

Functional dynamic factor model for intraday price curves

Piotr Kokoszka*, Hong Miao†, Xi Zhang‡

Abstract

This paper proposes a functional dynamic factor model for the evaluation of the impact of scalar- and curve-valued factors on the shapes of intraday price curves. The asymptotic theory leads to practically useful confidence intervals for the factor coefficients. The main findings pertain to the impact of the shapes of intraday oil futures on the shapes of intraday prices of blue chip stocks.

JEL Classification: C32, G14, C53, C12,

Key words: Functional factor model; Intraday price curves; Oil futures.

1 Introduction

In its simplest form, the Capital Asset Pricing Model (CAPM) is defined by the straight line regression

$$r_n = \alpha + \beta r_{m,n} + \varepsilon_n, \quad (1.1)$$

where

$$r_n = 100(\ln P_n - \ln P_{n-1}) \approx 100 \frac{P_n - P_{n-1}}{P_{n-1}} \quad (1.2)$$

is the excess return, in percent, over a unit of time, and $r_{m,n}$ is the analogously defined return on a relevant market index. The unit of time can be day, month or year, with the classical theory initiated by Markowitz (1959), Sharpe (1964), Lintner (1965) and Black (1972) pertaining to a static model for annual returns. The one-factor model (1.1) has been extended to multifactor models, but there has been little research on the extension of

*Department of Statistics, Colorado State University, Fort Collins, CO 80523

†Department of Finance and Real Estate, Colorado State University, Fort Collins, CO 80523

‡PNC Bank, Pittsburgh, PA 15222

the fundamental ideas of the theory of finance contained in the CAPM and factor models to the intraday price curves.

Traditional financial models, such as CAPM, deal with point-to-point returns (using the closing prices at day/month/year t and $t - 1$). This approach completely depends on the points at which the prices are taken and totally ignores the price curve, that is how price moves from $t - 1$ to t . Thus, if the price at t is equal to the price at $t - 1$, the return is just zero. However, the path from $t - 1$ to t certainly contain some information about the evolution of the price. In statistical terminology, this means that if price data are treated as discrete time series, important information about the smooth functional behavior of the price generating process may be ignored. *Functional Data Analysis* (FDA) methods can often extract additional information contained in the function that is not normally available from application of traditional statistical methods. In contrast to traditional high-frequency data analysis, FDA methods focus on low frequency components and extract information which may be ignored by the usual time series based high frequency methods. Because the FDA approach treats the whole curve as a single entity, there is also no concern about correlations between repeated measurements. This represents a change in philosophy towards the handling of time series and correlated data.

We study daily curves of intraday cumulative returns (CIDR's) defined as follows.

DEFINITION 1.1 Suppose $P_n(t_j), n = 1, \dots, N, j = 1, \dots, m$ is the price of a financial asset at time t_j on day n . The functions

$$R_n(t_j) = 100[\ln P_n(t_j) - \ln P_n(t_1)], \quad j = 1, 2, \dots, m, \quad n = 1, \dots, N,$$

are called the *cumulative intraday returns* (CIDR's).

The CIDR's were introduced by Gabrys *et al.* (2010) as a way of transforming the price curves to stationarity in the daily index n , Horváth *et al.* (2014) provide a formal test of their stationarity. Since $r_n(t_j) \approx 100(P_n(t_j) - P_n(t_1))/P(t_1)$, with $P_n(t_1)$ being a constant for a given day n , the price and the CIDR curves look very similar. We work with one minute averages, so $t_{j+1} - t_j = 1$ min, and $P(t_j)$ is the average of the maximum and minimum price within the j th minute. Our objective is to propose a regression type model in which the impact of curve-valued factors on the CIDR's on individual assets can be statistically evaluated. The central question we seek to answer is whether additional factors beyond CIDR's on a market index are statistically significant. The CIDR's, which are functions (or curves), have not been studied in economic literature, and no financial theory in the vein of Merton (1973) or Ross (1976) is available. Instead, we develop a statistical methodology in the framework of Functional Data Analysis (FDA), with some

hope that our research will stimulate relevant advances in financial theory. We treat the CIDR's as continuous curves, one curve per day. For every fixed day, these curves exhibit a specific pattern, typically with some upward or downward trend, which may reverse during a trading day. It is therefore natural to study the statistical behavior of the CIDR's within the framework of FDA which treats the curves as complete statistical objects, rather than as collections of a large number of individual observations. In the following, we describe only the tools of FDA we need; Horváth and Kokoszka (2012) and Hörmann and Kokoszka (2012) contain the relevant broader background on functional time series. Ramsay and Silverman (2005) provide a comprehensive introduction to the ideas of FDA. All applications of FDA assume that the data can be converted to continuous curves. The methodology presented in this paper assumes that the intraday price data are available at a fine temporal resolution throughout the trading day. This is the case for the data which motivate this research. Conversion of such data to smooth curves is not the focus of this paper. Technical details are explained in Ramsay *et al.* (2009).

An extension of the CAPM (1.1) to CIDR's can be defined by the functional regression

$$R_n(t) = \beta_0(t) + \beta_1 M_n(t) + \varepsilon_n(t). \tag{1.3}$$

We consider a more general factor model defined in (2.1). The statistical framework we develop allows us to test if additional factors beyond market CIDR's, $M_n(\cdot)$, are significant. The factors can be scalars or functions. For example, would the predictions be better if (1.3) were replaced by

$$R_n(t) = \beta_0(t) + \beta_1 M_n(t) + \beta_2 S_n + \beta_3 H_n + \varepsilon_n(t), \tag{1.4}$$

where S_n and H_n are the Fama–French factors (defined in Section 4)? For a stock of an energy company or an airline, it is natural to consider the model

$$R_n(t) = \beta_0(t) + \beta_1 M_n(t) + \beta_2 C_n(t) + \varepsilon_n(t), \tag{1.5}$$

where $C_n(t)$ are the CIDR's on oil futures. Is the coefficient β_2 significant for such companies? Is it significant for companies in the IT sector? The intuition we have regarding such questions is based on daily or lower frequency returns, see e.g. Jones and Kaul (1996), Park and Ratti (2008) and Narayan and Sharma (2011). Regression (1.5) quantifies the dependence between “standardized” price curves, essentially the dependence between their shapes. To obtain such insights, we develop an asymptotic theory for the estimation of the parameters in models like (1.4) and (1.5) which allows us to develop tests for the significance of specific regression coefficients, like β_2 in (1.5). Models with

several regression functions have not been studied in FDA literature; the focus has been on models with a single regressor, like (1.3), but with functional regression coefficients, that is in place of $\beta_1 M_n(t)$, $\int \beta_1(t, s) M_n(s) ds$ has been used see e.g. Chapter 8 of Horváth and Kokoszka (2012). Such models are difficult to estimate even if the functions M_n are independent. For financial data, the temporal dependence would require the estimation of a high dimensional covariance matrix, which is a difficult. The simpler model with scalar coefficients leads to accurate asymptotic confidence intervals for scalar coefficients which are readily interpretable. The sign of a coefficient like β_2 in (1.5) also has an obvious interpretation.

A different class of functional factor models was recently proposed by Hays *et al.* (2012). These models, introduced to forecast yield curves, say X_n , are of the form

$$X_n(t) = \sum_{k=1}^K \gamma_{nk} F_k(t) + \varepsilon_n(t). \quad (1.6)$$

In contrast to (1.4) or (1.5), the factors F_k do not depend on n and are *orthonormal* functions to be estimated. The dynamics are in the coefficients γ_{nk} which are assumed to follow Gaussian autoregressive processes (the ε_n are also Gaussian). Model (1.6) could be termed a statistical factor model. It is designed for temporal forecasting, while our model is designed for regression type prediction in which the correlation structure of factors plays a major role. Earlier contributions to curve forecasting include Diebold and Li (2001), Kargin and Onatski (2008), Bowsher and Meeks (2008), Shen (2009), Koopman *et al.* (2010). A functional approach was also used to model intraday volatility, Müller *et al.* (2011), and to test temporal predictability, Kokoszka and Reimherr (2013).

We note that it is, in principle, possible to replace the scalars β_1 and β_2 in (1.5) by integral operators, e.g. $\beta_1 M(t)$ could be replaced by $\int \beta_1(t, s) M(s) ds$. For model (1.3), such an extension leads to the extensively studied fully functional linear model, see e.g. Chapter 8 of Horváth and Kokoszka (2012). In the case of model (1.3), estimation of the kernel $\beta_1(t, s)$ has been extensively studied, but the results of Kokoszka and Zhang (2012) show that, *for intraday price data*, the more complex model does not lead to better predictions than a simple model which uses only a scalar coefficient.

The paper proceeds as follows. In Section 2, we propose an estimation procedure for our model and show that the estimators are consistent and asymptotically normal under general conditions which admit temporal dependence. Section 3 outlines how the theory presented in Section 2 can be used to evaluate the significance of the factors. This methodology is applied to U.S. Stocks in Section 4. The contribution of the paper and main conclusions are summarized in Section 5, while the proofs of the results of Section 2

are presented in Section 6.

2 The functional factor model and its estimation

It is convenient to write the model considered in this paper as

$$R_n(t) = \beta_0(t) + \sum_{j=1}^p \beta_j F_{nj}(t) + \varepsilon_n(t), \quad (2.1)$$

where the ε_n are mean zero error functions. All factors are nominally functional, but scalar factors are allowed, by treating scalars as constant functions. In our asymptotic setting, the interval on which all functions are defined is normalized to be the unit interval $[0, 1]$, so in various integrals appearing below, constant functions behave exactly as the constants they represent. The parameters of the model are the mean function $\beta_0(\cdot)$ and the vector of the coefficients:

$$\beta = [\beta_1, \dots, \beta_p]^T.$$

Before describing the estimation in model (2.1), we introduce briefly the requisite functional setting. We provide only the minimal required background; a reader interested in a broader perspective is referred to Bosq (2000) and Horváth and Kokoszka (2012). All functions are assumed to be random elements of the Hilbert space L^2 of square integrable functions on $[0, 1]$ with the inner product $\langle f, g \rangle = \int f(t)g(t)dt$. When no limits of integration are indicated, the integral is over the whole interval $[0, 1]$. Every element of L^2 is square integrable in a sense that $\|f\|^2 = \int |f(t)|^2 dt < \infty$. However, when X is a *random* function taking values in L^2 , its square integrability means that $E \|X\|^2 < \infty$. With this background, we can derive the estimators. Set

$$\begin{aligned} \bar{R}(t) &= \frac{1}{N} \sum_{n=1}^N R_n(t), & R_n^c(t) &= R_n(t) - \bar{R}(t); \\ \bar{F}_j(t) &= \frac{1}{N} \sum_{n=1}^N F_{nj}(t), & F_{nj}^c(t) &= F_{nj}(t) - \bar{F}_j(t); \\ \bar{\varepsilon}(t) &= \frac{1}{N} \sum_{n=1}^N \varepsilon_n(t), & \varepsilon_n^c(t) &= \varepsilon_n(t) - \bar{\varepsilon}(t). \end{aligned}$$

Computing sample means of both sides of (2.1), we get

$$\beta_0(t) = \bar{R}(t) - \sum_{j=1}^p \beta_j \bar{F}_j(t) - \bar{\varepsilon}(t).$$

Inserting the above into (2.1), we obtain

$$R_n^c(t) = \sum_{j=1}^p \beta_j F_{nj}^c(t) + \varepsilon_n^c(t). \quad (2.2)$$

Since

$$\beta_0(t) = \mu_R(t) - \sum_{j=1}^p \beta_j \mu_j(t),$$

where

$$\mu_R(t) = ER_n(t), \quad \mu_j(t) = EF_{nj}(t),$$

the mean function is estimated by

$$\hat{\beta}_0(t) = \bar{R}(t) - \sum_{j=1}^p \hat{\beta}_j \bar{F}_j(t), \quad (2.3)$$

with appropriate estimators $\hat{\beta}_j$. These can be derived using the method of moments as follows. In light of (2.2), set

$$\begin{aligned} S(\beta) &= E \left\| R_n^c - \sum_{j=1}^p \beta_j F_{nj}^c \right\|^2 \\ &= E \|R_n^c\|^2 + \sum_{j=1}^p \sum_{k=1}^p \beta_j \beta_k E \langle F_{nj}^c, F_{nk}^c \rangle - 2 \sum_{j=1}^p \beta_j E \langle R_n^c, F_{nj}^c \rangle. \end{aligned}$$

Note that

$$\frac{\partial S(\beta)}{\partial \beta_j} = 2 \left\{ \sum_{k=1}^p \beta_k E \langle F_{nj}^c, F_{nk}^c \rangle - E \langle R_n^c, F_{nj}^c \rangle \right\}.$$

The vector $\beta^{(m)}$ minimizing $S(\beta)$ thus satisfies

$$\mathbf{R}_c = \mathbf{F}_c \beta^{(m)}, \quad \beta^{(m)} = \mathbf{F}_c^{-1} \mathbf{R}_c,$$

where

$$\begin{aligned} \mathbf{F}_c &= [E \langle F_{nj}^c, F_{nk}^c \rangle, j, k = 1, 2, \dots, p] \quad (p \times p), \\ \mathbf{R}_c &= [E \langle R_n^c, F_{nj}^c \rangle, j = 1, 2, \dots, p]^T \quad (p \times 1). \end{aligned}$$

An β thus is

$$\hat{\beta} = \hat{\mathbf{F}}^{-1} \hat{\mathbf{R}}, \quad (2.4)$$

where

$$\hat{\mathbf{F}} = \left[N^{-1} \sum_{n=1}^N \langle F_{nj}^c, F_{nk}^c \rangle, j, k = 1, 2, \dots, p \right] \quad (p \times p), \quad (2.5)$$

$$\hat{\mathbf{R}} = \left[N^{-1} \sum_{n=1}^N \langle R_n^c, F_{nj}^c \rangle, j = 1, 2, \dots, p \right]^T \quad (p \times 1). \quad (2.6)$$

Estimator (2.4) is essentially a least square estimator, except that we must correct for the unknown mean function, which is estimated by the method of moments.

The remainder of this section is devoted to the consistency and the asymptotic normality of estimators (2.4) and (2.3). Our first assumption ensures the existence of moments required to justify the derivation above.

ASSUMPTION 2.1 *The factors F_{nj} and the errors ε_n are random elements of the space L^2 which satisfy $E \|F_{nj}\|^4 < \infty$, $E \|\varepsilon_n\|^4 < \infty$, and $E\varepsilon_n = 0$. Moreover, for each n , the error functions ε_n are independent of the factors $F_{nj}, 1 \leq j \leq p$.*

Classical multifactor models going back to the 1970's are essentially one period models, see e.g. Chapter 6 of Campbell *et al.* (1997), with the period being typically one year. In applications for which model (2.1) was designed, for each j , F_{nj} is a sequence of functions, one function per day. These are thus dynamic factors, and their temporal dependence cannot be ignored. Very little is known about the structure of this dependence, so an inclusive and general form of dependence must be postulated. The above discussion motivates our next assumption.

ASSUMPTION 2.2 *Suppose sequences $\{\delta_n\}$ and $\{\eta_n\}$ consist of iid random variables taking values in measurable spaces \mathcal{S}_δ and \mathcal{S}_η , respectively. Assume there are measurable functions f_j and e defined on the appropriate product spaces, such that*

$$F_{nj} = f_j(\delta_n, \delta_{n-1}, \dots), \quad \varepsilon_n = e(\eta_n, \eta_{n-1}, \dots).$$

Assumption 2.2 implies that the factors and the errors are stationary and ergodic random sequences taking values in L^2 . Ergodicity follows from the general results which states that Bernoulli shifts, i.e. nonlinear moving averages of the form defined above, are ergodic, see e.g. Theorem 36.4 of Billingsley (1995). Ergodicity implies that sample averages of integrable functionals of the sequences F_{nj} and ε_j converge a.s. to their expectations. This is needed to establish the consistency of the estimators. Assumptions similar to 2.2 have been used extensively in recent theoretical work, as all stationary time series models in practical use can be represented as Bernoulli shifts, see Wu (2005), Shao and Wu (2007), Aue *et al.* (2009), Hörmann and Kokoszka (2010), among many other contributions. It must be emphasized that they have been used in econometric research even earlier, and the work of Pötscher and Prucha (1997) contributed to their popularity. The representation

$\varepsilon_n = e(\eta_n, \eta_{n-1}, \dots)$ admits heteroskedastic models for the errors, like those studied in Hörmann *et al.* (2013) or used in Kokoszka and Reimherr (2013).

To formulate our consistency result, we introduce the matrix

$$\mathbf{F} = [E \langle F_{nj} - \mu_j, F_{nk} - \mu_k \rangle, j, k = 1, 2, \dots, p] \quad (p \times p). \quad (2.7)$$

THEOREM 2.1 *Suppose Assumptions 2.1 and 2.2 hold and the matrix \mathbf{F} defined by (2.7) is nonsingular. Then*

$$\hat{\beta} \xrightarrow{a.s.} \beta \quad \text{and} \quad \int \left(\hat{\beta}_0(t) - \beta_0(t) \right)^2 dt \xrightarrow{a.s.} 0.$$

Theorem 2.1 is proven in Section 6.

To establish the asymptotic normality, we must impose stronger moment and weak dependence conditions. The conditions we have chosen are justified by their generality and the fact that they have been shown to hold for all known models for temporally dependent functions, see Hörmann and Kokoszka (2010, 2012), or Chapter 16 of Horváth and Kokoszka (2012). Suppose H is a separable Hilbert space. Let $p \geq 1$ and let L_H^p be the space of H -valued random elements X such that

$$\nu_p(X) = (E \|X\|^p)^{1/p} < \infty.$$

DEFINITION 2.1 A sequence $\{X_n\} \in L_H^p$ is called L^p -*m*-approximable if each X_n admits the representation

$$X_n = f(u_n, u_{n-1}, \dots), \quad (2.8)$$

where the u_i are iid elements taking values in a measurable space S , and f is a measurable function $f : S^\infty \rightarrow H$. Moreover we assume that if $\{u'_i\}$ is an independent copy of $\{u_i\}$ defined on the same probability space, then letting

$$X_n^{(m)} = f(u_n, u_{n-1}, \dots, u_{n-m+1}, u'_{n-m}, u'_{n-m-1}, \dots) \quad (2.9)$$

we have

$$\sum_{m=1}^{\infty} \nu_p(X_n - X_n^{(m)}) < \infty. \quad (2.10)$$

The gist of Definition 2.1 is that the dependence of f in (2.8) on the innovations far in the past decays so fast that these innovations can be replaced by their independent copies. Such a replacement is asymptotically negligible in the sense quantified by (2.10).

THEOREM 2.2 Suppose Assumptions 2.1 and 2.2 hold, the sequences $\{\varepsilon_n\}$ and $\{F_{nj}\}$ are L^4 - m -approximable, and the matrix \mathbf{F} defined in (2.7) is nonsingular. Then,

$$\sqrt{N} \left(\hat{\beta}_0(\cdot) - \beta_0(\cdot), \hat{\beta} - \beta \right) \xrightarrow{d} (Z(\cdot) + \mathbf{f}^T(\cdot) \mathbf{F}^{-1} \mathbf{W}, \mathbf{F}^{-1} \mathbf{W}), \quad (2.11)$$

where the random elements Z and \mathbf{W} are jointly normal. The covariance matrix of \mathbf{W} is given by

$$\mathbf{\Gamma} = \sum_{h=-\infty}^{\infty} E[\boldsymbol{\xi}_0 \boldsymbol{\xi}_h^T] \quad (2.12)$$

with

$$\boldsymbol{\xi}_n = [\langle \varepsilon_n, F_{n1} - \mu_1 \rangle, \dots, \langle \varepsilon_n, F_{np} - \mu_p \rangle]^T. \quad (2.13)$$

The covariance function of Z is specified in Theorem 6.2. The R^p -valued function $\mathbf{f}(\cdot)$ is defined by

$$\mathbf{f}(t) = [EF_{n1}(t), EF_{n2}(t), \dots, EF_{np}(t)]^T. \quad (2.14)$$

Theorem 2.2 is proven in Section 6, which also contains the details of the asymptotic distribution of the sample mean function $\hat{\beta}_0(\cdot)$. For the purpose of testing the significance of the components of β it is enough to know the asymptotic covariance matrix of $\hat{\beta}$, which is given by (2.12). Finite sample implementation is discussed in Section 3.

3 Prediction and testing

A simple way of evaluating if additional regressors contribute to the explanatory power of a regression model is to see if they reduce the mean squared error of prediction. The functional regression model (2.1) leads to the following equation for predicting an asset's CIDR curve from the factors

$$\hat{R}_n(t) = \hat{\beta}_0(t) + \sum_{j=1}^p \hat{\beta}_j F_{nj}(t).$$

The quality of prediction can be evaluated by the integrated mean squared error defined as

$$\text{MSEP} = N^{-1} \sum_{n=1}^N \int (\hat{R}_n(t) - R_n(t))^2 dt. \quad (3.1)$$

Different measures could be used, but the estimators in Section 2 were defined with the goal of minimizing the MSEP; Gneiting (2011) emphasizes that using different measures to derive and evaluate predictions can lead to spurious findings.

In Section 4, we study models similar to (1.4) and (1.5). All of them have the CIDR on a market index, the curves M_n , as the first factor. The question we want to answer is whether adding additional factors makes the MSEP's smaller. We calculated relative predictive efficiency gains (in percent) defined as

$$E = 100 \left(\frac{\text{MSEP}_M}{\text{MSEP}_F} - 1 \right), \quad (3.2)$$

where MSEP_M is the MSEP computing using only M_n , and MSEP_F is the MSEP computed using all factors in the model. Theoretically, the values of E must be positive, but if the additional factors do not contribute to predictive power, they may be close to zero and even slightly negative due to rounding errors. Large positive values of E would indicate that additional factors reduce the MSEP.

While the values of E give an initial idea about the importance of additional factors, and can be computed without resorting to any asymptotic theory, they do not allow us to attach statistical significance to the conclusions. This can be done using Theorem 2.2. To test the null hypothesis $\beta_j = 0$, we construct a 95% confidence interval for β_j . If this interval contains zero, we cannot reject the null hypothesis. According to Theorem 2.2, $\hat{\beta}$ is asymptotically distributed as a normal vector with the mean β and the covariance matrix $N^{-1}\mathbf{F}^{-1}\mathbf{T}\mathbf{F}^{-1}$. The matrix \mathbf{F} is estimated by $\hat{\mathbf{F}}$ given by (2.5). The estimation of \mathbf{T} given by (2.12) is more complex. The random vectors ξ_n in (2.13) are not observable. They must be replaced by $\hat{\xi}_n$ defined analogously, but with μ_j replaced by \bar{F}_j and ε_n by $\hat{\varepsilon}_n$ defined by

$$\hat{\varepsilon}_n(t) = R_n(t) - \hat{\beta}_0(t) - \sum_{j=1}^p \hat{\beta}_j F_{nj}(t).$$

The matrix \mathbf{T} is estimated as the long run covariance matrix of the sequence $\hat{\xi}_n$. We used an R function `lrvar` (package “sandwich”) with default kernel and bandwidth values. Denote this estimate by $\hat{\mathbf{T}}$. The variance of $\hat{\beta}_j$ is the j th diagonal element of $N^{-1}\hat{\mathbf{F}}^{-1}\hat{\mathbf{T}}\hat{\mathbf{F}}^{-1}$. The confidence interval for β_j is constructed using standard normal quantiles.

4 Application to U.S. stocks

We now apply the methodology developed in the previous sections to a selection of U.S. stocks. Our objective in this section is not to present a comprehensive analysis of all reasonable functional dynamic factor models on a very broad collection of stocks. We merely aim at obtaining useful insights, both regarding the finite sample performance of the tests and the conclusions they lead to for a representative portfolio of stocks. Table 4.1

TABLE 4.1 Sectors and stocks used in this study.

Sector	Symbol	Full Name
Energy	XOM	Exxon Mobil Corporation
	CVX	Chevron Corporation
	COP	ConocoPhillips
Information Technology	MSFT	Microsoft Corporation
	IBM	IBM Corporation
	ORCL	Oracle Corporation
Financials	CITI	Citi Bank
	BAC	Bank of America Corporation
	JPM	JPMorgan Chase Co.
Consumer Staples	KO	Coca-Cola
	WMT	Wal-Mart Stores
	PG	Procter & Gamble Co.
Consumer Discretionary	MCD	McDonald's Corporation
	DIS	The Walt Disney Corporation
	CMCSA	Comcast Corporation
Transportation	FDX	FedEx Corporation
	JBLU	JetBlue Airways Corporation
	UPS	United Parcel Service, Inc.

lists the stocks we used. We selected two periods of time, from 09/03/2002 to 08/26/2004 and from 04/07/2005 to 04/02/2007. Each period contains 500 trading days. We removed a few outliers due to splits and data errors.

In addition to models (1.4) and (1.5), we consider the model

$$R_n(t) = \beta_0(t) + \beta_1 M_n(t) + \beta_2 L_{n-1} + \varepsilon_n(t), \tag{4.1}$$

where L_{n-1} is a scalar factor equal to the previous day's return on the asset or the index. To facilitate the presentation, we introduce the following notation:

- PA model (4.1) with L_{n-1} representing the asset return;
- PI model (4.1) with L_{n-1} representing the index return;
- FF the Fama–French model (1.4);
- OF model (1.5) with oil futures as the extra factor.

Market CIDR's, $M_n(t)$, are represented by the Standard & Poor's 100 index. In the Fama–French model, see Fama and French (1995) or Section 17.4 of Ruppert (2011), S_n

TABLE 4.2 Asymptotic 95% confidence intervals for the regression coefficients of the extra factors of the energy sector stocks. The extension 1 or 2 following the ticker symbol refers to the first and second 500 day long period. The arguments (1) and (2) in the FF model indicate the two Fama–French factors. **Confidence intervals that do not contain 0 are in bold.**

	PA	PI	FF(1)	FF(2)	OF
XOM1	(-.0429,.0440)	(-.0577,.0580)	(-.0019, .0019)	(-.0065, .0066)	(.0041, .0778)
XOM2	(-.0396, .0397)	(-.1101, .1057)	(-.0020, .0020)	(-.0045, .0044)	(.3302, .4129)
CVX1	(-.1172, .1187)	(-.0565, .0600)	(-.0020, .0020)	(-.0068, .0071)	(.0393, .1130)
CVX2	(-.0443,.0441)	(-.1243, .1211)	(-.0022, .0021)	(-.0054, .0053)	(.3551, .4468)
COP1	(-.1021, .1045)	(-.0654, .0675)	(-.0021, .0022)	(-.0083, .0078)	(.0389, .1417)
COP2	(-.0644, .0632)	(-.1722, .1660)	(-.0025, .0025)	(-.0071, .0071)	(.4875, .6083)

stands for the “small minus large” factor and H_n for the “high minus low” factor. The factor S_n is the the difference in returns on a portfolio of small stocks and a portfolio of large stocks; H_n is is the difference in returns on a portfolio of high book-to-market value (BE/ME) stocks and a portfolio of low BE/ME stocks. The oil futures functional factor $C_n(\cdot)$ in model OF is constructed using tick by tick futures prices for light sweet crude obtained from TickData. The Chicago Mercantile Exchange provides both open outcry (pit) and electronic (Globex) trading in oil futures, and we used both series to construct daily time series in one minute resolution.

As explained in Section 3, the initial step of our analysis was to compute the relative predictive efficiency gains E defined by (3.2). It turns out that for models PA, PI and FF, they are practically nonexistent; the values are of the order $10E-4$. This strongly suggests that the additional scalar factors in these models have no additional explanatory power. The situation is different for the OF model. CIDR’s on oil futures lead to values of E approaching 40 for oil companies and to single digit gains for many other companies. This leads us to conjecture that the intraday oil futures will be a significant factor, but to verify it a rigorous statistical analysis based on the confidence intervals is needed. The signs of the estimates of β_2 will also allow us to see the direction of the impact, if it is significant. The coefficients of the market factor are always positive and significant.

Confidence intervals for the stocks in the energy sector are displayed in Table 4.2. As already indicated by the analysis of the MSEF’s, the additional scalar factors in models PA, PI and FF are not significant. The CIDR’s on oil futures in model OF are significant. for all sectors and stocks listed in Table 4.1. Based on tables for other sectors, not shown to conserve space, we conclude that the additional factors in models PA, PI and FF are never

TABLE 4.3 Summary of conclusions for the OF model for the stocks listed in Table 4.1.
Key: + ($\beta_2 > 0$), - ($\beta_2 < 0$), -/0 (mostly -, some 0), 0/- (mostly 0, some -).

Sector	value of β_2
Energy	+
Information Technology	-
Financial	-/0
Consumer Staples	-/0
Consumer Discretionary	0/-
Transportation	-

statistically significant. The conclusions for model OF are summarized in Table 4.3. They indicate that the shapes of intraday oil futures are "positively correlated" with the shapes of intraday price curves for large U.S. Oil companies, and are mostly negatively correlated for the remaining companies. In the consumer discretionary sector, only the Walt Disney Corporation is negatively correlated with oil futures. A somewhat surprising finding is the negative correlation (in the sense studied in this paper) for the IT sector. Unlike the other sectors, the evidence is uniform. An ex post intuition behind this finding is not clear, but the result itself agrees with the empirical analysis of Narayan and Sharma (2011) who established a similar dependence for daily returns, and who offer some discussion.

To validate our conclusions, we performed additional checks. The conclusions reported in Table 4.3 do not change if the maximal lag used in the estimation of the long run covariance matrix is increased by 50 percent. While the end points of the confidence intervals obviously change, the inclusion of the value $\beta_2 = 0$ was not affected. We also performed a simulation study to find the empirical coverage probability. To make the discussion more concrete, consider the Microsoft stock over the first time period. We obtained $\hat{\beta}_2 = -.053676$, with the asymptotic confidence interval $(-.0930, -.0150)$. We calculated the residual functions

$$\hat{\varepsilon}_n(t) = R_n(t) - \hat{\beta}_0(t) - \hat{\beta}_1 M_n(t) - \hat{\beta}_2 C_n(t).$$

We thus have 500 curves M_n , 500 curves C_n and 500 curves $\hat{\varepsilon}_n$. For each set of curves, we draw 500 curves with replacement (bootstrap), which produces 500 curves M_n^* , 500 curves C_n^* , and 500 curves $\hat{\varepsilon}_n^*$. To assess the empirical size of the test, we generate bootstrap Microsoft CIDR's as $R_n^*(t) = \hat{\beta}_0(t) + \hat{\beta}_1 M_n^*(t) + \hat{\varepsilon}_n^*(t)$. To assess the empirical power, we use the DGP $R_n^*(t) = \hat{\beta}_0(t) + \hat{\beta}_1 M_n^*(t) + \hat{\beta}_2 C_n^*(t) + \hat{\varepsilon}_n^*(t)$ (with $\hat{\beta}_2 = -.053676$). We can replicate these procedures a large number of times and obtain MC empirical size and

power. The asymptotic test has empirical size of about 7 percent and the power of about 74 percent (for $\beta_2 = -.053676$). We obtained very similar results for other stocks.

5 Summary and conclusions

We have proposed a regression framework that allows us to evaluate quantitatively how the shapes of intraday price curves depend on the shapes of other curve-valued factors or on scalar factors. When applied to blue chip stocks, our methodology has shown that these shapes do not depend on scalar factors we considered (one number per day). In the hindsight, this is perhaps not surprising, as such factors can be expected to affect closing daily prices, not the particular intraday shape through which these closing prices are arrived at. (On the other hand, a priori, a positive return on the previous day, might to some extent determine the shape on the price curve.) In contrast, our methodology has revealed a significant impact of the intraday oil futures on several sectors, most notably a very strong positive impact on the shapes of intraday prices of U.S. oil companies. While this finding is not entirely surprising, it shows that even the intraday price evolution of the equities of these companies is much more strongly impacted by the intraday evolution of oil futures prices than the general market portfolio. For most other stocks, the impact is negative, if it is significant. Roughly speaking, a negative significant impact means that when oil futures prices increase during a trading day, the share prices tend to fall faster than the prices of a market portfolio. Our goal was to introduce a new quantitative framework and apply it to a relatively small selection of assets and functional factor models, to show its feasibility and potential. We hope that our ideas will be received with some interest, and our method will be applied to different asset classes and different functional factors.

We have advanced the theory of functional regression models in the following directions. Model (2.1) contains multiple regressor functions and a combination of a functional intercept and scalar slope coefficients. Such models have not been considered so far; only models with a single regression function have been considered. A main difficulty was to develop an asymptotic and practically useful framework in which parameters of very different structure, a function and a finite dimensional vector, could be accommodated. To do it, we derived a central limit theorem for dependent sequences in an abstract Hilbert space, Theorem 6.3. By stating our assumptions in terms of Bernoulli shifts, we were able to derive almost sure consistency of the estimators, Theorem 2.1, rather than the mean square consistency established prior to this work for models with a single regressor and functional coefficients. In both the asymptotic normality and consistency analysis, an im-

portant technical point was to carefully account for the effect of centering of all functions prior to computing the LSE of the vector parameter. This centering has a nonnegligible asymptotic impact.

6 Proofs of the results of Section 2

We begin with three general lemmas.

LEMMA 6.1 *Suppose X and Y are integrable random elements in a Hilbert space H . If X and Y are independent and $EX = 0$, then $E\langle X, Y \rangle = 0$.*

PROOF: This lemma essentially follows from the definition of the expectation in a Hilbert space, see e.g. Section 1.3 of Bosq (2000). First we use the relation $E\langle X, Y \rangle = E\{E[\langle X, Y \rangle | Y]\}$. Since the expectation commutes every linear operator, in particular with the inner product, $E[\langle X, Y \rangle | Y] = \langle E[X|Y], Y \rangle$. Since X is independent of Y and has mean zero, $E[X|Y] = EX = 0$, and the claim follows. ■

LEMMA 6.2 *Suppose $\{X_n\}$ and $\{Y_n\}$ are L^p - m -approximable sequences in Hilbert spaces H and K , respectively. Then, the sequence $\{(X_n, Y_n)\}$ is L^p - m -approximable in the product space $H \times K$.*

PROOF: Clearly,

$$(X_n, Y_n) = (f(u_n, u_{n-1}, \dots), g(v_n, v_{n-1}, \dots)) = h((u_n, v_n), (u_{n-1}, v_{n-1}), \dots)$$

has the required representation, and it follows that $(X_n, Y_n)^{(m)} = (X_n^{(m)}, Y_n^{(m)})$.

Since for $a, b \geq 0$ and $r > 0$, $(a + b)^r \leq 2^r(a^r + b^r)$, for any $X \in L_H^p$ and $Y \in L_K^p$,

$$\begin{aligned} \nu_p((X, Y)) &= \{E\|(X, Y)\|^p\}^{1/p} = \left\{E(\|X\|_H^2 + \|Y\|_K^2)^{p/2}\right\}^{1/p} \\ &\leq \sqrt{2}\{E\|X\|_H^p + E\|Y\|_K^p\}^{1/p} \leq \sqrt{2}2^{1/p}\{\nu_p(X) + \nu_p(Y)\}. \end{aligned}$$

It follows that the product analog of (2.10) holds. ■

LEMMA 6.3 *Suppose $\{X_n\}$ and $\{Y_n\}$ are L^4 - m -approximable sequences a Hilbert space H . Then $\{\langle X_n, Y_n \rangle\}$ is an L^2 - m -approximable real valued sequence*

PROOF: By the construction introduced in the proof of Lemma 6.2, we must show that

$$\sum_{m=1}^{\infty} \left\{ E \left| \langle X_n, Y_n \rangle - \langle X_n^{(m)}, Y_n^{(m)} \rangle \right|^2 \right\}^{1/2} < \infty.$$

Observe that

$$\left| \langle X_n, Y_n \rangle - \langle X_n^{(m)}, Y_n^{(m)} \rangle \right|^2 \leq 2 \|X_n - X_n^{(m)}\|^2 \|Y_n\|^2 + 2 \|Y_n - Y_n^{(m)}\|^2 \|X_n^{(m)}\|^2.$$

By the Cauchy–Schwarz inequality,

$$\begin{aligned} E \left| \langle X_n, Y_n \rangle - \langle X_n^{(m)}, Y_n^{(m)} \rangle \right|^2 &\leq 2 \left\{ E \|X_n - X_n^{(m)}\|^4 \right\}^{1/2} \left\{ E \|Y_n\|^4 \right\}^{1/2} \\ &\quad + 2 \left\{ E \|Y_n - Y_n^{(m)}\|^4 \right\}^{1/2} \left\{ E \|X_n^{(m)}\|^4 \right\}^{1/2} \\ &\leq C \left[\left\{ E \|X_n - X_n^{(m)}\|^4 \right\}^{1/2} + \left\{ E \|Y_n - Y_n^{(m)}\|^4 \right\}^{1/2} \right], \end{aligned}$$

where $C = 2 \left[\left\{ E \|X_1\|^4 \right\}^{1/2} + \left\{ E \|Y_1\|^4 \right\}^{1/2} \right]$. Therefore,

$$\left\{ E \left| \langle X_n, Y_n \rangle - \langle X_n^{(m)}, Y_n^{(m)} \rangle \right|^2 \right\}^{1/2} \leq \sqrt{C} \left[\left\{ E \|X_n - X_n^{(m)}\|^4 \right\}^{1/4} + \left\{ E \|Y_n - Y_n^{(m)}\|^4 \right\}^{1/4} \right],$$

and the claim follows from the assumed L^4 – m –approximability of the sequences $\{X_n\}$ and $\{Y_n\}$. ■

The following lemma is needed in the proofs of Theorems 2.1 and 2.2.

LEMMA 6.4 *If Assumptions 2.1 and 2.2 hold, then $\hat{\mathbf{F}} \xrightarrow{a.s.} \mathbf{F}$, where the matrices $\hat{\mathbf{F}}$ and \mathbf{F} are defined, respectively, by (2.5) and (2.7).*

PROOF: Consider the matrix

$$\tilde{\mathbf{F}} = \left[N^{-1} \sum_{n=1}^N \langle F_{nj} - \mu_j, F_{nk} - \mu_k \rangle, j, k = 1, 2, \dots, p \right] \quad (p \times p).$$

By the joint ergodicity of the F_{nj} , $\tilde{\mathbf{F}} \xrightarrow{a.s.} \mathbf{F}$. Using the decomposition

$$\begin{aligned} &\langle F_{nj} - \bar{F}_j, F_{nk} - \bar{F}_k \rangle \\ &= \langle F_{nj} - \mu_j + (\mu_j - \bar{F}_j), F_{nk} - \mu_k + (\mu_k - \bar{F}_k) \rangle \\ &= \langle F_{nj} - \mu_j, F_{nk} - \mu_k \rangle \\ &\quad + \langle F_{nj} - \mu_j, \mu_k - \bar{F}_k \rangle \\ &\quad + \langle \mu_j - \bar{F}_j, F_{nk} - \mu_k \rangle \\ &\quad + \langle \mu_j - \bar{F}_j, \mu_k - \bar{F}_k \rangle, \end{aligned}$$

we see that

$$\hat{F}(j, k) - \tilde{F}(j, k) = \langle \mu_j - \bar{F}_j, \mu_k - \bar{F}_k \rangle \xrightarrow{a.s.} 0.$$

PROOF OF THEOREM 2.1: Recall that $R_n^c = \sum_{k=1}^p \beta_k F_{nk}^c + \varepsilon_n^c$, with ε_n^c appearing in (2.2). Taking the inner product with F_{nj}^c and averaging over n , we obtain

$$\frac{1}{N} \sum_{n=1}^N \langle R_n^c, F_{nj}^c \rangle = \sum_{k=1}^p \left\{ \frac{1}{N} \sum_{n=1}^N \langle F_{nj}^c, F_{nk}^c \rangle \right\} \beta_k + \frac{1}{N} \sum_{n=1}^N \langle \varepsilon_n^c, F_{nj}^c \rangle. \quad (6.1)$$

In matrix notation, the above relation becomes $\hat{\mathbf{R}} = \hat{\mathbf{F}}\boldsymbol{\beta} + \boldsymbol{\varepsilon}^c$, with the vectors and matrices defined via (6.1). Consequently,

$$\hat{\mathbf{F}}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \boldsymbol{\varepsilon}^c, \quad (6.2)$$

with the k th component of $\boldsymbol{\varepsilon}^c$ given by

$$\begin{aligned} \frac{1}{N} \sum_{n=1}^N \langle \varepsilon_n^c, F_{nk}^c \rangle &= \frac{1}{N} \sum_{n=1}^N \{ \langle \varepsilon_n, F_{nk}^c \rangle - \langle \bar{\varepsilon}, F_{nk}^c \rangle \} \\ &= \left\{ \frac{1}{N} \sum_{n=1}^N \langle \varepsilon_n, F_{nk}^c \rangle \right\} - \langle \bar{\varepsilon}, \bar{F}_k \rangle \end{aligned} \quad (6.3)$$

In light of Lemma 6.4 and the assumed invertibility of \mathbf{F} , it remains to show that $\boldsymbol{\varepsilon}^c \xrightarrow{a.s.} \mathbf{0}$. This follows from the assumption $E\varepsilon_n = 0$ and the ergodicity of the sequence $\{(\varepsilon_n, F_{nk}^c)\}$. The claim $\hat{\boldsymbol{\beta}} \xrightarrow{a.s.} \boldsymbol{\beta}$ has thus been established.

To show that $\|\hat{\beta}_0 - \beta_0\| \rightarrow 0$ a.s., observe that

$$\hat{\beta}_0(t) - \beta_0(t) = \sum_{j=1}^p (\beta_j - \hat{\beta}_j) \bar{F}_j(t) + \bar{\varepsilon}(t) \quad (6.4)$$

and so

$$\|\hat{\beta}_0 - \beta_0\| \leq \sum_{j=1}^p |\beta_j - \hat{\beta}_j| \|\bar{F}_j\| + \|\bar{\varepsilon}\|.$$

Since the sequences $\{F_{jn}\}$ and $\{\varepsilon_n\}$ are ergodic with finite expectations in L^2 , we have $\bar{F}_j \xrightarrow{a.s.} EF_{j1}$ and $\bar{\varepsilon} \xrightarrow{a.s.} E\varepsilon_1 = 0$. Since the norm is a continuous function on L^2 , we conclude that $\|\bar{F}_j\| \xrightarrow{a.s.} \|EF_{j1}\|$ and $\|\bar{\varepsilon}\| \xrightarrow{a.s.} 0$. We established that $|\hat{\beta}_j - \beta_j| \xrightarrow{a.s.} 0$, so it follows that $\|\hat{\beta}_0 - \beta_0\| \xrightarrow{a.s.} 0$.

For ease of reference, we now state two limit theorems for L^2 - m -approximable vectors and functions, respectively. The following result was established in Aue *et al.* (2009).

THEOREM 6.1 Suppose $\{\xi_n\}$ is a d -dimensional L^2 - m -approximable mean zero sequence. Then the matrix valued sequence

$$\Gamma = \sum_{h=-\infty}^{\infty} E[\xi_0 \xi_h^T]$$

converges absolutely, and

$$N^{-1/2} \sum_{n=1}^N \xi_n \xrightarrow{d} \mathbf{W}, \quad (6.5)$$

where \mathbf{W} is a mean zero normal random vector with the covariance matrix Γ .

An analog of Theorem 6.1 for sequences of functions was established by Horváth *et al.* (2013).

THEOREM 6.2 Suppose $\{\varepsilon_n\}$ is an L^2 - m -approximable sequence of mean zero functions in $L^2([0, 1])$. Then the infinite sum

$$\gamma(t, s) = \sum_{h=-\infty}^{\infty} E[\varepsilon_0(t) \varepsilon_h(s)]$$

converges in $L^2([0, 1] \times [0, 1])$ (hence $\gamma(\cdot, \cdot)$ is square integrable). Moreover

$$N^{-1/2} \sum_{n=1}^N \varepsilon_n \xrightarrow{d} Z, \quad (6.6)$$

where Z is a mean zero Gaussian element of $L^2([0, 1])$ with the covariance function $E[Z(t)Z(s)] = \gamma(t, s)$.

None of the above theorems is exactly suited to our setting because the model parameters take values in the Hilbert space $H = L^2 \times R^p$ with the norm

$$\langle x, x^* \rangle = \int g(t)g^*(t)dt + \sum_{j=1}^p b_j b_j^*,$$

where

$$x = (g(\cdot), b_1, \dots, b_p), \quad x^* = (g^*(\cdot), b_1^*, \dots, b_p^*).$$

We therefore need a more abstract version of these theorems.

THEOREM 6.3 Suppose $\{X_n\}$ is an L^2 - m -approximable sequence of mean zero random elements in a separable Hilbert space H . Then the series

$$\Gamma = \sum_{h=-\infty}^{\infty} E[\langle X_h, \cdot \rangle X_0] \quad (6.7)$$

converges in the operator norm and

$$N^{-1/2} \sum_{n=1}^N X_n \xrightarrow{d} G,$$

where G is a mean zero Gaussian element of H with the covariance operator Γ given by (6.7).

The proof of Theorem 6.3 parallels that of Theorem 6.2, so we establish only the existence of the long run variance operator Γ whose form is an essential ingredient of the proof of Theorem 2.2.

PROOF OF THE CONVERGENCE IN (6.7): Under assumptions of Theorem 6.3, for every positive integer T ,

$$\Gamma(T) = \sum_{|h| \leq T} E[\langle X_h, \cdot \rangle X_0]$$

is a bounded linear operator acting on H , i.e. $\Gamma(T) \in \mathcal{L}(H)$. Recall that for every $L \in \mathcal{L}(H)$, its operator norm can be computed as

$$\|L\|_{\mathcal{L}} = \sup \{ |\langle Lx, y \rangle|, \|x\| = 1, \|y\| = 1 \}.$$

Therefore, for $T_1 < T_2$,

$$\begin{aligned} \|\Gamma(T_2) - \Gamma(T_1)\|_{\mathcal{L}} &= \sup_{\|x\|=1} \sup_{\|y\|=1} |\langle (\Gamma(T_2) - \Gamma(T_1))x, y \rangle| \\ &\leq \sup_{\|x\|=1} \sup_{\|y\|=1} \sum_{T_1 < |h| \leq T_2} |E[\langle X_h, x \rangle \langle X_0, y \rangle]|. \end{aligned}$$

Since

$$E[\langle X_h, x \rangle \langle X_0, y \rangle] = E[\langle X_h^{(h)}, x \rangle \langle X_0, y \rangle] + E[\langle X_h - X_h^{(h)}, x \rangle \langle X_0, y \rangle]$$

and $E[\langle X_h^{(h)}, x \rangle \langle X_0, y \rangle] = 0$, by (2.8) and (2.9), we see that

$$E[\langle X_h, x \rangle \langle X_0, y \rangle] = E[\langle X_h - X_h^{(h)}, x \rangle \langle X_0, y \rangle].$$

Notice further that

$$\begin{aligned} \left| E \left[\left\langle X_h - X_h^{(h)}, x \right\rangle \left\langle X_0, y \right\rangle \right] \right| &\leq \left\{ E \left\langle X_h - X_h^{(h)}, x \right\rangle^2 \right\}^{1/2} E \left\{ \left\langle X_0, y \right\rangle^2 \right\}^{1/2} \\ &\leq \left\{ E \left\| X_h - X_h^{(h)} \right\|^2 \right\}^{1/2} \left\{ E \left\| X_0 \right\|^2 \right\}^{1/2} \|x\| \|y\|. \end{aligned}$$

Combining the above, we see that

$$\|\Gamma(T_2) - \Gamma(T_1)\|_{\mathcal{L}} \leq \{E \|X_0\|^2\}^{1/2} \nu_2 \left(X_h - X_h^{(h)} \right).$$

By (2.10), the operators $\Gamma(T)$ form a Cauchy sequence in $\mathcal{L}(H)$, so series (6.7) converges in the norm of $\mathcal{L}(H)$. ■

PROOF OF THEOREM 2.2: By Lemma 6.3 and Lemma 6.1, vectors (2.13) form an L^2 - m -approximable sequence with $E\xi_n = \mathbf{0}$. By Lemma 6.2, the sequence $\{(\varepsilon_n(\cdot), \xi_n)\}$ is L^2 - m -approximable in the Hilbert space $H = L^2 \times R^p$. By Theorem 6.3, we conclude that

$$N^{-1/2} \sum_{n=1}^N (\varepsilon_n(\cdot), \xi_n) \xrightarrow{d} (Z(\cdot), \mathbf{W}), \quad (6.8)$$

where the random elements \mathbf{W} and Z are normal and have covariances specified in Theorems 6.1 and 6.2.

We first focus on the asymptotic distribution of $\sqrt{N}\varepsilon^c$, where the k th components of the vector ε^c is given by (6.3). Observe that k th components of $\sqrt{N}\varepsilon^c$ is

$$\begin{aligned} N^{-1/2} \sum_{n=1}^N \langle \varepsilon_n^c, F_{nk}^c \rangle &= N^{-1/2} \sum_{n=1}^N \langle \varepsilon_n, F_{nk} \rangle - \sqrt{N} \langle \bar{\varepsilon}, \bar{F}_k \rangle \\ &= N^{-1/2} \sum_{n=1}^N \langle \varepsilon_n, F_{nk} - \mu_k \rangle - N^{-1/2} \sum_{n=1}^N \langle \varepsilon_n, \bar{F}_k - \mu_k \rangle. \end{aligned}$$

Since $\|\bar{F}_k - \mu_k\| = o_P(1)$, the second term is asymptotically negligible, and so $\sqrt{N}\varepsilon^c$ has the same asymptotic distribution as the $N^{-1/2} \sum_{n=1}^N \xi_n$, with the vectors ξ_n given by (2.13). Consequently, by (6.2) and Lemma 6.4, $\sqrt{N}(\hat{\beta} - \beta)$ has the same asymptotic distribution as $\mathbf{F}^{-1} \left\{ N^{-1/2} \sum_{n=1}^N \xi_n \right\}$.

Next, note that by (6.4),

$$\sqrt{N}(\hat{\beta}_0(\cdot) - \beta_0(\cdot)) = N^{-1/2} \sum_{n=1}^N \varepsilon_n(\cdot) + \bar{\mathbf{f}}^T(\cdot) \sqrt{N}(\hat{\beta} - \beta),$$

where $\bar{\mathbf{f}}(t) = [\bar{F}_1(t), \bar{F}_2(t), \dots, \bar{F}_p(t)]^T$. By ergodicity, $\bar{\mathbf{f}} \xrightarrow{a.s.} \mathbf{f}$, where \mathbf{f} is defined by (2.14). It follows that the left-hand side of (2.11) has the same asymptotic distribution as

$$N^{-1/2} \sum_{n=1}^N \varepsilon_n(\cdot) + \mathbf{f}^T(\cdot) \mathbf{F}^{-1} \left\{ N^{-1/2} \sum_{n=1}^N \boldsymbol{\xi}_n \right\}, \quad \mathbf{F}^{-1} \left\{ N^{-1/2} \sum_{n=1}^N \boldsymbol{\xi}_n \right\}.$$

Relation (2.11) thus follows from (6.8). ■

References

- Aue, A., Hörmann, S., Horváth, L. and Reimherr, M. (2009). Break detection in the covariance structure of multivariate time series models. *The Annals of Statistics*, **37**, 4046–4087.
- Billingsley, P. (1995). *Probability and Measure*, 3rd edn. Wiley, New York.
- Black, F. (1972). Capital market equilibrium with restricted borrowing. *Journal of Business*, **45**, 444–454.
- Bosq, D. (2000). *Linear Processes in Function Spaces*. Springer.
- Bowsher, C. G. and Meeks, R. (2008). The dynamics of economic functions: Modeling and forecasting the yield curve. *Journal of the American Statistical Association*, **103**, 1419–1437.
- Campbell, J. Y., Lo, A. W. and MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton University Press, New Jersey.
- Diebold, F. and Li, C. (2001). Forecasting the term structure of government bond yields. *Journal of Econometrics*, **130**, 337–364.
- Fama, E. F. and French, K. R. (1995). Size and book-to-market factors in earnings and returns. *Journal of Finance*, **50**, 131–155.
- Gabrys, R., Horváth, L. and Kokoszka, P. (2010). Tests for error correlation in the functional linear model. *Journal of the American Statistical Association*, **105**, 1113–1125.
- Gneiting, T. (2011). Making and evaluating point forecasts. *Journal of the American Statistical Association*, **106**, 746–762.
- Hays, S., Shen, H. and Huang, J. Z. (2012). Functional dynamic factor models with application to yield curve forecasting. *The Annals of Applied Statistics*, **6**, 870–894.
- Hörmann, S., Horváth, L. and Reeder, R. (2013). A functional version of the ARCH model. *Econometric Theory*, **29**, 267–288.
- Hörmann, S. and Kokoszka, P. (2010). Weakly dependent functional data. *The Annals of Statistics*, **38**, 1845–1884.
- Hörmann, S. and Kokoszka, P. (2012). Functional time series. In *Time Series* (eds C. R. Rao and T. Subba Rao), Handbook of Statistics, volume 30. Elsevier.
- Horváth, L. and Kokoszka, P. (2012). *Inference for Functional Data with Applications*. Springer.

- Horváth, L., Kokoszka, P. and Reeder, R. (2013). Estimation of the mean of functional time series and a two sample problem. *Journal of the Royal Statistical Society (B)*, **75**, 103–122.
- Horváth, L., Kokoszka, P. and Rice, G. (2014). Testing stationarity of functional time series. *Journal of Econometrics*, **00**, 000–000; Forthcoming.
- Jones, C. M. and Kaul, G. (1996). Oil and the stock markets. *The Journal of Finance*, **51**, 463–491.
- Kargin, V. and Onatski, A. (2008). Curve forecasting by functional autoregression. *Journal of Multivariate Analysis*, **99**, 2508–2526.
- Kokoszka, P. and Reimherr, M. (2013). Predictability of shapes of intraday price curves. *The Econometrics Journal*, **16**, 285–308.
- Kokoszka, P. and Zhang, X. (2012). Functional prediction of intraday cumulative returns. *Statistical Modelling*, **12**, 378–398.
- Koopman, S. J., Mallee, M. I. P. and Van der Wel, M. (2010). Analyzing the term structure of interest rates using the dynamic nelson–siegel model with time-varying parameters. *Journal of Business & Economic Statistics*, **28**, 329–343.
- Lintner, J. (1965). The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, **47**, 13–37.
- Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investments*. John Wiley.
- Merton, R. (1973). An intertemporal capital pricing model. *Econometrica*, **41**, 867–887.
- Müller, H-G., Sen, R. and Stadtmüller, U. (2011). Functional data analysis for volatility. *Journal of Econometrics*, **165**, 233–245.
- Narayan, P. K. and Sharma, S. S. (2011). New evidence on oil price and firm returns. *Journal of Banking and Finance*, **35**, 3253–3262.
- Park, J. and Ratti, R. A. (2008). Oil price shocks and stock markets in the U.S. and 13 European countris. *Energy Economics*, **30**, 2587–2608.
- Pötscher, B. and Prucha, I. (1997). *Dynamic Non-linear Econonometric Models. Asymptotic Theory*. Springer.
- Ramsay, J., Hooker, G. and Graves, S. (2009). *Functional Data Analysis with R and MATLAB*. Springer.
- Ramsay, J. O. and Silverman, B. W. (2005). *Functional Data Analysis*. Springer.
- Ross, S. (1976). The arbitrage theory of capital asset pricing. *Journal of economic theory*, **13**, 341–360.
- Ruppert, D. (2011). *Statistics and Data Analysis for Financial Engineering*. Springer.
- Shao, X. and Wu, W. B. (2007). Asymptotic spectral theory for nonlinear time series. *The Annals of Statistics*, **35**, 1773–1801.
- Sharpe, W. (1964). Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance*, **19**, 425–442.

Shen, H. (2009). On modeling and forecasting time series of curves. *Technometrics*, **51**, 227–238.

Wu, W. (2005). *Nonlinear System Theory: Another Look at Dependence*. Proceedings of The National Academy of Sciences of the United States, volume 102.