

Porque as bolsas de valores quebram: a origem das caudas grossas nas distribuições de retornos

(*Why stock markets crash: the origin of fat tailed distributions of returns*)

H.B. do Nascimento Jr.¹, U.L. Fulco^{1,2}, M.L. Lyra¹, M. Serva^{1,3} e G.M. Viswanathan¹

¹*Instituto de Física, Universidade Federal de Alagoas, Maceió, AL, Brasil*

²*Departamento de Física, Universidade Federal do Piauí, Teresina, PI, Brasil*

³*Dipartimento di Matematica and I.N.F.M. Università dell'Aquila, L'Aquila, Italy*

Recebido em 21/12/2006; Aceito em 9/5/2007

Um problema importante na física concerne a origem dos grandes eventos na dinâmica de sistemas complexos, tais como terremotos, epidemias, extinção de espécies e quebras nas bolsas de valores. Aqui revisamos recentes avanços sugerindo que os grandes eventos na dinâmica dos mercados financeiros surgem dos efeitos de memória de longo alcance. Estudamos as distribuições nos mercados financeiros medidas em diferentes intervalos de tempo τ e comparamos esses resultados com aqueles que foram posteriormente obtidos embaralhando os dados com o objetivo de reduzir a memória. Especificamente, destruindo todas as correlações embaralhando a ordem dos retornos, mas sem mudar a distribuição $\tau = 1 d$, significa reduzir a probabilidade dos eventos raros. A distribuição dos retornos para $\tau > 1 d$ perde as caudas grossas e adquire um formato mais Gaussiano. Entretanto, embaralhando só os sinais - mas não o módulo - dos retornos permite as caudas grossas e os grandes eventos persistirem para $\tau > 1 d$. A partir desses resultados, podemos concluir que os grandes eventos são causados pela conhecida correlação de longo alcance multifractal no módulo das séries temporais financeiras.

Palavras-chave: econofísica, correlações de longo alcance, distribuições de caudas grossas.

An important problem in physics concerns the origin of very large events in the dynamics of complex systems, such as large earthquakes, pandemics, mass extinctions and financial crashes. Here we review recent advances that suggest that the largest events in the dynamics of stock markets are due to long-range memory effects. We have studied the distributions of stock returns measured over different time lags τ and compared the results with those obtained after shuffling the data to reduce the memory. Specifically, destroying all correlations by shuffling the order of the returns, but without changing the $\tau = 1 d$ distribution, significantly reduces the probability of very large events. The distribution of returns for $\tau > 1 d$ loses the fat tails and acquires a more Gaussian shape. However, shuffling only the signs—but not the modulus—of the returns allows the fat tails and large events to persist for $\tau > 1 d$. From these results, one can conclude that the very large events are caused by known multifractal long-range correlations in the modulus of the financial time series.

Keywords: econophysics, long-rang correlations, fat-tailed distributions.

1. Introdução

Nas últimas décadas, uma área interdisciplinar de pesquisa, hoje chamada de econofísica, desenvolveu-se com uma participação crescente de físicos, o que pode ser aferido pelo crescimento exponencial de trabalhos publicados na área. O mercado financeiro é um sistema aberto onde um grande número de operadores (*traders*) interagem entre si e reagem às informações externas para determinar o preço de um dado ativo financeiro (ações, títulos, *commodities*, etc). Assim, decisões de compra e venda influenciam nos preços de ati-

vos que influenciam novas decisões, apresentando características de cooperação e competição. A complexidade deste comportamento permite a utilização de conceitos e técnicas de física estatística, tais como teoria de fenômenos críticos, para a modelagem dos fenômenos econômicos [1].

Ultimamente diversos progressos foram feitos para descrever fenômenos não-Gaussianos, incluindo os fenômenos observados em sistemas econômicos [2] e sociais [3], que são tipicamente caracterizados pela presença de distribuições de caudas grossas [4]. O comportamento do mercado financeiro tem recentemente

¹E-mail: helciomaceio@gmail.com.

[2, 5-14] se tornou um foco de interesse para físicos tão bem como uma área de pesquisa ativa devido a sua dinâmica rica e complexa [15-30]. Uma questão que está em aberto diz respeito à distribuição dos retornos em mercados de ações. Sabemos muito bem que o modelo Gaussiano [31] não é bom para descrever probabilidades de grandes eventos, como por exemplo quebras nas bolsas de valores. De fato a distribuição dos retornos possui caudas grossas [2, 14, 15]. Tais caudas grossas podem surgir devido a distribuição de Lévy. De acordo com o teorema do limite central generalizado, a soma de τ variáveis aleatórias independentes de Lévy (isto é, não correlacionadas) apresenta também distribuição de Lévy, tal que a persistência das caudas grossas para grandes τ surge unicamente da distribuição original de Lévy ($\tau = 1$). Além disso, mesmo quando $\tau = 1$ a distribuição é de Lévy com truncamento exponencial na cauda, esperamos ainda um certo grau de persistência das caudas grossas [32]. Por outro lado, as caudas grossas podem persistir [11] para $\tau > 1$ a partir de correlações de longo alcance em uma “variável oculta” [9], tal como a volatilidade (isto é, medida local do desvio padrão) [7]. Assim as correlações de longo alcance frequentemente podem gerar caudas grossas [29]. Sabemos que os retornos não possuem correlações de longo alcance, porém o valor absoluto dos retornos possui correlações de longo alcance [9, 10, 12, 33-36]. O valor absoluto dos retornos são correlacionados com expoentes de escalas não-único [9, 12] (ou seja, multifractal). Neste trabalho nós estudamos como testar a hipótese (ver Ref. [11]) que a correlação de longo alcance da volatilidade é a origem das caudas grossas.

Mostraremos que para retornos no mercado de ações, a persistência das caudas grossas para grandes τ não pode ser explicada sem a correlação de longo alcance da volatilidade. Por isso utilizamos modelos já conhecidos de embaralhamento dos retornos no intuito de estudar a origem das caudas grossas. Primeiramente embaralhamos os retornos diários, assim destruimos todas as correlações sem mudar a distribuição $\tau = 1 d$. Para os dados embaralhados, mostraremos que a distribuição é de cauda grossa para $\tau = 1 d$ mas não para $\tau > 1 d$. Assim, descobriremos que as correlações na volatilidade são mais apropriadas para explicar a origem das caudas grossas do que a distribuição de Lévy para $\tau = 1 d$, que são responsáveis pela existência das caudas grossas para grandes τ . Posteriormente, mostraremos que embaralhando só os sinais dos retornos, as caudas grossas irão persistir para $\tau > 1 d$ e a distribuição não convergirá para a Gaussiana. Estes resultados provam que as correlações de longo alcance da volatilidade (em vez das correlações de curto alcance dos sinais [13]) são responsáveis pelas caudas grossas para

qualquer $\tau > 1 d$. Nós também mostraremos que as correlações de curto alcance dos sinais ($1-2d$) representam um papel importante nas propriedades da distribuição de pequenas variações de preços. Vamos ainda explicar porque os dados Gaussianos são mais apropriados para pequenos retornos do que para grandes retornos.

2. Modelos “ERD” e “ESR”

Nosso grupo de dados consistem nos log-retornos gerados a partir de 59 índices do mercado de ações (obtidos a partir do site yahoo.com: AEX, AORD, ATG, ATX, BFX, BSESN, BVL30, BVSP, CCSI, DJA, DJI, DJT, DJU, DOT, FCHI, FTSE, HEX, HSI, IBC, IGRA, IIX, IPSA, IXIC, JKSE, KFX, KLSE, KS11, KSE, MERV, MID, MTMS, MXX, N225, NDX, NTOT, NYA, NZ40, OEX, PSE, PSI, PX50, RUA, RUI, RUT, SAX, SETI, SML, SMSI, SOOX, SPC, SSEC, SSML, STI, TA100, TSE, TWII, VLIC, XMI, XU100). Os retornos $r(t)$ são definidos em termos dos preços $P(t)$ por

$$r(t) \equiv \log_{10} \frac{P(t)}{P(t-1)} . \quad (1)$$

Normalizamos os retornos para obter variância unitária, com isso somos capazes de comparar retornos medidos em diferentes escalas de tempo. Também definimos um retorno renormalizado r_τ

$$r_\tau(t) \equiv \frac{1}{\sqrt{\tau}} \sum_{t'=t-\tau+1}^t r(t') , \quad (2)$$

com τ medido em dias e $r_1 = r$. Os retornos diários e renormalizados representam um papel similar àqueles representados pelas quantidades “bare” e “dressed” em teoria de campos. Notamos que para retornos Gaussianos unitários não-correlacionados (independentes), suas variâncias serão idênticas de acordo com o teorema do limite central: $\sigma(r_1) = \sigma(r_\tau) = 1$. Similarmente, se $r_1(t)$ apresenta distribuição de Lévy, então $r_\tau(t)$ também apresentará uma distribuição de Lévy.

Mesmo para retornos Gaussianos, entretanto, a presença de correlações pode conduzir a comportamentos anômalos, tal que r_1 e r_τ podem ter distribuições não idênticas. Portanto desenvolvemos um método para “subtrair” os efeitos das correlações. Para cada uma das 59 séries temporais, geramos uma série temporal modificada, embaralhando a ordem dos retornos diários (Fig. 1). Este modelo que embaralha os retornos diários (ERD) terá distribuição idêntica aos dados reais para $\tau = 1 d$, mas sem possuir nenhum tipo de correlação. Portanto, para $\tau > 1 d$ os dados reais e o modelo ERD irão gerar distribuições não idênticas (Fig. 2) a não ser que os dados sejam descorrelacionados. Conseqüentemente, temos agora uma maneira

de testar a hipótese de que as caudas grossas em $p(\tau)$ persistem unicamente devido às correlações. Se a densidade de probabilidade $p(r_\tau)$ possui caudas grossas para os dados reais mas não para o modelo ERD, podemos concluir que as caudas grossas em $p(\tau)$ surgem devido às correlações.

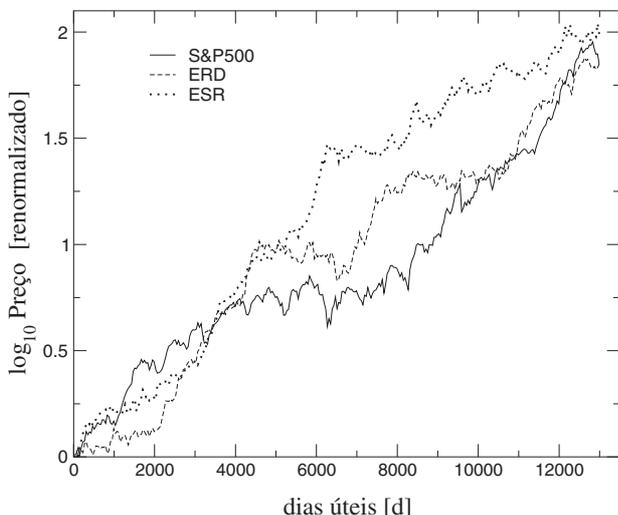


Figura 1 - Índice S&P 500, mostrado em escala logarítmica. Também são mostrados os modelos ERD, no qual os retornos são completamente decorrelacionados mas com uma distribuição $\tau = 1 d$ idêntica, e o modelo ESR, o qual possui uma distribuição do valor absoluto dos retornos diários idêntica, mas sem a correlação dos sinais [1].

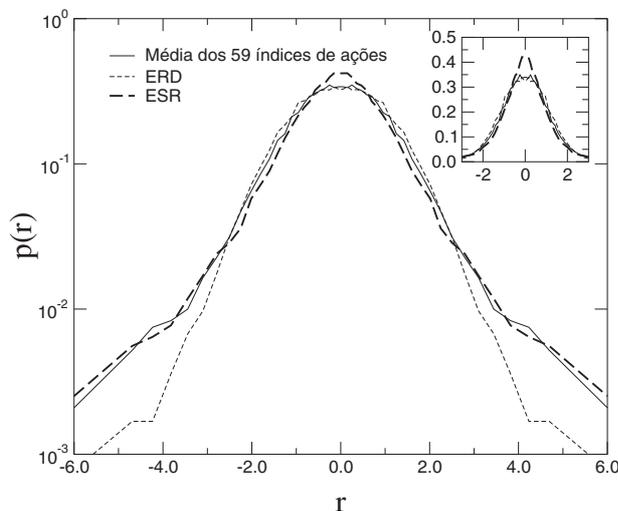


Figura 2 - Distribuição densidade de probabilidade simetrizada $p(r)$ dos retornos r_τ medida em um período de $\tau = 10 d$ para 59 índices do mercados de ações. Também mostramos os modelos ERD e ESR para $\tau = 10 d$ [1]. Essas distribuições são típicas para $\tau > 1 d$. Nós descobrimos que existe caudas grossas no modelo ESR mas não no modelo ERD, indicando que a origem das caudas grossas é causada pela correlação de longo alcance no valor absoluto dos retornos. O gráfico menor segue uma escala linear (não semilog). A distribuição de r_1 foi normalizada para obter variância unitária. Para $\tau = 1 d$ todas as três distribuições são idênticas (não mostrado).

Ao embaralharmos os dados, destruímos todos os tipos de correlações, dessa maneira os dados tornam-se números independentes. Especificamente, o embaralhamento destrói as correlações de longo alcance na volatilidade dos retornos tão bem como as correlações de curto alcance nos sinais dos retornos. Portanto desenvolvemos agora um método para “subtrair” só as correlações nos sinais dos retornos $r_1(t)$, enquanto preservamos as correlações no valor absoluto dos retornos (volatilidade) $|r_1(t)|$ (Fig. 1). Para cada uma das 59 séries temporais, geramos uma segunda série temporal modificada, embaralhando a ordem dos sinais, mas não do valor absoluto dos retornos diários. Este modelo que embaralha os sinais dos retornos (ESR) terá uma distribuição simetrizada idêntica aos dados reais e ao modelo ERD para $\tau = 1 d$, mas não necessariamente para $\tau > 1 d$.

3. Funções características modificadas

Posteriormente estudamos a função densidade de probabilidade simetrizada $p(r_\tau)$ dos retornos r_τ a partir dos 59 índices dos mercados de ações e comparamos estes aos modelos ERD e ESR. Focamos nosso estudo nas regiões de caudas grossas da distribuição, utilizando uma função característica modificada definida por

$$f(\tau) = \int dr_\tau p(r_\tau) \exp[-(|r_\tau| - r_0)^2] \quad (3)$$

$$\simeq (1/N) \sum_t \exp[-(|r_\tau| - r_0)^2], \quad (4)$$

onde t é o tempo em dias. Para estudar a região de caudas grossas e manter boas estatísticas, escolhemos um valor $r_0 = 5$ que corresponde a 5 desvios padrões. Similarmente, para estudar a região central da distribuição, definimos uma segunda função

$$g(\tau) = \int dr_\tau p(r_\tau) \exp(-r_\tau^2) \simeq (1/N) \sum_t \exp(-r_\tau^2), \quad (5)$$

onde $N = \sum_t 1$. Na prática, calculamos estas funções diretamente dos retornos e não a partir da distribuição, obtendo assim uma melhor estatística.

Descobrimos assim que as caudas grossas quase desaparecem para $\tau > 1 d$ no modelo ERD, mostrando que as correlações são necessárias para manter as caudas grossas para $\tau > 1$. Estes resultados são consistentes com aqueles relatados na Ref. [11] e elimina a possibilidade que a distribuição de Lévy seja responsável pela persistência das caudas grossas. Na realidade, se isto fosse verdadeiro, as caudas grossas teriam que persistir para $\tau > 1 d$ mesmo depois de embaralhar a ordem dos retornos $r_1(t)$, que contradiria nossos resultados.

Conseqüentemente, concluímos que as caudas grossas surgem principalmente das correlações. Notamos, portanto, que no modelo ERD, as caudas grossas não desaparecem completamente e $p(r_\tau)$ nunca torna-se Gaussiana mesmo para $\tau \rightarrow 100 d$ (dias úteis no mercado de ações), dessa maneira a distribuição de Lévy truncada de $r_1(t)$ pode em princípio surgir para $\tau = 1 d$ [14, 29]. Também pode surgir a distribuição sugerida na Ref. [9]

Nossa descoberta mais importante é que as caudas grossas permanecem completamente intactas para o modelo ESR, mostrando que as caudas grossas podem persistir para $\tau > 1 d$, quando os dados são descorrelacionados nos sinais mas possuem correlação de longo alcance no valor absoluto. Esta descoberta prova que, seja qual for a escolha da distribuição dos retornos diários $p(r_1)$, as correlações de longo alcance na volatilidade são necessárias para explicar o comportamento de $p(r_\tau)$. Uma importante consequência deste resultado é que devemos tomar muito mais cuidado ao estudar as distribuições $p(r_\tau)$ independentemente das correlações.

Nós também descobrimos que a região central da distribuição dos retornos é mais similar ao modelo ERD do que ao modelo ESR para $\tau > 1$, mostrando que nesta região os dados reais são mais parecidos a uma Gaussiana e que a correlação Markoviana dos sinais nos retornos são importantes para manter a aparência Gaussiana. Finalmente, descobrimos que o comportamento da distribuição $p(r_\tau)$ é bastante parecido para diferentes τ .

4. Discussão e conclusão

O ensino de física teórica interdisciplinar pode motivar os alunos a compreender melhor as leis da natureza, tão bem como relacionar a física com fenômenos complexos, econômicos e sociais, que são de grande importância no cotidiano. Os novos resultados relatados aqui são de grande interesse e cientificamente importantes devido a presença das correlações de longo alcance e das distribuições de caudas grossas em diversos fenômenos físicos, químicos e biológicos. Além disso, os preços da maioria dos derivativos financeiros dependem da distribuição dos retornos. A existência da correlação de longo alcance na volatilidade e a distribuição de caudas grossas em séries temporais financeiras já são conhecidas por algum tempo. O que não foi completamente entendido é a origem das caudas grossas. Notamos que as correlações com decaimento exponencial (isto é,

não sendo lei de potência) não podem produzir caudas grossas em grandes intervalos. Um estudo sistemático do índice S&P 500 *et al.* [11], encontrou que na escala da distribuição surgem dependências temporais. Aqui, mostramos que embaralhando só os sinais, mas não o valor absoluto, dos retornos as caudas grossas persistem, provando que as caudas grossas surgem das correlações da volatilidade e não de qualquer outro tipo de dependência temporal.

Comentamos sobre os resultados notando que os dados reais são mais parecidos ao modelo ERD na região central da curva. Já o modelo ESR é mais parecido na região das caudas grossas. Fig. 3 mostra que a ausência de correlações na volatilidade causa um colapso das caudas grossas no modelo ERD, mas que a ausência de correlações nos sinais causa grande desvio da região central da curva no modelo ESR, para $\tau > 1$. Os dados reais aparecem qualitativamente em algum lugar entre os modelos ERD e ESR. Isto implica que o erro de desprezar a correlação dos sinais pode de algum modo compensar o erro de ignorar a correlação de longo alcance da volatilidade para uma pequena variação de preços. Para $\tau > 1 d$ tiramos como um resultado interessante que a região central é enganosamente bem descrita por uma Gaussiana. Os dois erros cancelam-se um ao outro, por este motivo a teoria Gaussiana para retornos de ações foi um sucesso (a afirmação de que a distribuição dos retornos era Gaussiana, foi entretanto a base do prêmio Nobel de Economia em 1997). O principal erro da teoria Gaussiana é que ela não pode explicar eventos raros ou grandes eventos, tal como a grande quebra de 1987. Existem outros fenômenos conhecidos na qual o cancelamento de dois erros conduz a bons modelos, um exemplo clássico sendo a teoria de polímeros de Flory [37], no qual o erro em estimar as energias repulsivas e atrativas cancelam um ao outro, tal que o modelo torna-se melhor do que esperado.

Em resumo, nossos resultados indicam que as distribuições de caudas grossas dos retornos surgem principalmente da correlação da volatilidade. Mais geralmente, mostramos que as distribuições de caudas grossas podem ter origem a partir da correlação de longo alcance no valor absoluto de qualquer série temporal. Notamos que as técnicas de embaralhamento que revisamos aqui são gerais e podem ser aplicadas ao estudo de séries temporais geradas por outros sistemas complexos similares.

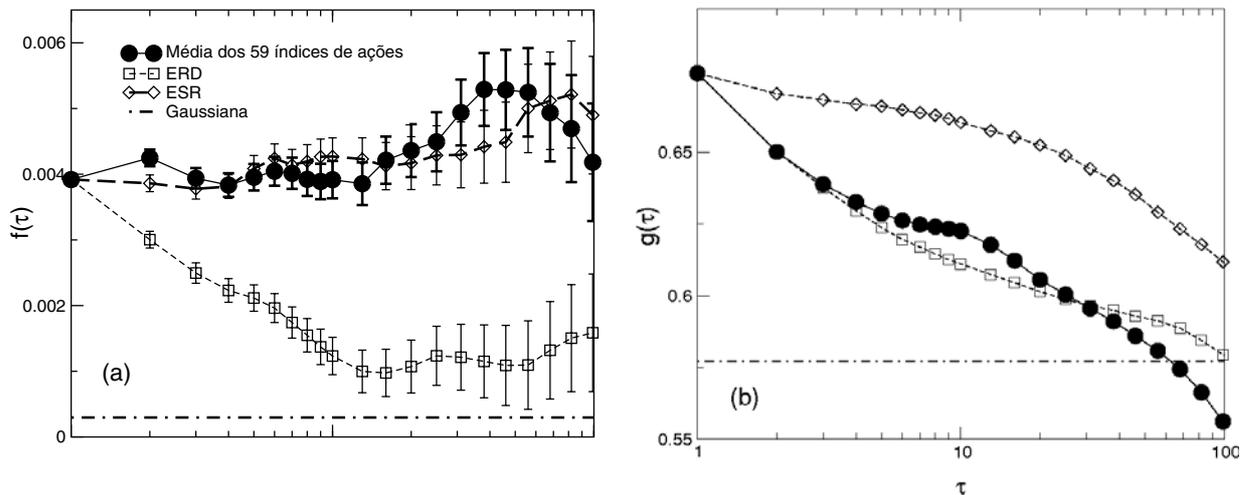


Figura 3 - (a) Função característica média $f(\tau)$ para os 59 índices do mercado de ações, no gráfico podemos ver também os modelos ERD e ESR. Os valores para o modelo ERD tornam-se muito pequenos com τ , indicando que as caudas são menos grossas para os dados embaralhados. Os valores para o modelo ESR, conseqüentemente, são consistentes com os dados originais, mostrando que a correlação de longo alcance da volatilidade é a causa real da distribuição de caudas grossas [1]. A perda da cauda no modelo ERD exclui a possibilidade da distribuição ser de Lévy estável. (b) Função característica média $g(\tau)$ para o mesmo grupo de dados. Notamos que no modelo ESR existem muito mais retornos próximos de zero para $\tau > 1 d$, conduz a um maior valor de g . Este resultado mostra que as correlações dos sinais nos dados reais representam um papel importante, neutralizando as correlações da volatilidade. Um outro resultado que pode ser visto em (a) e (b) é que uma aproximação Gaussiana é extremamente melhor para pequenos r (como podemos ver a partir de $\Delta g/g \simeq 20\%$), mas muito pior para grandes r (desde que $\Delta f/f \simeq 1000\%$), uma descoberta potencialmente importante para teoria de opções de preços.

Referências

- [1] G.M. Viswanathan, U.L. Fulco, M.L. Lyra and M. Serva, *Physica A* **329**, 273 (2003).
- [2] R.N. Mantegna and H.E. Stanley, *An Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge University Press, Cambridge, 2000).
- [3] E.W. Montroll and W.W. Badger, *Introduction to Quantitative Aspects of Social Phenomena* (Gordon and Breach, Nova Iorque, 1974).
- [4] M.F. Shlesinger, G. Zaslavsky and U. Frisch, (eds) *Lévy Flights and Related Topics in Physics* (Springer, Berlin, 1995).
- [5] J.-P. Bouchaud, A. Maticz and M. Potters, *Phys. Rev. Lett.* **87**, 228701 (2001).
- [6] V.M. Eguiluz and M.G. Zimmermann, *Phys. Rev. Lett.* **85**, 5659 (2000).
- [7] B. Podobnik, P. Ch. Ivanov, Y. Lee, A. Chessa and H.E. Stanley, *Europhys. Lett.* **50**, 711 (2000).
- [8] B. Podobnik, P. Ch. Ivanov, Y. Lee and H.E. Stanley, *Europhys. Lett.* **52**, 491 (2000).
- [9] M. Pasquini and M. Serva, *Euro. Phys. J. B* **16**, 195 (2000).
- [10] Y. Liu, P. Gopikrishnan, P. Cizeau, M. Meyer, C.-K. Peng and H.E. Stanley, *Phys. Rev. E* **60**, 1390 (1999).
- [11] P. Gopikrishnan, V. Plerou, L.A.N. Amaral, M. Meyer, and H.E. Stanley, *Phys. Rev. E* **60**, 5305 (1999).
- [12] M. Pasquini and M. Serva, *Econ. Lett.* **65**, 275 (1999).
- [13] R.N. Mantegna and H.E. Stanley, *Nature* **383**, 587 (1996).
- [14] R.N. Mantegna and H.E. Stanley, *Nature* **376**, 46 (1995).
- [15] B.B. Mandelbrot, *J. Bus.* **36**, 394 (1963).
- [16] L.P. Kadanoff, *Simulation* **16**, 261 (1971).
- [17] R.N. Mantegna, *Physica A* **179**, 232 (1991).
- [18] W. Li, *Int. J. Bifurcation and Chaos* **1**, 583 (1991).
- [19] H. Takayasu, H. Miura, T. Hirabayashi and K. Hamada, *Physica A* **184**, 127 (1992).
- [20] P. Bak, K. Chen, J.A. Scheinkman and M. Woodford, *Ricerche Economiche* **47**, 3 (1993).
- [21] J.-P. Bouchaud and D. Sornette, *J. Phys. I France* **4**, 863 (1994).
- [22] S. Ghashghaie, W. Breymann, J. Peinke, P. Talkner and Y. Dodge, *Nature* **381**, 767 (1996).
- [23] M. Mevy and S. Solomon, *Int. J. Mod. Phys. C* **7**, 595 (1996).
- [24] S. Galluccio, G. Caldarelli, M. Marsili and Y.-C. Zhang, *Physica A* **245**, 423 (1997).
- [25] P. Bak, M. Paczuski and M. Shubik, *Physica A* **246**, 430 (1997).
- [26] N. Vandewalle and M. Ausloos, *Physica A* **246**, 454 (1997).
- [27] Y. Liu, P. Cizeau, M. Meyer, C.-K. Peng and H.E. Stanley, *Physica A* **245**, 437 (1997).
- [28] P. Cizeau, Y. Liu, M. Meyer, C.-K. Peng and H.E. Stanley, *Physica A* **245**, 441 (1997).
- [29] B. Podobnik, K. Matia, A. Chessa, P. Ch. Ivanov, Y. Lee and H.E. Stanley, *Physica A* **300**, 300 (2001).

- [30] J.E. Ingersoll, *Theory of Financial Decision Making* (Rowman & Littlefield, Savage, 1987).
- [31] L. Bachelier, Ann. Sci. École Norm. Sup. **17**, 21 (1900).
- [32] R.N. Mantegna and H.E. Stanley, Phys. Rev. Lett. **73**, 2946 (1994).
- [33] M. Greene and B. Fielitz, J. Financial Econ. **4**, 339 (1977).
- [34] V. Akgiray, J. Bus. **62**, 55 (1989).
- [35] A.W. Lo, Econometrica **59**, 1279 (1991).
- [36] N. Crato and P. Rothman, Econ. Lett. **45**, 287 (1994).
- [37] P.J. Flory, *Principles of Polymer Chemistry* (Cornell University Press, Nova Iorque, 1953).