physica A: Statistical Mechanics and its Applications*, 246:454–459, 1997.

Vaughan Williams, L. and Paton, D. Why are some favourite-longshot biases positive and others negative? *Applied Economics*, 30:1505–1510, 1998.

Vaughan Williams, L., Sung, M.-C., Fraser-Mackenzie, P., Peirson, J., and Johnson, J. Towards an Understanding of the Origins of the Favourite-Longshot Bias: Evidence from Online Poker Markets, a Real-money Natural Laboratory. *Economica*, 2016.

Wachtel, S. Certain observations on seasonal movements in stock prices. *The Journal of Business of the University of Chicago*, 15(2), 1942.

Weisbuch, G., Deffuant, G., Amblard, F., and Nadal, J. P. Interacting agents and continuous opinions dynamics. In Fandel, G. and Trockel, W., editors, *Lecture Notes in Economics and Mathematical Systems - No. 521*, pages 225–242. Springer Verlag, Berlin, 2002.

Wolfers, J. and Zitzewitz, E. Prediction Markets. *Journal of Economic Perspectives*, 18(2):107–126, 2004.

Wolfers, J. and Zitzewitz, E. Prediction Markets in Theory and Practice. 2006.

Woodland, L. M. and Woodland, B. M. Market Efficiency and the Favorite-Longshot Bias: The Baseball Betting Market. *The Journal of Finance*, 49(1): 269, 1994.

Woodland, L. M. and Woodland, B. M. The Reverse Favourite-longshot Bias and Market Efficiency in Major League Baseball: An Update. *Bulletin of Economic Research*, 55(2):113–123, 2003.

Zhou, W.-X. and Sornette, D. Self-fulfilling Ising Model of Financial Markets. *European Physical Journal B*, pages 175–181.