Is there a relationship between the time scaling property of asset returns and the outliers? Evidence from international financial markets

González-Sánchez, Mariano*

Business and Accounting Department, School of Economics and Business,
Universidad Nacional de Educación a Distancia (UNED),
Paseo Senda del Reay, 11. 28040 Madrid (Spain)

Abstract

Stylized facts are statistical properties present in high frequency returns of financial assets. While some of them supposes that returns are not Gaussian, another, called time scaling, involves that decreasing the frequency of observation, the returns converge to normal distribution. This paper find evidence that the existence of scaling and outliers entails other stylized facts. Also, a methodology for identifying outliers is proposed and applied to both simulated series and 1,300 market assets. Results indicate that all market returns have time scaling (between 2 and 28 days) and, in 95% of cases, daily outliers represent less than 6% of observations.

Keywords: stylized facts; time scaling; outlier; heteroskedasticity; leptokurtosis.

JEL: C12; C13; G10; G17

1. Introduction and Background

The term *stylized facts* was coined to describe the characteristics of asset returns (Cont (2001)). In this way, all empirical work, regardless of its final objective, previously needs to consider these facts so that their estimates are consistent, making it is increasingly common to find complex methods and models within financial econometrics. Within these properties there are four characteristics which give researchers and investors the most trouble: asymmetry, leptokurtosis, autocorrelation and heteroskedasticity. But, while these properties suppose that asset returns probability distribution is not Gaussian, another fact known as scaling shows that as the term for estimating returns increases (or the frequency of observation decreases) the behavior of the theses is closer to a normal distribution. So, normality vs. non-normality dichotomy of the time series is a relevant question for empirical works in finance.

The financial literature has analyzed the effects of this time scaling property, for example Antypas et al. (2013) analyze two apparently contradictory empirical regularities of financial returns, namely, the fact that the empirical distribution of returns tends to normality as the frequency of observation decreases (aggregational Gaussianity or time scaling) and found evidence that aggregational Gaussianity and infinite variance can coexist. On the contrary, to our knowledge, there are no empirical studies that search what causes such stylized fact. In this way, this study would be a bridge between the financial and econometric literatures, in order to test whether time scaling is caused by outliers, which could lead to mistakes in the acceptance and rejection of the hypotheses about the statistical behavior of the time series. Then, a

*Corresponding author:
Email address: mariano.gonzalez@cee.uned.es (González-Sánchez, Mariano)

question arises: Why do the properties of a high frequency financial time series (daily, for example) disappear (consequence of scaling) for lower frequency (monthly, for example)? Our aim is to find an empirical answer to this question and its effects

Econometric literature, concerning the robustness of the usual statistical tests in presence of outliers, find like some test results which accept a structural break or non-linear relationships, now should be considered outliers, and that this has major implications regarding the validity of the tests and the consequent econometric models adjusted to the time series (see Koop & Potter (2000)). So, detecting outliers constitutes an entire field of statistical analysis (Barnett & Lewis (1978) and Lucas et al. (2002)), even reaches to financial markets, for example, Chiang et al. (2016) observe that if stocks prices embed some outliers, due to the impact of unusual financial and non-financial events then, it involves that the accuracy of parameter estimates and volatility forecasting (family GARCH models) cannot be guaranteed in the presence of these extreme observations.

The statistical disadvantages of the outliers are very diverse. Peña (1990) analyzes the sensitivity of ARIMA parameters to the presence of outliers and measures their influence on the Mahalanobis distance. About autorregresive process, Ahmad & Donayre (2016) find that tests can be severely distorted in the presence of large outlier observations. Moreover, when AR processes are more persistent, the distortion increases, especially with large samples. Baragona et al. (2016) find outlier effects on AR processes and propose an outlier detection test based on the empirical likelihood method. From van Dijk et al. (1999), Franses & Ghijsels (1999) and Franses et al. (2004), among others, we know that outliers may wrongly reject the null hypothesis on homoskedasticity. Besides, in the presence of outliers, Sakata & White (1998) observe a loss of precision in the estimate of parameters and biases of the parameter related to the persistence of volatility. As consequence of this, the GARCH model tends to provide poor volatility forecasts in the presence of additive outliers; for that, Park (2002) proposes a robust GARCH model estimates by least absolute deviation with superior forecasting out-of-sample than the usual GARCH models. However, the least absolute deviation estimate gives less weight to outliers and then, part of information about the variable behavior may be lost. Carnero et al. (2006) show that outliers in uncorrelated stationary series bias the sample square autocorrelation test and, also analyze the robust test for conditional heteroscedasticity proposed by van Dijk et al. (1999) and, find, for large samples, its size is distorted. Another proposal addressing the problems of outliers are Alih & Ong (1996), which propose a variation of the Goldfeld-Quandt test on heteroskedasticity to avoid false nulls in the presence of outliers.

Finally, Charles (2008) analyzes the effects of outliers on several econometric tests, which can lead to a poor specification of the model. This empirical work finds that the volatility forecast is improved (using Diebold & Mariano (1995) test) when the data are cleaned of outliers and indicates that although there are robust methods to estimate with outliers, they have several disadvantages: (i) sometimes the performance is inadequate or poor; (ii) since outlying observations are not adjusted, the outliers continue to impact forecasts; (iii) in most cases, only limited information about the outlier can be obtained (e.g. from the weights applied to the residuals).

In short, the stylized facts are statistical properties attributed to the asset returns and, as consequence, complex econometric models are required to model them. But when the estimate frequency of these returns becomes lower, some of these facts disappear. On the other hand, the econometric and statistical literature has shown that the existence of outliers entails unbiased estimates of the models parameters and the false positives in the hypothesis about the existence of heteroskedasticity. So, this paper aims is an empirical answer about the relationship between scaling and outliers in high frequency asset returns. Our proposal replaces the highest return, in absolute value, until the resulting sample overcomes the normality test, non-autocorrelation for raw series and squared data, and absence heteroskedasticity. So, on a sample of daily returns of 1,330 assets (with around 1,250 daily observations per asset) that include stocks, exchange rates and commodities, we found that replacing a few observations (between 1% and 6% or only 12 and 75 observations) the resulting sample is adjusted to a normal distribution. Additionally, we obtain a division of the information contained in the original data allows us to separately analyze the behavior random and the extraordinary or atypical (outliers or jumps) with objective to measure systematic and idiosyncratic risks.

The rest of paper is organized as follows: Section 2 shows a new approach to this question. Section 3 analysis the results for simulated series and financial market time series. Section 4 contains the studys main conclusions.

2. Methodology

The statistical definition of outlier is *a case that does not follow the same model as the rest of the data* (Weisberg (1985)). Since the seminal work on testing outliers (Fox (1972)), two types of outliers are usually studied: the additive outlier, which only affects a single observation, and the innovative outlier that affects several observations. As Chen & Liu (2011) point out, from a computational standpoint, the strategy of detecting outliers one by one may be the only feasible approach to dealing with multiple outliers. There are different methods to detect outliers, some graphic (for example, box-plot) and other empirical. Aguinis et al. (2013) review 14 outlier definitions, 39 outlier identification techniques and 20 different ways of handling outliers. Among the most common tests are Grubbs (1950), Dean & Dixon (1951), Hampel (1974), Tukey (1977) and Thompson (2006), although all of them suffer from the same problem: (i) they do not determine the number of outliers that exist in a series a priori and; (ii) do not guarantee that the clean series of outliers is Gaussian, not autocorrelated and without heterokedasticity and therefore, the resulting time series maintains complex statistical characteristics to adjust. That they are simple to estimate is their main advantage since they use mean, standard deviation, median or interquartile distance to describe the performance of a time series, but the drawback of selecting an adequate value to the fence persists.

Another problem with outlier detection is imputation data. Usually outliers are replaced with zero, the mean, median, percentiles or even random values (Tabachnick & Fidell (2007)). More elaborate techniques (regressions and others) are used with multiple imputations (Elliott & Stettler (2007) and Dang & Serfling (2011)), Liu et al. (2004) use Kalman filters for the AR model and Weekley et al. (2010) employ an ARMA model. However, as seen in the literature reviewed above, outliers give false positives for autocorrelation and heteroscedasticity, so that using econometric models to replace extreme values would only increase the problems.

Our proposal solves problem-(ii) of usual outliers tests since our clean time series is *i.i.d.* Gaussain, ie, it shows neither autocorrelation nor heterocedasticity. For that, we define t as a temporary moment of observation from a sample size T, where p is the asset log-price and r is asset return estimated for frequency j as $r_{t,j} = p_t - p_{t-j}$. Note that any frequency return higher than 1 ($T > j > 1$) can be expressed as the sum of the daily returns ($r_{t,j} = \sum_{i=0}^{j-1} r_{t+j-i,1}$). This assumes: (i) the return sum is non-overlapping since it is from 0 to $j-1$ and, (ii) it shows the re-scaling of the sample as a sum. Then, we enunciate the stylized fact of scaling as:

$$\begin{aligned} r_{t,1} &\sim \mathcal{F} \neq \mathcal{N}(\mu_1, \sigma_1^2) \\ r_{t,j} &\sim \mathcal{N}(\mu_j, \sigma_j^2) \end{aligned} \tag{1}$$

Where $\mathcal{F}$ is any distribution different from normal distribution ($\mathcal{N}$). Note 1 is a result of the Central Limit Theorem (CLT), and then if there exists a frequency j that satisfies 1 for the asset return, then there is a limited number of outliers (ω) in original frequency (usually daily), such that if these outliers are eliminated the original series is also Gaussian ($r_1^* \sim \mathcal{N}(\mu_1^*, \sigma_1^{*2})$). Therefore, we express original time series as:

$$\begin{aligned} r_{t,1} &= r_{t,1}^* + \omega_t^+ + \omega_t^- \\ \omega_t^{+/-} &= \begin{cases} r_{t,1} \pm rp & \text{if } u_t^{+/-} \leq \lambda^{+/-} \quad \text{where } u_t^{+/-} \sim \mathcal{U}(0,1) \\ 0 & \text{otherwise} \end{cases} \\ r_{t,1}^* &= \begin{cases} r_{t,1} & \text{if } u_t^{+/-} > \lambda^{+/-} \quad \text{where } u_t^{+/-} \sim \mathcal{U}(0,1) \\ rp & \text{otherwise} \end{cases} \end{aligned} \tag{2}$$

Where $\mathcal{U}$ is uniform distribution, λ is the frequency of positive or negative jumps [1] or outliers and rp is the replacement value when an outliers is identified, which is usually zero, mean, median or a percentile. As $r_{t,1}^*$, ω_t^+ and ω_t^- are independent then $\mathrm{E}(r_{t,1}) = \left[(1 - \lambda^+ - \lambda^-) \cdot \mathrm{E}(r_{t,1}^*) + \lambda^+ \cdot \mathrm{E}(\omega_t^+) + \lambda^- \cdot \mathrm{E}(\omega_t^-)\right]$ and, $\sigma^2(r_{t,1}) = \left[(1 - \lambda^+ - \lambda^-)^2 \cdot \sigma^2(r_{t,1}^*) + (\lambda^+)^2 \cdot \sigma_+^2 + (\lambda^-)^2 \cdot \sigma_-^2\right]$.

So, the methodology proposed to adjust 2 is:

1. We select τ_n, τ_a, $\tau_{a,2}$ and τ_h as the tests of normality, autocorrelation on raw and square of the data and heteroskedasticity, respectively.
2. For $j = 1$ to $j = J < T$, we calculate $r_{t,j}$ and $\tau_{n,j}$, $\tau_{a,j}$, $\tau_{a,2,j}$ and $\tau_{h,j}$. If for any j, all of these tests show $p-values$ higher than confidence level α (for example, 0.05) then the scaling property is true and there is an outliers set. Otherwise, the usual econometric AR-GARCH is the most efficient estimate method.
3. Next, when scaling characteristic is observed, we identify outliers as:
   a) For $t = 1$ to T, we search $max(|r_{t,j}|)$.
   b) If this data is positive, we do $\omega_t^+ = r_{t,1} - rp$, but if it is negative then $\omega_t^- = r_{t,1} + rp$.
   c) For both cases, we replace the original data: $r_{t,1} = rp$.

[1] Note that this approach is an econometric version (discrete time) of Merton (1976) jump-diffusion model (continuous time).

d) Finally, after replacement, we define a new variable r_1^* and we estimate τ_n, τ_a, $\tau_{a,2}$ and τ_h on this time series. Then:

i) If we accepted the hypotheses on normality, non-autocorrelation and non-heteroskedasticity, we estimate the parameters of 2 as: $\hat{\mu_1^*} = \frac{1}{T}\sum_{t=1}^{T} r_{t,1}^*$, $\hat{\sigma_1^{*2}} = \frac{1}{T-1}\sum_{t=1}^{T}(r_{t,1}^*)^2 - (\hat{\mu_1^*})^2$ and $\hat{\lambda}^{+/-} = \frac{n^{+/-}}{T}$. Where $n^{+/-}$ is the number of observations in the set of positive or negative outliers.

ii) Otherwise, we go to back step 3a.

3. Empirical analysis and results

First, a study of the effects of the proposal on simulated series is carried out, in order to know how it behaves. Subsequently, it is applied to data of international financial markets, including stocks, commodities and exchange rates.

3.1. Experimental exercise

According to Berry-Essen Theorem (Berry (1941) and Esseen (1942)), the convergence in CLT depends on moments higher than 2, so that, we use the Jarque-Bera test to compare the normality of a time series, since moments of order 3 and 4 are considered. We also apply the usual tests to detect autocorrelation in raw and square data (Ljung & Box (1978)) and heteroskedasticity (LM-ARCH of Engle (1982)).

First, we check outliers effect on stationary tests. To do so, 10,000 time series of 2,000 observations each are simulated as follows:

$$
\begin{aligned}
&P_t = P_{t-1} + \kappa \cdot \epsilon_t \quad \epsilon_t \sim \mathcal{N}(0,1) \\
&\kappa = \begin{cases} 0.0015 & \text{if } u_t > \lambda \quad \text{where} u_t \sim \mathcal{U}(0,1) \\ 0.03 & \text{otherwise} \end{cases}
\end{aligned}
\tag{3}
$$

Where $P_0 = 100$ (starting price) and $\lambda = 0.005$ (outlier rate or intensity). We estimate the Augmented Dickey-Fuller (ADF) test on stationarity for simulated prices. At a 1% confidence level, we find 117 cases for which we accept that the series is stationary and at a 5% confidence level, we find 509 cases. Therefore, a few outliers (in our case approximately 10 out of 2,000) may give a false positive regarding the stationarity of a time series or around 5.09% [2]

Next, on previously simulated daily returns from 3, we analyze the overlapping samples effect on scaling.[3] Our question is if an overlapping sample of asset returns shows scaling. Table-1 shows that samples overlapping, unlike non-overlapping, are not always scaling, and then our proposal is only applicable in non-overlapping cases.

[2] Also, we simulate with $\kappa = (0.015, \quad 0.06)$ to check the outlier sizes effect, and the results are 103 false positives at 1% and 481 at 5%. Additionally, we compared the outlier rate ($\lambda = 0.01$), and the results at 1% and 5% are 112 and 510, respectively. In short, if there are outliers, we found a false positivity rate of stationarity of over 5%.

[3] For example, if we estimate weekly returns from Monday to Monday on a time series of T-size, we obtain $\frac{T}{7}$ returns, but we can also calculate it from Tuesday to Tuesday and so on. In this last case, the result is $T - 7$ observations. The first sample

Table 1: Scaling analysis for overlapping sample

Scaling Term	Overlapping	Non-overlapping
Non-scaling	2471	0
1	7529	7529
2	0	1573
3	0	674
4	0	132
5	0	52
6	0	27
7	0	9
8	0	1
Total	10000	10000

Now, we analyze the replacement of outliers. This is an important problem in detecting and handling outliers, since our proposal is an iterative method to avoid removing an excessive number of observations we could use to replace different options: fixed or dynamic values. In the first case, all outliers are replaced by the same value, while in the second, each outlier is substituted by a different value in each iteration. To do so, we use simulated daily returns from 3. Table-2 shows, for each type of substituion, the root square mean error (RSME) between real outliers and the identified outlier, applying our proposal. Note that zero is the replacement with the lowest error prediction of the outlier number, instead mean and median values needed to eliminate more outliers to achieve the normality of the series, which is a consequence of the jump size effect with respect to the central moment values, since a priori, without other external information, it is impossible to disaggregate each observation into a typical and atypical value or outlier. In addition, the dynamic replacements error is higher than the respective fixed values. These results are a consequence of the endogeneity introduced into the data generating process, when the replacement is adjusted to the new samples or to the sign of the substituted outlier; this effect is observed through higher values in the autocorrelation tests (raw and square data), so more outliers must be eliminated to overcome them. Therefore, we use $rp = 0$ in the rest of paper.

Table 2: Replacement of outliers

Replacement	RMSE
Zero	8.0226
Mean fix	8.8777
Median fix	8.8881
Q1-Q3 fix	9.152
Mean dynamic	8.8505
Median dynamic	8.8617
Q1-Q3 dynamic	9.5715

Next, we compare our proposal results to other outlier tests on simulated prices according to 3. Table-3 shows that Grubbss, Dixons and Thompson tests have the worst results for the identification of outliers

is non-overlapping, while the second is overlapping. The disadvantage of the second sample is that each observation collects information from the previous 7 and therefore involves a moving average.

(higher RMSE). Although the Hampel and Tukeys tests have a lower error in the identification of outliers than our proposal, the clean series of outliers shows higher percentage of series with autocorrelation and heteroskedasticity and non-normality than our proposal.

Table 3: Results of checking different methods for detecting outliers

Method	Normality	Non-Autorregresive on Raw	Non-Autorregresive on Square	Non Heterokedasticity	RMSE
Dixon	5%	2%	39%	31%	10.5164
Grubbs	1%	95%	95%	100%	11.9321
Hampel	88%	95%	95%	100%	7.9879
Thompson	3%	95%	95%	100%	11.9321
Tukey	45%	95%	95%	90%	7.9416
New proposal	100%	100%	100%	100%	8.0226

Finally, we check if it is possible that a autoregressive and heteroskedastic time series can also display the scaling property. For that, we simulate 10,000 time series with 2,000 observations each as:

$$
\begin{aligned}
& i = 1, \ldots, 10000 \quad t = 1, \ldots, 2000 \\
AR(1): \quad & r_{i,t} = \rho_i \cdot r_{i,t-1} \quad \rho_i \sim \mathcal{U}(0.1, 0.5) \quad r_{i,0} \sim \mathcal{N}(0, 0.015^2) \\
GARCH(1,1): \quad & r_{i,t} = \sigma_{i,t} \cdot \epsilon_{i,t} \\
& \sigma^2_{i,t} = \alpha_{0,i} + \alpha_{1,i} \cdot \epsilon^2_{i,t-1} + \beta_{i,1} \cdot \sigma^2_{i,t-1} \\
& \alpha_{i,0} \sim \mathcal{U}(0.01^2, 0.02^2) \quad \alpha_{i,1} \sim \mathcal{U}(0.05, 0.15) \quad \beta_{i,1} \sim \mathcal{U}(0.7, 0.9)
\end{aligned} \tag{4}
$$

Table-4 shows the scaling results. Note that the autoregressive model does not display scaling and the heteroskedastic model shows a higher scaling than the usual value for asset returns. In most cases, we observe scaling of over 6 months (125 or more days) for simulated times series.

Table 4: Scaling analysis for autoregressive and heteroskedastic process

Scaling	AR(1)	GARCH(1,1)	AR(1)+GARCH(1,1)
0	0	0	0
5	0	0	0
10	0	2	0
20	0	32	0
30	0	71	0
40	0	128	0
50	0	325	0
100	0	874	5
125	0	2054	42
150	0	4395	173
175	0	5895	192
200	0	8799	227
225	32	9800	281
(+)250	10000	10000	10000

Note: the econometric models of financial assets returns do not usually show more than one lag, but we have verified that by increasing the number of lags the results are even more conclusive about the outliers, ie, scaling value is very high.

3.2. Financial market data

3.2.1. Sample and stylized facts

Our sample covers the components of the stock market indexes for different countries: CAC-40 (France), DAX-30 (Germany), IBEX-35 (Spain), SWISS-20 (Switzerland), FTSE-100 (UK), NASDAQ-100 (USA), DJI-30 (USA), IBOVESPA (Brazil), IPSA-30 (Chile), IPC-35 (Mexico), MERVAL-20 (Argentina), IGBC (Colombia), JSE (South Africa), HANG SENG-50 (Hong Kong), ASX-200 (Australia), NIKKEI-225 (Japan) and SHANGHAI-300 (China). Additionally, we include 16 exchange rates to the USA dollar: euro (EUR), Great Britain pound (GBP), Canadian dollar (CAD), Australian dollar (AUD), New Zealand dollar (NZD), Swiss franc (CHF), Danish crown (DKK), Norwegian crown (NOK), Swedish crown (SEK), Singapore dollar (SGD), South Korean won (KRW), Taiwan dollar (TWD), South African rand (ZAR), Mexican peso (MXN), Brazilian real (BRL) and Colombian peso (COP). Different commodity spot prices are also included: gold, silver, nickel, WTI crude, Brent, steel, Bitcoin, coffee, wheat, milk, cotton and corn. The period studied is from 01/01/2014 to 12/31/2018 (daily closing prices) and all data are from Bloomberg. We exclude stocks with less than 300 consecutive trading days (more than one year) within the period. The sample is made up of 1,330 assets for 5 years of daily returns [4].

Table-5 show, for each stock market index, exchange rate and commodity, corresponding assets with the lowest, the median and the highest number of outliers. We compare normality, autocorrelation and heteroskedasticity tests of the original daily return and the clean data after replacing outliers with zero. Note that the maximum number of outliers is 172, the lowest 7 and the median 70 and, after replacing them, all the series are Gaussian *i.i.d.*

Table-6 compares jumps up with jumps down. Also, it shows detection outliers with our proposal and using Tukeys test (inter-quartile). We observe that Tukeys test overestimates the number of outliers for the minimums, is similar in the median and underestimates them in the maximums. In addition, we note that the minimum, median and maximun jumps of both signs (negative and positive) is 3 (intensity of 0.26%-0.24%), 34-35 (intensity of 2.69%-2.85%) and 83 (intensity 6.81%) respectively. As a result, the descriptive statistics for both types of jumps are very similar.

Table-7 shows the frequency of negative and positive outliers and respective intensity rates for all the samples. Note that the usual jumps (median or 50%) are 2.69% (down) and 2.85% (up), and jumps higher than 6.79% (down) and 6.71% (up) are less than 1%.

Now, we apply the identification of outliers to check if the shocks are the same in different markets for the same asset. However, the asynchrony of the data or the complexity of the model used to adjust them makes this analysis complicated. Instead, the methodology used in this paper is simple and fast, and we have a particular case to compare, since IAG is included on both the IBEX-35 and the FTSE-100. Figure-1,

[4] We have preferred to include a wide and varied type of assets, at the cost of not having more than 5 years of daily market prices. Cont (2001) also employ similar periods in their study. Additionally, in 3.1, we analyze simulated samples with 2,000 observations (about 8 years of daily data) and 10,000 assets

Table 5: Tests for original daily return and clean data

Market	Ticker	Observ.	Scaling	Outliers	Original data				Clean data			
					JB	AR raw	AR square	ARCH)	JB	AR raw	AR square	ARCH
F. Exchange	SEK	1303	4	21	150.689[**]	5.008	7.55	7.632	5.89	3.852	1.247	1.359
	CHF	1303	25	43	66899.564[**]	16.242[**]	0.532	0.522	1.765	4.21	7.995	7.395
	MXN	1303	11	101	4948.393[**]	12.247[*]	67.932[**]	54.438[**]	5.981	0.958	3.747	3.789
CAC	UG	1277	9	29	47.938[**]	3.5	7.743	7.291	2.042	3.115	6.57	6.968
	MT NA	1277	24	87	57.778[**]	1.594	66.028[**]	49.814[**]	0.107	7.165	5.91	5.592
	SW	1277	28	124	1987.844[**]	3.747	5.905	5.359	3.178	7.836	7.585	7.726
DAX	ADS	1262	9	23	8448.308[**]	7.946	0.7	0.696	5.112	4.912	7.311	7.052
	FME	1262	39	70	16280.617[**]	3.787	4.556	4.504	2.462	6.127	7.275	6.964
	DAI	1262	7	172	290.065[**]	14.306[**]	63.479[**]	48.192[**]	1.121	1.83	8.388	7.729
IBEX	ENC	1277	8	24	229.602[**]	4.508	21.514[**]	18.616[**]	1.487	3.92	7.769	7.227
	SAN	1277	11	73	16684.717[**]	4.212	8.345	7.566	0.719	3.758	7.358	6.816
	REP	1277	23	151	1104.987[**]	8.804	171.309[**]	121.066[**]	0.192	2.235	7.813	7.177
SWISS	GIVN	1253	6	39	5340.36[**]	2.907	71.507[**]	66.536[**]	4.716	1.885	8.236	7.275
	ABBN	1253	17	72	1729.387[**]	25.577[**]	56.403[**]	51.924[**]	1.115	7.509	8.168	7.368
	USBG	1253	11	113	2734.112[**]	14.905[**]	96.18[**]	81.445[**]	2.368	9.459	5.177	5.374
IBOVESPA	CYRE3	1233	9	13	1103.724[**]	2.986	28.046[**]	23.181[**]	4.743	8.259	7.732	7.185
	BBSE3	1233	10	65	296.799[**]	8.624	100.005[**]	89.052[**]	2.831	6.516	8.531	7.606
	MRVE3	1233	4	105	165.402[**]	5.579	22.878[**]	19.587[**]	5.971	1.865	0.379	0.382
IPSA	CCU	1242	7	25	219.3[**]	5.544	21.331[**]	18.328[**]	5.202	3.408	5.015	4.693
	ILC	1242	21	74	9733.448[**]	5.473	13.304[**]	11.524[**]	2.248	3.601	8.645	7.652
	SMCHILEB	1242	21	98	17836.818[**]	21.769[**]	3.415	3.367	0.036	8.87	2.185	2.082
IPC	KIMBERA	1255	6	9	85.419[**]	7.355	9.938[*]	9.124[*]	1.983	6.377	7.67	7.394
	WALMEX	1255	7	60	742.822[**]	7.858	6.331	5.77	2.946	8.165	7.491	6.693
	GFINBURO	1255	6	137	159.311[**]	25.197[**]	40.122[**]	39.89[**]	3.007	9.142	4.168	4.049
MERVAL	TGSU2	1213	7	27	200.278[**]	4.588	40.357[**]	35.243[**]	3.553	4.738	6.37	6.136
	FRAN	1217	8	95	209.128[**]	10.193[*]	210.171[**]	131.211[**]	0.14	2.404	7.927	7.11
	TXAR	1217	6	132	144.61[**]	36.82[**]	86.96[**]	70.906[**]	2.03	8.965	2.419	2.388
IGBC	PFAVAL	1215	13	45	980.291[**]	2.82	99.109[**]	86.279[**]	5.569	6.918	6.939	7.02
	ECOPETL	1215	11	86	415.721[**]	7.726	65.952[**]	54.553[**]	0.213	6.882	8.279	7.471
	PFGRUPOA	1204	13	145	251.993[**]	19.084[**]	98.541[**]	74.492[**]	0.69	8.606	6.247	5.906
JSE	WHL	1247	7	7	701.255[**]	8.774	96.383[**]	89.22[**]	4.483	4.083	5.889	5.316
	SHP	1247	10	66	261.898[**]	8.237	15.399[**]	13.724[**]	2.345	1.488	7.006	6.797
	DSY	1247	6	101	1780.423[**]	7.236	133.685[**]	110.524[**]	4.986	8.934	6.26	6.508
HANG SENG	101 HK	1230	6	14	212.691[**]	2.038	22.082[**]	21.718[**]	4.619	2.218	6.565	7.242
	11 HK	1230	22	65	1308.406[**]	1.138	45.739[**]	35.504[**]	0.686	4.714	8.233	7.742
	669 HK	1230	9	121	269.011[**]	26.017[**]	38.232[**]	31.508[**]	5.951	3.759	3.145	3.019
NASDAQ	JBHT	1257	5	17	64.472[**]	0.816	5.893	5.322	4.797	1.448	6.891	6.391
	AMGN	1257	11	81	583.951[**]	6.268	14.395[**]	12.503[**]	0.578	1.595	5.358	5.161
	CHTR	1257	7	107	921.751[**]	3.795	5.608	5.448	2.68	9.393	4.78	4.767
DOW JONES	VZ	1257	9	48	616.123[**]	1.917	16.903[**]	15.013[**]	0.592	1.188	6.09	5.891
	DIS	1257	7	82	2446.251[**]	2.365	18.723[**]	15.926[**]	1.357	7.971	8.22	7.656
	V	1257	14	139	1307.896[**]	14.472[**]	43.506[**]	34.986[**]	5.271	9.277	6.337	5.695
ASX	SCG	1145	4	7	5.456	9.355	23.431[**]	18.601[**]	1.517	9.199	9.345	7.267
	CSR	1264	5	58	451.558[**]	4.437	14.779[**]	13.227[**]	0.793	8.043	6.167	5.962
	TLS	1264	16	153	5669.74[**]	4.248	19.778[**]	17.861[**]	5.991	3.042	4.664	4.105
FSTE	CPG	1263	4	15	248.404[**]	3.565	5.718	5.594	5.735	1.687	8.224	7.75
	ITRK	1263	5	58	2876.052[**]	2.586	2.25	2.14	5.898	0.91	6.955	6.827
	PSN	1263	25	142	181606.523[**]	28.248[**]	54.815[**]	54.009[**]	5.938	4.603	1.888	1.821
NIKKEI	5707 JT	1224	3	16	193.92[**]	1.693	18.696[**]	17.673[**]	5.539	0.885	7.452	6.843
	4506 JT	1224	17	67	13517.088[**]	3.263	13.691[**]	14.725[**]	4.436	5.858	7.061	7.158
	5541 JT	1224	8	164	1741.257[**]	18.269[**]	8.922	8.486[*]	3.355	8.795	6.822	7.256
SHANGHAI	002241 CH	1205	9	35	288.703[**]	1.985	162.731[**]	103.962[**]	5.534	2.679	4.202	4.055
	000709 CH	1180	28	129	789.73[**]	16.838[**]	512.369[**]	262.359[**]	5.638	0.927	4.328	4.151
	603288 CH	1197	18	148	7812.481[**]	9.203	156.075[**]	148.884[**]	4.508	8.812	8.177	7.602
Commodities	Gold	1303	5	30	376.524[**]	4.534	5.382	5.181	5.416	6.732	4.299	4.579
	Cottom	1303	7	53	165.677[**]	0.845	16.966[**]	17.015[**]	5.747	1.447	2.899	2.931
	Brent	1303	17	112	322.427[**]	4.358	273.179[**]	175.73[**]	1.472	1.651	6.188	6.127

Note: JB is Jarque-Bera normality test. [*] and [**] mean that value test rejects null hypothesis at 5% and 1% respectively.

in its two top graphs, shows the relationship between the original daily returns for IAG in both markets and the relationship between the same data but without outliers. The limits of the relationship diminish considerably when eliminating the outliers. In addition, the correlation between the original returns is 0.97, while once the outliers are eliminated it is 0.90; the difference cannot only be attributed to outliers. The

Table 6: Analysis outliers: jumps down and up

Market/index	Ticker	Jumps Down					Jumps Up				
		Tukey(-)	Outliers(-)	$\lambda(-)$	Mean	Std. Dev.	Tukey(+)	Outliers(+)	$\lambda(+)$	Mean	Std. Dev.
	SEK	22	11	0.84%	-0.0185	0.0026	17	10	0.77%	0.0194	0.0058
Foreign Exchange	CHF	28	23	1.77%	-0.0234	0.0365	23	20	1.53%	0.0159	0.004
	MXN	14	46	3.53%	-0.0132	0.0041	15	55	4.22%	0.0136	0.0072
	UG	26	12	0.94%	-0.0825	0.0402	29	17	1.33%	0.0693	0.0189
CAC-40	MT NA	29	42	3.29%	-0.0689	0.0202	33	45	3.52%	0.0711	0.0172
	SW	24	58	4.54%	-0.022	0.0157	16	66	5.17%	0.0208	0.011
	ADS	20	10	0.79%	-0.0644	0.0353	24	13	1.03%	0.067	0.0204
DAX-30	FME	29	34	2.69%	-0.0441	0.0267	32	36	2.85%	0.0377	0.0113
	DAI	38	86	6.81%	-0.0329	0.0113	35	86	6.81%	0.0304	0.0065
	ENC	21	13	1.02%	-0.0773	0.0168	19	11	0.86%	0.0793	0.0194
IBEX-35	SAN	27	36	2.82%	-0.0537	0.0348	27	37	2.90%	0.0465	0.0094
	REP	46	77	6.03%	-0.0388	0.0151	40	74	5.79%	0.0378	0.013
	GIVN	30	22	1.76%	-0.0372	0.0187	25	17	1.36%	0.034	0.0084
SWISS-20	ABBN	31	39	3.11%	-0.0366	0.0156	23	33	2.63%	0.0307	0.0049
	USBG	33	54	4.31%	-0.0419	0.0205	28	59	4.71%	0.036	0.01
	CYRE3	16	7	0.57%	-0.0939	0.0342	15	6	0.49%	0.0881	0.0125
IBOVESPA	BBSE3	16	33	2.68%	-0.0515	0.0165	19	32	2.60%	0.0526	0.0137
	MRVE3	12	49	3.97%	-0.0361	0.0141	10	56	4.54%	0.0356	0.0118
	CCU	20	13	1.05%	-0.042	0.0078	26	12	0.97%	0.0428	0.0086
IPSA	ILC	24	24	1.93%	-0.0406	0.0167	50	50	4.03%	0.04	0.0198
	SMCHILEB	22	47	3.78%	-0.0134	0.0063	16	51	4.11%	0.0137	0.0084
	KIMBERA	13	4	0.32%	-0.0579	0.0087	14	5	0.40%	0.0615	0.008
IPC	WALMEX	19	28	2.23%	-0.0422	0.0151	15	32	2.55%	0.0374	0.0131
	GFINBURO	14	65	5.18%	-0.0305	0.0107	20	72	5.74%	0.0305	0.0084
	TGSU2	14	10	0.82%	-0.101	0.0237	22	17	1.40%	0.0891	0.017
MERVAL	FRAN	28	38	3.12%	-0.0762	0.0177	40	57	4.68%	0.0779	0.0186
	TXAR	20	64	5.26%	-0.042	0.0172	29	68	5.59%	0.0443	0.016
	PFAVAL	41	29	2.39%	-0.0385	0.0095	24	16	1.32%	0.0358	0.0124
IGBC	ECOPETL	31	45	3.70%	-0.0532	0.0149	33	41	3.37%	0.0562	0.0152
	PFGRUPOA	38	74	6.15%	-0.0293	0.0118	24	71	5.90%	0.0269	0.0099
	WHL	11	4	0.32%	-0.0806	0.0149	18	3	0.24%	0.0918	0.0286
JSE	SHP	21	31	2.49%	-0.0469	0.0098	24	35	2.81%	0.0488	0.0145
	DSY	16	51	4.09%	-0.0293	0.0134	16	50	4.01%	0.0306	0.0128
	101 HK	15	7	0.57%	-0.0551	0.0118	20	7	0.57%	0.0517	0.011
HANG SENG	11 HK	39	37	3.01%	-0.0282	0.0059	25	28	2.28%	0.0319	0.0128
	669 HK	22	61	4.96%	-0.0317	0.0114	39	60	4.88%	0.035	0.0147
	JBHT	22	10	0.80%	-0.0411	0.0031	14	7	0.56%	0.0432	0.0076
NASDAQ-100	AMGN	33	41	3.26%	-0.0398	0.013	25	40	3.18%	0.0393	0.0094
	CHTR	27	45	3.58%	-0.0329	0.0148	30	62	4.93%	0.0313	0.0116
	VZ	25	23	1.83%	-0.0325	0.0073	27	25	1.99%	0.0305	0.0098
DOW JONES	DIS	35	40	3.18%	-0.0342	0.0138	22	42	3.34%	0.0287	0.0104
	V	40	71	5.65%	-0.0236	0.0093	24	68	5.41%	0.0222	0.0116
	SCG	11	3	0.26%	-0.0358	0.0004	11	4	0.35%	0.038	0.0018
ASX-200	CSR	21	30	2.37%	-0.0527	0.0177	18	28	2.22%	0.0495	0.0111
	TLS	18	79	6.25%	-0.0185	0.0121	16	74	5.85%	0.0172	0.0078
	CPG	14	8	0.63%	-0.043	0.007	12	7	0.55%	0.0417	0.0067
FSTE-100	ITRK	35	28	2.22%	-0.0424	0.0163	32	30	2.38%	0.0449	0.0171
	PSN	22	69	5.46%	-0.0323	0.0293	23	73	5.78%	0.0293	0.0113
	5707 JT	13	8	0.65%	-0.1015	0.0263	21	8	0.65%	0.0872	0.0097
NIKKEI-225	4506 JT	35	32	2.61%	-0.0643	0.0356	35	35	2.86%	0.0656	0.0318
	5541 JT	26	83	6.78%	-0.0295	0.0163	22	81	6.62%	0.0305	0.0156
	002241 CH	33	25	2.07%	-0.0937	0.0112	20	10	0.83%	0.0887	0.0061
SHANGHAI-300	000709 CH	51	66	5.59%	-0.0658	0.0234	44	63	5.34%	0.0638	0.0202
	603288 CH	33	72	6.02%	-0.0329	0.0182	36	76	6.35%	0.0347	0.0222
	Gold	26	15	1.15%	-0.0248	0.0037	24	15	1.15%	0.0282	0.0075
Commodities	Cottom	32	21	1.61%	-0.0345	0.0054	39	32	2.46%	0.0354	0.0047
	Brent	34	54	4.14%	-0.0509	0.0129	42	58	4.45%	0.0518	0.0149

original series are not Gaussian so a linear correlation coefficient does not seem an adequate measure of dependence between two non-normal serie (see Kim et al. (2015)). The two bottom graphs in Figure-1 show the outliers detected (up and down). We observe two types of patterns in the jumps: those that take place at the same time in both markets and those that only happen in one of the two markets. The jumps seem to have two possible origins, either systematic or idiosyncratic.

Table 7: Frequency of jumps down and up

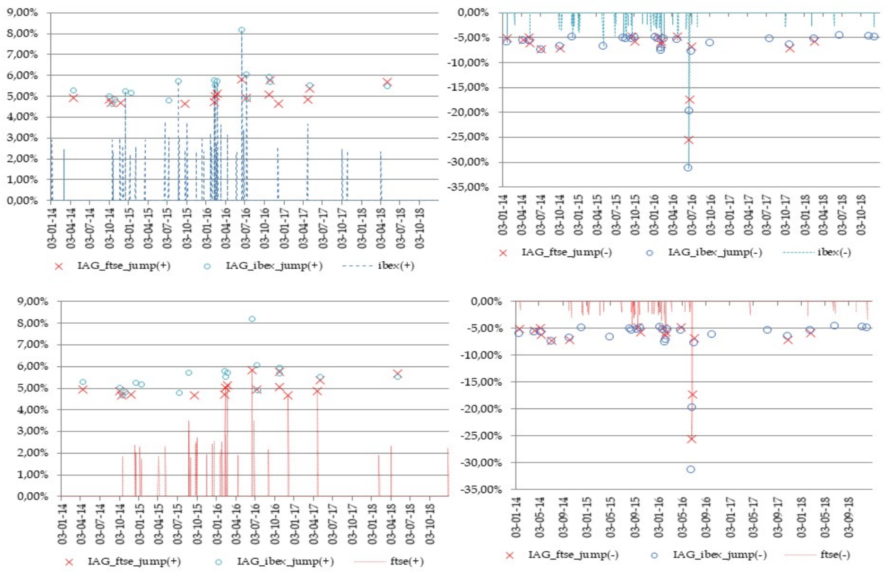
Percentile	Outliers(-)	$\lambda(-)$	Outliers(+)	$\lambda(+)$
1%	4	0.32%	4	0.35%
5%	6	0.56%	6	0.49%
25%	15	1.15%	16	1.32%
50%	34	2.69%	35	2.85%
75%	54	4.14%	59	4.71%
95%	77	6.03%	74	5.79%
99%	84	6.79%	83	6.71%

Figure 1: Outliers for IAG company from FTSE-100 and IBEX-35 data

So, being able to differentiate if an outlier has a systemic or idiosyncratic origin is important since, in the first case, we analyze contagion effects, while in the second we look for the cause: reputational, credit or liquidity risks, among others. Then we check whether each jump or outlier for IAG is idiosyncratic or systematic. To do so, we identify systematic jumps on a date from daily return of indexes. Figure-2 shows up and down jumps for IAG and for both markets, IBEX-35 and FTSE-100. We observe that the most outliers detected correspond to outliers for the Spanish market (IBEX-35), and there are dates in which both markets have outliers (contagion effect), so they are systematic outliers. There are also dates with independent market outliers, for example in October 2016 (up) or February 2018 (down). For the latter, note that some financial newspapers said that IAG did not attain its expected profits (falling 4.65% in the stock market) since it only gained 3.5% more than the previous year when more was expected for fiscal year 2017.

Another important use for this outlier detection methodology is asset pricing and selection portfolio. Thus, if we estimate the beta coefficient as the slope of the linear regression between the daily asset returns and the market index (Capital Asset Pricing Model), we can also estimate this parameter for the data without outliers, with positive jumps and with negative jumps. We select IBEX-35 and all its components

Figure 2: Outliers for IAG vs. outliers of markets (FTSE-100 and IBEX-35)



that have quoted the full sample period and, we estimate four betas, as shown in Table-8. Note that stock with the highest beta of original data and clean data is SAN or Santander Bank with 1.4542 (original) and 0.5861 (clean). However, if we analyze the betas of the jumps or outliers detected, we find that the highest beta for positive jumps in the market (IBEX-35) is 0.6704 for ANA (Acciona) and, for the negative ones, IAG or International Airlines Group (0.6738). Finally, the sum of clean, up, and down betas show that SGRE (Siemmens-Gamesa) has the highest value (1.4579). There are remarkable differences between the estimated beta of the original series and the sum of the clean, up, and down betas. REP (Repsol) shows the highest overvaluation (+0.5056) and ENG (Enagás) the highest underestimate (-0.5505). Therefore, we can value and select financial assets based not only on the expected value, using the beta clean (estimated on Gaussian sample data), but we can also make decisions based on the jumps and their sign.

Finally, we check the utility our prosopal comparing the results of risk estimation on the FTSE-100 index. We divided the sample period into two parts: the estimate period (from January 1, 2014 to December 31, 2017) and the check period (from January 1, 2018 to December 31, 2018). Then, we calculate the risk for 1-day ahead at confidence level of 99% (long position) and 1% (short position). The procedure is rolling: every day the parameters are estimated using a sample that replaces the oldest date with the newest one. Additionally, following the same procedure, the risk is estimated for a temporal horizon equal to the FTSE-100 time scaling. For 1-day horizon, we use to estimate market outlier detection[5] and GARCH t-student (the best AIC value) and, for term horizon of time scaling, we use Gaussian distribution. Table-9 shows that the risk underestimate (fault) is higher in the outlier method than in GARCH-t, but the overestimate

[5]As original data consists of three independent components (2), we estimate the risk at confidence level as weighted sum ($\lambda^{+/-}$) of Gaussian variable (clean data) and Cornish-Fisher expansion for outliers, since these jumps show asymmetry and kurtosis.

Table 8: Beta (slope on daily data) of components IBEX-35

Ticker	original beta	clean beta	beta up	beta down	clean+up+down	Original - Sum
ACS	1.0389	0.4934	0.2778	0.1491	0.9203	0.1185
ACX	0.8623	0.3659	0.1895	0.2555	0.811	0.0513
AMS	0.5908	0.2376	0.0983	0.1473	0.4831	0.1077
ANA	0.8914	0.3082	0.6704	0.3129	1.2915	-0.4
BBVA	1.3028	0.5156	0.3192	0.3577	1.1924	0.1103
BKIA	1.2177	0.4694	0.2159	0.1327	0.8181	0.3996
BKT	0.9885	0.4338	0.3278	0.3675	1.1291	-0.1406
CABK	1.2802	0.5187	0.2849	0.2429	1.0465	0.2337
CIE	0.7828	0.3056	0.176	0.3021	0.7837	-0.0009
COL	0.7138	0.2263	0.3138	0.1751	0.7152	-0.0014
ENC	1.0257	0.1654	0.1955	0.1732	0.5342	0.4916
ENG	0.5561	0.5141	0.2279	0.3646	1.1066	-0.5505
ENO	0.3286	0.2516	0.1592	0.2014	0.6123	-0.2836
FER	0.7509	0.3474	0.1534	0.1891	0.6899	0.061
GRF	0.668	0.356	0.0228	0.0732	0.452	0.216
IAG	1.1973	0.3837	0.0926	0.6738	1.1501	0.0472
IBE	0.6828	0.2303	0.263	0.1559	0.6492	0.0336
IDR	0.8545	0.3234	0.271	0.2527	0.8471	0.0074
ITX	0.8364	0.3114	0.4037	0.2658	0.9809	-0.1445
MAP	1.0007	0.4198	0.1462	0.4022	0.9681	0.0325
MEL	0.7017	0.3169	0.0877	0.0797	0.4843	0.2174
MTS	1.2052	0.3281	0.319	0.321	0.9681	0.2371
REE	0.5818	0.3007	0.1977	0.2556	0.754	-0.1723
REP	1.0656	0.2459	0.1512	0.1629	0.56	0.5056
SAB	1.3563	0.3337	0.619	0.4196	1.3723	-0.016
SAN	1.4542	0.5861	0.1291	0.2492	0.9644	0.4897
SGRE	1.0699	0.5538	0.4549	0.4492	1.4579	-0.3879
TEF	1.0637	0.4585	0.5379	0.2746	1.271	-0.2073
TL5	0.8633	0.3246	0.5065	0.3924	1.2235	-0.3602
TRE	0.6974	0.3743	0.1912	0.0628	0.6283	0.069
VIS	0.4573	0.3003	0.3188	0.1356	0.7546	-0.2973

is higher in GARCH-t. If we observe the backtesting cases (fault/overestimate case %) then we verify that the outlier method adjusts better than GARCH-t to long position (1.2% errors vs. 1.61%) and is the same for short positions. Also note that errors do not exceed 1% for the time scaling horizon. When comparing it to amounts (excess or overestimate amount), however, GARCH-t shows the highest percentages for both the long and short positions, and this implies an excess of capital requirement with respect to risk, with the consequent decrease in the yield per unit of capital at risk.

Table 9: Results for risk estimate by GARCH-t and outliers methods for 1-day ahead and normal estimate for 13-days ahead or time scaling of FTSE-100

Estimation	Model	99%	1%
Fault mean	GARCH-t 1 day	-22.7	17.84
	OUTLIERS 1 day	-79.57	50.95
	Normal scaling time 13 days	-62.61	0.00
Overestimate mean	GARCH-t 1 day	147.25	-161.8
	OUTLIERS 1 day	114.58	-120.51
	Normal scaling time 13 days	478.01	-638.77
Fault number	GARCH-t 1 day	4	2
	OUTLIERS 1 day	3	2
	Normal scaling time 13 days	2	0
Overestimate number	GARCH-t 1 day	249	251
	OUTLIERS 1 day	250	251
	Normal scaling time 13 days	239	241
Fault/Over amount %	GARCH-t 1 day	15.42%	11.03%
	OUTLIERS 1 day	69.44%	42.28%
	Normal scaling time 13 days	13.10%	0.00%
Fault/Over Cases %	GARCH-t 1 day	1.61%	0.80%
	OUTLIERS 1 day	1.20%	0.80%
	Normal scaling time 13 days	0.84%	0.00%

4. Conclusions

The stylized facts are statistical properties attributed to the asset returns. An enormous development of increasingly complex econometric models are required to model them. But when the estimate frequency of these returns becomes lower, some of these facts disappear. This characteristic is called time scaling. So while the high frequency returns are not Gaussian, when the calculate frequency of asset return decreases, their behavior converges to a normal distribution. Evidently, it is the result of the Central Limit Theorem, since for non-overlapping samples the lower frequency returns (e.g. weekly) are the result of the sum of the most frequent (daily). The econometric and statistical literature has shown that the existence of outliers entails unbiased estimates of the models parameters and the false positives in the hypothesis about the existence of heteroskedasticity. In this context, our aim is to found an empirical answer about the relationship between time scaling and outliers in high frequency asset returns. Our proposal replaces the highest return, in absolute value, until the resulting sample overcomes the tests of normality, non-autocorrelation for raw series and squared data, and absence heteroskedasticity.

Our empirical study has two parts. First, we check the results of this identification outliers procedure on simulated time series and, we observe that in presence of outliers: (i) the hypothesis of stationarity was erroneously accepted for around 5% of the cases of simulated prices; (ii) the autoregressive and/or heteroskedastic processes show scaling for frequency much higher than those usual in the financial markets (6 months or more); (iii) replacing the outliers with different values in each iteration leads to autocorrelation problems and using zero as a substitute had the least errors and avoids distinguishing between what part of the returns is usual and which part is atypical; (iv) only two traditional methods of detecting outliers (Tukey and Hampel) showed a slightly higher capacity to detect outliers than our proposal, however, these methods do not eliminate the remaining statistical problems so the resulting sample is not Gaussian; and (v) the overlapping, unlike non-overlapping, samples do not have time scaling.

As a consequence of these experimental results, we consider the possibility of modeling high frequency returns in way other than the usual AR-GARCH. So, in the second part of the experimental analysis, we apply our proposal to a sample of daily returns of 1,330 assets (with around 1,250 daily observations per asset) that include stocks, exchange rates and commodities. The main conclusions of this second part of study are: (i) by replacing a few observations (between 1% and 6% or only 12 and 75 observations) the resulting sample is adjusted to a normal distribution (non-autoregressive and non heteroskedasticity); (ii) the division of the information contained in the original data allows us to separately analyze the behavior random and the extraordinary or atypical (outliers or jumps); (iii) as a consequence, we can determine the frequency and intensity of outliers and distinguish if they are idiosyncratic or systematic; (iv) this also opens up other possibilities in asset management, specifically in the identification of assets with lower or higher reactions to systematic outliers and also facilitate to simpler risk estimate with more satisfactory results than complex model as GARCH t-student, in addition to allowing to estimate risk through to simple normal distribution for temporal horizons equal to the time scaling of the original series.

Acknowledgement

This work has been supported by the Spanish Ministry of Economics and Competitiveness under grant MINECO/FEDER ECO2015-65826-P, and Cátedra Universidad CEU San Pablo-Mutua Madrileña insurance company (grant ARMEG 060516-USPMM-01/17).

References

Aguinis, H., Gottfrendson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, *16*, 270–301.

Ahmad, Y., & Donayre, L. (2016). Outliers and persistence in threshold autoregressive processes. *Studies in Nonlienar Dynamics & Econometrics*, *20*, 37–56.

Alih, E., & Ong, H. C. (1996). An outlier-resistant test for heteroskedasticity in linear models. *Journal of Applied Statistics*, *48*, 617–634.

Antypas, A., Koundouri, P., & Kourogenis, N. (2013). Aggretational gaussianity and barely infinnite variance in financial returns. *Journal of Empirical Finance*, *20*, 102–108.

Baragona, R., Battaglia, F., & Cucina, D. (2016). Empirical likelihood for outlier detection and estimation in autoregressive time series. *Journal of Time Series Analysis*, *37*, 315–336.

Barnett, V., & Lewis, T. (1978). *Outliers in Statistical Data*. New York, USA: Wiley.

Berry, A. C. (1941). The accuracy of the gaussian approximation to the sum of independent variates. *Transactions of the American Mathematical Society*, *49*, 122–136.

Carnero, A., Peña, D., & Ruiz, E. (2006). Effects of outliers on the identification and estimation of garch models. *Journal of Time Series Analysis*, *28*, 471–497.

Charles, A. (2008). Forecasting volatility with outliers in garch models. *Journal of Forecasting*, *27*, 551–565.

Chen, C., & Liu, L.-M. (2011). Joint estimation of model parameters and outlier effects in time series. *Journal of the American Statistical Association*, *88*, 284–297.

Chiang, M.-H., Chou, R., & Wang, L.-M. (2016). Outlier detection in the lognormal logarithmic conditional autoregressive range model. *Oxford Bulletin of Economic and Statistics*, *78*, 126–144.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, *1*, 223–236.

Dang, X., & Serfling, R. (2011). A numerical study of multiple imputation methods using nonparametric multivariate outlier identifiers and depth-based performance criteria with clinical laboratory data. *Journal of Statistical Computation and Simulation*, *85*, 547–570.

Dean, R. B., & Dixon, W. J. (1951). Simplified statistics for small numbers of observations. *Analytical Chemistry*, *23*, 636–638.

Diebold, F., & Mariano, R. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, *13*, 134–144.

van Dijk, D., Franses, P. H., & Lucas, A. (1999). Testing for smooth transition nonlinearity in the presence of outliers. *Journal of Business and Economics Statistics*, *17*, 217–235.

Elliott, M. R., & Stettler, N. (2007). sing a mixture model for multiple imputation in the presence of outliers: U the 'healthy for life' project. *Applied Statistics*, *56*, 63–78.

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, *50*, 987–1007.

Esseen, C. G. (1942). On the liapunoff limit of error in the theory of probability. *Arkiv för Matematik, Astronomi och Fysik*, *A28*, 1–19.

Fox, A. J. (1972). Outliers in time series. *Journal of the Royal Statistical Society B*, *34*, 350–363.

Franses, P. H., & Ghijsels, E. (1999). Additive outliers, garch and forecasting volatility. *International Journal of Forecasting*, *15*, 1–9.

Franses, P. H., Van Dijk, D., & Lucas, A. (2004). Short patches of outliers, arch and volatility modeling. *Applied Financial Economics*, *14*, 221–231.

Grubbs, F. E. (1950). Sample criteria for testing outlying observations. *Annals of Mathematical Statistics*, *21*, 27–58.

Hampel, F. R. (1974). The influence curve and its role in robust estimation. *Journal of the American Statistical Association*, *69*, 383–393.

Kim, Y., Kim, T.-H., & Ergün, T. (2015). The instability of the pearson correlation coefficient in the presence of coincidental outliers. *Finance Research Letters*, *13*, 243–257.

Koop, G., & Potter, S. (2000). *Nonlinearity, Structural Breaks or Outliers in Economic Time Series*. Cambridge, UK: Barnett, W. and Hendry, D. and Hylleberg, S. and Terasvirta, T. and Tjostheim, D. and Wurtz, A. Cambridge University Press.

Liu, H., Shah, S., & Jiang, W. (2004). On-line outlier detection and data cleaning. *Computers and Chemical Engineering*, *28*, 1635–1647.

Ljung, G., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, *66*, 67–72.

Lucas, A., Franses, P. H., & van Dijk, D. (2002). *Outlier Robust Analysis of Economic Time Series*. Oxford, UK: Oxford University Press Inc.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, *3*, 125–144.

Park, B.-J. (2002). An outlier robust garch model and forecasting volatility of exchange rate returns. *Journal of Forecasting*, *21*, 381–393.

Peña, D. (1990). Influential observations in time series. *Journal of Business & Economic Statistics*, *8*, 235–241.

Sakata, S., & White, H. (1998). High breakdown point conditional dispersion estimation with application to s&p500 daily returns volatility. *Econometrica*, *66*, 529–567.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Boston, USA: Pearson/Allyn & Bacon.

Thompson, G. L. (2006). An spss implementation of the nonrecursive outlier deletion procedure with shifting z score criterion. *Behavior Research Methods*, *38*, 344–352.

Tukey, J. (1977). *Exploratory data analysis*. London, UK: Addison-Wesely.

Weekley, R. A., Goodrich, R. K., & Cornman, L. B. (2010). An algorithm for classification and outlier detection of time-series data. *Journal of Atmospheric and Oceanic Technology*, *27*, 94–107.

Weisberg, S. (1985). *Applied Linear Regression*. New York, USA: John Wiley & Sons.