Journal of Economics and International Finance Vol. 3(7), pp. 418-423, July 2011
Available online at http://www.academicjournals.org/JEIF
ISSN 2006-9812 ©2011 Academic Journals

Full Length Research Paper

Modelling stock price behaviour: The Kernel approach

John Muteba Mwamba

Department of Economics and Econometrics, Office: D - Ring 218, University of Johannesburg, P. O Box 524, Auckland Park, South Africa. E-mail: johnmu@uj.ac.za Tel: +27 11 559 4371.

Accepted 13 June, 2011

The efficient market hypothesis asserts that financial markets are always efficient and therefore cannot be predicted in order to make abnormal returns. This paper investigates the predictability of stock prices in more efficient and developed markets (U.S and U.K) using two econometric methods namely, the random walk and the non-parametric methods. Based on the out-of-sample predicted mean square error, and the resampled confidence interval and volatility we found that both U.S and U.K stock prices are predictable with more accuracy when a nonparametric method is used.

Key words: Kernel regression, Epanechnikov Kernel, bandwidth, dynamic random walk, bootstrapping, non-parametric model, F-test.

INTRODUCTION

Predictability of future behaviour of stock price has been a challenging puzzle that many financial economists attempted to solve during the past decade. Wu and Hu (1997) argue that "a long standing puzzle to financial economists is their difficulty to out-perform the benchmark random walk model in out-of-sample contests". In this paper, data from U.S and U.K stock markets (assumed to be efficient markets) are used in order to assess the out-of-sample performance of the kernel regression method over the traditional random walk model. Results indicate that by relaxing normality assumption on stock price distribution; it is possible to outperform the random walk model based on the predicted mean square error (the widely used criterion for comparison of two econometric models) and the bootstrapped mean. These results contribute towards improving predictability of stock price behaviour in financial market as discussed in the past by many researchers (Campbell and Shiller, 1988; Greene and Hodges, 2002; Zitzewitz, 2003; Chalmers et al., 2001) who have found that future stock price movements can be predictable in order to generate scenarios for buy and sell signals.

Predictability of future stock price behaviour has become an increasingly attractive area of research in financial econometrics for both academics and practitioners as the need for adequate and precise market timing increases. Traders in stock markets need to know exactly when to buy and/or sell in order to generate profit for their companies. It is in this regard that the random walk model which is considered as being consistent with the weak form of efficient market hypothesis (Anderson and lauvsnes, 2007) has been used for decades now by academics and professionals to model the behaviour of stock price. This method assumes that tomorrow's stock price is equal to today's price plus some random error. It also assumes that the price generating process is of the Gaussian family and can be modelled by a linear model. Such assumptions should only be drawn if imbalances between stock buyers and sellers are temporary, and occur in an efficient market where price reflects all known and unknown information. Linear models such as the random walk, autoregressive moving average (Box and Jenkins, 1994) and nonlinear econometric models are used to predict future behaviour of stock prices. As noted by Fan and Huang (2001), such models present significant drawbacks, the major one being the normality assumption on which they are built. This strong assumption may neglect the occurrence of extreme events during market crashes. Moreover, Adya and Collopy (1998), Chatfield (1995), and Tkacz (2001) found no clear evidence on whether nonlinear models might provide better forecasts than linear and random walk models when applied to financial time series data.

This paper uses a different approach in modelling stock price behaviour namely the nonparametric kernel regression

JEL: C14, C15, C53, C58, and G17

method which overcomes some of limitations of parametric models.

The paper deals specifically with predictability of stock price behaviour using univariate nonparametric kernel method. In this method, the normality assumption for example, is relaxed to let the data speak for itself about the distribution that it follows. Nonparametric kernel method has been used in the past by Skabar (2008) to estimate density distributions of returns in order to generate out-of sample forecasts. He found that the method can predict future direction of change in stock price and reduce the risk of under or overfitting the time series. His performance results in the out-of-sample period are consistent with those obtained in this paper.

Niglio and Perna (2003) applied the Kernel method with the corrected version of generalized cross validation for optimal bandwidth selection to two climatic time series data collected from South Italy (Scafari) from January 1960 to December 2000.

They showed that improving the bandwidth selection in Kernel method can overcome undersmoothing of climatic data and hence predict accurately the future behaviour of temperature in Scafari. Other works on univariate nonparametric kernel regression include Auestad and Tjostheim (1990), and Hardle and Vieu (1991) who used the kernel estimator proposed by Nadaraya (1964) to estimate the conditional mean and conditional variance of time series. Hardle and Tsybakov (1992) estimated both the conditional mean and variance of a time series using local linear estimates.

METHODOLOGY

In this paper a non-linear autoregressive kernel model for a univariate time series was built in five steps, namely, the estimation of the density function (kernel function), the search for an optimal bandwidth for the kernel function, the determination of the exact number of lags to be included in the regression equation, and the estimation of the conditional mean and volatility.

A specification test was lastly used in order to assess the correctness of the model. These steps are now discussed as follows.

Density estimation

The smooth density function that is dealt with in this paper is known as the Kernel density function estimator $\hat{f}(x)$ given by the following expression:

$$\hat{f}(x) = (nh)^{-1}\sum_{i=1}^{n} K\{(X_i - x)/h\} \quad (1)$$

where h represents the smoothing parameter known as the bandwidth, and $K(.)$ the Kernel function satisfying the following properties:

$$\int_{-\infty}^{+\infty} k(x)dx = 1,\ \int_{-\infty}^{+\infty} xK(x)dx = 0 \quad \text{and}$$

$$\int_{-\infty}^{+\infty} x^2 K(x)dx = \delta < \infty$$

Usually $K(.)$ is chosen to be a unimodal probability density function that is symmetric about zero. This ensures that $\hat{f}(x)$ is itself also a density. In this case notice that $K(.)$ is simply the $N(0,h^2)$ density so that h plays the role of a scaling factor which determines the spread of the kernel.

The Kernel estimator is thus constructed by centering a scaled kernel at each observation. The value of the kernel estimator at a point, x for instance; is simply the average of the n kernel ordinates at that point. One can think of the kernel as spreading a "probability mass" of size i/n associated with each data point about its neighbourhood. The list of kernel functions used in practice includes the Epanechnikov; Gaussian, triangular, biweight; rectangular kernel, and so on. Practically, the choice of the shape of the Kernel function is less important than the choice of the optimal bandwidth.

Optimal bandwidth selection

The integrated mean square error of the kernel estimator $(imse\hat{f}(x))$ is used to obtain the optimal bandwidth. Hall et al. (2004) demonstrated that:

$$imse\hat{f}(x) \approx \frac{\phi_0}{nh} + \frac{h^4}{4}\kappa_2^2\phi_1 \quad (2)$$

where $\phi_0 = \int K^2(z)dz$ and $\phi_1 = \int \{f''(x)\}^2 dx$

The optimal bandwidth (h_{opt}) can now be obtained by minimizing the $imse\hat{f}(x)$ aforestated with respect to the bandwidth h :

$$h_{opt} = \left\{ \frac{\int K^2(z)dz}{(\int z^2 K(z)dz)^2 \int \{f''(x)\}^2 dx} \right\}^{1/5} n^{-1/5} \quad (3)$$

The optimal kernel density estimator corresponding to the aforestated optimal bandwidth has been suggested by Epanechnikov (1969), and is known as the Epanechnikov kernel, and is given by:

$$K_{Ep}(x) = \begin{cases} \frac{3}{4}\left(1 - x^2\right) & for|x| < 1 \\ 0 & otherwise \end{cases}$$

Lag selection

The exact number of lags to include in a non-linear autoregressive kernel regression equation is obtained using the estimated final predictor error (FPE) criterion (Auestad and Tjostheim, 1990).

Estimation of conditional mean and volatility

Let Y_t be the stock price value at time t, and $y_{t-1}; y_{t-2}; y_{t-3};......y_{t-l}$ the stock price in previous periods. It is assumed that there is a non-parametric and non-linear relationship between the current and the previous stock prices. This relationship is modelled by a non-linear autoregressive heteroskedastic process of the form:

$$Y_t = m(y_{t-1}, y_{t-2},..., y_{t-l}) + \varepsilon_t \; ; \; t = 1,2,..........T \quad (4)$$

where $\varepsilon_t = \sigma_t e_t$, and

$$\sigma_t^2 = Var(\varepsilon_t /(Y_{t-1}, Y_{t-2},..., Y_{t-l}) = x)$$

In Equation (4) $m(y_{t-1}, y_{t-2},..., y_{t-l})$ represents the conditional mean whereas σ_t represents the conditional volatility of the stock price. In order to plot the conditional mean and volatility functions in a three-dimensional space, it is assumed in this paper that two lags only have been selected using the final prediction error during the lag selection stage. The kernel regression in (4) can also be rewritten in terms of the conditional mean and the conditional volatility as follows:

$$Y_t = m(Y_t /(Y_{t-1}, Y_{t-l}) = x) + \sigma_t e_t \quad (5)$$

This model is different from the classical autoregressive (AR) in two ways. Firstly, the AR model assumes linear dependence on past stock prices, whereas the aforestated model assumes a non-linear dependence on past stock prices. Secondly, the classical GARCH models of volatility assume normality and symmetrical behaviour of volatility.

In the model presented earlier, no such assumption is made about the distribution of the stock price and error terms. The estimation of the conditional mean and volatility is now presented as follows. Let p be the degree of the polynomial being fit to the stock price data set, and h_{opt} the optimal bandwidth selected. The estimator of the conditional mean denoted by $\hat{m}(x_t, p, h_{opt})$ is obtained by fitting a polynomial of the form $\beta_0 + \beta_1(x_1. - x) + ... + \beta_p(x_p. - x)^p$ to the stock price series using the least square cross-validation method in which the kernel estimator is considered as the weight function:

$$\hat{m}(x_t, p, h_{opt}) = (X_t^T K_{h_{opt}} X_t)^{-1} X_t^T K_{h_{opt}} Y_t \quad (6)$$

where X_t is a square matrix containing one in the first column, Y_{t-1} in the second column, and Y_{t-2} in the last column. The Nadaraya-Watson conditional mean estimator (Nadaraya and Watson, 1964) is obtained when the degree of the polynomial being fitted to the stock price data is zero $(p = 0)$:

$$\hat{m}(x_t, 0, h_{opt}) = \frac{\sum_{t=1}^{T} K_{h_{opt}}(X_t - x) Y_t}{\sum_{t=1}^{T} K_{h_{opt}}(X_t - X)} \quad (7)$$

The local linear estimator of the conditional mean is obtained when the degree of the polynomial being fit to the stock price series is one $(p = 1)$:

$$\hat{m}(x_t; 1, h_{opt}) = n \sum_{t=1}^{T} \frac{\{\hat{S}_2(x_t, h_{opt}) - \hat{S}_1(x_t, h_{opt})(x_t - x)\} K_{h_{opt}}(x_t - x) Y_t}{\hat{S}_2(x_t, h_{opt}) \hat{S}_0(x_t, h_{opt}) - \hat{S}_1(x_t, h_{opt})^2} \quad (8)$$

The conditional volatility is specified the same way as earlier. First the form of the estimator for the conditional mean must be specified (Nadaraya-Watson or the local linear estimator).

Assume that the Nadaraya-Watson estimator of the conditional mean is specified; then the stock price behaviour is modelled by the following expression; $Y_t = \hat{m}(x_t, 0, h_{opt}) + \varepsilon_t$ residual squared are then computed:

$$\varepsilon_t^2 = (Y_t - \hat{m}(x_t, 0, h_{opt}))^2 \quad (9)$$

Using the least square cross-validation method with $K(.)$ considered as weight function, the Nadaraya-Watson estimator of the conditional volatility is then obtained:

$$\sigma^2(x_{\sigma,t}, 0, h_{opt}) = (X_{\sigma,t}^T . K_{h_{opt}} . X_{\sigma,t})^{-1} . X_{\sigma,t}^T . K_{h_{opt}} . \varepsilon_t^2 \quad (10)$$

where $X_{\sigma,t}$ is a square matrix whose first column entries are equal to one, the second column entries are ε_{t-1}^2, and the last column entries are equal to ε_{t-l}^2 (with l representing the largest lag). The Nadaraya-Watson estimator for the conditional volatility is obtained when the degree of the polynomial being fitted to the stock price conditional volatility is zero $(p = 0)$:

$$\sigma^2(x_{\sigma,t}, 0, h_{opt}) = \frac{\sum_{t=l}^{T} K_{h_{opt}}(X_{\sigma,t} - x_\sigma) \varepsilon_t^2}{\sum_{t=l}^{T} K_{h_{opt}}(X_{\sigma,t} - x_\sigma)} \quad (11)$$

Specification tests

The error terms e_t in (5) need to be estimated in order to check whether the fitted conditional volatility is appropriate.

$$\hat{e}_t = \frac{Y_t - \hat{m}(x_t, 0, h_{opt})}{\sqrt{\sigma^2(x_{\sigma,t}, 0, h_{opt})}} = \frac{\hat{\varepsilon}_{t,opt}}{\sqrt{\sigma^2(x_{\sigma,t}, 0, h_{opt})}} \quad (12)$$

Note that if the conditional mean $\hat{m}(x_t, 0, h_{opt})$ and the

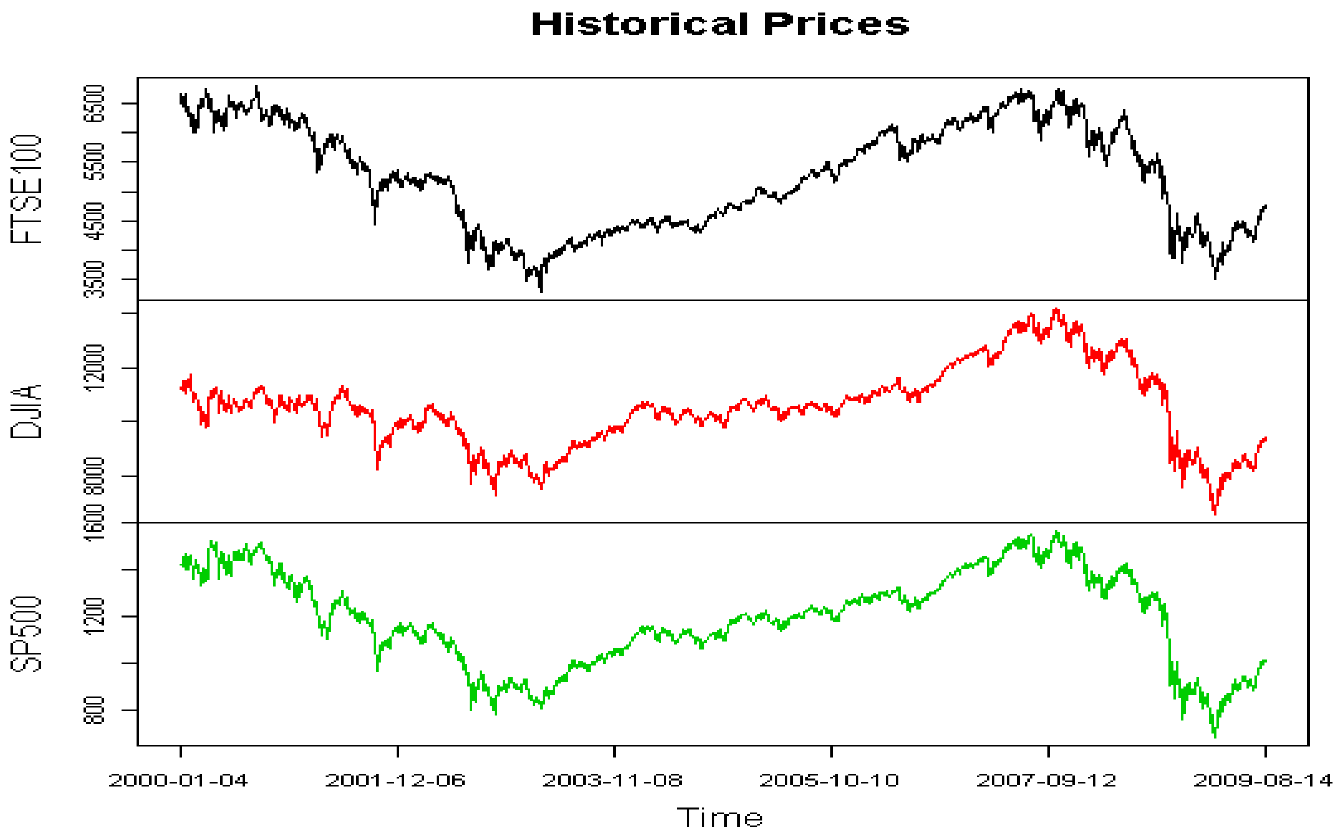

Figure 1. Historical prices.

conditional volatility $\sigma^2(x_{\sigma,t}, 0, h_{opt})$ are correctly specified, then the estimates $\hat{e}_t$ would result in white noise random variables.

EMPIRICAL RESULTS

Daily stock prices of S&P 500, DOWJONES (DJIA), and FTSE 100, have been collected from Yahoo finance, these markets are more efficient than emerging markets; our aims is to investigate whether these efficient markets can be predictable. The dataset encompasses 02 January 2001 to 31 December 2010. Figure 1 shows that the U.K and U.S equity markets commove together during both bull and bear market periods. In order to compare the forecasting power of the kernel and the random walk model, the root mean square error criterion as well the bootstrap confidence interval and the volatility are used. Firstly, the sample period is divided into two consecutive periods, namely the in-sample period (from 4 January 2000 to 06 January 2008, representing 86% of the sample data) and the out-of-sample period (from 7 January 2008 to 27 April 2009). The Epanechnikov kernel has been used as the estimator of the univariate density function corresponding to individual distribution of all the major stock indices used in this paper (S&P 500, DJIA and the FTSE 100). An automated optimal bandwidth selection using jmulti software (www.jmulti.com) has been used for lag selection as well as for the estimation of conditional mean and volatility for each one of the aforementioned stock indices.

Two lags only are considered for the non-linear autoregressive kernel regression using the corrected asymptotic final prediction error suggested by Tschernig and Yang (2000). Firstly, local linear estimators of conditional mean for DJIA and S&P 500 (using lag one and three) and for FTSE 100 (using lag one and five) are found. Previous stock price (lag one) is found to have a significant impact on the current price for both the U.S and U.K stocks, whereas lag three and five were found to have a significant impact on U.S and U.K stocks respectively. Secondly, the conditional volatility of S&P 500, DJIA, and the FTSE 100 respectively was modelled using lag one and lag five. Both lag one and five were found to be having a significant impact on U.S and U.K stocks respectively. Forecasts are generated using the kernel regression and the random walk models. In this paper, two parametric forecasts are employed using the random walk model. The first is the dynamic random forecasting, which is a multi-stage forecasting method (from the start of the out-of-sample forecast); forecasts are computed recursively using the lagged value of the

Table 1. Predicted mean square errors and volatilities.

	Kernel model	Dynamic random walk	Static random walk
DJIA	207.4939 **(1804)**	2378.1900 **(461.966)**	208.3989 **(1801.396)**
S&P 500	23.9196 **(218.473)**	300.3200 **(40.108)**	24.9600 **(216.7371)**
FTSE 100	96.7723 **(746.7752)**	1044.7910 **(183.7762)**	96.9913 **(745.2531)**

Table 2. Decision rules and probabilities (in bold) of F-test two-sample for variance.

Method	DJIA	S&P 500	FTSE 100
Non-parametric Kernel	Do not reject **(0.4876)**	Do not reject **(0.4407)**	Do not reject **(0.4756)**
Dynamic Random Walk	Reject **0**	Reject **0**	Reject **0**
Static Random walk	Do not reject **(0.4996)**	Do not reject **(0.499)**	Do not reject **(0.4944)**

dependent variable. The second is the static random walk forecasting model which performs a series of one-step ahead of the dependent variable. Table 1 contains the predicted mean square error and volatility (in bold) of forecasts generated from the kernel model, the dynamic, and the static random walk methods.

Table 1 reveals that the pmse alone cannot be used to assess the performance of these forecasting methods. It can be seen in Table 1, that non-parametric kernel method has the lowest pmse; however its volatility is not minimal. An investment analyst would be confused depending on his risk aversion, whether he should use the kernel method or the random walk models (dynamic or statistic) in order to generate scenarios for potential buy or sell signals in short term. To further assess the performance of the kernel method; another statistical criterion - the F-test of out-of-sample volatility is used. The underlying idea here is that, if the null hypothesis of an F-test ($H_0 : \sigma_1^2 = \sigma_2^2$; according to which the variance of the actuals is the same as the variance of the forecasts) is not rejected, then the method used to generate these forecasts is an empirically good model to generate future scenarios for long or short positions in the stock indices aforementioned.

Table 2 presents decision rule about the null hypothesis ($H_0 : \sigma_1^2 = \sigma_2^2$) as well as the probability of one tail F-test (two-sample for variances) for the three methods, namely the non-parametric kernel, the dynamic random walk, and the static random walk models.

The rejection of the null hypothesis as well as the higher pmse provides enough strong evidence to discard the dynamic forecasting method in this study. From Table 2 one can realize that forecasts generated from dynamic random walk model are not close to the actual values, this is consistent with what has been found with the higher pmse in Table 1.

Further investigation is needed to assess the performance of the kernel model over the static random walk model. Both forecasts (non-parametric kernel and the statistic random walk) are bootstrapped in order to create ten thousand samples without replacement for each of the earlier mentioned stock indices.

Table 3 presents the confidence interval (C.I) of the ten thousand sample means, average volatility and the mean of the ten thousand sample means of the bootstrapped samples. Table 3 shows that the kernel method provides forecasts with higher mean returns and a slightly high volatility.

A slight difference in terms of confidence interval is noted.

The kernel method seems to possess a slightly larger confidence interval for the ten thousand sample means than the random walk one; however, based on the predicted mean square error criterion and the bootstrapped mean, the kernel model outperforms the random walk model.

Table 3. Bootstrapping results of static random walk and kernel method (in bold).

	C.I of 10000 means	Average volatility of 10000 samples	Mean of 10000 samples
DJIA	[9449.1 - 9831.3] (**9450.3 – 9833.3**)	1797.8 **(1801.2)**	9640.2 **(9641.8)**
S&P 500	[1012.9 – 1058.9] (**1013.9 – 1060.3**)	216.3681 **(218.1315)**	1035.9 **(1037.1**
FTSE 100	[4651.0 – 4809.2] **(4651.9 – 4810.3)**	743.9685 **(745.4792)**	4730.1 **(4731.1)**

Conclusion

This paper has investigated the predictability of stock price behaviour using parametric (random walk) and nonparametric (kernel) models. The study was limited only to stock price indices (S&P 500, DJIA, and FTSE 100). A sample of two thousand and eighty five data points has been divided into two periods; the in-sample and out-sample periods respectively. Preliminary analysis of these stock indices showed that volatility in U.S and U.K stock market was significantly influenced by the price of previous trading (lag one) day and that of the trading from the past fifth day (lag five).

Based on the predicted mean square error, the results suggest that forecasts generated from nonparametric method are closer to actual or observed prices than those generated from the parametric model. Parametric models assume normality and (non-)linearity in the underlying stock price; by relaxing these assumptions we can improve predictability of stock price and be able to make decisions that can help traders to generate scenarios for buy and sell signals in short term.

REFERENCES

Adya M, Collopy F (1998). How effective are neural networks at forecasting and prediction? A review and evaluation; J. Forecasting, 17: 481–495.

Anderson J, Lauvsness SO (2007). Forecasting stock index prices and domestic credit: does cointegration help? Working paper presented at the Nordic Econometric Meeting, www.mtk.ut.ee/198293.

Auestad B, Tjøstheim D (1990). Identification of nonlinear time series: first order characterization and order estimation. Biometrika, 77: 669-687.

Box G, Jenkins G (1994). Time Series analysis forecasting and control, Third edition, Prentice-Hall.

Campbell JY, Shille RJ (1988). The dividend-price ratio and expectations of future dividends and discount factors. Rev. Financ. Study, 1:195-227.

Chalmers JMR, Edelen R, Kadlec GB (2001). On the perils of financial intermediaries setting security prices: The mutual fund wild card option. J. Financ., 56: 2209-2236.

Chatfield C (1995). Positive or negative? Int. J. Forecasting, 11: 501–502.

Epanechnikov V (1969). Nonparametric estimation of multivariate probability densities. Theory of probability and its applications 14: 153-158.

Fan J, Huang L (2001). Goodness-of-fit test for parametric regression models. J. Am. Stat. Assoc. 96: 454.

Greene JT, Hodges CW (2002). The dilution impact of daily fund flows on open-end mutual funds. J. Fin. Econo., 65: 131-158.

Hall P, Racine J, Li Q (2004). Cross-validation and the estimation of conditional probability densities. J. Am. Stat. Ass. 99(486): 1015-1026.

Hardle W, Tsybakov A (1992). Robust locally adaptive nonparametric regression, data analysis statistical inference, Festschrifts in Honour of Friedhelm Fiker Schach, S. and Trenkler, G. pp. 127-144.

Hardle W, Vieu P (1991). Kernel regression smoothing of time series, J. Time Series Anal., 13: 209-23.

Nadaraya W (1964). On estimating regression. Probability Theory and its Applications 10: 186-190.

Niglio M, Perna C (2003). Kernel smoothing for the analysis of climatic data, Quaderni di statistica, p. 5.

Skabar A (2008). A kernel-based technique for diretion-of-change financial time series forecasting. J. Fin. Time Series Forecasting, 14(4): 275-284.

Tkacz G (2001). Neural network forecasting of Canadian GDP growth, Int. J. Forecasting, 17: 57–69.

Tschernig C, Yang W (2000). Nonparametric lag selection for time series, J. Time series Anal., 21(4): 457-487.

Wu JL, Hu YH (1997). The Predictability of U.S. stock-price movements. Working paper, Institute of economics, National Sun Yat-Sen University; www.econ.ntu.edu.tw/sem-paper/97_1/macro_971016.pdf.

Zitzewitz E (2003). Who cares about shareholders? Arbitrage-proofing mutual funds, J. Law, Econ. Organ., 19: 245-280.