Central European Journal of Economic Modelling and Econometrics

Information Flows Around the Globe: Predicting Opening Gaps from Overnight Foreign Stock Price Patterns

Jan G. De Gooijer*, Cees G. H. Diks†, Łukasz T. Gątarek‡

Submitted: 30.05.2012, Accepted: 30.10.2012

Abstract

This paper describes a forecasting exercise of close-to-open returns on major global stock indices, based on high-frequency price patterns that have become available in foreign markets overnight. Generally speaking, out-of-sample forecast performance depends on the forecast method as well as the information that the forecasts are based on. In this paper both aspects are considered. The fact that the close-to-open gap is a scalar response variable to a functional variable, namely an overnight foreign price pattern, brings the prediction exercise in the realm of functional data analysis. Both parametric and non-parametric functional data analysis are considered, and compared with a simple linear benchmark model. The information set is varied by dividing global markets into three clusters, Asia-Pacific, Europe and North-America, and including or excluding price patterns on a per-cluster basis. The overall best performing forecast is nonparametric using all available information, suggesting the presence of nonlinear relations between the overnight price patterns and the opening gaps.

Keywords: Close-to-open gap forecasting, functional data analysis, international stock markets, nonparametric modeling.

JEL Classification: C14, C53, F37, G17.

*University of Amsterdam; e-mail: j.g.degooijer@uva.nl
†University of Amsterdam; e-mail: c.g.h.diks@uva.nl
‡University of Amsterdam; e-mail: gatarek@tlen.pl

1 Introduction

Empirical research in finance has traditionally focused on the analysis of daily stock returns, usually measured as changes in closing prices. However, since trading can be thought of as a continuous-time process, it is also natural to consider returns over other than daily intervals. Recently, some interest has been developed into dividing daily returns into overnight (close-to-open) returns and daytime returns. There is considerable empirical evidence that return dynamics are different over non-trading periods than during trading periods (French and Roll, 1986; Lockwood and Linn, 1990; Hasbrouck, 1991, 1993; and Madhavan, Richardson, Roomans 1997; George and Wang, 2001; Cliff, Cooper, Gulen, 2008). Accordingly, a number of models have been proposed to quantify this phenomenon, often using stocks traded on a particular stock market; see, e.g., Oldfield and Rogalski (1980) and Hong and Wang (2000).
The information revealed in consecutive overnight and day-time returns can also be employed for prediction. In this vein, Zhong (2007) considered predicting daytime volatility of stock prices based on the preceding overnight returns. As far as we know, there have been no attempts to explore the price evolution in a set of foreign stock markets as a result of the information content revealed during non-trading periods of a home market. The economic justification for potential existence of such relations is based on abundant evidence in causality and linkages among the markets based on trading hours returns, see for instance De Gooijer and Sivarajasingham (2007). Furthermore there are a few attempts to explain the relations between overnight returns and subsequent daily returns within the same market, see Branch and Ma (20012), Cliff, Cooper, Gulen (2008), Berkman, Koch, Tuttle, Zhang (2012). The natural consequence is an attempt to broaden the scope of this research to the intra-markets day-night return causalities.
Accurate forecasts of the opening gap, based on foreign information, could be exploited, for instance, by trading futures prior to the opening of the market. With this in mind, one of the aims of this study is to predict the overnight return on an individual stock index of a home market, based on the information content revealed in a set of foreign markets during non-trading hours of the home market.
The formation of price in the non-trading hours might have some implications for the understanding of potential features of prediction mechanism for opening gap. For general introduction to the literature on price discovery and informed trading in pre-open period we refer to Gerety and Mulherin (1994), Biais, Hillion, Spatt (1997) and (1999) or Barclay and Hendershott (2003) and (2004).
In addition to predictability of the opening gaps, we investigate if global markets are informationally efficient in the sense that adding information from clusters of stock indices traded further in the past into the information set does not improve predictive ability.
The final objective is to test the performance of the functional data analysis with respect to its applicability in modeling of financial time series. In the study we employ the benchamrk model of linear regression and we contrast it with complexity

of parametric and nonparametric functional data analysis (P-FDA and NP-FDA) techniques. FDA is a natural alternative to linear regression in this setting, because overnight foreign price patterns can be viewed as continuous functions of time. For background on FDA we refer to Ramsay and Silverman (2005). The appplications of FDA to economic time series can be found in Malfait and Ramsay (2003) and Ramsay and Ramsey (2001). Within empirical finance, where the explanatory variables often depend on some continuous parameter (e.g. price patterns), functional data often arise. For instance, Benko (2006) applies parametric FDA techniques to the analysis of implied volatility functions and yield curve dynamics. Also Mueller, Sen, Stadtmueller (2001) and Mueller, Stadtmueller, Yao (2006) discuss the methods of functional modeling of financial volatility.

This paper is organized as follows. Section 2 formalizes the prediction problem. Section 3 describes the various FDA methods considered in this paper, as well as their corresponding predictive intervals (PIs). The measures used for evaluating the out-of-sample predictions are described in Section 4, and Section 5 provides a description of the data. Section 6 describes the results obtained, and Section 7 discusses the results and concludes. To avoid confusion we like to stress that the adjective ‘functional’ refers to the form of the data and ‘parametric/nonparametric’ to the form of the constraints imposed on the model.

2 The prediction problem

To formalize the prediction problem some notation and definitions are introduced. Let the random variable $Y_{i,s}$ denote the close-to-open gap for stock market i at trading session (trading day) s. The aim is to predict $Y_{i,s}$ based on a pattern of prices (an ordered set of curves), $\chi_{i,s}$, that has been realized within a specific collection of markets other than i overnight. Note that by concatenating the price patterns of the different markets in a fixed order we may represent all prices that $\chi_{i,s}$ contains by a single piece-wise continuous curve $\chi_{i,s}(t)$ indexed by a single time variable t. In practice $\chi_{i,s}(t)$ can not be observed at all possible times t, and we have to use a version of it which is sampled at (typically regular) intervals. Following Ferraty and Vieu (2006) we will denote this discretized version of $\chi_{i,s}$, which mathematically is just a finite-dimensional vector, by $\boldsymbol{x}_{i,s}$. Here data discretized at regular five-minute intervals are used throughout. Also the realized close-to-open returns $y_{i,s}$ are calculated based on discretized observations. If the daily 5-minute price quotes in market i and session s are given by $p_{i,s}(t)$, $t = 1, 2 \dots, T_i$, where T_i is the number of five-minute intervals in a trading day on market i, the close-to-open return for stock index i in trading session s is given by $y_{i,s} = (p_{i,s}(1) - p_{i,s-1}(T_i))/p_{i,s-1}(T_i)$.

To specify the information set on which predictions of close-to-open gaps are to be based, it is convenient to introduce a universal ‘background’ time variable that measures time globally, as opposed to the within-trading session time variable. Time is measured in five-minute units again, and in addition we assume that the universal clock does not run in weekends, between Central European Time (CET) Sat 00:00 and CET Mon 00:00; a period during which all markets are closed simultaneously. Since a 24 hour day contains 288 five-minute intervals, each observation can be represented as $x_{i,s}(t) = \tilde{x}_i(288 \times s + c_i + t)$, where $288 \times s + c_i + t$ is the universal time corresponding to a quote in session s of market i at trading-session time t. The shift $c_i \in 0, \ldots, 287$ represents the opening time of market i, again in five-minute units. Note that $\tilde{x}_i(\cdot)$ is only defined for universal times t at which market i is open, and is not available otherwise.

The prediction variable of interest is the close-to-open return in market i for trading session s, given by $y_{i,s} = (x_{i,s}(1) - x_{i,s-1}(T_i))/x_{i,s-1}(T_i)$, where $s-1$ denotes the last session prior to session s during which market i was open. In universal time, $y_{i,s}$ materializes at time $t_{i,s}^{\text{open}} = 288 \times s + c_i$.

Below, several different specifications are considered, which use various amounts of the information available. Based on trading hours, three global clusters of markets can be distinguished; in order, the Asian-Pacific markets, the European markets, and the American markets. If we include or exclude information on a per-cluster basis, for each home market there are three possible information sets that can be extracted from the overnight price patterns. For instance, forecasts for the opening gap in each of the Asian-Pacific markets can be based on the previous overnight price pattern in the North-American markets, the European markets, or both. In an analogue way, three different information sets can be defined for the European and the North-American markets. In general we denote these information sets as 'cluster(-1)' (only the previous cluster), 'cluster(-2)' (the second-most recent cluster) and 'cluster(-1)-cluster(-2)' (both the previous cluster and the second most recent cluster). In cases of missing data as a result of a holiday in either one of the explanatory variables or the home market of interest, the corresponding explanatory data and opening gap are excluded from the analysis.

The original sample for each market is split into two sub-samples: a learning sample containing the units $\{(\boldsymbol{x}_{i,s}, y_{i,s})_{s=1,\ldots,k_i}\}$, and a testing sample containing the units $\{(\boldsymbol{x}_{i,s}, y_{i,s})_{s=k_i+1,\ldots,S}\}$ where k_i $(i \in 1, \ldots, M)$ denotes the number of observations in the learning sample. The learning sample allows us to construct the various models, while the testing sample is used for making actual predictions and evaluating predictive performance.

3 Functional data analysis

In our description of functional data analysis we consider predicting the opening gap for a specific market i. For notational convenience, the subscript i is dropped from the respective random variables. Let $(\boldsymbol{\chi}_s, Y_s)_{s=1,\ldots,k}$ be $k = k_i$ pairs of random variables, identically distributed as $(\boldsymbol{\chi}_i, Y_i)$ but not necessarily independent, and taking values in $\mathcal{E} \times \mathbb{R}$, where $(\mathcal{E}, d)$ is a semi-metric space with semi-metric d. In addition, it is assumed that $(\boldsymbol{\chi}_s, Y_s)$ is strictly stationary. The aim is to predict the unobserved scalar response variable Y_s from the curve(s) $\boldsymbol{\chi}_s$ (covariates). The idea behind FDA is that similar patterns χ lead to similar responses y. This can be either modeled parametrically, for instance through

$$Y = f(\chi; \beta) + \text{error},$$

where β is a finite-dimensional parameter, or it can be modeled nonparametrically.
The nonparametric approach does not rely on any specific finite-dimensional specification, but only exploits the fact that similar patterns have similar responses. This leads to prediction by analogy: for a given observed pattern χ_s of overnight returns, use the data to identify sessions s' for which the pattern $\chi_{s'}$ was close to χ_s. The responses $y_{s'}$ observed for these 'close' patterns are good representatives of Y_s given χ_s. To make this operational it is convenient to introduce a notion of closeness between patterns, for which many different semi-metrics are being used in practice. Given a semi-metric, standard nonparametric methods, such as kernel-weighing methods are readily available to construct predictors consisting of weighted averages of the responses $y_{s'}$ for which $\chi_{s'}$ was closest to χ_s. It is beyond the scope of the present paper to give an in-depth introduction to functional data analysis. For an exposition of the various available techniques and many applications, please refer to Ferraty and Vieu (2006).

3.1 Nonparametric FDA

Ferraty and Vieu (2006) emphasize that the object of interest in a nonparametric functional data analysis need not always be the conditional mean $E(Y|\chi = \chi)$, and they propose forecasts based on the conditional median and conditional mode as well. This problem can be viewed as follows. Suppose that there exists a function $r(\cdot)$ modeling the relationship between Y and $\boldsymbol{\chi}$ and that $r(\cdot)$ is defined through the conditional distribution. Given a convex loss function $\ell(\cdot)$ with a unique minimum, define $r(\cdot)$ such that it minimizes the mean $\mathbb{E}(\ell(Y - a)|\boldsymbol{\chi} = \chi)$ with respect to a.
A nonparametric estimator of $r(\cdot)$ provides a nonparametric predictor $\hat{Y}$ in terms of $\boldsymbol{\chi}$. Using this principle, we consider three nonparametric predictors, based on different loss functions. Assuming $(\boldsymbol{\chi}_s, Y_s)$ is α-mixing, Ferraty and Vieu (2005, 2006) proved almost complete convergence, and established the corresponding rates of convergence, of these three nonparametric functional predictors, which are also considered here. Also these authors established the rates of convergence of the predictors.

Conditional mean: It is well-known that taking $\ell(u) = u^2$ leads to the conditional mean function $r(\chi) = \mathbb{E}(Y|\boldsymbol{\chi} = \chi)$.
Recall that the model is to be based on the observed k pairs $(\boldsymbol{x}_s, y_s)_{s=1,\ldots,k}$ of identically distributed random variables, where $\boldsymbol{x}_s$ is a discretized version of the pattern $\boldsymbol{\chi}_s$. Let $\boldsymbol{x}$ be an observed curve (overnight foreign price pattern) at which

one would like to estimate the regression function. Then, using the Nadaraya-Watson kernel density estimator, the one-step-ahead conditional mean predictor (measured in five-minute units) is defined as:

$$\hat{y}^{\text{mean}} = \sum_{s=1}^{k} y_s W(\boldsymbol{x}_s, \boldsymbol{x}).$$

Here $W(\cdot)$, the so-called kernel weight, is defined as:

$$W(\boldsymbol{x}_s, \boldsymbol{x}) = K(d(\boldsymbol{x}_s, \boldsymbol{x})/h) \Big/ \sum_{r=1}^{k} K(d(\boldsymbol{x}_r, \boldsymbol{x})/h),$$

where h denotes the bandwidth, $K(\cdot)$ the kernel function, and $d(\boldsymbol{x}_s, \boldsymbol{x})$ is any semi-metric between $\boldsymbol{x}_s$ and $\boldsymbol{x}$.

Conditional median: In this case the loss function is given by $\ell(u) = |u|$. Then the conditional median function is given by:

$$r(\chi) = \inf\{y : F(y|\boldsymbol{\chi} = \chi) \geq 1/2\},$$

where $F(\cdot|\cdot)$ is the conditional distribution function of Y given $\boldsymbol{\chi} = \chi$. Consequently, the one-step-ahead nonparametric functional predictor of the conditional median is defined as

$$\hat{y}^{\text{med}} = \inf\left\{y : \hat{F}(y|\chi) \geq 1/2\right\},$$

where

$$\hat{F}(y|\chi) = \sum_{s=1}^{k} W(\boldsymbol{x}_s, \boldsymbol{x}) \mathbf{1}_{\{y_s \leq y\}},$$

with $\mathbf{1}_{\{A\}}$ denoting the indicator function of set $\{A\}$, is the estimated conditional cumulative distribution function (CDF) of Y given $\boldsymbol{\chi} = \chi$.

Conditional mode: In this case we have a non-convex loss function with a unique minimum $\ell(u) = 0$ when $u = 0$, and $\ell(u) = 1$ otherwise. The loss function becomes:

$$r(\chi) = \arg\max_{y \in \mathbb{R}} f(y|\boldsymbol{\chi} = \chi),$$

where $f(\cdot|\cdot)$ denotes the conditional density function of Y given $\boldsymbol{\chi} = \chi$. Hence, g Given the observed data, the nonparametric functional predictor of the conditional mode is given by:

$$\hat{y}^{\text{mode}} = \arg\max_{y \in \mathbb{R}} \sum_{s=1}^{k} K(|y - y_s|/h) W(\boldsymbol{x}_s, \boldsymbol{x}),$$

where, for ease of notation, we assume that the same kernel function $K(\cdot)$ and bandwidth h apply in the y direction.
Thus the predictive conditional density of y given the observed curve $\boldsymbol{x}$ can be represented as the probability measure over the observations $y_{s_{\{s=1,\dots,k\}}}$. This measure distributes the probability over $y_{s_{\{s=1,\dots,k\}}}$ proportionaly to the distance of respective $\boldsymbol{x}_s$ from $\boldsymbol{x}$ quantified via $W(\boldsymbol{x}_s, \boldsymbol{x})$. Thus we can predict by conditional mean, median or mode of this distribution. If the sample is dense enough, any quantile (percentile) predictor can be specified in similar way as prediction via conditional median. Due to the construction of the predictors, in the testing subsample we can apply any prediction evaluation criteria. In each period the observed curve $\boldsymbol{x}$. Thus the prediction can be obtained conditional on this curve. Then the predicted y can be compared with the observable counterpart in the testing sample.

Semi-metric: As a measure of closeness a standard semi-metric based on functional principal component analysis (FPCA) is used. FPCA builds upon ideas from classical principal component analysis. Assuming $\mathbb{E}(\int \boldsymbol{\chi}^2(t)dt) < \infty$, it can be shown that the functional random variable $\boldsymbol{\chi}$ can be written as:

$$\boldsymbol{\chi} = \sum_{k=1}^{\infty} \int \chi_s(t) e_k(t) dt e_k,$$

where e_k are orthonormal eigenfunctions of the covariance operator

$$\Gamma_{\boldsymbol{\chi}}(t, t') = \mathbb{E}(\boldsymbol{\chi}(t)\boldsymbol{\chi}(t'))$$

associated with the eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq 0$. A truncated version of this expansion forms the basis of the FPCA semi-metric. In particular, the empirical version of this semi-metric is defined, in the case $\boldsymbol{x}$ is an observed pattern of a single variable consisting of T consecutive observations, as

$$\boldsymbol{d}_q(\boldsymbol{x}_s, \boldsymbol{x}) = \sqrt{\sum_{k=1}^{q} \Big(\sum_{j=1}^{T} (\boldsymbol{x}_s(j) - \boldsymbol{x}(j))[\hat{\boldsymbol{e}}_k]_j \Big)^2}, \tag{1}$$

where, q denotes the number of retained principal components in the FPCA expansion, with q much smaller than T. It is straightforward to generalize (1) into a semi-metric suitable for multiple covariates. In that case, the parameter q need to be chosen. Comparing the prediction performance using the evaluation measures introduced in the next section, we noticed that values of $q \geq 6$ did not alter the results. Hence, we fixed q at 5 for each market.

Implementation: For general NP-FDA prediction `R/S+`-routines are available at the website: *http://www.lsp.ups-tlse.fr/staph/npfda*; see also Ferraty and Vieu

(2006, Chapter 7) for some details. We modified these routines for our purpose. The resulting `R`-codes, the datasets, and a brief description can be obtained from the authors. Two relatively "simple" practical aspects, concern the choice of the kernel function and the associated bandwidth. Throughout the analysis we employed the quadratic kernel: $K(u) = 1\frac{1}{2}(1-u^2)_{[0,1]}(u)$. The quadratic kernel was proven to be the leading one with respect to the statistical significance of the results investigated by bootstrapping. The bandwidth choice follows the data-driven procedure as described in Ferraty, Vieu (2005, Section 4.4), i.e. h is chosen in order to minimize $\sum_{s=1}^{k}\left|\hat{y}_s^{(.)} - y_s^{(.)}\right|$, where $\hat{y}_s^{(.)}$ denotes the value of a predictor based on one of the FDA methods discussed above.

Another practical issue is that price levels may differ considerably across trading days. To obtain price patterns that are comparable across trading days all price patterns in $\boldsymbol{x}$ are expressed relative to the opening price of that day.

3.2 Functional Parametric Regression

To enable comparisons with a predictor based on parametric functional data analysis we consider a functional parametric regression model. The simplest of these models establishes a linear relationship between a functional covariate $\chi_s(t)$ and the response variable Y_s, according to

$$Y_s = \beta_0 + \int_0^T \chi_s(t)\beta(t)dt + \varepsilon, \tag{2}$$

where $\{\varepsilon\}$ is a sequence of i.i.d. random variables such that $\mathbb{E}(\varepsilon|\chi_s(t)) = 0$ and $\mathbb{E}(\varepsilon^2|\chi_s(t)) = \sigma^2 < \infty$.

A popular approach to reduce the number of degrees of freedom in (2) is to use a truncated functional basis expansion, similar to the truncation applied in the NP-FDA case. There are three prominent examples of functional bases: Fourier, Polynomial and B-spline. Here, following Ramsay and Silverman (2005, Chapter 15), we adopt a set of Fourier (orthonormal) basis function $\theta_k(t)$, i.e.

$$\beta(t) = \sum_{k=1}^{K_\beta} b_k\theta_k(t) = \boldsymbol{b}'\boldsymbol{\theta}(t), \tag{3}$$

where $\theta_{2k-1}(t) = \sin k\omega t$ and $\theta_{2k}(t) = \sin k\omega t$, K_β denotes the length of the set, and where $\boldsymbol{\theta}(t) = (\theta_1(t), \ldots, \theta_{K_\beta}(t))'$ and $\boldsymbol{b} = (b_1, \ldots, b_{K_\beta})'$. Similarly, $\chi_s(t)$ can be expanded in another set of Fourier basis function $\psi_{k,s}(t)$ of length K_z as follows

$$\chi_s(t) = \sum_{k=1}^{K_z} c_{s,k}\psi_k(t) = \boldsymbol{c}_s'\boldsymbol{\psi}(t), \tag{4}$$

where $\boldsymbol{\psi}(t) = (\psi_1(t), \ldots, \psi_{K_z}(t))'$ and $\boldsymbol{c}_s = (c_{s,1}, \ldots, c_{s,K_z})'$. Inserting (3) and (4) into (2), and using the data in the learning sample, yields

$$Y_s = \beta_0 + \boldsymbol{C}_s \boldsymbol{J} \boldsymbol{b} + \varepsilon_s, \quad (s \in 1, \ldots, k),$$

where $\boldsymbol{J}$ is a $K_z \times K_\beta$ matrix defined by $\boldsymbol{J} = \int \boldsymbol{\psi}(t)\boldsymbol{\theta}'(t)dt$, and where the $k \times K_z$ matrix is given by $\boldsymbol{C} = \{c_{s,k} : s = 1, \ldots, k, k = 1, \ldots, K_z\}$. The notation can be further simplified by defining a $(K_\beta+1)$-vector $\boldsymbol{\xi} = (\beta_0\ \boldsymbol{b}')'$ and a $k \times (K_\beta+1)$ matrix $\boldsymbol{Z} = [\mathbf{1}\ \boldsymbol{CJ}]$. Thus, the resulting functional regression model has the same structure as the classical linear regression model. Consequently, the augmented parameter vector $\boldsymbol{\xi}$ can be estimated by least squares, i.e. $\hat{\boldsymbol{\xi}} = (\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}\boldsymbol{y}$. Clearly, the above setup can be easily modified into a specification with more than one functional covariate. The choice of the numbers of basis functions, K_z and K_β, is a trade-off between information loss and computational costs. In the present study K_z and K_β were set equal to 15. For the specific data used in the present study we verified that there was no gain in performance by increasing the number of basis functions beyond this number.
Given $\hat{\boldsymbol{\xi}}$ and the set of basis functions $\boldsymbol{\theta}(t)$, estimates $\hat{\beta}(t)$ of $\beta(t)$ can be obtained. Then, using (2), the predictor for Y_s may be constructed as

$$\hat{y}_s^0 = \hat{\beta}_0 + \int_0^T \chi_s(t)\hat{\beta}(t)dt,$$

where s runs over the collection of available out-of-sample sessions, denoted as $\mathcal{S}$. The number of evaluation sample points available (size) will be denoted by $|\mathcal{S}|$. One feature of above setup is that the conditioning takes place on the same information set as used in predicting a response variable via NP-FDA.

3.3 Linear regression model

To include a benchmark model to compare both the parametric and nonparametric FDA results with, we consider a linear regression model which uses as the explanatory variables a constant plus the total overnight returns of the foreign stock indices. The use of overnight returns rather than the complete set of 5-minute observations ensures that the model is parsimonious. Although there are other ways to construct a parsimonious linear model, we have chosen to focus on overnight returns, since, at least if markets are close to being informationally efficient, these are expected to reflect most of the relevant information regarding the development of fundamentals underlying the respective indices.

3.4 Predictive intervals

Following De Gooijer and Gannoun (2000) we consider two types of PIs: the conditional percentile interval (CPI) and the shortest conditional modal interval

(SCMI). The CPI with nominal coverage probability γ is given by $\left(\xi_{\frac{1-\gamma}{2}}(\chi), \xi_{\frac{1+\gamma}{2}}(\chi)\right)$, where $\xi_\alpha(\chi)$ denotes the α-th quantile of the conditional distribution of Y given $\boldsymbol{\chi} = \chi$ i.e. the solution of $F(\xi_\alpha(y|\boldsymbol{x})) = \alpha$ with respect to y. A natural estimator for the CPI is $\left(\hat{\xi}_{\frac{1-\gamma}{2}}(\boldsymbol{x}), \hat{\xi}_{\frac{1+\gamma}{2}}(\boldsymbol{x})\right)$, where the estimated quantiles satisfy $\hat{F}(\hat{\xi}_\alpha(\boldsymbol{x})|\boldsymbol{x}) = \alpha$.
The SCMI with nominal coverage probability γ is:

$$(a, b) = \underset{(c,d)}{\arg\min}\, \{d - c \mid F(d|\boldsymbol{x}) - F(c|\boldsymbol{x}) \geq \gamma\},$$

and a natural estimator for the SCMI is:

$$(\hat{a}, \hat{b}) = \underset{(c,d)}{\arg\min} \left\{d - c \mid \hat{F}(d|\boldsymbol{x}) - \hat{F}(c|\boldsymbol{x}) \geq \gamma\right\},$$

where, as before, the estimated conditional CDF is given by

$$\hat{F}(y|\chi) = \sum_{s=1}^{k} W(\boldsymbol{x}_s, \boldsymbol{x}) \mathbf{1}_{\{y_s \leq y\}}.$$

The SCMI is particularly suitable when the predictive density is asymmetric. For symmetric and unimodal distributions SCMI reduces to CPI.

4 Prediction evaluation measures

Four prediction evaluation criteria will be adopted. The first two criteria are measures for evaluating point predictions, while the latter two are concerned with evaluation of the PIs. The first measure is the mean-squared prediction error, given by

$$\text{MSE} = |\mathcal{S}|^{-1} \sum_{s \in \mathcal{S}} (\hat{y}_s^{(\cdot)} - y_s)^2,$$

where $\hat{y}_s^{(\cdot)}$ denotes the value of a predictor based on the respective NP-FDA, P-FDA approaches discussed above, or based on predictions obtained from the benchmark linear (multivariate) regression model.
From a practical point of view there often is some interest in predicting the sign of a return on the index rather than its precise value. For instance, many trading strategies are based on sign predictions. The following measure evaluates the point predictions by comparing the predicted and realized signs of the close-to-open gaps:

$$d^{\text{sgn}} = |\mathcal{S}|^{-1} \sum_{s \in \mathcal{S}} \mathbf{1}_{\{\text{sgn}(\hat{y}_s^{(\cdot)}) \neq \text{sgn}(y_s)\}}.$$

This measures the fraction of cases where the sign of the predicted close-to-open return and the actual return differ. Up to a constant factor, this is a generalization

of the mean-squared prediction error applied to the signs of $\hat{y}_{i,s}^{(\cdot)}$ and $y_{i,s}$ rather than the actual values.
The remaining two criteria are used to evaluate the PIs. We compute the average width of PIs, as well as the empirical coverage probability. Denote by $\hat{\ell}_i$ and $\hat{u}_i$ the estimated lower and upper interval limits, respectively, obtained with one of the specific methods described in section 3.4 (CPI or SCMI). The (root-mean-squared) average width of the corresponding PIs is calculated as

$$v = \sqrt{|\mathcal{S}|^{-1} \sum_{s \in \mathcal{S}} (\hat{u} - \hat{\ell})^2}.$$

The empirical coverage probability is computed as

$$p_c = |\mathcal{S}|^{-1} \sum_{s \in \mathcal{S}} \mathbf{1}_{\{y_s \in (\hat{\ell}, \hat{u})\}}.$$

Ideally, a PI has coverage probability equal to the nominal coverage probability, while having the smallest possible average width. As an overall measure of the capability of the intervals to 'capture much probability' while having a small width, we also calculate the average PI length divided by the average coverage probability, $q = \bar{v}/\bar{p}_c$.

5 Data

The data consist of intra-day quotations of the following nine ($M = 9$) major stock market indices: the All Ordinary Composite Stock Index (AU), the Nikkei 225 Stock Index (JP), the Hang Seng Stock Index (HK), the FTSE 100 Share Index (UK), the Frankfurt DAX 30 Composite Stock Index (DE), the CAC 40 Composite Stock Index (FR), the Zurich Swiss Market Composite Index (CH), the Dow Jones Industrial Average (US), and the Toronto 300 Composite Stock Index (CA). All indices are retrieved from the Bloomberg data bank. The period covered is from 24th September 2007 to 8th May 2008. Bloomberg offers one, five, and 15-minute quotations. In the case of one-minute quotations the market microstructure noise is more pronounced. Hence, we decided to use five-minute quotes.
For each of the y-variables, i.e. a close-to-open gap of one of the 9 stock indices, we consider various specifications, differing in terms of the information included in $\boldsymbol{x}$. To limit the number of possible specifications, information is added to the $\boldsymbol{x}$-variable cluster-wise, where the three global clusters are the Asian-Pacific cluster (JP, AU, HK), the European cluster (DE, FR, CH, UK), and the North-American cluster (US, CA). The first specification only contains the stock index patterns from the 'previous' cluster, for instance using the Asian-Pacific cluster to predict the opening gap of FR. This specification is referred to as 'Cluster(-1)'. The second specification only uses information from the before-last cluster. For instance, using the North-American cluster to predict the opening gap of FR. This is denoted by 'Cluster(-2)'. Finally,

specification 'Cluster(-1)–Cluster(-2)' contains the patterns from the last two clusters in the $\boldsymbol{x}$-variable.

For specifications with one explanatory functional variable the dataset is organized in the form of a matrix. The predictions are based on the explanatory functional variable $(x_{i,s}(t)/x_{i,s}(1)) \times 100$ which resulted in MSEs that were at least twice as small as for three alternative transformations. When the session overlaps the opening gap is predicted with the information set spanning over the quotation in the foreign market up to 5-minutes before the opening. Table 1 provides an overview of the respective

Table 1: Trading times expressed in CET

		$y_{i,s}$								
$\boldsymbol{x}_{is}$	CET	AU	JP	HK	UK	DE	FR	CH	US	CA
AU	00:00-06:05				74 (118)	74 (116)	74 (118)	74 (115)	74 (115)	74 (117)
JP	01:00-06:35				57 (110)	57 (109)	57 (109)	57 (110)	57 (108)	57 (110)
HK	03:00-09:00				50 (114)	50 (113)	50 (115)	50 (113)	51 (111)	51 (113)
UK	09:00-17:30	103 (86)	103 (84)	103 (82)					66 (117)	66 (119)
DE	09:00-17:35	104 (86)	104 (83)	104 (82)					66 (115)	66 (117)
FR	09:00-17:25	102 (86)	102 (84)	102 (82)					66 (117)	66 (119)
CH	09:00-17:30	101 (85)	101 (83)	101 (81)					66 (114)	66 (116)
US	14:30-21:00	79 (85)	79 (84)	79 (81)	79 (86)	79 (85)	79 (85)	79 (85)		
CA	14:30-21:05	80 (86)	80 (84)	80 (82)	80 (87)	80 (86)	80 (86)	80 (87)		

Total number T_i of 5-minute quotes per trading session s, when predictions are based on a single explanatory variable, and (in parentheses) the total number k_i of 5-minute quotes in the learning sample.

trading times (expressed in CET). Additionally, it shows information on the total number T_i of five-minute quotes per trading session when predictions are based on one explanatory functional variable. In the case of two- or more explanatory functional variables, the total number of five-minute quotes varies with specifications and trading times. To safe space, we have not included this information in the paper. Further, note that Table 1 contains the total number k_i of five-minute quotes (in parentheses) in the learning sample. The testing sample for each specification contains 35 days (curves). The complete dataset was prepared with great care, taking into account national holidays in all markets by considering overnight returns.

6 Results

In this section the observed performance measures, MSE and d^{sgn}, are presented and discussed. To facilitate the interpretation it is convenient to start at the aggregate level and then look for particular differences between clusters and individual stock indices.

6.1 Out-of-sample MSEs

Table 2 shows aggregate out-of-sample MSEs observed for each of the three NP-FDA and the two P-FDA specifications. Standard errors are given in brackets. The specifications that performed best in terms of the MSE criterion is indicated by an underlined entry. The aggregate MSEs quoted are obtained by averaging MSEs for identical specifications across the varoius y-variable. Global aggregate MSE values are provided, as well as individual aggregates for the three 'clusters' Asia-Pacific, Europe and North-America. The standard errors of the aggregates are calculated from the standard errors of the individual MSEs, where these were assumed to be uncorrelated. The idea behind presenting aggregate MSEs across certain groups of ($\boldsymbol{x}$,y)-pairs is that they provide a measure of a method's average accuracy across ($\boldsymbol{x}$,y)-pairs randomly selected from that group.
At the overall aggregate level, the best specification in terms of MSE turned out to be the mean-based NP-FDA method using information from both overnight clusters (i.e. all information that has become available overnight). Notably, the average MSE (0.47) observed for that specification is considerably smaller than the average MSE values observed for the linear model specifications, which suggests the presence of a nonlinear relation between $\boldsymbol{x}$ and y. For each of the three NP-FDA methods one can observe that the specification using only information from cluster(-1) performs better than using that from cluster(-2). This is improved upon by using cluster(-1)-cluster(-2) suggesting that the information revealed by cluster(-1) is not reflecting all available information regarding the opening gap. We next consider the aggregate results at the cluster level. The observed pattern for Europe coincides with that of the global aggregate level. This is not the case for North-America. Although the mean-based NP-FDA is again performing very well, the linear model is performing just as well for North-America, but based on a different information set (information from Europe only, rather than from Europe and Asia-Pacific). A possible interpretation might be that although there is extra information in the patterns of the Asia-Pacific markets that could be exploited for prediction, the linear model is more parsimonious and therefore able to achieve equal out-of-sample performance for the small dataset considered here. The aggregate NP-FDA and P-FDA results for Asia-Pacific suggest that the opening gaps in the Asian markets are determined by the North-American stock index patterns as well as the European. Note that when just a linear regression specification is used, the best performing forecast seems to suggest that Asia-Pacific is only affected by the patterns in the European markets, which would be highly counterintuitive. The presence of nonlinear dependence of the Asian-Pacific opening gaps on the observed overnight price patterns may explain this. Indeed, the mean-based NP-FDA and the P-FDA method achieve smaller out-of-sample MSEs based on trading patterns in both European and North-American markets. Table 3 shows the out-of-sample MSE values observed for the individual markets. It can be observed that the results for the individual European markets roughly coincide with that of the global aggregate. The best model is the mean-

Table 2: Average MSEs with standard deviations in parentheses.

	nonparametric			parametric	
$\boldsymbol{x}_s$	Mean	Median	Mode	PFDA	Lin. reg
			Overall aggregate		
Cluster(-1)	0.57 (0.05)	0.80 (0.08)	0.65 (0.06)	0.73 (0.11)	0.64 (0.07)
Cluster(-2)	0.70 (0.07)	1.12 (0.11)	0.78 (0.09)	0.67 (0.11)	0.61 (0.07)
Cluster(-1)-cluster(-2)	<u>0.47</u> (0.05)	0.78 (0.08)	0.65 (0.07)	0.50 (0.12)	0.75 (0.08)
			Europe		
Cluster(-1)	0.55 (0.09)	0.59 (0.09)	0.61 (0.10)	0.64 (0.06)	0.60 (0.11)
Cluster(-2)	0.60 (0.11)	0.77 (0.12)	0.66 (0.11)	0.69 (0.05)	0.63 (0.11)
Cluster(-1)-cluster(-2)	<u>0.39</u> (0.07)	0.61 (0.10)	0.53 (0.10)	0.44 (0.05)	0.63 (0.11)
			North-America		
Cluster(-1)	0.45 (0.08)	0.77 (0.14)	0.45 (0.10)	0.53 (0.05)	<u>0.39</u> (0.08)
Cluster(-2)	0.56 (0.13)	1.12 (0.19)	0.53 (0.11)	0.53 (0.05)	0.49 (0.10)
Cluster(-1)-cluster(-2)	<u>0.39</u> (0.10)	0.52 (0.09)	0.45 (0.10)	0.48 (0.07)	0.57 (0.12)
			Asia-Pacific		
Cluster(-1)	0.67 (0.10)	1.11 (0.18)	0.85 (0.11)	0.98 (0.31)	0.87 (0.15)
Cluster(-2)	0.93 (0.13)	1.60 (0.25)	1.11 (0.19)	0.76 (0.32)	0.67 (0.13)
Cluster(-1)-cluster(-2)	0.64 (0.09)	1.20 (0.18)	0.96 (0.15)	<u>0.61</u> (0.35)	1.02 (0.16)

based NP-FDA, except for DE, for which P-FDA performs slightly, but insignificantly, better. The best-performing NP-FDA specification is that using information from both clusters, while the linear model performs best with a parsimonious specification, based on information from the latest available cluster only. The best performing model for the US opening gap is the linear regression model, based on the information revealed by the European markets overnight, in line with what one would expect for informationally efficient markets. A similar result holds for Canada, although in that case the mean-based NP-FDA performed slightly (very insignificantly) better. The opening in Japan seems to be affected by the North-American markets only, both in terms of the linear benchmark and the NP-FDA. The P-FDA results might indicate that Europe also has some effect on Japan, but this is insignificant. The best performing model for Australia is the linear model based on information from Europe. However, the observed MSEs for several of the other specifications are almost as small, and well within the standard error. Likewise, for Hong Kong one of the linear models, one of the NP-FDA and one of the P-FDA methods perform practically equally well.

Table 3: Market specific MSEs with standard deviations in parentheses.

		nonparametric			parametric	
$\boldsymbol{x}_s$	y_s	Mean	Median	Mode	PFDA	Lin. reg
HK-AU-JP	FR	0.60 (0.18)	0.79 (0.19)	0.68 (0.22)	0.75 (0.12)	0.67 (0.25)
US-CA		0.71 (0.25)	0.97 (0.27)	0.71 (0.25)	0.85 (0.14)	0.73 (0.23)
HK-AU-JP-US-CA		<u>0.44</u> (0.11)	0.65 (0.22)	0.64 (0.22)	0.61 (0.09)	0.75 (0.22)
HK-AU-JP	UK	0.44 (0.11)	0.39 (0.08)	0.44 (0.11)	0.44 (0.10)	0.43 (0.12)
US-CA		0.50 (0.16)	0.93 (0.29)	0.59 (0.19)	0.57 (0.09)	0.49 (0.14)
HK-AU-JP-US-CA		<u>0.33</u> (0.08)	0.47 (0.11)	0.44 (0.11)	0.31 (0.07)	0.51 (0.14)
HK-AU-JP	CH	0.63 (0.21)	0.72 (0.27)	0.73 (0.27)	0.79 (0.11)	0.72 (0.30)
US-CA		0.68 (0.29)	0.67 (0.25)	0.75 (0.29)	0.74 (0.10)	0.73 (0.30)
HK-AU-JP-US-CA		<u>0.45</u> (0.18)	0.84 (0.30)	0.57 (0.27)	0.58 (0.10)	0.71 (0.30)
HK-AU-JP	DE	0.54 (0.19)	0.45 (0.12)	0.59 (0.18)	0.59 (0.12)	0.56 (0.17)
US-CA		0.49 (0.14)	0.51 (0.13)	0.57 (0.16)	0.58 (0.10)	0.56 (0.13)
HK-AU-JP-US-CA		0.33 (0.19)	0.46 (0.12)	0.47 (0.18)	<u>0.26</u> (0.12)	0.56 (0.17)
DE-FR-CH-UK	US	0.39 (0.07)	0.83 (0.20)	0.40 (0.08)	0.48 (0.07)	<u>0.35</u> (0.09)
HK-AU-JP		0.56 (0.15)	1.05 (0.26)	0.53 (0.13)	0.57 (0.07)	0.51 (0.13)
HK-AU-JP-DE-FR-UK-CH		0.37 (0.09)	0.48 (0.10)	0.42 (0.09)	0.40 (0.08)	0.57 (0.16)
DE-FR-CH-UK	CA	0.50 (0.15)	0.71 (0.19)	0.49 (0.19)	0.57 (0.08)	0.42 (0.14)
HK-AU-JP		0.55 (0.20)	1.18 (0.29)	0.53 (0.18)	0.49 (0.07)	0.47 (0.16)
HK-AU-JP-DE-FR-UK-CH		<u>0.41</u> (0.17)	0.56 (0.14)	0.47 (0.17)	0.55 (0.11)	0.57 (0.18)
US-CA	JP	<u>0.35</u> (0.08)	0.63 (0.15)	0.44 (0.10)	0.45 (0.10)	0.50 (0.09)
DE-FR-CH-UK		0.67 (0.10)	1.07 (0.22)	0.70 (0.12)	0.57 (0.14)	0.56 (0.08)
DE-FR-CH-UK-US-CA		0.50 (0.08)	0.63 (0.12)	0.65 (0.13)	0.41 (0.14)	0.63 (0.12)
US-CA	AU	0.32 (0.10)	0.42 (0.16)	0.41 (0.13)	0.33 (0.07)	0.28 (0.08)
DE-FR-CH-UK		0.37 (0.08)	0.52 (0.11)	0.39 (0.09)	0.39 (0.12)	<u>0.26</u> (0.07)
DE-FR-CH-UK-US-CA		0.29 (0.08)	0.28 (0.07)	0.28 (0.07)	0.35 (0.07)	0.32 (0.08)
US-CA	HK	1.34 (0.26)	2.28 (0.49)	1.70 (0.29)	2.17 (0.93)	1.83 (0.42)
DE-FR-CH-UK		1.76 (0.36)	3.20 (0.70)	2.25 (0.56)	1.31 (0.95)	1.20 (0.36)
DE-FR-CH-UK-US-CA		1.14 (0.26)	2.68 (0.52)	1.95 (0.41)	<u>1.07</u> (1.03)	2.10 (0.45)

Note that for none of the y-variables the median-based NP-FDA or the mode-based NP-FDA method performed best in terms of the MSE criterion.

Table 4: Aggregate performance measure d^{sgn} for the sign predictions

$\boldsymbol{x}_s$	Mean	Median	Mode	PFDA	Lin. reg
			Overall aggregate		
Cluster(-1)	0.27 (0.02)	0.31 (0.03)	0.40 (0.03)	0.32 (0.03)	<u>0.26</u> (0.02)
Cluster(-2)	0.36 (0.03)	0.39 (0.03)	0.28 (0.02)	0.37 (0.03)	0.35 (0.03)
Cluster(-1)-cluster(-2)	0.31 (0.03)	0.30 (0.03)	0.37 (0.03)	0.36 (0.03)	0.38 (0.03)
			Europe		
Cluster(-1)	0.22 (0.03)	0.25 (0.04)	0.54 (0.04)	0.32 (0.04)	0.19 (0.03)
Cluster(-2)	0.31 (0.04)	0.36 (0.04)	<u>0.15</u> (0.03)	0.33 (0.04)	0.34 (0.04)
Cluster(-1)-cluster(-2)	0.29 (0.04)	0.23 (0.04)	0.33 (0.04)	0.37 (0.04)	0.37 (0.04)
			North-America		
Cluster(-1)	0.43 (0.06)	0.39 (0.06)	0.41 (0.06)	0.44 (0.06)	0.39 (0.06)
Cluster(-2)	0.43 (0.06)	0.45 (0.06)	<u>0.26</u> (0.05)	0.42 (0.06)	0.39 (0.06)
Cluster(-1)-cluster(-2)	0.29 (0.05)	0.37 (0.06)	0.33 (0.05)	0.34 (0.06)	0.46 (0.06)
			Asia		
Cluster(-1)	<u>0.25</u> (0.04)	0.31 (0.05)	0.31 (0.05)	0.26 (0.04)	0.27 (0.04)
Cluster(-2)	0.38 (0.05)	0.41 (0.05)	0.48 (0.05)	0.38 (0.05)	0.33 (0.05)
Cluster(-1)-cluster(-2)	0.34 (0.05)	0.36 (0.05)	0.47 (0.05)	0.34 (0.05)	0.35 (0.05)

Standard deviations in parentheses.

6.2 Sign forecast performance

Interestingly, this picture changes rather substantially if we consider the performance measure, d^{sgn}, for the sign predictions of the opening gap. Table 4, providing the aggregate values, show that at the global aggregate level the linear model based on information from the last cluster performs best, closely (with an insignificant difference) followed by the mode-based NP-FDA using information of the before-last cluster. The sign of the close-to-open gap in the European stock indices appears to be mainly determined by the overnight price patterns in the North-American

markets, and the mode-based NP-FDA method picks up this structure best. Across the different prediction methods, the sign of the opening gap of the North-

Table 5: Market specific performance measure d^{sgn} for the sign predictions

$\boldsymbol{x}_s$	y_s	Mean	Median	Mode	PFDA	Lin. reg
HK-AU-JP	FR	<u>0.17</u> (0.06)	0.29 (0.08)	0.54 (0.08)	0.34 (0.08)	0.17 (0.06)
US-CA		0.31 (0.08)	0.34 (0.08)	0.26 (0.07)	0.31 (0.08)	0.31 (0.08)
HK-AU-JP-US-CA		0.29 (0.08)	0.26 (0.07)	0.34 (0.08)	0.40 (0.08)	0.37 (0.08)
HK-AU-JP	UK	0.23 (0.07)	0.23 (0.07)	0.49 (0.08)	0.29 (0.08)	0.20 (0.07)
US-CA		0.37 (0.08)	0.37 (0.08)	<u>0.06</u> (0.04)	0.40 (0.08)	0.34 (0.08)
HK-AU-JP-US-CA		0.29 (0.08)	0.17 (0.06)	0.31 (0.08)	0.43 (0.08)	0.31 (0.08)
HK-AU-JP	CH	0.23 (0.07)	0.31 (0.08)	0.57 (0.08)	0.34 (0.08)	<u>0.20</u> (0.07)
US-CA		0.34 (0.08)	0.37 (0.08)	<u>0.20</u> (0.07)	0.31 (0.08)	0.37 (0.08)
HK-AU-JP-US-CA		0.34 (0.08)	0.31 (0.08)	0.34 (0.08)	0.37 (0.08)	0.46 (0.08)
HK-AU-JP	DE	0.23 (0.07)	0.17 (0.06)	0.54 (0.08)	0.29 (0.08)	0.20 (0.07)
US-CA		0.23 (0.07)	0.34 (0.08)	<u>0.09</u> (0.05)	0.31 (0.08)	0.34 (0.08)
HK-AU-JP-US-CA		0.23 (0.07)	0.17 (0.06)	0.31 (0.08)	0.29 (0.08)	0.34 (0.08)
DE-FR-CH-UK	US	0.51 (0.08)	0.43 (0.08)	0.51 (0.08)	0.51 (0.08)	0.49 (0.08)
HK-AU-JP		0.37 (0.08)	0.46 (0.08)	<u>0.29</u> (0.08)	0.43 (0.08)	0.40 (0.08)
HK-AU-JP-DE-FR-UK-CH		<u>0.29</u> (0.08)	0.34 (0.08)	0.54 (0.08)	0.34 (0.08)	0.46 (0.08)
DE-FR-CH-UK	CA	0.34 (0.08)	0.34 (0.08)	0.31 (0.08)	0.37 (0.08)	0.29 (0.08)
HK-AU-JP		0.49 (0.08)	0.43 (0.08)	0.23 (0.07)	0.40 (0.08)	0.37 (0.08)
HK-AU-JP-DE-FR-UK-CH		0.29 (0.08)	0.40 (0.08)	<u>0.11</u> (0.05)	0.34 (0.08)	0.46 (0.08)
US-CA	JP	<u>0.20</u> (0.07)	0.26 (0.07)	0.29 (0.08)	0.20 (0.07)	0.23 (0.07)
DE-FR-CH-UK		0.51 (0.08)	0.51 (0.08)	0.54 (0.08)	0.37 (0.08)	0.37 (0.08)
DE-FR-CH-UK-US-CA		0.43 (0.08)	0.37 (0.08)	0.57 (0.08)	0.29 (0.08)	0.31 (0.08)
US-CA	AU	<u>0.26</u> (0.07)	0.37 (0.08)	0.34 (0.08)	0.29 (0.08)	0.31 (0.08)
DE-FR-CH-UK		0.34 (0.08)	0.31 (0.08)	0.43 (0.08)	0.43 (0.08)	0.37 (0.08)
DE-FR-CH-UK-US-CA		0.29 (0.08)	0.34 (0.08)	0.43 (0.08)	0.43 (0.08)	0.40 (0.08)
US-CA	HK	<u>0.29</u> (0.08)	0.31 (0.08)	0.31 (0.08)	0.29 (0.08)	0.26 (0.07)
DE-FR-CH-UK		0.29 (0.08)	0.40 (0.08)	0.46 (0.08)	0.34 (0.08)	0.26 (0.07)
DE-FR-CH-UK-US-CA		0.31 (0.08)	0.37 (0.08)	0.40 (0.08)	0.31 (0.08)	0.34 (0.08)

Standard deviations in parentheses.

Table 6: Coverage probabilities, predictive interval widths, and q

	Cov. prob.		Width		q	
	CPI	SCMI	CPI	SCMI	CPI	SCMI
			Overall aggregate			
Cluster(-1)	0.76	0.72	2.02	1.87	2.65	2.58
Cluster(-2)	0.72	0.71	1.89	1.79	2.62	2.52
Cluster(-1)-cluster(-2)	0.67	0.63	1.57	1.47	2.35	2.35
			Europe			
Cluster(-1)	0.80	0.73	1.96	1.75	2.45	2.39
Cluster(-2)	0.67	0.69	1.50	1.43	2.25	2.07
Cluster(-1)-cluster(-2)	0.67	0.65	1.31	1.27	1.96	1.95
			North-America			
Cluster(-1)	0.74	0.76	1.96	1.83	2.66	2.43
Cluster(-2)	0.83	0.80	2.03	1.92	2.45	2.40
Cluster(-1)-cluster(-2)	0.83	0.77	1.68	1.56	2.02	2.03
			Asia-Pacific			
Cluster(-1)	0.73	0.69	2.13	2.03	2.94	2.96
Cluster(-2)	0.73	0.68	2.33	2.18	3.21	3.23
Cluster(-1)-cluster(-2)	0.73	0.50	1.82	1.82	3.30	3.34

q denotes overall predictive interval quality measure. Nominal coverage is 0.90.

American indices appears to be determined by the Asian-Pacific as well as the European patterns, although the prediction method that performed best (mode-based NP-FDA) did so using the Asian-Pacific stock index patterns only. The Asian-Pacific aggregate results show that both the linear model and the mean-based NP-FDA both perform well, using patterns from the North-American indices only.

The results for the individual indices, shown in Table 5 roughly follow the structure already reflected by the aggregate results. An exception is FR, as it is the only European market for which the sign of the opening gap is determined by the Asian-Pacific patterns only. For the other European indices the sign is determined by the North-American index patterns. A possible explanation for the fact that the conditional mean based and median-based NP-FDA sign forecasts are outperformed by the mode-based NP-FDA forecast is skewness in the conditional distribution of the opening gap.

6.3 CPI and SCMI

Table 6 shows the results obtained for the CPI and the SCMI predictive intervals. For ease of presentation only the aggregate results are provided, which closely coincide with the individual results. It can be observed that all coverage probabilities are smaller than the nominal value of 90%. In all cases the coverage probabilities of CPI are better in the sense that they are closer to the nominal value. On the other hand, on average the SCMIs are shorter than the CPIs. This indicates that the CPI is more sensitive to the position in the state-space from which predictions are being made than the SCMI. The overall quality measure q corresponding with the ratio of the average length and the average coverage probabilities are very similar for both types of intervals.

7 Summary and conclusions

The aggregate results for the MSE show that the best FDA specification, mean-based NP-FDA with Cluster(-1)-Cluster(-2), on average performs much better than any of the linear models. This suggests that the NP-FDA method successfully exploits nonlinearities in the relation between $\boldsymbol{x}$ and y. This result is in line with the huge empirical evidence for nonlinear dependence in daily stock returns. In a recent systematic model-based prediction exercise, Guidolin, Hyde, McMillan (2009), found that stock and bond returns from the G7 countries, and in particular UK and US, appear to require nonlinear modelling.

The three clusters seem to be governed by different types of dynamics. For the European and Asian-Pacific markets the specification using information from Cluster(-1) gives the smallest MSE among the linear models. For the US this is also the overall best performing specification. In contrast, for the European markets the MSE obtained with the mean-based NP-FDA using the Cluster(-1)-Cluster(-2) specification was substantially better than any linear specification, suggesting the presence of nonlinear dynamics. The best specification for JP is based on Cluster(-1) only. Also for JP a substantial improvement in the MSE is obtained in going from the linear to the mean-based NP-FDA specification, suggesting the presence of a nonlinear relation between the North-American price patterns and JP. Further, the results have shown that in none of the cases P-FDA outperforms NP-FDA. This holds for the MSE as well as for the d^{sgn} measure. Among the NP-FDA methods considered, the best MSEs were obtained with the mean-based NP-FDA, while the mode-based NP-FDA performed best in terms of the sign forecast evaluation measuer d^{sgn} in many cases.

As far as we are aware, exploring the information in the intra-day stock price patterns in foreign markets to predict the opening of an index in a home-market, using NP-FDA, has not been a topic of earlier research. The present study recognizes the fact that traders in a home market may use any information revealed overnight in a foreign market due to fast transmission of information worldwide.

Although our present data set did not allow us to take into account trading in stocks of the home market after closing hours, we think this would be be an interesting future extension of our research. In addition, the nonlinear dependence between global markets suggested by our results could be investigated in more detail in a future study. For instance, one might test for causal relationships using nonparametric Granger causality tests before and after linear filtering of the data, as was done in an exchange rate setting by Bekiros and Diks (2008).

Acknowledgements

The authors would like to thank participants of the 2010 SNDE meeting in Novara, Italy, and the CEF2010 conference in London for their useful comments and suggestions.

References

[1] Barclay M. J., Hendershott T. (2003) Price Discovery and Trading After Hours, *Review of Financial Studies* 16, 1041-1073

[2] Barclay M. J., Hendershott T. (2004) Liquidity Externalities and Adverse Selection: Evidence from Trading After Hours, *Journal of Finance* 59, 681-710

[3] Biais B., Hillion P. Spatt Ch. S. (1999) Price Discovery and Learning during the Preopening Period in the Paris Bourse, Tepper School of Business, Working Paper 480.

[4] Biais B., Hillion P., Spatt Ch. S. (1997) Price Discovery and Learning during the Preopening Period in the Paris Bourse, Manuscript. Toulouse: Toulouse Univeristy

[5] Bekiros, S.D. and Diks C.G.H. (2008) The nonlinear dynamic relationship of exchange rates: Parametric and nonparametric causality testing, Nonparametric conditional predictive regions for time series, *Journal of Macroeconomics* 30, 1641-1650

[6] Berkman H., Koch P.D., Tuttle L., Zhang Y. (2012) Paying Attention: Overnight Returns and the Hidden Cost of Buying at the Open, *Journal of Financial and Quantitative Analysis 47* 715-741

[7] Benko M. (2006) *Functional Data Analysis with Applications in Finance*, PhD Humboldt-Universität, Berlin

[8] Branch B.S., Ma A. (2012), Overnight Return, the Invisible Hand behind Intraday Returns? *Journal of Applied Finance*, 22

[9] Cliff M., Cooper M.J., Gulen H. (2008) Return differences between trading and non-trading hours: Like night and day, Working paper *http://ssrn.com/abstract=1004081*

[10] De Gooijer J.G., Gannoun A. (2000) Nonparametric conditional predictive regions for time series, *Computational Statistics & Data Analysis* 33, 359–275

[11] De Gooijer J.G., Sivarajasingham S. (2007) Parametric and nonparametric Granger causality testing: Linkages between international stock markets, *Physica* A387, 2547-2560

[12] Ferraty F., Vieu P. (2005) Conditional quantiles for dependent functional data with application to the climatic *El Niño* phenomenon, *Sankhyā: The Indian Journal of Statistics* 67, 378–398

[13] Ferraty F., Vieu P. (2006) *Nonparametric Functional Data Analysis*, Springer-Verlag, New York

[14] French K.R., Roll R. (1986) Stock return variances: The arrival of information and the reaction of traders, *Journal of Financial Economics*, 17, 5-26

[15] Gerety M. S., Mulherin J. H. (1994) Price formation on stock exchanges: the evolution of trading within the day, *Review of Financial Studies* 7, 609-629

[16] George T.J., Hwang C.Y. (2001) Information flow and pricing errors: A unified approach to estimation and testing, *Review of Financial Studies* 14, 979-1020

[17] Guidolin, M., Hyde, S., McMillan, D. and Ono, S. (2009), Non-linear predictability in stock and bond returns: When and where is it exploitable?, *International Journal of Forecasting* 25, 373-399

[18] Hasbrouck J. (1991) Measuring the information content of stock trades, *Journal of Finance* 46, 179-207

[19] Hasbrouck J. (1993) Assessing the quality of a security market: A new approach to transaction-cost measurement, *Review of Financial Studies* 6, 191-212

[20] Hong H., Wang J. (2000) Trading and return under periodic market closures, *Journal of Finance* 55, 297-354

[21] Lockwood L.J., Lin S. (1965) An examination of stock market return volatility during overnight and intraday periods, 1964-1989, *Journal of Finance* 45, 591-601

[22] Malfait N., Ramsay J. O. (2003) The historical functional linear model, *Canadian Journal of Statistics* 31, 115-128

[23] Madhavan A., Richardson M., Roomans M. (1997) Why do security prices change? A transaction level analysis of NYSE stocks, *Review of Financial Studies* 10, 1035-1064

[24] Mueller H.G., Sen R., Stadtmueller U. (2011) Functional data analysis for volatility, *working paper*

[25] Mueller H.G., Stadtmueller U., Yao F. (2006) Functional variance processes, *Journal of American Statistical Association* 101, 1007-1018

[26] Oldfield G.S., Rogalski R.J. (1980) A theory of common stock returns over trading and non-trading periods, *Journal of Finance* 35, 729-751

[27] Ramsay J.O., Silverman B.W. (2005) *Functional Data Analysis* (2nd ed.) Springer Verlag, New York

[28] Ramsay J. O., Ramsey J. B. (2001) Functional data analysis of the dynamics of the monthly index of non durable goods production, *Journal of Econometrics* 107, 327-344

[29] Zhong Z. (2007) *The predictability of overnight information*, M.Sc. thesis, Singapore Management University.