The stylized facts of prediciton markets: analysis of price changes

Valerio Restocchi[a,b], Frank McGroarty[a], Enrico Gerding[b]

[a]Southampton Business School, University of Southampton, SO17 1BJ, Southampton, UK
[b]Electronics and Computer Science, University of Southampton, SO17 1BJ, Southampton, UK

Abstract

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 means of validation for quantitative models of prediction markets. To address this shortcoming, in this paper we compile a list of empirical regularities (stylized facts) of price changes we find by analyzing daily price changes from 3385 prediction markets on political events, a data set provided by PredictIt. We find that price changes in prediction markets show characteristics similar to emerging markets, with some small differences.

Keywords: Prediction markets; Political markets; Stylized facts; Long memory; Power-law behavior.

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

*Corresponding author
Email addresses: v.restocchi@soton.ac.uk (Valerio Restocchi), f.j.mcgroarty@soton.ac.uk (Frank McGroarty), eg@ecs.soton.ac.uk (Enrico Gerding)

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 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 2 we present our data set and outline the methods we employed to conduct analysis on it. In Section 3 we show in detail why raw returns, and not log returns, should be used to examine price time series in prediction markets. Section 4 and 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 6.

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

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

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. 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.

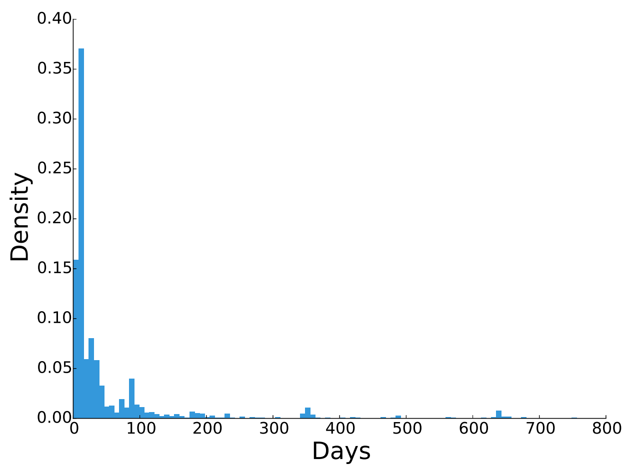


Figure 1: Distribution of market durations in days.

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

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

Similarly, the log return at the same time is:

$$r_t^l = ln(P_t) - ln(P_{t-1}) \tag{2}$$

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

$$r_t^{\%} = \frac{P_t - P_{t-1}}{P_t} \tag{3}$$

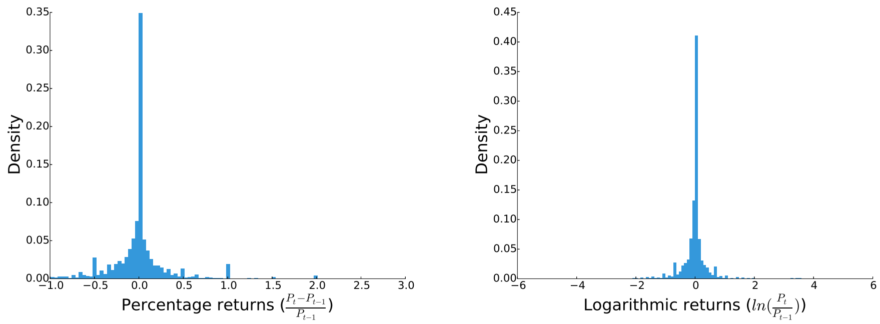
3.1. Why raw returns

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:

$$r_{0,T} = r_{0,t} + r_{t,T} \tag{4}$$

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



(a) (b)

Figure 2: Figure (a) displays the distribution of relative returns, for $-1 \leq r_t^{\%} \leq 3$. Figure (b) displays the distribution of log returns r_t^l.

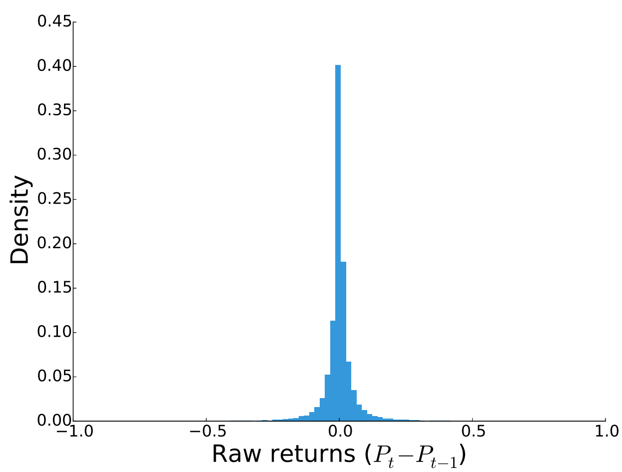


Figure 3: Distribution of raw returns.

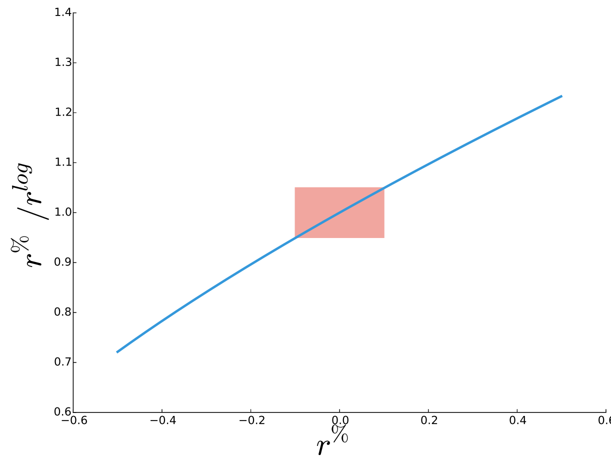


Figure 4: Function of the ratio $\frac{r^{\%}}{r^{l}}$ depending on the value of percentage returns. The shaded area represents the range in which $r_t^l = r_t^{\%} \pm 5\%$ holds, and corresponds to $|r^{\%}| \leq 0.1$.

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. 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 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 $r_t^{\%} \leq 0.5 \vee r_t^{\%} \geq 0.5$, 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 $-1 \leq r_t^{\%} \leq 99$, generating a strongly asymmetric and irregular distribution as shown in Fig. 2(a). In this case, using log returns only solves the latter problem (see Fig.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. 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. 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, 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.

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, 1999). 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, 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 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 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.

Table 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

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:

$$p(x) \propto x^{-\alpha} \tag{5}$$

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

$$\zeta(\alpha, x_{min}) = \sum_{i=0}^{\infty} \frac{1}{(i + x_{min})^{\alpha}} \tag{6}$$

whose value can only be found numerically. Third, our return distributions are characterized by another constraint. That is, since $0.01 \leq |r_t| \leq 0.99$, 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:

$$p(x) = \frac{x^{-\alpha}}{\Delta\zeta} \tag{7}$$

where

$$\Delta\zeta \equiv \zeta(\alpha, x_{min}) - \zeta(\alpha, x_{max}) \tag{8}$$

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

$$L(\alpha) = -\alpha\left(\sum_{i=0}^{N} ln(x_i)\right) - N ln\big(\Delta\zeta\big) \tag{9}$$

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

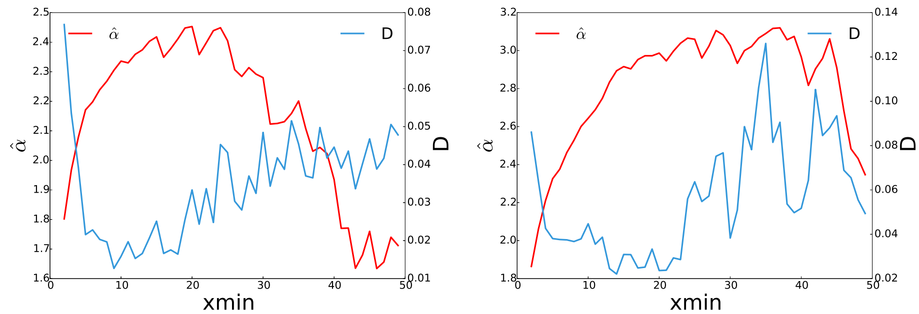
$$\hat{\alpha} = \underset{\alpha}{\operatorname{argmax}}[L(\alpha)] \tag{10}$$

Since, in this case, there exists no closed-form solution for $\hat{\alpha}$, we find the value that maximizes Eq. (9) numerically.

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



(a) (b)

Figure 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)**.

gives us a measure of the error, and is given by

$$\Delta\hat{\alpha}^2 = \frac{1}{n}\frac{\Delta\zeta^2}{\Delta\zeta''\Delta\zeta^2 - \Delta\zeta'^2} \tag{11}$$

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).

4.3. Gain/loss asymmetry

We find two significantly different values $\hat{\alpha}$ for positive and negative returns. As we showed in Table 1, the number of observations and the value of the skewness (Table 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. 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. 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

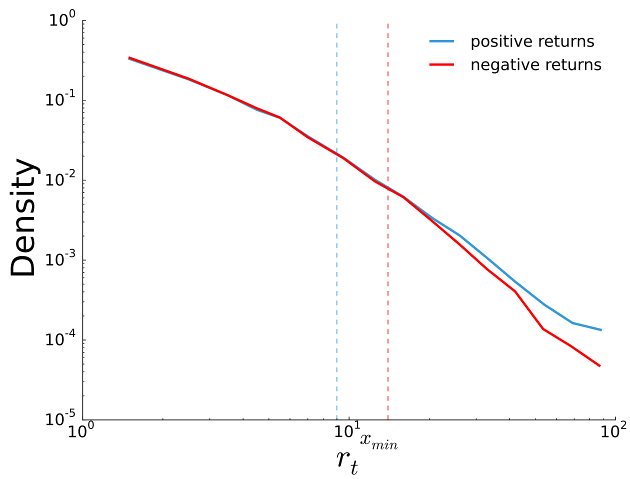


Figure 6: Comparison of the PDFs between positive and negative returns.

Table 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

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.

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

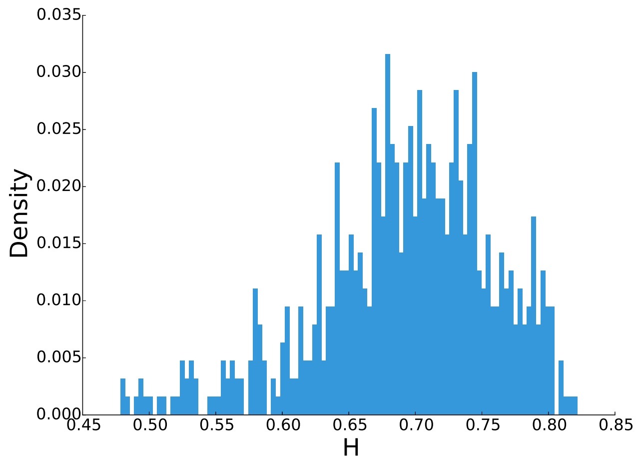


Figure 7: Distribution of H for the markets which last more than 30 days.

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. 1), which implies that it is unlikely to mistake short-range memory for long-range effect, and Lo's modified R/S statistic 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

Table 3: Summary statistics for the distribution of H.

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

range, and S, the standard deviation of the time series. R is given by:

$$R_\tau = max(z_1, z_2, ..., z_\tau) - min(z_1, z_2, ..., z_\tau) \quad (12)$$

where

$$z_\tau = \sum_{t=1}^{\tau}(r_t - \bar{r}_\tau) \quad (13)$$

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

$$S_\tau = \sqrt{\frac{1}{\tau}\sum_t (r_t - \bar{r}_\tau)^2} \quad (14)$$

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

$$(R/S)_\tau = \left(\frac{\tau}{2}\right)^H \quad (15)$$

Fig. 7 shows the distribution of H, and Table 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.

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.

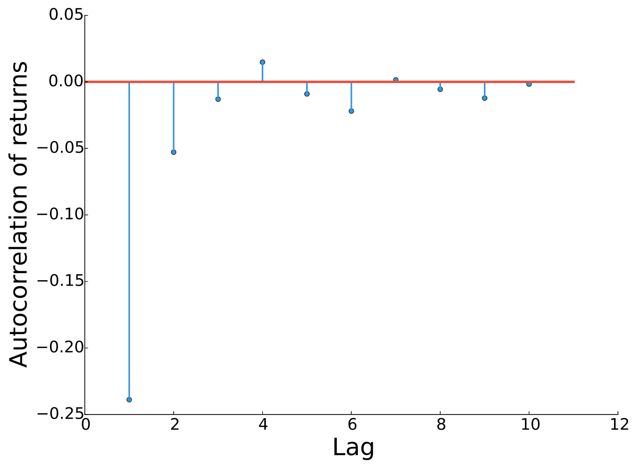


Figure 8: Autocorrelation of raw returns for lags in the range $0 > \tau \geq 10$ days.

5.1. Autocorrelation of returns

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

$$C(\tau) = corr(r(t, \Delta t), r(t + \tau, \Delta t)) \tag{16}$$

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. 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).

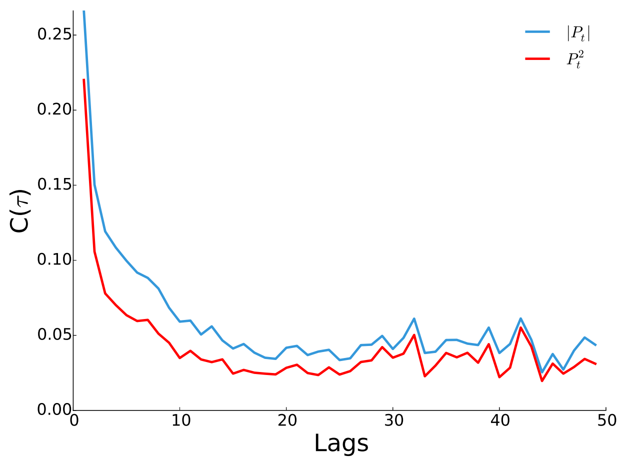


Figure 9: Autocorrelation of raw returns for lags in the range $0 > \tau \geq 50$ days.

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:

$$C(\tau)^{sq} = corr(r(t)^2, r(t+\tau)^2) \tag{17}$$

or:

$$C(\tau)^{abs} = corr(|r(t)|, |r(t+\tau)|) \tag{18}$$

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

$$C(\tau) \propto \frac{K}{\tau^{\alpha}} \tag{19}$$

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. 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

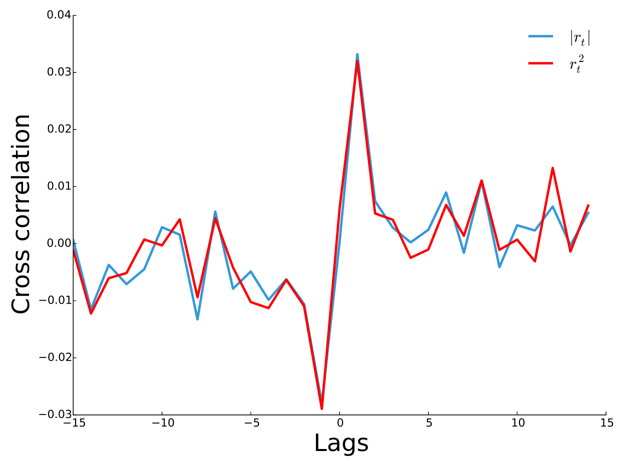


Figure 10: Correlation between two measures of volatility and past returns.

possible. However, as shown by Fig. 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 $C(\tau)^{abs}$ 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.

5.3. 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. 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.

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.

Acknowledgments

We would like to thank all participants at the 2nd Workshop of the Econophysics Network and seminar attendees at the London Institute of Mathematical Sciences for precious comments. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

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

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

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

Bariviera, A. F., Basgalla, M. J., Hasperu, W., Naiouf, M., 2017. Some stylized facts of the Bitcoin market. Physica A: Statistical Mechanics and its Applications 484, 82–90.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Mandelbrot, B., 1971. 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.

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

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

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

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

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

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

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

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

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

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

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

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