ELSEVIER

Physica A 300 (2001) 551–557

PHYSICA A

www.elsevier.com/locate/physa

Correlations and multi-affinity in high frequency financial datasets

Roberto Baviera[a], Michele Pasquini[b,*], Maurizio Serva[b], Davide Vergni[c], Angelo Vulpiani[c]

[a] *Departement Finance & Economie, Groupe HEC F-78351, rue de la Liberation 1, Jouy-en-Josas Cedex, France*

[b] *Dipartimento di Matematica and INFM, Ist. Nazionale Fisica della Materia, Università dell'Aquila, I-67010 Coppito, L'Aquila, Italy*

[c] *Dipartimento di Fisica and INFM, Università di Roma "La Sapienza" I-00185 P.le A. Moro 2, Roma, Italy*

Received 25 June 2001

Abstract

In this paper we perform a quantitative check of long term correlations and multi-affinity in Deutsche Mark/US Dollar exchange rates using high frequency data. We show that the use of *business* time, i.e., the ranking of the quotes in the sequences, eliminates most of the seasonality in financial-time series, allowing a precise estimation of some return anomalies.

PACS: 89.65.Gh

Keywords: Structure function; Correlation; Foreign exchange market

1. Introduction

Is it possible to provide a unique statistical description of financial series on different time scales? In this paper we address to this question and we show how the use of *business* time (the ranking of the quotes in the sequences) leads to a precise scaling behavior for correlations and structure functions. The use of business time, in fact, filters out most of the seasonality which affects results obtained by using ordinary calendar time.

* Corresponding author. Tel.: +39-862-433-153; fax: +39-862-433-180.
E-mail address: pasquini@serva.dm.univaq.it (M. Pasquini).

PII: S0378-4371(01)00363-6

We focus in this paper on lags longer than few minutes (where returns are no longer correlated), but shorter than two weeks (after which statistical meaning is lost): in this sense we deal with *long term* return anomalies. Recent research on low frequency data has pointed out that absolute returns not only exhibit long range power law decaying correlations [1–4], but that the typical exponent is not unique [5].

Currency exchange seems to be the natural subject for a statistical description. In fact, the foreign exchange (FX) involves thousands of market participants and there are not a privileged role in the market, such as the *specialist* in an equity market (New York Stock Exchange [6]). Furthermore, the FX is a very liquid market, and since there is an high competition between brokers, prices cannot be easily manipulated. Finally, it is a worldwide never-closing market and then no opening/closing procedures occur.

For these reasons we have decided to analyze a one-year high frequency data set of the Deutsche Mark/US Dollar exchange, the most liquid market. Our data, made available by Olsen and Associates, contains all worldwide 1,472,241 bid–ask Deutsche Mark/US Dollar exchange rate quotes registered by the inter-bank Reuters network over the period Oct. 1, 1992 to Sept. 30, 1993.

The presence of correlations in Deutsche Mark/US Dollar exchange returns is a well-known fact. For example De Jong et al. [7], who consider the same data set we use, show that returns are negatively correlated for about 3 mins. Long correlations for absolute returns in FX have been first pointed out by Müller et al. [8], however, results present a strong daily and weekly seasonality. One of the main problems in tick data analysis, in fact, is the irregular spacing of quotes [8–10].

In this paper we try to characterize some statistical properties common to the whole sample, rather than to focus the attention on daily or intra-daily patterns. Therefore, also in order to minimize the effects of seasonality, we consider business time. This seems to be a reasonable way to consider time in a worldwide time series, where long lags are due to geographical reasons. In this way we obtain the same statistical behavior at different scales, which is also stationary over all the financial series, since the same power law decay for three order of magnitude in time is observed.

The paper is organized as follows. In Section 2, a structure functions analysis shows that, it is not possible to rescale properly the distribution functions at different lags [11]. In Section 3, we perform a direct independence test, using powers of absolute returns. In Section 4, we summarize and discuss the results.

2. Structure functions

Following Ref. [12], we call "random walk" the financial models where the returns

$$r_t \equiv \ln \frac{S_{t+1}}{S_t} \tag{1}$$

are independent variables. We define S_t as the average between bid and ask price. The business time t is the ranking on *all* bid/ask proposals in the period considered.

At present, it is commonly accepted that the variables

$$r_t^{(\tau)} \equiv \sum_{t'=t+1}^{t+\tau} r_{t'} = \ln \frac{S_{t+\tau}}{S_t} , \tag{2}$$

do not behave according to Gaussian for small τ, while the Gaussian behavior is recovered for large τ. A return r_t distributed according to a Levy distribution, as suggested by Mandelbrot [13], is stable under composition and then also $r_t^{(\tau)}$ would follow the same distribution for every τ after a proper rescaling.

A way to study these features, which is common within the fully developed turbulence theory [14], is to study the *structure functions*

$$F_q(\tau) \equiv \langle |r_t^{(\tau)}|^q \rangle , \tag{3}$$

where $\langle \cdot \rangle$ denotes the temporal average on a time window T.

In the simple case where r_t is an independent random process, one has (for a certain range of τ)

$$F_q(\tau) \sim \tau^{\alpha q} , \tag{4}$$

where $\alpha > 1/2$ is the Levy-stable case while the Gaussian behavior corresponds to $\alpha = 1/2$: "random walk" models present always a unique scaling exponent and the process is called self-affine (sometimes uni-fractal). A recent proposal is the truncated Levy distribution model introduced by Mantegna and Stanley [15]. In case of i.i.d. returns, it corresponds to $\alpha > 1/2$ for τ sufficiently small and to $\alpha = 1/2$ for large τ.

In the most general case, instead of (4) one has

$$F_q(\tau) \sim \tau^{\xi_q} , \tag{5}$$

where ξ_q are called scaling exponents of order q. If ξ_q is not linear in q, the process is called multi-affine (sometimes multi-fractal). Using simple arguments it is possible to see that ξ_q must be a convex function of q [16]. The larger is the difference of ξ_q from the linear behavior in q, the wilder are the fluctuations and the correlations of returns. In this sense the deviation from a linear shape for ξ_q gives an indication of the relevance of correlations.

In the FX there is some evidence that the process $r_t^{(\tau)}$ cannot be described in terms of a unique scaling exponent [17], nevertheless seasonality affects these results, due to the choice of calendar time. For example in Ref. [17] the scaling exponent ξ_q does not satisfy the convexity property discussed above: this is an effect of the non-stationarity induced by the irregular spacing of the quotes.

In Fig. 1, we plot the $F_q(\tau)$ for three different values of q. A multi-affine behavior is exhibited by different slopes of $1/q \log_2(F_q)$ vs. $\log_2(\tau)$, at least for the business time τ between 2^4 and 2^{15} (which roughly correspond to few minutes and two weeks in calendar time). In the inset we plot the ξ_q estimated by standard linear regression of $\log_2 F_q(\tau)$ vs. $\log_2(\tau)$ for the values of τ mentioned before.

To give an estimation of errors due to the non-stationarity of the signal, the most natural way turns out to be a partition of the one-year data set in two semesters. This is

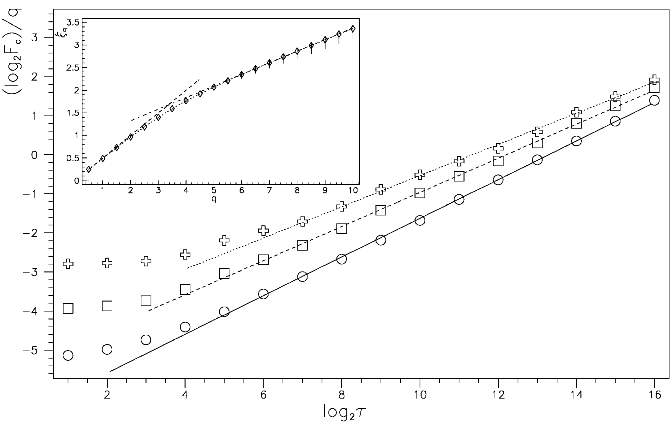

Fig. 1. Structure functions $1/q \log_2 F_q(\tau)$ vs. $\log_2 \tau$ for Deutsche Mark/US Dollar exchange rate quotes. The three plots correspond to different value of $q: q=2$ (○), $q=4$ (□) and $q=6$ (+). In the inset we show ξ_q vs. q. We estimate with linear regression two different regions in this graph. The first one is a line of slope 0.5 (dashed line), and the second has a slope 0.256 (dash dotted line).

natural in the financial context, since it is a measure of reliability of the second semester forecasting based on the first one. We observe that the traditional stock market theory (Brownian motion for returns), gives a reasonable agreement with $\xi_q \simeq q/2$ only for $q < 3$, while for $q \geqslant 5$ one has $\xi_q \simeq \alpha q + c$ with $\alpha = 0.256$ and $c = 0.811$. We stress once again that such a behavior cannot be explained by uni-fractal models.

Also notice that plots in Fig. 1 shows that the same phenomenological scaling laws are present at the intra-daily, daily and weekly time scales, by virtue of the absence of the seasonal effects in the business time context.

3. Long term autocorrelation

In recent literature [1–4] the absolute returns series $\{|r_t|\}$ are shown to be long range correlated. The generalized covariances

$$C_q(\tau) \equiv \langle |r_t|^q |r_{t+\tau}|^q \rangle - \langle |r_t|^q \rangle \langle |r_{t+\tau}|^q \rangle \tag{6}$$

are a useful tool to improve the study of this problem [5], since they allow to investigate separately the correlations of absolute returns of different size: small sizes are emphasized by small q, large sizes are emphasized by large q (the usual definition of covariance for absolute returns is recovered for $q = 1$).

Following the definitions of [18], let us suppose to have a long memory for the absolute returns series, i.e., the covariances $C_q(\tau)$ approaches zero very slowly at increasing τ, i.e., $C_q(\tau)$ is a power law

$$C_q(\tau) \sim \tau^{-\beta_q} .$$

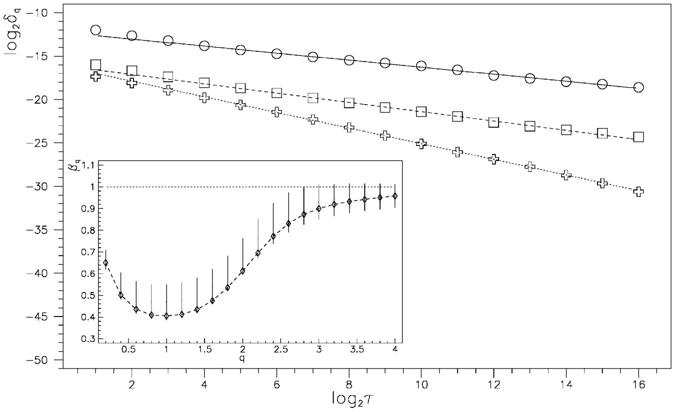

Fig. 2. $\log_2 \delta_q$ vs. $\log_2 \tau$. The three plots correspond to different value of $q: q=1$ (○), $q=1.8$ (□) and $q=3$ (+). In the inset we show β_q vs. q, the horizontal line shows value $\beta=1$ corresponding to independent variable.

Instead of directly computing covariances $C_q(\tau)$ of single returns we consider rescaled sums of returns. This is a well established trick, which is very useful to drastically reduce statistical errors if one is interested only in long term analysis [19]. Therefore, let us introduce the *generalized cumulative absolute returns* [5]

$$\chi_{t,q}(\tau) \equiv \frac{1}{\tau} \sum_{i=0}^{\tau-1} |r_{t+i}|^q \tag{7}$$

and their variance

$$\delta_q(\tau) \equiv \langle \chi_{t,q}(\tau)^2 \rangle - \langle \chi_{t,q}(\tau) \rangle^2 . \tag{8}$$

After some standard algebra (see for example Ref. [5]), one can show that $\delta_q(\tau)$ is a power law with exponent β_q if $C_q(\tau)$ is a power law with the same exponent

$$C_q(\tau) \sim \tau^{-\beta_q} \Rightarrow \delta_q(\tau) \sim \tau^{-\beta_q} \quad \beta_q < 1 .$$

If $|r_t|^q$ is an uncorrelated process one has that $\delta_q(\tau)$ scales with $\beta_q = 1$.

In other words long range memory for absolute returns ($\beta_q < 1$), can be detected via the numerical analysis of $\delta_q(\tau)$.

In Fig. 2 we plot the δ_q vs. τ in log–log scale, for three different values of q. The variance $\delta_q(\tau)$ is affected by small statistical errors, and it confirms the persistence of a long range memory for a business time τ larger than 2^4 and up to 2^{15} (the same lag considered in the structure function analysis).

The exponent β_q can be estimated by standard linear regression of $\log_2(\delta_q(\tau))$ vs. $\log_2(\tau)$, and the errors are estimated in the same way we did for structure functions.

In the inset we plot β_q vs. q for the Deutsche Mark/US Dollar exchange, which exhibits a non-trivial behavior; it is remarkably different from the "random walk" hypothesis of independent variables ($\beta_q = 1$), which could be reasonable only for $q > 3$

(large fluctuations are practically independent). This implies the presence of strong correlations for smaller sizes, furthermore, the exponent of the power law correlations strongly depends on q, which confirms the multi-affine nature of the underlying process.

Let us stress once again that the choice of business time is crucial. In fact, it permits to filter out most of the seasonal effects, and the results show the same behavior over three orders of magnitude in time. This is a strong evidence that ARCH [20] or GARCH [21] process are not adequate to describe FX markets, since they imply an exponential decay for autocorrelations. This remark has been already pointed out by [9], but the use of calendar time keeps seasonal effects in time series and therefore it does not permit to distinguish an exponential from a power law decay.

4. Conclusions

In this paper we have considered the long term anomalies in the Deutsche Mark/US Dollar quotes in the period from October 1, 1992 to September 30, 1993. We have shown the presence of long term anomalies with two techniques: structure functions and a generalization of the usual correlation analysis.

In a round-the-clock market, business time allows one to avoid most of the seasonality in the results: the same statistical features are obtained for three orders of magnitude in time (between few minutes and two weeks) and are stationary over the sample.

In particular, we have pointed out that an exponential decay in autocorrelations, as suggested by ARCH and GARCH models to explain the *clustering of volatility*, cannot describe the power law behavior observed in autocorrelations.

Furthermore, we have shown that "random walk" models (or other uni-fractal models) cannot describe these features. The new approach allows to distinguish clearly uni-fractal and multi-fractal behavior (different scaling of different absolute returns), while the usual one (based on the use of calendar time) fails in detecting a different exponent in the decay of autocorrelations for absolute value and squared returns (See e.g. Ref. [8]).

In conclusion, since most of the seasonality is filtered out, structure functions and generalized correlations are very clearly computed, and therefore business time seems to be the proper unit of time in FX markets.

Acknowledgements

We thank Luca Biferale and Rosario Mantegna for many useful discussions at an early stage of this research. R.B. thanks the European Community TMR Grant: "Financial Market Efficiency".

References

[1] Z. Ding, C.W.J. Granger, R.F. Engle, A long memory property of stock market returns and a new model, J. Empirical Finance 1 (1993) 83.

[2] R.T. Baillie, Long memory processes and fractional integration in economic and finance, J. Econometrics 73 (1996) 5.

[3] P. De Lima, N. Crato, Long range dependence in the conditional variance of stock returns, Econ. Lett. 45 (1994) 281.

[4] A. Pagan, The econometrics of financial markets, J. Empirical Finance 3 (1996) 15–102.

[5] M. Pasquini, M. Serva, Multiscale behavior of volatility autocorrelations in a financial market, Econ. Lett. 65 (1999) 275–279.

[6] Bank of International Settlements, Triennial Central Bank Survey of Foreign Exchange and Derivatives Market Activity, 1996, pp. 1–50.

[7] F. De Jong, R. Mahieu, P. Schotman, Price discovery in the foreign exchange market: an empirical analysis of the yen/dmark rate, J. Int. Money Finance 17 (1998) 5–27.

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

[9] M.M. Dacorogna, U.A. Müller, R.J. Nagler, R.B. Olsen, O.V. Pictet, A geografical model for the daily and weekly seasonal volatility in the foreign exchange market, J. Int. Money Finance 12 (1993) 51–63.

[10] T.G. Andersen, T. Bollerslev, Deutsche Mark–Dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies, J. Finance 53 (1998) 219–265.

[11] B.B. Mandelbrot, Fractals and Scaling in Finance, Springer, Berlin, 1997.

[12] E.F. Fama, Efficient capital markets: a review of theory and empirical work, J. Finance 25 (1970) 383–417.

[13] B.B. Mandelbrot, The variation of certain speculative prices, J. Business 36 (1963) 394–419.

[14] U. Frisch, Turbulence: the Legacy of Kolmogorov, Cambridge University Press, Cambridge, 1995.

[15] R. Mantegna, H.E. Stanley, Scaling behavior in the dynamics of an economics index, Nature 376 (1995) 46–49.

[16] W. Feller, An Introduction to Probability Theory and its Applications, Wiley, New York, 1971.

[17] S. Ghashghaie, W. Breymann, J. Peinke, P. Talkner, Y. Dodge, Turbulent cascades in foreign exchange markets, Nature 381 (1996) 767–770.

[18] P.J. Brockwell, R.A. Davis, Time Series: Theory and Methods, Springer, New York, 1991.

[19] A. Papoulis, Probability, Random Variables and Stochastic Processes, McGraw-Hill, New York, 1977.

[20] P.F. Engle, Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation, Econometrica 50 (1982) 987–1008.

[21] T. Bollerslev, Generalized autoregressive conditional heteroskedasticity, J. Econometrics 31 (1986) 307–327.