HSC Research Report

HSC/98/02

Scaling in currency exchange: A Conditionally Exponential Decay approach

Szymon Mercik*
Rafał Weron**

* Institute of Physics, Wrocław University of Technology, Poland
** Institute of Mathematics, Wrocław University of Technology, Poland

Scaling in currency exchange:
A Conditionally Exponential Decay approach

Szymon Mercik[a], Rafal Weron[b,1]

[a]Institute of Physics,
[b]Institute of Mathematics,
Wroclaw University of Technology, 50-370 Wroclaw, Poland

Abstract:

We use the Conditionally Exponential Decay (CED) model to explain the scaling behavior in currency exchange (FX) rates. This approach enables us not only to show that FX returns satisfy scaling with an exponent qualitatively different from that of a random walk, but also to identify the distributions of these returns corresponding to the empirical scaling laws. The study is conducted via three different estimation methods and using intra-daily FX data which offers the great advantage of large samples and high significance.

1   Introduction

Typical data sets employed by economists do not exceed a few hundred or thousand observations per series. However, in the last decade data sets containing tick-by-tick observations have become available. The studies of these data have turned up new and interesting facts about the pricing of assets. For example, it has been recognized recently that financial markets display scaling properties [1, 2, 3, 4, 5] similar to those of complex systems found in diverse areas of science [6, 7, 8, 9].

In this paper we show that currency exchange (FX) rate returns satisfy scaling with an exponent significantly different from that of a random walk. But what is more important, we also show that the Conditionally Exponential Decay (CED) model [10, 11] can be used to solve a long standing problem in the analysis of intra-daily data [3, 12], i.e. it can be used to identify the mathematical structure of the distributions of FX returns corresponding to the empirical scaling laws.

The CED model is based on asymptotic behavior of complex stochastic systems and current developments of chaos theory. In particular, it is consistent with the Fractal and the Heterogeneous Market Hypotheses [1, 12, 13] which emphasize the

---

[1]Corresponding author; e-mail: rweron@im.pwr.wroc.pl, fax: +48 71 320-26-54.

impact of information and investment horizons on the behavior of investors. The basic assumptions of these hypotheses are the following: the market is made up of many individuals with a large number of different investment horizons and is heterogeneous in the geographic location of the participants; information has a different impact on different investment horizons; prices reflect a combination of short-term technical trading and long-term fundamental valuation.

The CED model clarifies the ideas of the Fractal and the Heterogeneous Market Hypotheses and provides a rigorous mathematical framework for further analysis of financial complex processes. In this model [5, 11] each ith investor is related to a cluster of agents acting simultaneously on common markets. The influence of this cluster of agents is of the type of short-range interactions and is reflected by a random risk-aversion factor A_i. Interactions of the long-range type are imposed on the ith investor by the inter-cluster relationship manifested by random risk factors B_j^i for all $j \neq i$. They reflect how fast the information flows to the ith investor. These assumptions lead to the following CED probability density of returns

$$f(r) = \alpha\lambda(\lambda r)^{\alpha-1}\left[1 - \exp\left(-\frac{(\lambda r)^{-\alpha}}{k}\right)\right]\exp\left[-\frac{1}{k}\int_0^{k(\lambda r)^\alpha}\left(1 - e^{-1/s}\right)ds\right], \quad (1)$$

for $r > 0$, which exhibits the two power-laws property

$$f(r) = \begin{cases} C_1(\lambda r)^{\alpha-1} & \text{for} \quad 0 < \lambda r \ll 1, \\ C_2(\lambda r)^{-\frac{\alpha}{k}-1} & \text{for} \quad \lambda r \gg 1, \end{cases} \quad (2)$$

where $C_1 = \alpha\lambda$ and C_2 is a function of all three CED parameters: α – the shape, λ – the scale, and k – a parameter that decides how fast the information flow is spread out in the market.

Now, that we have a model which can be used to identify the mathematical structure of the distributions of FX returns corresponding to the empirical scaling laws, the basic question to ask is: *How can we fit the CED model to financial data?* In what follows we present three different estimation methods and use the best one in a detailed analysis of the FX data.

But before we start we want to emphesize the fact that the CED probability density is defined only on the positive half line. As a consequence, in all three presented methods, we have to carry out the same analysis for positive returns (CED$^+$) and then for absolute value of negative returns (CED$^-$). This results in obtaining two sets of estimators: $\{\widehat{\alpha}^+, \widehat{\lambda}^+, \widehat{k}^+\}$ and $\{\widehat{\alpha}^-, \widehat{\lambda}^-, \widehat{k}^-\}$. To make notation simpler in the next two sections we describe how to obtain generic estimators $\{\widehat{\alpha}, \widehat{\lambda}, \widehat{k}\}$ without specifying if we are using positive or absolute value of negative returns.

One might criticise the CED model by saying that it is possible to approximate pretty well any empirical distribution with a six parameter law. However, as we will show later, we can reduce the number of parameters to a three element set: $\{\widehat{\lambda}, \widehat{k}^+, \widehat{k}^-\}$, because $\widehat{\alpha}^+ = \widehat{\alpha}^- = 1$ and $\widehat{\lambda}^+ = \widehat{\lambda}^- = \widehat{\lambda}$ both for positive and for negative FX returns.

2   Direct approach

Probably the most natural method of estimating α, λ and k makes use of the two power-laws property (2) of the CED density. In this method we have to calculate the

kernel density estimator $\widehat{f}(r)$ of the empirical returns and then plot it on a double logarithmic paper (log-log plot), see Fig. 1. If the empirical density of returns is unimodal then to reduce estimation errors we first have to center the density around its mode. Otherwise it would be impossible to "glue" the positive and the negative part of the CED density.
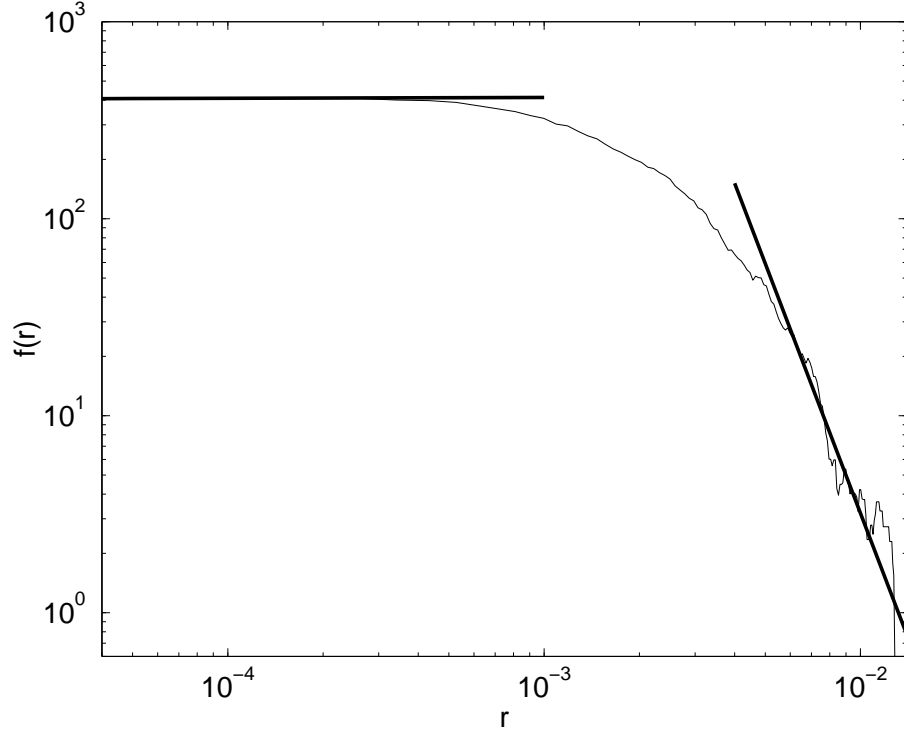

Figure 1: Double logarithmic plot of the empirical density (kernel estimator) of the USD/JPY exchange rate 150 minute returns. Bold lines represent the two power laws. For positive returns the direct approach yielded the following values: $\alpha^+ = 1.0044$, $\lambda^+ = 412.78$ and $k^+ = 0.3130$.

If the empirical data follows the CED law than the plot should be linear both for small and for large λr. Namely, if for $0 < \lambda r \ll 1$ the best linear fit of $\widehat{f}(r)$ is given by

$$\log\left(\widehat{f}(r)\right) = a + b\log(r),$$

than comparing it with the logarithm of (2) we obtain estimators of α and λ

$$\widehat{\alpha} = b + 1, \qquad \widehat{\lambda} = \left(\frac{\exp(a)}{b+1}\right)^{1/(b+1)}. \tag{3}$$

Similarly, if for $\lambda r \gg 1$ the best linear fit is given by

$$\log\left(\widehat{f}(r)\right) = c + d\log(r),$$

than comparing it with the logarithm of (2) we obtain

$$\widehat{k} = -\frac{\widehat{\alpha}}{d+1}. \tag{4}$$

3

Unfortunately some serious problems arise when we apply this method. Namely, we need a precise, yet as smooth as possible, estimate of the empirical density. This can be done – only up to a certain degree – by manipulating the window and the kernel of the density estimator. For a given sample $R = (R_1, R_2, \ldots, R_n)$, e.g. 30 minute returns of a specified exchange rate, and for any real r the *kernel density estimator* [14, 15] provides an approximate value of the density in the form

$$\widehat{f}_n(r) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{b_n} K\left(\frac{r - R_i}{b_n}\right).$$

The kernel $K(u)$ is a continuous, nonnegative and symmetric function satisfying

$$\int_{-\infty}^{\infty} K(u)du = 1,$$

whereas the window $\{b_n\}_{n \in \mathbf{N}}$ is a sequence of positive real numbers such that $\lim_{n \to \infty} b_n = 0$ and $\lim_{n \to \infty} nb_n = \infty$. In our estimation procedures we used the Bertlett kernel

$$K(u) = \frac{3}{4}(1 - u^2)\mathbf{1}_{[-1,1]}(u),$$

with $b_n = 2.3\sigma n^{-1/5}$, where σ is the standard deviation of the sample, as it seemed to give the best results.

Another weakness of the presented method is caused by the fact that a small number of observations in the tails of the distribution (i.e. for $0 < \lambda r \ll 1$ and for $\lambda r \gg 1$) may introduce a large bias to the estimators. Ironically, to estimate α, λ and k we only use values from the tails of the distribution! Thus we lose most of the information carried by the sample. Moreover, there is still a problem of selecting data points for the linear fit on the log-log plot. We were unable to automate this procedure and the estimators depended on visual inspection.

3   Maximum likelihood method

To overcome weak points of the direct approach we turned to more classical estimation methods: minimization of distance in the L^p norm and the maximum likelihood method. In our case the former one reduces to finding a minimum of a function of three variables

$$\int_0^\infty \left|\widehat{f}(r) - f(r; \alpha, \lambda, k)\right|^p dr. \tag{5}$$

Unfortunately, it also uses an estimate of the empirical density $\widehat{f}(r)$ and thus is subject to unnecessary estimation errors.

On the other hand the maximum likelihood method is free of this flaw. In short, it is a recipe for producing an estimator $\widehat{\theta}$ of the vector of parameters $\theta = (\alpha, \lambda, k)$, called the *maximum likelihood estimator* (MLE). The MLE is defined as an estimator such that it is the value of the argument θ which maximizes the likelihood function

$$L_\theta(R_1, \ldots, R_n) = \prod_{k=1}^{n} f(R_k),$$

4

where $R = (R_1, \ldots, R_n)$ is the sample and f is the probability density, as a function of θ.

Notice that $L_\theta(R)$ obtains a maximum exactly for the same values of θ as the so called log-likelihood function

$$\log L_\theta(R_1, \ldots, R_n) = \sum_{k=1}^{n} \log f(R_k).$$

As it happens in our case, this observation usually leads to much easier maximization algorithms. Namely, from formula (1) we can calculate the log-likelihood function of the CED model

$$
\begin{aligned}
\log L_\theta(R_1, \ldots, R_n) \; = \; & \sum_{k=1}^{n} \left\{ \log \alpha \lambda^\alpha + (\alpha - 1) \log(R_k) + \log \left[1 - \exp \left(-\frac{(\lambda R_k)^{-\alpha}}{k} \right) \right] \right\} \\
& - \sum_{k=1}^{n} \left\{ \frac{1}{k} \int_0^{k(\lambda R_k)^\alpha} \left(1 - e^{-1/s} \right) ds \right\}.
\end{aligned}
$$

Then using the Nelder-Mead simplex minimization procedure (MATLAB implementation) applied to the function

$$- \log L_\theta(R_1, \ldots, R_n)$$

we obtain estimates of α, λ and k. Note that this method uses information carried by all returns and not only those in the tails of the distribution. This is illustrated in Fig. 2 where we plot the empirical and the approximating CED densities (using both direct and maximum likelihood estimation) for 150 minute returns of the USD/JPY exchange rate. For positive returns the maximum likelihood estimation yielded the following values: $\alpha^+ = 1.0413$, $\lambda^+ = 500.75$ and $k^+ = 0.1564$. Clearly, MLEs are much better than those obtained using the direct approach, see Fig. 1, or even those obtained by minimizing the distance in the L^p norm, see eq. (5). This superiority of the maximum likelihood estimation is indeed true in general, i.e. for almost all FX rates and sampling intervals (Δt's).

4   Empirical analysis

The empirical studies were conducted on a data set released by Olsen & Associates for the Second International Conference on High Frequency Data in Finance, Zurich, April 1-3, 1998. The data set included exchange rates of all major currencies from January 1, 1996 to December 31, 1996. The data came in files where GMT time and FX rates were reported sequentially in 30 minute intervals, thus the number of data was 17520 for each exchange rate. As in [3], data are quotations of foreign currencies available from international vendors like Reuters, Knight-Ridder and Telerate, and do not correspond to real prices in the global FX market. The actual deals are usually made over the telephone and the transaction prices may differ from the offers or even no transactions may take place at the offered prices.

One of the main features of the FX spot market is the fact that it is a 24 hours global market, which is mostly inactive during weekends and national holidays. The
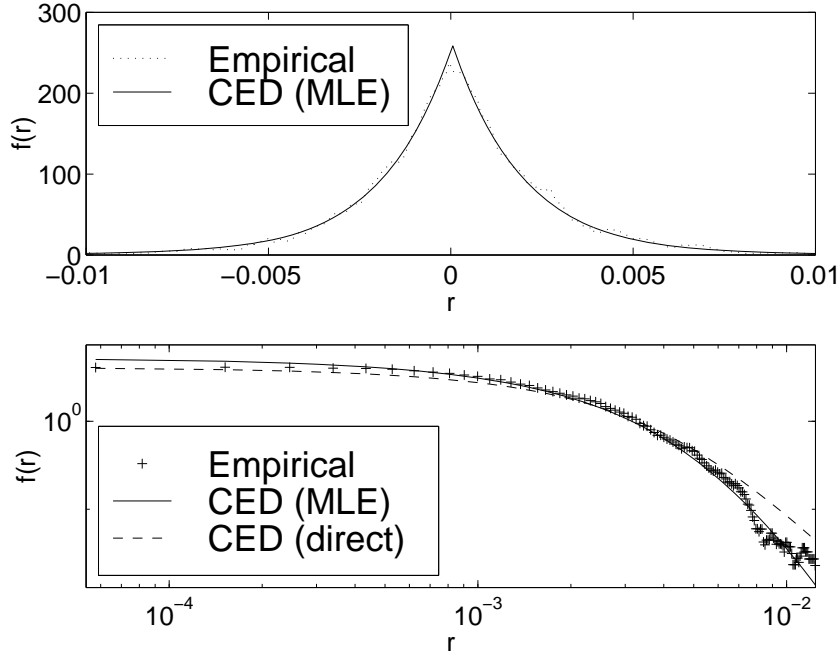
Figure 2: Plot of the empirical density for returns of the USD/JPY exchange rate and the approximating CED density (MLE) for $\Delta t = 150$ minutes (*top*). Log-log plot of the above densities and the approximating CED density (obtained using the direct approach) for positive returns only (*bottom*).

first observation of the week arrives at 22:30 Greenwich Mean Time (GMT) on Sunday with the opening of the Asian markets and the last observation comes from the West Coast of the USA at about 22:30 GMT on Friday [12], see Fig. 3. In contrast to traditional high frequency analysis [1], we exclude inactive periods from the calculations, because they introduce a large bias to the estimators and make comparison of scaling laws for different instruments much more difficult. In the second column of Table 1 we give the percent of zero returns in the data, which corresponds to the inactivity of a given market. The inactivity ranges from 20% for the most actively traded in 1996 exchange rate – USD/DEM, to 48% for the DEM/FIM exchange rate.

Analysis of high frequency (intra-daily) data relies on definitions of the variables under study – in our case – the price and the volatility. Although these definitions are probably well known, we give them for the sake of completeness. The (logarithmic) price at time t_i is defined as [11, 12]

$$x(t_i) \equiv x(t_i, \Delta t) \equiv \frac{1}{2}\Big(\log p_{bid}(t_i) + \log p_{ask}(t_i)\Big),$$

where $\{t_i\}$ is the sequence of the regular spaced in time data, Δt is the time interval ($\Delta t = 30$ min., $\Delta t = 1$ hour, etc.) and $p_{bid}(t_i)$ ($p_{ask}(t_i)$) is the arithmetic average of the bid (ask) quotes just prior to and just after time t_i. The definition takes the average of the bid and ask price rather than either the bid or the ask series as a better approximation of the transaction price. The volatility $v(t_i)$ at time t_i is
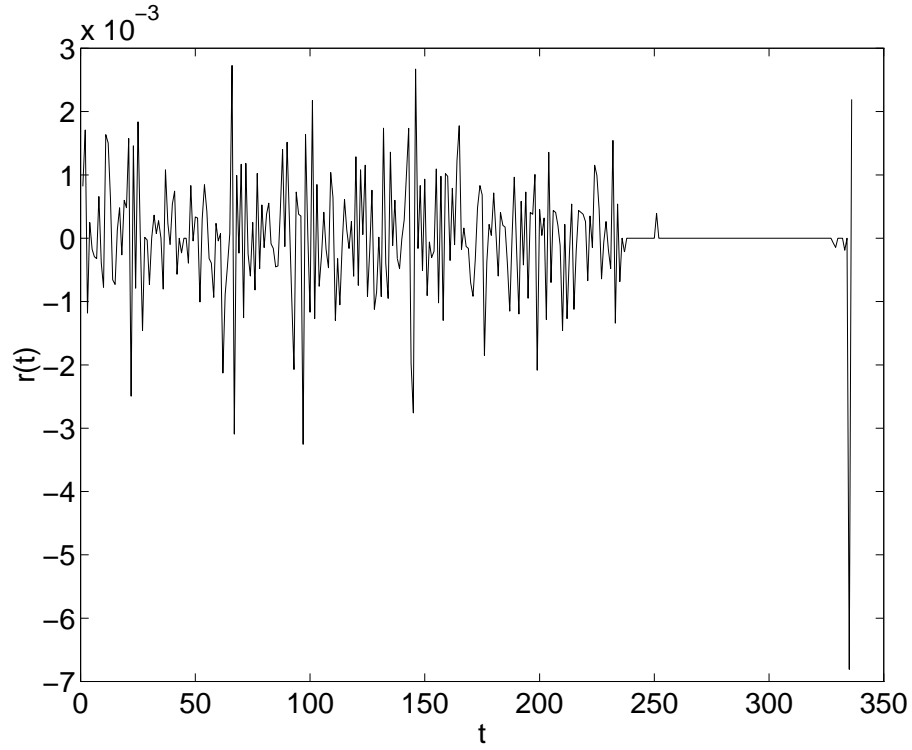
6

Figure 3: Half hour returns of the DEM/FRF exchange rate during the second week of 1996, i.e. January 8th till January 14th. Note that one day consists of 48 returns.

Table 1: Empirical and CED drift exponents for the twelve analyzed FX rates.

FX rate	$\frac{N_0}{N}$	D	D_0	$\frac{D_0-D}{D}$	c_λ	D_{CED}	$\frac{D_{CED}-D}{D}$
USD/DEM	20%	0.4614	0.5117	10.9%	0.4777	0.4595	-0.4%
GBP/USD	22%	0.4226	0.4783	13.2%	0.4172	0.4209	-0.4%
USD/JPY	22%	0.4308	0.4859	12.8%	0.4197	0.4300	-0.2%
USD/CHF	23%	0.4621	0.5200	12.5%	0.4788	0.4564	-1.2%
AUD/USD	25%	0.4182	0.4835	15.6%	0.4159	0.4171	-0.3%
DEM/JPY	26%	0.4454	0.5138	15.4%	0.4475	0.4419	-0.8%
GBP/DEM	27%	0.4835	0.5485	13.4%	0.4534	0.4682	-3.2%
USD/FRF	27%	0.4232	0.4921	16.3%	0.4281	0.4172	-1.4%
CAD/USD	28%	0.4046	0.4720	16.7%	0.4115	0.3934	-2.8%
DEM/FRF	29%	0.3249	0.4022	23.8%	0.3477	0.3109	-4.3%
DEM/ITL	29%	0.4334	0.5024	15.9%	0.4519	0.4309	-0.6%
DEM/FIM	48%	0.3426	0.4817	40.6%	0.3527	0.3388	-1.1%

defined as the average of unsigned changes of the logarithmic price (return)

$$v(t_i) \equiv v(t_i, \Delta t, T) \equiv \frac{1}{n} \sum_{k=1}^{n} \left| x(t_{i-k}) - x(t_{i-k} - \Delta t) \right|, \tag{6}$$

where T is the sample period on which the volatility is computed (e.g. one day, one year) and n is a positive integer with $T = n\Delta t$.

A statistical study of financial data from the fractal point of view is based on the analysis of time intervals Δt of different sizes. A reasonable question to ask is: *What is the relation between volatility and the size of time intervals?* The answer to this question is the scaling law for volatility [1, 11]

$$v(t_i) = c(\Delta t)^D, \tag{7}$$

where c is an empirical constant and D is the empirical drift exponent. In spite of its elementary nature, a scaling law study is immediately able to reject the Gaussian hypothesis and reveal an important property of financial time series. For the Gaussian case the above formula is true with a drift exponent of 0.5. In the third and fourth column of Table 1 we give the empirical drift exponent for data without inactive periods (D) and for data with inactive periods (i.e. with zero returns, D_0), whereas the difference (in percent) between these two exponents is given in the fifth column. We can clearly see that for all FX rates exponent D is significantly smaller than 0.5. Moreover, the difference between D and D_0 increases with the number of zero returns, thus showing that the inclusion of inactive periods causes overestimation of the drift exponent.

Intensive studies of all twelve FX rates and all 48 sampling intervals ($\Delta t = 30, 60, ..., 1440$ minutes) lead us to the following conclusions:

- The shape parameter α is equal to one for positive and for absolute value of negative returns. This is illustrated in Fig. 4 for the DEM/FRF exchange rate. Moreover, in Table 2 the mean values of α^+ and α^- for all FX rates are given.

- The scale parameter λ is the same for positive and for absolute value of negative returns. Moreover, it is related to the time interval by a power law

$$\lambda(\Delta t) = \Lambda(\Delta t)^{-c_\lambda}.$$

This is illustrated in Fig. 5 for the USD/CHF exchange rate, separately for positive (λ^+) and for absolute value of negative returns (λ^-). The straight line represents the best linear fit (linear regression) jointly for all returns. The slope is given by $-c_\lambda = -0.4788 \pm 0.0102$ and the coefficient of determination $R^2 = 0.9587$. Drift exponents $c_{\lambda+}$ and $c_{\lambda-}$ of the scale parameters λ^+ and λ^-, respectively, are given in Table 2. The values of these exponents are almost equal and their mean value c_λ closely approximates the empirical drift exponent D, see Table 1.

- As a consequence of the two above observations the CED density has finite and non-zero limits at point zero. Moreover, the left-hand limit is equal to the

8

Table 2: Mean values of the shape parameters α^+ and α^- and drift exponents of the scale parameters λ^+ and λ^- for the twelve analyzed FX rates.

FX rate	$\langle \alpha^+ \rangle$	$\langle \alpha^- \rangle$	c_{λ^+}	c_{λ^-}
USD/DEM	1.0442	1.0030	0.4743	0.4811
GBP/USD	0.9999	1.0084	0.4368	0.3998
USD/JPY	1.0185	0.9991	0.4470	0.4018
USD/CHF	1.0093	0.9946	0.4939	0.4637
AUD/USD	1.0374	1.0568	0.4374	0.3958
DEM/JPY	1.0686	1.0273	0.4651	0.4313
GBP/DEM	1.0095	0.9884	0.4964	0.4234
USD/FRF	1.0554	1.0343	0.4248	0.4323
CAD/USD	0.9802	1.0190	0.4083	0.4146
DEM/FRF	0.9923	0.9734	0.3446	0.3508
DEM/ITL	0.9980	1.0062	0.4367	0.4678
DEM/FIM	0.9589	0.9936	0.3375	0.3679

right-hand limit and we can make the CED density continuous on the whole real line by setting $f(0) = \lambda$. This is visualized in Fig. 5, where values of the kernel density estimator at point zero are almost equal to the λ's. Note that in general kernel estimators flatten approximated functions and as a result for small Δt's we can see that $\widehat{f}(0) < \lambda$.

- The third parameter k, which decides how fast the information flow is spread out in the market, is qualitatively different for positive and for absolute value of negative returns. This difference is responsible for the assymetry of the density of returns. Since k defines the tails of the distribution its estimator is very fragile. This causes a large dispersion of estimates and prevents us from identifying k as a function of Δt, see Fig. 6.

We can see from Table 1 that the drift exponent c_λ closely approximates the empirical drift exponent D. Yet we can do much better. Recall that if a random variable R has density $f(r)$ then its mean value is given by

$$\langle R \rangle = \int_A r f(r) dr,$$

where A is the support of $f(r)$. Unfortunately in our case it not clear how to integrate $\int_0^\infty r f(r) dr$, because the CED probability density function itself is a quite complicated function. However, the mean value of R can also be obtained [16] by taking the limit

$$\langle R \rangle = - \left. \frac{d}{dt} \right|_{t=0} \varphi(t),$$

where $\varphi(t)$ is the moment generating function $\varphi(t) = E\{e^{-tR}\}$. This leads us to the
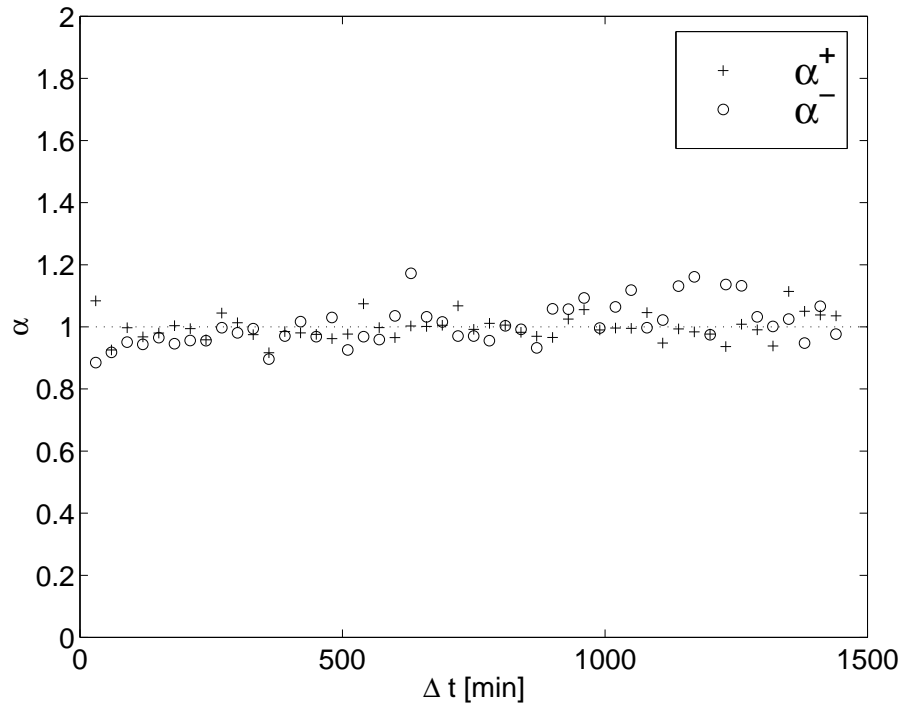
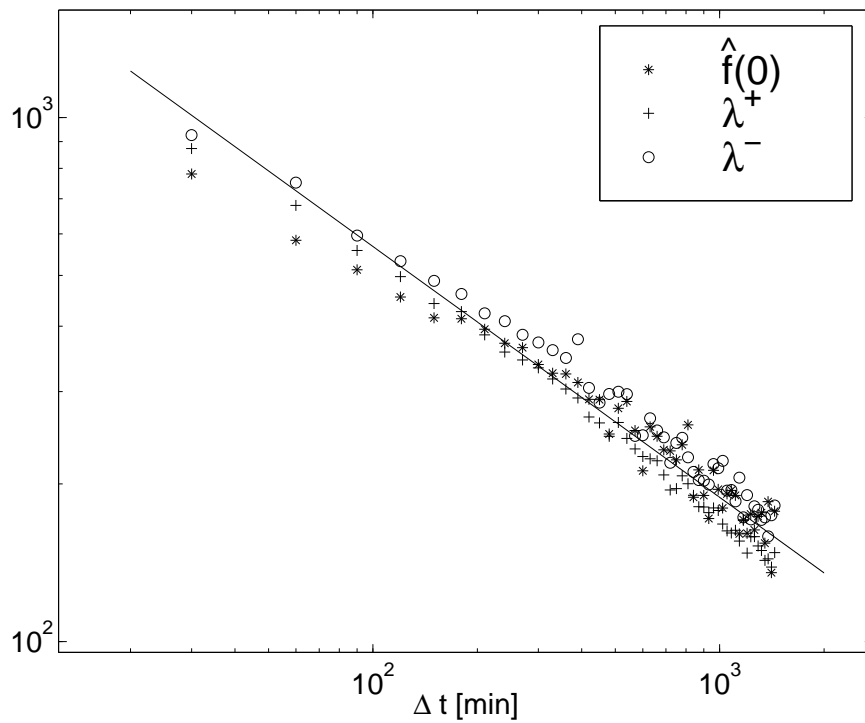Figure 4: Shape parameter α for the DEM/FRF exchange rate and for all Δt's.

Figure 5: Scale parameter λ for the USD/CHF exchange rate and for all Δt's. Values of the kernel density estimator at point zero are added for comparison.
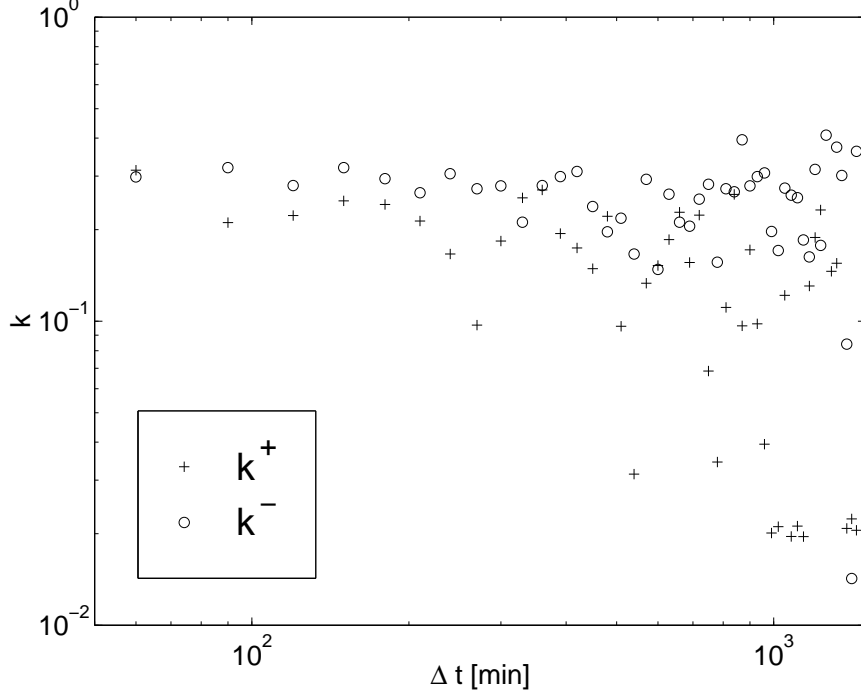
Figure 6: Parameter k for the AUD/USD exchange rate and for all Δt's.

following

$$\langle R \rangle = - \left. \frac{d}{dt} \right|_{t=0} \int_0^\infty f(r) e^{-tr} dr = \frac{1}{\lambda} \int_0^\infty \exp\left\{ -\frac{1}{k} \int_0^{kr^\alpha} \left[1 - \exp(-s^{-1}) \right] ds \right\} dr. \qquad (8)$$

Introducing the function

$$\Psi(\alpha, k) = \int_0^\infty \exp\left\{ -\frac{1}{k} \int_0^{kr^\alpha} \left[1 - \exp(-s^{-1}) \right] ds \right\} dr,$$

we can rewrite (8) as $\langle R \rangle = \frac{1}{\lambda} \Psi(\alpha, k)$ and the mean of the absolute value of return is given by

$$\langle |R| \rangle = \frac{\langle R^+ \rangle + \langle -R^- \rangle}{2}.$$

Analytic analysis of $\Psi(\alpha, k)$ is quite difficult. However, from our earlier empirical studies we know that $\alpha = 1$. This simplifies things and in Fig. 7 we can see Ψ as a function of $1/k$. Clearly for small k (large $1/k$) $\Psi(k)$ is very close to one. This justifies our earlier use of c_λ as an approximation of the empirical drift exponent D. But if we include $\Psi(k)$ in our calculations then the approximation is even better, see Table 1 where the CED estimate D_{CED} and the difference in percent between D_{CED} and D is given in the last two columns. This is also illustrated in Fig. 8, where almost a perfect match is obtained for the USD/FRF exchange rate.
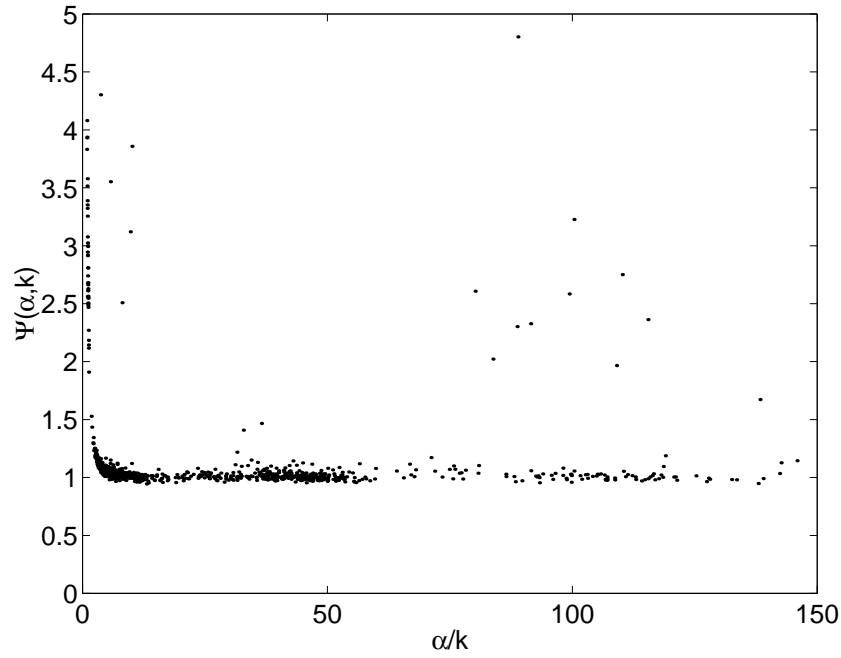
11

Figure 7: Function $\Psi(\alpha, k)$ for all twelve exchange rates and for all Δt's.
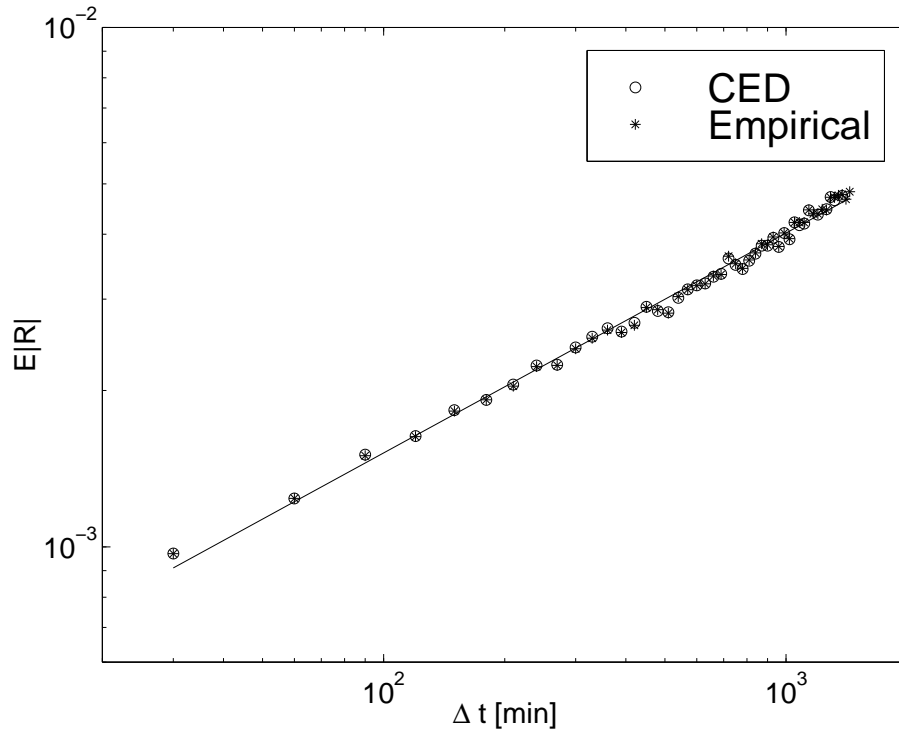

Figure 8: Scaling law for volatility. A comparison of the CED and empirical volatilities for the USD/FRF exchange rate and for all Δt's. The straight line represents the best linear fit to the CED volatilities.

References

[1] U.A. Müller, M.M. Dacorogna, R.B. Olsen, O.V. Pictet, M. Schwarz, C. Morgenegg, J. Banking & Finance 14 (1990) 1189.

[2] R.N. Mantegna, H.E. Stanley, Nature 376 (1995) 46.

[3] S. Galluccio, G. Caldarelli, M. Marsili, Y.-C. Zhang, Physica A 245 (1997) 423.

[4] P. Gopikrishnan, M. Meyer, L.A.N. Amaral, H.E. Stanley, Euro. Phys. J. B 3 (1998) 139.

[5] A. Weron, Sz. Mercik, R. Weron, Physica A (1998) to appear.

[6] G.K. Zipf, Human Behavior and the Principle of Least Effort, Addison-Wesley, Reading MA, 1949.

[7] K. Weron, A. Jurlewicz, J. Phys. A: Math. Gen. 26 (1993) 395.

[8] A.K. Jonscher, Universal Relaxation Law, Chelsea Dielectrics, London, 1996.

[9] I. Peterson, The Jungles of Randomness, Viking, New York, 1997.

[10] A. Jurlewicz, A. Weron, K. Weron, Appl. Math. 23 (1996) 379.

[11] R. Weron, K. Weron, A. Weron, Physica A (1998) to appear.

[12] D.M. Guillaume, M.M. Dacorogna, R.R. Dave, U.A. Müller, R.B. Olsen, O.V. Pictet, Finance Stochast. 1 (1997) 95.

[13] E.E. Peters, Fractal Market Analysis: Applying Chaos Theory to Investment and Economics, Wiley & Sons, New York, 1994.

[14] L. Devroye, A Course on Density Estimation, Birkhäuser, Boston, 1987.

[15] A. Janicki, A. Weron, Simulation and Chaotic Behavior of α-Stable Stochastic Processes, Marcel Dekker, New York, 1994.

[16] W. Feller, An Introduction to Probability Theory and Its Applications, 2nd edn. Wiley, New York, 1971.

HSC Research Report Series 1998

For a complete list please visit http://ideas.repec.org/s/wuu/wpaper.html