

# Statistical Collapse of Excessive Market Losses

M. DENYS<sup>a</sup>, M. JAGIELSKI<sup>a,b,c,\*</sup>, T. GUBIEC<sup>a</sup>, R. KUTNER<sup>a</sup> AND H.E. STANLEY<sup>b</sup>

<sup>a</sup>Faculty of Physics, University of Warsaw, L. Pasteura 5, PL-02-093 Warsaw, Poland

<sup>b</sup>Center for Polymer Studies and Dept. of Physics, Boston Univ., Boston, MA 02215, USA

<sup>c</sup>Department of Management, Technology and Economics, ETHZ, Scheuchzerstr. 7, CH-8092 Zürich, Switzerland

We analytically derive superstatistics (or complex statistics) that accurately model empirical market activity data (supplied by Bogachev, Ludescher, Tsallis, and Bunde) exhibiting transition thresholds. We measure the interevent times between excessive losses (that is, greater than some threshold) and use the mean interevent time as a control variable to derive a universal description of empirical data collapse. Our superstatistic value is a power-law corrected by the lower incomplete gamma function, which asymptotically tends toward robustness but initially gives an exponential. We find that the scaling shape exponent that drives our superstatistics subordinates themselves and a “superscaling” configuration emerges.

DOI: [10.12693/APhysPolA.129.913](https://doi.org/10.12693/APhysPolA.129.913)

PACS/topics: 89.65 Gh, 05.40.–a, 89.75.Da

## 1. Introduction and principal goal

Financial markets fluctuate (quite often strongly) as traders estimate risk levels and attempt to gain a profit. The interevent interval between times when market returns are producing excessive losses and times when they are producing excessive profits can be described by using waiting-time distributions (see Refs. [1–4] and references therein).

Empirical market data on excessive profits and losses [5–8] define excessive profits as those greater than some positive fixed threshold  $Q$  and excessive losses as those below some negative threshold  $-Q$ . The mean interevent time<sup>†</sup> between losses versus  $Q$  has been used in this paper as an aggregated basic variable.

Interevent times constitute a universal stochastic measurement of market activity on time-scales that range from one minute to one month [5, 6]. The mean interevent time can be used as a control variable that produces a universal description of empirical data collapse [7], i.e., the distribution of interevent times for a fixed mean interevent time is a universal statistical quantity unaffected by time-scale, type of market, asset, or index.

This distribution can be described using (i) the continuous-time random walk (CTRW) valley model (see Refs. [2, 4] and references therein), which treats time intervals as random variables, and (ii) generalized extreme value statistics<sup>‡</sup> for stochastic dependent basic processes [13].

\*corresponding author; e-mail: [zagielski@gmail.com](mailto:zagielski@gmail.com)

<sup>†</sup>The term “interevent time” appears in the literature under such names as “pausing time”, “waiting time”, “intertransaction time” and “interoccurrence time” in the context of different versions of the continuous-time random walk formalism [4, 9–12].

<sup>‡</sup>Whether the value of losses or profits in the basic stochastic process are statistically independent is irrelevant because any possible correlations between them are absent in our derivations.

Our principal goal is to model the empirical data (provided in Refs. [5–8]) associated with single-variable statistics, i.e., (i) the mean interevent time period  $R_Q$  between extreme (excessive) losses, defined as those below a negative threshold  $-Q$ , as a function of the  $Q(> 0)$  value<sup>§</sup> and (ii) the distribution  $\psi_Q(\Delta_Q t)$  of interevent times between losses  $\Delta_Q t$ , previously described using ad hoc  $q$ -exponentials [5, 6]. Note that the  $q$ -exponentials used in Refs. [5–8] cannot produce the key empirical data in item (i), and thus in our approach we use superstatistics. Although small losses (and also profits) are of little concern to traders, we focus both on small and medium to high  $Q$ -values. Our goal is to provide market superstatistics that have universality.

## 2. Basic results

We here find an analytically closed form of the mean interevent time period  $R_Q$  between excessive (extreme) losses that is greater than some threshold  $Q$ , i.e.,

$$R_Q^{-1} = P(-\varepsilon \leq -Q) = P(\varepsilon \geq Q) = \int_Q^\infty D(\varepsilon) d\varepsilon, \quad (1)$$

where  $D(\varepsilon)$  is the density of returns given by the Weibull distribution of extreme (or excessive) losses<sup>¶</sup> [14–16],

$$D(\varepsilon) = \frac{\eta}{\bar{\varepsilon}} \left( \frac{\varepsilon}{\bar{\varepsilon}} \right)^{\eta-1} \exp \left( - \left( \frac{\varepsilon}{\bar{\varepsilon}} \right)^\eta \right), \quad \bar{\varepsilon}, \quad \eta > 0. \quad (2)$$

Note that we consider random variable  $\varepsilon$  to be an increment of some underlying stochastic process. In general, values of this random variable can be dependent [13]. It is sufficient to consider herein the case when  $\eta < 1$  (see Table I) which means that distribution  $D(\varepsilon)$  is, for  $\varepsilon/\bar{\varepsilon} \gg 1$ , a decreasing truncated power-law [17].

<sup>§</sup>For the sake of simplicity, we will treat losses as positive quantities.

<sup>¶</sup>The remaining extreme distributions, i.e., the Gumbel and Fréchet ones, are unable to give a proper description of empirical data.

Reference [18] uses the Weibull distribution to describe the statistics of interevent times between subsequent transactions for a given asset. We use the Weibull distribution and the conditional exponential distribution of the CTRW valley model to derive superstatistics or complex statistics associated with the threshold of excessive losses.

Substituting (2) into (1), we obtain

$$R_Q = \exp\left(\left(\frac{Q}{\bar{\varepsilon}}\right)^\eta\right), \quad (3)$$

i.e.,  $\ln R_Q$  increases vs. the relative (dimensionless) variable  $Q/\bar{\varepsilon}$  according to a power-law.

The solid curves in Fig. 1 indicate the predictions generated by (3) and fit the empirical data (the points are represented by different marks). This basic agreement enables us to construct the corresponding superstatistics and allows us to study the successive empirical data. Because the statistical error is low we are able to determine  $\eta$  and  $\bar{\varepsilon}$  (see Table I) and derive the subsequent parameters that define the shape of the superstatistics.

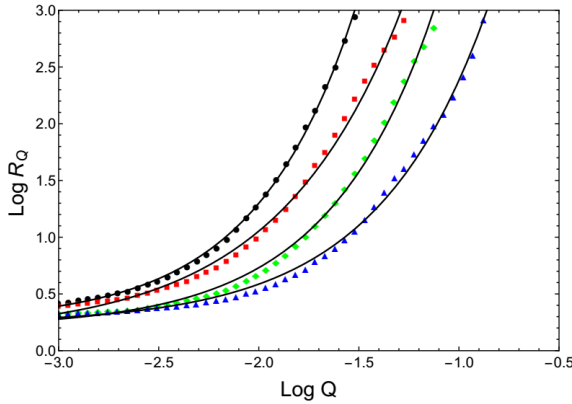


Fig. 1. Mean interevent time period  $R_Q$  vs. threshold  $Q$  for four typical classes of quotations. Black circles, red squares, green rhomboids, and blue triangles concern US/GBP exchange rate, S&P 500 index, IBM stock, and WTI (the crude oil) empirical data (from January 2000 to June 2010), respectively, taken from Fig. 2 in Ref. [5] (plotted from the top curve down to the bottom one). The solid, well fitted curves present predictions of our formula (3).

TABLE I

Values of exponent  $\eta$  and quantity  $\bar{\varepsilon}$  obtained from the fit of predictions of formula (3) to the empirical data (all of them plotted in Fig. 1).

Index/Par.	$\eta$	$\bar{\varepsilon}$
US/GBP	$0.8756 \pm 0.0156$	$0.0037 \pm 0.0003$
S&P500	$0.6981 \pm 0.0292$	$0.0035 \pm 0.0005$
IBM	$0.8246 \pm 0.0236$	$0.0078 \pm 0.0007$
WTI	$0.7855 \pm 0.0182$	$0.0131 \pm 0.0008$

### 3. Superstatistics and its empirical verification

We next construct an unnormalized, unconditional distribution  $\psi_Q(\Delta_Q t)$  of the interevent time stochastic variable,  $\Delta_Q t$ , in the form of superstatistics based on the Weibull distribution used in Sect. 2,

$$\psi_Q(\Delta_Q t) = \int_Q^\infty \psi_Q(\Delta_Q t|\varepsilon) D(\varepsilon) d\varepsilon. \quad (4)$$

Here we assume the conditional distribution  $\psi_Q(\Delta_Q t|\varepsilon)$  is in the exponential form<sup>‡</sup>

$$\psi_Q(\Delta_Q t|\varepsilon) = \frac{1}{\tau_Q(\varepsilon)} \exp\left(-\frac{\Delta_Q t}{\tau_Q(\varepsilon)}\right). \quad (5)$$

Because it is conditional, the next (subsequent) loss is exactly  $\varepsilon$ , and the relaxation time is given by the stretched exponential

$$\tau_Q(\varepsilon) = \tau_Q(0) \exp((B_Q \varepsilon)^\eta) \quad (6)$$

as a straightforward extension of the exponential relaxation time used in the CTRW valley model introduced by [1, 19–21] in the context of photocurrent relaxation in amorphous films. Here  $\tau_Q(0)$  is a free ( $\varepsilon$ -independent) relaxation time, and quantity  $B_Q (> 0)$  is independent of variable  $\varepsilon$  which is an analog of the valley depth. Exponents in (6) (Ockham's razor principle) and to derive superstatistics  $\psi_Q(\Delta_Q t)$  in an exact closed analytical form. Note that the stochastic dependence of interevent time  $\Delta_Q t$  on loss  $\varepsilon$  assumed in (5) is confirmed when smaller losses appear more frequently than larger ones. This is described by definition (6) in which conditional mean time  $\langle \Delta_Q t|\varepsilon \rangle = \tau_Q(\varepsilon)$  is a monotonically increasing function of  $\varepsilon$ . This creates an expanding hierarchy of interevent times where larger losses and profits appear less frequently than smaller ones.

Substituting (6) and (5) into (4) we finally derive a superstatistics in the searched form

$$\psi_Q(\Delta_Q t) = \frac{1}{\tau_Q(Q)} \frac{\alpha_Q}{(\Delta_Q t/\tau_Q(Q))^{1+\alpha_Q}} \times \gamma_{\text{Euler}}(1 + \alpha_Q, \Delta_Q t/\tau_Q(Q)), \quad (7)$$

where the scaling shape exponent

$$\alpha_Q = \frac{1}{(B_Q \bar{\varepsilon})^\eta} = \frac{1}{\ln(\tau_Q(\bar{\varepsilon})/\tau_Q(0))}, \quad (8)$$

and the lower incomplete gamma function

$$\gamma_{\text{Euler}}(1 + \alpha_Q, \Delta_Q t/\tau_Q(Q)) = \int_0^{\Delta_Q t/\tau_Q(Q)} y^{\alpha_Q} \exp(-y) dy. \quad (9)$$

It is possible to analytically prove that  $R_Q$  (or  $\ln R_Q$ ) is the control variable that allows a universal form of (7) that depends solely on  $R_Q$  (or  $\ln R_Q$ ). This variable was used previously in the analogous context in connection with the  $q$ -exponential [5]. Moreover, we obtain the

<sup>‡</sup>The exponential form of the conditional distribution (5) assumes that the losses of a fixed value  $\varepsilon$  are statistically independent, which is generally not valid for different values of losses.

superscaling of the scaling variable  $\ln R_Q$  or the scaling of scaling, i.e., the scaling of the scaling exponent  $\alpha_Q$  (for details see [22]).

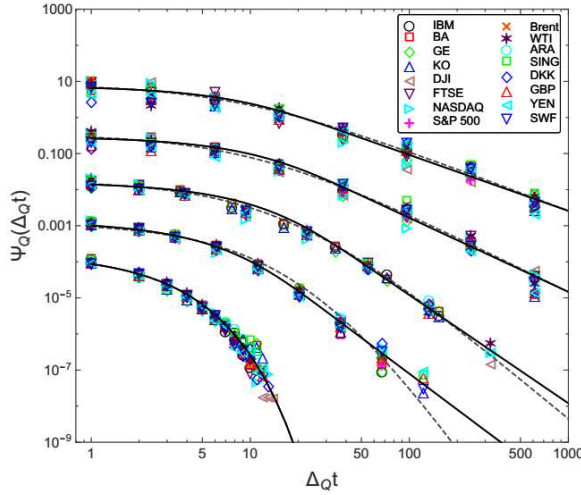


Fig. 2. Collected plots of empirical distributions (colored marks) and theoretical superstatistics,  $\psi_Q(\Delta_Q t)$ , (black solid curves), which are predictions of our formula (7) (while the dashed curves were given by  $q$ -exponential shown by Eq. (3) in Ref. [5]) vs. interevent time,  $\Delta_Q t$ , for  $R_Q = 2, 5, 10, 30$ , and  $70$  (going from the bottom curve to the top one — for the relative daily price returns for sixteen typical examples of financial data in the period 1962–2010 all empirical data were drawn from Refs. [5, 6]).

Figure 2 shows the agreement between the predictions of (7) and the empirical data for IBM for  $R_Q = 2, 5, 10, 30$ , and  $70$  (going from the bottom curve to the top one).

Table II shows the corresponding fits of quantities  $\alpha_Q$  and  $\tau_Q(Q)$  (for more empirical results see Ref. [22], especially Figs. 2 and 5 there). Thus by analytically and empirically proving the  $R_Q$ -dependence of the unconditional superstatistics,  $\psi_Q(\Delta_Q t)$ , we explain the empirical data collapse shown in Fig. 2 (for details see Ref. [22]).

TABLE II

Values of exponent  $\alpha_Q$  and quantity  $\tau_Q(Q)$  obtained from the fit of formula (7) to the empirical data representing companies shown in Fig. 2 for  $R_Q = 2, 5, 10, 30, 70$ .

$R_Q$	$\alpha_Q$	$\tau_Q(Q)$
2	1000	1.5436
5	2.30	2.70
10	2.0	5.0
30	1.050	5.560
70	0.550	4.760

#### 4. Concluding remarks

We find an explicitly closed form of the threshold interevent time superstatistics (7) that is valid for excessive losses, is the foundation of the continuous-time random walk valley model, and that is useful in the study of a double action market (see [12] and refs. therein). These superstatistics are more credible than the  $q$ -exponential distribution that is applied *ad hoc* in this context in Refs. [5, 6], and they agree with the key empirical relation between the mean interevent time  $R_Q$  and the threshold  $Q$  (see Fig. 1).

We model the empirical data collapse (cf. Fig. 2) using superstatistics as a function of a single aggregated variable  $R_Q$ .

Note that using our microscopic model to simulate agent behavior [23, 24] gives results very close to those predicted by Eq. (7). An approach using agent-based modeling in this context was recently explored by other authors [25].

#### Acknowledgments

Two of us (M.J. and T.G.) are grateful to the Foundation for Polish Science for financial support. The work of H.E.S. was supported by NSF Grant CMMI 1125290, DTRA Grant HDTRA1-14-1-0017 and ONR Grant N00014-14-1-0738. One of us (R.K.) is grateful for inspiring discussions with Shlomo Havlin (during The Third Nikkei Econophysics Symposium, Tokyo 2004), Armin Bunde (during his visit at Faculty of Physics, University of Warsaw in 2011), and Constantino Tsallis (during the SMSEC2014 in Kobe).

#### References

- [1] G. Pfister, H. Scher, *Adv. Phys.* **27**, 747 (1978).
- [2] J.W. Haus, K.W. Kehr, *Phys. Rep.* **150**, 263 (1987).
- [3] R. Kutner, F. Świtała, *Quant. Finance* **3**, 201 (2003).
- [4] T. Sandev, A. Chechkin, H. Kantz, R. Mentzler, *Frac. Calc. Appl. Anal.* **18**, 1006 (2015).
- [5] J. Ludescher, C. Tsallis, A. Bunde, *Eur. Phys. Lett.* **95**, 68002 (2011).
- [6] J. Ludescher, A. Bunde, *Phys. Rev. E* **90**, 062809 (2014).
- [7] M.I. Bogachev, A. Bunde, *Phys. Rev. E* **78**, 036114 (2008).
- [8] M.I. Bogachev, A. Bunde, *Phys. Rev. E* **80**, 026131 (2009).
- [9] K.W. Kehr, R. Kutner, K. Binder, *Phys. Rev. B* **23**, 4931 (1981).
- [10] J. Perelló, J. Masoliver, A. Kasprzak, R. Kutner, *Phys. Rev. E* **78**, 036108 (2008).
- [11] A. Kasprzak, R. Kutner, J. Perelló, J. Masoliver, *Eur. Phys. J. B* **76**, 513 (2010).
- [12] T. Gubiec, R. Kutner, *Phys. Rev. E* **82**, 046119 (2010).

- [13] E. Bertin, M. Clusel, *J. Phys. A Math. Gen.* **39**, 7607 (2006).
- [14] Y. Malevergne, D. Sornette, *Extreme Financial Risks. From Dependence to Risk Management*, Springer-Verlag, Berlin 2006.
- [15] P. Embrechts, C. Klüppelberg, Th. Mikosch, *Modelling Extremal Events for Insurance and Finance*, Springer-Verlag, Berlin 1997.
- [16] J. Franke, W. Härdle, Ch. Hafner, *Statistics of Financial Markets*, Springer-Verlag, Berlin 2004.
- [17] B.B. Mandelbrot, *Fractals and Scaling in Finance*, Springer-Verlag, New York 1997.
- [18] P.Ch. Ivanov, A. Yuen, B. Podobnik, L.Y. Youngki, *Phys. Rev. E* **69**, 056107 (2004).
- [19] E.W. Montroll, G.H. Weiss, *J. Math. Phys.* **6**, 167 (1965).
- [20] *Aspects and Applications of the Random Walk, Random Materials and Processes, Vol. 5*, Ed. G.H. Weiss, North-Holland, Amsterdam 1994.
- [21] H. Scher, E.W. Montroll, *Phys. Rev. B* **12**, 2455 (1975).
- [22] M. Denys, M. Jagielski, T. Gubiec, R. Kutner, H.E. Stanley, [arXiv:1509.06315v1](https://arxiv.org/abs/1509.06315v1) [q-fin.ST] (2015).
- [23] M. Denys, T. Gubiec, R. Kutner, *Acta Phys. Pol. A* **123**, 513 (2013).
- [24] M. Denys, T. Gubiec, R. Kutner, [arXiv:1411.1689v1](https://arxiv.org/abs/1411.1689v1) [q-fin.ST] (2014).
- [25] V. Gontis, S. Havlin, A. Kononovicius, B. Podobnik, H.E. Stanley, [arXiv:1507.05203](https://arxiv.org/abs/1507.05203) [q-fin.GN] (2015).