Supporting Information - Memory effects in stock price dynamics: evidences of technical trading

Federico Garzarelli[1], Matthieu Cristelli[2], Gabriele Pompa[3], Andrea Zaccaria[2], Luciano Pietronero[1,2,4]

[1] "Sapienza", Università di Roma, Dip. Fisica, P.le A. Moro 2, 00185, Roma, Italy
[2] Institute of Complex Systems, CNR, Roma
[3] IMT - Institutions Market Technology, Lucca, Italy
[4] London Institute for Mathematical Sciences, South Street 22, Mayfair London, UK

1 Statistical framework: bounce modeling

Our statistical framework is the following: we call *trial* an event in which the price enters in a stripe. The price may only exit upward or downward from the stripe[1]. We model this random event as an elementary Bernoullian event as the price may either bounce on the stripe (*success*), or break it (*failure*).
Now consider a generic stripe, suppose that we have previously observed $b_{prev} = i$ bounces on that stripe and that another trial has just happened. We call this event

E_i : $\{b_{prev} = i$ bounces on a stripe previously happened$\}$ *and* $\{$one more Bernoulli *trial* happened right now$\}$

and define the elementary probability of trial success in this event (i.e. a bounce b) as

$$\pi_i = p(b|b_{prev} = i) \tag{1}$$

Now we assume no prior knowledge on the parameter π_i and treat it as a random variable uniformly distributed in the interval $[0, 1]$:

$$Prior(\pi_i) = \mathcal{U}([0, 1]) \tag{2}$$

Suppose now that, out of N_i events E_i observed, we have measured n_i trial successes, or in other words, out of N_i times in which the price entered again a stripe - having previously already made i bounces on it - overall it bounced back n_i times and consequently broke the stripe $N_i - n_i$ times.
We modeled the likelihood of having observed n_i successes out of N_i events E_i (each governed by the same elementary probability π_i) as

$$L(n_i|N_i; \pi_i) = Binom(n_i|N_i; \pi_i) = \binom{N_i}{n_i}(\pi_i)^{n_i}(1 - \pi_i)^{N_i - n_i} \tag{3}$$

The goal of this framework is to infer the posterior distribution of π_i

$$Posterior(\pi_i|N_i, n_i) \tag{4}$$

i.e. which is the value of π_i that most likely governs the bernoullian trials belonging to events E_i, having observed a certain amount N_i of them and having counted out n_i successes.
Therefore, with the choice of an uniform prior for the parameter π_i we are avoiding to give priority to any particular choice of the probability that the bernoullian trial of any event E_i will eventually lead to a success. Moreover, it is possible to see the uniform distribution as a particular case of Beta distribution:

$$Beta(\pi_i|1; 1) \equiv 1 \quad \forall \pi_i \in [0, 1]$$

[1] we are then ignoring the remote possibility of an infinite permanence within the stripe.

and this allows us easily to know the form of the Posterior distribution of π_i, as a Beta *prior* is conjugated with a Beta *posterior* with respect to a binomial *likelihood* (refer to appendix A for details):

$$if: \begin{cases} Prior(\pi_i) & = & \mathcal{U}([0,1]) \\ L(n_i|N_i;\pi_i) & = & Binom(n_i|N_i;\pi_i) \end{cases} \quad then: \quad Posterior(\pi_i|N_i,n_i) = Beta(\pi_i|n_i+1;N_i-n_i+1)$$

whose expected value $\mathrm{E}[p(b|b_{prev})]$ and variance $\mathrm{Var}[p(b|b_{prev})]$ are those in Eq. (1) and (2) of main text.

2 Statistical validation of the measured memory effect

In this section we provide statistical evidences of the difference between the memory effect measured in the stock series and the lack of memory shown by the reshuffled returns series.
Let us first introduce the same quantities as before, N_i^{res} and n_i^{res}, but calculated for the reshuffled series. We consider that if the randomly reshuffled series were to be significantly different from the stock series, their empirical bounce frequency

$$f_i^{res} = f^{res}(b|b_{prev}=i) = \frac{n_i^{res}}{N_i^{res}}$$

which is the empirical proxy for the probability

$$\pi_i^{res} = p(b|b_{prev}=i)$$

should be very unlikely obtained as an extraction from the posterior distribution of the π_i of the stock series.
We formalize this intuition setting a Kolmogorov-Smirnov (K-S) test, following these steps:

- we generate an ensemble $M = 100$ reshuffled return series[2], from which we obtained a sample of empirical bounce frequencies for any kind of bounce $i = 1,2,3,4$: $\left(f_{i,1}^{res}, f_{i,2}^{res}, ..., f_{i,M}^{res}\right)$.
- for any kind of bounce $i = 1,2,3,4$, from the empirical values of N_i and n_i computed for the stock series, we evaluate the theoretical posterior distribution $Posterior(\pi_i|N_i,n_i)$ of the bounce probability π_i from equations (6) and (7) in Appendix A.
- we perform a K-S test to compare the sample distribution of $\left(f_{i,1}^{res}, f_{i,2}^{res}, ..., f_{i,M}^{res}\right)$ with the reference distribution $Posterior(\pi_i|N_i,n_i)$.

The K-S test returns the p-value as well a D-value which consists in the absolute value of the maximum (supremum) between the sample CDF and the theoretical one. The closer this number is to 1 the less likely it is that the sample was extracted from the theoretical distribution.
Results are shown in Table 1 for Resistances and in Table 2 for Supports. As most of the p-values are even under the numerical precision, it is widely confirmed that for what concerns the measured memory effect the randomly reshuffled series are significantly different from the stock series.

[2] we set $M = 100$.

Resistances		1^{st} *bounce*	2^{nd} *bounce*	3^{rd} *bounce*	4^{th} *bounce*
$T = 1$	p-value	0.0	0.0	0.0	0.0
	D-value	1.0	1.0	1.0	1.0
$T = 15$	p-value	0.0	0.0	0.0	0.0
	D-value	1.0	1.0	1.0	1.0
$T = 30$	p-value	0.0	0.0	0.0	0.0
	D-value	1.0	1.0	> 0.999	> 0.999
$T = 45$	p-value	0.0	0.0	0.0	0.0
	D-value	> 0.999	> 0.999	> 0.999	> 0.999
$T = 60$	p-value	0.0	0.0	0.0	0.0
	D-value	0.996	> 0.999	0.994	0.996
$T = 90$	p-value	0.0	0.0	0.0	0.0
	D-value	0.923	> 0.999	0.990	0.793
$T = 120$	p-value	0.0	0.0	0.0	0.0
	D-value	0.970	0.976	0.808	0.871
$T = 150$	p-value	0.0	0.0	0.0	0.0
	D-value	0.851	0.837	0.681	0.876
$T = 180$	p-value	0.0	0.0	0.0	0.0
	D-value	0.993	0.909	0.914	0.866

Table 1: *Resistances*: Kolmogorov-Smirnov results (p-value and D-value) for comparison between CDF of sample frequency of reshuffled returns series and theoretical posterior of bounce probability π_i of stock series.

Supports		1^{st} *bounce*	2^{nd} *bounce*	3^{rd} *bounce*	4^{th} *bounce*
$T = 1$	p-value	0.0	0.0	0.0	0.0
	D-value	1.0	1.0	1.0	1.0
$T = 15$	p-value	0.0	0.0	0.0	0.0
	D-value	1.0	1.0	1.0	1.0
$T = 30$	p-value	0.0	0.0	0.0	0.0
	D-value	1.0	1.0	> 0.999	> 0.999
$T = 45$	p-value	0.0	0.0	0.0	0.0
	D-value	1.0	1.0	> 0.999	> 0.999
$T = 60$	p-value	0.0	0.0	0.0	0.0
	D-value	> 0.999	> 0.999	> 0.999	> 0.999
$T = 90$	p-value	0.0	0.0	0.0	0.0
	D-value	> 0.999	0.973	0.983	0.912
$T = 120$	p-value	0.0	0.0	0.0	4.32×10^{-13}
	D-value	0.999	0.997	0.725	0.374
$T = 150$	p-value	0.0	0.0	0.0	1.45×10^{-6}
	D-value	0.996	0.729	0.503	0.262
$T = 180$	p-value	0.0	0.0	2.24×10^{-13}	4.35×10^{-14}
	D-value	0.996	0.909	0.378	0.388

Table 2: *Supports*: Kolmogorov-Smirnov results (p-value and D-value) for comparison between CDF of sample frequency of reshuffled returns series and theoretical posterior of bounce probability π_i of stock series.

Appendix A

In this appendix we will show the intercurrent conjugation between a Beta *prior* and a Beta *posterior* with respect to a binomial *likelihood*, in the case considered here of a uniform prior.
We begin recalling the definition of a distribution of the random variable $\pi \in [0,1]$ belonging to a Beta family described by parameters $a, b \in [0,1]$:

$$Beta(\pi|a;b) = \frac{\pi^{a-1}(1-\pi)^{b-1}}{\mathcal{B}(a,b)}$$

where the Beta function $\mathcal{B}(a,b)$ is defined as the integral

$$\mathcal{B}(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \tag{5}$$

We first notice that if $a = 1$ and $b = 1$, then the Beta distribution is uniform on the $[0,1]$ interval:

$$Beta(\pi|a=1;b=1) = \frac{1}{\mathcal{B}(1,1)} \equiv 1.$$

In this particular case we now show the conjugation between the uniform prior of π and its posterior distribution having observed N trials and, out of them, n successes, binomially distributed. This consists to show that if:

$$\begin{cases} Prior(\pi) & = & \mathcal{U}([0,1]) \\ L(n|N;\pi) & = & Binom(n|N;\pi) \end{cases}$$

then:

$$Posterior(\pi|N,n) = Beta(\pi|n+1;N-n+1)$$

Indeed, by the Bayes theorem:

$$\begin{aligned} Posterior(\pi|N,n) &= \frac{L(n|N;\pi)Prior(\pi)}{\int L(n|N;\pi)Prior(\pi)d\pi} \\ &= \frac{L(n|N;\pi)\mathcal{U}([0,1])}{\int L(n|N;\pi)\mathcal{U}([0,1])d\pi} \\ &= \frac{\binom{N}{n}\pi^n(1-\pi)^{N-n}}{\int_0^1 \binom{N}{n}\pi^n(1-\pi)^{N-n}d\pi} \\ &= \frac{\pi^n(1-\pi)^{N-n}}{\mathcal{B}(n+1,N-n+1)} \\ &= Beta(\pi|n+1;N-n+1) \end{aligned} \tag{6}$$

and we conclude noticing that the shape of the posterior is:

$$Beta(\pi|n+1;N-n+1) = (N+1)\binom{N}{n}\pi^n(1-\pi)^{N-n} \tag{7}$$

which is easily derived from equation (5) and the properties of the Γ function and that has been adopted in the computation of the theoretical distribution in the Kolmogorov-Smirnov test in section 2.