Anomalous Impact in Reaction-Diffusion Financial Models

I. Mastromatteo,[1] B. Tóth,[2] and J.-P. Bouchaud[2]
[1]Centre de Mathématiques Appliquées, CNRS, UMR7641, Ecole Polytechnique, 91128 Palaiseau, France
[2]Capital Fund Management, 23-25 Rue de l'Université, 75007 Paris, France
(Received 19 March 2014; published 31 December 2014)

We generalize the reaction-diffusion model $A + B \to \emptyset$ in order to study the impact of an excess of A (or B) at the reaction front. We provide an exact solution of the model, which shows that the linear response breaks down: the average displacement of the reaction front grows as the square root of the imbalance. We argue that this model provides a highly simplified but generic framework to understand the square-root impact of large orders in financial markets.

DOI: 10.1103/PhysRevLett.113.268701 PACS numbers: 89.65.Gh, 02.50.Ey, 05.40.-a, 82.40.Ck

In most systems, small perturbations induce proportionally small responses: this is the linear response regime. Critical systems are exceptions to this general rule: long-range correlations make these systems particularly fragile. A well known example of an anomalously large response is the magnetic susceptibility close to the para- or ferromagnetic transition. In fact, exactly at the transition, the magnetization M is zero in the absence of an external force (the magnetic field H), but behaves when $H \to 0$ as $M \sim H^\delta$ with $\delta < 1$ (for example $\delta = 1/3$ in mean-field, see, e.g., Ref. [1]). The fact that $\delta < 1$ is tantamount to saying that the linear response coefficient $\lim_{H\to 0} M/H$ diverges at criticality, indicating anomalous fragility. Conversely, the observation of a diverging linear response suggests a nontrivial underlying organization of the system. This is partly the reason why the recently reported universal anomalous impact of small trades in financial markets has triggered a spree of activity (see, e.g., Refs. [2–5] and references therein). Market impact is not only a problem of paramount importance for finance practitioners (for whom market impact amounts to trading costs), it also relates to one of the most fundamental questions in theoretical economics: Why and how do prices change? Market impact is at the core of the sophisticated mechanism through which markets absorb trading information as an input and produce prices as an output [6,7]. The failure of such a mechanism can have dramatic consequences for society, ranging from market inefficiencies to full-fledged crashes (see, e.g., Ref. [8] and references therein).

More precisely, by "market impact" we mean the average price change $\mathcal{I}$ after the sequential execution of a total volume Q of contracts (which we call metaorder). Contrary to the models customarily employed in the field of theoretical economics [9], in which $\mathcal{I}$ is traditionally assumed to be a linear function of Q, a growing consensus in the empirical literature indicates that impact follows a concave law, which is well described by the so called "square-root" impact formula

$$\mathcal{I} = Y\sigma_D\left(\frac{Q}{V_D}\right)^\delta, \tag{1}$$

where δ is an exponent in the range 0.4–0.7, and σ_D and V_D are, respectively, the daily price fluctuations and the daily traded volume [2,4,5,10–12]. Y is a dimensionless coefficient, which is found to be of order 1.

As mentioned above, the fact that $\delta < 1$ indicates that markets are inherently fragile: vanishingly small traded volumes are expected to have a disproportionate impact on prices. Even more surprisingly, the law appears to be universal, as it is to a large degree independent of details such as the type of contract traded, the geographical position of the market venue (U.S., Europe, Asia), the time period (1995 → 2014), or the strategy used to execute the metaorder [4]. It appears to be extremely robust against microstructural changes; for example, the rise of high-frequency trading (HFT) in the last ten years seems to have had very little effect on the validity of Eq. (1) (compare Refs. [2,11], which use pre-2004 data, with Refs. [4,5,10], which use post-2007 data). Such a universality is the main reason why one should expect simple models to be able to reproduce the square-root law. If the relevant properties of the market are included in a stylized model, the low-frequency properties of the dynamics (say, from some hours to a few days) should be correct even if the high-frequency (say, below 1 min) description is inaccurate or not realistic.

In this spirit, we propose here a coarse-grained model of the market much inspired by Refs. [13,14], which relies on two fundamental ingredients in order to describe market dynamics: (i) participants place and update orders to buy (sell) at prices as low (high) as possible, and (ii) market clearing (buy orders and sell orders annihilate each other when at the same price). Therefore, we postulate, as in Refs. [4,5], the existence of a latent order book (modeled as a one-dimensional grid of length L) encoding the trading intentions of the market participants: in this setting each price level x can be populated by particles of two types (B and A), representing, respectively, the intended orders to buy (bids) and to sell (ask). In practice, such a book can be seen as a proxy for the supply and demand curves at the intraday scale. Such a point of view tacitly assumes segregation among the two particle types (i.e., supply and demand curves do not overlap), and implicitly enforces the presence of a finite spread separating the highest bid (the rightmost B particle) and the lowest ask (the leftmost A particle) through a market clearing condition.

The stochastic dynamics that we propose for the particles populating the book consists of a hopping process (all particles can jump either right or left with probability D per unit time) and of a reaction process mimicking the market clearing condition: particles at the same site have a probability λ per unit time to start a reaction process (we will eventually consider the limit $\lambda \to +\infty$). The reaction process may have three different outcomes, chosen at random according to the value of two parameters p and m:

$$A + B \to \emptyset \quad \text{with probability} \quad 1 - p, \tag{2}$$

$$A + B \to B \quad \text{with probability} \quad p\frac{1+m}{2}, \tag{3}$$

$$A + B \to A \quad \text{with probability} \quad p\frac{1-m}{2}. \tag{4}$$

For $p = 0$, this boils down to the model studied in Refs. [13,14], but this setting is too restrictive as it does not allow one to introduce a bias m, which is of course a crucial ingredient to study impact. In fact, the events associated with $p > 0$ can be interpreted as due to the action of an additional agent, who adds to the system an extra bid particle [with probability $(1+m)/2$] or an extra ask particle [with probability $(1-m)/2$]. The lack of a conservation law for the difference between the number of buy and sell particles is then explained by the imbalance introduced by such an extra agent. Finally, we suppose that a flux of particles per unit time $J_B = J_A = J$ (of type B and A) is inserted at the boundaries (respectively, at sites 1 and L). Hence, the system lies in a nonequilibrium state due to the presence of an external particle pressure, representing the flux of orders coming from new participants, that can become interested in entering the market. The model will only make sense if the results do not depend on L, which is to a large extent arbitrary.

The master equation for this system yields an equation for the average density of A and B particles, which in the regime $1 \ll x \ll L$ can be approximated by the continuous dynamics:

$$\frac{\partial\langle b(x,t)\rangle}{\partial t} = D\frac{\partial^2\langle b(x,t)\rangle}{\partial x^2} - \lambda u_A\langle a(x,t)b(x,t)\rangle, \tag{5}$$

$$\frac{\partial\langle a(x,t)\rangle}{\partial t} = D\frac{\partial^2\langle a(x,t)\rangle}{\partial x^2} - \lambda u_B\langle a(x,t)b(x,t)\rangle, \tag{6}$$

where $a(x,t)$ and $b(x,t)$ are the densities of particles of type A and B, and $u_A = 1 - p((1+m)/2)$ and $u_B = 1 - p((1-m)/2)$. In this limit the boundary conditions are of Neumann type:

$$J = -D\frac{\partial\langle b(x,t)\rangle}{\partial x}\bigg|_{x=0}, \qquad 0 = -D\frac{\partial\langle b(x,t)\rangle}{\partial x}\bigg|_{x=L}, \tag{7}$$

$$0 = -D\frac{\partial\langle a(x,t)\rangle}{\partial x}\bigg|_{x=0}, \quad -J = -D\frac{\partial\langle a(x,t)\rangle}{\partial x}\bigg|_{x=L}. \tag{8}$$

This model is extremely hard to solve in one dimension due to the presence of strong correlations among the particle positions [15,16]. Although in higher dimension (or in the small coupling regime $\lambda J^{-1/2}D^{-1/2} \ll 1$) the mean-field approximation $\langle ab\rangle = \langle a\rangle\langle b\rangle$ is quite accurate, in one dimension and in the large coupling regime $\lambda J^{-1/2}D^{-1/2} \gg 1$ (which is relevant here), interactions are too strong for the mean-field prediction to be even qualitatively correct [15,16]. In that case, even in the simpler case $p = 0$, it is necessary to rely on approximate results obtained by using sophisticated renormalization group techniques [17,18] or to resort to numerical simulations [19,20].

In our setting, the symmetric case $p = 0$ corresponds to the case in which the flux of the market is balanced; i.e., no metaorder is being executed. Hence, it represents the market unperturbed state, and it is then worth discussing its main features. First, we remark that in the symmetric case $u_A = u_B$, due to the conservation law for the difference of A and B particles, the combination $\varphi = b - a$ follows the diffusion equation $\partial_t\varphi = D\partial^2_{xx}\varphi$, subject to the boundary condition $-D\partial_x\varphi|_{x=0,L} = J$. Its associated stationary state is a linear density profile:

$$\varphi_{\rm st}(x) = -(J/D)(x - L/2). \tag{9}$$

Second, the interface of the model x_t^* (corresponding to the traded price) diffuses anomalously: while at large times the boundaries obviously confine the system between $x = 0$ and $x = L$, in the small time regime $tD/L^2 \ll 1$ the interface diffuses very slowly, as shown in the bottom curve of Fig. 1. We numerically find the law of $|x_t^* - x_0^*|$ to be compatible with $\sim \log t$, as opposed to the case $J = 0$ considered in Ref. [17] for which $|x_t^* - x_0^*| \sim t^{1/4}$. In particular for $L \to \infty$ the interface—and hence the midprice—is subdiffusive. Despite being at odds with empirical observations of actual financial markets, subdiffusion of the price within the model is expected from the confining effect of the order book itself. Reproducing the diffusive behavior of prices in financial markets (namely, the scaling $|x_t^* - x_0^*| \sim t^{1/2}$ for times larger than a few trades) would require additional terms in our model, accounting for the strategic interactions of the traders (see Refs. [4,5] for a detailed discussion of this point). Summarizing, Eq. (5) alone provides an appropriate description for the linear

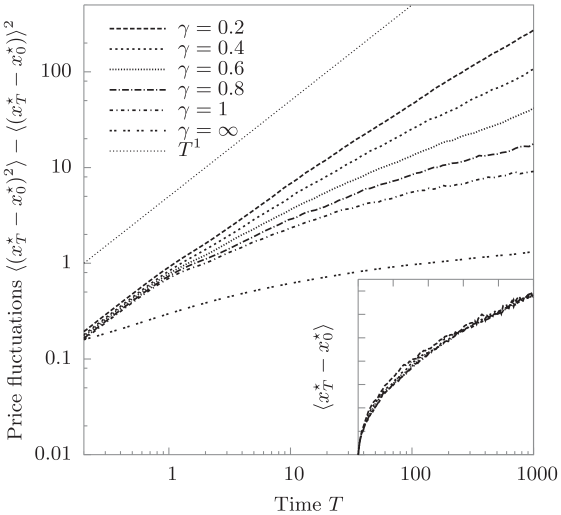

FIG. 1. Fluctuations in the interface position for a modified model in which the terms u_A and u_B are random variables. In particular we change Eq. (2) by choosing with probability $1-p$ the sign of the reaction ($A+B \to$ either A or B) according to a zero-mean, long-range correlated process with tail exponent γ. We find that the diffusion properties of the model change even though the impact properties are unaffected. We plot the variance of the interface position for different values of γ for the set of parameters $L=400$, $J=D=1$, $\lambda=1000$, and $p=m=0$. The inset shows a perfect data collapse for the impact curves (as a function of T) obtained at different γ.

dynamics of price, but fails to describe its quadratic variations.

The goal of the present discussion is to investigate the change in the interface position due to an imbalance in the order flux, i.e., the case $p \neq 0$, $m \neq 0$. We model such an imbalance by supposing that the system, after being prepared in the symmetric stationary state at time $t=0$, is subject to a sudden change of the values p and/or m controlling the imbalance parameters u_A, u_B until a time $t=T$. In that case, it is convenient to study the evolution of the linear combination $\psi = u_B b - u_A a$, which again follows the diffusion equation

$$\frac{\partial \langle \psi(x,t) \rangle}{\partial t} = D \frac{\partial^2 \langle \psi(x,t) \rangle}{\partial x^2} \tag{10}$$

with boundary conditions

$$J u_B = -D \frac{\partial \langle \psi(x,t) \rangle}{\partial x}\bigg|_{x=0}, \qquad J u_A = -D \frac{\partial \langle \psi(x,t) \rangle}{\partial x}\bigg|_{x=L}, \tag{11}$$

where in particular $u_A \neq u_B$ for $0 < t < T$. The interest in the field ψ lies in the fact that its zeros coincide with the zeros of the field $\varphi = b - a$, because of the reduced size of the reaction zone in the regime $\lambda \to \infty$. Hence, by identifying the average price change $\langle x_t^* \rangle$ with the point verifying $\langle \varphi(\langle x_t^* \rangle, t) \rangle = 0$, it is possible to connect the solution of Eq. (10) with the expected position of the interface at a time $t=T$ after the initial perturbation. The solution of Eq. (10) subject to the boundary conditions (11) and the initial conditions (9) is

$$\begin{aligned} f(y,\tau) = {}& \frac{1}{12}(u_B - u_A) - \frac{u_B + u_A}{2} y + \frac{u_B - u_A}{2} y^2 \\ & + (u_B - u_A)\tau - \frac{u_B - u_A}{2} \sum_{n=1}^{\infty} \frac{\cos(2\pi n y)}{\pi^2 n^2} e^{-4\pi^2 n^2 \tau}, \end{aligned} \tag{12}$$

where we have defined the dimensionless variables

$$\tau = DT/L^2, \tag{13}$$

$$y = x/L - 1/2, \tag{14}$$

$$f(y,\tau) = \frac{D}{JL} \psi(y(x), T(\tau)). \tag{15}$$

An inspection of Eq. (12) at $\tau = 0$ reveals that the motion of the interface is due to the discontinuous shape of $\psi(x,0)$ right after the perturbation: the smooth stationary shape of $\varphi_{\rm st}(x)$ is mapped into the piecewise linear function $\psi(x,0)$. Additionally, the boundary conditions for ψ are asymmetric, implying that in the modified coordinates a current is induced due to the mismatch of the incoming fluxes. The trajectory of the average midpoint $\langle x_T^* \rangle = L(1/2 + y_\tau^*)$ can be computed by exploiting the relation

$$0 = \frac{d}{d\tau} f(y_\tau^*, \tau) = \frac{\partial f}{\partial y} \dot{y}_\tau^* + \frac{\partial f}{\partial \tau}, \tag{16}$$

while the partial derivatives can be extracted from Eq. (12), which implies

$$\begin{aligned} \frac{\partial f}{\partial \tau} &= (u_B - u_A)\Theta_3(\pi y, e^{-4\pi^2 \tau}), \\ \frac{\partial f}{\partial y} &= -\frac{u_B + u_A}{2} + (u_B - u_A) \int_0^y dy' \Theta_3(\pi y', e^{-4\pi^2 \tau}), \end{aligned} \tag{17}$$

where $\Theta_3(z,q)$ is the Jacobi theta function of the third kind. The above expressions can be used to solve Eq. (16) with respect to $\dot{y}_\tau^*$. A small τ expansion for $\Theta_3(\pi y, e^{-4\pi^2 \tau})$ leads finally to a differential equation for the trajectory y_τ^*, whose solution is

$$y_\tau^* = 2\alpha(u_B/u_A)\tau^{1/2}, \tag{18}$$

where the function $\alpha(z)$ satisfies the transcendental equation

$$\alpha(z)\left(\frac{z+1}{z-1}-\mathrm{erf}[\alpha(z)]\right)-\frac{1}{\sqrt{\pi}}e^{-\alpha^2(z)}. \qquad (19)$$

Equations (18) and (19) are our central result: they state that the average change in the interface position grows as the square root of the rescaled time. Moreover, when putting back the original units, one finds that $\langle x_T^* \rangle - L/2 = 2\alpha(DT)^{1/2}$, independent of L. Hence, in the infinite size limit $DT/L^2 \to \infty$ the impact is unaffected by the long size behavior of the system. Finally, in this regime Eq. (18) becomes exact, as the large L regime corresponds to the small τ limit. Numerical simulations of the model have been performed in this regime, finding perfect agreement with Eq. (18) (see Fig. 2).

In order to relate these findings to empirical results on market impact [Eq. (1)], we need to link the variation of the midprice $\langle x_T^* \rangle$ to the executed volume Q. According to the financial interpretation suggested above, $p > 0$ represents the action of an additional agent, which for $m \neq 0$ is introducing a bias in the volume imbalance. Hence, it is natural to identify such a bias as the volume Q executed by the agent. Its average is equal to $\langle Q \rangle = \int dx \langle (b-a) \rangle = D \int dt (\langle \partial_x a \rangle_{x=x^{*,+}} + \langle \partial_x b \rangle_{x=x^{*,-}})$, the average number of A particles that reached the interface minus the number of B particles that touched the reaction zone within $t=0$ and $t=T$. Another quantity of interest is $\langle V \rangle = D \int dt (\langle \partial_x a \rangle_{x=x^{*,+}} - \langle \partial_x b \rangle_{x=x^{*,-}})$, which is equal to the total number of particles that reacted within that same time interval. One can accurately approximate $\langle Q \rangle$ and $\langle V \rangle$ by mapping Eq. (17) on the original coordinate system, so as to integrate in time the fluxes through the interface. Exploiting again the properties of the Jacobi theta function of the third kind, one finds

$$\langle Q \rangle = \beta_0(u_B/u_A)(JT), \qquad (20)$$

$$\langle V \rangle = \beta_1(u_B/u_A)(JT), \qquad (21)$$

where the functions $\beta_n(z)$ are given by

$$\beta_n(z) = \frac{(z^{2-n}-1)^{1+n} - \mathrm{erf}[\alpha(z)](z^{1+n}-1)^{2-n}}{2z}. \qquad (22)$$

Equation (20) leads to an approximate estimate of the impact of the type $\mathcal{I} = 2\alpha(QD/\beta_0 J)^{1/2}$, in agreement with the simulation results shown in the inset of Fig. 2. Equation (22) characterizes the imbalance parameter $z = u_B/u_A$ as a function of the local participation rate of the additional agent $\phi = 2Q/(Q+V)$, whose average is equal in the mean-field approximation to

$$\langle \phi(z) \rangle = \frac{2\beta_0(z)}{\beta_0(z)+\beta_1(z)}. \qquad (23)$$

Equation (22) also identifies the Y term appearing in Eq. (1) with the combination $Y(z) = 2\alpha(z)\beta_0^{-1/2}(z)$. For small ϕ, this tends to $Y \approx (\phi/\pi)^{1/2}$, whereas at large ϕ, Y slowly tends to $Y_\infty = 2^{1/2} \approx 1.41$. Interestingly, empirical observations indeed suggest that Y is roughly independent of ϕ.

The above results hold in an extremely broad context. (i) If drift terms of the type $\mu\langle \partial_x a \rangle, \mu \langle \partial_x b \rangle$ or if decay terms $-\nu\langle a \rangle, -\nu \langle b \rangle$ are added to Eq. (5), then an additional time scale appears in the model. In this case Eqs. (18) and (22) still provide a correct description of the system in the regime of small times. Second, (ii) when changing the reaction term $\lambda u_{A/B} ab$ to any other symmetric combination of a and b, the equation for ψ will be unaltered. This implies the same equation for x_t^*, as for infinite λ the zeros of φ and ψ will still coincide. Hence, by appropriately tuning the reaction term (see Fig. 1) it is possible to change the diffusion properties of the system all the way from $\log t$ to $t^{1/2}$ without affecting the square-root impact law (18).

In this Letter, we have provided an analytically tractable implementation of the type of system proposed in Ref. [4]: in our model market clearing indeed induces a locally linear (V-shaped) liquidity profile close to the traded price, which in turn induces a square-root impact shape, as suggested by the mean-field argument in Ref. [4]. However, it is highly nontrivial that such a mean-field argument gives the correct answer since the fluctuations in the interface position are in fact found to be much larger than the impact itself. It is

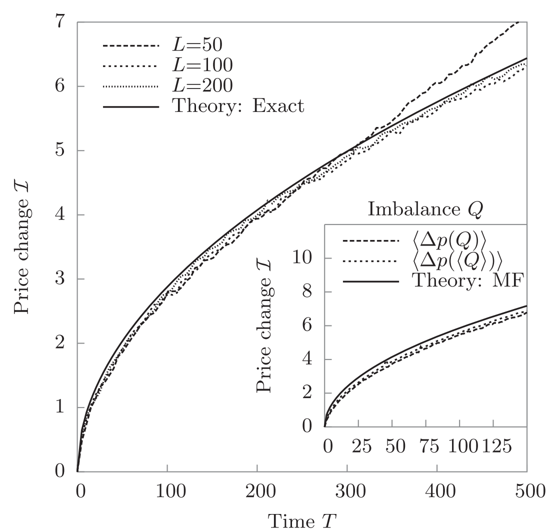

FIG. 2. Main figure: average change in the position of the midpoint $\mathcal{I} = \langle x_t^* \rangle - L/2$ after a perturbation of duration T. We compare the results of simulations of systems of different length (dashed lines) with the analytical prediction valid in the limit $L \to \infty$ (solid line), finding very good agreement. We used the parameters $J = D = 1$, $p = 0.5$, and $m = 0.75$. The limit $\lambda \to \infty$ is enforced by setting $\lambda = 10^3$. The crossover of the curve to the linear regime (indicating $Dt/L^2 \gtrsim 1$) appears in the curve for $L = 50$. Inset: average change in the midpoint position $\mathcal{I}$ against the volume imbalance Q for a simulated system of length $L = 100$ (dashed lines). The solid line indicates the mean-field (MF) estimate predicted by Eq. (20). We have chosen the parameters $D = J = 1$, $\lambda = 1000$, $p = 1$, and $m = 0.5$.

therefore very important to exhibit a model where the square-root impact can be established analytically (rather than numerically, as in previous papers [4,5]). Even though the exact predictions of our stylized model might depend on the actual choice of the reaction parameters, our results suggest that in a one-dimensional system of annihilating particles, a concave dependence of the interface position on the flux imbalance should be regarded as the rule, rather than as the exception. This confirms that very generic features (diffusion and market clearing condition) are, as surmised in Ref. [4], sufficient to explain the anomalous reaction of prices to volume imbalances. As emphasized in previous papers and recalled in the introduction, this also means that markets are "critical," i.e., generically close to an instability, since the liquidity is vanishingly small in the vicinity of the current price. Liquidity fluctuations setting the noise level for $\psi(x,t)$ are thus bound to play a crucial role, and we expect these fluctuations to be at the heart of the turbulent dynamics of financial markets [4,8,21].

We thank R. Benichou, J. de Lataillade, C. Deremble, J. Donier, D. Farmer, J. Gatheral, J. Kockelkoren, P. Kyle, Y. Lempérière, F. Lillo, M. Potters, and H. Waelbroeck for many discussions on these issues. This research benefited from the support of the "Chair Markets in Transition," under the aegis of the "Louis Bachelier Finance and Sustainable Growth" laboratory, a joint initiative of École Polytechnique, Université d'Évry Val d'Essonne, and Fédération Bancaire Française.

[1] N. Goldenfeld, *Lectures on Phase Transitions and the Renormalization Group* (Addison-Wesley, Reading, MA, 1992).
[2] R. Almgren, C. Thum, E. Hauptmann, and H. Li, Risk **18**, 57 (2005).
[3] J. D. Farmer, A. Gerig, F. Lillo, and H. Waelbroeck, Quant. Finance **13**, 1743 (2013).
[4] B. Toth, Y. Lempérière, C. Deremble, J. de Lataillade, J. Kockelkoren, and J.-P. Bouchaud, Phys. Rev. X **1**, 021006 (2011).
[5] I. Mastromatteo, B. Toth, and J.-P. Bouchaud, Phys. Rev. E **89**, 042805 (2014).
[6] J.-P. Bouchaud, J. D. Farmer, and F. Lillo, in *Handbook of Financial Markets: Dynamics and Evolution* (Elsevier, New York, 2008), pp. 57–156.
[7] J.-P. Bouchaud, in *Encyclopedia of Quantitative Finance* (John Wiley & Sons, New York, 2010).
[8] J.-P. Bouchaud, in *Lessons from the 2008 Crisis*, edited by A. Berd (Risk Books, London, 2011).
[9] A. S. Kyle, Econometrica: Journal of the Econometric Society **53**, 1315 (1985).
[10] C. Gomes and H. Waelbroeck, "Is Market Impact a Measure of the Information Value of Trades? Market Response to Liquidity vs Informed Trades," SSRN Report No. 2291720, 2013 (to be published).
[11] N. Torre and M. Ferrari, *Market Impact Model Handbook* (BARRA Inc., Berkeley, 1997), http://www.barra.com/newsletter/nl166/miminl166.asp.
[12] E. Moro, J. Vicente, L. G. Moyano, A. Gerig, J. D. Farmer, G. Vaglica, F. Lillo, and R. N. Mantegna, Phys. Rev. E **80**, 066102 (2009).
[13] P. Bak, M. Paczuski, and M. Shubik, Physica (Amsterdam) **246A**, 430 (1997).
[14] L. H. Tang and G. S. Tian, Physica (Amsterdam) **264A**, 543 (1999).
[15] E. Ben-Naim and S. Redner, J. Phys. A **25**, L575 (1992).
[16] S. Cornell and M. Droz, Phys. Rev. Lett. **70**, 3824 (1993).
[17] G. T. Barkema, M. J. Howard, and J. L. Cardy, Phys. Rev. E **53**, R2017 (1996).
[18] J. M. Park and M. W. Deem, Eur. Phys. J. B **10**, 35 (1999).
[19] S. J. Cornell, Phys. Rev. E **51**, 4055 (1995).
[20] M. Araujo, H. Larralde, S. Havlin, and H. E. Stanley, Phys. Rev. Lett. **71**, 3592 (1993).
[21] F. Lillo and J. D. Farmer, Fluctuations and Noise Letters **05**, L209 (2005).