

Article

Not peer-reviewed version

Opinion Dynamics Explain Price Formation in Prediction Markets

[Valerio Restocchi](#)*, [Frank McGroarty](#), [Enrico Gerding](#), [Markus Brede](#)

Posted Date: 21 July 2023

doi: 10.20944/preprints202307.1444.v1

Keywords: Opinion dynamics; Price formation; Econophysics; Prediction markets; Complex networks; Agent-based modelling.



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Opinion Dynamics Explain Price Formation in Prediction Markets

Valerio Restocchi ^{1,*}, Frank McGroarty², Enrico Gerding ³, and Markus Brede ³

¹ School of Informatics, The University of Edinburgh

² School of Business, University of Southampton

³ School of Electronics and Computer Science, University of Southampton

* Correspondence: v.restocchi@ed.ac.uk

Abstract: Prediction markets are heralded as powerful forecasting tools, but models to describe them often fail to capture the full complexity of the underlying mechanisms that drive price dynamics. To address this issue, we propose a model in which agents belong in a social network, 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 neighbours in the network, following the Deffuant model of opinion dynamics. Our results suggest that a simple market model that takes into account opinion formation dynamics is capable to replicate the empirical properties of historical prediction market time series, including volatility clustering and fat-tailed distribution of returns. Interestingly, the best results are obtained when there is the right level of variance in the opinions of the agents. Moreover, this paper provides a new way to indirectly validate opinion dynamics models against real data by using historical data obtained from PredictIt, an exchange platform whose data has never been used before to validate models of opinion diffusion.

Keywords: opinion dynamics; econophysics; prediction markets; complex networks; agent-based modelling

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 [1,2]. Although prediction markets are often heralded as an effective mechanism to make highly accurate predictions [3,4], it has been shown that they are prone to bias [5] and manipulation [6], and that these phenomena can spread to financial markets with dreadful consequences [6,7].

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 [5,8–10], 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 [11]. Opinion dynamics have 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, [11–15] for seminal opinion dynamics models and [16] for a thorough review of the topic). Research has also been done using opinion dynamics models with binary opinions to model agents in the stock market [17–20]. 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, a common (and often necessary) practice for models of financial markets [21,22]. 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.

This paper makes three important contributions. First, we introduce a model of prediction markets that uses 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, as 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 show that our model reproduces historical data best when opinions are heterogeneous but revolve around a single opinion cluster. This suggests that participants in prediction markets tend to have a similar, yet not identical view on the outcome, especially in markets that have a limited duration. Third, our results support the presence of Deffuant-like opinion dynamics. This contribution helps tackling an important challenge that Sobkowicz posed by arguing that opinion dynamics models are often disconnected from the real world and lack of empirical validation [23]. Since his paper, the explosion of popularity social media experienced has certainly offered means of compelling validation [24,25]. However, the availability and popularity of such vast data sets resulted in most opinion dynamics models being validated only on social media data, potentially introducing a bias. In this paper, we use market data to provide evidence a continuous opinion diffusion models, and specifically the Deffuant model, is compatible with the exchange of opinions in prediction markets. 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 underlying opinion diffusion process in social networks when opinions are scalar.

The remainder of the paper is organized as follows. In Section 2 we define the model of opinion formation and market exchange. In Section 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 4.

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. Markets for which multiple options are available, e.g., *Who will win the presidential elections?* can be seen as markets in which there are N events, where N is the number of candidates running for president. The payoff of a contract on the i -th event is 1 if the corresponding event occurs and 0 otherwise. Let us denote with $\tilde{\pi}_i \in (0,1)$ 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 [11]. 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:

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

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 [26,27]. In general, ε affects the number of clusters at equilibrium (i.e., the number of opinions that coexist), while μ drives the convergence time [11].

To reflect the temporal volume dynamics observed by Restocchi et al. [28], 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 represents the number of days elapsed from the beginning of the market (with $0 \leq \tau < T$), and $\gamma = 2.44$ is a scaling parameter, estimated on the same dataset [28]. 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 $\pi = o_i$. Following influential agent-based models of financial markets, we assume that price is purely driven by excess demand. Agents can only buy/sell at the current trading price and contracts are created whenever necessary. The demand of agent i , D_i , is proportional to the distance between their opinion and the price, and is described by

$$D_i(t) = o_i(t) - \pi(t) \quad (2)$$

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 $v \sim N(0, \sigma)$, as follows:

$$ED(t) = |v| \sum_i D_i(t) \quad (3)$$

where, following [29], $\sigma^2 = 0.05$. 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. To ensure that this decision does not affect the model's robustness, we run extensive simulations where the equation for the excess demand is $ED(t) = \beta \sum_i D_i(t) + v$. We found that, for $0 < \beta < 0.25$, the two models return qualitatively similar results. 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:

$$\pi(t+1) = \begin{cases} \min\{1, \pi(t) + ED(t)\} & ED(t) \geq 0 \\ \max\{0, \pi(t) + ED(t)\} & ED(t) < 0 \end{cases} \quad (4)$$

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

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 $0 \leq \mu \leq 0.5$ and $0 \leq \varepsilon \leq 1$, which represent the whole possible space for the Deffuant model, with a precision of 0.02 units. Our results show that this model provides a particularly accurate and robust description of prediction markets, because even under the worst-performing conditions, the synthetic time series produced by our simulations capture (at least to some degree) the emerging properties of prediction markets, such as volatility clustering and absence of autocorrelation of returns [30].

We tested the model on three different network topologies: a random network, a scale-free network, and a complete graph. We show results obtained on a complete graph (i.e. any agent is free to exchange opinions with any other agent), since we find that these replicate historical data more accurately. However, we do not find any significant difference in qualitative behaviour of the results, suggesting that i) price formation does not heavily depend on the network structure and that ii) our model is robust to topological changes. Results on the other two network topologies are shown in Appendix A.

Although it has been demonstrated that, at least under some circumstances, diffusion processes display finite-size effects [14,31], we decide to run experiments with 1000 agents to reflect the scarce liquidity prediction markets usually exhibit [28], with a median of approximately 300 trades per day per market. 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. The choice of using Barabasi-Albert networks for our model is motivated by the fact that social networks exhibit hubs and that network topology does not have an influence on the equilibria of the Deffuant model [32].

To specify the other parameters of the simulations, we follow [28,30], who provide a comprehensive quantitative analysis of the empirical properties of prediction markets using 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. [28,30] 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. We initialised our agents with a uniform opinion distribution with an average of $\frac{1}{2}$. This is for three reasons: First, these are the initial conditions that were first used in the paper that introduced the Deffuant model, and are the most widely used in practice. Second, we believe that a uniform distribution with an average of $\frac{1}{2}$ constitutes the most “uninformative prior”, and as such, it should not introduce a bias in our model. That is, other assumptions, and especially those with a different average opinion, could change the results. Given the impossibility to calibrate the initial opinion distribution with the existing dataset, we decided to remove what would have been another free parameter of the model. Third, 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 event *Will \mathbb{E} happen?* but also on the event *Will \mathbb{E} not happen?*).

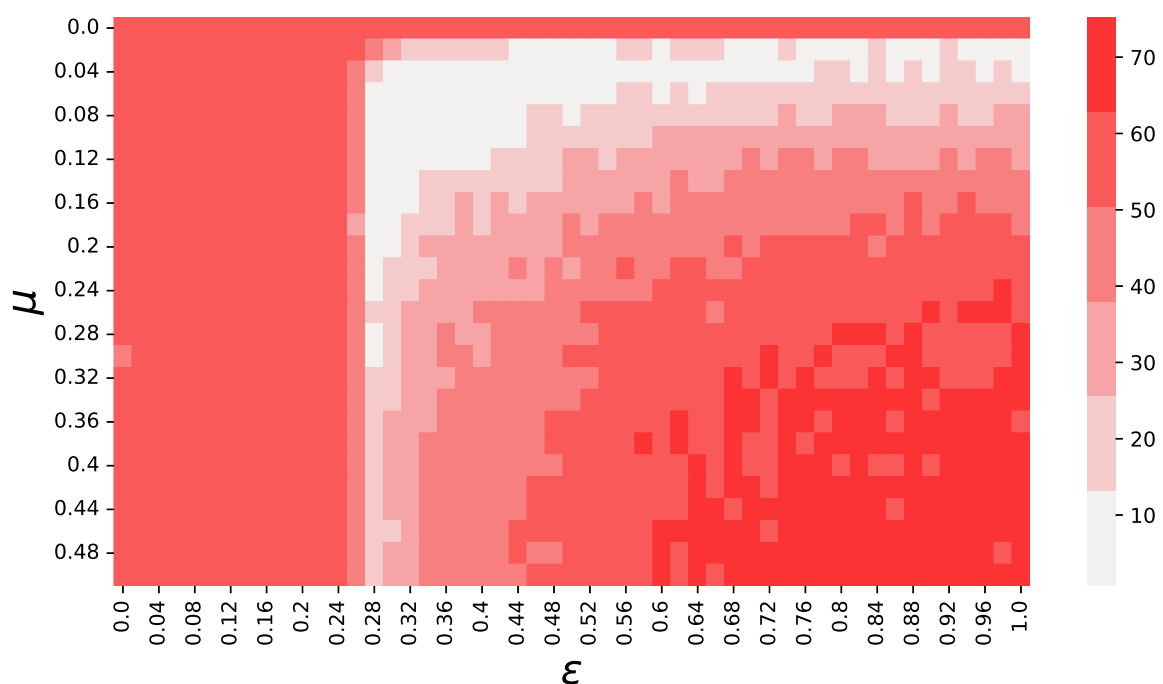


Figure 1. Objective function values (see Equation (5)) depending on μ and ϵ . These results show that there is a region, approximately delimited by the area $(\mu, \epsilon) \in [0.04, 0.16] \times [0.26, 0.5]$, where the objective function f reaches its minima. Each color used in this figure represents an interval of 12.5 for f , starting with $0.7 < f < 13.2$. We chose to discretize the colors to smooth our results over noise, and make the regions easily recognizable.

To find the optimal values for the pair μ, ϵ , we follow [33] and define the following objective function:

$$f = |k_{sim} - k_{emp}| + \lambda |\alpha_{sim} - \alpha_{emp}| \quad (5)$$

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. 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 in (5) can significantly differ in magnitude, to ensure that no component in the objective function outweighs the other, Gilli and Winker suggest that the second term, in our case $|\alpha_{emp} - \alpha_{sim}|$, is multiplied by a constant $\lambda = \frac{k_{emp}}{\alpha_{emp}}$ that rescales its magnitude. For our data set, Restocchi et al. [30] 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 Figure 2.

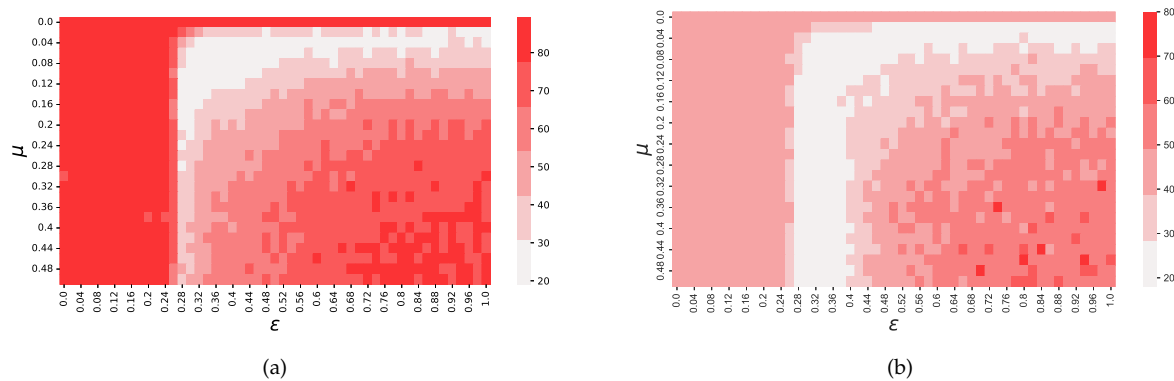


Figure 2. Heatmaps for the objective functions $f_2 = |k_{sim} - k_{emp}| + \lambda|\alpha_{sim} - \alpha_{emp}| + \lambda^a|a_{sim} - a_{emp}|$ and $f_3 = |k_{sim} - k_{emp}| + \lambda^a|a_{sim} - a_{emp}|$, respectively. These figures show that including a in the objective function f does not add information (Figure 2a), whereas removing α reduces the granularity of f (Figure 2b).

The objective function values computed from our simulations for each pair μ, ε are displayed by the heatmap in Figure 1. It is clear to see from this figure that there exists one region, approximately within the interval $(\mu, \varepsilon) \in [0.04, 0.16] \times [0.26, 0.5]$ where f is significantly lower than in the rest of the space. This region also includes the global minimum, which is found in $\mu = 0.1$ and $\varepsilon = 0.32$.

It is also interesting to note that f displays a regular behavior when varying μ and ε . To better visualize this, in Figure 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 see two regularities. First, from the right-hand side plot one can note that, in the range $0.26 \lesssim \varepsilon \lesssim 0.3$, 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 approximately in the range $0.04 \lesssim \mu \lesssim 0.1$, but, depending on ε , this minimum is more or less pronounced, and it almost disappears for $\varepsilon \lesssim 0.26$.

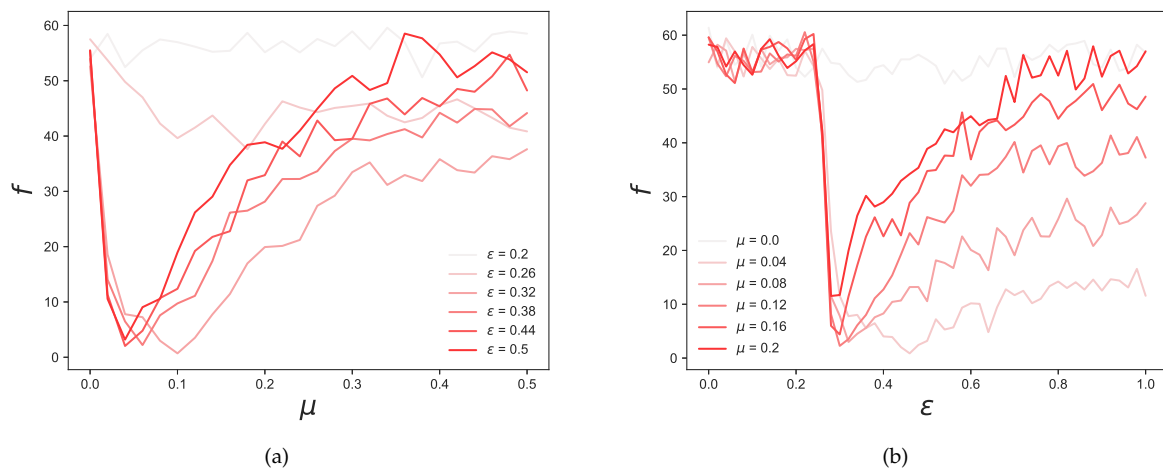


Figure 3. Detailed view of the objective function f values, depending on μ (ϵ), for five values of ϵ (μ) in the neighborhood of the optimal values. The objective function shows strong dependence on both μ and ϵ , but these figures suggest that ϵ has a greater impact on f than μ . Importantly, we can see that for all values of μ except from $\mu = 0$, there is a phase transition in the value of f in the region around $\epsilon = 0.25$, which is the threshold after which the Deffuant model starts converging to a single opinion cluster.

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 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 $\mu \gtrsim 0.25 \cup \epsilon \gtrsim 0.5$. However, our results suggest that ϵ has a greater impact on f than μ . Specifically, we observe that there is one region, delimited by $\epsilon \lesssim 0.26$, for which the objective function value becomes particularly high. Interestingly, this is the same value of ϵ below which consensus on a single opinion is not reached in the Deffuant model, and two or more opinions coexist at equilibrium [11], suggesting a relation between the accuracy of our model and whether there coexist one or more opinion clusters in the underlying opinion dynamics model. These results suggest that the quantitative accuracy at reproducing k_{sim} and α_{sim} heavily depends on whether there is a single opinion cluster existing at equilibrium, and on the time it takes for opinions to converge. Further evidence is represented by the results shown in Figure 4, in which we show the dependence of k and α (Figure 4b) on μ and ϵ .

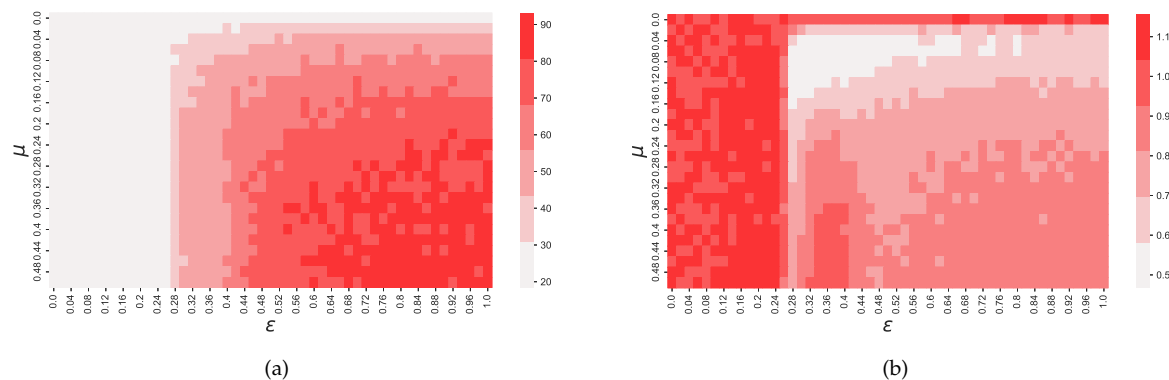


Figure 4. Values of k (panel (a)) and α (panel (b)) depending on μ and ϵ . These figures show that the kurtosis of the return distribution depends on both μ and ϵ equally, but that the value of α is heavily affected by ϵ , suggesting a link with the number of coexisting opinions at equilibrium.

The results shown in Figure 4a 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 $\varepsilon < 0.26$, 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, Figure 4b shows that α seems to depend mostly on ε , but exhibits a far sharper phase transition around $\varepsilon = 0.5$ than k , significantly decreasing its value for $\varepsilon > 0.5$. Also, Figure 4b shows that, similarly to k , in the region $\varepsilon > 0.5$, α depends mainly on μ , and its value becomes the larger the faster opinions converge.

These results are further evidence that the accuracy of our model depends on whether there is a single opinion cluster or more, except for the region approximately delimited by $(\mu, \varepsilon) \in [0.25, 0.5] \times [0.5, 1]$, where consensus is reached early in the market and little or no trading exists after a certain point. Finally, Figure 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. However, it is important to note that combinations of parameters far from the optimum still show the emergence of the discussed stylized facts, but these are quantitatively far off from those observed in historical data.

Figure 5 shows two comparisons between historical data and simulation results, obtained with both the best and worst configurations, found for the pairs with $\mu = 0.1, \varepsilon = 0.32$ and $\mu = 0.48, \varepsilon = 0.82$, respectively. The comparisons are based on the autocorrelation of absolute return and the probability density function of absolute returns, which are commonly used metrics to describe time series in financial markets [30,34]. From this figure we observe that the both the best and worst configurations generate distributions of returns which are similar to the historical one, only with slightly heavier tails. Results displayed in Figure 5 suggest that, whereas time series generated by the best configuration perfectly match the decay slope 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 exhibits a long tail, its values quickly drop when the lag considered is greater than 10 days. This is because, for the pair of values $\mu = 0.48, \varepsilon = 0.82$, 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 very small or zero, in contrast to the beginning of the market, when price changes are larger due to the higher heterogeneity of opinions.

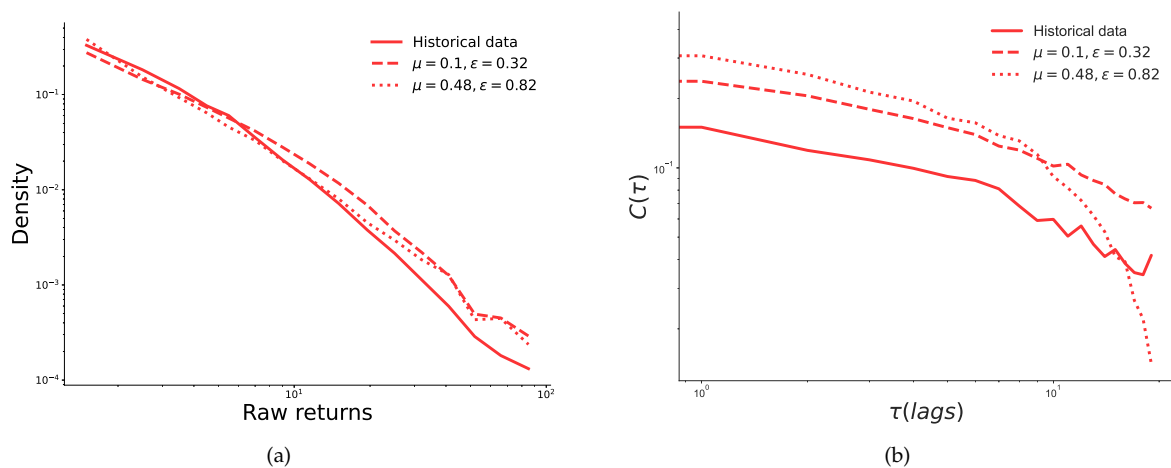


Figure 5. Comparison between historical data and simulation data generated by the best and worst pairs of μ, ϵ . Figure 5a shows the comparison between the probability density distributions of the distributions of raw returns. We note that the distribution of returns spans two decades. This is the maximum range for prediction markets in our data, since the minimum price increment is 0.01 [30], and prices are between 0 and 1. Also, this limits the negative impact of the time series being non-stationary, since large price changes are not allowed by default. Figure 5b shows the decay of the autocorrelation functions. These figures show that the both the best and worst fit of our model generate realistic returns, and that the best fit of our model also generates an autocorrelation of returns function that decays with the same exponent as the historical data $\alpha = 0.54$.

4. Conclusions

In this paper we propose a model that is able to capture salient features of price formation in prediction markets. To achieve this, we propose 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 of 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 suitable to represent prediction markets. At the same time, using historical data, our findings corroborate the validity of the Deffuant model as a representation of real-world phenomena such as opinion formation with continuous opinions. Additionally, our results show that prediction markets are best reproduced by our model when opinions are scattered around a single opinion cluster. Conversely, when two or more opinions exist at equilibrium ($\epsilon < 0.26$) or all agents have exactly the same opinion (large values of μ , for which consensus is reached early), the model yields quantitatively worse results. This suggests that, most of the times, single participants in prediction markets do not reach perfect consensus on the outcome, but still get close to it. This supports empirical evidence of mispricing in prediction markets, which tends to decrease over time and is higher when market duration is shorter [35]. Our results suggest that this might be a consequence of market participants not having enough time to reach consensus, so that their opinions are scattered around a cluster.

However, our model has some limitations that should be addressed further in future work: most importantly, our model does not include shocks. The availability of new information during a financial market can have a significant impact on prices, and this is true for prediction markets too. The analysis of the impact of shocks on the underlying opinion dynamics process and their impact on price would be of great interest to understand how prediction markets can be made more stable, therefore improving

their predictive power and reducing market manipulation. Finally, some of the model's assumptions were made to simplify this initial exploration of the impact of opinion dynamics on price formation in prediction markets. Although all our assumptions were consistent with the literature, we believe that a thorough stress test could improve the understanding of price formation in prediction markets and perhaps lead to the creation of ad hoc opinion dynamics models for prediction markets.

Appendix A

In this section, we present results obtained by using different network topologies for the opinion dynamics process. Specifically, we have tested our model on a random network (ER), built by following the Erdős-Rényi algorithm with $p = 0.04$, and on a scale-free network (BA), built by following the Barabasi-Albert preferential attachment algorithm [36], with each node forming 4 links when added to the network.

Here we show that our model, when agents are connected through a network, still displays the same qualitative behaviour. However, the main difference between these and the model on a complete graph, is that the latter gives a better minimum of the fitness function.

First, Figure A1 shows the objective function values for the model using a BA and a ER network. The regions in which the model best fits historical data remain approximately the same as with the model on the complete graph we presented in the main text. However, results on networks are worse than for the case of a complete graph case, as is evident from the higher minimum value and worse average fitness. The distributions of returns displayed in Figure A2 suggest that the model on a BA network yields realistic results for both the best and worst fit of μ and ε , whereas on the ER network the worst performing pair of parameters does not seem to follow the historical data accurately (see Figure 2b). Finally, the autocorrelation functions seem to decay with a similar exponent to that of the historical data for both networks. However, this is not the case for the worst-fit combinations, which follow a different path and, after a lag of 10 days, decay much more rapidly.

The presented results for small connectivity suggest that sparse networks are less suitable to reproduce price formation in prediction markets than a complete graph.

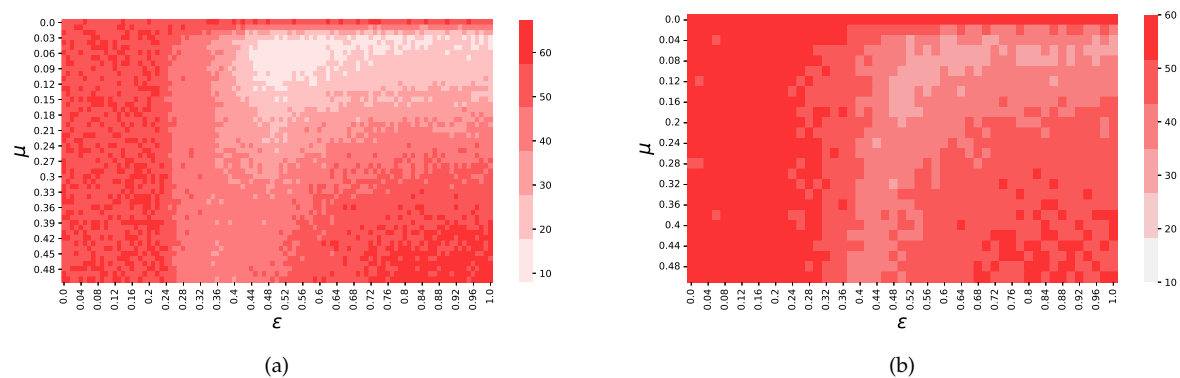


Figure A1. Heatmaps for the objective function $f = |k_{sim} - k_{emp}| + \lambda |\alpha_{sim} - \alpha_{emp}|$. The left-hand side figure has been obtained using a BA network, whereas the right-hand side figure has been obtained using a ER network.

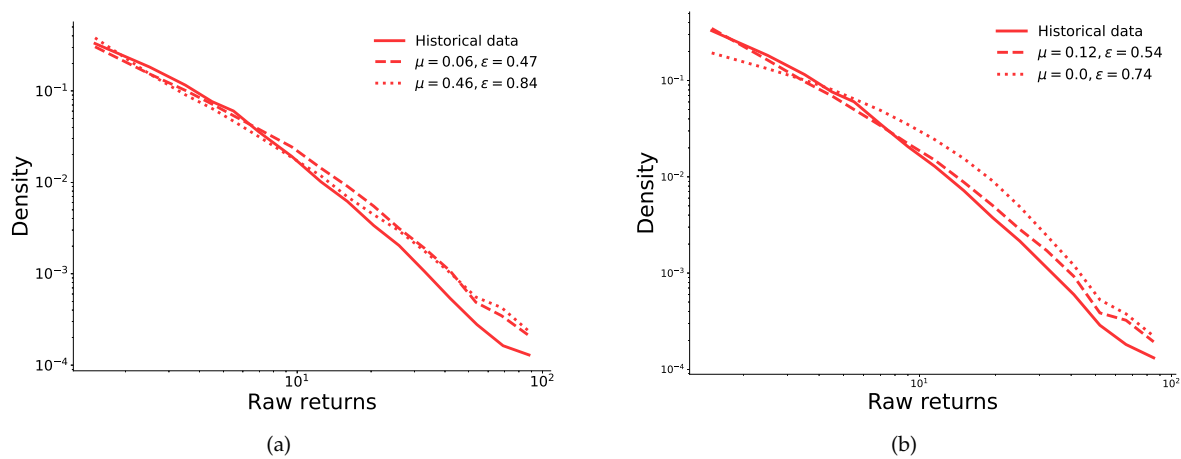


Figure A2. Comparisons of the distribution of returns for simulations run using a BA network (left-hand side), and a ER network (right-hand side). For each network, the figures show both the best and worst fit pair of values μ and ϵ .

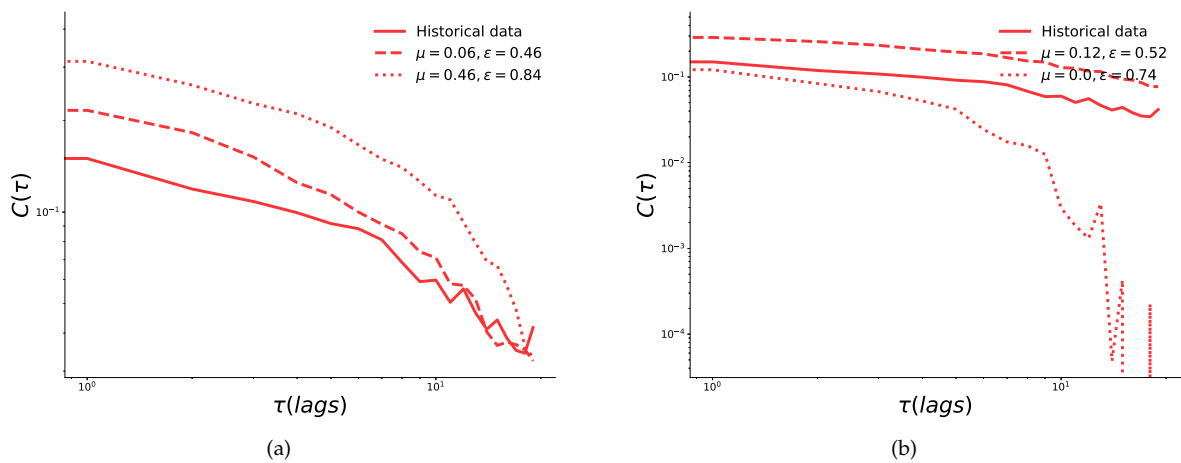


Figure A3. Comparisons the behavior of the autocorrelation function of absolute returns, on a BA network (left-hand side), and on a ER network (right-hand side).

Appendix B

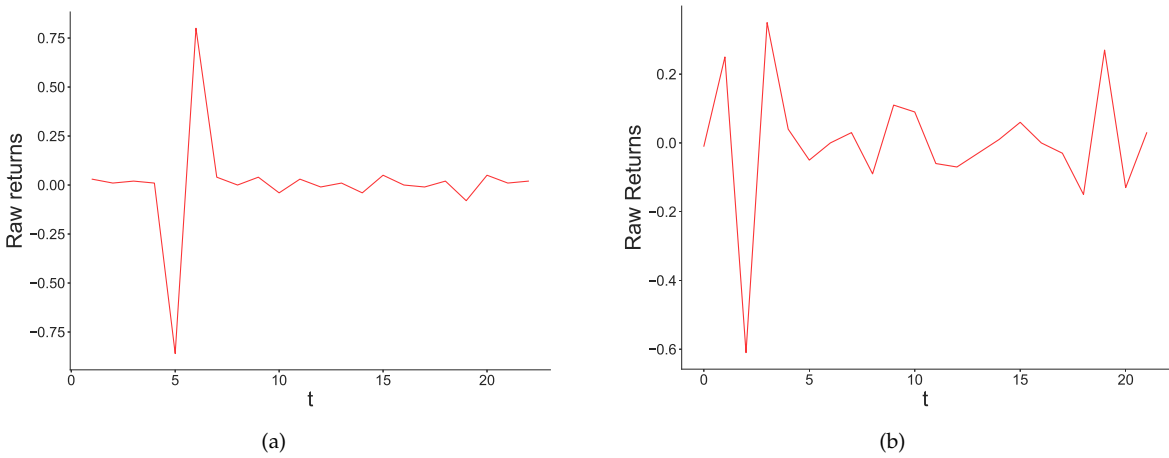


Figure A4. Historical time series of raw returns (left-hand side panes), and time series of raw returns generated by our model with $\mu = 0.1$ and $\varepsilon = 0.32$ (right-hand side panes). These represent a market duration of 23 days.

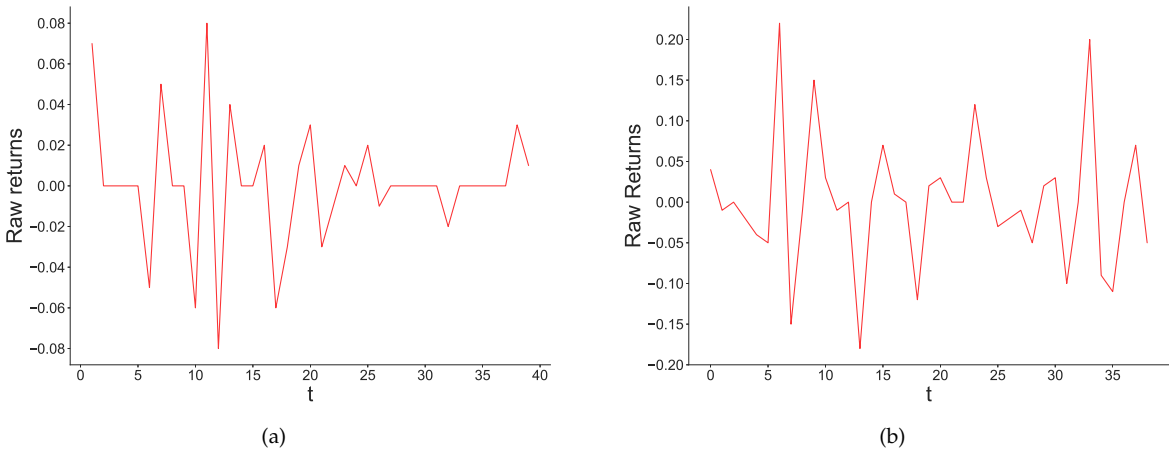


Figure A5. Historical time series of raw returns (left-hand side panes), and time series of raw returns generated by our model with $\mu = 0.1$ and $\varepsilon = 0.32$ (right-hand side panes). These represent a market duration of 40 days.

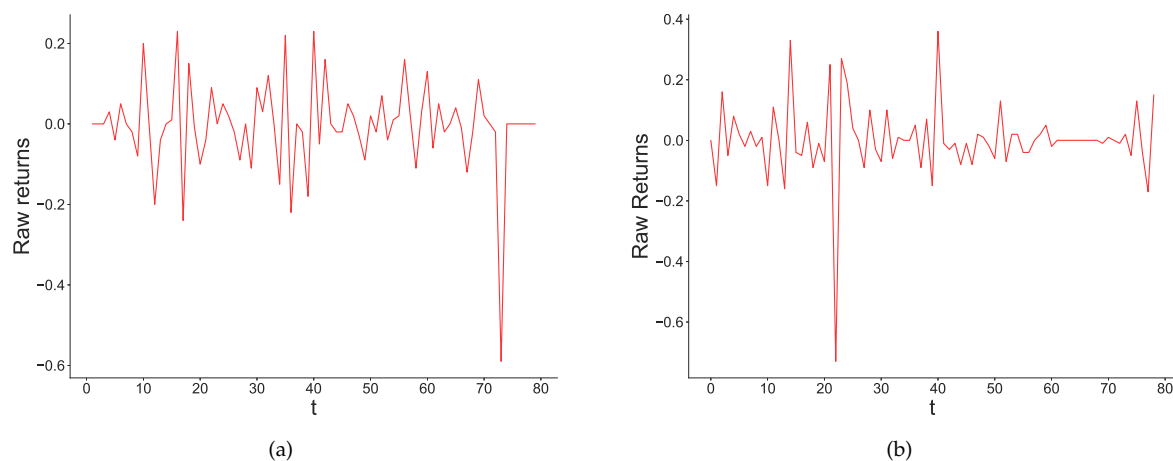


Figure A6. Historical time series of raw returns (left-hand side panes), and time series of raw returns generated by our model with $\mu = 0.1$ and $\varepsilon = 0.32$ (right-hand side panes). These represent a market duration of 80 days.

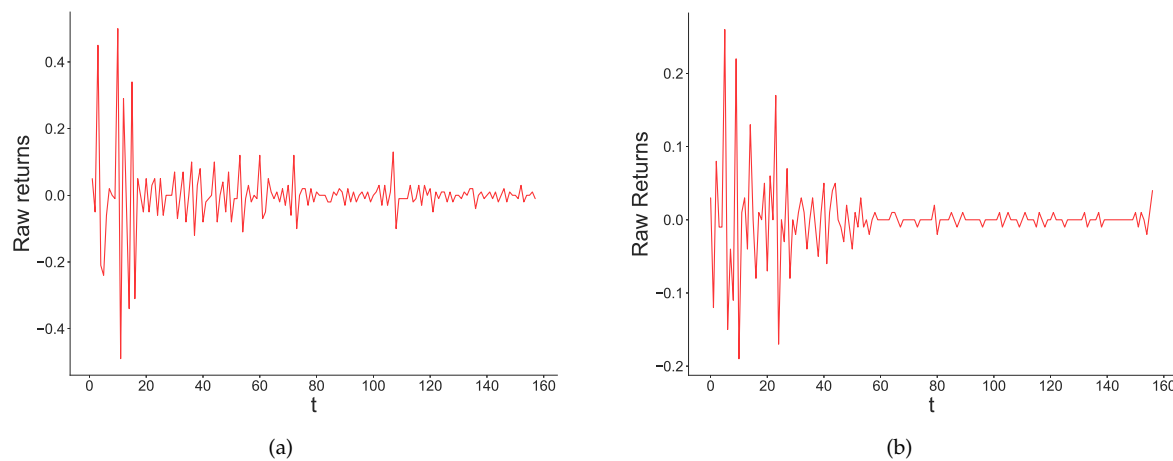


Figure A7. Historical time series of raw returns (left-hand side panes), and time series of raw returns generated by our model with $\mu = 0.1$ and $\varepsilon = 0.32$ (right-hand side panes). These represent a market duration of 158 days.

In this section, we show some historical trajectories of price changes together with similar trajectories generated by our model. Consistent with the financial literature, and especially with literature on agent-based models for financial markets, we focus on the time series of returns. To select time series, we employed the following procedure: First, to give a good illustration of different timespans, we selected four market durations, namely $T = \{20, 40, 80, 160\}$. Second, we randomly picked a historical market whose duration is within 5 days from the selected durations. In doing so, we selected the following time series:

- Will Mary Landrieu be defeated in the Louisiana Senate election? (duration $T = 23$ days, Figure A4b)
- Will Rick Perry win the 2015 Iowa Straw Poll? (duration $T = 40$ days, Figure A5a)
- Will Ed Miliband be Prime Minister after the next British election? (duration $T = 80$ days, Figure A6a)
- Will the Republican party win Barbara Boxer's Senate seat in California in 2016? (duration $T = 158$ days, Figure A7a)

Third, for each duration T , we generated time series from our model. We did so 10 times for each T , and chose the time series that minimized the mse with the historical time series of same duration.

Results are presented in Figures A4–A7. From this figure, we see that our model generates realistic time series that exhibit similar trajectories to those observed in historical markets. Figure A4a (historical) and Figure A4b (simulated) show that, for markets with short durations, price changes tend to be stable, with occasional bursts of much higher magnitude than the rest of the time series. This can be explained by the low volumes short-term markets attract (short-term markets tend to be about less important events), which can give rise to such high volatility. Occasional bursts are also observed in markets with duration $T = 80$ days, shown by Figure A6a (historical) and Figure A6b (simulated). The main difference of these markets with the ones with duration $T = 23$ is that they display a higher volatility throughout. In contrast, longer-term markets, such as the ones displayed in Figure A7a (historical) and Figure A7b (simulated), display a high volatility at the beginning, where volumes are low, to then stabilize later on in the market, thanks to the participation of more traders.

Author Contributions: Conceptualization, V.R., F.M. and E.G.; methodology, V.R., E.G.; software, V.R.; validation, All authors; formal analysis, V.R.; data curation, V.R.; writing—original draft preparation, V.R.; writing—review and editing, All authors.; visualization, V.R.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding

Data Availability Statement: The data was provided by PredictIt and are available for academic research from PredictIt upon request. The code will be made publicly available on www.comses.net upon publication.

Acknowledgments: The authors would like to thank Claudio J. Tessone, Pietro Panzarasa, and Tiejun Ma for their feedback and insightful discussions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Plott, C.; Chen, K. Information aggregation mechanism: Concept, design, and implementation for a sales forecasting problem.
2. Cowgill, B.; Wolfers, J.; Zitzewitz, E. Using prediction markets to track information flows evidence from Google. In *Auctions, Market Mechanisms and Their Applications*; Das, S.; Ostrovsky, M.; Pennock, D.; Szymanski, B., Eds.; Springer: Berlin, 2009.
3. Berg, J.; Nelson, F.; Rietz, T. Prediction market accuracy in the long run. *Int. J. Forecast.* **2008**, *24*.
4. Arnesen, S.; Bergfjord, O. Prediction markets vs polls—a examination of accuracy for the 2008 and 2012 elections. *J. Predict. Mark.* **2014**, *8*, 24–33.
5. Restocchi, V.; McGroarty, F.; Gerding, E.; Johnson, J.E. It Takes All Sorts: A Heterogeneous Agent Explanation for Prediction Market Mispicing. *European Journal of Operational Research* **2018**, pp. 556–569. <https://doi.org/https://doi.org/10.1016/j.ejor.2018.04.011>.
6. Goodell, J.; McGroarty, F.; 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* **2015**, pp. 397–410.
7. Goodell, J.W.; Bodey, R.A. Price-earnings changes during US presidential election cycles: Voter uncertainty and other determinants. *Public Choice* **2012**, *150*, 633–650. <https://doi.org/10.1007/s11127-010-9720-8>.
8. Shin, H.S. Measuring the Incidence of Insider Trading in a Market for State-Contingent Claims. *Econ. J.* **1993**, *103*, 1141–1153.
9. Ottaviani, M.; Sørensen, P.N. The Favorite-Longshot Bias : An Overview of the Main Explanations. In *Handbook of Sports and Lottery markets*; Hausch, D.B.; Ziemba, W.T., Eds.; North-Holland, 2008; chapter 6, pp. 83–101.
10. Snowberg, E.; Wolfers, J. Explaining the Favorite-Long Shot Bias: Is it Risk-Love or Misperceptions? *J. Political Econ.* **2010**, *118*, 723–746. <https://doi.org/10.1086/655844>.
11. Deffuant, G.; Neau, D.; Amblard, F.; Weisbuch, G. Mixing beliefs among interacting agents. *Adv. Complex Syst.* **2000**, *87*, 11.
12. Sznajd-Weron, K.; Sznajd, J. Opinion evolution in closed community. *Int. J. Mod. Phys. C* **2000**, *11*, 1157, [arXiv:cond-mat/0101130]. <https://doi.org/10.1142/S0129183100000936>.
13. Clifford, P.; Sudbury, A. Minority Opinion Spreading in Random Geometry. *Eur. Phys. J. B* **2002**, *25*, 403–406.

14. Tessone, C.J.; Toral, R.; Amengual, P.; Wio, H.S.; Miguel, M.S. Neighborhood models of minority opinion spreading. *Eur. Phys. J. B* **2004**, *39*, 535–544, [arXiv:cond-mat/0403339]. <https://doi.org/10.1140/epjb/e2004-00227-5>.
15. Baronchelli, A.; Dall'Asta, L.; Barrat, A.; Loreto, V. Topology Induced Coarsening in Language Games. *Phys. Rev. E* **2006**, *73*.
16. Castellano, C.; Fortunato, S.; Loreto, V. Statistical physics of social dynamics. *Rev. Mod. Phys.* **2009**, *81*, 591–646.
17. Galam, S. The invisible hand and the rational agent are behind bubbles and crashes. *Chaos, Solitons Fractals* **2016**, *88*, 209–217. Complexity in Quantitative Finance and Economics, <https://doi.org/https://doi.org/10.1016/j.chaos.2016.03.011>.
18. Vilela, A.L.; Wang, C.; Nelson, K.P.; Stanley, H.E. Majority-vote model for financial markets. *Phys. A: Stat. Mech. Its Appl.* **2019**, *515*, 762–770. <https://doi.org/https://doi.org/10.1016/j.physa.2018.10.007>.
19. Kononovicius, A.; Gontis, V. Agent based reasoning for the non-linear stochastic models of long-range memory. *Phys. A: Stat. Mech. Its Appl.* **2012**, *391*, 1309–1314. <https://doi.org/https://doi.org/10.1016/j.physa.2011.08.061>.
20. Alfarano, S.; Lux, T.; Wagner, F. Estimation of Agent-Based Models: The Case of an Asymmetric Herding Model. *Comput. Econ.* **2005**, *26*, 19–49. <https://doi.org/https://doi.org/10.1007/s10614-005-6415-1>.
21. Lux, T.; Marchesi, M. Scaling and criticality in a stochastic multi-agent model of a Financial market. *Nature* **1999**, *397*.
22. Zhou, W.X.; Sornette, D. Self-fulfilling Ising Model of Financial Markets **2005**. *181*, 175–181, [arXiv:physics/0503230]. <https://doi.org/10.1140/epjb/e2006-00391-6>.
23. Sobkowicz, P. Modelling opinion formation with physics tools: Call for closer link with reality. *J. Artif. Soc. Soc. Simulations* **2009**, *12*.
24. Del Vicario, M.; Bessi, A.; Zollo, F.; Petroni, F.; Scala, A.; Caldarelli, G.; Stanley, H.E.; Quattrociocchi, W. The spreading of misinformation online. *Proc. Natl. Acad. Sci.* **2016**, *113*, 554–559, [http://www.pnas.org/content/113/3/554.full.pdf]. <https://doi.org/10.1073/pnas.1517441113>.
25. Del Vicario, M.; Zollo, F.; Caldarelli, G.; Scala, A.; Quattrociocchi, W. Mapping social dynamics on Facebook: The Brexit debate. *Soc. Networks* **2017**, *50*, 6–16. <https://doi.org/https://doi.org/10.1016/j.socnet.2017.02.002>.
26. Weisbuch, G.; Deffuant, G.; Amblard, F.; Nadal, J.P. Interacting Agents and Continuous Opinions Dynamics. In *Lecture Notes in Economics and Mathematical Systems - No. 521*; Fandel, G.; Trockel, W., Eds.; Springer Verlag, Berlin, 2002; pp. 225–242.
27. Lorenz, J. *Managing Complexity: Insights, Concepts, Applications*; Springer: Berlin, 2008.
28. Restocchi, V.; McGroarty, F.; Gerding, E. Statistical properties of volume and calendar effects in prediction markets. *Physica A* **2018**.
29. Lux, T.; Marchesi, M. Volatility Clustering in Financial Markets: A Microsimulation of Interacting Agents. *Int. J. Theor. Appl. Financ.* **2000**, *03*, 675–702. <https://doi.org/10.1142/S0219024900000826>.
30. Restocchi, V.; McGroarty, F.; Gerding, E. The stylized facts of prediction markets: Analysis of price changes. *Physica A* **2018**.
31. Toral, R.; Tessone, C.J. Finite size effects in the dynamics of opinion formation. *Commun. Comput. Phys.* **2006**, *2*, 177–195, [arXiv:physics/0607252].
32. Fortunato, S. Universality of the Threshold for Complete Consensus for the Opinion Dynamics of Deffuant et al. *Int. J. Mod. Phys. C* **2004**, *15*, 1301–1307.
33. Gilli, M.; Winker, P. A global optimization heuristic for estimating agent based models. *Comput. Stat. Data Anal.* **2003**, *43*, 299–312.
34. Chakraborti, A.; Toke, I. Econophysics review: II. Agent-based models. *Quant. ...* **2011**, *1041*, 1013–1041.

35. Restocchi, V.; McGroarty, F.; Gerding. The temporal evolution of mispricing in prediction markets. *Financ. Res. Lett.* **2019**, *29*, 303–307. <https://doi.org/https://doi.org/10.1016/j.frl.2018.08.003>.
36. Barabási, A.L.; Albert, R. Emergence of scaling in random networks. *Science* **1999**, *5439*, 509–512.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.