Potential Applications of Function Data Analysis in High-frequency Financial Research

Hong Miao*
Finance and Real Estate, Colorado State University, USA

High-Frequency and intraday financial data in general, have been an important focus of research in finance, econometrics and statistics for over two decades. There are mainly two streams of studies in this area: 1) time series properties at tick by tick or fixed high frequency resolution, 2) inference based on diffusion processes. The distinction is not sharp. The focus of the statistical analysis has been on building of time series models at fine resolution, possibly at non-evenly spaced times, or on estimating quantities, i.e. the noise covariance's, implied by the continuous time financial theory.

Factor models, i.e. Capital Asset Pricing Model (CAPM), are commonly used at daily, monthly, and annual frequency financial research. However, since the intraday return series are very noisy, it is almost impossible to use traditional approaches such as OLS regressions to examine the relations between high frequency intraday return series. Thus, the lack of proper tools puts a limit on studies in this area. The Functional Data Analysis (FDA) technique might be a potential solution. FDA is about the analysis of information on curves or functions.

FDA is particularly applicable in high-frequency studies since statistical tools of FDA typically rely on some form of smoothing to transform high dimensional or incomplete data building up a curve into a smoother curve that can be described by a smaller number of parameters. For detailed background of FDA, see Ramsay [1]. It is not yet popularly used in finance. However, there are several working papers available. Müller et al. [2] introduce a functional volatility process for modeling volatility trajectories for high frequency observations in financial markets and describe functional representations and data-based recovery of the process from repeated observations. They show that the model is useful to identify recurring patterns of volatility and for successful prediction of future volatility, through the application of functional regression and prediction techniques. Dablemont et al. [3] apply FDA to intraday financial time series modeling and forecasting. They show that this approach is powerful to particularly effective when the observations are sparse, irregularly spaced, occur at different time points for each curve, or when only fragments of the curves are observed; standard methods completely fail in these circumstances.

The following simple example can show the potential power of FDA in high-frequency financial research.

As mentioned above, intraday return series are very noisy. However, we can study a different aspect, the shape of various intraday curves. If we denote the asset price on day n and at time t within that day by \( P_n(t) \). We define the Cumulative Intraday Return (CIDR) as

\[
\text{CIDR}(t) = \ln[P_n(t)] - \ln[P_n(0)]
\]

The CIDR's were introduced by Gabrys et al. [4] as a way of transforming the price curves to stationary. By describing how the return accumulates throughout the day, they are also more relevant to intraday trading than the price curves. It has been shown by Gabrys et al. [4] that the CIDR curve is stationary and thus the functional regression and statistical tools can be applied. For example, instead of studying the intraday relations between returns of individual securities and the market, we can study the relations between the CIDRs of individual securities and the market. Thus, functional factor models can be developed, such as:

\[
R_n(t) = \beta_0(t) + \beta_1 M_n(t) + \epsilon_n(t),
\]

where, \( \beta_i(\cdot) \) is a function, \( \beta_1 \) is a scalar which is similar to the traditional CAPM beta, and \( M_n(t) \) refers to the CIDR of the market on day \( n \) and at time \( t \) within that day. Similarly, we can have:

\[
R_n(t) = \beta_0(t) + \beta_1 M_n(t) + \beta_2 S_n + \beta_3 H_n + \epsilon_n(t),
\]

where, \( S_n \) and \( H_n \) are the “small (market capitalization) minus big” and for “high (book-to-market ratio) minus low” factors, and \( \beta_2 \) and \( \beta_3 \) are scalars to be estimated. Kokoszka et al. [5] develop statistical tools to estimate the function \( \beta_1(\cdot) \) and the scalar \( \beta_3 \), and tools to estimate and test significances of the parameters in a more general factor model [6].

In summary, FDA can transform very noisy intraday returns into smooth and stationary curves. The curves can thus be studied using the available statistical tools. This is a very interesting and potential area which is worth more attentions and awareness from the finance researchers.

References
1. Ramsay J, Silverman BW (2005) Functional Data Analysis, (2nd edn). Springer.
2. Müller H, Seri R, Stadtmüller U (2011) Functional Data Analysis for Volatility, Working paper, University of California, Davis.
3. Dablemont S, Bellegem SV, Verleysen M (2007) Modelling and Forecasting financial time series of “tick data” by functional analysis and neural networks.
4. Gabrys R, Horváth L, Kokoszka P (2010) Tests for error correlation in the functional linear model. Journal of the American Statistical Association 105: 1113–1125.
5. Kokoszka P, Miao H, Zhang X (2012) Functional multifactor regression for intraday price curves. Working paper, Colorado State University.
6. Hays S, Shen H, Huang JZ (2012) Functional dynamic factor models with application to yield curve forecasting. The Annals of Applied Statistics 6: 870–894.

*Corresponding author: Hong Miao, Finance and Real Estate, Colorado State University, USA. Tel: 970-491-2356; E-mail: Hong.Miao@business.colostate.edu

Received: January 07, 2013; Accepted: January 08, 2013; Published: January 09, 2013

Copyright: © 2013 Miao H. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.