Correlation structures in short-term variabilities of stock indices and exchange rates

Tomomichi Nakamura\textsuperscript{a,b,*}, Michael Small\textsuperscript{b}

\textsuperscript{a}Sony Computer Science Laboratories Inc., 3-14-13 Higashigotanda, Shinagawa-ku, Tokyo 141-0022, Japan
\textsuperscript{b}Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

Available online 10 May 2007

Abstract

Financial data usually show irregular fluctuations and some trends. We investigate whether there are correlation structures in short-term variabilities (irregular fluctuations) among financial data from the viewpoint of deterministic dynamical systems. Our method is based on the small-shuffle surrogate method. The data we use are daily closing price of Standard & Poor’s 500 and the volume, and daily foreign exchange rates, Euro/US Dollar (USD), British Pound/USD and Japanese Yen/USD. We found that these data are not independent.

\textcopyright{} 2007 Elsevier B.V. All rights reserved.

Keywords: Correlations structures; Econophysics; Financial data; Surrogate data

1. Introduction

It is usually considered that market prices are interconnected or interrelated in some way or another to varying degrees. That is (for example), change of Euro/US Dollar (USD) exchange rate may influence Japanese Yen/USD exchange rate. The aim of this paper is to examine whether there are correlation structures in short-term variabilities (irregular fluctuations) of stock indices and foreign exchange rates, where all of these data show irregular fluctuations and some trends (See Figs. 1 and 2 in Section 4).

Recently to investigate correlations of price changes, random matrix theory (RMT) is often applied\cite{1–3}. RMT predictions represent an average over all possible interactions\cite{1}. In the approach of RMT the distribution of the eigenvalues of a cross correlation matrix is compared with that of relevant ensemble of the random matrix. Broadly speaking, when there are eigenvalues that are outside the RMT predictions, it is considered that the deviations might suggest the presence of information of the marketwise effects\cite{1}. This result shows that there are correlations among data used in the matrix. This information is useful, but also rather limited, because it does not specifically indicate which data exhibits the correlation. That is, we still do not know which data are related to one another. Moreover, to apply RMT we need to use very large matrices\cite{3}.

\textsuperscript{*}Corresponding author. Current address: Sony Computer Science Laboratories Inc., 3-14-13 Higashigotanda, Shinagawa-ku, Tokyo 141-0022, Japan.
E-mail addresses: tomo@cs.sony.co.jp (T. Nakamura), ensmall@polyu.edu.hk (M. Small).
One of the useful statistics that can directly investigate some kind of relations among data and do not need very large matrices is the cross-correlation function (CC). This statistic can investigate correlations in short-term variabilities of two signals. However, the statistic is also restrictive because correlation is only a useful measure of linear similarity [4]. That is to say, even when two signals are not similar, there are still possibilities that these systems have some kind of correlation structures (that is, the systems are interconnected or interrelated). This means that it is not enough to use only statistics such as CC in all cases. To investigate correlation structures more reliably, an approach from the viewpoint of a deterministic dynamical system is useful and necessary. However, such an approach is not easy because it is difficult to treat data which exhibit short-term variabilities and some trends. Financial data usually have such features.

To accomplish this, we apply a method based on the small-shuffle surrogate (SSS) method [4]. We apply our method to daily closing price of Standard & Poor’s 500 (S&P500) and the volume, and to daily exchange rates, Euro (EUR)/US dollar (USD), British Pound (GBP)/USD and Japanese Yen (JPY)/USD. The analysis of these data is the major novel contribution of this paper.

In Section 2 the SSS algorithm and the hypothesis will be discussed. Our method is based on this algorithm. In Section 3 we present our choice of discriminating statistics and describe how to test our null hypothesis. In Section 4 we apply the method to financial data.

2. The small-shuffle surrogate method

To investigate whether temporal correlations in data are absent or data are independently distributed (ID) random variables even if it exhibits trends, the SSS method is useful [5]. Moreover, the SSS method does not depend on the specific data distribution. The SSS method has proven to be effective for tackling data exhibiting short-term variabilities and long-term trends [5–8].

SSS data are generated as follows; Let the original data be \(x(t)\), let \(i(t)\) be the index of \(x(t)\) (that is, \(i(t) = t\), and so \(x(i(t)) = x(t)\)), let \(g(t)\) be Gaussian random numbers and \(s(t)\) will be the surrogate data.

(i) Obtain \(\hat{i}(t) = i(t) + Ag(t)\), where \(A\) is an amplitude.
(ii) Sort \(\hat{i}(t)\) by the rank-order and let the index of \(\hat{i}(t)\) be \(\hat{i}(t)\).
(iii) Obtain the surrogate data \(s(t) = x(\hat{i}(t))\).

We have found that choosing \(A = 1.0\) is adequate for nearly all purposes. In the SSS data, local structures or correlations in irregular fluctuations (short-term variabilities) are destroyed and the global behaviors (trends) are preserved. Further details of the method and the mechanism are provided in Refs. [4–7]. The null hypothesis (NH) addressed by this algorithm is that irregular fluctuations (short-term variabilities) are ID random variables or time-varying random variables (in other words, there is no short-term dynamics or determinism) [4–7]. The SSS method also can be applied to multivariate data, irrespective of whether the data have similar or different long-term trends. The NH is that there is no short-term correlation structure among data or that the irregular fluctuations are independent [4].

3. When to reject a null hypothesis

Discriminating statistics are necessary for surrogate data hypothesis testing. The SSS method changes the flow of information in the data. Hence, we choose to use the cross-correlation function (CC) and the average mutual information (AMI) as discriminating statistics. These statistics can determine, on average, how much one learns about one signal by observing the other [9].

After the calculation of these statistics, we need to inspect whether a NH shall be rejected. We employ Monte Carlo hypothesis testing and inspect whether the estimated statistics of the original data fall within or outside the statistical distribution of the surrogate data [10]. When the statistics fall within the distributions of the surrogate data, we conclude that the hypothesis may not be rejected. In this paper, we generate 99 SSS data and hence the significance level is between 0.01 and 0.02 for a one-sided test with two non-independent statistics [4].
It should be noted that although the multiple comparison problem is common in surrogate data applications, we use the CC and the AMI as complementary statistics. This is because some of the test systems are robust to one or the other of our two primary statistical test (the CC and the AMI). Also, although we show plots of both the CC and the AMI as a function of time lag, the hypothesis testing is robustly conducted for a fixed small value of lag (for example, we use only lag $= 1$ or $-1$). The plots of the CC and the AMI as a function of lag are provided for information only. For more details see Refs. [4,7].

4. Applications

We apply our method to financial data in different markets: (1) daily closing price of S&P500 and the volume from 3 January 1950 to 30 December 2005, and (2) daily foreign exchange rates, EUR/USD, GBP/USD and JPY/USD from 5 July 1995 to 30 June 2006. We use $A = 1.0$ and generate 99 SSS data in all cases as mentioned above.

It is generally believed that these are not independent but rather interconnected or interrelated. For stock markets, it is said that the volume is a good reference indicator to know the market force. As a general trend, stock participants have the following four expectations: (1) when stock prices rise and the volume increases or abounds, as the buying pressure is strong, expectation for the higher quotation becomes strong, (2) when stock prices rise and the volume decreases or is few, as the buying pressure is weak, uncertainty over the higher quotation is fomented or the end of the higher quotation is expected, (3) when stock prices fall and the volume increases or abounds, as the selling pressure is strong, uncertainty over the lower quotation becomes strong, and (4) when stock prices fall and the volume decreases or is few, as downward trend for the quotation is weak, expectation for the end of the lower quotation is fomented or the market upturn is expected. For exchange rate markets, it is shown that there is triangular arbitrage opportunities in foreign exchange market and that these generate an interaction among foreign exchange rates [11]. Furthermore, a hierarchical taxonomy of various currencies is derived [12]. The above mentioned examples indicate that financial data are not independent.

As a preliminary test, we investigate whether temporal correlations of the first difference data of the closing prices and three exchange rates are absent. If so, we consider that the dynamics is essentially the same as

![Fig. 1. Daily closing price of S&P500 and the volume: (a) and (c) show the first difference data, and (b) and (d) the volume, where (c) and (d) show an enlargement of (a) and (b). The horizontal axis shows the trading day and the ordinate axis is arbitrary in (a)-(d). For the purposes of this study, physical units carry no significant meaning. Hence, we use arbitrary unit for the ordinate axis.](image-url)
random walk (RW), and the essence of the data is the first difference data (return). Further details are provided in Ref. [8]. We found that the first differences are 1D random variables or time-varying random variables, and then we could not reject the RW hypothesis [8]. Hence, we decided to use the first difference data except the volume of stock data.

Fig. 1 shows the return of daily closing price of S&P500 and the volume. Figs. 1(a) and (b) show that when the volatility is large the volume is also large. From this, we expect that there are long-term correlations between them. However, we cannot find clear similarities in short-term variabilities (irregular fluctuations) in Figs. 1(c) and (d). Fig. 2 shows the return of EUR/USD, GBP/USD and JPY/USD. In Fig. 2 we find neither long- nor short-term similarities among them.

We apply the SSS method to these data. When the data are the return of S&P500 and the volume, although Fig. 3(a) shows that the CC of the original data falls within the distribution of the SSS data, the AMI of the original data falls outside the distribution of the SSS data. As there is an argument that the change of the volume has an important roll as mentioned above, we also investigate the case of the first difference data. The result is essentially the same as when using the original volume data. Hence, we conclude that the return of

---

Fig. 2. The first difference data of daily closing price of EUR/USD and GBP/USD: (a) and (b) show EUR/USD, (c) and (d) GBP/USD, and (e) and (f) JPY/USD, where (b), (d) and (f) show an enlargement of (a), (c) and (e), respectively. The horizontal axis shows the trading day and the ordinate axis is arbitrary in (a)–(f). For the purposes of this study, physical units carry no significant meaning. Hence, we use arbitrary unit for the ordinate axis.
daily closing price of S&P500 and the volume have correlation structures. When the data are exchange rates, Figs. 3(c)–(h) show that the CC and the AMI of the original data fall outside the distribution of the SSS data. Hence, we conclude that these exchange rates are not independent.
5. Conclusion

We have applied an algorithm based on the small-shuffle surrogate method to multivariate financial data to investigate whether there are correlation structures in the irregular fluctuations. Applying our method, we find that all financial data examined in this paper are not independent.

Acknowledgments

Tomo Nakamura would like to thank Managing Director Y. Ibuki (Ichiyoshi International (Hong Kong) Ltd) for his fruitful discussions. This research is supported by a Hong Kong University Grants Council Competitive Earmarked Research Grant (CERG) No. PolyU 5235/03E.

References