

TRENDING TIME-VARYING COEFFICIENT MARKET MODELS

by

CHONGSHAN ZHANG

(Under the Direction of Xiangrong Yin)

ABSTRACT

The market model, also known as single factor model or the β -representation in Capital Asset Pricing Model (CAPM) context, is a purely statistical model used to explain the behavior of asset returns. In this paper we study time-varying coefficient models with time trend function to characterize non-linear, non-stationary and trending phenomenon in time series. Compared with the Nadaraya-Watson method and the local linear approach, the general local polynomial approach is developed to estimate the time trend and coefficient functions. Bandwidth is selected based on the nonparametric version of the Akaike information criterion (AIC). In our general local polynomial method, we plot the derivatives of the parameters vs. time. These graphs provide a useful way to estimate what local polynomial orders we should use for data analysis. Finally, we conduct some Monte Carlo experiments to examine the finite sample performances of the proposed modeling procedure, and an empirical example is discussed by checking the regression coefficients between individual stocks and S&P 500 index (market portfolio).

KEY WORDS: Bandwidth selection; Functional coefficient models; Nadaraya-Watson estimation; Local linear estimate; local polynomial estimate, Nonlinearity; Nonstationarity; Stationarity; Time series errors.

TRENDING TIME-VARYING COEFFICIENT MARKET MODELS

by

CHONGSHAN ZHANG

Bachelor of Science, Beijing University, People's Republic of China, 1988

Master of Science, Beijing University, People's Republic of China, 1991

Master of Science, University of Georgia, 2005

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2006

© 2006

Chongshan Zhang

All Rights Reserved

TRENDING TIME-VARYING COEFFICIENT MARKET MODELS

by

CHONGSHAN ZHANG

Major Professor: Xiangrong Yin

Committee: Lynne Seymour
Cheolwoo Park

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2006

DEDICATION

This is dedicated to my loving wife, Ruoyan Zhang.

ACKNOWLEDGEMENTS

I am very thankful for the members of my thesis committee. I want to express my thanks to Dr. Seymour and Dr. Park for their time reviewing my thesis and their insightful advises. Many thanks to my advisor, Dr. Yin, to whom I owe the most overwhelming debt of gratitude for his support and his invaluable guidance for my thesis research.

I thank Zongwu Cai very much, who has given me tremendous advice and help in my academic endeavor.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	v
LIST OF FIGURES	viii
CHAPTER	
1 INTRODUCTION	1
2 SOME BACKGROUND	5
2.1 Basic knowledge of prices and returns	5
2.2 Market model	6
2.3 Stationarity and linearity of time series	8
3 MODEL PROCESSING	9
3.1 Local linear and constant estimation	9
3.2 Kernel and bandwidth selection	11
3.3 Estimation of variance	12
3.4 Checking order- k in local polynomial method	13
4 SIMULATION RESULTS	14
4.1 Simulation example 1 – A constant model	15
4.2 Simulation example 2 – A linear model	21
4.3 Simulation example 3 – A parabolic model	27
4.4 A real example	34
5 CONCLUSIONS	39
REFERENCES	40

LIST OF TABLES

Table 4.1 The mean and standard deviation of the 100 MADE values for the constant model	16
Table 4.2 The mean and standard deviation of the 100 MADE values for the linear model	21
Table 4.3 The mean and standard deviation of the 100 MADE values for the parabolic model	27

LIST OF FIGURES

	Page
Fig. 4.1 The Nadarya-Watson estimator (solid line) of β_1 with 95% confidence intervals (dotted line) for $\beta_1 = 1$	17
Fig. 4.2 The local linear estimator (solid line) of β_1 with 95% confidence intervals (dotted line) for $\beta_1 = 1$	17
Fig.4.3 The local parabolic estimator (solid line) of β_1 with 95% confidence intervals (dotted line) for $\beta_1 = 1$	18
Fig. 4.4 The local linear estimator (solid line) of β_1' with 95% confidence intervals (dashed line) for $\beta_1 = 1$	18
Fig. 4.5 The local parabolic estimator (solid line) of β_1' with 95% confidence intervals (dashed line) for $\beta_1 = 1$	19
Fig. 4.6 The local linear estimator (solid line) of β_1'' with 95% confidence intervals (dashed line) for $\beta_1 = 1$	19
Fig. 4.7 Y vs X for $\beta_1 = 1$	20
Fig. 4.8 Y (solid line) and estimated Y (dotted line) for $\beta_1 = 1$	20
Fig.4.9 The Nadarya-Watson estimator (solid line) of β_1 with 95% confidence intervals (dotted line) for $\beta_1 = 2t$	23
Fig. 4.10 The local linear estimator (solid line) of β_1 with 95% confidence intervals (dotted line) for $\beta_1 = 2t$	23

Fig.4.11 The local parabolic estimator (solid line) of β_1 with 95% confidence intervals (dotted line) for $\beta_1 = 2t$	24
Fig. 4.12 The local linear estimator (solid line) of β_1' with 95% confidence intervals (dashed line) for $\beta_1 = 2t$	24
Fig. 4.13 The local parabolic estimator (solid line) of β_1' with 95% confidence intervals (dashed line) for $\beta_1 = 2t$	25
Fig. 4.14 The local linear estimator (solid line) of β_1'' with 95% confidence intervals (dashed line) for $\beta_1 = 2t$	25
Fig. 4.15 Y vs X for $\beta_1 = 2t$	26
Fig. 4.16 Y (solid line) and estimated Y (dotted line) for $\beta_1 = 2t$..	26
Fig.4.17 The Nadarya-Watson estimator (solid line) of β_1 with 95% confidence intervals (dotted line) for $\beta_1 = 8(t - 0.5)^2$	29
Fig. 4.18 The local linear estimator (solid line) of β_1 with 95% confidence intervals (dotted line) for $\beta_1 = 8(t - 0.5)^2$	29
Fig.4.19 The local parabolic estimator (solid line) of β_1 with 95% confidence intervals (dotted line) for $\beta_1 = 8(t - 0.5)^2$	30
Fig. 4.20 The local linear estimator (solid line) of β_1' with 95% confidence intervals (dashed line) for $\beta_1 = 8(t - 0.5)^2$..	30
Fig. 4.21 The local parabolic estimator (solid line) of β_1' with 95% confidence intervals (dashed line) for $\beta_1 = 8(t - 0.5)^2$..	31

Fig. 4.22 The local linear estimator (solid line) of β_1'' with 95% confidence intervals
(dashed line) for $\beta_1 = 8(t - 0.5)^2$ 31

Fig. 4.23 Y vs X for $\beta_1 = 8(t - 0.5)^2$ 32

Fig. 4.24 Y (solid line) and estimated Y (dotted line) for $\beta_1 = 8(t - 0.5)^2$ 32

Fig. 4.25 $\text{var}(\hat{\beta}_0(t))_{\text{sample}}$ (dash line), and the mean of $\text{var}(\hat{\beta}_0(t))_{\text{ls}}$ (solid line) and its 95%
confidence intervals (dotted lines), maximum and minimum (dash-dot lines).33

Fig. 4.26 $\text{var}(\hat{\beta}_0(t))_{\text{sample}}$ (dash line), and the mean of $\text{var}(\hat{\beta}_0(t))_{\text{ls}}$ (solid line) and its 95%
confidence intervals (dotted lines), maximum and minimum (dash-dot lines).33

Fig. 4.27 The time series plot of LOW stock return35

Fig. 4.28 The time series plot of S&P 500 index return35

Fig. 4.29 The local polynomial estimator (solid line) of β_1 with 95% confidence intervals
(dashed lines) and least-square estimator (dotted line).36

Fig. 4.30 The local polynomial estimator (solid line) of β_1' with 95% confidence
intervals (dashed lines)36

Fig. 4.31 The local polynomial estimator (solid line) of β_1'' with 95% confidence
intervals (dashed lines)37

Fig.4.32 The residual plot of LOW stock return.....37

CHAPTER 1

INTRODUCTION

The analysis of nonlinear and nonstationary time series particularly with time trend has been very active during the last two decades, because most of time series, in particular, observed from economics and business, are nonlinear or nonstationary or trending (Granger and TerÄasvirta 1993; Franses 1996, 1998; Phillips 2001; Tsay 2002). For example, 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. Another example is the term structure of interest rates in which the time evolution of the relationship between interest rates with different maturities is investigated. For more examples in macroeconomic activity, see the survey paper by Phillips (2001). To characterize these phenomena, during the recent years there have been proposed several nonlinear and nonstationary parametric, semi-parametric and nonparametric time series models with/without time trend in the econometrics, finance and statistics literature. For more detailed discussions on this aspect, see the papers by Park and Phillips (1999, 2002) and Chang, Park and Phillips (1999), Phillips (2001), and the books by Granger and TerÄasvirta (1993), Franses (1996, 1998) and Tsay (2002) and the references therein. 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.

There are several ways to explore the nonlinearity and nonstationarity and one of the most attractive models is the time-varying coefficient time series models with time.

Recently, there are some new developments. Roussas (1989) studied the following fixed design time series model without exogenous variables

$$Y_i = \beta_0(t_i) + \varepsilon_i \quad . \quad (1.1)$$

He considered a linear estimator and obtained consistency of the proposed estimator. Indeed, this model deals with only the time trend function $\beta_0(t)$.

To allow for the presence of exogenous (explanatory) variables (covariates) $X_i = (X_{i1}, \dots, X_{id})$ which might be important in an econometric context or other fields, we consider the following time-varying coefficient time series models with time trend

$$Y_i = \beta_0(t_i) + \sum_{j=1}^d \beta_j(t_i) X_{ij} + \varepsilon_i = \mathbf{X}_i^T \boldsymbol{\beta}(t_i) + \varepsilon_i, \quad (1.2)$$

where $\mathbf{X}_i = (X_{i0}, X_{i1}, \dots, X_{id})^T$ with $X_{i0} = 1$, $\boldsymbol{\beta}(t) = (\beta_0(t), \beta_1(t), \dots, \beta_d(t))$, and $E(\varepsilon_i | \mathbf{X}_i) = 0$. To include the heteroscedasticity in the model, $E(\varepsilon_i^2 | \mathbf{X}_i)$ is allowed to be a function of X_i . This is particularly appealing in economics and finance. Finally, $\{(\varepsilon_i, \mathbf{X}_i)\}$ is assumed to be stationary. Clearly, both models (1.1) and (1.2) include the deterministic time trend function $\beta_0(t)$, the time series $\{Y_i\}$ is not stationary, and (1.2) covers (1.1) as a special case.

Robinson (1989) studied model (1.2) under the assumptions that the time series $\{X_i\}$ is stationary α -mixing and the errors $\{\varepsilon_i\}$ are iid and independent of X_i and developed the Nadaraya-Watson (local constant) method to estimate the coefficient functions and studied the asymptotic properties of the proposed estimator. For more discussions on this point, see the paper by Robinson (1989) and Section 3.1. Later,

Robinson (1991) considered a more general model and relaxed the iid assumption on $\{\varepsilon_i\}$ to α -mixing.

Cai (2006) considered model (1.2) by using the local linear estimation to estimate the coefficient functions and made a comparison with the Nadaraya-Watson method. It is showed that the estimators based on both the local linear fitting and the Nadaraya-Watson method share the exact same asymptotic behavior at the interior points but not at boundaries. Also, Cai's results showed that the consistency of the proposed estimators can be obtained without specifying the error distribution and the asymptotic variance of the proposed estimator depends on not only the variance of the error but also the autocorrelations. Further, to choose the data-driven fashioned bandwidth, Cai proposed a new bandwidth selector based on the nonparametric version of the AIC.

However, some recent studies show that the beta-coefficients might vary over time non-linearly for many applications in economics and finance; see Ghysels (1998) and Akdeniz *et.al.* (2003). Motivated by this, we extend the local linear estimation to local polynomial estimation. We do simulation for constant model, linear model and parabolic model to demonstrate whether local polynomial estimation is necessary by checking the significance of high order of derivation of $\beta_j(t)$, and whether local polynomial estimation is superior to local constant estimation and local linear by checking their bias. Finally, we apply our local polynomial estimation method to market data to show its usefulness.

The thesis is organized as follows. Chapter 2 introduces some basic background about stock return and market model. Chapter 3 is devoted to the presentation of the estimation

methods. Chapter 4 reports some results from numerical simulations and application of market model. Finally, Chapter 5 is the conclusions and a brief discussion.

CHAPTER 2

SOME BACKGROUND

2.1 Basic knowledge of prices and returns

2.1.1 Where to get price data

Any empirical analysis of dynamics of asset prices through time requires price data. There are many sources of data including web sites, commercial vendors, university research centers, and financial markets. Here are some of them, list below:

CRSP: www.crsp.com (US stocks)

Commodity System Inc: www.csidata.com (Futures)

Datastream: www.datastream.com/product/has/ (Stocks, bonds, currencies, etc.)

Trades and Quotes DB: www.nyse.com/marketinfo (US stocks)

US Federal Reserve: www.federalreserve.gov/resleases (Currencies, etc.)

Yahoo! (free): <http://biz.yahoo.com/r/> (Stock, many countries)

Before abstracting the price data, one also should plan what the frequency of data is, such as daily, weekly, monthly, *etc.*, how many periods of data needed for analysis, and what kind of price needed.

2.1.2 Frequency of observations

It depends on the researcher's interests. The price interval between prices should be sufficient to ensure that trade occurs in most intervals and it is preferable that volume of

trade is substantial. Also, it is important to distinguish the price data indexed by transaction counts from the data indexed by time of associated transactions.

2.1.3 Definitions of Returns

The statistical inference on asset prices is complicated because of non-stationary behavior of asset price (upward and downward movements). One can transform asset prices into returns, which empirically display more stationary behavior. Also, returns are score-free, which depend on the ratio of the change of the price, but not on the absolute values of the price.

Return of a financial asset (stock) with prices P_t at date t that produces no dividends over the period $(t, t+H)$ is defined as:

$$r(t, t+H) = \frac{P_{t+H} - P_t}{P_t} \quad (2.1)$$

Very often, we will investigate returns at a fixed unitary horizon. In this case $H=1$ and return is defined as:

$$r(t, t+1) = \frac{P_{t+1} - P_t}{P_t} \quad (2.2)$$

Returns in (2.1) and (2.2) are sometimes called the simple net return. In this paper, we use simple net return.

2.2 Market model

The market model is used to explain the behavior of asset returns. It is also known as Shapoe's single index model (SIM) or the single factor model or the β -representation in

Capital Asset pricing model (CAPM) context. The market model has the form of the simple bivariate linear regression model:

$$r_i = \alpha(t) + \beta(t)r_{\text{MKT},i} + \varepsilon_i, \quad (2.3)$$

where r_i is the continuously compounded return on some asset between time period t and $t+1$ and $r_{\text{MKT},i}$ is the continuously compounded return on a market index portfolio or an individual stock return. Comparing the notations with (1.2), r is Y , r_{MKT} is X and α is β_0 .

The intuition behind the market model is as follows. The market index $r_{\text{MKT},i}$ captures macro or market-wide systematic risk factor. This type of risk is called systematic risk or market risk, which can not be eliminated in a well diversified portfolio. The random error term ε_i captures micro or firm-specific risk factors that affect an individual asset return and that are not related to macro events. This type of risk is called firm specific risk, idiosyncratic risk or non-marked risk. This type of risk can be eliminated in a well-diversified portfolio.

The market model (single factor model) can be extended to capture multiple factors:

$$r_i = \alpha(t) + \sum_{j=1}^d \beta_j(t)f_{j,i} + \varepsilon_i, \quad (2.4)$$

where $f_{j,i}$ denotes the j th systematic factor, $\beta_j(t)$ denotes asset's loading on the j th factor, and ε_i denotes the random component independent of all the systematic factors. This is called multiple factor model. And when $\beta(t) = 0$ in (2.3), the model (2.3) becomes

$r_i = \alpha(t) + \varepsilon_i$, which is a special case of the market model, called constant expected return (CER) model.

The market model is heavily used in empirical finance. It is used to estimate expected returns, variance and covariance that are needed to implement portfolio theory. It is used as a model to explain normal or usual rate of return on an asset for the use in event studies.

2.3 Stationarity and linearity of time series

A time series $\{r(t)\}$ is said to be strictly stationary if the joint distribution of $(r(t_1), r(t_2), \dots, r(t_k))$ is invariant under time shift. This is a very strong condition that is hard to verify empirically. A weaker version of stationarity is often assumed. A time series $\{r(t)\}$ is weakly stationary if the mean of $r(t)$ and the covariance between $r(t)$ and $r(t-l)$ are time-invariant, where l is an arbitrary integer.

A time series is said to be linear if it can be written as

$$r(t) = \mu + \sum_{i=0}^{\infty} \varphi_i a(t-i) \quad (2.5)$$

where μ is the mean of $r(t)$, φ_i are real number with $\varphi_0 = 1$, and $\{a(t)\}$ is a white noise series.

For the market model (2.3), $\alpha(t)$ and $\beta(t)$ are general function of t , and the return time series $\{r(t)\}$ could have any pattern, therefore, $\{r(t)\}$ usually do not have to satisfy the condition as stationary and/or linear time series, thus should be considered as non-stationary and non-linear time series.

CHAPTER 3

MODEL PROCESSING

3.1 Local polynomial estimation

For estimating $\{\beta_j(\cdot)\}$ in (1.2), a local linear method is previously employed by Cai (2006), however a local polynomial method is generally more applicable. Assuming $\{\beta_j(\cdot)\}$ have a continuous $k+1$ order derivative, then $\{\beta_j(\cdot)\}$ can be approximated by a linear function near and at any fixed time point t as follows:

$$\beta_j(t_i) \approx a_{0j} + a_{1j}(t_i - t) + a_{2j}(t_i - t)^2 + \cdots + a_{kj}(t_i - t)^k \quad (3.1)$$

where \approx denotes the Taylor approximation and $a_{0j} = \beta_j(t)$, $a_{lj} = \beta_j^{(l)}(t)$ $l = 1, \dots, k$.

Hence (1.2) is approximated by

$$Y_i \approx \mathbf{Z}_i^T \boldsymbol{\theta} + \varepsilon_i, \quad (3.2)$$

where $\mathbf{Z}_i = (\mathbf{X}_i^T, \mathbf{X}_i^T(t_i - t), \mathbf{X}_i^T(t_i - t)^2, \dots, \mathbf{X}_i^T(t_i - t)^k)^T$,

and $\boldsymbol{\theta} = \boldsymbol{\theta}(t) = (\boldsymbol{\beta}^T(t), \boldsymbol{\beta}^T(t), \boldsymbol{\beta}^T(t), \dots, \boldsymbol{\beta}^{(k)T}(t))$.

Therefore the locally weighted least square is

$$\sum_{i=1}^n \{Y_i - \mathbf{Z}_i^T \boldsymbol{\theta}\}^2 K_h(t_i - t), \quad (3.3)$$

where $K_h(u) = K(u/h)/h$, $K(\cdot)$ is the kernel function and $h = h_n > 0$ is the bandwidth satisfying $h \rightarrow 0$ and $nh \rightarrow \infty$ when $n \rightarrow \infty$ which controls the amount of smoothing used in the estimation.

By minimizing (3.3) with respect $\boldsymbol{\theta}$, we obtain the local polynomial estimate of $\beta_j(t)$, denoted by $\hat{\beta}_j(t)$, which is the first $(d+1)$ elements of $\hat{\boldsymbol{\theta}}$, and the l -order derivatives of $\beta_j(t)$, denoted by $\hat{\beta}_j^{(k)}(t)$, which is the $(l+1)^{\text{th}}$ $(d+1)$ elements of $\hat{\boldsymbol{\theta}}$. If $k = 0$, then it is local constant estimation (Robinson, 1989), if $k=1$, then it is local linear estimation (Cai, 2006).

Note that the local polynomial estimator can be viewed as the least square estimator of the following working linear model

$$K_h^{1/2}(t_i - t)Y_i = K_h^{1/2}(t_i - t)\mathbf{Z}_i^T \boldsymbol{\theta}_1 + K_h^{1/2}(t_i - t)\mathbf{Z}_i^T (t_i - t)\boldsymbol{\theta}_2 + \mu_i \quad (3.4)$$

Therefore, the computational implementation can be easily carried out by any standard statistical software.

Note that many other nonparametric smoothing methods can be used here. The locally weighted least square method or the Nadaraya-Watson approach is just one of the choices. There is a vast literature in theory and empirical study on the comparison of different nonparametric smoothing methods (see Härdle 1990; Fan and Gijbels 1996; Pagan and Ullah 1999).

The restriction to the locally weighted least square method suggests that the normality is at least being considered as a baseline. However, when the non-normality is clearly present, a robust approach would be considered. Cai and Ould-Said (2003) considered this aspect in nonparametric regression estimation for time series. If some of \mathbf{X}_i 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. For example, we can apply the two-stage local linear technique proposed by Cai, Das, Xiong and Wu (2006).

3.2 Kernel and bandwidth selection

Generally choosing kernels is not an issue. One can choose kernels as the usual way, such as Gaussian kernel, Biweight kernel and Epanechnikov kernel in non-parameter estimation. We use Epanechnikov kernel, $K(u) = 0.75(1 - u^2)I(|u| \leq 1)$, as it is easy for computation.

The bandwidth is selected using the same method described in Cai (2006). The selection of the bandwidth is similar to the model selection for linear models. The selection procedure can be regarded as a nonparametric version of the AIC to be attentive to the structure of time series data and the over-fitting or under-fitting tendency. The basic idea is described as follows.

For given observed values $\{Y_t\}_{t=1}^n$, the fitted values can be expressed as $\hat{\mathbf{Y}} = \mathbf{H}_\lambda \mathbf{Y}$, where $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)^T$ and \mathbf{H}_λ is called the $n \times n$ smoother (or hat) matrix associated with the smoothing parameter λ . Motivated by the classical AIC for linear models under the likelihood setting

$$-2 (\text{maximized log likelihood}) + 2 (\text{number of estimated parameters}), \quad (3.5)$$

here we propose the following nonparametric version of AIC to select h by minimizing

$$\text{AIC}(\lambda) = \log\{SSE\} + \psi(n_\lambda, n); \quad (3.6)$$

where $SSE = \sum_{t=1}^n (Y_t - \hat{Y}_t)^2$, regarded as the replacement of the first term in (3.5) and n_λ

is the trace of the hat matrix \mathbf{H}_λ , called to be the effective number of parameters or the nonparametric version of degrees of freedom by Hastie and Tibshirani (1990, Section 3.5) for nonparametric models. Particularly, we choose $\psi(n_\lambda, n)$ to be the form of the bias-corrected version of the AIC, due to Hurvich and Tsai (1989),

$$\psi(n_\lambda, n) = 2(n_\lambda + 1)/(n - n_\lambda - 2). \quad (3.7)$$

Alternatively, one might use some existing methods in the literature although they may require more computing, for example, see the papers by Robinson (1989), Yao and Tong (1994), View(1994), and Tschernig and Yang (2000).

3.3 Estimation of variance

Cai (2006) suggested method to estimate the asymptotic variance of $\hat{\beta}_j(t)$ to construct pointwise confidence intervals, where the variance of $\hat{\beta}_j(t)$ is independent of t . In other words, the variance of $\hat{\beta}_j(t)$ is a constant for all t . However, it is not always true. Near the boundary of the time series, or when $\beta_j(t)$ changes quickly, its variance usually is bigger than other time period.

Practically, we construct the confidence intervals of $\beta_j(t)$ using the sample variance. In the simulation, for each data set $\{X_i, Y_i\}$ generated randomly, we can calculate $\hat{\beta}_j(t)$. Same processes are repeated, for example, m times. Then at any t , we can obtain the sample mean and variance of $\hat{\beta}_j(t)$, denoted as $\overline{\hat{\beta}_j(t)}$ and $\text{var}(\hat{\beta}_j(t))_{\text{sample}}$.

In practice, however, it is desirable to estimate the variance of $\beta_j(t)$ with one data set, because we can not replicate more data sets under the exact same condition. To solve this problem, we notice the linear model (3.4) provides a least square estimator of the variance which can be calculated easily, denoted as $\text{var}(\hat{\beta}_j(t))_{\text{ls}}$. We check its performance using the simulation data. At any t , there is one value of $\text{var}(\hat{\beta}_j(t))_{\text{sample}}$ and m values of $\text{var}(\hat{\beta}_j(t))_{\text{ls}}$. If $\text{var}(\hat{\beta}_j(t))_{\text{sample}}$ is in the confidence interval of m values of $\text{var}(\hat{\beta}_j(t))_{\text{ls}}$, we use the $\text{var}(\hat{\beta}_j(t))_{\text{ls}}$ as a replacement for an estimate of $\text{var}(\hat{\beta}_j(t))_{\text{sample}}$.

3.4 Checking order- k in local polynomial method

In real data, since we do not know that one should use whether local constant estimation, local linear estimation, or higher order local polynomials. Of course, if the bandwidth is small enough, then local constant estimation may always be satisfied in theory. However practically, data size may prohibit this (together with the result in 3.2 about the selection of bandwidth). To decide the order of the local polynomials, say, k , one can graphically check the plots of $\beta^{(l)}(t)$ vs. t . If the plot shows a flat line, i.e. $\beta^{(l)}(t)$ is a constant, then $k = l$ should be good enough.

Though we suggest a theory of general local polynomial estimation method, in this thesis we only evaluate the case when $k = 2$ in our simulation, and make comparison with $k = 0$ and $k = 1$. Practically we believe $k = 2$ is perhaps good enough in most cases.

CHAPTER 4

SIMULATIONAL RESULTS

To illustrate our modeling procedure, we consider three simulated examples and further investigate the empirical relationship between the Microsoft stock return and the S&P 500 return. In both simulations and real data application, the Epanechnikov kernel $K(u) = 0.75(1-u^2)I(|u| \leq 1)$ is used and optimal bandwidth h_{opt} is selected as follows. For a predetermined sequence of h 's from a wide range, say from 0.1 to 0.5 with an increment 0.05, based on the AIC bandwidth selector described in Section 3.2, we compute $\text{AIC}(h)$ for each h and choose h_{opt} to minimize $\text{AIC}(h)$.

For the simulated examples, the performance of the proposed estimators is evaluated by the mean absolute deviation error (MADE)

$$e_j = n_0^{-1} \sum_{i=1}^{n_0} |\hat{\beta}_j(t_i) - \beta_j(t_i)| \quad (4.1)$$

The variance of the $\hat{\beta}_j(t)$ is calculated by the 100 independent simulations. We check the performance of the variance of the $\hat{\beta}_j(t)$ determined by the least square method when we analyze the parabolic example.

In the simulated examples, we will illustrate the finite sample performances of the local polynomial estimators by comparing them with the local constant and local linear estimator, and test with a simulated example of the time-varying coefficient time series models.

To demonstrate the analysis and conclusion more clearly and directly, we will consider a constant model, a linear model and a parabolic model. The higher order models can be extended in a similar way.

4.1 Simulated example 1 --- A constant model

First, we consider a constant model, which is generated as follows:

$$Y_i = \beta_0(t_i) + \beta_1(t_i)X_i + \varepsilon_i, \quad (4.2)$$

where $\beta_0(t_i) = 0.0$, $\beta_1(t_i) = 1.0$. X_i is the daily S&P index for 2003 and 2004. ε_i is generated from $N(0, 0.02)$ independently. The simulation is repeated 100 times with sample size $n = 516$. The optimal bandwidth and 95% confident intervals are calculated according to the description of Chapter 3.

The mean (top in block) and standard deviation (bottom in block) of the 100 MADE values are summarized in table 4.1.

From table 4.1, we can see that the overall bias for local constant (Nadarya-Watson), local linear and local parabolic estimations increases, but within the one time standard deviation. It means even when β is constant, such local constant estimation is enough, local linear and local parabolic estimations do not make the results worse.

Fig. 4.1-4.6 show the mean of the 100 estimated values of β_1 , β_1' , β_1'' and their sample standard deviation using local constant (Nadarya-Watson), local linear and local parabolic estimations. Fig 4. 15 is the plot of Y vs X , and Fig 4.16 shows Y and estimated Y vs time.

Table 4.1 The mean and standard deviation of the 100 MADE values for the constant model

	e_0	e_1	$e = e_0 + e_1$
Local constant	0.0005726276	0.0251420735	0.0257147
	0.0002215662	0.0095122872	0.009545667
Local linear	0.000654793	0.035535783	0.03619058
	0.0002361304	0.0119087754	0.0119623
Local parabolic	0.0008487866	0.0404588940	0.04130768
	0.0002326101	0.0122197248	0.01221564

From Fig 4.1-4.3, we can see that for the three local estimates, the estimated values of β_1 are very close to the true value, 1, and all estimated values are within the 95% confidence intervals. Although we assume β_1 is a constant, thus local constant estimation is enough while local linear and local parabolic estimations are unnecessary, they do not worsen the estimations. From Fig. 4.4-4.6, we can see that the values of β_1', β_1'' using local linear and local parabolic estimations do not significantly differ from 0, just as expected, which may conclude that local constant estimation method is good enough. At both ends, we still have accurate estimates for $\beta_j(t)$, but its variance gets much bigger than that in the middle, which is mainly because of the boundary effects.

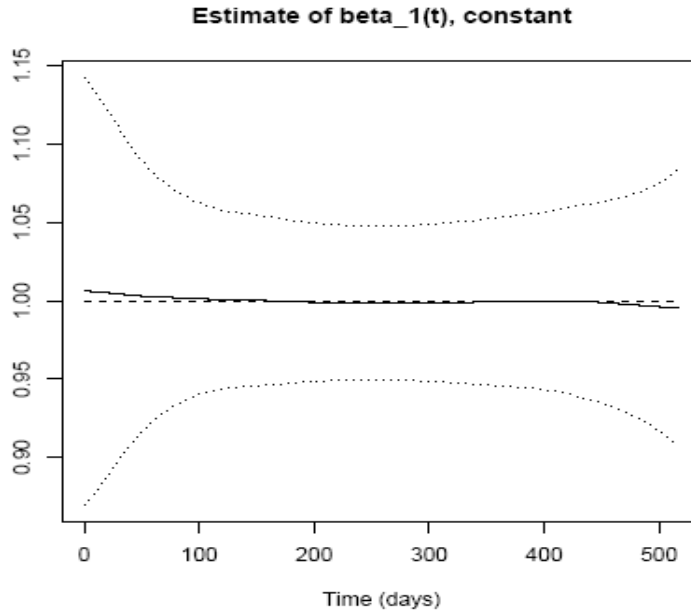


Fig.4.1 The Nadarya-Watson estimator (solid line) of β_1 with 95% confidence intervals (dotted line) for $\beta_1 = 1$.

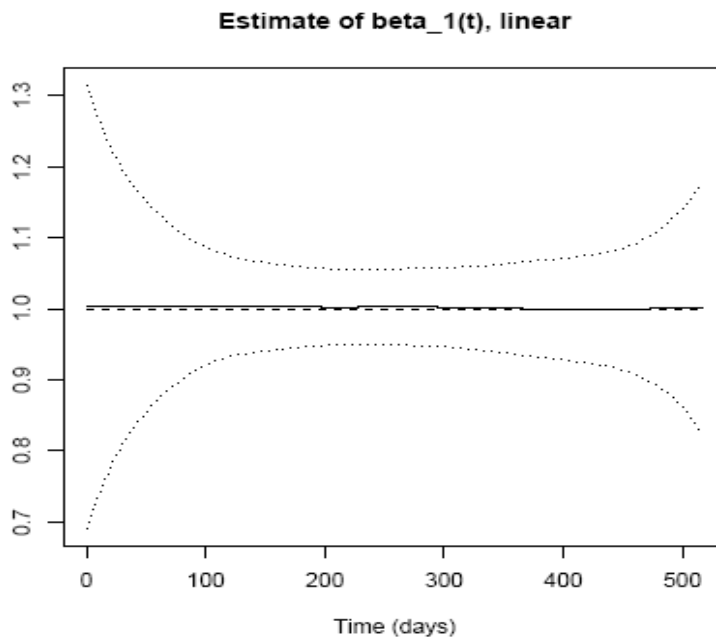


Fig. 4.2 The local linear estimator (solid line) of β_1 with 95% confidence intervals (dotted line) for $\beta_1 = 1$.

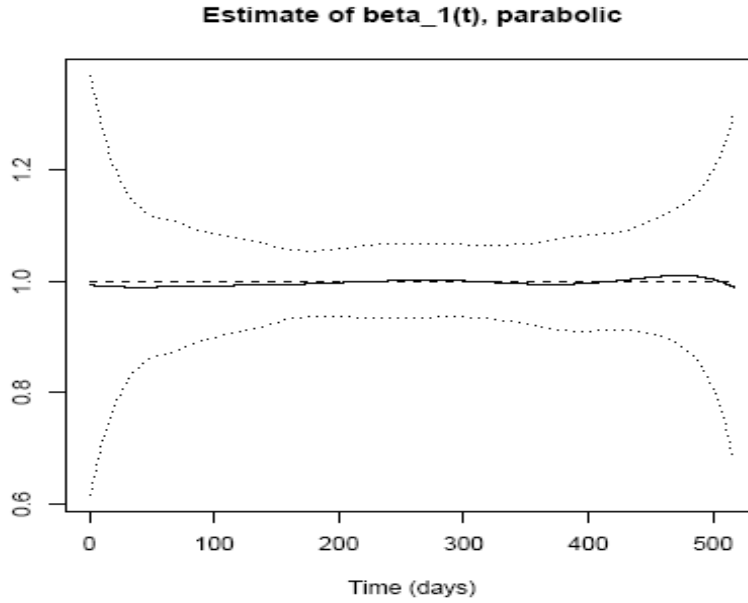


Fig. 4.3 The local parabolic estimator (solid line) of β_1 with 95% confidence intervals (dotted line) for $\beta_1 = 1$.

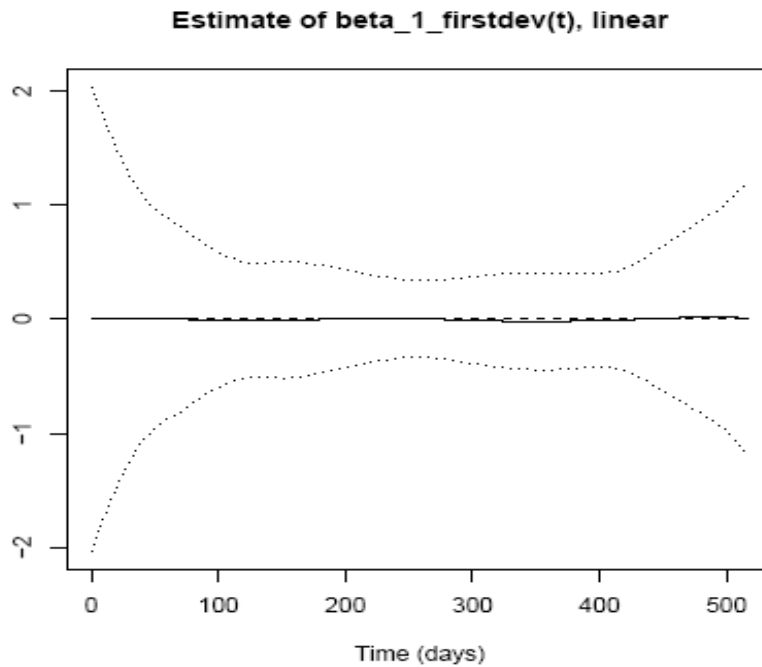


Fig. 4.4 The local linear estimator (solid line) of β_1' with 95% confidence intervals (dotted line) for $\beta_1 = 1$.

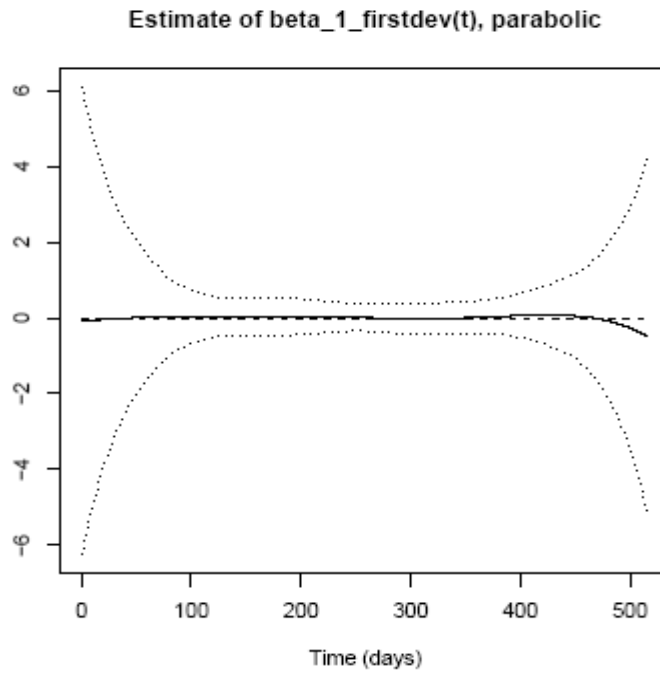


Fig. 4.5 The local parabolic estimator (solid line) of β_1' with 95% confidence intervals (dotted line) for $\beta_1 = 1$.

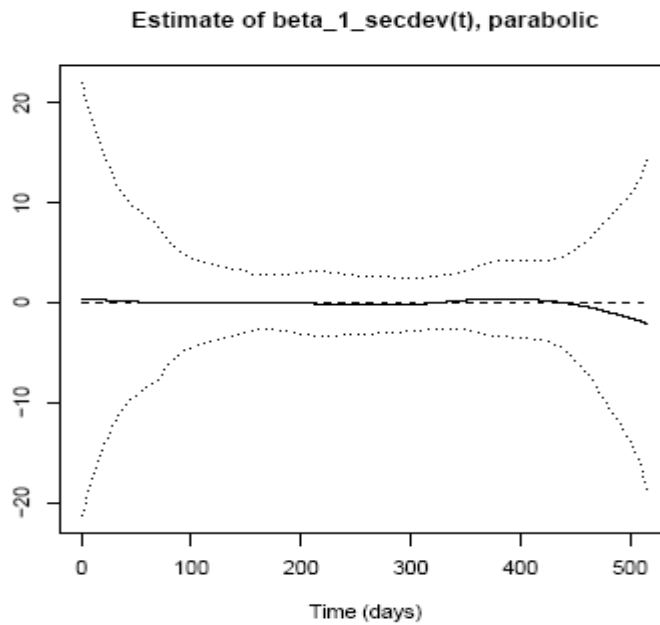


Fig. 4.6 The local parabolic estimator (solid line) of β_1'' with 95% confidence intervals (dotted line) for $\beta_1 = 1$.

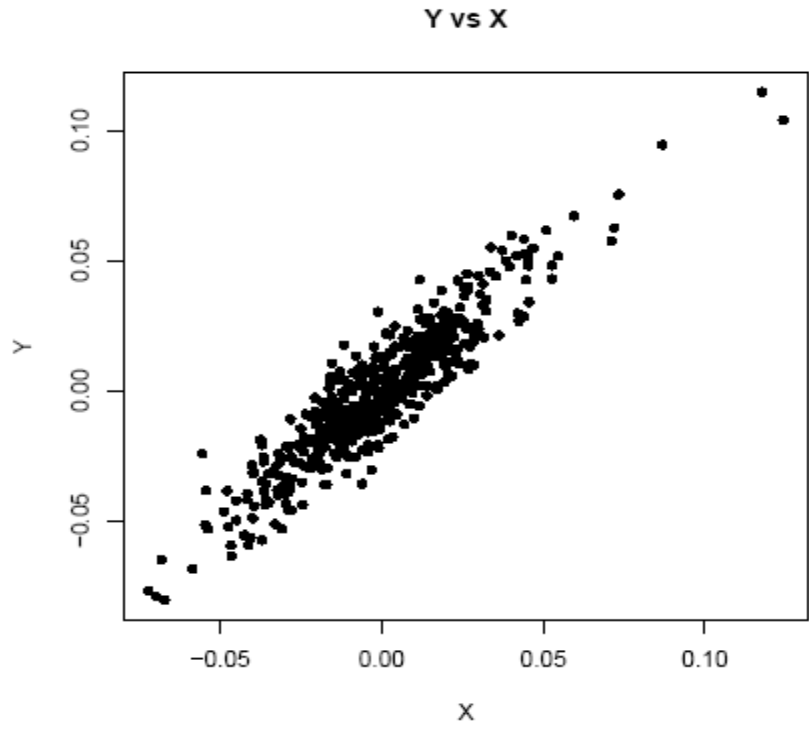


Fig. 4.7 Y vs X for $\beta_1 = 1$.

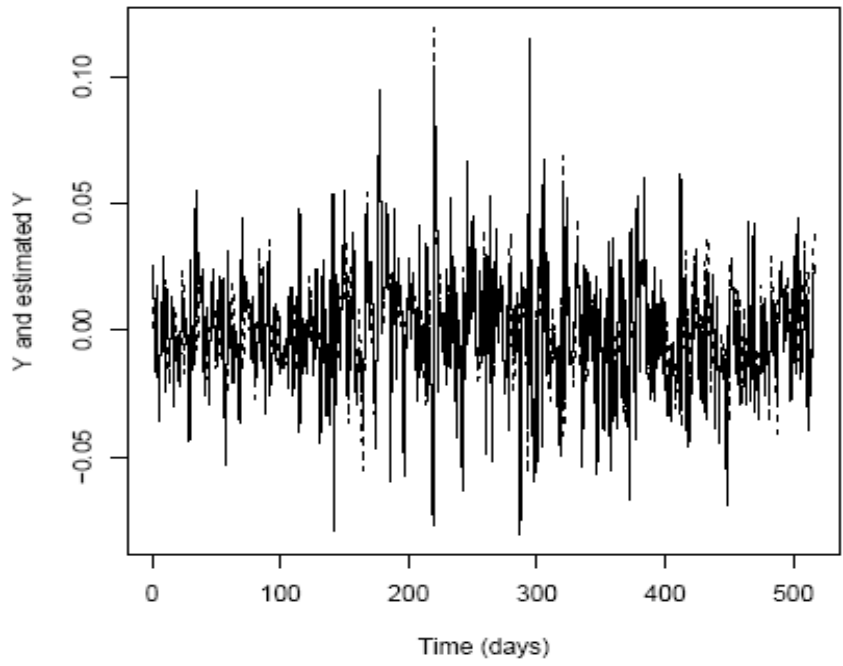


Fig. 4.8 Y (solid line) and estimated Y (dotted line) for $\beta_1 = 1$.

4.2 Simulated example 2 --- A linear model

Second, we consider a linear model, which is generated as follows:

$$Y_i = \beta_0(t_i) + \beta_1(t_i)X_i + \varepsilon_i, \quad (4.3)$$

where $\beta_0(t_i) = 0.0$, $\beta_1(t_i) = 2t_i$. The data X_i and the simulation process are the same as Section 4.1.

The mean (top in block) and standard deviation (bottom in block) of the 100 MADE values are summarized in table 4.2.

Table 4.2 The mean and standard deviation of the 100 MADE values for the linear model

	e_0	e_1	$e = e_0 + e_1$
Local constant	0.0006919575	0.0988648660	0.09955682
	0.0002544398	0.0143103939	0.01438828
Local linear	0.0006431996	0.0333548272	0.03399803
	0.0002207628	0.0129108871	0.01293842
Local parabolic	0.0008494827	0.0398192934	0.04066878
	0.000252125	0.011736765	0.01178139

From table 4.2, we can see that the overall bias for local constant (Nadarya-Watson) estimation is larger than that for the local linear estimation. The local parabolic

estimation does not worsen results too much although it does not improve the results. As expected, the local linear estimation is the best.

Fig. 4.9-4.14 show the mean of the estimated values of β_1 , β_1' , β_1'' and their sample standard deviation using local constant (Nadarya-Watson), local linear and local parabolic estimations. Fig 4. 15 is the plot of Y vs X , and Fig 4.16 shows Y and estimated Y vs time.

From Fig. 4.10-4.11, we can see that for local linear and local parabolic estimates, the estimated values of β_1 are very close to the true values and all estimated values are within the 95% confidence intervals. However the local constant estimates of β_1 in Fig 4.9 obviously deviate from the true values of β_1 .

From Fig. 4.12 and 4.13, we can see that the values of β_1' using local linear and local parabolic estimations significantly differ from 0. It means the β_1 changes over time, therefore local constant estimation is not sufficient. On the other hand, the values of β_1'' using local parabolic estimations do not significantly differ from 0 (Fig. 4.14), just as expected, which again implies that local linear method is good enough.

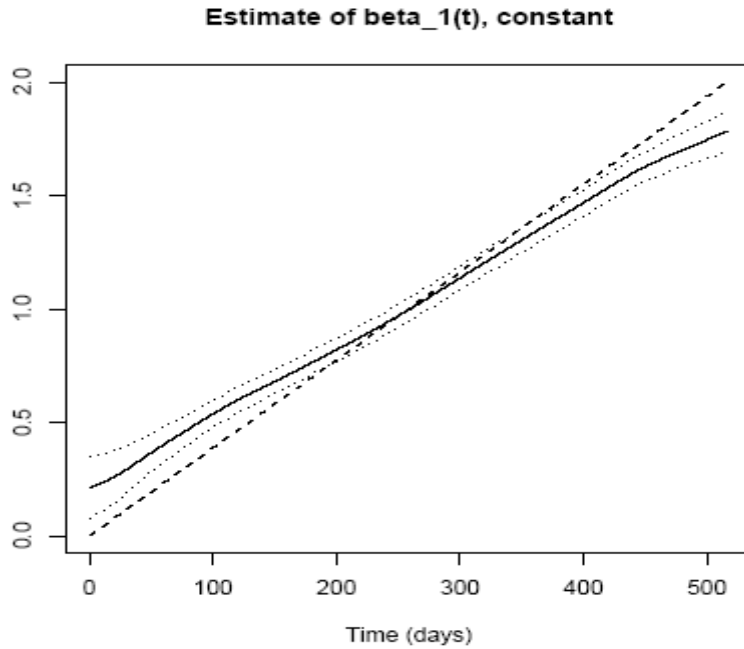


Fig. 4.9 The Nadarya-Watson estimator (solid line) of β_1 with 95% confidence intervals (dotted line) for $\beta_1 = 2t$ (dashed line).

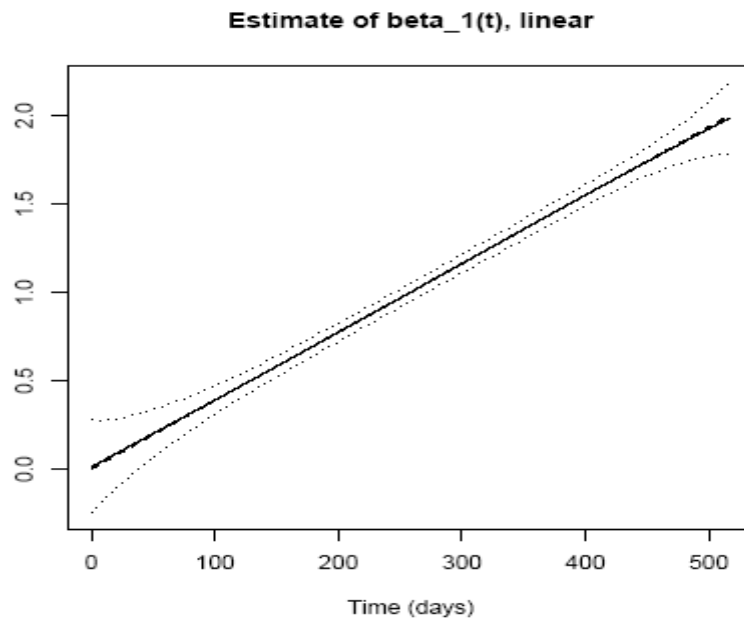


Fig. 4.10 The local linear estimator (solid line) of β_1 with 95% confidence intervals (dotted line) for $\beta_1 = 2t$ (dashed line).

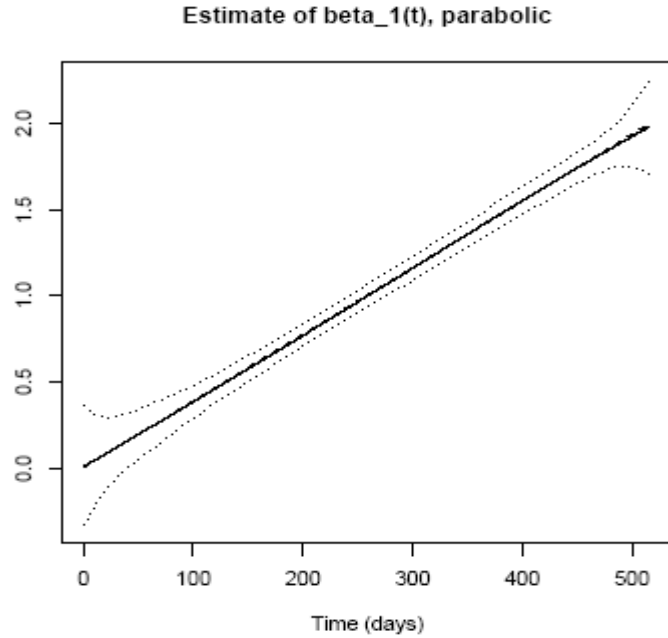


Fig. 4.11 The local parabolic estimator (solid line) of β_1 with 95% confidence intervals (dotted line) for $\beta_1 = 2t$ (dashed line).

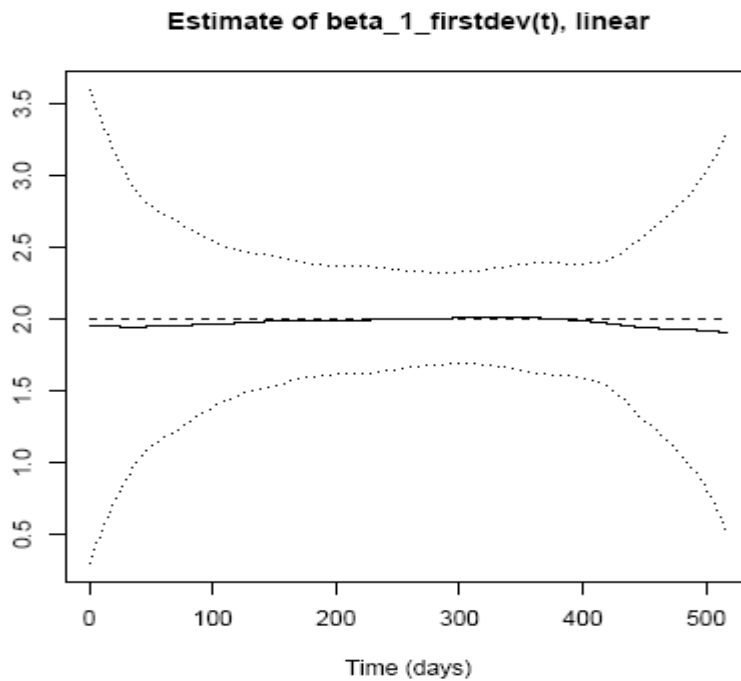


Fig. 4.12 The local linear estimator (solid line) of β_1' with 95% confidence intervals (dashed line) for $\beta_1 = 2t$.

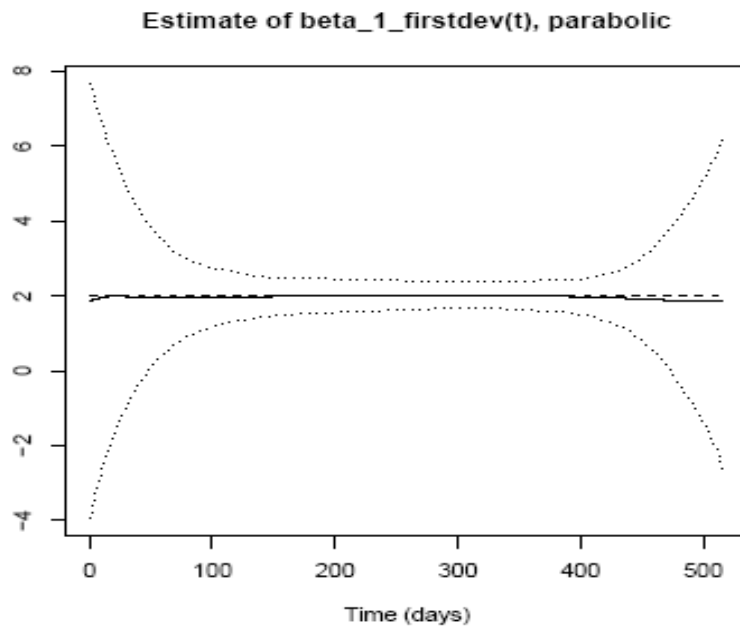


Fig. 4.13 The local parabolic estimator (solid line) of β_1' with 95% confidence intervals (dashed line) for $\beta_1 = 2t$.

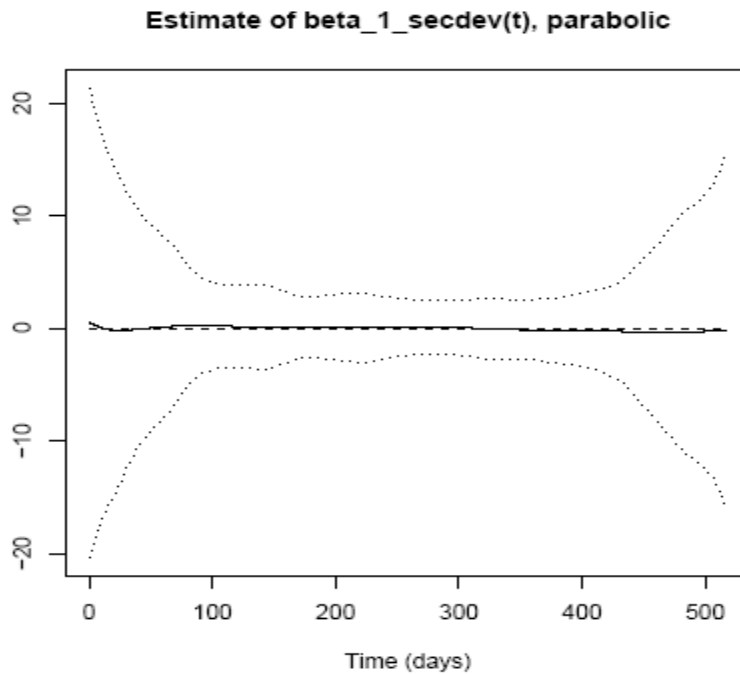


Fig. 4.14 The local parabolic estimator (solid line) of β_1'' with 95% confidence intervals (dashed line) for $\beta_1 = 2t$.

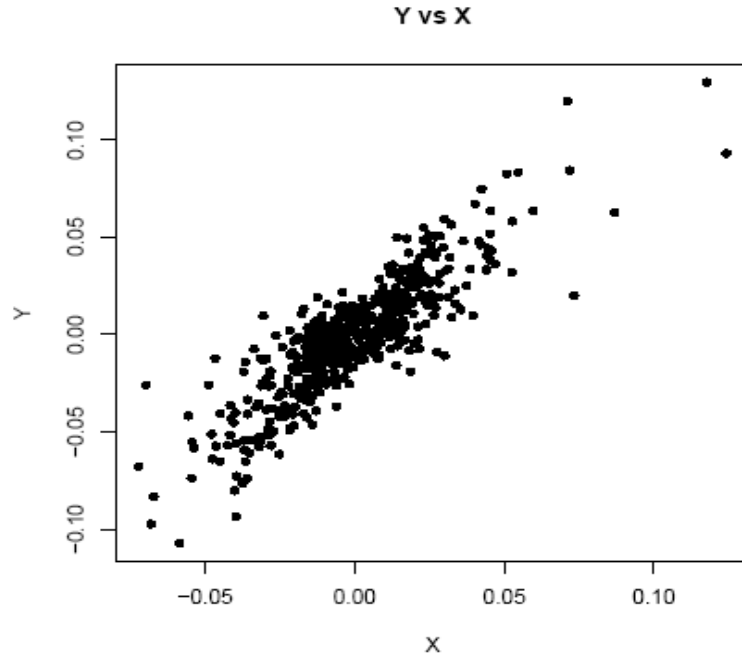


Fig. 4.15 Y vs X for $\beta_1 = 2t$.

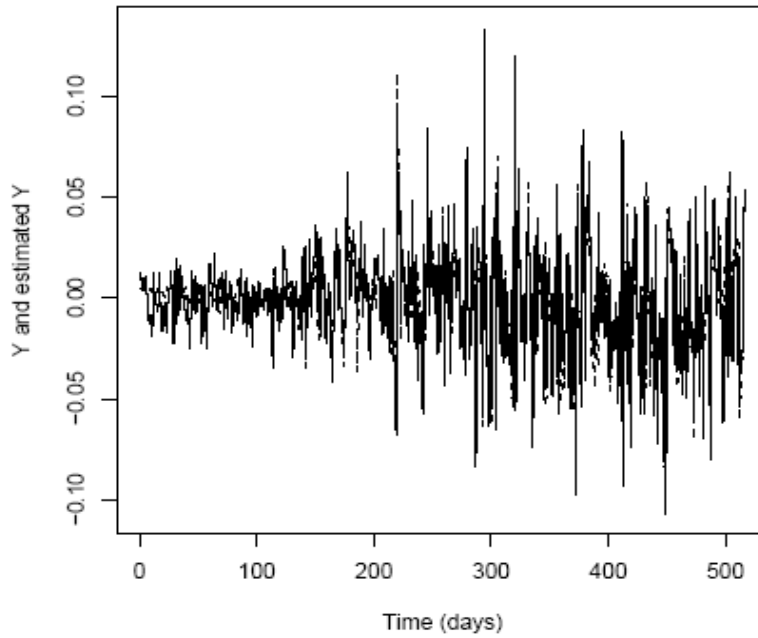


Fig. 4.16 Y (solid line) and estimated Y (dotted line) for $\beta_1 = 2t$.

4.3 Simulated example 3 --- A parabolic model

Third, we consider a linear model, which is generated as follows:

$$Y_i = \beta_0(t_i) + \beta_1(t_i)X_i + \varepsilon_i, \quad (4.4)$$

where $\beta_0(t_i) = 0.0$, $\beta_1(t_i) = 8(t_i - 0.5)^2$. The data X_i and the simulation process are the same as Section 4.1.

The mean (top in block) and standard deviation (bottom in block) of the 100 MADE values are summarized in table 4.3.

Table 4.3 The mean and standard deviation of the 100 MADE values for the parabolic model

	e_0	e_1	$e = e_0 + e_1$
Local constant	0.0007542788	0.1987400918	0.1994944
	0.0002766998	0.0129962653	0.01302159
Local linear	0.0006736791	0.0750514298	0.07572511
	0.0002369932	0.0156451021	0.01569648
Local parabolic	0.0008544916	0.0409885485	0.04184304
	0.0002473328	0.0119249394	0.01194850

From table 4.3, we can see that the overall bias for local linear estimation is less than that for the local constant estimation, and the bias for the local parabolic estimations is least.

That is what we expected.

Fig. 4.17-4.22 show mean of the 100 estimated values of β_1 , β_1' , β_1'' and their sample standard deviations using local constant, local linear and local parabolic estimations. Fig. 4. 23 is the plot of Y vs X , and Fig 4.24 shows Y and estimated Y vs time.

From Fig 4.17-4.19, we can see that the estimated values of β_1 for local constant estimation are the worst (Fig. 4.17); improvements are made by local linear estimation (Fig. 18); the best, as expected, is for local parabolic estimation (Fig 19).

From Fig. 4.20-4.21, we can see that the values of β_1' using local linear and local parabolic estimations significantly differ from 0 and not a constant. It means the β_1 changes over time and non-linearly, thus local constant estimation is not sufficient. Fig. 4.21 is much better than Fig. 4.20, which implies local parabolic estimation is better than local linear estimation. On the other hand, the values of β_1'' using local parabolic estimations significantly differ from 0 (Fig. 4.22), but is a constant, which implies that local parabolic estimation is necessary and it should be good enough.

For this example, we also calculate the $\text{var}(\hat{\beta}_j(t))_{\text{ls}}$ by the least square method, as shown in Fig. 4.25 for $\hat{\beta}_0(t)$ and Fig 4.26 for $\hat{\beta}_1(t)$. From Fig. 4.25 and 4.26, we can find that in most of region of t , the $\text{var}(\hat{\beta}_j(t))_{\text{ls}}$ calculated using the least square method are close to $\text{var}(\hat{\beta}_j(t))_{\text{sample}}$ calculated from the 100 simulations. That means, in practice, we can estimate the variance of $\hat{\beta}_j(t)$ using the least square estimate, $\text{var}(\hat{\beta}_j(t))_{\text{ls}}$. Note that the approximation in Fig 4.26 is better than that in Fig. 4.25. This is because β_0 is a constant, thus the departure in Fig 4.25 is expected, but still reasonable.

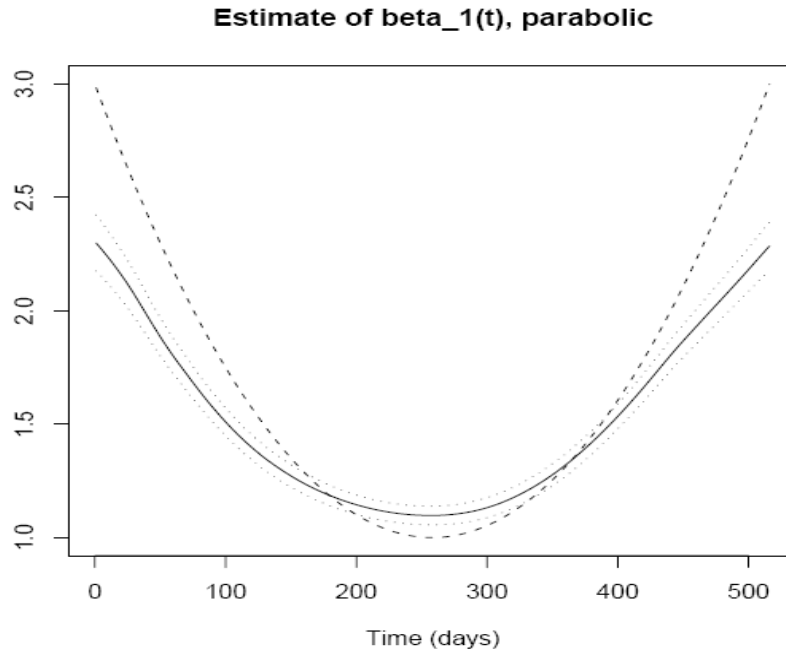


Fig. 4.17 The Nadarya-Watson estimator (solid line) of β_1 with 95% confidence intervals (dotted line) for $\beta_1 = 8(t - 0.5)^2$ (dashed line).

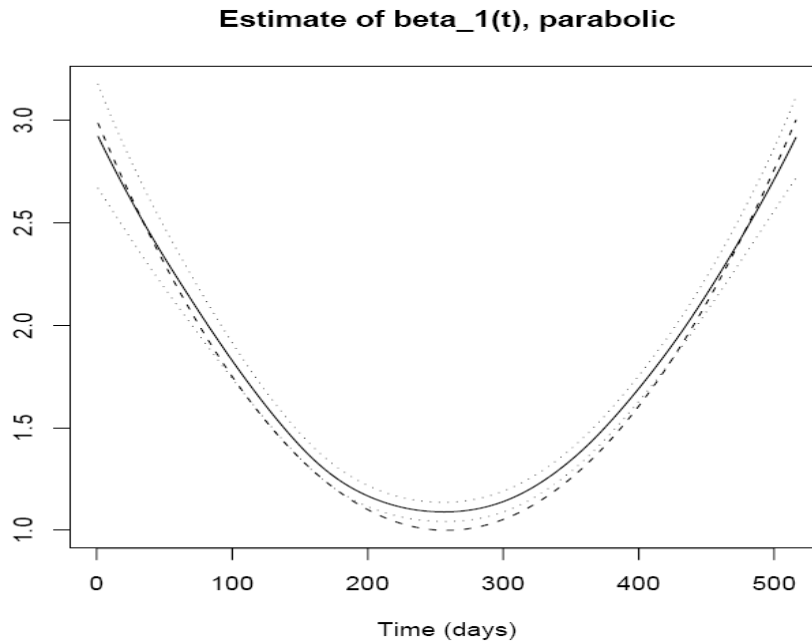


Fig. 4.18 The local linear estimator (solid line) of β_1 with 95% confidence intervals (dotted line) for $\beta_1 = 8(t - 0.5)^2$ (dashed line).

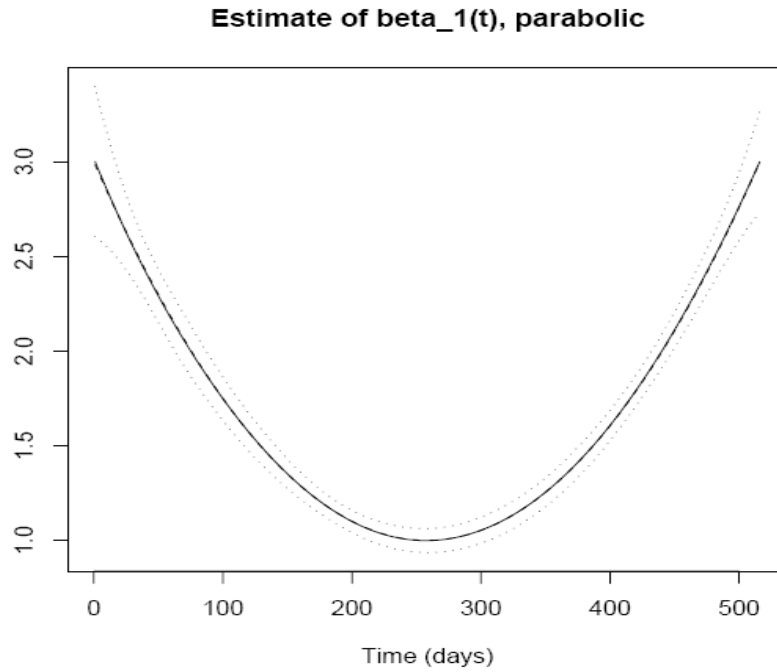


Fig. 4.19 The parabolic estimator (solid line) of β_1 with 95% confidence intervals (dotted line) for $\beta_1 = 8(t - 0.5)^2$ (dashed line).

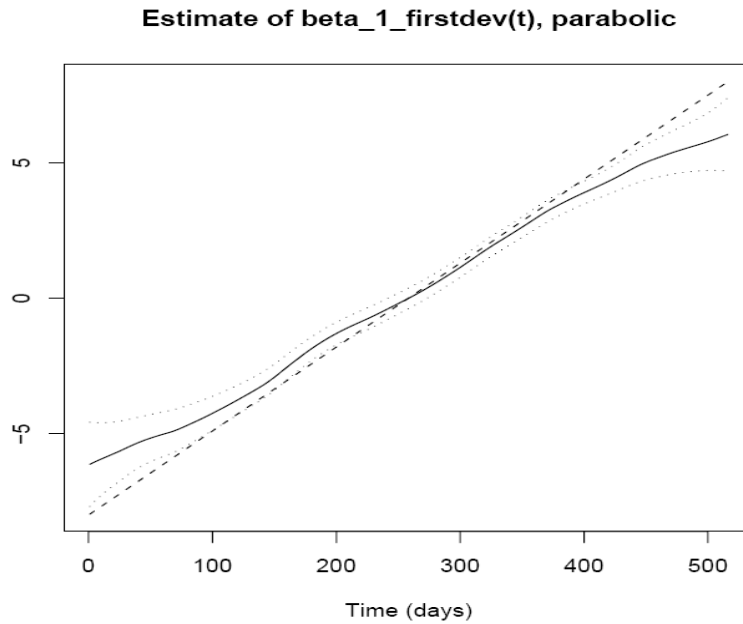


Fig. 4.20 The local linear estimator (solid line) of β_1' with 95% confidence intervals (dashed line) for $\beta_1 = 8(t - 0.5)^2$.

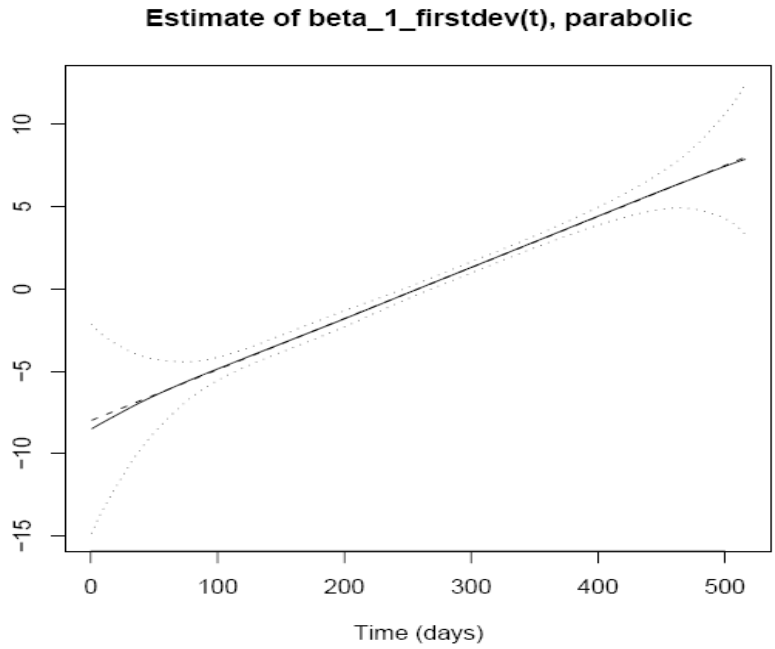


Fig. 4.21 The local parabolic estimator (solid line) of β_1' with 95% confidence intervals (dashed line) for $\beta_1 = 8(t - 0.5)^2$.

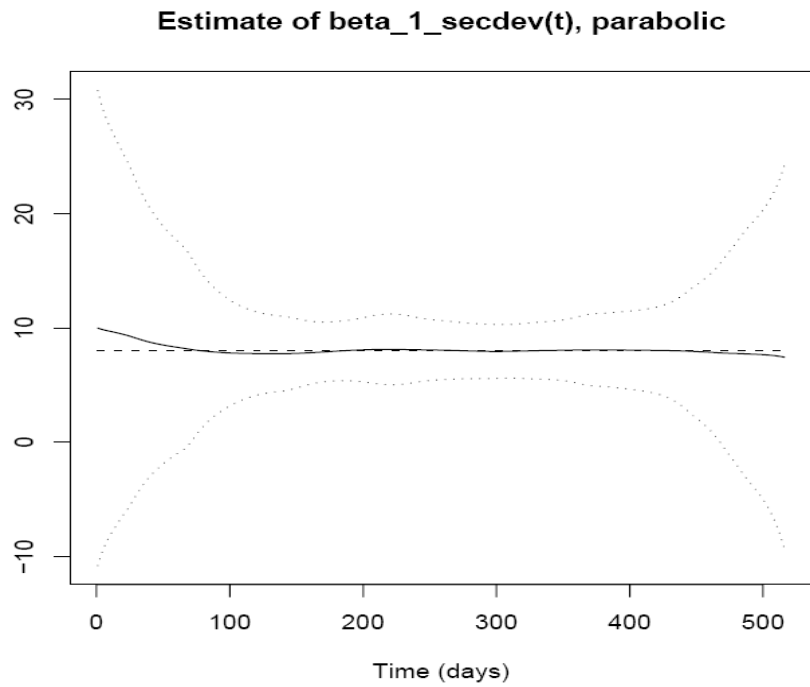


Fig. 4.22 The local parabolic estimator (solid line) of β_1'' with 95% confidence intervals (dashed line) for $\beta_1 = 8(t - 0.5)^2$.

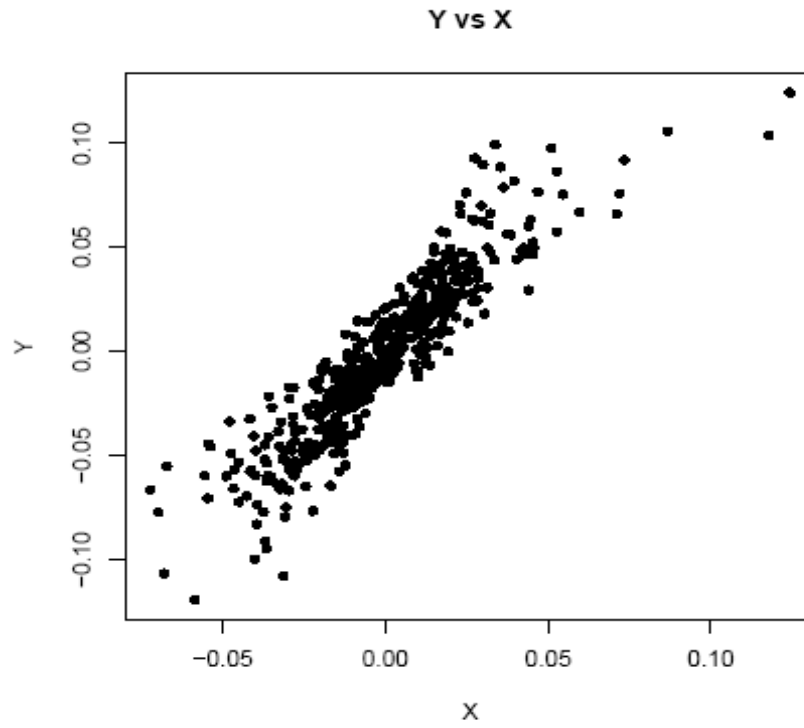


Fig. 4.23 Y vs X for $\beta_1 = 8(t - 0.5)^2$

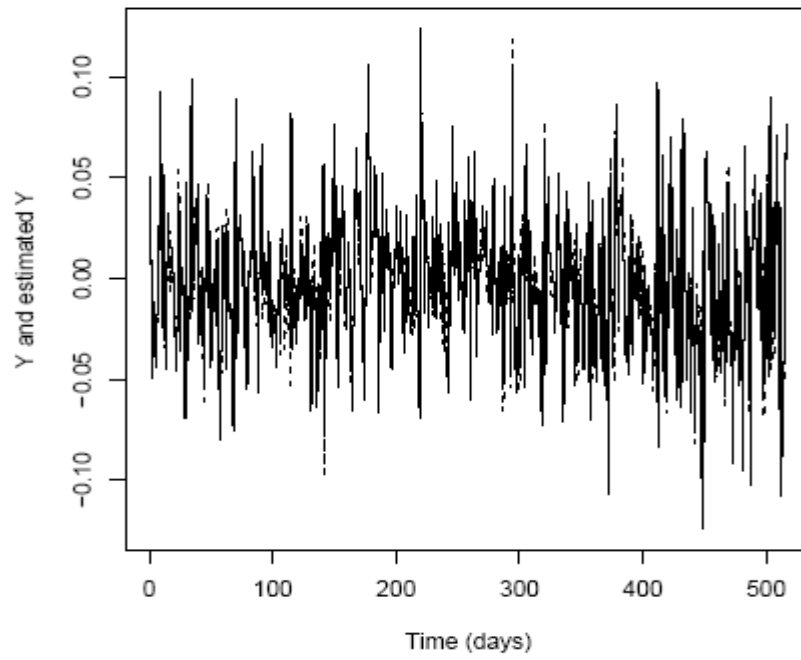


Fig. 4.24 Y (solid line) and estimated Y (dotted line) for $\beta_1 = 8(t - 0.5)^2$.

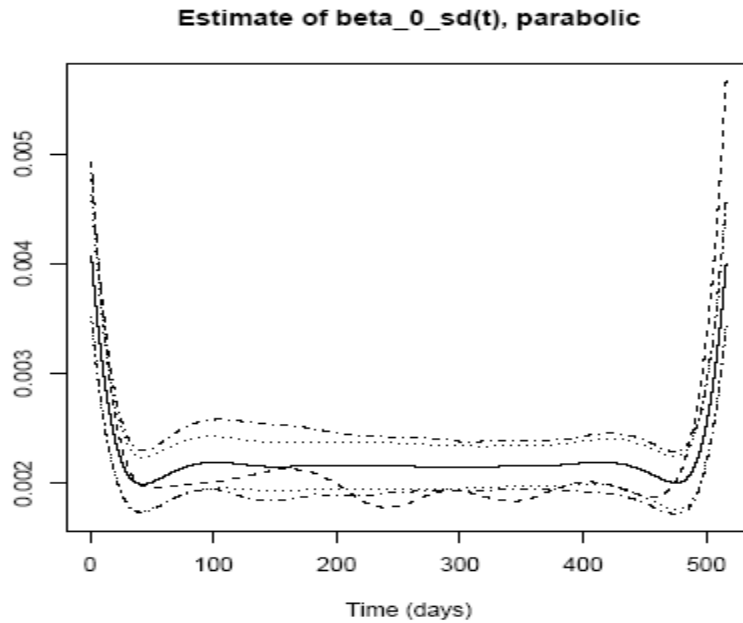


Fig. 4.25 $\text{var}(\hat{\beta}_0(t))_{\text{sample}}$ (dash line), and the mean of $\text{var}(\hat{\beta}_0(t))_{\text{ls}}$ (solid line) and its 95% confidence intervals (dotted lines), maximum and minimum (dash-dot lines)

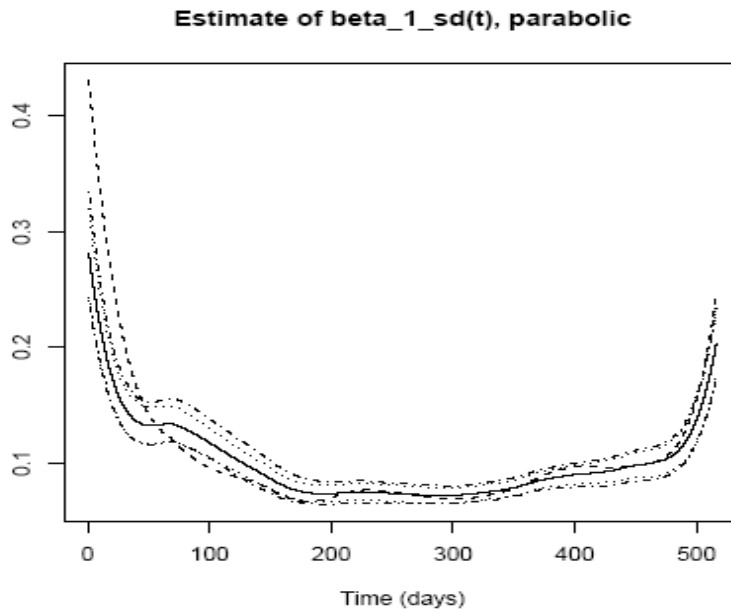


Fig. 4.26 $\text{var}(\hat{\beta}_1(t))_{\text{sample}}$ (dash line), and the mean of $\text{var}(\hat{\beta}_1(t))_{\text{ls}}$ (solid line) and its 95% confidence intervals (dotted lines), maximum and minimum (dash-dot lines)

4.4 A real example

In finance and security analysis, one often measures the risk of an individual stock as its regression coefficient against a market index or portfolio. If this coefficient, usually called a beta-coefficient in CAPM, 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 considered to be more risky; see the books by Cochrane (2001) and Tasy (2002) for more details. However, some recent studies show that the beta-coefficients might vary over time for many applications in economics and finance; see Ghysels (1998), and Akdeniz *et. al.* (2003).

We apply the proposed method and its modeling procedures to analyze the common stock price of LOW during the year 2003 and 2004 using the daily closing prices. To measure its risk relative to the market of U. S. blue chip stocks, we take the standard and Poor's 500 (P&S 500) as a proxy to this market. Following the convention in the finance literature, here we consider the simple daily stock returns. The time series plots of LOW stock returns and S&P 500 index returns are displayed in Fig. 4.27 and Fig. 4.28. It should be noted that the returns of S&P 500 index may be as nonstationary as the returns of LOW stock.

To establish the empirical relationship between the return of LOW stock and S&P 500 index, first we fit a simple regression model as

$$r_i = \beta_0(t) + \beta_1(t)r_{\text{MKT},i} + \varepsilon_i, \quad (4.5)$$

where β_0, β_1 are assumed as constants. The respective least-square estimates of β_1 are plotted in Fig. 4.29 (dashed line).

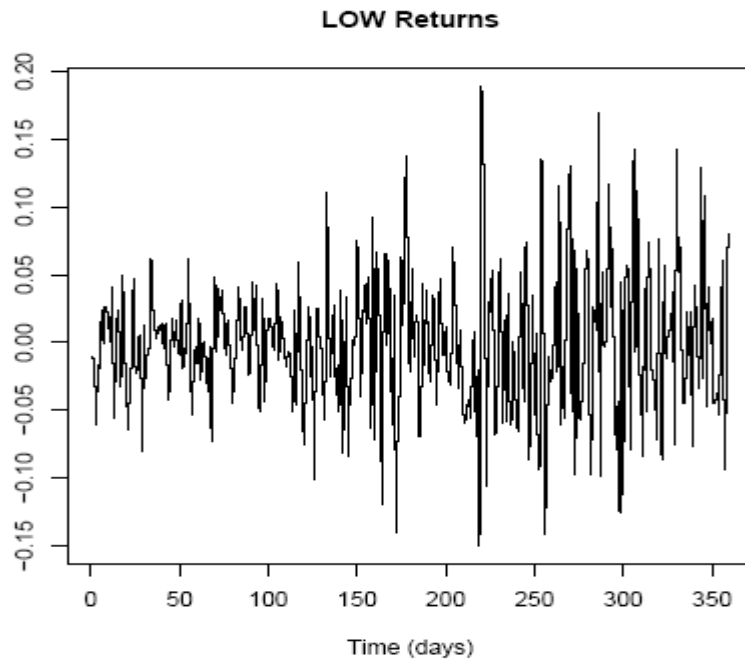


Fig. 4.27 The time series plot of LOW stock return

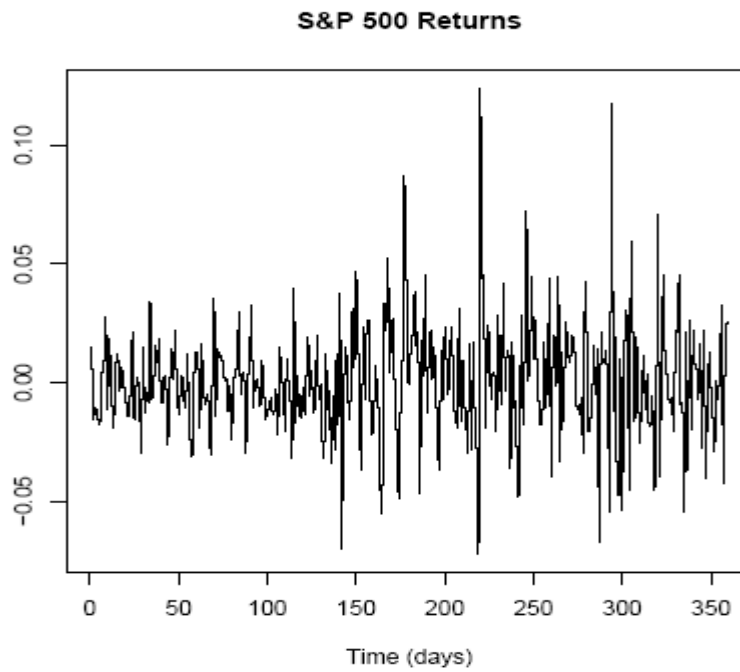


Fig. 4.28 The time series plot of S&P 500 index return

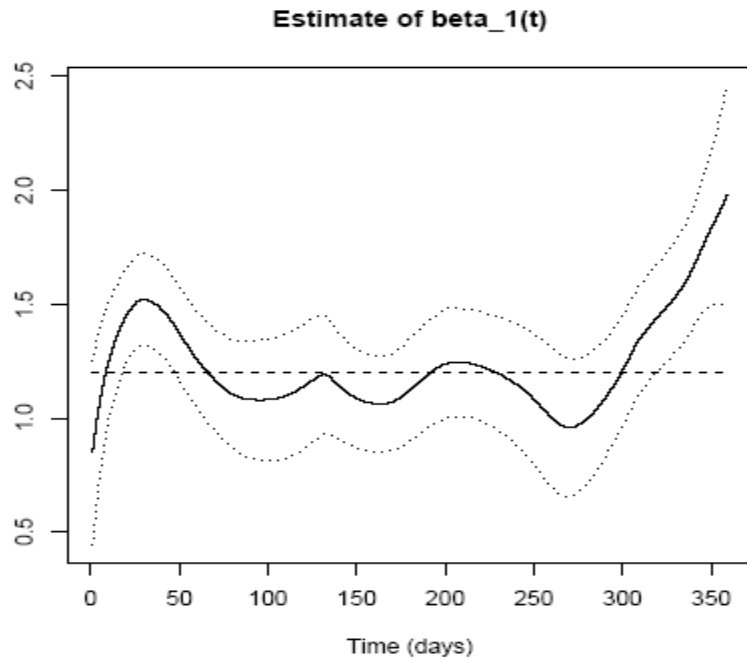


Fig. 4.29 The local polynomial estimator (solid line) of β_1 with confidence intervals (dotted lines) and least-square estimator (dashed line)

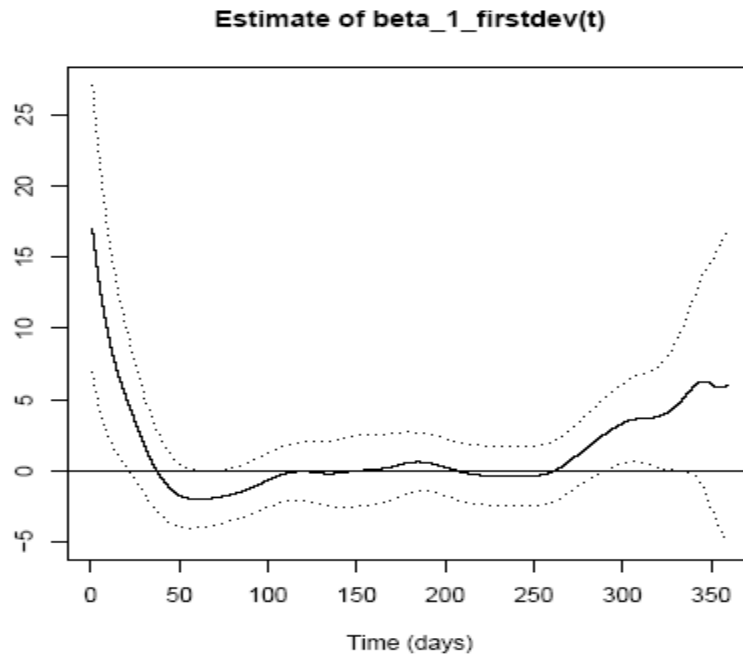


Fig. 4.30 The local polynomial estimator (solid line) of β_1' with 95% confidence intervals (dashed lines)

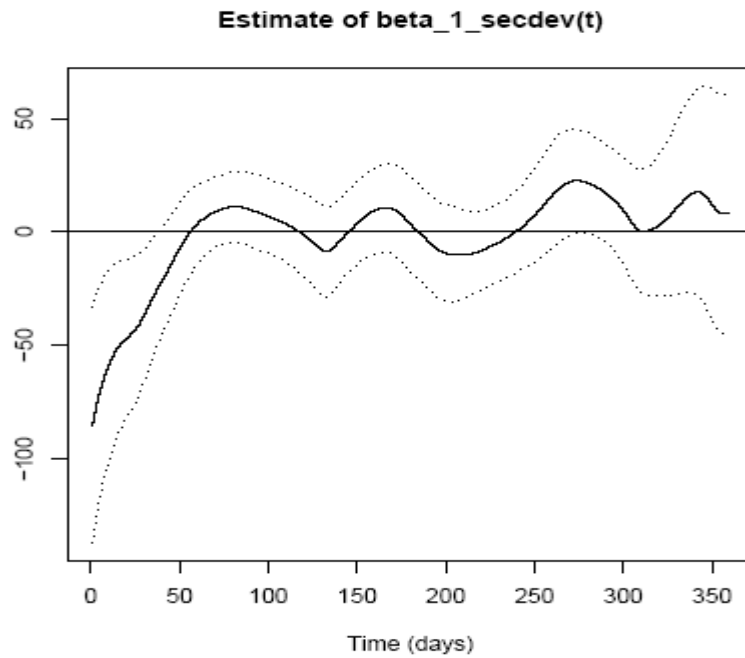


Fig. 4.31 The local polynomial estimator (solid line) of β_1'' with 95% confidence intervals (dashed lines)

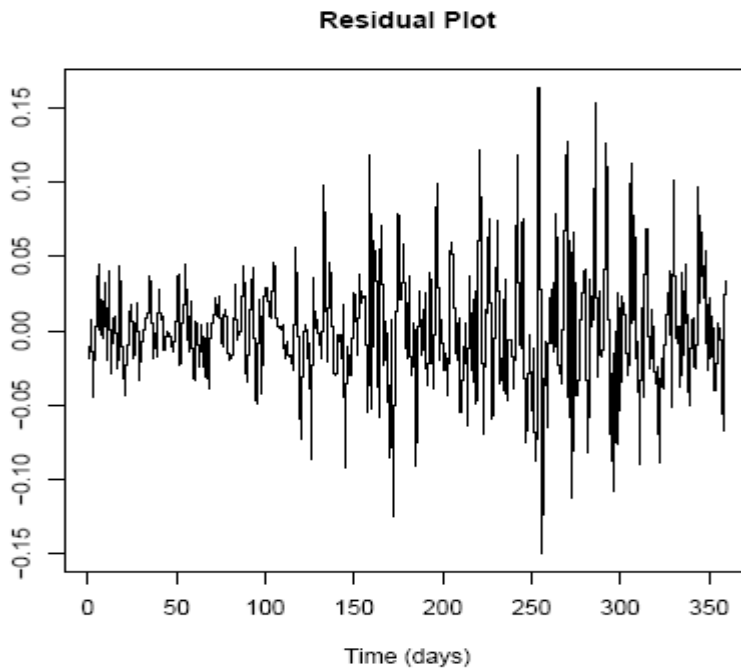


Fig.4.32 The residual plot of LOW stock return

However, we might suspect that the coefficients β_0, β_1 might change over time. Therefore, we fit the following time varying coefficient model with our proposed method:

$$r_i = \beta_0(t) + \beta_1(t)r_{\text{MKT},i} + \varepsilon_i, \quad (4.6)$$

where $\beta_0(t), \beta_1(t)$ are assumed as some functions with t . The optimal bandwidth and confident intervals are calculated according to the description of chapter 3. The local polynomial estimates of β_1 and β_1', β_1'' are depicted in Fig 4.29-4.31, and the residual is plotted in Fig. 4.32. The residual plot shows the errors are close to normal distribution so the local parabolic estimation method seems efficient.

It is evident from Fig. 4.29 that β_1 does change over time. It roughly keeps constant for first year; then it increases afterwards and finally decreases. The β_1', β_1'' (Fig. 4.30, 4.31) are significant different from 0. This further proves that β_1 does change over time, and shows that the local polynomial estimate method are necessary. The values of β_1 are large than 1 over all time. Therefore, one can conclude that LOW is a stock that was more volatile than the U. S. blue chip market as a whole.

CHAPTER 5

CONCLUSIONS

We developed a useful class of time series models, the time-varying coefficient time series models for modeling nonlinear, nonstationary and trending time series. We obtained some insights about the modeling methods and we demonstrated that when the coefficients change with time non-linearly, which is common in finance and security analysis, the local constant and local linear estimator may not be sufficient. For this case, we show that the local polynomial estimator (in this paper, we take parabolic estimation as an example) is necessary and show that it is superior to the local constant and local linear estimator. We simulated a constant model, a linear model and a parabolic model, and demonstrate the usefulness of the models by a real example. To make the model practically useful, we proposed an easily implemented bandwidth selector.

Finally, the predictive utility of using the time-varying coefficient time series models studied in this paper needs definitely a further investigation due to its importance in various applications in economics and finance.

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(4), Article 1.
- Cai, Z. (2006) Varying coefficient dynamic panel data models. *Journal of Econometrics*, in Press.
- Cai, Z., M. Das, H. Xiong and X. Wu (2006). Functional-coefficient instrumental variables models. *Journal of Econometrics*, in Press.
- Cai, Z. and Ould-Said, E. (2003). Local robust regression estimation for time series. *Statistics and Probability Letters* 65, 433-499.
- Chang, Y., Park, J.Y. and Phillips, P.C.B. (2001). Nonlinear econometric models with cointegrated and deterministically trending regressors. *Econometrics Journal*, 4, 1-36. 21.
- Cochrane, J.H. (2001) *Asset Pricing*. Princeton University Press, Englewood Cliffs, NJ.
- Fan, J., Zhang, C. and Zhang, J. (2001). Generalized likelihood ratio statistics and Wilks phenomenon. *The Annals of Statistics*, 29, 153-193.
- Franses, P.H. (1996). *Periodicity and Stochastic Trends in Economic Time Series*. New York: Cambridge University Press.
- Franses, P.H. (1998). *Time Series Models for Business and Economic Forecasting*. New York: Cambridge University Press.

- Ghysel, E. (1998) On stable factor structures in the pricing of risk: do time varying betas help or hurt. *Journal of Finance* 53, 549-574.
- Gorodetskii, V.V. (1977). On the strong mixing property for linear sequences. *Theory of Probability and Its Applications*, 22, 411-413.
- Granger, C.L.J. and TerÄasvirta, T. (1993). *Modelling Nonlinear Economic Relationships*. Oxford: Oxford University Press.
- Pagan, A. and Ullah, A. (1999). *Nonparametric Econometrics*. New York: Cambridge University Press.
- Park, J.Y. and Phillips, P.C.B. (1999). Asymptotics for nonlinear transformations of integrated time series. *Econometric Theory*, 15, 269-298.
- Phillips, P.C.B. (2001). Trending time series and macroeconomic activity: Some present and future challenges. *Journal of Econometrics*, 100, 21-27.
- Robinson, P.M. (1989). Nonparametric estimation of time-varying parameters. In *Statistical Analysis and Forecasting of Economic Structural Change* (P. Hackl, ed.). Springer-Verlag, Berlin, 1989, 253-164.
- Roussas, G.G. (1989). Consistent regression estimation with fixed design points under dependence conditions. *Statistics and Probability Letters*, 8, 41-50.
- Tsay, R.S. (2002). *Analysis of Financial Time Series*. New York: John Wiley & Sons.