

FUNCTIONAL COEFFICIENT MODELS FOR ECONOMIC AND FINANCIAL DATA¹

Zongwu CAI^{a,b}

^aDepartment of Mathematics & Statistics, University of North Carolina at
Charlotte, Charlotte, NC 28223, USA. E-mail: zcai@uncc.edu

^bThe Wang Yanan Institute for Studies in Economics, Xiamen University,
Xiamen, Fujian 361005, China

Abstract

This paper gives a selective overview on the functional coefficient models with their particular applications in economics and finance. Functional coefficient models are very useful analytic tools to explore complex dynamic structures and evolutions for functional data in various areas, particularly in economics and finance. They are natural generalizations of classical parametric models with good interpretability by allowing coefficients to be governed by some variables or to change over time, and also they have abilities to capture nonlinearity and heteroscedasticity. Furthermore, they can be regarded as one of dimensionality reduction methods for functional data exploration and have nice interpretability. Due to their great properties, functional coefficient models have had many methodological and theoretical developments and they have become very popular in various applications.

Key words: Bandwidth selection; Bootstrap; Capital asset pricing model; Dimensionality reduction; Endogeneity; Functional linear process; Functional varying coefficient model; Generalized likelihood ratio test; Instrumental variables model; Local linear estimation; Locally stationary model; Longitudinal data; Misspecification test; Piecewise stationary process; Structural change model; Threshold model; Time-varying model; Trending panel model.

Forthcoming in *Oxford Handbook on Statistics and Functional Data Analysis* (Eds: F. Ferraty, Y. Romain). Oxford University Press, Oxford, UK.

¹The author thanks Frederic Ferraty and two referees for their constructive comments and suggestions. This research was supported, in part, by the National Nature Science Foundation of China grant #70871003, and funds provided by the Cheung Kong Scholarship from Chinese Ministry of Education, the Minjiang Scholarship from Fujian Province, China, and Xiamen University.

1 Introduction

There is a swift growing literature in methodological and theoretical research on the functional coefficient models in the recent two decades. Particularly, due to their great flexibility and interpretability, the functional coefficient models have been extensively applied to economics and finance to capture the dynamic changes and evolutions in economic and financial phenomena. Given space limitations, it is impossible to survey all the important recent developments and applications in functional coefficient models. Therefore, I choose to limit my focus on the following areas. In Section 2, I review the recent developments of nonparametric estimation and testing of functional coefficient models. In Section 3 is devoted to two major real applications in economics and finance. For more about the methodology, theory and applications of functional coefficient models in statistics, economics, finance as well as other fields, the reader is referred to the review papers by Fan and Zhang (2008), Cai and Hong (2009) and Cai and Li (2009).

2 Functional Coefficient Models

2.1 The models

A general nonparametric regression model (see (1) in Chapter 1) can be written as

$$Y_t = g(\mathbf{X}_t, \mathbf{Z}_t) + \varepsilon_t, \quad 1 \leq t \leq T, \quad (1)$$

where Y_t is the dependent variable, both $\mathbf{X}_t \in \mathbb{R}^p$ and $\mathbf{Z}_t \in \mathbb{R}^q$ are regressors, and the regression function $g(\mathbf{x}, \mathbf{z})$ is a \mathbb{R}^{p+q} -dimensional surface. A functional (varying) coefficient model (see (20) in Chapter 1) has the following particular form

$$g(\mathbf{X}_t, \mathbf{Z}_t) = \sum_{j=0}^p a_j(\mathbf{Z}_t) X_{jt} = \mathbf{a}(\mathbf{Z}_t)^\top \mathbf{X}_t, \quad (2)$$

which is linear in \mathbf{X}_t and nonlinear in \mathbf{Z}_t and the nonparametric coefficient functions are in \mathbb{R}^q rather than in \mathbb{R}^{p+q} . Here, $X_{0t} = 1$, $\mathbf{a}(\mathbf{Z}_t) = (a_0(\mathbf{Z}_t), \dots, a_p(\mathbf{Z}_t))^\top$, $\mathbf{X}_t = (1, X_{1t}, \dots, X_{pt})^\top$, and \mathbf{A}^\top denotes the transpose of a matrix or vector \mathbf{A} . As in (21) in Chapter 1, one can assume that

$$E[\varepsilon_t | \mathbf{X}_t, \mathbf{Z}_t] = 0. \quad (3)$$

Then, both variables \mathbf{X}_t and \mathbf{Z}_t are exogenous. However, (3) might not be true for many applications in economics and finance. In such a case,

some components of \mathbf{X}_t are called endogenous variable; see (13) and (14) later for the detailed setting. Notice that the functional coefficient model given in (2) and (3) was proposed in Cleveland, Grosse and Shyu (1991) and studied extensively by Hastie and Tibshirani (1993). For more about statistical properties of functional coefficient models, the reader is referred to the statistical survey paper by Fan and Zhang (2008).

It is clear from (2) and (3) that the identification condition can be derived as follows

$$E \left[\mathbf{X}_t \mathbf{X}_t^\top | \mathbf{Z}_t \right] \mathbf{a}(\mathbf{Z}_t) = E[Y_t \mathbf{X}_t | \mathbf{Z}_t]$$

and

$$\mathbf{a}(\mathbf{z}) = \boldsymbol{\Omega}(\mathbf{z})^{-1} E[Y_t \mathbf{X}_t | \mathbf{Z}_t = \mathbf{z}]$$

provided that $\boldsymbol{\Omega}(\mathbf{z}) = E \left[\mathbf{X}_t \mathbf{X}_t^\top | \mathbf{Z}_t = \mathbf{z} \right]$ is nonsingular for all \mathbf{z} . Therefore, the sufficient and necessary condition to identify $\mathbf{a}(\mathbf{z})$ is $\boldsymbol{\Omega}(\mathbf{z}) > 0$ for all \mathbf{z} , which is an identification condition. When (3) does not hold, to estimate and identify functionals in (2), one need instrumental variables (IV); see Cai, Das, Xiong and Wu (2006, CDXW) for details. In what follows, it is assumed that the model is identified.

As elaborated by CDXW (2006), functional coefficient models are appropriate and flexible enough for many applications, in particular when additive separability of covariates is unsuitable for the problem at hand. For ease of notation, we assume here that $p = 1$ and $q = 1$. Indeed, by assuming that $g(x, z)$ has a higher order partial derivative with respect to x and applying Taylor expansion to $g(x, z)$, one obtains

$$g(x, z) = \sum_{j=1}^{\infty} \frac{\partial^j g(0, z)}{\partial x^j} \frac{x^j}{j!} \approx \sum_{j=0}^d a_j(z) x_j \quad (4)$$

for some d (large), where $a_j(z) = (j!)^{-1} \partial^j g(0, z) / \partial x^j$ and $x_j = x^j$. Equation (4) implies that a functional coefficient model in (2) might be a good approximation to a general nonparametric model in (1).

More importantly, functional coefficient model in (2) and (3) has an ability to capture heteroscedasticity. To get insights about this, it is easy to see that

$$\text{Var}(Y_t | \mathbf{Z}_t) = \mathbf{a}(\mathbf{Z}_t)^\top \text{Var}(\mathbf{X}_t | \mathbf{Z}_t) \mathbf{a}(\mathbf{Z}_t) + \sigma_\varepsilon^2(\mathbf{Z}_t),$$

where $\sigma_\varepsilon^2(\mathbf{Z}_t) = \text{Var}(\varepsilon_t | \mathbf{Z}_t)$. Therefore, the first term in the above expression behaves as an ARCH type model. Furthermore, the functional coefficient approach allows appreciable flexibility on the structure of fitted models without

suffering from the “curse of dimensionality” since the nonparametric estimation is conducted in \mathbb{R}^q instead of \mathbb{R}^{p+q} . Finally, functional coefficient model can be used as a tool to study covariate adjusted regression for situations where both predictors and response in a regression model are not directly observable, but are contaminated with a multiplicative factor that is determined by the value of an unknown function of an observable covariate (confounding variable); See Şentürk and Müller (2005) and Cai and Xu (2008) for more details.

Model in (2) and (3) covers many familiar models popularly used in the literature. For example, if \mathbf{Z}_t is time, it becomes to the following trending time varying coefficient model

$$Y_t = \sum_{j=0}^p a_j(t) X_{jt} + \varepsilon_t, \quad 1 \leq t \leq T, \quad (5)$$

which has an ability to capture the dynamic changes and evolutions with time. A trending time varying time series model given in (5) has gained a lot of attention during the last two decades due to many applications in economics and finance. Following are some examples. The market model in finance is an example that relates the return of an individual stock to the return of a market index or another individual stock and the coefficient usually is called a beta coefficient in the capital asset pricing model (CAPM); see the papers by Cochrane (1996) and Cai (2007) and the book by Tsay (2005) for more details on theory and real examples. However, some recent studies show that the beta coefficients might vary over time; see Cai (2007) and the references therein.² The term structure of interest rates is another example in which the time evolution of the relationship between interest rates with different maturities is evidenced; see Tsay (2005) for details. The last example is the relationship between the prototype electricity demand and other variables such as the income or production, the real price of electricity, and the temperature. Indeed, Chang and Martinez-Chombo (2003) found that this relationship may change over time based on the empirical study of the demand equation using monthly Mexican electricity data for residential, commercial and industrial sectors. Although the literature is already vast and continues to grow swiftly, as pointed out by Phillips (2001), the research in this area is just beginning.

²There are many recent developments on time varying beta coefficient CAPMs; see Bansal, Hsieh and Viswanathan (1993), Bansal and Viswanathan (1993), Jagannathan and Wang (1996), Ghysels (1998), Reyes (1999), Cui, He and Zhu (2002), Akdeniz, Altay-Salih, Caner (2003), Wang (2002, 2003) and You and Jiang (2007).

Also, notice that if X_{jt} in (5) is a lagged variable (say, $X_{jt} = Y_{t-j}$) and $\{a_j(t)\}$ satisfy some conditions, model (5) becomes the well known locally stationary time series model, which, proposed by Dahlhaus (1997), has a great ability to capture nonstationarity and nonlinearity; see Dahlhaus (1997) and Dahlhaus and Subba Rao (2006) for details on the theory of locally stationary time series models and their applications in finance.

If $\{a_j(t)\}$ are piecewise constant functions of time; that is

$$a_j(t) = \sum_{l=1}^m a_{jl} \mathbb{I}(T_{l-1} \leq t < T_l), \quad (6)$$

where $1 = T_0 < T_1 < \dots < T_m < T$, $\mathbb{I}(A)$ is the indicator function of the event A and $\{T_j\}$ are the structural change points which might be known or unknown, model in (6) includes the class of structural multiple change models popular in economics and finance and can characterize parameter instability. Parameter instability for economic and financial models is a common phenomenon. This is particularly true for time series data covering an extended period, as it is more likely for the underlying data generating mechanism to be disturbed over a longer horizon by various factors such as policy regime shift, macroeconomic announcements, global or regional financial crises, an unusually large unemployment announcement by a government, and a dramatic interest rate cut by the Federal Reserve and so on. For example, for the empirical problem discussed in the paper by Bai (1997), the finding is that the response pattern of interest rates to the changes in discount rates varies over time. The timing of variation is consistent with the timing of changes in the Fed's operating procedures. It is well known that failure to take into account parameter changes, given their presence, may lead to incorrect policy implications and predictions. On the other hand, proper treatment of parameter changes can be useful in uncovering the underlying factors that fostered the changes, in identifying misspecification, and in analyzing the effect of a policy change. There are a vast literature on structural change models; see the papers by Bai (1997), Bai and Perron (1998) and Bai and Perron (2003) and the references therein. Finally, if X_{jt} is a lagged variable ($X_{jt} = Y_{t-j}$), the model becomes the piecewise autoregressive models (see Davis, Lee and Rodriguez-Yam, 2006, for recent advances), which have an ability to depict nonstationarity and can approximate well the locally stationary time series models of Dahlhaus (1997); see Davis, Lee and Rodriguez-Yam (2006) for more discussion.

If $\{a_j(\mathbf{Z}_t)\}$ in (2) have the following particular parametric forms as

$$a_j(\mathbf{Z}_t) = \sum_{k=1}^m a_{jk} \mathbb{I}(\mathbf{Z}_t \in \Omega_k),$$

where $\{\Omega_k\}$ form a (non-overlapping) partition of the whole domain of \mathbf{Z}_t . This model is called threshold model, which is a special case of a nonlinear model. Theoretical properties and practical implementations of threshold modeling have been covered by Tong (1990). The threshold regression theory has gained a lot of momentum recently, for some of the selected studies in economics and finance; see Hansen (2000), Caner and Hansen (2001), Akdeniz, Altay-Salih and Caner (2003), and Caner and Hansen (2004).

Notice that since a structural change model or a threshold model is a parametric model and my focus is on nonparametric models, I will not review structural change models, locally stationary time series models, threshold models and piecewise stationary processes. The reader is referred to the aforementioned literatures on these models.

Finally, it is worth to pointing out that a functional coefficient model in (2) and (3) can be used to analyze the functional data as in Ramsay and Silverman (1997). For example, Ramsay and Silverman (1997), Wei, Pere, Koenker and He (2006), Wei and He (2006) and Şentürk and Müller (2008) extended model in (2) and (3) to the following form

$$Y_i(t_{ij}) = \mathbf{a}(t_{ij})^\top \mathbf{X}_i(t_{ij}) + \varepsilon_i(t_{ij}), \quad 1 \leq j \leq n_i, \quad 1 \leq i \leq n, \quad (7)$$

where $Y_{ij} = Y_i(t_{ij})$, $\mathbf{X}_{ij} = \mathbf{X}_i(t_{ij})$ and $\varepsilon_i(t_{ij})$ is a zero-mean process with covariance function $\delta(t, s) = \text{Cov}(\varepsilon_i(t), \varepsilon_i(s))$, and they used this model for longitudinal growth studies. Model (7) might have a potential application in economics and finance; see Cai, Hsiao and Zhu (2009) for studying trending panel data. Cardot and Sarda (2008) considered a generalization of the functional coefficient regression model which takes the form

$$Y = \int \mathbf{a}(\mathbf{Z}, t)^\top \mathbf{X}(t) dt + \varepsilon, \quad (8)$$

where \mathbf{Z} and ε are real random variables such that $E(\varepsilon|\mathbf{Z}) = 0$, $E(\mathbf{X}\varepsilon|\mathbf{Z}) = 0$ and $\text{Var}(\varepsilon|\mathbf{Z}) = \sigma^2$, and they used this model for ozone pollution forecasting. Indeed, model (8) can be modified and generalized to be a functional volatility process which provides a new tool for modeling volatility trajectories in financial markets; see Müller, Sen and Stadtmüller (2007) for more discussion.

2.2 Nonparametric modeling procedures

There are three major approaches in estimating the $\{a_j(\cdot)\}$ in model (2) if they are assumed to be continuous. The first one is kernel-local polynomial smoothing; see Cai, Fan and Li (2000) and Cai, Fan and Yao (2000). The second one is polynomial spline; see Huang and Shen (2004). The last one is smoothing spline; see Hastie and Tibshirani (1993). A local linear fitting has several great properties such as high statistical efficiency in an asymptotic minimax sense, adaptive design, and automatic edge correction (Fan and Gijbels, 1996). Therefore, in what follows, I am going to outline only the kernel local polynomial smoothing method and for other methods, the reader is referred to the aforementioned papers.

2.2.1 Nonparametric estimation of functional coefficients

For simplicity, in what follows, I assume that $q = 1$. I estimate the functions $\{a_j(\cdot)\}$ in (2) using the local linear regression method from observations $\{(\mathbf{X}_t, Y_t, Z_t)\}_{t=1}^T$. It is assumed throughout the article that $a_j(\cdot)$ has a continuous second derivative. Notice that one may approximate $a_j(Z_t)$ locally at any grid point $z \in \mathbb{R}$ by a linear function $a_j(Z_t) \approx a_j + b_j(Z_t - z)$. The local linear estimator is defined as $\hat{a}_j(z) = \hat{a}_j$, where $\{(\hat{a}_j, \hat{b}_j)\}$ minimize the sum of locally weighted least squares

$$\sum_{t=1}^T \left[Y_t - \sum_{j=0}^p \{a_j + b_j(Z_t - z)\} X_{jt} \right]^2 K_h(Z_t - z), \quad (9)$$

where $K_h(x) = h^{-1}K(x/h)$, $K(\cdot)$ is a kernel function on \mathbb{R} and $h > 0$ is a bandwidth which controls the degree of smoothing in estimation, and it satisfies $h \rightarrow 0$ and $hT \rightarrow \infty$ as $T \rightarrow \infty$.

Notice that the local linear estimator can be viewed as the least squares estimator of the following working linear (parametric) model

$$K_h^{1/2}(Z_t - z)Y_t = K_h^{1/2}(Z_t - z) \sum_{j=0}^p \{a_j + b_j(Z_t - z)\} X_{jt} + u_t.$$

Therefore, the estimator $\hat{a}_j(z)$ is a linear estimator of $a_j(z)$ (a linear combination of Y_1, \dots, Y_T) and computational implementation can be easily carried out by any standard statistical software.

Remark 1 *The restriction to the locally weighted least squares method suggests that normality is at least being considered as a baseline. However, when abnormality is clearly present, a local quasi likelihood approach can be used; see Cai (2003). If there are any outliers, one can use a robust local linear fitting scheme; see Cai and Ould-Said (2003). If some of \mathbf{X}_t are endogenous variables, the various instrumental variable type estimates of linear and nonlinear simultaneous equations and transformation models can be easily applied here with some modifications; see CDXW (2006) and Das (2005). Although such methods appropriately modified can be applied to the current setting, their asymptotic properties are not obvious (see Section 3 later).*

When Z_t is random, Cai, Fan and Yao (2000) showed that under some regularity conditions, $\hat{\mathbf{a}}(z)$ is asymptotically normally distributed; that is

$$\sqrt{hT} \left[\hat{\mathbf{a}}(z) - \mathbf{a}(z) - \frac{h^2}{2} \mu_2 \mathbf{a}''(z) + o_p(h^2) \right] \rightarrow N(0, \nu_0 \boldsymbol{\Sigma}(z)), \quad (10)$$

where $\mu_2 = \int u^2 K(u) du$, $\nu_0 = \int K(u)^2 du$ and $\boldsymbol{\Sigma}(z) = \boldsymbol{\Omega}(z)^{-1} \boldsymbol{\Omega}_1(z) \boldsymbol{\Omega}(z)^{-1} / f_z(z)$. Here, $f_z(z)$ is the marginal density of Z_t and $\boldsymbol{\Omega}_1(z) = E [\sigma^2(\mathbf{X}_t, Z_t) \mathbf{X}_t \mathbf{X}_t^\top | Z_t = z]$, where $\sigma^2(\mathbf{X}_t, Z_t) = \text{Var}(\varepsilon_t | \mathbf{X}_t, Z_t)$. When Z_t is time (indeed, Z_t is normalized as $Z_t = s_t = t/T$; see Cai, 2007, for details), Cai (2007) showed that under some regularity conditions, for any $s \in [0, 1]$,

$$\sqrt{hT} \left[\hat{\mathbf{a}}(s) - \mathbf{a}(s) - \frac{h^2}{2} \mu_2 \mathbf{a}''(s) + o_p(h^2) \right] \rightarrow N(0, \nu_0 \boldsymbol{\Sigma}_s), \quad (11)$$

where $\boldsymbol{\Sigma}_s = \boldsymbol{\Omega}^{-1} \boldsymbol{\Omega}_1(s) \boldsymbol{\Omega}^{-1}$, $\boldsymbol{\Omega} = E [\mathbf{X}_t \mathbf{X}_t^\top]$ and $\boldsymbol{\Omega}_1(s) = \sum_{k=-\infty}^{\infty} \mathbf{R}_k(s)$ with $\mathbf{R}_k(t) = \text{Cov}(\sigma(\mathbf{X}_{i+k}, t) u_{i+k} \mathbf{X}_{i+k}, \sigma(\mathbf{X}_i, t) u_i \mathbf{X}_i)$, $\varepsilon_t = \sigma(\mathbf{X}_t, t) u_t$ and $\{u_t\}$ is stationary and $\text{Var}(u_t) = 1$.

From (10) and (11), it is easy to drive the asymptotic mean squares error (AMSE) which is the asymptotic variance plus the square of the asymptotic bias and to drive the optimal bandwidth by minimizing the AMSE. Clearly, the optimal bandwidth is $h_{opt} = cT^{-1/5}$ for some unknown positive constant c which can be estimated based on a data-driven fashion, described below.

2.2.2 Bandwidth selection

It is well known that the bandwidth plays an essential role in the trade-off between reducing bias and variance. By following a similar idea in Cai and

Tiwari (2000), Cai (2002) and Cai (2007), here I adapt a simple and quick method to select the bandwidth for the foregoing estimation procedures, described as follows. For the given observed values $\{Y_t\}_{t=1}^T$, the fitted values $\{\hat{Y}_t\}_{t=1}^T$ can be expressed as $\hat{\mathbf{Y}} = \mathbf{H}_h \mathbf{Y}$, where $\mathbf{Y} = (Y_1, \dots, Y_T)^\top$ and \mathbf{H}_h is called the $T \times T$ smoother (or hat) matrix associated with the smoothing parameter h . Motivated by the ideas in Cai and Tiwari (2000) and Cai (2002), I use the following nonparametric version of AIC to select the optimal bandwidth h_{opt} by minimizing

$$\text{AIC}(h) = \log(\hat{\sigma}^2) + 2(T_h + 1)/(T - T_h - 2),$$

where $\hat{\sigma}^2 = \sum_{t=1}^T (\hat{Y}_t - Y_t)^2 / T$ and T_h is the trace of the hat matrix \mathbf{H}_h . This selection criterion counteracts the over/under-fitting tendency of the generalized cross-validation and the classical AIC; see Cai and Tiwari (2000) and Cai (2002) for more details. Alternatively, one might use some existing methods in the time series literature although they may require more computing; see Fan and Gijbels (1996), Cai, Fan and Yao (2000) and Cai (2007). This bandwidth selection criterion will be used in Section 3 for real examples.

2.3 Misspecification testing

An important econometric question in fitting model (2) or (5) is that there is a need to test the following scenarios: (1) whether all coefficient functions are actually varying (namely, if a linear model is adequate); (2) more generally, if a parametric model fits the given data such as testing for structural breaks as in Bai (1997) and Bai and Perron (1998, 2003) or testing a threshold model as in Hansen (2000), Caner and Hansen (2001), Akdeniz, Altay-Salih and Caner (2003), and Caner and Hansen (2004) or a specific parametric form as in Ferson and Harvey (1998, 1999), Cai, Fan and Yao (2000) and Cai, Fan and Li (2000); (3) if there is no $a_0(\cdot)$ at all; and (4) whether there are some economic variables not statistically significant. This amounts to testing whether some or all coefficient functions are constant or zero or in a certain parametric form. This testing problem can be formulated as

$$H_0 : a_j(z) = a_j^*(z, \gamma), \tag{12}$$

where $a_j^*(z, \gamma)$ is a given family of functions indexed by an unknown parameter vector γ . Some tests similar to (12) have been considered in the econometrics and finance literature; see, for example, Ghysels (1998) by using the supreme lagrange multiplier (LM) test proposed by Andrews (1993)

for testing the structural break and Akdeniz, Altay-Salih and Caner (2003) by applying the heteroskedasticity consistent LM test for a threshold as in Hansen (1996).

For an easy implementation purpose, I adapt a misspecification test based on comparing the residual sum of squares (RSS) from both parametric (H_0) and nonparametric fittings (H_a), described as follows. Let $\hat{\gamma}$ be a consistent estimator of γ (say MLE or LSE). The RSS under the null hypothesis is $\text{RSS}_0 = T^{-1} \sum_{t=1}^T e_{t,0}^2$, where $e_{t,0} = Y_t - \sum_{j=0}^p a_j^*(Z_t, \hat{\gamma}) X_{tj}$ and the RSS under H_a is $\text{RSS}_1 = T^{-1} \sum_{t=1}^T e_{t,1}^2$, where $e_{t,1} = Y_t - \sum_{j=0}^p \hat{a}_j(Z_t) X_{tj}$. The test statistic is defined as

$$J_T = (\text{RSS}_0 - \text{RSS}_1) / \text{RSS}_1 = \text{RSS}_0 / \text{RSS}_1 - 1,$$

which can be regarded as a generalized F -test statistic (see Cai and Tiwari, 2000; Cai, 2002) and a generalized likelihood ratio test statistic (see Fan, Zhang and Zhang, 2001). The null hypothesis (12) is rejected for a large value of J_T . For simplicity, the p -value is computed by using the following nonparametric wild bootstrap approach that can accommodate heteroscedasticity in the model. Notice that this kind of test has been used in the statistics and econometrics literature by several authors; see, for example, Cai, Fan and Yao (2000), Cai and Tiwari (2000), and Cai (2007) for various applications in economics and finance and Cai, Fan and Li (2000) and Fan, Zhang and Zhang (2001) for applications in other areas.

The steps for the wild bootstrap sampling scheme are described as follows.

1. Generate the residuals $\{e_t^b\}_{t=1}^T$ from the centered nonparametric residuals $\{e_t^0\}_{t=1}^T$, where $e_t^b = e_{t,1} - \bar{e}_{t,1}$ with $\bar{e}_{t,1} = T^{-1} \sum_{t=1}^T e_{t,1}$.
2. Define the bootstrap sample $Y_t^b = \sum_{j=0}^p a_j^*(Z_t, \hat{\gamma}) X_{tj} + e_t^b$. In practice, one can define $e_t^b = e_t^0 \cdot \eta_t$, where $\{\eta_t\}$ is a sequence of iid random variables with mean zero and unit variance.
3. Calculate the bootstrap test statistic J_T^* based on the bootstrap sampling sample $\{(Y_t^b, \mathbf{X}_t, Z_t)\}_{t=1}^T$. Notice that for simplicity, the same bandwidth might be used in calculating both J_T^* and J_T .
4. Compute the p -value of the test based on the relative frequency of the event $\{J_T^* \geq J_T\}$ in the replications of the bootstrap sampling.

Remark 2 *At the first step, the reason why one bootstraps the centralized residuals from the nonparametric fit instead of the parametric fit is that the nonparametric estimate of residuals is always consistent, no matter whether the null or the alternative hypothesis is correct. Therefore, the method should provide a consistent estimator of the null distribution even when the null hypothesis does not hold. The consistency issue addressed in Cai, Fan and Yao (2000) can be applied to the setting here (see Cai, Fan and Yao, 2000, for details). This testing procedure will be used in Example 2 in Section 3 for a real application.*

3 Applications in Economics and Finance

There are many applications of functional coefficient models in economics and finance but I only present two real examples in the next two subsections due to space limitations. For more empirical examples in economics and finance, the reader is referred to the aforementioned papers and the additional papers by Li, Huang, Li and Fu (2001), Hong and Lee (2003), Fan and Zhang (2003), Fan, Jiang, Zhang and Zhou (2005), Cai, Kuan and Sun (2009) and Cai and Wang (2009) and the references therein.

3.1 Functional coefficient instrumental variables models

Functional coefficient models are appropriate for many economics applications. For example, here is a labor economics problem. A large body of work has established that while positive, marginal returns to education vary with the level of schooling (see Schultz, 1997), if work experience is also an attribute valued by employers, then the marginal returns to education should vary with experience. In fact, Card (2001) suggested that if a wage model assumes the additive separability of education and experience, the returns to education will be understated at higher levels of education because the marginal return to education is plausibly increasing in work experience. This setting is therefore a natural one for a functional coefficient model.

Under a functional coefficient representation, the nonparametric structural model no longer exhibits the ill-posed problem of Newey and Powell (2003). CDXW (2006) showed that under standard regularity conditions the model is identified and the estimators are well-defined with known asymptotic distribution. It is also shown that under this representation the estimators obtain faster convergence rates relative to analogous structural models that do not satisfy a functional coefficient representation.

Das (2005) considered a nonparametric IV model with discrete endogenous variables,

$$Y_i = g(X_i, \mathbf{Z}_{1i}) + \varepsilon_i, \quad (13)$$

where X_i is a discrete endogenous variable and \mathbf{Z}_{1i} is an exogenous variable. Here $E[\varepsilon_i | X_i, \mathbf{Z}_{1i}] \neq 0$. Without loss of generality, I assume that $X_i = 0$ or 1. Then, $g(x, \mathbf{z}_1)$ in (13) can be rewritten as

$$g(x, \mathbf{z}_1) = g(0, \mathbf{z}_1)\mathbb{I}(x = 0) + g(1, \mathbf{z}_1)\mathbb{I}(x = 1) = a_0(\mathbf{z}_1) + a_1(\mathbf{z}_1)x,$$

where $a_0(\mathbf{z}_1) = g(0, \mathbf{z}_1)$ and $a_1(\mathbf{z}_1) = g(1, \mathbf{z}_1) - g(0, \mathbf{z}_1)$. Therefore, $g(x, \mathbf{z}_1)$ is linear in endogenous variable but nonlinear in exogenous variable \mathbf{Z}_1 . Clearly, this model belongs to the class of functional coefficient models. Therefore, CDXW (2006) studied the following functional coefficient IV model

$$Y_i = \sum_{j=1}^d a_j(\mathbf{Z}_{i1}) X_{ij} + u_i = \mathbf{a}(\mathbf{Z}_{i1})^\top \mathbf{X}_i + u_i, \quad E[u_i | \mathbf{Z}_i] = 0, \quad (14)$$

where Y_i is an observable scalar random variable, $\{a_j(\cdot)\}$ are the unknown structural functions of interest, $X_{i0} \equiv 1$, $\mathbf{X}_i = (X_{i0}, X_{i1}, \dots, X_{id})^\top$ is a $(d+1)$ -dimension vector consisting of d endogenous regressors, $\mathbf{a}(\mathbf{Z}_{i1}) = (a_0(\mathbf{Z}_{i1}), \dots, a_d(\mathbf{Z}_{i1}))^\top$, and \mathbf{Z}_i is a $(d_1 + d_2)$ -dimension vector consisting of a d_1 -dimension vector \mathbf{Z}_{i1} of exogenous variables and a d_2 -dimension vector \mathbf{Z}_{i2} of instrumental variables.

Model (14) includes the following nonparametric IV model with binary endogenous variable D_i as a special case:

$$Y_i = a_0(\mathbf{Z}_{i1}) + a_1(\mathbf{Z}_{i1})D_i + \varepsilon_i,$$

which, as noted above, is analyzed in Das (2005). Further, if $a_j(\cdot)$ is a threshold function such as

$$a_j(\mathbf{z}) = a_{j1}\mathbb{I}(\mathbf{z} \leq \mathbf{r}_j) + a_{j2}\mathbb{I}(\mathbf{z} > \mathbf{r}_j)$$

for some \mathbf{r}_j , then model (14) may describe a threshold IV regression model. Indeed, Caner and Hansen (2004) considered a threshold model related to this with endogenous covariates. In this way, the class of models in (14) includes some interesting special cases that arise commonly in empirical research in economics.

To estimate $\{a_j(\mathbf{z}_1)\}$ in (14) nonparametrically, I propose using a two-stage nonparametric method as in CDXW (2006), described as follows. I begin with the first stage, where I obtain $\hat{\pi}_j(\mathbf{Z}_i)$, the fitted value for $\pi_j(\mathbf{Z}_i) = E[X_{ij}|\mathbf{Z}_i]$ ($1 \leq j \leq d; 1 \leq i \leq n$). To this end, I apply the local linear fitting technique and the jackknife (leave-one-out) idea as follows. Assuming that $\{\pi_j(\cdot)\}$ has a continuous second derivative, when \mathbf{Z}_k falls in a neighborhood of \mathbf{Z}_i , a Taylor expansion approximates $\pi_j(\mathbf{Z}_k)$ by

$$\pi_j(\mathbf{Z}_k) \approx \pi_j(\mathbf{Z}_i) + (\mathbf{Z}_k - \mathbf{Z}_i)^\top \pi_j'(\mathbf{Z}_i) = \alpha_{ij} + (\mathbf{Z}_k - \mathbf{Z}_i)^\top \boldsymbol{\beta}_{ij}.$$

The jackknife idea is to use the all observations except the i th observations in estimating $\pi_j(\mathbf{Z}_i)$. Then, the least squares estimator with a local weight (i.e., locally weighted least squares) is given by

$$\sum_{k \neq i}^n \left\{ X_{kj} - \alpha_{ij} - (\mathbf{Z}_k - \mathbf{Z}_i)^\top \boldsymbol{\beta}_{ij} \right\}^2 K_{h_1}(\mathbf{Z}_k - \mathbf{Z}_i).$$

Minimizing the above locally weighted least square with respect to α_{ij} and $\boldsymbol{\beta}_{ij}$ gives the local linear estimate of $\pi_j(\mathbf{Z}_i)$ by $\hat{\pi}_{j,-i}(\mathbf{Z}_i) = \hat{\alpha}_{ij}$. Now, I derive the local linear estimator of $\{a_j(\cdot)\}$. The local linear estimators \hat{b}_j and $\hat{\mathbf{c}}_j$ are defined as the minimizers of the sum of weighted least squares

$$\sum_{i=1}^n \left[Y_i - \sum_{j=0}^d \left\{ b_j + (\mathbf{Z}_{i1} - \mathbf{z}_1)^\top \mathbf{c}_j \right\} \hat{\pi}_{j,-i}(\mathbf{Z}_i) \right]^2 L_{h_2}(\mathbf{Z}_{i1} - \mathbf{z}_1),$$

and $\hat{a}_j(\mathbf{z}_1) = \hat{b}_j$. CDXW (2006) showed that under some regularity conditions, $\hat{a}_j(\mathbf{z}_1)$ is asymptotically normally distributed. Also, CDXW (2006) suggested an ad hoc bandwidth selection procedure to select two bandwidths in a data-driven fashion; see CDXW (2006) for details. At the first step, the bandwidth is chosen as small as possible and at the second step, one can use the data-driven method mentioned in Section 2.2.2 to choose the optimal bandwidth. This bandwidth selection criterion will be used in Example 1 below.

Example 1. I investigate the empirical relation between wages and education, using a random sample of young Australian female workers from the 1985 wave of the Australian Longitudinal Survey. The endogeneity of education in a wage model due to unobservable heterogeneity in schooling choices is well known in the literature; see e.g., the review in Card (2001). I consider the following functional coefficient specification:

$$Y = \delta_0(Z_{12})^\top \mathbf{Z}_{11} + g_0(Z_{12}) + g_1(Z_{12}) X + \varepsilon$$

and $E(X | \mathbf{Z}_{11}, Z_{12}, Z_2) = \pi(\mathbf{Z}_{11}, Z_{12}, Z_2)$, where Y is the natural logarithm of the hourly wage, \mathbf{Z}_{11} includes indicators for marital status, government employed, union status and Australian-born, Z_{12} is a measure of work experience measured in years, X is the measure of (endogenous) education (“Schooling”), Z_2 is an instrumental variable, and $g_0(\cdot)$, $g_1(\cdot)$ and $\pi(\cdot)$ are unknown functions. The object of interest is $g_1(\cdot)$, the functional coefficient of education, that depends on the level of experience.

The main results from estimation of this model are summarized in Figure 1 which plots the estimators of the functional coefficient $g_1(\cdot)$ correcting for endogeneity (the smooth solid line), and without correcting for endogeneity (the dashed line). First, notice that the profile without correcting

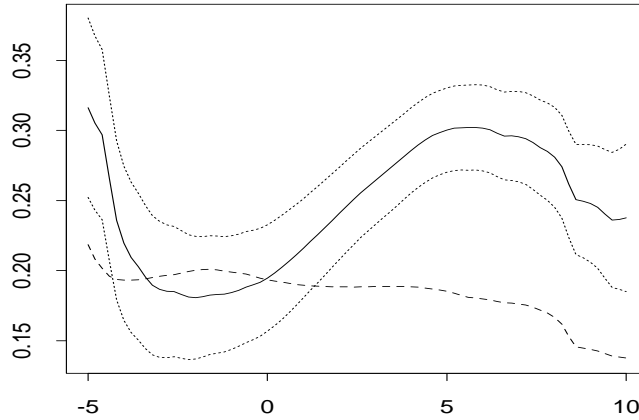


Figure 1: Functional coefficient estimates. The figure corresponds to the functional coefficient $g_1(\cdot)$, graphing the two-stage local linear estimate (solid line) with pointwise 95% confidence intervals (dotted lines), and the ordinary nonparametric estimate (dashed line).

for endogeneity is almost constant. In the profile correcting for endogeneity, one can find that the range of $\hat{g}_1(\cdot)$ is positive and nonlinear for all values of experience in the sample. This implies that holding experience fixed at any level, the marginal wage returns to schooling (given by the functional coefficient) are strictly positive although. In addition, I provide the 95 percent pointwise confidence interval (dotted lines) for the profile, which shows clearly that the pointwise confidence interval does not contain a constant function. This implies that $\hat{g}_1(\cdot)$ indeed is not a constant function. This result shows that the functional coefficient model captures the known non-

linear effect of education on wages as discussed in Card (2001). However, the confidence intervals indicate that the results correcting for endogeneity are statistically significant only for experience between 0 and 15. Indeed, CDXW (2006) gave details on how to construct a pointwise confidence interval for functional coefficient IV models.

Finally, notice also that the derivative of $\hat{g}_1(\cdot)$ changes over its range, being negative at both low and high levels of experience but positive in the middle range of experience. This suggests that while the marginal returns to education are positive, these returns are themselves declining in experience for both low experienced and high experienced workers.

3.2 Functional coefficient beta models

Although there is a vast amount of empirical evidences on time variation in betas and risk premia, there is no theoretical guidance on how betas and risk premia vary with time or variables that represent conditioning information. Many recent studies focus on modelling the variation in betas using continuous approximation and the theoretical framework of the conditional CAPM; see Cochrane (1996), Jaganathan and Wang (1996, 2002), Wang (2002, 2003) and Ang and Liu (2004) and the references therein. Recently, Ghysels (1998) discussed the problem in detail and stressed the impact of misspecification of beta risk dynamics on inference and estimation and argued that betas change through time very slowly and linear factor models like the conditional CAPM may have a tendency to overstate the time variation. Further, he showed that among several well known time varying beta models, a serious misspecification produces time variation in beta that is highly volatile and leads to large pricing errors.

To combine the aforementioned varying coefficient beta models under a unified framework, I consider a general nonparametric econometric model

$$r_{i,t} = \beta_{i,0}(\mathbf{Z}_{i,t}) + \beta_{i,1}(\mathbf{Z}_{i,t})^\top \mathbf{r}_{m,t} + \varepsilon_{i,t}, \quad 1 \leq i \leq N \quad \text{and} \quad 1 \leq t \leq T, \quad (15)$$

where $r_{i,t}$ is the i -th excess return on any asset or portfolio, $\beta_{i,1}(\cdot)$ is a $p \times 1$ vector of the varying coefficient betas, the prime denotes the transpose of a matrix or vector, $\mathbf{r}_{m,t}$ represents a $p \times 1$ vector of the excess returns on the market portfolios or indices, and $\mathbf{Z}_{i,t}$ is a set of instruments. Here, $\mathbf{Z}_{i,t}$ is either a set of instruments or time, both $\beta_{i,0}(\cdot)$ and $\beta_{i,1}(\cdot)$ are unknown functions, and $\varepsilon_{i,t}$ is the error term satisfying $E[\varepsilon_{i,t} | \mathbf{r}_{m,t}, \mathbf{Z}_{i,t}] = 0$. It is common in the finance literature to assume that the market return $\mathbf{r}_{m,t}$ and

the state variable $\mathbf{Z}_{i,t}$ are uncorrelated with the error term $\varepsilon_{i,t}$; see Akdeniz, Altay-Salih and Caner (2003). Here, I allow that the error terms $\{\varepsilon_{i,t}\}$ might be autocorrelated among both i and t and the conditional variance $\sigma_{i,t}^2 = \text{Var}[\varepsilon_{i,t} | \mathbf{r}_{m,t}, \mathbf{Z}_{i,t}]$ might not be constant. This time varying conditional heteroscedasticity can commonly be seen in many financial applications; see Reyes (1999) and Cho and Engle (2000) and the references therein. Further, I assume that the series $\{(u_{i,t}, \mathbf{r}_{m,t})\}$ is strictly stationary α -mixing, where $u_{i,t} = \varepsilon_{i,t}/\sigma_{i,t}$. Clearly, model (15) covers the aforementioned models as a special case and some other models in the finance literature.

Some specific examples of model (15) were studied by several authors in the literature; see, for example, Ferson and Harvey (1998, 1999) in which the betas are a linear function of instruments (an index model). To measure the risk of an individual stock against the market of US stocks, Cui, He and Zhu (2002) considered the following structural change model

$$r_t = \beta_0(t, T_0) + \beta_1 r_{m,t} + \varepsilon_t,$$

where r_t is the daily return of the Microsoft stock, $r_{m,t}$ is the Standard & Poor's 100 index, as a proxy to this market, and $\beta_0(t, T_0)$ is a structure change function with unknown change point T_0 , and You and Jiang (2007) extended the above model to a semi-parametric setting

$$r_t = \beta_{0,1} \mathbb{I}(t \leq T_0) + \beta_{0,2} \mathbb{I}(t > T_0) + \beta_1(t) X_t + \varepsilon_t$$

with $T_0 = 64$, where $\beta_1(\cdot)$ is an unknown smooth function. Recently, Ferson and Harvey (1998, 1999) and Harvey (1989) studied some parametric models by assuming the betas to be linear combinations of the world market-wide information variables and/or the attributes for the security, whereas Akdeniz, Altay-Salih and Caner (2003) investigated the threshold CAPM with economic variable(s)

$$r_t = \beta_0 + (\beta_{11} \mathbb{I}(\xi_t \leq \lambda) + \beta_{12} \mathbb{I}(\xi_t > \lambda)) r_{m,t} + \varepsilon_t,$$

where λ is an unknown threshold parameter and ξ_t is one of some economic variables such as one month real t-bill rate, dividend yield of the CRSP value weighted NYSE stock index, de-trended stock price level, measure of the slope of the term structure, and quality related yield spread in the corporate bond market.

If $Z_{i,t}$ is just time t , $\beta_{i,j}(t)$ depends on time t . To estimate $\beta_{i,j}(t)$ non-parametrically, as argued by Robinson (1989), it is necessary to assume

$\beta_{i,j}(t)$ to depend on the sample size T to provide the asymptotic justification for any nonparametric smoothing estimators. Following this convention, I assume that $\beta_{i,j}(t) = \beta_{i,j}(Z_t)$, $0 \leq j \leq 1$, where $Z_t = t/T$ and $\beta_{i,j}(\cdot)$ is an unknown function. The intuitive explanation to this “intensity” assumption is that it is an increasingly intense sampling of data points to derive the consistent estimation; see Robinson (1989) and Cai (2007) for more discussion. Similarly, I might assume that $\sigma_{i,t} = \sigma_i(r_{m,t}, Z_t)$ for some unknown function $\sigma_i(\cdot, \cdot)$. To estimate the beta functions $\beta_{i,j}(z_0)$ nonparametrically at any given grid point z_0 , one can apply the formulation in (9) to this setting with a minor modification but the details are omitted due to similarity. The bandwidth selection criterion described in Section 2.2.2 can be applied here, in particular in our implementation in Example 2 below.

Example 2. I apply the proposed time-varying beta model in (15) and its modeling procedures to analyze the common stock price (P_{1t}) of Microsoft (MSFT) during the year 2000 using the daily closing prices. To measure its risk relative to the market of U.S. blue chip stocks, I take the Standard & Poor’s 100 index (P_{2t}) as a proxy to this market. For the first 10-month period with 206 observations, Cui, He and Zhu (2002) modelled the stock price gains Y_t (the price at t -th day divided by the price on day one) and X_t the change in the market index from day one to the t -th day through the following threshold model

$$Y_t = \beta_0(t, T_0) + \beta X_t + \varepsilon_t,$$

where $\beta_0(t, T_0)$ is a threshold function with unknown change point T_0 with the estimated value $\hat{T}_0 = 64$. Recently, You and Jiang (2007) extended the above model to a semiparametric setting

$$Y_t = \beta_{0,1} \mathbb{I}(t \leq 64) + \beta_{0,2} \mathbb{I}(t > 64) + \beta_1(t) X_t + \varepsilon_t$$

and they used a penalized spline method to estimate the unknown slope function $\beta_1(t)$. Following the convention in the finance literature, here I consider the simple daily stock returns $r_t = P_{1t}/P_{1,t-1} - 1$ for the MSFT price and $r_{m,t} = P_{2t}/P_{2,t-1} - 1$ for the S&P 100 Index. It should be noted that the returns of the S&P 100 index may not be as nonstationary as the stock returns of Microsoft.

To establish the empirical relationship between the returns of MSFT and S&P 100 Index, similar to Tsay (2005) who considered the linear relationship

between the 1-year Treasury constant maturity rate and the 3-year Treasury constant maturity rate, I first fit the following simple beta model

$$r_t = \alpha_0 + \alpha_1 r_{m,t} + \varepsilon_t.$$

Notice that in finance and security analysis, α_1 measures the risk of an individual stock or portfolio as its (standardized) beta coefficient in CAPM against a market index or portfolio. If α_1 is greater than 1, the change in this stock price is expected to be more than that in the market index and thus the stock is regarded as one risky stock. As a result, the least squares estimates of α_0 and α_1 are $-0.0027(0.0018)$ and $1.3612(0.1243)$ respectively, which are plotted (dashed line) in Figure 2. By comparing these results with

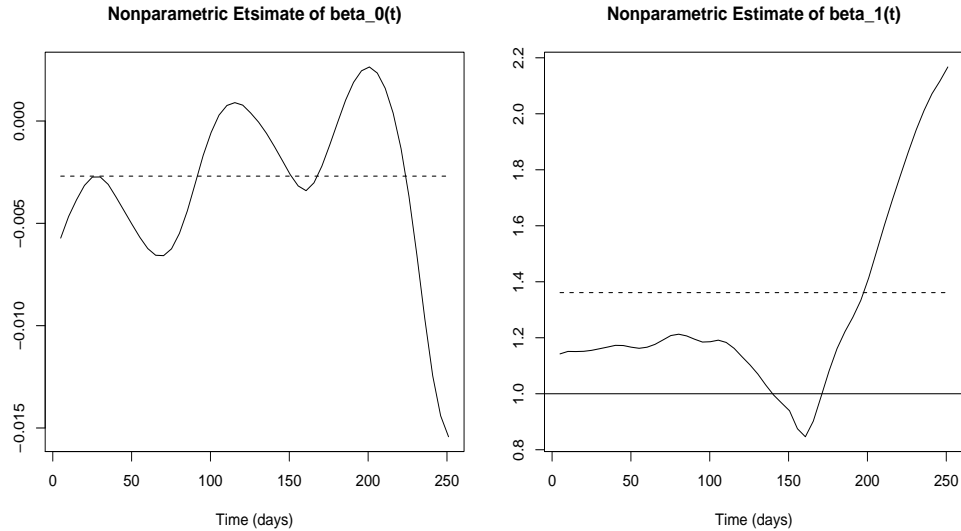


Figure 2: Results for Example 2. Left panel: The local linear estimator (solid line) of the trend function $\beta_0(\cdot)$ and the least square estimate of α_0 (dashed line). Right panel: The local linear estimator (solid line) of the beta coefficient $\beta_1(\cdot)$ and the least square estimate of α_1 (dashed line).

those from Cui, He and Zhu (2002) and You and Jiang (2007), I suspect that the coefficients α_0 and α_1 might change over time. To provide more empirical evidence, I examine the covariance between r_t and $r_{m,t}$, and I find that the covariance does change over time, which is not presented due to space limitations. Therefore, due to sufficient reasons, I fit the following time varying beta model

$$r_t = \beta_0(t) + \beta_1(t) r_{m,t} + \varepsilon_t.$$

The local linear estimators $\hat{\beta}_0(\cdot)$ and $\hat{\beta}_1(\cdot)$ are computed. The estimated curves $\hat{\beta}_0(\cdot)$ (left panel) and $\hat{\beta}_1(\cdot)$ (right panel) are depicted in Figure 2.

It is evident from Figure 2 that both the trend $\beta_0(\cdot)$ (left) and the slope $\beta_1(\cdot)$ (right) do change over time and the slope $\beta_1(\cdot)$ is almost above 1 except the period of the trading days from 141 to 171 (The corresponding calendar days are from July 26, 2000 to September 7, 2000). For the trend function $\beta_0(\cdot)$, reflecting the dynamic change for MSFT itself, although it is up and down during this period, the overall trend increases slightly for the first three quarters. But the trend decreases dramatically for the last quarter. In contrast, the beta function $\beta_1(\cdot)$ keeps a constant (around 1.17) during the first 111 trading days of the year (June 13, 2000) and it decreases afterwards until the 161st trading day (August 23, 2000) and finally, it increases afterwards (to the end of the year). Therefore, it concludes that MFST is a stock that was more volatile than the U.S. blue chip market as a whole.

Finally, to support the aforementioned conclusions statistically, I consider the testing null hypothesis $H_0 : \beta_0(\cdot) = \alpha_0$ and $\beta_1(\cdot) = \alpha_1$, the testing procedure described in Section 2.3 is used with the bootstrap sampling 1000 times. As a result, the p -value is less than 0.001. Therefore, this test result further supports the finding that both the trend function $\beta_0(\cdot)$ and the beta function $\beta_1(\cdot)$ do change over time.

4 Concluding Remarks

In this paper, I have presented a selective overview on the recent developments on the functional coefficient models with particular applications in economics and finance. Indeed, there are numerous papers addressing various types of functional coefficient models in the past two decades. My citation in this paper is not exhaustive due to space limitations. In addition to the applications to economics and finance, the functional coefficient models have also been used in other subjects in statistics such as time series, longitudinal data analysis and survival analysis (see Fan and Zhang, 2008). Finally, Cai and Li (2009) surveyed some recent developments in non-parametric econometric models, including some applications of functional coefficient models in economics, while Cai and Hong (2009) gave a review on the recent developments of nonparametric estimation and testing of financial econometric models, including functional coefficient diffusion models

that are frequently used to describe the dynamics of an underlying process including stock and bond prices and various interest rates.

References

- AKDENIZ, L., ALTAY-SALIH, A., CANER, M. (2003). Time-varying betas help in asset pricing: the threshold CAPM. *Studies in Nonlinear Dynamics and Econometrics*, **6**, Article 1.
- ANDREWS, D.W.K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica*, **61**, 821-856.
- ANG, A., LIU, J. (2004). How to discount cashflows with time-varying expected return. *The Journal of Finance*, **59**, 2745-2783.
- BAI, J. (1997). Estimation of a change point in multiple regression models. *Review of Economic and Statistics*, **79**, 551-563.
- BAI, J., PERRON, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, **66**, 47-78.
- BAI, J., PERRON, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, **18**, 1-22.
- BANSAL, R., HSIEH, D.A., VISWANATHAN, S. (1993). A new approach to international arbitrage pricing. *The Journal of Finance*, **48**, 1719-1747.
- BANSAL, R., VISWANATHAN, S. (1993). No arbitrage and arbitrage pricing: A new approach. *The Journal of Finance*, **47**, 1231-1262.
- CAI, Z. (2002). A two-stage approach to additive time series models. *Statistica Neerlandica*, **56**, 415-433.
- CAI, Z. (2003). Local quasi-likelihood approach to varying-coefficient discrete-valued time series models. *Journal of Nonparametric Statistics*, **15**, 693-711.
- CAI, Z. (2007). Trending time varying coefficient time series models with serially correlated errors. *Journal of Econometrics*, **136**, 163-188.
- CAI, Z., DAS, M., XIONG, H., WU, X. (2006). Functional coefficient instrumental variables models. *Journal of Econometrics*, **133**, 207-241
- CAI, Z., FAN, J., LI, R. (2000). Efficient estimation and inferences for varying-coefficient models. *Journal of American Statistical Association*, **95**, 888-902.
- CAI, Z., FAN, J., YAO, Q. (2000). Functional-coefficient regression models for nonlinear time series. *Journal of American Statistical Association*, **95** 941-956.
- CAI, Z., HONG, Y. (2009). Some recent developments in nonparametric finance. Forthcoming in *Advances in Econometrics*.

- CAI, Z., HSIAO, C., ZHU, Y. (2009). Trending panel models. *Working paper*, Department of Mathematics and Statistics, University of North Carolina at Charlotte.
- CAI, Z., KUAN, C.-M., SUN, L. (2009). Nonparametric pricing kernel models. *Working paper*, Department of Mathematics and Statistics, University of North Carolina at Charlotte.
- CAI, Z., LI, Q. (2009). Recent developments in nonparametric econometrics. Forthcoming in *Advances in Econometrics*.
- CAI, Z., OULD-SAID, E. (2003). Local robust regression estimation for time series. *Statistics and Probability Letters*, **65**, 433-449.
- CAI, Z., TIWARI, R.C. (2000). Application of a local linear autoregressive model to BOD time series. *Environmetrics*, **11**, 341-350.
- CAI, Z., WANG, Y. (2009). Instability of predictability of asset returns. *Working paper*, Department of Mathematics and Statistics, University of North Carolina at Charlotte.
- CAI, Z., XU, X. (2008). Nonparametric quantile estimations for dynamic smooth coefficient models. *Journal of the American Statistical Association*, **103**, 1596-1608.
- CANER, M., HANSEN, B.E. (2001). Threshold autoregressions with a near unit root. *Econometrica*, **69**, 1555-1597.
- CANER, M., HANSEN, B.E. (2004). Instrumental variable estimation of a threshold model. *Econometric Theory*, **20**, 813-843.
- CARD, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica*, **69**, 1127-1160.
- CARDOT, H., SARDA, P. (2008). Varying-coefficient functional linear regression models. *Communications in Statistics: Theory and Methods*, **37**, 3186-3203.
- CHANG, Y., MARTINEZ-CHOMBO, E. (2003). Electricity demand analysis using cointegration and error-correction models with time varying parameters: the Mexican case. *Working paper*, Department of Economics, Rice University.
- CHO, Y.-H., ENGLE, R.F. (2000). Time-varying betas and asymmetric effects of news: Empirical analysis of blue chip stocks. *Working Paper*, Department of Finance, New York University.
- CLEVELAND, W.S., GROSSE, E., SHYU, W.M. (1991). Local regression models. In *Statistical Models in S*, (Eds, Chambers, J.M. and Hastie, T.J.), 309-376. Wadsworth & Brooks, Pacific Grove.
- COCHRANE, J.H. (1996). A cross-sectional test of an investment-based asset pricing model. *Journal of Political Economy*, **104**, 572-621.

- CUI, H., HE, X., ZHU, L. (2002). On regression estimators with de-noised variables. *Statistica Sinica*, **12**, 1191-1205.
- DAHLHAUS, R. (1997). Fitting time series models to nonstationary processes. *The Annals of Statistics*, **25**, 1-37.
- DAHLHAUS, R., SUBBA RAO, S. (2006). Statistical inference for locally stationary ARCH models. *The Annals of Statistics*, **34**, 1075-1114.
- DAS, M. (2005). Instrumental variables estimators for nonparametric models with discrete endogenous regressors. *Journal of Econometrics*, **124**, 335-361.
- DAVIS, R.A., LEE, T.C.M., RODRIGUEZ-YAM, G.A. (2006). Structural break estimation for nonstationary time series models. *Journal of American Statistical Association*, **101**, 223-238.
- FAN, J., GIJBELS, I. (1996). *Local Polynomial Modelling and Its Applications*. Chapman and Hall, London
- FAN, J., JIANG, J., ZHANG, C., ZHOU, Z. (2003). Time-dependent diffusion models for term structure dynamics and the stock price volatility. *Statistica Sinica*, **13**, 965-992.
- FAN, J., ZHANG, C. (2003). A re-examination of diffusion estimators with applications to financial model validation. *Journal of the American Statistical Association*, **98**, 118-134.
- FAN, J., ZHANG, C., ZHANG, J. (2001). Generalized likelihood ratio statistics and Wilks phenomenon. *The Annals of Statistics*, **29**, 153-193.
- FAN, J., ZHANG, W. (2008). Statistical methods with varying coefficient models. *Statistics and Its Interface*, **1**, 179-195.
- FERSON, W.E., HARVEY, C.R. (1998). Fundamental determinants of national equity market returns: A perspective on conditional asset pricing. *Journal of Banking and Finance*, **21**, 1625-1665.
- FERSON, W.E., HARVEY, C.R. (1999). Conditional variables and the cross section of stock return. *The Journal of Finance*, **54**, 1325-1360.
- GHYSELS, E. (1998). On stable factor structures in the pricing of risk: do time varying betas help or hurt. *Journal of Finance*, **53**, 549-574.
- HANSEN, B.E. (1996). Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica*, **64**, 413-430.
- HANSEN, B.E. (2000). Sample splitting and threshold estimation. *Econometrica*, **68**, 575-605.
- HARVEY, C.R. (1989). Time-varying conditional covariances in tests of asset pricing models. *Journal of Financial Economics*, **24**, 289-317.
- HASTIE, T.J., TIBSHIRANI, R.J. (1993). Varying coefficient models (with discussion). *Journal of the Royal Statistical Society, Series B*, **44**, 646-685.

- HONG, Y., LEE, T.-H. (2003). Inference and forecast of exchange rates via generalized spectrum and nonlinear time series models. *Review of Economics and Statistics*, **85**, 1048-1062.
- HUANG, J.Z., SHEN, H. (2004). Functional coefficient regression models for nonlinear time series: A polynomial spline approach. *Scandinavian Journal of Statistics*, **31**, 515-534.
- JAGANNATHAN, R., WANG, Z. (1996). The conditional CAPM and the cross-section of expected returns. *Journal of Finance*, **51**, 3-53.
- JAGANNATHAN, R., WANG, Z. (2002). Empirical evaluation of asset pricing models: A comparison of the SDF and beta methods. *The Journal of Finance*, **57**, 2337-2367.
- LI, Q., HUANG, C., LI, D., FU, T. (2002). Semiparametric smooth coefficient models. *Journal of Business and Economic Statistics*, **20**, 412-422.
- MÜLLER, H.-G., SEN, R., STADTMÜLLER, U. (2007). Functional data analysis for volatility. *Working paper*, Department of Statistics, University of California at Davis.
- NEWKEY, W.K., POWELL, J.L. (2003). Nonparametric instrumental variables estimation. *Econometrica*, **71**, 1565-1578.
- PHILLIPS, P.C.B. (2001). Trending time series and macroeconomic activity: some present and future challenges. *Journal of Econometrics*, **100**, 21-27.
- RAMSAY, J.O., SILVERMAN, B.W. (1997). *Functional Data Analysis*. Springer, New York.
- REYES, M.G. (1999). Size time-varying beta and conditional heteroscedasticity in UK stock return. *Review of Financial Economics*, **8**, 1-10.
- ROBINSON, P.M. (1989). Nonparametric estimation of time-varying parameters. In *Statistical Analysis and Forecasting of Economic Structural Change (Eds, Hackl, P. and Westland, A.H.)*, Springer-Verlag, Berlin, 253-164.
- SCHULTZ, T.P. (1997). *Human Capital, Schooling and Health*, IUSSP, XXIII, General Population Conference, Yale University.
- ŞENTÜRK, D., MÜLLER, H.G. (2005). Covariate adjusted correlation analysis via varying coefficient models. *Scandinavian Journal of Statistics*, **32**, 365-383.
- ŞENTÜRK, D., MÜLLER, H.G. (2008). Generalized varying-coefficient models for longitudinal data. *Biometrika*, **95**, 653-666.
- TONG, H. (1990). *Nonlinear Time Series: A Dynamical System Approach*. Oxford University Press, Oxford, UK.
- TSAY, R.S. (2005). *Analysis of Financial Time Series*. Wiley, New York.
- WANG, K.Q. (2002). Nonparametric tests of conditional mean-variance efficiency of a benchmark portfolio. *Journal of Empirical Finance*, **9**, 133-169.

- WANG, K.Q. (2003). Asset pricing with conditioning information: a new test. *Journal of Finance*, **58**, 161-196.
- WEI, Y., HE, X. (2006). Conditional growth charts (with discussion). *The Annals of Statistics*, **34**, 2069-2097.
- WEI, Y., PERE, A., KOENKER, R., HE, X. (2006). Quantile regression methods for reference growth charts. *Statistics in Medicine*, **25**, 1369-1382.
- YOU, J., JIANG, J. (2007). Inferences for varying-coefficient partially linear models with serially correlated errors. In *Advances in Statistical Modeling and Inference: Essays in Honor of Kjell A. Doksum (Ed. Nair, V.)*. *Series in Biostatistics*, **3**, 175-195. World Scientific Publishing Co. Pte. Ltd., Singapore.