Invariance in the Recurrence of Large Returns and the Validation of Models of Price Dynamics*

Lo-Bin Chang[†], Stuart Geman[‡], Fushing Hsieh[§], and Chii-Ruey Hwang[¶]

March, 2013

*keywords and phrases: market dynamics, distribution-free statistics, market time, GARCH, geometric Brownian motion

[†]Department of Applied Mathematics, National Chiao Tung University

[‡]Division of Applied Mathematics, Brown University

[§]Department of Statistics, University of California, Davis

[¶]Institute of Mathematics, Academia Sinica

Contents

Abstract. Starting from a robust, nonparametric, definition of large returns ("excursions"), we study the statistics of their occurrences focusing on the recurrence process. The empirical waiting-time distribution between excursions is remarkably invariant to year, stock, and scale (return interval). This invariance is related to self-similarity of the marginal distributions of returns, but the excursion waiting-time distribution is a function of the entire return process and not just its univariate probabilities. GARCH models, market-time transformations based on volume or trades, and generalized (Lévy) random-walk models all fail to fit the statistical structure of excursions.

1 Introduction

Given a sequence of stock prices $s_0, s_1, \ldots$ recorded at fixed intervals, say every five minutes, let $r_n \doteq \log \frac{s_n}{s_{n-1}}$, $n = 1, 2, \ldots$, be the corresponding sequence of returns. Fix N and define an excursion to be a return that is large, in absolute value, relative to the set $\{r_1, r_2, \ldots, r_N\}$. Specifically, following Hsieh et al. (2012), define the *excursion process*, $z_1, z_2, \ldots, z_N$:

$$z_n = \begin{cases} 1 & \text{if } r_n \leq l \text{ or } r_n \geq u \\ 0 & \text{if } r_n \in (u, l) \end{cases}$$

where l and u are, respectively, the 10'th and 90'th percentiles of $\{r_1, \ldots, r_N\}$. We call the event $z_n = 1$ an excursion, since it represents a large movement of the stock relative to the chosen set of returns. We will study the distribution of waiting times between large stock returns by studying the distribution of the number of zeros between successive ones of the excursion process. Our motivation includes:

1. The empirical observation (cf. Chang et al., 2013) that this waiting-time distribution is nearly invariant to time scale (e.g. thirty-second, one-minute, or five-minute returns), to stock (e.g. IBM or Citigroup), and to year (e.g. 2001 or 2007).

2. The waiting-time to large returns is of obvious interest to investors, and much easier to study if, and to the extent that, it is invariant across time scale, stock, and year.

3. The particular waiting-time distribution found in the data and its invariance to time scale have implications for models of price and volatility movement. For instance, Lévy processes, "market-time" models based on volume or trades, and GARCH models are each one way or another inconsistent with the empirical data.

4. Overwhelmingly, the evidence for self-similarity comes from studies of the univariate (marginal) return distributions (e.g. evidence for a stable-law distribution), but marginal distributions leave data models underspecified. Waiting-time distributions provide additional, explicitly temporal, constraints, and these appear to be nearly universal.

Larger returns can be studied by using more extreme percentiles. Although we have not experimented extensively, the empirical results we will report on appear to be qualitatively robust to the chosen percentiles and hence the definition of "large return." In general, the upper and lower percentiles index a family of waiting-time distributions that might prove useful to systematically constrain the dynamics of price and volatility models.

In §2, we study the invariance of the empirical waiting-time distribution. Starting with the Lévy type models, we first make a connection between the model-based distribution and the geometric distribution. To be concrete, let $S(t)$ follow the "Black-Scholes model" (geometric Brownian motion) as an example: $d \log S(t) = \mu dt + \sigma dw(t)$, where $w(t)$ is a standard Brownian motion. Because of the independent increments property of Brownian motion $w(t)$, the return sequence under this model is exchangeable (i.e. the distribution of any permutation remains the same). Therefore, the empirical waiting-time distribution under this model is provably invariant to time scale and to time period. More specifically, the probability of getting a "large" return, with l =10'th percentile and u =90'th percentile, is exactly 0.2 at each return interval and the empirical waiting-time distribution is, therefore, nearly a geometric distribution with parameter 0.2 (see §2.1 for more detail). We emphasize the these considerations apply without modification not just to the geometric Brownian motion but to all of its popular generalizations as geometric Lévy processes.

Not surprisingly (cf. "stochastic volatility"), the *actual* (i.e. empirical) waiting-time distribution is different from geometric. But what is surprising is the invariance of this distribution across time scale, stock, and year. In §2.2 we make an exhaustive comparison of empirical waiting-time distributions, using trading prices of approximately 300 stocks from the S&P 500 observed over the eight years from 2001 through 2008. Invariance to timescale is strong in all eight years; invariance to stock is strong in years 2001–2007 and less strong in 2008; and invariance across years is stronger for pairs of years that do not include 2008. (We have not studied the years since 2008.) In §2.3, we will connect waiting-time invariance to self-similarity, being careful to distinguish a self-similar *process* from a process having self-similar increments (i.e. distinguish dynamics from marginal distributions).

Which of the state-of-the-art models of price dynamics are consistent with the empirical distribution of the excursion process? The existence of a nearly invariant waiting-time distribution between excursions provides a new tool for evaluating these models, through which questions of consistency with the data can be addressed using statistical measures of fit and hypothesis tests. In general, we will advocate for permutation and other combinatorial statistical approaches that robustly and efficiently exploit symmetries shared by large classes of models, supporting exact hypothesis tests as well as exploratory data analysis. In §3 we introduce some combinatorial tools for hypothesis testing and explore the implications of waiting-time distributions to the time scale of volatility clustering. We continue with this approach, in §4, with a discussion of stochastic volatility modeling, as well as "market-time" and other stochastic time-change models. We conclude, in §5, with a summary and some proposals for price and volatility modeling.

2 Waiting Times Between Large Returns

There were 252 trading days in 2005. The traded prices of IBM stock (s_n, $n = 0, 1, \ldots, 18{,}899$) at every 5-minute interval from 9:40AM to 3:50PM (seventy five prices each day), throughout the 252 days, are plotted in Figure 1, Panel A.[1] Often, activities near the opening and closing are not representative. To mitigate their influence, we exclude prices in the first ten minutes (9:30 to 9:40) and last 10 minutes (3:50 to 4:00) of each day. The corresponding intra-day returns, $r_n \doteq \log \frac{s_n}{s_{n-1}}$, $n = 1, 2, \ldots, 18{,}648$ (seventy four returns per day) are plotted in Panel B. Overnight returns are not included.

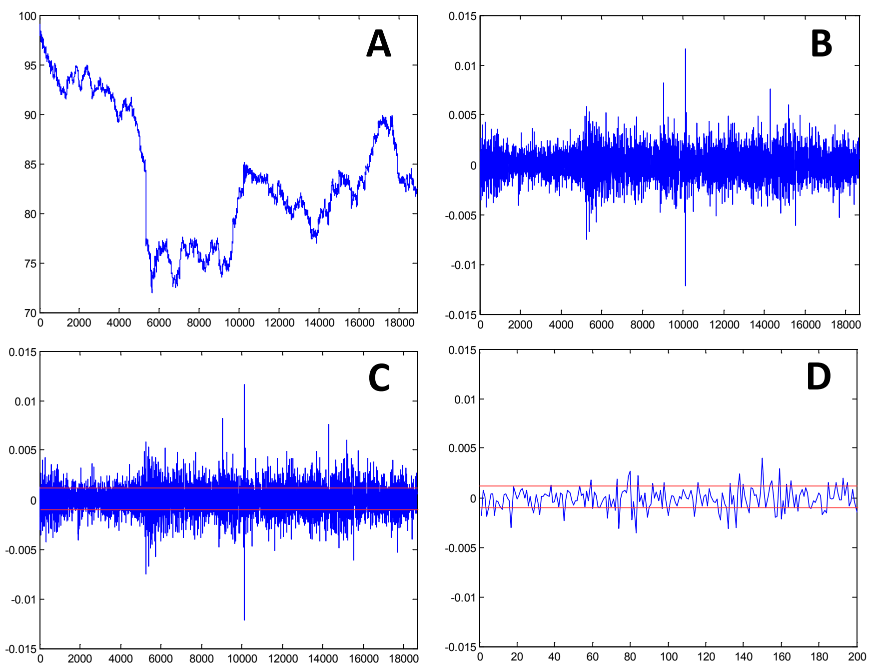


Figure 1: **Returns, percentiles, and the excursion process. A.** IBM stock prices, every 5 minutes, during the 252 trading days in 2005. The opening (9:30 to 9:40) and closing (3:50 to 4:00) prices are excluded, leaving 75 prices per day (9:40,9:45,. . .,15:50). **B.** Intra-day 5-minute returns for the prices displayed in A. There are 252×74=18,648 data points. **C.** Returns, with the 10'th and 90'th percentiles superimposed. **D.** Zoomed portion of C with 200 returns. The "excursion process" is the discrete time zero-one process that signals (with ones) returns above or below the selected percentiles.

We declare a return "rare" if it is rare relative to the interval of study, in this case the calendar year 2005. We might, for instance, choose to study the largest and smallest returns in the interval, or the largest 10% and smallest 10%. Panel C shows the 2005 intra-day returns with the tenth and ninetieth percentiles superimposed. More generally, given any fractions $f, g \in [0, 1]$ (e.g. 0.1 and

[1]The price at a specified time is defined to be the price of the most recent trade.

0.9), define

$$l_f = l_f(r_1, \ldots, r_N) = \inf\{r : \#\{n : r_n \le r, 1 \le n \le N\} \ge f \cdot N\} \tag{1}$$

$$u_g = u_g(r_1, \ldots, r_N) = \sup\{r : \#\{n : r_n \ge r, 1 \le n \le N\} \ge (1-g) \cdot N\} \tag{2}$$

where, presently, $N = 18{,}648$. The lower and upper lines in Panel C are $l_{.1}$ and $u_{.9}$, respectively. Panel D is a magnified view, covering $r_{1001}, \ldots, r_{1200}$, but with $l_{.1}$ and $u_{.9}$ still figured as in equations (1) and (2) from the entire set of 18,648 returns.[2]

The *excursion process* is the zero-one process that signals large returns, meaning returns that either fall below l_f or above u_g:

$$z_n = 1_{r_n \le l_f \text{ or } r_n \ge u_g}$$

Hence $z_n = 1$ for at least 20% of $n \in \{1, 2, \ldots, 18{,}648\}$ in the Figure 1 example. Obviously, many generalizations are possible, involving indicators of single-tale excursions (e.g. $f = 0$, $g = .9$ or $f = .1$, $g = 1$) or many-valued excursion processes (e.g. z_n is one if $r_n \le l_f$, two if $r_n \ge u_g$, and zero otherwise). Or we could be more selective by choosing a smaller fraction f and a larger fraction g, and thereby move in the direction of truly rare events. (There is, then, an inevitable tradeoff between the magnitude of the excursions and the sample size; more rare events are studied at the cost of statistical power.) Here we will work with the special case $f = .1$ and $g = .9$, but a similar exploration could be made of these other excursion processes.

2.1 The role of the geometric distribution

As with the Black-Scholes model discussed in the introduction, any stochastic process with stationary and independent increments (i.e. any Lévy process) has exchangeable increments, and hence exchangeable returns if used as a model for the log-price distribution. What would the excursion waiting-time distribution look like under a geometric Brownian-motion model, or one of its generalizations to geometric Lévy?

Specifically, assume

$$d \log S(t) = \mu dt + \sigma dw(t),$$

where $w(t)$ is a Lévy process. Then the return sequence

$$R_k = \log S(t_0 + k\delta t) - \log S(t_0 + (k-1)\delta t), \ \forall k = 1, 2, 3, \cdots, n \tag{3}$$

is exchangeable. With the particular percentiles used here, the sequence $z_1, z_2, \ldots, z_N$ has 20% 1's and 80% 0's. If real returns were exchangeable then the excursion process would be as well, since the percentiles l_f and u_g (equations 1 & 2) are symmetric functions of the returns. Hence, the probability that a 1 is followed immediately by another 1 (waiting time zero) is very nearly 0.2. (Not *exactly* 0.2, even ignoring edge effects, because there are a finite number of 1's – the first 1 of the pair uses one of them up.) The probability that exactly one 0 intervenes is very nearly (0.8)(0.2)=0.16, two 0's very nearly (0.8)(0.8)(0.2)=0.128, and so-forth following the geometric distribution.

In general, the waiting-time distribution for an exchangeable process converges to the geometric distribution as the number of excursions (number of return intervals) goes to infinity (Diaconis &

[2] To break ties and to mitigate possible confounding effects from "micro-structure," prices are first perturbed, independently, by a random amount chosen uniformly from between $\pm\$0.005$.

Freedman, 1980, Chang et al., 2013). In this sense, the KS distance[3] to the geometric distribution is a measure of departure of a return process from exchangeability, and can be used as a statistic to calibrate the temporal structure of real price data as well as proposed models of prices and returns (as will be discussed more deeply in §3 & §4). Figure 2 compares the empirical waiting-time distribution generated by 93,240 one-minute 2005 IBM returns to the geometric distribution with parameter 0.20. Obviously there is a substantial departure, characterized by high probabilities of short and long waits in the real data as compared to the geometric distribution. (The *slope* of the P-P curve is greater than one or less than one as waiting-time probabilities are respectively larger than or smaller than geometric.) Thus, for example, the empirical probability that the waiting time is zero ($z_{n+1} = 1$ given that $z_n = 1$) is about 0.32 instead of 0.20. Indeed estimates of this probability reliably fall in a narrow range, from about 0.32 to 0.33, independent of the time interval with respect to which returns are defined, the stock from which the returns are derived, and the year from which the data is collected. In fact, the entire empirical waiting-time distribution is a near invariant to time scale, stock, and year, as we shall now demonstrate.

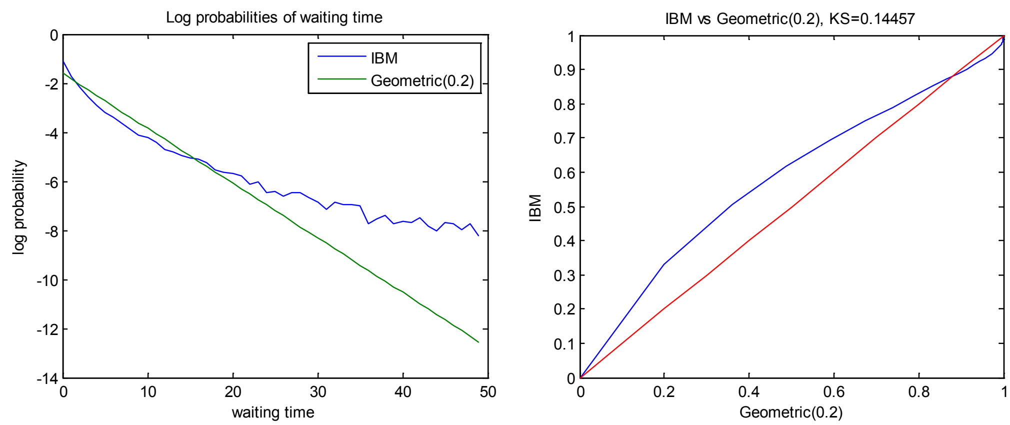


Figure 2: **Geometric(0.2) and empirical waiting times.** The empirical waiting-time distribution of 1-minute returns of IBM stock in 2005 was compared with the geometric distribution with parameter 0.2. **Left panel:** Log plots for the geometric distribution and the empirical waiting-time distribution. The x-axis is the waiting times and the y-axis is the log probabilities of the waiting times. **Right Panel:** P-P plots for the geometric distribution versus the empirical waiting-time distribution. The KS distance is the maximum horizontal (= maximum vertical) distance between the P-P curve (shown in blue) and the diagonal (shown in red).

2.2 Empirical evidence for invariance

Chang et al. (2013) and Hsieh et al. (2012) studied the waiting-time distribution between excursions, i.e. the distribution on the number of zeros between two ones. The empirical waiting-time

[3]Given two cumulative distribution functions, F_1 and F_2, the P-P plot is the two-dimensional curve from $(0, 0)$ to $(1, 1)$ defined by $\{(F_1(t), F_2(t) : t \in R\}$. The Kolmogorov Smirnov (KS) distance is the maximum vertical (and horizontal) distance between the diagonal and the P-P plot, which is also the maximum distance between F_1 and F_2:

$$d_{KS}(F_1, F_2) = \sup_t |F_1(t) - F_2(t)|$$

distribution from 2005, for the 18,648 5-minute returns, the 93,240 1-minute returns, and the 186,480 30-second returns of IBM are shown across the top of Figure 3. They are remarkably similar.

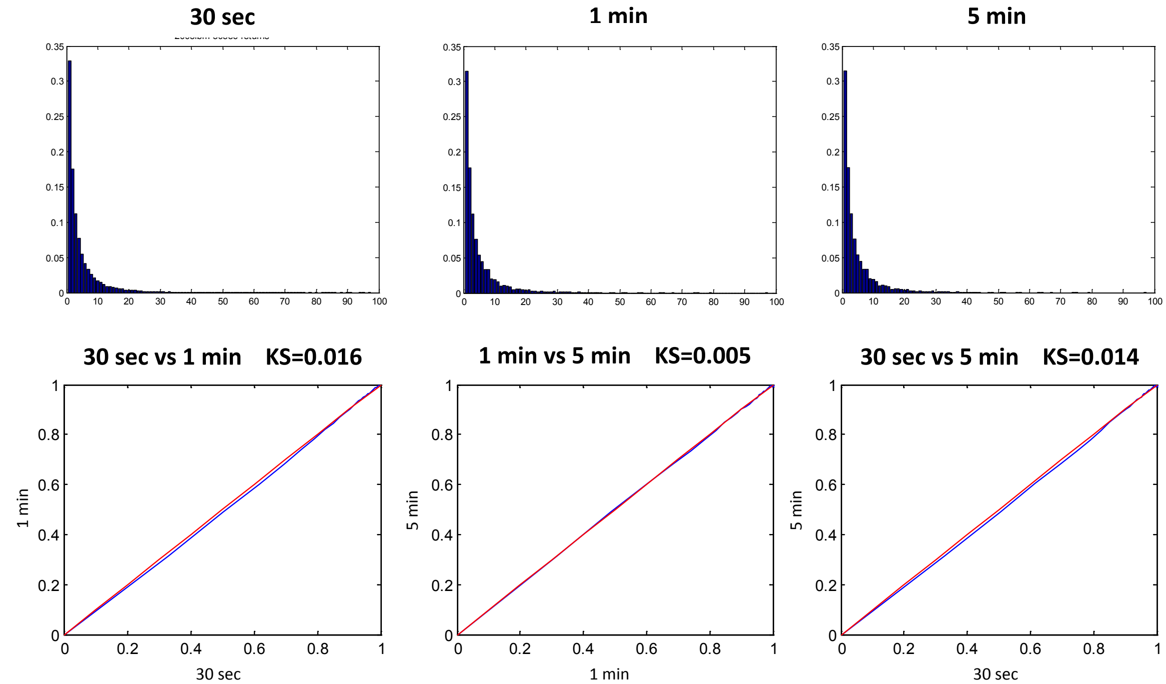


Figure 3: **Scale invariance. Top row:** Empirical waiting-time distributions captured from 30-second, 1-minute, and 5-minute returns of IBM in 2005. **Bottom row:** P-P plots for the three waiting-time distributions taken two at a time, and their corresponding Kolmogorov-Smirnov distances.

Invariance to scale. The bottom row of Figure 3 has three P-P plots that come from taking the three waiting-time distributions (30-second, 1-minute, and 5-minute, shown in the top row) two at a time. The KS distances, one for each comparison, are also shown. The distribution of waiting times between excursions for IBM 2005 returns is strikingly invariant to the return interval. (We are using d_{KS} here as a descriptive statistic, and not for the purpose of hypothesis testing. These waiting times are not *precisely* invariant, and many pairs that look well matched will nevertheless have small p-values, simply because of the large sample sizes.)

Table 1: **Scale invariance, aggregate data.** Approximately 300 stocks were tested. Table shows median KS distances for pairwise comparisons of three time scales (30 seconds, 1 minute, 5 minutes) for each of the years 2001 through 2008.

year	2001	2002	2003	2004	2005	2006	2007	2008
30 sec vs 1 min	0.0199	0.0109	0.0148	0.0148	0.0163	0.0128	0.0113	0.0103
1 min vs 5 min	0.0253	0.0203	0.0197	0.017	0.0175	0.017	0.0194	0.0143
30 sec vs 5 min	0.0348	0.0247	0.0268	0.0259	0.0264	0.0223	0.0242	0.0172

The phenomenon is not unique to IBM, nor to the year 2005. We tested approximately 300 of the S&P 500 stocks for the years 2001 through 2008. The results are summarized in Table 1. In this regard, 2008 is not an outlier, as can be seen from the last column of the table, and from the three histograms of KS distances, one for each pair of return intervals, over all stocks tested in 2008 (Figure 4).

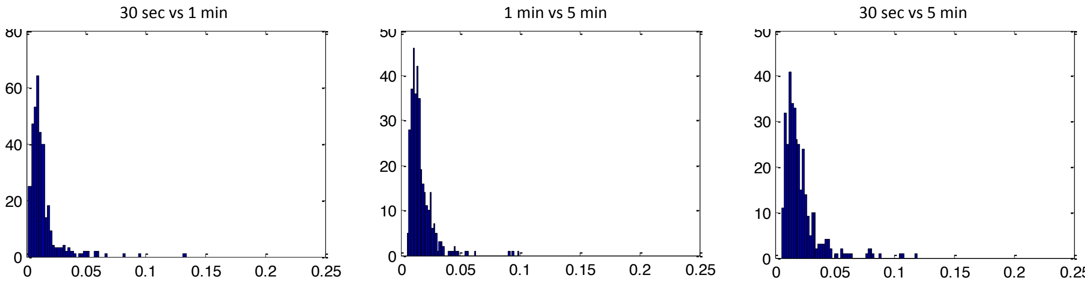


Figure 4: **Histogram of KS distances, 2008.** Each panel shows the histogram of Kolmogorov-Smirnov distances between excursion waiting-time distributions at different time scales in 2008, for approximately 300 stocks.

As we will see shortly, self-similar processes have excursion waiting-time distributions that are invariant to scale. It is interesting, then, to note that the empirical evidence for waiting-time invariance is substantially weaker at larger intervals, e.g. using hourly or daily returns. This same progression is often observed in studies of self-similarity (cf. Mantegna and Stanley, 2000). Possibly, it can be traced to sample size. Because the return sequences are derived from a single calendar year, larger return intervals have smaller numbers of returns, and hence a larger *variance* of the empirical waiting-time distribution. For example, as a rough estimate, we can expect hourly returns to multiply the spread of a five-minute-return across-stock histogram of empirical KS distributions (as in the lower-left panel of Figure 5) by about $\sqrt{60/5} \approx 3.5$, which would substantially obscure the evidence for invariance. It is also possible that invariance systematically breaks down for larger return intervals. We have not explored either hypothesis.

Invariance to stock and year. How do the excursion waiting-time distributions of one stock compare to those of another? For each of the eight years studied we compared the waiting-time distributions, for 5-minute returns, between all pairs of the 300 or so stocks in our data set. See Figure 5 and the accompanying table. With the possible exception of 2008, excursion waiting-time distributions are nearly invariant across stocks.

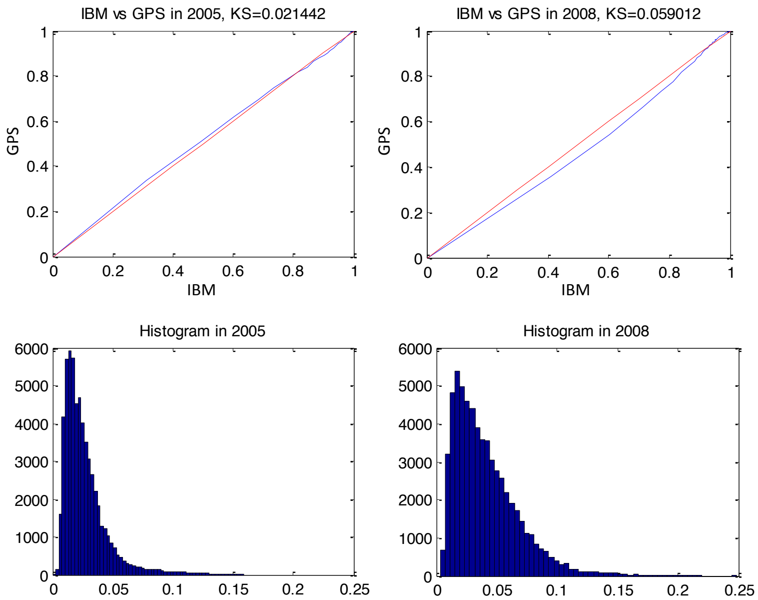
Finally, we examined the change in waiting-time distributions from year to year. For each stock and each return interval (30-seconds, 1-minute, 5-minutes), we compared distributions between pairs of years. Table 2 indicates that waiting-time distributions were typically unchanged during the period 2001 to 2007, but considerably different during the financial crises of 2008.

2.3 Connections to self-similarity

Recall that $P(t)$, $t \geq 0$, is a self-similar process if there exists $H \geq 0$ ("Hurst index") such that

$$\mathcal{L}\{P(\delta t), t \geq 0\} = \mathcal{L}\{\delta^H P(t), t \geq 0\}$$

for all $\delta \geq 0$, where $\mathcal{L}\{Q(t), t \geq 0\}$ denotes the probability distribution ("law") of the *process* $Q(\cdot)$. In other words, the joint distributions of $(P(\delta t_1), P(\delta t_2), \ldots, P(\delta t_m))$ and $\delta^H(P(t_1), P(t_2), \ldots, P(t_m))$ are the same, for all m, $t_1, t_2, \ldots, t_m$, and δ (e.g. Embrechts and Maejima, 2002). Let $S(t)$, $t \geq 0$,



Year	Median
2001	0.0264
2002	0.0293
2003	0.0228
2004	0.0209
2005	0.022
2006	0.0208
2007	0.0293
2008	0.035

Figure 5: **Invariance to stock.** Comparisons of excursion waiting-time distributions for 5-minute returns between IBM and GPS in 2005 (top-left panel) and 2008 (top-right panel). Histograms of KS distributions for all pairs of stocks (bottom panels) show a breakdown of invariance across stocks in 2008 as compared to 2005. **Table:** Summary of year-by-year comparisons of waiting-time distributions across stocks. With the exception of 2008, waiting times are nearly invariant to stock.

be the price of a stock at time t. Beginning with Mandelbrot (1963,1967), it has often been observed that the marginal distribution of the (drift-corrected) increments in price, or more typically log price, is nearly self-similar, e.g. $\log S(\delta t) - \log S(\delta(t-1))$ has nearly the same distribution as $\delta^H \log S(t) - \delta^H \log S(t-1)$, although different methods for estimating the exponent H give different values. Many authors (e.g. Calvet & Fisher, 2002 and Xu & Gencay, 2003) argued that the exponent is not constant (generally decreasing at larger scales) or that there are actually multiple exponents, as in the more general multi-fractal models. Within the framework of (single-exponent) self-similarity, the estimation method of Mantegna and Stanley (1995) is among the most convincing since it focuses on the centers of return distributions rather than their tails. Mantegna and Stanley

Table 2: **Year-to-year changes in excursion waiting-time distributions. Left column:** Medians of KS distances, over all stocks and all pairs of years, 2001 through 2007. **Right column:** Median distances over all stocks from the single pair of years, 2005 and 2008. Waiting-time distributions in 2008 differ substantially from those of previous years.

year	2001 through 2007	2005 vs 2008
30-second returns	0.0236	0.0623
1- minute returns	0.0219	0.0681
5-minute returns	0.0228	0.0811

reported a Hurst index of about 0.71 for the S&P 500, with evidence for self-similarity spanning three orders of magnitude in the return interval, though as they and others (e.g. Bouchaud, 2001) pointed out, scaling breaks down at larger intervals.

Additionally, many authors have studied empirical scaling through a variety of statistics that can be derived from, but are not directly equivalent to, self-similarity. For example, Gopikrishnan et al. (1999) investigated scaling properties of normalized returns, while Wang and Hui (2001) studied scaling phenomena using returns divided by their daily average returns. Gencay et al. (2001) explored wavelet variance, Matteo (2007) used R/S analysis, and Glattfelder et al. (2011) described 12 scaling laws in high-frequency FX data. Wang et al. (2006) studied the return interval between big volatilities and showed the persistence of scaling for a range of time resolution scales ($\delta t = 1, 5, 10, 15, 30$ min).

Here we give a brief explanation of the mathematical relationship between self-similarity and scale invariance of the excursion waiting-distribution. Assume that the drift-corrected log price, $P(\cdot)$, is a self-similar process. Then, as for the return process, at scale δ with drift coefficient r,

$$\begin{aligned} R_t^{(\delta)} &\doteq \log \frac{S(\delta t)}{S(\delta(t-1))} \\ &= P(\delta t) - P(\delta(t-1)) + \delta r \\ \Rightarrow \mathcal{L}\{R_t^{(\delta)}, t \geq 1\} &= \mathcal{L}\{\delta^H R_t^{(1)} + (\delta - \delta^H) r, t \geq 1\} \\ &= \mathcal{L}\{G^{(\delta)}(R_t^{(1)}), t \geq 1\} \end{aligned}$$

where $G^{(\delta)}(x)$ is the monotone function $\delta^H x + (\delta - \delta^H) r$. Now let $Z_n^{(\delta)}$, $n = 1, 2, \ldots, N$, be the excursion process corresponding to the return process $R_n^{(\delta)}$, $n = 1, 2, \ldots, N$, for some scale (interval) δ (e.g. thirty seconds or five minutes). Since percentages are unchanged by monotone transformations, it follows that $\mathcal{L}\{Z_n^{(\delta)}, n = 1, 2, \ldots, N\} = \mathcal{L}\{Z_n^{(1)}, n = 1, 2, \ldots, N\}$, for all $\delta > 0$. In short, self-similarity of the process $P(t)$, $t \geq 0$, implies that the excursion process, and therefore its waiting-time distribution, is invariant to scale.

One family of self-similar models for P, made popular in finance by Mandelbrot's 1963 paper, is the family of stable Lévy processes, i.e. the processes with stable, stationary, and independent increments. But the corresponding returns, $R_1^{(\delta)}, R_2^{(\delta)}, \ldots$, are then iid for all $\delta > 0$, and this violates volatility clustering. This shortcoming (already apparent to Mandelbrot in 1963) has led to the consideration of other self-similar models, that have stationary, and possibly stable, but not-necessarily-independent increments. One way to construct such processes is through random time changes of Brownian motion (Mandelbrot and Taylor, 1967, Clark, 1973, Anderson, 1996, Heyde, 1999, H. Geman et al., 2001). We will return to this approach in §4.3. A more direct approach is with fractional Brownian motion (FBM), which we will briefly discuss now as an illustration of the application of the excursion waiting-time distribution in the study of price fluctuations and their models.

The FBMs are a family of self-similar Gaussian processes, one for each Hurst index $H \in (0, 1]$. The particular value $H = 1/2$ is the ordinary Brownian motion. Which value of H best describes the 5-minute excursion waiting-time distribution of the 2005 IBM data? We explored different values of H. For each value, we generated 500 samples of the process P and extracted 18,648 returns, along with the corresponding excursion processes and their waiting-time distributions. (As discussed, in light of the fact that FBM is self-similar, the waiting-time distribution is invariant to δ.) Each waiting-time distribution has a KS distance to the distribution extracted from the real data. The

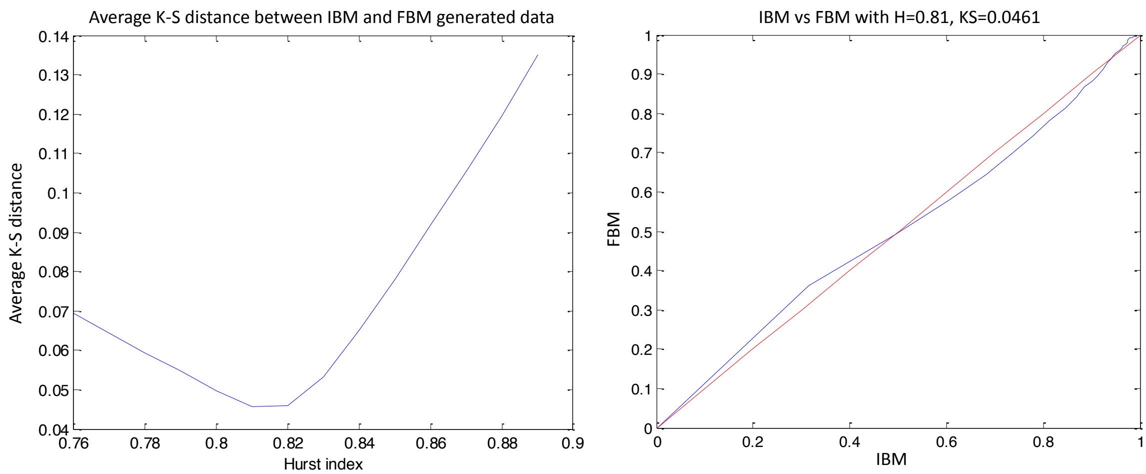


Figure 6: **Fractional Brownian motion and excursion waiting times. Left panel:** For each Hurst index $H = 0.76, 0.77, \ldots 0.89$ we generated 500 FBM samples and extracted 18,648 returns, matching the 18,648 returns in the 5-minute 2005 IBM data. The average KS distances between the FBM excursion waiting times and the empirical IBM waiting times are plotted. The best fit, with KS about 0.046, is at $H = 0.81$. **Right panel:** P-P plot of excursion waiting-time distribution for IBM versus a sample from the best-fitting FBM. FBM overestimates the probabilities of short and very long waiting times.

averages of the 500 KS distances, for each of $H = 0.76, 0.77, \ldots, 0.89$, are shown in the left-hand panel of Figure 6. The smallest KS distance over all examined H values was approximately 0.046, at $H = 0.81$. As can be seen from the right-hand panel of the figure, in comparison to real returns the fitted FBM model has too many short and too many long waiting times.

3 Conditional inference, permutations, and hypothesis testing

Our purpose in this section is to introduce some statistical tools that relate the near-invariance of the excursion waiting-time distribution to the temporal characteristics of the empirical return data, focusing particularly on the time scale of volatility clustering. In the following section, §4, these tools will be used to explore some familiar themes in price-dynamics modeling, including implied volatility, GARCH models, and various approaches to stochastic time change, a.k.a. market time. The statistical characterization of price and volatility fluctuations is obviously very complicated. Under the circumstance, model-free statistical methods can be particularly effective tools for probing dynamics and discerning spatial and temporal patterns. The excursion process itself is an example, in that it avoids absolute thresholds and model-based parameter estimates. Permutation tests are another example, and are particularly suitable for relating the excursion process to the time scales operating in price fluctuations, as we shall now discuss.

3.1 Permutation tests

Returns are not exchangeable. If they were, there would be no stochastic volatility. Whereas we anticipated a failure of exchangeability, what is not apparent is the time scales involved in this departure of real dynamics from the basic random-walk models encapsulated by the geometric Lévy processes. Are the five-minute returns of IBM *locally* exchangeable? What if we were to permute the 12 five-minute returns in each hour; would the price process look any different, either visually or statistically? As for visually, there is certainly no obvious "tell," judging from a comparison of Panels B and C in Figure 7. Panel B plots the prices of IBM at five-minute intervals from 9:45AM to 3:45 PM, on a randomly selected day in 2005. Panel C plots a surrogate price sequence, derived from the original (i.e. the trajectory in Panel B) by permuting, randomly and independently, each set of twelve returns within each of the six hours. The surrogate sequence is started at the same price as the original and therefore again has the same price as the original at each ensuing hour. There is no visual clue that separates the real from the surrogate price sequence, and by our experience there never is one.

How about statistically? Can we detect a difference in the dynamics? Is there any indication that separates a real trajectory from its permutation surrogates? If so, how does this separation depend on time scale? We could as easily permute the set of five-minute returns within each week, each day, each hour, or each thirty-minute interval. At what time scale does exchangeability break down? Put differently, at what time scales does volatility clustering operate? These questions can be systematically and robustly answered through a permutation test, and the resulting departure of the excursion waiting times between the permuted and original trajectories as measured through the KS distance.

Let $r_1, r_2, \ldots, r_{18648}$ be the 18,648 five-minute intra-day returns, as defined in §2. Consider any statistic T (function of these returns), such as the KS distance between the excursion waiting-time distribution and the geometric distribution, as examined in Figure 7, Panel A. And consider the particular "null hypothesis," H_o, that $\mathcal{L}\{(R_{\rho(1)}, R_{\rho(2)}, \ldots, R_{\rho(18648)})\}$ is invariant to the permutations ρ in a set Π, where $R_1, R_2, \ldots, R_{18648}$ are the random variables associates with the observed returns. The point is not that we actually believe H_o (among other things, it violates volatility clustering), but rather that it leads to a measure of departure from exchangeability as determined by the particular statistic being examined, and the particular set of permutations Π. Under the null hypothesis a sequence of M iid permutations, $\rho_1(\cdot), \rho_2(\cdot), \ldots, \rho_M(\cdot)$, chosen from the uniform distribution on the set of permutations in Π, produces a sequence of $M+1$ conditionally iid T's, namely the observed $T_{obs} = T(r_1, r_2, \ldots, r_{18648})$ together with one additional value for each permutation:

$$T_{\rho_m} = T(r_{\rho_m(1)}, r_{\rho_m(2)}, \ldots, r_{\rho_m(18648)}) \quad m = 1, 2, \ldots, M$$

It follows that under H_o

$$Pr\{\#\{m = 1, 2, \ldots, M : T_{\rho_m} \geq T_{obs}\} \geq N\} \leq \frac{N+1}{M+1} \tag{4}$$

In other words, if $N = \#\{m = 1, 2, \ldots, M : T_{\rho_m} \geq T_{obs}\}$ then $(N+1)/(M+1)$ is an exact p-value for H_o, in the direction of the alternative H_a that T_{obs} is larger than would be expected under H_o.[4]

Panel D of Figure 6 illustrates the test with M = 5,000 and Π unrestricted, i.e. the entire permutation group on the sequence $1, 2, \ldots, 18648$. Since T_{obs} is larger than any of the values of T evaluated for the surrogate (i.e. permuted) sequences, $N = 0$ and the test has a p-value of $\frac{1}{5001} \approx 0.0002$. As expected, the waiting-time distribution of real returns is not consistent with exchangeability, and in fact produced the largest deviation from geometric among all of the 5,001 sequences. Suppose now that we restrict Π to include only *local* permutations, say within each day, or hour, or twenty-minute period. Then selecting from the uniform distribution on Π is the

[4]This is an instance of *conditional inference*, in that the test is conditioned on the particular realization. The correctness of the p-value follows from its correctness for any realization.

same thing as independently choosing a permutation for each (non-overlapping) day, or hour, or twenty-minute period, providing a mechanism for systematically exploring the time scale of volatility clustering.

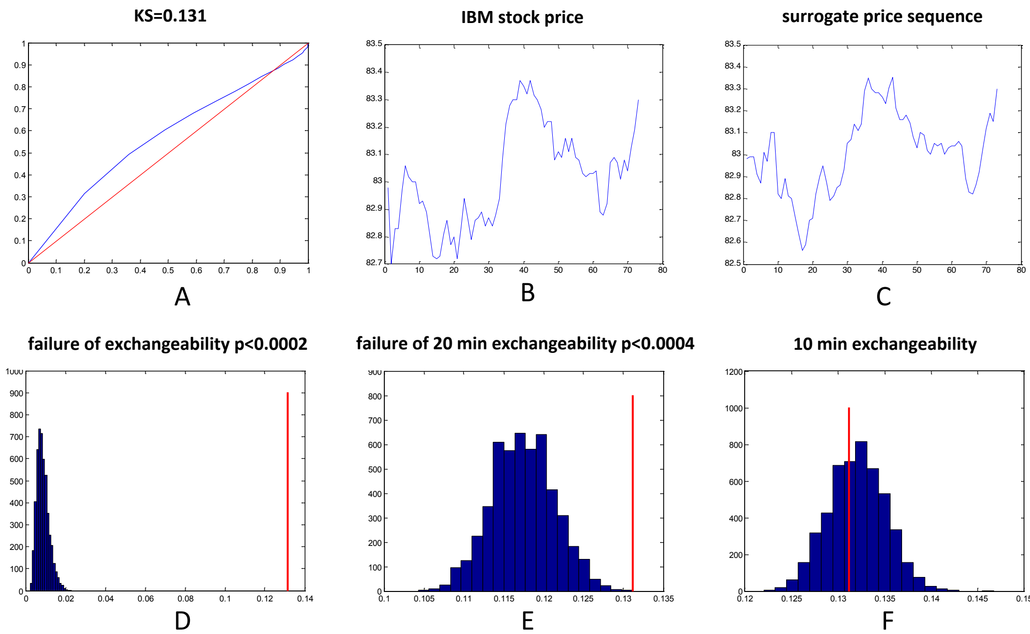


Figure 7: **Exchangeability and time scale.** The five-minute returns on IBM stock in 2005 were tested for their departure from exchangeability, as reflected in the excursion waiting-time distribution. **Panel A:** P-P plot of the excursion waiting-time distribution of the IBM returns versus the geometric distribution (corresponding to the waiting time between successes in a Bernoulli sequence with probability 0.2 of success). The distributions would be nearly identical if the returns were exchangeable. **Panel B:** Trajectory of IBM prices from 9:45 to 3:45, sampled every five minutes, for a randomly selected day in 2005. **Panel C:** Same starting price as in Panel B, but with the twelve five-minute returns in each of the six one-hour intervals randomly and independently permuted. Since the returns within a given hour are exactly preserved, the stock prices in B and C are the same at 10:45, and at each hour thereafter. The dynamics governing the trajectories in B and C are not apparently different. **Panel D:** Distribution of KS distances to the geometric distribution, obtained from 5,000 surrogate return sequences corresponding to five thousand random permutations of the 18,648 IBM five-minute returns. Vertical line (in red) marks the KS distance (0.131) of the original sequence of returns. **Panel E:** Test for local exchangeability. Surrogates were produced by independently permuting every disjoint 20-minute block of four five-minute returns. The distribution of KS distances was again computed from 5,000 surrogates. In general, tests employing larger time intervals produce still lower p values. Thus, despite appearances, the evidence strongly points to a highly significant difference between the trajectories in B and C. **Panel F:** The ensemble of surrogates derived from permutations of pairs of returns, for every ten-minute block, are indistinguishable from the original sequence, with respect to the departure of their excursion waiting-time distributions from geometric.

3.2 Exploring time scale

Clearly we cannot treat the entire set of 18,648 IBM five-minute returns from 2005 as exchangeable (Panel D, Figure 7). In practice, traders adjust for changes in volatility, as measured by σ (the standard deviation of logarithmic returns); returns should only be considered exchangeable within a time period. But how often should volatility be updated? Are the returns, at least approximately, exchangeable within days, or perhaps within one-hour or one-half-hour intervals? In general, consider a partitioning of the index set $\{1, 2, \dots, 18648\}$ into disjoint intervals of length λ, where λ is a time span, measured in units of five minutes, over which the returns are presumed to be essentially exchangeable. We would use $\lambda = 74$ to test exchangeability within single days (recall that the first and last ten minutes of each day of prices are excluded), and $\lambda = 12$, 6, 4, and 2, respectively, to test exchangeability in one-hour, thirty-minute, twenty-minute, and ten-minute intervals. By virtue of equation (4), these hypotheses can be tested and exact p-values can be computed by generating ensembles of surrogate return sequences from ensembles of random permutations, and then comparing the corresponding values of the KS statistic to its observed value. For fixed λ, permutations are drawn iid from the uniform distribution on the set of permutations, Π, that preserve membership in the designated intervals.

Figure 7, Panels E and F, show the results of testing for local exchangeability of the excursion process in the five-minute IBM data, over twenty-minute ($\lambda = 4$, Panel E) and ten-minute ($\lambda = 2$, Panel F) intervals. Intervals longer than twenty minutes result in smaller p-values. Evidently, if time-varying volatility is the source of the breakdown in exchangeability, then it is operating at an extremely high frequency.

In line with the near-invariance of the waiting-time distribution, we find that other intervals, other stocks, and other years lead to similar results.

4 Time Scale and Stochastic Volatility Models

These observations of non-geometric waiting times and remarkably rapid changes in volatility suggest mechanisms for evaluating the validity of models of price and return dynamics. Which models and mechanisms are consistent with the observed properties of the excursion process? Stock dynamics are highly non-stationary, and stochastic volatility is a compelling modeling tool through which non-stationarity can be accommodated. We examined implied volatility, GARCH volatility models, and market-time transformations (trade and volume based) for their consistency with the invariance of excursion waiting-times and the empirical characteristics of local and global exchangeability. We were unable to match the data from any one of these points of view, as discussed in the following paragraphs.

4.1 Implied volatility

One place to look for a non-stationary volatility process that is commensurate with the breakdown of exchangeability is in the volatility implied by the pricing of options. Implied volatilities are forward looking and, as such, not a model for $\sigma \to \sigma_t$ in Black-Scholes. But the question here is not whether they reflect the actual minute-to-minute or hour-to-hour volatilities of their underlying stocks, but rather whether they include sufficiently rapid changes in amplitude to support the lack of global and even local exchangeability in the return process.

Eight days of minute-by-minute Citigroup 2008 stock and option prices were sampled from 9:35 AM until 3:55 PM (381 prices per day), and used to compute the minute-by-minute volatilities implied by the 2008 April 19 put with strike price 22.5 (left-hand panel, Figure 8). This sequence was used to produce a corresponding return process, from which an empirical excursion waiting-time distribution was extracted.[5] The volatility trajectory includes substantial fluctuations across multiple time scales, as is evident from the plot in Figure 8, and it would be reasonable to expect a failure of exchangeability in the derived return process. To the contrary, the waiting-time distribution was surprisingly similar to geometric (middle panel, KS=0.02), and in fact the return sequence was indistinguishable from globally exchangeable, based on the KS statistic and full-interval permutations (right-hand panel). Results for local exchangeability were similar. The experiment again makes the point that extreme high-frequency fluctuations in volatility might be needed to match the properties of the real excursion process in the context of a Black-Scholes model with time varying σ. Implied volatilities evidently do not take into account these strong intra-day volatility fluctuations.

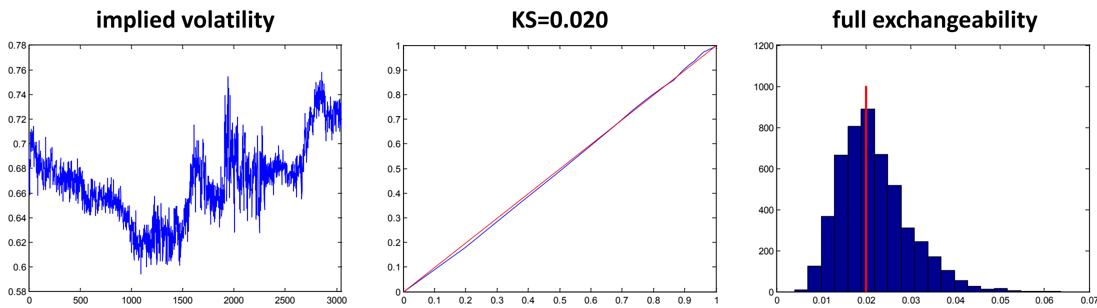


Figure 8: **Implied volatility generates exchangeable returns.** Eight days of minute-by-minute 2008 Citigroup stock and put prices (strike price 22.5, maturing on April 19, 2008) were used to calculate the minute-by-minute implied volatility, and to generate simulated minute-by-minute returns from a geometric Brownian motion with volatility function ($\sigma = \sigma_t$) equal to the implied volatility. **Left panel:** Minute-by-minute implied volatility. **Center panel:** The excursion waiting-time distribution of the simulated returns closely resembles the geometric distribution, unlike the real one-minute returns for which the P-P plot against the geometric is essentially identical to the one shown in the upper-left panel of Figure 7 (five-minute returns of IBM). **Right panel:** Simulated returns were not distinguishable from exchangeable returns through the KS statistic, despite substantial fluctuations in the implied volatility at multiple time scales.

4.2 GARCH

We examined the suitability of Engle's (1982) autoregressive conditional heteroskedasticity (ARCH) model and its generalization, GARCH (Bollerslev, 1986), for producing excursion processes that match the statistics of the excursions of real stock returns. We explored a collection of ARCH and GARCH models by fitting to the one-minute returns from the 2005 IBM stock prices. Over a wide range of values for the moving-average and auto-regressive orders (q and p, respectively), we found that GARCH(p, q) models provide a nearly perfect fit to empirical waiting-time distributions, but fail to match the invariance properties of these distributions across return intervals. We will show results for the particular model GARCH(10,10), but emphasize that virtually identical results were obtained for the more commonly used GARCH(1,1) model, as well as every combination of $1 \leq p \leq 10$ and $1 \leq q \leq 10$ that we tested. Given the ample amount of data (93,240 one-minute

[5] The scale of the volatility process is irrelevant, since the excursion process is invariant to multiplication of the returns.

returns), and given that for $1 \leq p, q \leq 10$ the GARCH(p,q) model is included in the GARCH(10,10) model, we chose to show the results for GARCH(10,10).

After fitting the GARCH parameters (see Table 3 for estimated parameters and their standard errors), the model was used to produce a full year of simulated one-minute returns. The excursion waiting-time distribution of the simulated data matches the distribution extracted from the real data, as indicated by the P-P plot in the upper-left panel of Figure 9, and the small KS distance. Furthermore, as with the real data, and in contrast to experiments with implied volatility (§4.1), GARCH simulated returns are not exchangeable, even under permutations confined to two-minute intervals – see upper-right, lower-left, and lower-right panels, respectively, for results on full exchangeability, and four-minute and two-minute exchangeability. In general, the match between simulated and actual returns was excellent.

Table 3: **GARCH parameter estimation.** The GARCH(10,10) model ($\sigma_t^2 = \omega + \sum_{i=1}^{10} \alpha_i R_{t-i}^2 + \sum_{i=1}^{10} \beta_i \sigma_{t-i}^2$) was estimated from 93,240 2005 one-minute returns of IBM stock ($R_t,\ t = 1, 2, \ldots, 93240$), using the UCSD Garch Matlab toolbox. The table shows the estimated values and standard errors of the twenty-one parameters. (Zero values are common due to stability and positivity constraints.)

	estimated value	standard error		estimated value	standard error
ω	1.5111	0.0000			
α_1	0.1445	0.0044	β_1	0.0646	0.3472
α_2	0.0758	0.0508	β_2	0	0.3514
α_3	0.0427	0.0404	β_3	0	0.3200
α_4	0.0368	0.0406	β_4	0.2385	0.2767
α_5	0	0.0285	β_5	0	0.2305
α_6	0	0.0283	β_6	0	0.2140
α_7	0	0.0260	β_7	0.1892	0.1739
α_8	0	0.0179	β_8	0	0.1994
α_9	0	0.0178	β_9	0	0.1862
α_{10}	0	0.0163	β_{10}	0.1613	0.1038

On the other hand, real stocks produce excursion waiting times that are nearly scale invariant, as already documented in §2, and illustrated in Figure 3 for the 2005 IBM data. For comparison, the left-hand panel of Figure 10 reproduces the bottom middle panel of Figure 3, whereas the right-hand panel shows the corresponding P-P plot for the GARCH simulated data. The KS distance between one-minute and five-minute waiting-time distributions for the IBM data is 0.005, whereas the GARCH generated one-minute returns, aggregated to produce five-minute returns, produce a KS distance of 0.05. In general, GARCH models have poor scaling properties, as already noted in the discussion of intra-day return intervals in §4 of Andersen & Bollerslev, 1997. In fact, GARCH models, though elegant and apparently suitable for fitting volatility, are inconsistent in the sense that in general a process can not obey a GARCH model for both one-minute and k-minute returns, for any $k = 2, 3, \ldots$, as is easily demonstrated analytically.

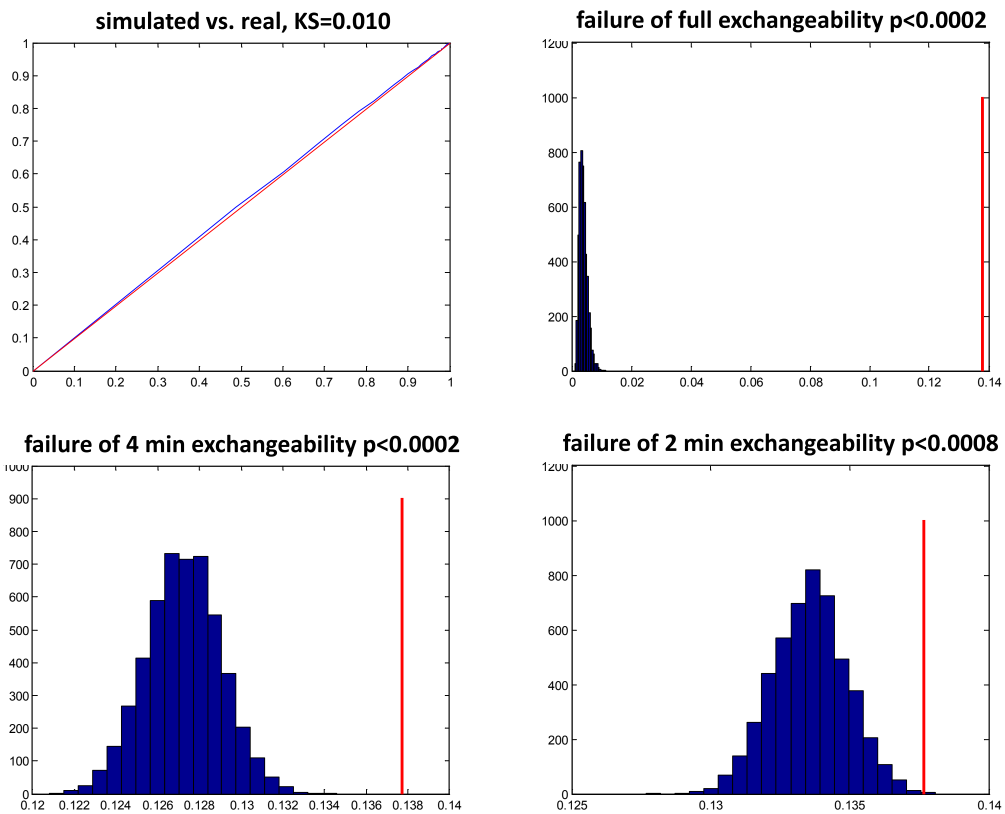


Figure 9: **Simulated one-minute IBM returns using GARCH.** One-minute returns on IBM for all of 2005 were used to fit a GARCH model, with autoregressive and moving average terms each of order 10 ($p = q = 10$). Waiting-time distribution between excursions in the simulated returns was a near-perfect match to the empirical distribution (upper-left panel). Results of permutation tests for global exchangeability (upper-right panel) and local exchangeability (four-minute intervals, lower-left panel, and two-minute intervals, lower-right panel) were essentially identical to the results for the real returns (not shown).

4.3 Market time

There is no reason to believe that a good model for the logarithm of stock prices should be homogeneous in time. To the contrary, the random walk model suggests that the variance of a return should depend on the number or volume of transactions (the number of "steps") rather than the number of seconds. The compelling idea that "market time" is measured by accumulated activity rather than the time on the clock seems to have been suggested first by Mandelbrot and Taylor (1967) and then worked through, more formally, by Clark (1973). It has been re-visited in several influential papers since then; see the discussions by H. Geman (2005) and Shephard (2005) for reviews and references.

Here we employ a simple yet definitive test that rules out the possibility that *any* function of volume or number of transactions can render the return process compatible with a geometric Brownian motion, or for that matter any of its Lévy generalizations. In particular, time changes based on volume or trade numbers do not transform returns into exchangeable sequences. The key, then, to ruling out these simple market-time transformations lies in the *dynamics*; it is not enough to simply match the marginal distributions of the returns, as we now demonstrate.

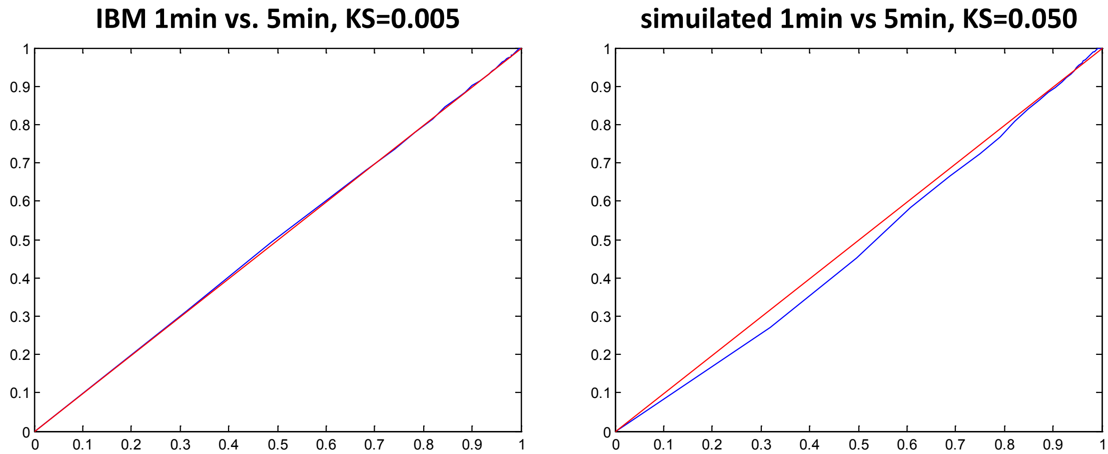


Figure 10: **Failure of GARCH model to match waiting-time scale invariance of real returns. Left panel:** P-P plot matching excursion waiting-time distributions for the IBM one-minute to the IBM five-minute returns (2005). Distributions are nearly identical. **Right panel:** Same comparison, using GARCH-generated one-minute returns, aggregated to make a record of five-minute returns. Although there is an excellent fit to the one-minute data (see Figure 9), the model fails to scale across different return intervals.

Formally, let $D(t) \doteq \log S(t)$, and start with the customary model $D(t) = \mu t + \sigma w(t)$, where w is a standard Brownian motion or a more general process with stationary and independent increments (i.e. a Lévy process). Volatility clustering is inconsistent with the resulting stationarity and/or independence of the increments of D (and hence the modeled returns). One remedy is to introduce a volatility process, $\sigma \to \sigma(t)$, as in the well-known models of Hull and White (1987) and Heston (1993), or any of a variety of other models for stochastic volatility (cf. Shephard, 2005). Another remedy is to introduce a market-time process $\tau(t)$, usually independent of w, and write $w(\tau(t))$ in place of $w(t)$. (Actually, the two models are oftentimes equivalent—see, e.g. H. Geman et al., 2001, Veraart and Winkel, 2010.) Depending on the details of the model for S and for τ, $D(t)$ becomes $\mu t + \sigma w(\tau(t))$ or $\mu\tau(t) + \sigma w(\tau(t))$.

Assuming that τ is independent of w, Clark (1973) experimented with various functions of the volume as measures of market time:

$$\tau(t) - \tau(s) = f(V(t) - V(s)) \quad \forall s \leq t \tag{5}$$

where $V(t)$ is accumulated volume and f is monotone increasing. More recently Easley et al. (2012) provided support for equation (5) by demonstrating "partial recovery of Normality" using equal-volume returns. On the other hand, Ané and H. Geman (2000) have argued that the number of trades as opposed to the accumulated volume, is the fundamental determinant of τ (hence $f(T(t) - T(s))$ in (5), where $T(t)$ is accumulated number of transactions). Mandelbrot and Taylor (1967) raised both possibilities.

The typical test shows that the normal distribution is a better approximation of the distribution of returns when returns are defined by equal intervals of τ rather than equal intervals of "clock time." But this is a weak test. The marginal distribution of a process caries no information about its temporal statistics. Dynamics are more important, but not as easily explored. The excursion waiting-time distribution is fundamentally about dynamics, and provides an easy and sensitive test of whether a time-transformed price process is, even approximately, a geometric Lévy process (e.g. geometric Brownian motion).

Whether volume-based (viz. equation 5) or trade based ($V(t) \to T(t)$), let $0 < t_1 < t_2 < \dots$ be an increasing sequence yielding equal increments of τ: $\tau(t_k) - \tau(t_{k-1}) = \tau(t_l) - \tau(t_{l-1})$, $\forall k \neq l$, $k, l > 0$. If $D(t) = \mu\tau(t) + \sigma w(\tau(t))$ then set $R_k = D(\tau_k) - D(\tau_{k-1})$, and otherwise, if $D(t) = \mu t + \sigma w(\tau(t))$, set $R_k = D(\tau_k) - D(\tau_{k-1}) - \mu(t_k - t_{k-1})$. (The difference is negligible for short intervals.) For either model of D and either model of τ (volume-based or trade-based), if the market-time corrected process is geometric Brownian motion (or more generally Lévy), then the return sequence $R_1, R_2, \dots$ constructed in this manner is necessarily iid and therefore exchangeable.

Consider for example Figure 11, where we examine equal-market-time returns on IBM 2005 stock, under the assumption that τ is determined by the number of trades. In particular, returns were defined on successive intervals containing 110 trades each (corresponding, on average, to five minutes of clock time). Thus

$$R_k = D(\tau_k) - D(\tau_{k-1}) = \log S(t_{\tau_k}) - \log S(t_{\tau_{k-1}})$$

where t_{τ_k} is the time when the τ_k-th trade occurred and $\tau_k = 110 \cdot k$ for all $k = 0, 1, 2, \dots$. Obviously, the process $R_1, R_2, \dots$ is far from exchangeable (right-hand panel) and the waiting-time distribution is a poor approximation of the geometric distribution (left-hand panel). We examined all combinations of models for D and τ (volume-based and trade-based). Each case produces a figure essentially identical to Figure 11; these market-time transformations fail to render the returns exchangeable.

By the evidence, neither the number of trades nor the accumulated volume is, in and of itself, a viable measure of market time. The dynamics of the return process, following a volume or trade-based time change, do not resemble those of a geometric Brownian motion or any other Lévy process.

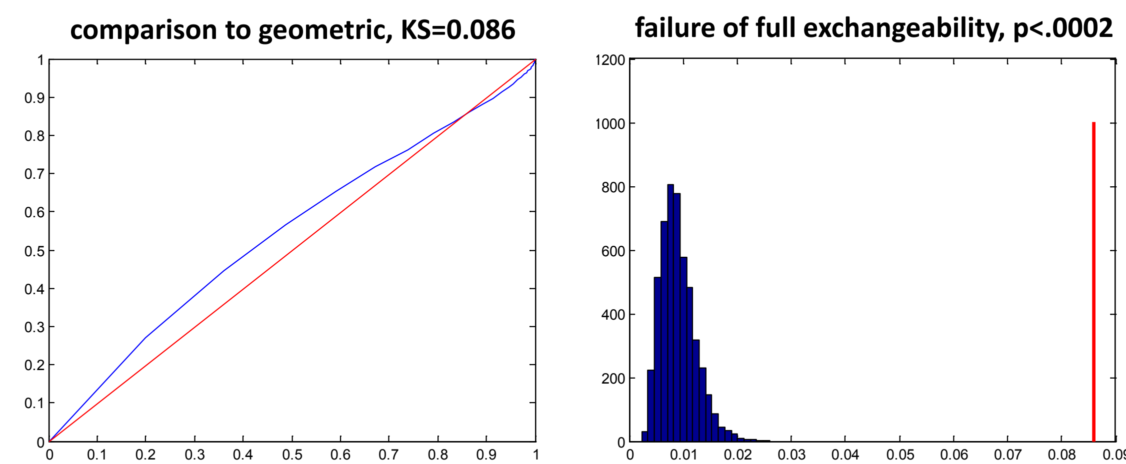


Figure 11: **Interval time measured by number of trades.** In 2005, there was an average of about 110 trades of IBM stock every five minutes. If "market time" were measured by the number of trades, and were adequate to transform prices into a Lévy process, then returns over 110-trade intervals would be exchangeable. **Left panel:** Excursion waiting-time distribution for equal market-time intervals (110 trades) does not match the geometric distribution. **Right panel:** Equal-market-time returns are not exchangeable, as evidenced by the distribution of KS values under permutations. Market time measured by volume instead of trades also fails to render returns exchangeable.

5 Summary and concluding remarks

We have given empirical evidence for a new invariant in the price movements of stocks. The waiting-time distribution between large returns ("excursions") is nearly invariant to scale (length of the return interval), stock, and the year of observation. The clustering of excursions is a manifestation of the well-studied clustering of volatility. The invariance in the clustering of excursions therefore constrains proposed models and mechanisms for volatility clustering. Self-similar (log) price processes have invariant waiting times between excursions, but the evidence for self-similarity is confined to the distributions on log price increments, and not the processes themselves. Furthermore, scaling indices estimated from return data vary from study to study (Bouchaud, 2001), and are extremely sensitive to statistical methodology, as might be expected given that most approaches focus on the tail behavior of the return distributions. By contrast, waiting-time distributions rely on percentiles, which are robust and non-parametric, and evidently stable given the weight of evidence for invariance presented in §2.

We have illustrated the possible utility of excursion waiting times by examining some models for price and volatility dynamics. In general, the failure of even local exchangeability of excursions (and therefore returns) points to rapid changes in volatility. Thus implied volatility, for example, is much too smooth (despite its appearance—see Figure 8). ARCH and GARCH models, even of low order, track volatility sufficiently well to produce simulated returns with excursion waiting times that are a near-perfect match to empirical waiting times. But unlike real returns, aggregating the simulated one-minute returns into simulated five-minute returns produces a different waiting-time distribution. This might have been anticipated (though not guaranteed) by the observation that these models themselves lack scale invariance. Finally, we examined the appealing idea of a market-activity based time change in an effort to remove volatility clustering and restore exchangeability to the random-walk model. Returns were re-defined with respect to equal increments of market time, as opposed to clock time, under both volume-based and trade-based measures of market activity. Neither definition of market time rendered an exchangeable sequence of excursions.

The usual caution about the distinction between statistical significance and scientific significance bears repeating here. We have introduced exact hypothesis tests that produce very small p-values. In and of themselves, these values are not particularly interesting given the large sample sizes involved (e.g. almost 20,000 five-minute returns on IBM stock from 2005). Our focus, instead, was on the trajectory of p-values under a sequence of global-to-local exchangeability tests, and on the comparison of p-values between data produced by real returns and data simulated from models.

A more subtle statistical issue concerns the use of aggregated data for inference about temporal dynamics, especially scaling properties, such as self-similarity. Consider using a year's worth of price data $(S(t), t \in [0, T])$ for estimating the joint distribution on successive returns $R_1^{(\delta)}, \ldots, R_n^{(\delta)}$ over intervals of length δ, where

$$R_k^{(\delta)} = \log \frac{S(\delta k)}{S(\delta(k-1))}$$

(Typically $n = 1$ and the goal is to study the distribution on returns and its relationship to δ.) To keep things simple, assume that

$$\log S(t) = \int_0^t \sigma(s) dw(s)$$

where w is an alpha-stable Lévy process ($\alpha = 2$ when w is Brownian motion), consistent with the basic geometric random-walk framework, but accommodating non-constant volatility.[6] The alpha-

[6]The other general approach to time-varying volatility is through a time change of w: $w(t) \to w(\tau(t))$. As

stable Lévy processes are self-similar, with scaling exponent $\alpha \in (0, 2]$ (i.e. Hurst index $H = 1/\alpha \in [.5, \infty)$). But given the year under study, with its particular sample path of $\sigma(t)$, $t \in [0, T]$, $\log S(t)$ is *not* self-similar: $\mathcal{L}\{\log S(\delta t)\} \neq \mathcal{L}\{\delta^{1/\alpha} \log S(t)\}$.[7] Nevertheless, an experimental study, such as Mandelbrot (1963), Muller et al. (1990), or Mantegna and Stanley (2000), to name just a few, might well lead to the opposite conclusion, as follows.

Assume for the time being that $\sigma(t)$ is independent of $w(t)$, and path-wise smooth enough to have negligible fluctuations in intervals of length $n\delta$, which is reasonable for all δ sufficiently small. What properties should be expected of the empirical joint distribution, $\hat{F}$, on $R_1^{(\delta)}, \ldots, R_n^{(\delta)}$:

$$\hat{F}_{R_1^{(\delta)},\ldots,R_n^{(\delta)}}(r_1, \ldots, r_n) = \frac{\delta}{T} \sum_{k=1}^{T/\delta} \prod_{i=1}^{n} 1_{R_{k+i}^{(\delta)} < r_i}$$

derived from a year of returns? In particular, which, if any, of the scaling properties of w are inherited by the empirical return distribution? Under the smoothness assumption on σ, a straightforward calculation shows that

$$\hat{F}_{R_1^{(\delta)},\ldots,R_n^{(\delta)}}(r_1, \ldots, r_n) \approx \hat{F}_{R_1^{(1)},\ldots,R_n^{(1)}}(\delta^{-1/\alpha}(r_1, \ldots, r_n)) \tag{6}$$

which is in fact the property that characterizes the increments of a self-similar process, with scaling index α, such as the increments of w itself. The fact that $\sigma = \sigma(t)$ is lost in the aggregation. *The returns $R_1^{(\delta)}, \ldots, R_n^{(\delta)}$ appear to come from a self-similar process even though they do not.*

The implicit assumption behind aggregation is stationarity. In its absence, the aggregated estimator is a mixture of distributions, each generated by w, but mixed with respect to the occupation measure of $\sigma(t)$[8] over the year-long observation $t \in [0, T]$. Chang and Geman (2013) demonstrated that the convergence is quite rapid and the approximation in (6) typically holds even when the return interval, δ, is large relative to the fluctuations of σ. What does the same reasoning say about the empirical waiting-time distribution for excursions, as computed over the same time interval? This is a substantially harder calculation, but in one regard the conclusion is likely to be the same: if we accept the geometric random-walk model then scale invariance of the *empirical* waiting-time distribution for all δ sufficiently small is a foregone conclusion. On the other hand, the particular invariant distribution, including for example the empirical probability of zero wait between excursions (approximately 0.32), very much depends on the particular occupation measure of σ.

In light of these observations, empirical scale invariance in the timing of excursions and for self-similarity of the price process is at least consistent with the geometric random-walk model, if not in fact further support for its basic soundness, whether or not the volatility process is stationary. What is more, the near-invariance of the excursion waiting-time distribution across stocks and years points to a volatility-generating process with occupation measure that is surprisingly reproducible, modulo a constant scale. Notice that if non-constant market activity were the source of stochastic volatility, then its strong correlations across stocks would begin to explain invariance of waiting times across stocks. Notice also that most days begin and end with relatively high activity, a daily rhythm which might contribute to the invariance from one era to another.

In light of the results in §4.3, however, we would need to look beyond any simple function of trades or volume for the relevant measure of market activity (and hence market time). It might be

mentioned in §4.3, in most models the two approaches, $\sigma dw(t) \to \sigma(t)dw(t)$ and $w(t) \to w(\tau(t))$, come down to the same thing. For more on conditions of equivalence see Veraart & Winkel (2010).

[7] $\log S(t) = \int_0^t \sigma(s)dw(s)$ but $\mathcal{L}\{\log S(\delta t)\} = \mathcal{L}\{\delta^{1/\alpha} \int_0^t \sigma(\delta s)dw(s)\} \neq \mathcal{L}\{\delta^{1/\alpha} \log S(t)\}$

[8] The distribution of the random variable $\sigma(X)$ when X is uniform on $[0, T]$.

sensible, for example, to view trades as indicating the time of a step in the random walk, and volume as determining the scale of the distribution on the step size (relating to the ideas of Gabaix et al., 2003). There is no reason to believe that the relationship between the volume v of a trade and the scale $\sigma = \sigma(v)$ of the resulting random step would be linear (though presumably it is monotonic). To the contrary, it would depend on the complexities of supply and demand as might be reflected in the state and dynamics of the collective order book. In any case, it might be feasible to estimate $\sigma(v)$, non-parametrically, by maximum likelihood. The test of the model would then be the same: are returns over equal market-time intervals exchangeable?

Acknowledgments. The authors gratefully acknowledge insightful discussions with Hélyette Geman and Matthew Harrison, as well as financial support from the Office of Naval Research under contract N000141010933, and the National Science Foundation under grants ITR-0427223, DMS-1007593 and DMS-1007219, and the Defense Advanced Research Projects Agency under contract FA8650-11-1-7151, and Center of Mathematical Modeling & Scientific Computing, and the National Center for Theoretical Science, Hsinchu, Taiwan, and the National Science Council under grant 100-2115-M-009-007-MY2.

References

[1] T. G. Andersen and T. Bollerslev. Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, 4:115–158, 1997.

[2] T.G. Andersen. Return volatility and trading volume: An information flow interpretation of stochastic volatility. *Journal of Finance*, 51:169–204, 1996.

[3] T. Ané and H. Geman. Order flow, transaction clock, and normality of asset returns. *Journal of Finance*, 55(5):2259–2284, 2000.

[4] T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327, 1986.

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

[6] L. Calvet and A. Fisher. Multifractality in asset returns: theory and evidence. *The Review of Economics and Statistics*, 84:381–406, 2002.

[7] L.-B. Chang and S. Geman. Empirical scaling laws and the aggregation of non-stationary data. *Physica A*, to appear, 2013.

[8] L.-B. Chang, A. Goswami, F. Hsieh, and C.-R. Hwang. An invariance for the large-sample empirical distribution of waiting time between successive extremes. *Bulletin of the Institute of Mathematics, Academia Sinica*, 8:31–48, 2013.

[9] P.K. Clark. A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, 41(1):135–155, 1973.

[10] P. Diaconis and D. Freedman. Finite exchangeable sequences. *Annals of Probability*, 8(4):745–764, 1980.

[11] D. Easley, M.M. López de Prado, and M. O'Hara. The volume clock: insights into the high frequency paradigm. *Journal of Portfolio Management*, 39:19–29, 2012.

[12] P. Embrechts and M. Maejima. *Selfsimilar Processes*. Princeton University Press, Princeton, NJ, 2002.

[13] R.F. Engle. Autoregressive conditional heteroscedasticity with estimates of variance of united kingdom inflation. *Econometrica*, 50:987–1008, 1982.

[14] X. Gabaix, P. Gopikrishnan, V. Plerou, and H.E. Stanley. A theory of power-law distributions in financial market fluctuations. *Nature*, 423:267–270, 2003.

[15] H. Geman. From measure changes to time changes in asset pricing. *Journal of Banking & Finance*, 29:2701–2722, 2005.

[16] H. Geman, D.B. Madan, and M. Yor. Time changes for Lévy processes. *Mathematical Finance*, 11(1):79–96, 2001.

[17] R. Gencay, F. Selcuk, and B. Whitcher. Scaling properties of foreign exchange volatility. *Physica A*, 289:249–266, 2001.

[18] J.B. Glattfelder, A. Dupuis, and R.B. Olsen. Patterns in high-frequency fx data: discovery of 12 empirical scaling laws. *Quantitative Finance*, 11:599–614, 2011.

[19] S.L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2):327–343, 1993.

[20] C.C. Heyde. A risky asset model with strong dependence through fractal activity time. *J. Appl. Prob.*, 36:1234–1239, 1999.

[21] F. Hsieh, S.-C. Chen, and C.-R. Hwang. Discovering stock dynamics through multidimensional volatility-phases. *Quantitative Finance*, 12:213–230, 2012.

[22] J. Hull and A. White. The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42:281–300, 1987.

[23] B. Mandelbrot. The variation of certain speculative prices. *Journal of Business*, 36:394–419, 1963.

[24] B. Mandelbrot and H. Taylor. On the distribution of stock price differences. *Operations Research*, 15:1057–1062, 1967.

[25] R.N. Mantegna and H.E. Stanley. Scaling behaviour in the dynamics of an economic index. *Nature*, 376:46–49, 1995.

[26] R.N. Mantegna and H.E. Stanley. *An Introduction to econophysics*. Cambridge University Press, 2000.

[27] T.D. Matteo. Multi-scaling in finance. *Quantitative Finance*, 7:21–36, 2007.

[28] U.A. Muller, M.M. Dacorogna, R.B. Olsen, O.V. Pictet, M. Schwarz, and C. Morgenegg. Statistical study of foreign exchange rates, empirical evidence of a price change scaling law, and intraday analysis. *Journal of Banking and Finance*, 14:1189–1208, 1990.

[29] N. Shephard. *Stochastic Volatility*. Oxford University Press, Oxford, 2005.

[30] A.E. Veraart and M. Winkel. Time change. In Rama Cont, editor, *Encyclopedia of Quantitative Finance*. Wiley Online Library, 2010.

[31] B.H. Wang and P.M. Hui. The distribution and scaling of fluctuations for hang seng index in hong kong stock market. *The European Physical Journal B*, 20:573–579, 2001.

[32] F. Wang, K. Yamasaki, S. Havlin, and H.E. Stanley. Scaling and memory of intraday volatility return intervals in stock markets. *Physical Review E*, 73:026117–1–023117–8, 2006.

[33] Z. Xu and R. Gencay. Scaling, self-similarity and multifractality in fx markets. *Physica A*, 323:578–590, 2003.