



LIBOR troubles: Anomalous movements detection based on maximum entropy



Aurelio F. Bariviera^{a,*}, María T. Martín^b, Angelo Plastino^{b,c}, Victoria Vampa^d

^a Department of Business, Universitat Rovira i Virgili, Av. Universitat 1, 43204 Reus, Spain

^b IFLP-CONICET-UNLP, C. C. 727, 1900 La Plata, Argentina

^c SThAR - EPFL Innovation Park, Lausanne, Switzerland

^d Math Department, Engineering Faculty, UNLP, Argentina

HIGHLIGHTS

- We study Libor and Sonia rates between 1999 and 2015.
- A forecast method based on Maximum Entropy is developed.
- Changes in forecast accuracy are detected in Libor rate.
- Findings are consistent with Libor manipulation.

ARTICLE INFO

Article history:

Received 11 November 2015

Available online 8 January 2016

Keywords:

Maximum Entropy
LIBOR manipulation
Interest rates

ABSTRACT

According to the definition of the London Interbank Offered Rate (LIBOR), contributing banks should give fair estimates of their own borrowing costs in the interbank market. Between 2007 and 2009, several banks made inappropriate submissions of LIBOR, sometimes motivated by profit-seeking from their trading positions. In 2012, several newspapers' articles began to cast doubt on LIBOR integrity, leading surveillance authorities to conduct investigations on banks' behavior. Such procedures resulted in severe fines imposed to involved banks, who recognized their financial inappropriate conduct. In this paper, we uncover such unfair behavior by using a forecasting method based on the Maximum Entropy principle. Our results are robust against changes in parameter settings and could be of great help for market surveillance.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

London Interbank Offered Rate (LIBOR) was established in 1986 by the British Banking Association (BBA), who defines LIBOR as “the rate at which an individual Contributor Panel bank could borrow funds, were it to do so by asking for and then accepting inter-bank offers in reasonable market size, just prior to 11:00 [a.m.] London time”. Every London business day each bank within the Contributor Panel (selected banks from BBA) makes a blind submission (each banker does not know what the quotes of the other Banks are) and a compiler (Thomson Reuters) averages the second and third quartiles. In other words, LIBOR is the trimmed average of the expected borrowing rates of leading banks. LIBOR rates are published for several maturities and currencies.

* Corresponding author.

E-mail addresses: aurelio.fernandez@urv.net (A.F. Bariviera), teremartin.map@gmail.com (M.T. Martín), angeloplastino@gmail.com (A. Plastino), victoriavampa@gmail.com (V. Vampa).

<http://dx.doi.org/10.1016/j.physa.2016.01.005>

0378-4371/© 2016 Elsevier B.V. All rights reserved.

Over the time LIBOR became a fundamental interest rate with three main characteristics: (i) it was viewed as an (intended) measure of the borrowing cost in the interbank market, (ii) before the financial crisis, it was interpreted as a risk free rate and (iii) it is a signal of global credit market conditions. Libor is enormously influential due to its use for the valuation of financial products worth trillions of dollars [1].

The way in which LIBOR is fixed is peculiar, because it does not arise from actual transactions. It is not the result of the competing forces of supply and demand. There is a panel of banks selected by the BBA. Each of them should submit their best estimate according to the following question: “At what rate could you borrow funds, were you to do so by asking for and then accepting inter-bank offers in a reasonable market size just prior to 11 am?” [2]. At some point, individual bank LIBOR submissions are often regarded as a proxy for the financial health of the submitting entity. Usually, an employee or group of employees responsible for cash management in a bank are in charge of the daily submission to BBA. They should base their submission on the money market conditions for the bank, and should not be influenced by other bank divisions such as the derivatives trading desks. A fair Libor could signals the state of the interbank money market, and the central banks could act to alleviate frictions in it.

Until May 29, 2008 LIBOR was presumed a pretty honest estimation of the borrowing costs of prime banks. On that day, [3] published an article on the Wall Street Journal casting doubts on the transparency of LIBOR's setting, implying that published rates were lower than those implied by credit default swaps (CDS). Investigations conducted by several market authorities such as US Department of Justice, the European Commission, and the Financial Services Authority (FSA)¹ detected data manipulation and imposed severe fines to banks involved in such illegal procedure.

Several leading banks applied for leniency. Jurists use to say “*confessio est probatio probatissima*”, i.e. confession is the best proof. Therefore, we can accept that, at least, there was some kind of unfair individual submissions or even worse, a collusion attempt by a cartel of banks. This manipulation had two main objectives. On the one hand, low submissions were intended to give the market a signal of the bank's own good financial health. If a bank steadily submits greater rates, this could indicate problems in raising money, generating concerns regarding an underlying solvency problem. On the other hand, some low submissions could be aimed to earn money from certain portfolio positions, whose assets are valued according to LIBOR.

The effect of erroneous LIBOR extends beyond the financial markets. In addition to provide a biased interbank lending cost, [4] affirms that it corrupts a “key variable in the first stage of the monetary transmission mechanism”.

The importance of a good pricing system is based on its usefulness for making decisions. As [5] affirmed “we must look at the price system as such a mechanism for communicating information if we want to understand its real function”. If the price system is contaminated, but perceived as pure, the effect could reach also the real economy, making it difficult to find a way out the financial crisis.

This rate-rigging scandal made economists to examine the evolution of LIBOR rates and compare it with other market rates. [6] documented the decoupling of the LIBOR rate from other market rates such as the Overnight Interest Swap (OIS), Effective Federal Fund (EFF), Certificate of Deposits (CDs), Credit Default Swaps (CDS), and Repo rates. They hypothesize that the reasons for the divergent behavior were due to expectations of future interest rates and in the accompanying counter-part risk. [7] study individual quotes in the LIBOR bank panel and corroborate the claim by [3] that LIBOR quotes in the US are not strongly related to other bank borrowing cost proxies. In their model, the incentive for misreporting borrowing costs is profiting from a portfolio position. Consequently, the misreporting could point upwards in one currency and downwards in another one, depending on the portfolio exposition. The evidence of such behavior is detected with the formation of a compact cluster of the different panel bank quotes around a given point. [8] track daily LIBOR rates over the period 1987 to 2008.

In particular, this paper analyzes the empirical distribution of the Second Digits (SDs) of the Libor interest rate, and compares them with the uniform and Benford's distributions. Taking into account the whole period, the null hypothesis that the empirical distribution follows either the uniform or Benford's distribution cannot be rejected. However, if only the period after the sub-prime crisis is taken into account, the null hypothesis is rejected. This result puts into question the “aseptic” setting of LIBOR. Recently, [9,10] uncover strange changes in the information endowment of LIBOR time series, as measured by different information theory quantifiers, namely permutation entropy, permutation statistical complexity and permutation Fisher information measure. Their results allow to infer some degree of manipulation or, at least, changes in the underlying stochastic process that governs interest rate's time series.

Antitrust law enforcement is complex, because manipulation and fraud can be elegantly camouflaged. A statistical procedure could hardly be accepted as legal proof in a court of law. However, its use by surveillance authorities makes the attempted manipulation more costly and more difficult to be maintained. Consequently, we view our proposal as a market watch mechanism that could make manipulation and/or collusion attempts more difficult in the future. Additionally, an efficient overseeing mechanism could increase the incentives to apply for leniency at earlier stages of the manipulation [11].

The aim of this paper is to show that a forecasting method based on Maximum Entropy Principle (MaxEnt) is very useful not only to produce accurate forecasts, but also to detect some anomalous situations in time series. In particular, we claim that, in absence of data manipulation, forecast accuracy should be approximately the same at all times under examination. On the contrary, manipulation would produce more predictable consequences, increasing the predictive-power of our model, that we apply here to LIBOR and other UK interest rates.

¹ It is noteworthy that the Financial Services Act 2012 renamed FSA as Financial Conduct Authority (FCA), raising the importance of “fair conduct” in financial markets.

This paper is organized as follows. Section 2 describes our methodology based on the Maximum Entropy method. Section 3 describes the data used in the paper and deals with the results obtained with the proposed methodology. Finally, Section 4 draws the main conclusions.

2. MaxEnt approach for predictions in time-series

In a recent paper [12] developed an information theory based method for time series prediction. Given its outstanding results in approaching the true dynamics underlying a given time series, we believe that it ‘provides a suitable methodology to apply in the present circumstances. In order to make the paper self-contained, we review below the description of the method, taken from Ref. [12].

The behavior of a dynamical system can be registrated as a time series i.e. a sequence of measurements $\{v(t_n), n = 1, \dots, N\}$ of an observable of the system at discrete times t_n , where N is the length of the time series.

The Takens theorem of 1981 asserts that, for $T \in \mathbb{R}, T > 0$, there exists a functional form of the type,

$$v(t + T) = F(\mathbf{v}(t)), \tag{1}$$

where

$$\mathbf{v}(t) = [v_1(t), v_2(t), \dots, v_d(t)], \tag{2}$$

and $v_i(t) = v(t - (i - 1)\Delta)$, for $i = 1, \dots, d$. Δ is the time lag and d is the embedding dimension of the reconstruction. T represents the *anticipation time* and it is of fundamental importance for a prediction model.

We will consider (as in [12] and references therein) a particular representation for the mapping function of Eq. (1), expressing it, using Einstein’s summation notation, as an expansion of the form

$$F^*(\mathbf{v}(t)) = a_0 + a_{i_1} v_{i_1}(t) + a_{i_1 i_2} v_{i_1}(t) v_{i_2}(t) + a_{i_1 i_2 i_3} v_{i_1}(t) v_{i_2}(t) v_{i_3}(t) + \dots + a_{i_1 i_2 \dots i_{n_p}} v_{i_1}(t) v_{i_2}(t) \dots v_{i_{n_p}}(t), \tag{3}$$

where $1 \leq i_k \leq d$ with $1 \leq k \leq n_p$ and n_p being an adequately chosen polynomial degree so as to series-expand the mapping F^* . The number of parameters in Eq. (3) corresponding to the terms of degree k depends on the embedding dimension and can be calculated using combination with repetitions,

$$\binom{d}{k}^* = \frac{(d + k - 1)!}{k!(d - 1)!}. \tag{4}$$

Accordingly, the length of the vector of parameters \mathbf{a} , N_c , is

$$N_c = \sum_{k=1}^{n_p} \binom{d}{k}^*. \tag{5}$$

The computations are made on the basis of a specific information supply, given by M points of the series

$$\{\mathbf{v}(t_n), v(t_n + T)\}, \quad n = 1, \dots, M. \tag{6}$$

Given the data set in Eq. (6), the parametric mapping in Eq. (3) will be determined by the following condition,

$$F^*(\mathbf{v}(t_n)) = v(t_n + T) \quad n = 1, \dots, M, \tag{7}$$

which can be expressed in matrix form as,

$$W\mathbf{a} = \mathbf{v}_T, \tag{8}$$

where W is a matrix of size $M \times N_c$, whose n th row is

$$[1, v_{i_1}(t_n), v_{i_1}(t_n)v_{i_2}(t_n), \dots, v_{i_1}(t_n)v_{i_2}(t_n) \dots v_{i_{n_p}}(t_n)]$$

(Cf. Eq. (3)) and $(\mathbf{v}_T)_n = v(t_n + T)$, for $n = 1, \dots, M$. Shannon’s entropy, defined for a discrete random variable, can be extended to situations for which the random variable under consideration is continuous.

In order to infer coefficients which are consistent with the data we shall assume that each set \mathbf{a} is realized with probability $P(\mathbf{a})$. Thus, a normalized probability distribution over the possible sets \mathbf{a} is introduced,

$$\int_I P(\mathbf{a}) \, d\mathbf{a} = 1, \tag{9}$$

where $d\mathbf{a} = da_1 da_2 \dots da_{N_c}$ and N_c is the number of parameters of the model.

The problem then becomes that of finding $P(\mathbf{a})$ subject to the requirement that the associated entropy H be maximized, since this is the best way of avoiding any bias. The expectation value of \mathbf{a} , is defined by

$$\langle \mathbf{a} \rangle = \int_I P(\mathbf{a}) \mathbf{a} \, d\mathbf{a}. \tag{10}$$

Consider the continuous random variable \mathbf{a} with probability density function $p(\mathbf{a})$ on I and $I = (-\infty, \infty)$. The entropy is given by

$$H(\mathbf{a}) = - \int_I P(\mathbf{a}) \ln P(\mathbf{a}) \, d\mathbf{a}, \quad (11)$$

whenever it exists, and the relative entropy reads

$$H = - \int_I P(\mathbf{a}) \ln \frac{P(\mathbf{a})}{P_0(\mathbf{a})} \, d\mathbf{a}, \quad (12)$$

where $P_0(\mathbf{a})$ is an appropriately chosen a priori distribution.

This measure exhibits many of the properties of a discrete entropy but, unlike the entropy of a discrete random variable, that for a continuous random variable may be infinitely large, negative, or positive (Ash, 1965 [2]). We characterize, via the entropic maximum principle, various probability distributions, subject to the constraints Eqs. (9) and (8) for the expectation $\langle \mathbf{a} \rangle$ of \mathbf{a} . The method of solving this constrained optimization problem is to use Lagrange multipliers for each of the operating constraints and maximize the following functional with respect to $P(\mathbf{a})$,

$$J = - \left[\int_I P(\mathbf{a}) \ln \frac{P(\mathbf{a})}{P_0(\mathbf{a})} \, d\mathbf{a} + \lambda_0 \left[\int_I P(\mathbf{a}) \, d\mathbf{a} - 1 \right] + \lambda^t \int_I [WP(\mathbf{a})\mathbf{a} - \mathbf{v}_T] \, d\mathbf{a} \right], \quad (13)$$

where λ_0 and λ are Lagrange multipliers associated, respectively, with the normalization condition and with the constraints, Eqs. (9) and (8).

Taking the functional derivative with respect to $P(\mathbf{a})$ we get

$$\frac{\partial J}{\partial P(\mathbf{a})} = \ln \left(\frac{P(\mathbf{a})}{P_0(\mathbf{a})} \right) + 1 + \lambda_0 + \lambda^t W\mathbf{a} = 0, \quad (14)$$

which implies that the maximum entropy distribution must have the form

$$P(\mathbf{a}) = \exp \left[-(1 + \lambda_0) \exp(\lambda^t W\mathbf{a}) P_0(\mathbf{a}) \right]. \quad (15)$$

If the a priori probability distribution $P_0(\mathbf{a})$ is chosen to be proportional to $\exp(-\frac{1}{2}\mathbf{a}^t[\sigma^2]^{-1}\mathbf{a})$, where σ^2 is the covariance matrix, a Gaussian form for the probability distribution $P(\mathbf{a})$ is obtained, with

$$\langle \mathbf{a} \rangle = -\sigma W^t \lambda. \quad (16)$$

Considering Eq. (8), the Lagrange multipliers λ can be eliminated:

$$\lambda = -\sigma^{-1}(WW^t)^{-1}\mathbf{v}_T, \quad (17)$$

and one can thus write

$$\langle \mathbf{a} \rangle = W^t(WW^t)^{-1}\mathbf{v}_T. \quad (18)$$

The matrix $W^t(WW^t)^{-1}$ is known as the Moore–Penrose pseudo-inverse of the matrix W (see [12] and references therein). Consequently, this result shows that the maximum entropy principle coincides with a least square criterion. Once the pertinent parameter vector \mathbf{a} is determined, it is used to predict *new series'* values, $\widehat{v}(t_n + T)_{n=1, \dots, M_p}$, according to

$$(\widehat{v}(t_n + T))_{n=1, \dots, M_p} = \widehat{W}\mathbf{a}, \quad (19)$$

where \widehat{W} is the matrix of size $M_p \times N_c$ (see Eq. (8)), obtained using $\widehat{v}(t_n)$ values.

3. Data and results

We analyze the Libor in pound Sterling, and we compare the results with the Sterling OverNight Interbank Rate (SONIA). The data span is from 01/01/1999 until 07/04/2015, with a total of 4243 datapoints. All data were retrieved from DataStream.

The main difference between Libor and SONIA is that, the former is at best an unbiased self estimation of banks' borrowing costs, whereas the latter reflects actual borrowing costs based on actual transactions. Consequently, SONIA is less prone to manipulation, since it reflects the weighted average rate of all unsecured overnight sterling transactions brokered in London by Wholesale Markets Brokers' Association (WMBA).

As described in Section 1, interest rates in a competitive market should follow approximately a random walk. We claim that data manipulation produces a contamination of the original stochastic process: an exogenous deterministic artifact is introduced into the time series. According to Wold's theorem [13], a time series can be separated into a deterministic part and a stochastic part. If the stochastic part dominates the behavior of the time series, forecast is unsuccessful. If the manipulation is strong enough, it dominates the time series behavior, resulting in a more predictable sequence.

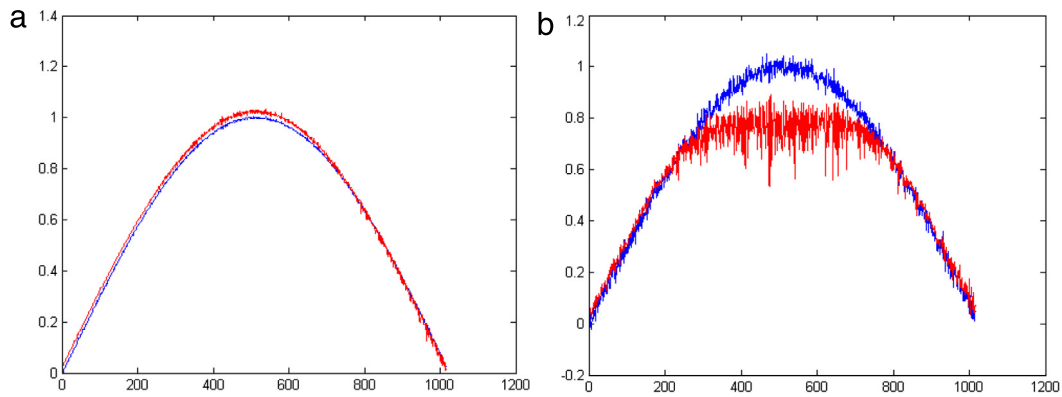


Fig. 1. Simulation of forecast accuracy of a deterministic function contaminated with (a) weak noise and (b) a strong noise signal. Blue lines are the original signals and red lines are the forecasted ones.

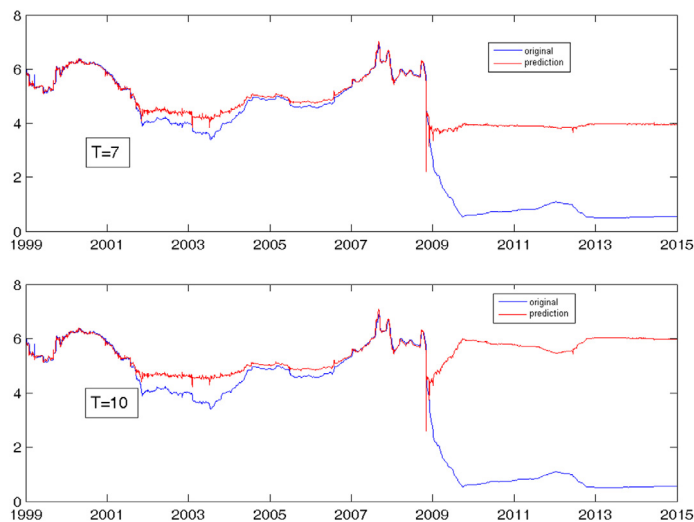


Fig. 2. Original (blue lines) and forecasted (red lines) Libor time series for different anticipation times. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

In order to describe this phenomenon we simulate a deterministic function (sinus function) and we add some white noise-level. Then, we train our model and proceed to forecast future values. The results of this experiment is shown in Fig. 1. We see that, when low noise is introduced into the system (Fig. 1(a)), the prediction is very accurate, since the deterministic part of the time series dominates. However, if we increase the noise level (Fig. 1(b)), the random behavior of the white noise dominates the behavior of the time series as a whole. Consequently the predictive performance falls.

Returning to our problem, we present the results obtained using the methodology proposed in Section 2 on Libor and Sonia time series. We consider the embedding dimension $d = 4$ and the polynomial degree $n_p = 2$. The length of the vector of parameters, according to Eq. (5) is $N_c = 15$.

We fit our model with $M = 700$ data-points, corresponding to approximately two and a half years beginning in 01/01/1999. Once the model's parameters are determined, we forecast the rest of the time series, up to the end of it.

In Figs. 2 and 3 the original time series values and the predicted ones are overlapped (blue and red refer to original and predicted values, respectively) for different anticipation time values.

In order to prove the robustness of our proposal we made predictions for different anticipation times ($T = 7, 10$ days) in the case of Libor and $T = 7$ for SONIA.

We see in Fig. 2 that, as expected, during the model interval period (until the middle of 2001), the original and the predicted time series are very close. This is the consequence of the adequate fitting power of the model. As is the case for any forecast method, one tries to mimic the behavior of the time series to be estimated. When we move into the (out of the sample) prediction interval, we note that during the first months, our method behaves rather well, but is not as good as it is for the training period. We expect that, as economic theory affirms, competitive prices should behave randomly [14]. Consequently, if we assume that the time series under study is generated by a memoryless stochastic process, accurate forecasts are not possible. This is reflected in the vertical difference between the original and the predicted time series. This difference disappears

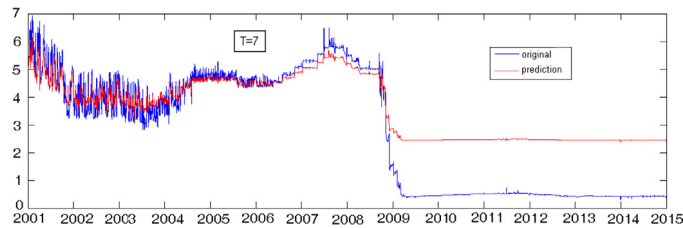


Fig. 3. Original (blue line) (red line) and forecasted SONIA time series. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

between 2007 and 2009. During this period, our model is able to capture almost completely the dynamics of the time series. We claim that this 2007–09 section corresponds to a more deterministic behavior, which is compatible with the alleged manipulation uncovered by regulatory authorities in the UK. After 2009, the predictions of our model diverge from the actual time series, indicative that a more random behavior dominates the system dynamics, and reflecting the end of manipulation.

The behavior of our model using SONIA time series is different. On the one hand, the fitting during training period (the first 700 data) is not as accurate as in the case of Libor. During the out-of-sample forecast, we observe that the original time series is very volatile and the model cannot fully capture this feature. Consequently there is always a vertical difference between the original time series and the predicted time series. This difference is more or less constant from 2003 until 2009 and increases afterwards. We cannot detect using our method differences in prediction accuracy during the period 2007–2009.

The uneven behavior between Libor and SONIA time series, reinforces our claim that Libor was somewhat contaminated during the period 2007–2009 by a device that made prediction more accurate. Given that the prediction model is the same for both periods, we conjecture that the time series is dominated by a deterministic process in such period. Recalling the literature review of Section 1, we can state that this result is an indirect proof of LIBOR manipulation. We emphasize that such “manipulation” necessarily comprises the contamination of the time series with a deterministic device, which was detected by the MaxEnt approach.

4. Conclusions

In this paper we present a novel predictive methodology based on the MaxEnt principle. Taking into account its previous performance [12], we believe it is suitable for the study of the “Libor Case”. We studied Libor time series between 1999 until 2015. Based on the predictive accuracy of our method, we were able to detect two distinct regimes. The first one, extends between 2002–2006 and 2009–2015. For these periods, the time series behaves as predicted by standard economic theory, reflecting the random character of prices in competitive environments. The prediction power is, consequently, poor. The second time-period comprises years 2007 and 2008. In it, the time series changes its regime, becoming a more predictable one. We can safely think that a deterministic device was introduced into the Libor setting. This situation takes place at the time that what was called by the newspapers as the “Libor manipulation” one. Additionally, we test the same model in SONIA time series. This interest rate is less prone to manipulation because it reflects the average of all effective transactions in the London wholesale lending market. SONIA time series predictions are less accurate than Libor predictions, and at the same time do not present remarkable differences during the conflictive years 2007 and 2008. Both time series should reflect the same economic fact: the true lending cost in the interbank lending market. Consequently, both time series should follow the same stochastic dynamics. The remarkable difference between them, is an indirect proof of Libor misconduct. As a consequence, our approach is able to detect such manipulation, using exclusively data from Libor time series. We would like to emphasize the relevance of advanced statistical models in market’s watch mechanisms. Our results could be of interest to surveillance authorities, given the importance of fair market conditions in free market economies.

References

- [1] Bank For International Settlements, 2014. Statistical release. otc derivatives statistics at end-december 2013. monetary and economic department. http://www.bis.org/publ/otc_hy1405.pdf.
- [2] BBA Trent Ltd., 2015. Historical perspective. <http://www.bbatrent.com/explained/historical-perspective>. (Online; accessed 20-Dec-2014).
- [3] C. Mollenkamp, M. Whitehouse, Study casts doubt on key rate: WSJ analysis suggests banks may have reported flawed interest data for libor, *Wall Street J.* (2008).
- [4] A. Stenfors, Libor deception and central bank forward (mis-)guidance: Evidence from norway during 2007–2011, *J. Int. Financ. Mark. Inst. Money* 32 (2014) 452–472.
- [5] F.A. Hayek, The use of knowledge in society, *Am. Econ. Rev.* 35 (1945) 519–530.
- [6] J.B. Taylor, J.C. Williams, A black swan in the money market, *Am. Econ. J. Macroecon.* 1 (2009) 58–83.
- [7] C.A. Snider, T. Youle, Does the libor reflect banks’ borrowing costs? Available at SSRN: <http://ssrn.com/abstract=1569603> or <http://dx.doi.org/10.2139/ssrn.1569603>.
- [8] R.M. Abrantes-Metz, S.B. Villas-Boas, G. Judge, Tracking the libor rate, *Appl. Econ. Lett.* 18 (2011) 893–899.
- [9] A.F. Bariviera, M. Guercio, L. Martinez, O. Rosso, The (in)visible hand in the libor market: an information theory approach, *Eur. Phys. J. B* 88 (2015) 208.
- [10] A.F. Bariviera, M. Guercio, L. Martinez, O. Rosso, A permutation information theory tour through different interest rate maturities: the libor case, *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* 373 (2015) 20150119.

- [11] R. Abrantes-Metz, D.D. Sokol, The lessons from libor for detection and deterrence of cartel wrongdoing, *Harv. Bus. Law Rev. Online* 3 (2012) 10–16. <http://www.hblr.org/2012/10/the-lessons-from-libor-for-detection-and-deterrence-of-cartel-wrongdoing/>.
- [12] M. Martín, A. Plastino, V. Vampa, G. Judge, A parametric, information-theory model for predictions in time series, *Phys. A* 405 (2014) 63–69.
- [13] H.O.A. Wold, *A Study in the Analysis of Stationary Time Series*, Almqvist & Wiksell, Stockholm, Sweden, 1954.
- [14] P.A. Samuelson, Proof that properly anticipated prices fluctuate randomly, *Ind. Manag. Rev.* 6 (1965) 41–49.