Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Physica A 344 (2004) 194–197
PHYSICA A
www.elsevier.com/locate/physa

Real prices from spot foreign exchange market

Filippo Petroni*, Maurizio Serva

Dipartimento di Matematica and INFM, Universitá dell'Aquila, via Vetoio, 67010 L'Aquila, Italy

Received 11 December 2003
Available online 23 July 2004

Abstract

In this work we discuss the problem of price definition when using high frequency foreign exchange data. If one uses the spot mid price a strong autocorrelation of returns, at one lag, is found which is only due to microstructure effect and does not capture the real behavior of price dynamics. This autocorrelation increases the intraday volatility estimated from this type of data. To solve this problem we introduce an algorithm which is able, by using the no-arbitrage principle, of eliminating every microstructure effects.

PACS: 89.65.Gh; 05.45.Tp

Keywords: Econophysics; Exchange market; High frequency

We analyze Deutsche mark/US dollar exchange quotes taken from Reuters' EFX pages (the dataset has been provided to us by Olsen & Associates) during a period of one year from January to December 1998. In this period 1,620,843 quotes entries in the EFX system were recorded. The dataset provides a sequence of bid and ask exchange quotation pairs from individual institutions whose names and locations are also recorded. The dataset does not contain any information on traded volume and on the lifetime of quotes. Furthermore, these quotes are indicative and they do not imply that any amount of currency has been actually traded. If one assumes that the asset price is the bid or ask quote, or an average of them, one finds an indeterminacy

*Corresponding author. Tel.: +39-086-341-5841.
E-mail address: fil_petroni@tin.it (F. Petroni).

0378-4371/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2004.06.115

[1] at very short time scale and anticorrelated returns at one time step [2–6]. One possible explanation for this indeterminacy is in the presence of a noise contribution due to microstructure effects. In this work we introduce an algorithm which, reducing the spread between bid and ask quotes, is able to determine the real price from the observed one and solve the indeterminacy. The key of our work is that we are able to do so with a parameter free algorithm which uses only the no-arbitrage principle [7].

Throughout this work we use business time (basically a tick time) as our time flow index [8]. We indicate with $S_t^{(b)}$ and $S_t^{(a)}$, respectively, bid and ask quotes at time t. We consider mid price as given by the geometric average of bid and ask quotes $S_t = \sqrt{S_t^{(a)} \times S_t^{(b)}}$ [4]. We stress that the same results can be found if bid or ask quotes are used [2]. It has been suggested [1,9] that the mid price is the composition of two different stochastic processes: a real price change and a noise contribution which is the result of microstructure effect. Given that S_t is the mid price at business time t we can express the two contributions as $S_t = \tilde{S}_t e^{\varepsilon_t}$ where $\tilde{S}_t$ is the real price and ε_t is the error contribution to the real price ($\varepsilon_t \equiv \ln(S_t/\tilde{S}_t)$). The relation between returns is then given by $r_t(1) = \tilde{r}_t(1) - \varepsilon_t + \varepsilon_{t+1}$, where returns are defined as $r_t(\tau) = \ln(S_{t+\tau}/S_t)$, $\tilde{r}_t(\tau) = \ln(\tilde{S}_{t+\tau}/\tilde{S}_t)$. To verify the validity of this approach we estimated, from real data, the following statistical quantities: the τ-dependent variance of returns defined as

$$\langle r_t^2(\tau)\rangle = 2\langle \varepsilon_t^2\rangle + \langle \tilde{r}_t^2(1)\rangle\tau \, , \tag{1}$$

where $\langle .. \rangle$ indicates an average over the probability distribution. It has been assumed that ε_t and $\tilde{r}_t(1)$ are uncorrelated random variables. The non-overlapping first-order covariance of two consecutive returns after τ business time

$$\langle r_{t+s}(\tau) r_t(\tau)\rangle = -\langle \varepsilon_t^2\rangle \tag{2}$$

$s = 1$, and the non-overlapping higher-order covariances of returns

$$\langle r_{t+s}(\tau) r_t(\tau)\rangle = 0 \tag{3}$$

with $s > 1$. The above picture corresponds exactly to what one can see in Fig. 1 in fact the variance of returns is a linear function of time lags τ, as expected, and it is different from zero in the limit $\tau \to 0$. This implies the existence of an implicit indeterminacy in the price estimation for vanishing time lags. The same indeterminacy is responsible for the negative autocorrelation of two consecutive returns at one time lag (results not shown) [2–6]. We stress that the correlation we find here is not due to the bid-ask spread [10,11] given that in our analysis we are not using the transaction price but the mid-quote. Even the discreteness [12,13] of prices cannot be invoked to explain the increase of volatility and this because in our dataset price changes are restricted to less than 1×10^{-4} of the actual price and this effect, if exist, is very small [2,5,6].

In this work we find an algorithm which is able to separate the two contributions in the mid price without any assumption on the nature of the mid quote return process and without fixing any arbitrary parameters [2,4,5]. As a result we obtain the

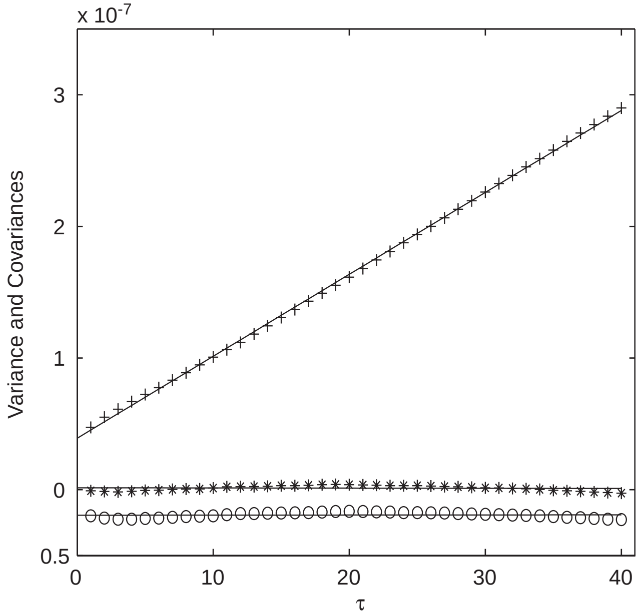

Fig. 1. DEM/USD spot exchange rates: variance (1)(crosses) compared with a linear fit $2\mathrm{A}+\mathrm{B}\tau$, non-overlapping first-order covariance (2)(circles) compared with -A, non-overlapping higher-order covariance (3)(stars) compared with zero. A and B are identified with $\langle \varepsilon_t^2 \rangle$ and $\langle \tilde{r}_t^2 \rangle$.

real prices sequence. If the dataset contained information on the time duration of each quote, or the life time of each quote was a known constant, there would be no problem in establishing real price at each time: it would be the best bid and ask quotes valid at that time. But this information is not available and a different strategy has to be found to establish real price at each time. The algorithm we propose is the following: let us suppose that we are observing the bid and ask price of a given currency and that we are able to detect each quotes from all the financial institution in the business time t. We define the spread between bid and ask as: $D_t = S_t^{(a)} - S_t^{(b)}$. Notice that for the no-arbitrage principle this quantity is greater than or equal to zero. Considering k time lags previous to business time t we consider the following effective spread $D_{t,k} = \min_{i\in\{t-k,t\}} S_i^{(a)} - \max_{i\in\{t-k,t\}} S_i^{(b)}$. For each t our algorithm find $\tilde{k}$ which gives $D_{t,\tilde{k}} \geqslant 0$ and $D_{t,\tilde{k}+1} < 0$. The real price is then given by $\tilde{S}_t = \sqrt{\tilde{S}_t^{(a)} \times \tilde{S}_t^{(b)}}$. Note that the number of steps we have to go backwards in time is only given by the no-arbitrage principle and it is different for every t. We stress that the no-arbitrage principle is needed to fix an upper bound on the life time of each quote given the efficient market hypothesis. Once we have obtained $\tilde{S}_t$ we can define $\tilde{r}_t(\tau) = \ln(\tilde{S}_{t+\tau}/\tilde{S}_t)$ and compute all quantities (variances and correlations) already computed for the observed price. Results for this analysis are presented in Fig. 2. From this figure, it is obvious that the new price definition is not effected by

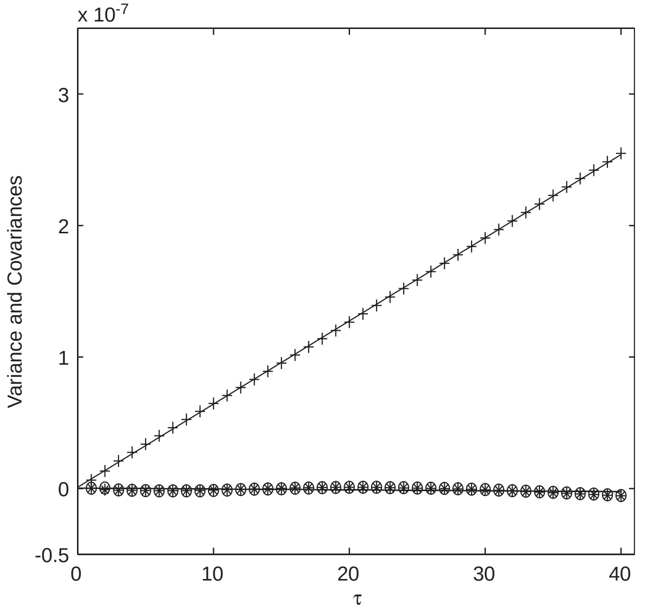

Fig. 2. DEM/USD real exchange rates: variance (1)(crosses) compared with a linear fit $B\tau$, non-overlapping first-order covariance (2)(circles) compared with zero, non-overlapping higher-order covariance (3)(stars) compared with zero, B is identified with $\langle \tilde{r}_t^2 \rangle$.

microstructure effect as for the previous case. Also, the autocorrelation of returns for the real price is uncorrelated at every step.

The idea we have used here is indeed very simple, we assume that old quotes are still valid until they produce arbitrage. In spite of the simplicity we are able to remove all artifacts in the data without the need of any parameter.

References

[1] M. Pasquini, M. Serva, Physica A 277 (1–2) (2000) 228–238.
[2] T. Bollerslev, I. Domowitz, J. Finance 48 (4) (1993) 1421–1443.
[3] F. De Jong, R. Mahieu, P. Schotman, J. Int. Money Finance 17 (1998) 5.
[4] M.M. Dacorogna, R. Genay, U.A. Müller, R.B. Olsen, O.V. Pictet, An Introduction to High-Frequency Finance, Academic Press, New York, 2001.
[5] F. Corsi, G. Zumbach, U.A. Müller, M. Dacorogna, Econ. Notes 30 (2) (2001) 183–204.
[6] B. Zhou, J. Business Econ. Stat. 14 (1) (1996) 45–52.
[7] F. Petroni, M. Serva, Eur. Phys. J. B 34 (4) (2003) 495–500.
[8] L. Berardi, M. Serva, Submitted to IJTAF (2004).
[9] J. Hasbrouck, Rev. Financial Stud. 6 (1) (1993) 191–212.
[10] R. Roll, J. Finance 39 (1984) 1127–1139.
[11] M. Blume, S. Stambaugh, J. Financial Econ. 12 (1983) 387–404.
[12] G. Glottlieb, A. Kalay, J. Finance 40 (1985) 135–153.
[13] L. Harris, J. Financial Quantitative Anal. 25 (3) (1990) 291–306.