

Article

Efficient Estimation in Heteroscedastic Varying Coefficient Models

Chuanhua Wei ^{1,*} and Lijie Wan ²

¹ Department of Statistics, Minzu University of China, Beijing 100081, China

² Department of Statistics, University of Kentucky, Lexington, KY 40508, USA;

E-Mail: lijie.wan@uky.edu

* Author to whom correspondence should be addressed; E-Mail: chweisd@163.com;

Tel: +86-10-6893-3910 (ext. 8062).

Academic Editor: Kerry Patterson

Received: 16 January 2015 / Accepted: 2 July 2015 / Published: 15 July 2015

Abstract: This paper considers statistical inference for the heteroscedastic varying coefficient model. We propose an efficient estimator for coefficient functions that is more efficient than the conventional local-linear estimator. We establish asymptotic normality for the proposed estimator and conduct some simulation to illustrate the performance of the proposed method.

Keywords: heteroscedasticity; local linear; varying coefficient models

JEL classifications: C13; C14

1. Introduction

Recently, the varying coefficient model has attracted much attention among econometricians and statisticians. One attractive feature of this model is its ability to capture the nonlinearity of the data without suffering from the “curse of dimensionality”. In general, it is of the form

$$Y_i = \mathbf{X}_i^T \boldsymbol{\alpha}(U_i) + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (1.1)$$

where Y_i 's are responses; $\mathbf{X}_i = (X_{i1}, X_{i2}, \dots, X_{ip})^T$ and U_i are associated covariates; $\boldsymbol{\alpha}(\cdot) = (\alpha_1(\cdot), \alpha_2(\cdot), \dots, \alpha_p(\cdot))^T$ is a p -dimensional vector of unknown functions; ε_i 's are independent and identically distributed random errors with $E(\varepsilon_i | \mathbf{X}_i, U_i) = 0$ and $\text{Var}(\varepsilon_i | \mathbf{X}_i, U_i) = \sigma^2(\mathbf{X}_i, U_i)$.

Due to its flexibility, the varying coefficient model has been studied in many different contexts and has been successfully applied to nonlinear time series analysis, longitudinal and functional data analysis, panel data analysis, spatial data analysis, and time-varying models in finance. See, for example, the work of Cai *et al.* [1], Cai [2], Cai and Li [3], Cai *et al.* [4], Fan and Zhang [5], Fan *et al.* [6], Fotheringham *et al.* [7], Hoover *et al.* [8], Li *et al.* [9] and Xiao [10], among others.

In the above models, the varying coefficient model is generally estimated by the local-linear approach. Usually, the errors are assumed to be i.i.d. to start. However, in applications, heteroscedasticity is often found in residuals from both cross-sectional and time series modelling. In the context of the linear regression model, it is well known that if the errors are heteroscedastic, then the generalized least-squares (GLS) estimator is more efficient than ordinary least-squares (OLS) estimator. To the best of our knowledge, there has been no work on the problem of designing an efficient estimation method for varying coefficient models with heteroscedastic errors. In this paper, we propose an efficient estimator for varying coefficients based on the local linear approach.

The paper is structured as follows. We introduce an efficient estimator in Section 2, and their asymptotic properties are given in Section 3. We report the results of some Monte Carlo simulations in Section 4.

2. Efficient Estimation

Without considering heteroscedasticity, we apply a local linear regression technique to estimate the varying coefficient functions. For each given u , the local linear estimator $\hat{\alpha}(u)$ of $\alpha(u)$ is the part corresponding to \mathbf{a} of the minimizer of

$$\sum_{i=1}^n [Y_i - \mathbf{X}_i^T \mathbf{a} - \mathbf{X}_i^T (U_i - u) \mathbf{b}]^2 K_h(U_i - u_0), \quad (2.1)$$

where K is a kernel function, h is a bandwidth and $K_h(\cdot) = K(\cdot/h)/h$. Then we have

$$\hat{\alpha}(u) = [\hat{\alpha}_1(u), \dots, \hat{\alpha}_p(u)]^T = (\mathbf{I}_p \quad \mathbf{0}_p) \{ \mathbf{D}_u^T \mathbf{W}_u^\delta \mathbf{D}_u \}^{-1} \mathbf{D}_u^T \mathbf{W}_u^\delta \mathbf{Y}. \quad (2.2)$$

where

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1^T \\ \mathbf{X}_2^T \\ \vdots \\ \mathbf{X}_n^T \end{bmatrix}, \mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \mathbf{D}_u = \begin{bmatrix} \mathbf{X}_1^T & \frac{U_1-u}{h} \mathbf{X}_1^T \\ \mathbf{X}_2^T & \frac{U_2-u}{h} \mathbf{X}_2^T \\ \vdots & \vdots \\ \mathbf{X}_n^T & \frac{U_n-u}{h} \mathbf{X}_n^T \end{bmatrix},$$

and $\mathbf{W}_u = \text{diag}\{K_h(U_1 - u), K_h(U_2 - u), \dots, K_h(U_n - u)\}$.

The estimator $\hat{\alpha}(u)$ ignores the information contained in the variance matrix and it is inefficient. To overcome this, we propose a class of efficient estimators in the following.

Denote $\sigma_i = \sqrt{\sigma^2(\mathbf{x}_i, u_i)}$, for the moment where we assume that σ_i is known. Multiply both sides of model (1.1) by $1/\sigma_i$, we have the following homoscedastic varying coefficient model

$$Y_i^* = \mathbf{Z}_i^T \alpha(u_i) + e_i, \quad i = 1, 2, \dots, n, \quad (2.3)$$

where $Y_i^* = y_i/\sigma_i$, $\mathbf{Z}_i = \mathbf{X}_i/\sigma_i$ and $e_i = \varepsilon_i/\sigma_i$ with $E(e_i|\mathbf{x}_i, u_i) = 0$, $\text{Var}(e_i|\mathbf{x}_i, u_i) = 1$.

Applying the local linear approach to model (2.3), the efficient estimator of $\alpha(u)$ is given as follows

$$\tilde{\alpha}(u) = (\mathbf{I}_p \mathbf{0}_p) \{ \mathbf{B}_u^T \mathbf{W}_u \mathbf{B}_u \}^{-1} \mathbf{B}_u^T \mathbf{W}_u \mathbf{Y}^*. \quad (2.4)$$

where $\mathbf{Y}^* = (Y_1^*, Y_2^*, \dots, Y_n^*)$,

$$\mathbf{B}_u = \begin{bmatrix} \mathbf{Z}_1^T & \mathbf{Z}_1^T \frac{U_1 - u}{h} \\ \mathbf{Z}_2^T & \mathbf{Z}_2^T \frac{U_2 - u}{h} \\ \vdots & \vdots \\ \mathbf{Z}_n^T & \mathbf{Z}_n^T \frac{U_n - u}{h} \end{bmatrix}.$$

3. Asymptotic Property

First, we make the following assumptions. Let $\mu_i = \int_0^\infty t^i K(t) dt$, $\nu_i = \int_0^\infty t^i K^2(t) dt$.

Assumption 1. The errors $\varepsilon_i (i = 1, 2, \dots, n)$ are independent and $0 < E(\varepsilon_i^2) < \infty$ and $\text{Var}(\varepsilon_i^2) > 0$.

Assumption 2. The random variable U has a bounded support Π . Its density function $f(\cdot)$ is Lipschitz continuous and bounded away from 0 on its support.

Assumption 3. The $p \times p$ matrixes $E[\mathbf{X}\mathbf{X}^T|U]$ and $E[\mathbf{X}\mathbf{X}^T/\sigma^2(\mathbf{X}, U)|U]$ are non-singular for each $U \in \Pi$.

Assumption 4. There is an $s > 2$ such that $E \|\mathbf{X}\|^{2s} < \infty$ and for some $k < 2 - s^{-1}$ such that $n^{2k-1}h \rightarrow \infty$ as $n \rightarrow \infty$.

Assumption 5. $\{\alpha_j(\cdot), j = 1, \dots, p\}$ have continuous second derivatives in $U \in \Pi$.

Assumption 6. The function $K(\cdot)$ is a symmetric density function with compact support and the bandwidth h satisfies $nh^8 \rightarrow 0$ and $nh^2/(\log n)^2 \rightarrow \infty$ as $n \rightarrow \infty$.

For the estimator $\hat{\alpha}(u)$, Cai *et al.* [1] proved the following result:

Theorem 1 Under the assumptions 1–6, the estimator $\hat{\alpha}(u)$ is asymptotically normal, namely,

$$\sqrt{nh} \left(\hat{\alpha}(u) - \alpha(u) - \frac{1}{2} h^2 \mu_2 \alpha''(u) \right) \rightarrow N(\mathbf{0}, \nu_0 \Psi / f(u)),$$

where $\Psi = \Gamma(U)^{-1} E[\mathbf{X}\mathbf{X}^T \sigma^2(\mathbf{X}, U) | U] \Gamma(U)^{-1}$, $\Gamma(U) = E[\mathbf{X}\mathbf{X}^T | U]$.

For the estimator $\tilde{\alpha}(u)$, we obtain the following result by the Theorem 1 directly.

Theorem 2 Under the assumptions 1–6, the estimator $\tilde{\alpha}(u)$ is asymptotically normal, namely,

$$\sqrt{nh} \left(\tilde{\alpha}(u) - \alpha(u) - \frac{1}{2} h^2 \mu_2 \alpha''(u) \right) \rightarrow N(\mathbf{0}, \nu_0 \Phi^{-1} / f(u)),$$

where $\Phi = E(\mathbf{Z}\mathbf{Z}^T|U) = E(\mathbf{X}\mathbf{X}^T/\sigma^2(\mathbf{X}, U)|U)$.

Denote $\bar{\mathbf{X}} = (\bar{\mathbf{X}}_1, \dots, \bar{\mathbf{X}}_n)^T$, $\bar{\mathbf{X}}_i = \sqrt{K_h(U_i - u)}\mathbf{X}_i$, $\Sigma = \text{diag}\{\sigma^2(\mathbf{X}_1, U_1), \dots, \sigma^2(\mathbf{X}_n, U_n)\}$. By the proof of Theorem 1 in Cai *et al.* [1], we have

$$\frac{1}{n} \sum_{i=1}^n K_h(U_i - u) \mathbf{X}_i \mathbf{X}_i^T = \bar{\mathbf{X}}^T \bar{\mathbf{X}} \xrightarrow{p} E(\mathbf{X}\mathbf{X}^T|U)f(u),$$

$$\frac{1}{n} \sum_{i=1}^n K_h(U_i - u) \mathbf{X}_i \mathbf{X}_i^T \sigma^2(\mathbf{X}, U) = \bar{\mathbf{X}}^T \Sigma \bar{\mathbf{X}} \xrightarrow{p} E(\mathbf{X}\mathbf{X}^T \sigma^2(\mathbf{X}, U)|U)f(u),$$

and

$$\frac{1}{n} \sum_{i=1}^n K_h(U_i - u) \mathbf{X}_i \mathbf{X}_i^T / \sigma^2(\mathbf{X}, U) = \bar{\mathbf{X}}^T \Sigma^{-1} \bar{\mathbf{X}} \xrightarrow{p} E(\mathbf{X}\mathbf{X}^T / \sigma^2(\mathbf{X}, U)|U)f(u).$$

Since $(\bar{\mathbf{X}}^T \bar{\mathbf{X}})^{-1} \bar{\mathbf{X}}^T \Sigma \bar{\mathbf{X}} (\bar{\mathbf{X}}^T \bar{\mathbf{X}})^{-1} \geq (\bar{\mathbf{X}}^T \Sigma^{-1} \bar{\mathbf{X}})^{-1}$, then we have

$$\Psi \geq \Phi^{-1}.$$

This implies that $\tilde{\alpha}(u)$ is asymptotically more efficient than $\hat{\alpha}(u)$ in terms of asymptotic covariance matrix.

Remark 1. Since $\tilde{\alpha}(u)$ depends on the unknown parameters $\sigma^2(\mathbf{X}, U)$, it is infeasible. To provide a feasible efficient estimator of $\alpha(u)$, we need to estimate $\sigma^2(\mathbf{X}, U)$ consistently. It is not difficult to show that the resultant feasible estimator has the asymptotic property as $\tilde{\alpha}(u)$.

Remark 2. To obtain the consistent estimator of the variance function $\sigma^2(\mathbf{z}, u)$, it is important to model $\sigma^2(\mathbf{z}, u)$. Several kinds of variance function have been proposed. Discussion on the parametric variance function can be found in Carroll and Ruppert [11]. Muller and Stadtmuller [12], Chiou and Muller [13] and Ruppert *et al.* [14] studied nonparametric variance estimation. Muller and Zhao [15] proposed a general semiparametric variance function model in a fixed design regression setting. Keilegom and Wang [16] considered a general class of mean-variance regression models, in which both the mean function and the variance function were semiparametrically modeled. Zhu *et al.* [17] consider a single-index structure to study heteroscedasticity in a single-index regression model with high-dimensional predictors.

4. Simulation Studies

In this section we compare the behavior of the conventional estimator $\hat{\alpha}(u)$ with that of the new estimator $\tilde{\alpha}(u)$, given in (2.2) and (2.4), respectively, when the sample size is finite. The data are generated from the following varying coefficient model

$$y_i = x_i \alpha(u_i) + \sigma(x_i, u_i) \varepsilon_i, i = 1, 2, \dots, n, \quad (4.1)$$

where $x_i \sim N(0, 1)$, $u_i = i/n$, $\alpha(u_i) = u_i + \sin(2\pi u_i)$. Firstly, we consider the following four known variance functions:

$$(A) : \sigma(x_i, u_i) = e^{x_i}; \quad (B) : \sigma(x_i, u_i) = e^{u_i};$$

$$(C) : \sigma(x_i, u_i) = 1 + x_i; \quad (D) : \sigma(x_i, u_i) = 1 + u_i.$$

Secondly, we consider the case that the variance function is unknown. For simplicity, the variance function is assumed to have the following parametric structure,

$$(E) : \quad E(\varepsilon_i^2 | x_i, u_i) = \sigma^2(x_i, u_i) = \exp(\gamma_0 + \gamma_1 u_i),$$

with $\gamma_0 = 1, \gamma_1 = 2$. Obviously, we can build the following linear regression model

$$\ln \varepsilon_i^2 = \gamma_0 + \gamma_1 u_i + \xi_i, \quad (4.2)$$

with $E\xi_i = 0$. In practice, ε_i is not available, but it may be estimated by $\hat{\varepsilon}_i = y_i - x_i^T \hat{\alpha}(u_i)$, where $\hat{\alpha}(u_i)$ are the local linear estimates of model (4.1) without considering the heteroscedasticity structure. Applying the least squares approach to linear model (4.2) with ε_i was replaced by $\hat{\varepsilon}_i$, we can obtain the estimators of γ_0, γ_1 , denoted by $\hat{\gamma}_0$ and $\hat{\gamma}_1$ respectively. Accordingly, we get the estimator of $\sigma^2(\mathbf{x}_i, \mathbf{z}_i, u_i)$, as $\hat{\sigma}^2(\mathbf{x}_i, \mathbf{z}_i, u_i) = \exp(\hat{\gamma}_0 + \hat{\gamma}_1 u_i)$.

To study the effect of the distribution of the error for our method, we take the following three different types of the error distribution, (1) $\varepsilon_i \sim N(0, 0.5^2)$, (2) $\varepsilon_i \sim U(-\sqrt{3}/2, \sqrt{3}/2)$, (3) $\varepsilon_i \sim \frac{1}{8}\chi_8^2 - 1$. The Gauss kernel function and $h = n^{-1/5}$ are used in our simulation studies.

We compare the proposed efficient estimator $\tilde{\alpha}(u)$ with that of the ordinary local linear estimator $\hat{\alpha}(u)$ by using the estimated mean average squared error (MASE),

$$\text{MASE}\{\hat{\alpha}(\cdot)\} = \frac{1}{1000 * n} \sum_{l=1}^N \sum_{i=1}^n [\hat{\alpha}_j^l(u_i) - \alpha_j(u_i)]^2,$$

where $\hat{\alpha}^l(u_i), l = 1, 2, \dots, N$, are the estimate of the coefficient $\alpha(u_i)$ in $N = 1000$ replications. The simulation results are presented in Table 1, and for all the scenarios we studied, the proposed efficient estimators outperform the ordinary local linear estimators.

Table 1. Mean average squared error (MASE) index for the estimators of varying coefficients.

| Variance Function | Sample n | $N(0, 0.5^2)$ | | $U(-\sqrt{3}/2, \sqrt{3}/2)$ | | $\frac{1}{8}\chi_8^2 - 1$ | |
|----------------------|---------------|-----------------------|-------------------------|------------------------------|-------------------------|---------------------------|-------------------------|
| | | $\hat{\alpha}(\cdot)$ | $\tilde{\alpha}(\cdot)$ | $\hat{\alpha}(\cdot)$ | $\tilde{\alpha}(\cdot)$ | $\hat{\alpha}(\cdot)$ | $\tilde{\alpha}(\cdot)$ |
| A | 30 | 0.4873 | 0.1354 | 0.5125 | 0.1277 | 0.5109 | 0.1374 |
| | 50 | 0.4063 | 0.0805 | 0.3820 | 0.0804 | 0.3686 | 0.0805 |
| | 80 | 0.3146 | 0.0539 | 0.3054 | 0.0523 | 0.2910 | 0.0525 |
| B | 30 | 0.2018 | 0.1936 | 0.1872 | 0.1791 | 0.1835 | 0.1752 |
| | 50 | 0.1129 | 0.1077 | 0.1174 | 0.1126 | 0.1114 | 0.1072 |
| | 80 | 0.0728 | 0.0696 | 0.0734 | 0.0706 | 0.0733 | 0.0706 |
| C | 30 | 0.1469 | 0.1302 | 0.1502 | 0.1269 | 0.1571 | 0.1304 |
| | 50 | 0.1024 | 0.0920 | 0.1039 | 0.0893 | 0.1015 | 0.0884 |
| | 80 | 0.0670 | 0.0679 | 0.0731 | 0.0660 | 0.0734 | 0.0645 |
| D | 30 | 0.1446 | 0.1422 | 0.1390 | 0.1368 | 0.1487 | 0.1474 |
| | 50 | 0.0841 | 0.0831 | 0.0844 | 0.0832 | 0.0819 | 0.0807 |
| | 80 | 0.0612 | 0.0606 | 0.0556 | 0.0549 | 0.0595 | 0.0589 |
| E | 30 | 0.3950 | 0.3936 | 0.4227 | 0.4133 | 0.3996 | 0.3913 |
| | 50 | 0.3094 | 0.3039 | 0.3137 | 0.3068 | 0.3077 | 0.3015 |
| | 80 | 0.2549 | 0.2508 | 0.2513 | 0.2508 | 0.2544 | 0.2520 |

5. Conclusions

In this paper, we focus on the estimation problem of the varying coefficient model with heteroscedastic errors. Based on the local linear method, we develop a simple approach to estimate the nonparametric coefficient functions by taking the estimated error heteroscedasticity into account. The resulting estimators are shown to have smaller asymptotic variances than the conventional local-linear estimators. The asymptotic normality of the proposed estimator is established. Furthermore, some simulation experiments are performed to evaluate the finite sample behaviors of the proposed estimators.

Acknowledgments

We are grateful to two anonymous referees for helpful comments on this paper. Chuanhua Wei's research was supported by the National Natural Science Foundation of China (No.11301565), and Beijing Higher Education Young Elite Teacher Project(No.YETP1316).

Author Contributions

The authors contributed equally to this work.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Cai, Z.W.; Fan, J.Q.; Yao, Q.W. Functional-coefficient regression models for nonlinear times series. *J. Am. Stat. Assoc.* **2000**, *95*, 941–956.
2. Cai, Z.W. Trending time-varying coefficient time series models with serially correlated errors. *J. Econom.* **2007**, *136*, 163–188.
3. Cai, Z.W.; Li, Q. Nonparametric Estimation of Varying Coefficient Dynamic Panel Data Models. *Econom. Theory* **2008**, *24*, 1321–1342.
4. Cai, Z.W.; Li, Q.; Park, J.Y. Functional-coefficient models for nonstationary time series data. *J. Econom.* **2009**, *148*, 101–113.
5. Fan, J.Q.; Zhang, J.T. Functional linear models for longitudinal data. *J. R. Stat. Soc. Ser. B* **2000**, *62*, 303–322.
6. Fan, J.Q.; Jiang, J.C.; Zhang, C.; Zhou, Z. Time-dependent diffusion models for term structure dynamics and the stock price volatility. *Stat. Sin.* **2003**, *13*, 965–992.
7. Fotheringham, A.S.; Charlton, M.; Brunson, C. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*; Wiley: New York, NY, USA, 2002.
8. Hoover, D.R.; Rice, J.A.; Wu, C.O.; Yang, L.P. Nonparametric smoothing estimation of time-varying coefficient models with longitudinal data. *Biometrika* **1998**, *85*, 809–822.
9. Li, Q.; Huang, C.J.; Li, D.; Fu, T.T. Semiparametric smooth coefficient models. *J. Bus. Econ. Stat.* **2002**, *20*, 412–422.
10. Xiao, Z. Functional-coefficient cointegration models. *J. Econ.* **2009**, *152*, 81–92.

11. Carroll, R.J.; Ruppert, D. *Transformation and Weighting in Regression*; Chapman and Hall: London, UK, 1988.
12. Muller, H.G.; Stadtmuller, U. Estimation of Heteroscedasticity in Regression Analysis. *Ann. Stat.* **1987**, *15*, 610–625.
13. Chiou, J.M.; Muller, H.G. Nonparametric quasi-likelihood. *Ann. Stat.* **1999**, *27*, 36–64.
14. Muller, H.G.; Zhao, P.L. On a semiparametric variance function model and a test for heteroscedasticity. *Ann. Stat.* **1995**, *23*, 946–967.
15. Ruppert, D.; Wand, M.P.; Host, U.; Hossjer, O. Local polynomial variance-function estimation. *Technometrics* **1997**, *39*, 262–273.
16. Keilegom, I.V.; Wang, L. Semiparametric modeling and estimation of heteroscedasticity in regression analysis of cross-sectional data. *Electron. J. Stat.* **2010**, *4*, 133–160.
17. Zhu, L.; Dong, Y.; Li, R. Semiparametric estimation of conditional heteroscedasticity through single index modeling. *Stat. Sin.* **2013**, *24*, 1235–1256.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).