JETEM

Journal of Economic Theory and Econometrics

The Journal of the Korean Econometric Society

( 계 량 경 제 학 보 )

March 2010

Volume 21 Issue 1. pp.23-42

# Local Linear Estimation of Nonparametric Cointegrating Regression[*]

**Jihyun Kim**[†], **Chang Sik Kim**[‡]

**Abstract** Nonparametric estimation of a nonlinear cointegrating regression model is stud-ied. This paper offers the asymptotic theory of the local linear estimator for a nonlinear cointegrating regression model. It is shown that the local linear estimator has the same asymptotic distribution as the Nadaraya-Watson estimator in a nonparametric cointegrat-ing regression, and the asymptotic orders for the biases of the two nonparametric estimators will be provided. Monte Carlo simulation shows that the finite sample performances of the local linear estimator outperform those of Nadaraya-Watson estimator.

**Keywords** Local Time, Cointegration, Nonparametric Regression, Nadaraya-Watson Es-timation, Local Linear Estimation

**JEL Classification** C14, C22

[*] We would like to thank three anonymous referees for many helpful discussions and suggestions. This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD), Basic Research Promotion Fund 2007-B00042.

[†] School of Economics, Sungkyunkwan University, E-mail address: qup33@skku.edu.

[‡] Associate Professor, School of Economics, Sungkyunkwan University, E-mail address: skimcs@skku.edu

*Received April 24, 2009, Revised August 11, 2009, Accepted October 6, 2009*

# 1. Introduction

Since the concept of cointegration was introduced by the seminal paper of Engle and Granger (1987), there have been numerous studies in various fields of econometrics. Even though the linear cointegration model is convenient for practical implementations and theoretical developments based on the linear framework of partial summation, it is argued that the linear models with integrated processes appear to be too restrictive to be much useful in many applications. In fact, many economic theories often suggest nonlinear responses without being specific regarding functional form, and many economic time series show only locally bounded or nonstationary. There would be no doubt that more flexible tools would be available for modeling economic and financial variables in nonlinear framework. On the other hand, the flexibility of nonlinear functions can be nicely equipped with nonparametric methods.

The most popular and commonly used estimation procedure in nonparametric regression models is the Nadaraya-Watson estimator that was introduced originally by Nadaraya (1964) and Watson (1964), and after about a decade, Stone (1977) introduced the local linear fitting approach. The nonparametric estimation methods have been used successfully to various fields such as economics and finance due to their advantage of requiring little prior information on the data generating process; see, for example, Pagan and Ullah (1999) and Li and Racine (2006) for the various applications in economics and finance.

For stationary time series data, the theories of nonparametric estimations and inferences are well developed and the methods are widely used in practice. It is also well known that the local linear estimator has significant advantages over the Nadaraya-Watson regression estimator. It has been shown that the local linear estimator reduces the bias and mean squared error; when the underlying probability density has a compact support in stationary time series.

The nonstationary nonparametric models, however, are rather undeveloped, and some theoretical aspects of the model have been explored very recently. Karlsen, Myklebust and Tjøstheim (2007, hereafter KMT) and Wang and Phillips (2008a, hereafter WP) provided asymptotic theories for the Nadaraya-Watson estimator of the following nonparametric cointegrating regression model:

$$y_t = m(x_t) + u_t, \quad t = 1, 2, ..., n, \tag{1}$$

where $x_t$ is a nonstationary regressor. Let $K(x)$ be a non-negative real function and write $K_h(s) = \frac{1}{h}K(s/h)$ where $h \equiv h_n \to 0$. The Nadaraya-Watson kernel estimator of $m(x)$ in model (1) is given by

$$\hat{m}_{NW}(x) = \frac{\sum_{t=1}^{n} y_t K_h(x_t - x)}{\sum_{t=1}^{n} K_h(x_t - x)}.$$

KMT investigated $\hat{m}_{NW}(x)$ in a nonstationary cointegrating regression model where $x_t$ is a recurrent Markov chain; and WP considered an alternative treatment by making use of local time limit theory and, instead of recurrent Markov chains, worked with partial sum representations of the type $x_t = \sum_{j=1}^{t} \nu_j$ where $\nu_j$ is a general linear process. More recently,* Wang and Phillips (2008b, 2009) further investigated the nonparametric estimation of a structural regression model with a nonstationary regressor. They showed that the limit theory for $\hat{m}_{NW}(x)$ has links to traditional nonparametric asymptotics for stationary models even though the rates of convergence are different and typically slower when $x_t$ is nonstationary. They established that the nonparametric estimation has much slower rate of convergence, and hence the finite sample performances seem poor despite the flexibility of the nonparametric formulations with nonstationary time series.

The reduction in convergence rate of nonparametric estimation with nonstationary processes can not be recovered with a different choice of higher order kernel functions. For the parametric approach, the wandering characteristics, however, increase the signal for some classes of functions if the regression function is increasing, and the convergence rate is given by the behavior of the function at infinity. However, this is not the case for the nonparametric estimation in which the regression function can only be evaluated locally. The nonparametric estimation would not take the effect of magnifying signals generated by the nonstationary regressors into account, and therefore the convergence rates will be reduced accordingly.

The present paper has a similar goal to KMT and WP but differs in the sense that it offers an asymptotic theory of the local linear estimator. While KMT uses the framework of null recurrent Markov chains, we use a direct local time density argument that suggested

*The first version of this paper was written independently of these recent works on nonparametric estimation of regression models with nonstationary covariates.

by WP. This method makes the approach more closely related to conventional nonparametric arguments. In a nonparametric cointegrating regression, we find that the local linear estimator has the same rate of convergence and variance as the Nadaraya-Watson estimator in WP. The bias term, however, has a different form when we compare it with that of the Nadaraya-Watson estimator. In fact, the linear term is eliminated from the asymptotic bias as shown in Wang and Phillips (2009c). In consequence and in contrast to the stationary case, the Nadaraya-Watson estimator has the same limit distribution (to the second order including bias) as the local linear nonparametric estimator. However, we closely analyze the bias term and derive the explicit orders of bias, and show that the local linear estimator leads to a better performance in estimating nonparametric cointegrating regression. We also investigated the finite sample performances of the Nadaraya-Watson estimator and the local linear estimator. It provides clearer evidence to show the advantages of the local linear estimator over the Nadaraya-Watson estimator.

This paper is organized as follows. Section 2 provides the model and assumptions. Section 3 presents our main results on the consistency and the limit distribution of the local linear estimator in a nonlinear cointegrating regression model. Section 4 reports a simulation experiment exploring the finite sample properties of the Nadaraya-Watson estimator and the local linear estimator. Section 5 concludes by discussing these results and some possible extensions. Proofs are given in Section 6 as Appendices to the paper.

## 2. Model and Assumptions

We consider the nonlinear cointegrating regression model given in (1), and the regression error $u_t$ is assumed to be a zero mean stationary error and $x_t$ is a nonstationary regressor generated by

$$x_t = x_{t-1} + \nu_t$$

where $\nu_t$ is assumed to be stationary and satisfies the invariance principle as in the assumptions below. $m$ is assumed to be a unknown function to be estimated with the observed data $\{y_t, x_t\}_{t=1}^n$.

The local linear regression smoothing can be obtained by expanding $m(x_t)$ around $x$ to get

$$\begin{aligned} y_t &\simeq m(x) + \frac{\partial m}{\partial x}(x)(x_t - x) + \frac{\partial^2 m}{\partial x^2}(x^*)\frac{(x_t - x)^2}{2} + u_t \\ &= z_t^{'}\delta + \frac{\partial^2 m}{\partial x^2}(x^*)\frac{(x_t - x)^2}{2} + u_t, \quad t = 1, 2, ..., n \end{aligned} \tag{2}$$

where $z_t^{'} = (1,\ x_t - x)$, $\delta^{'} = \left(m(x),\ \frac{\partial m}{\partial x}(x^*)\right)$ and $x^* \in [x_t, x]$. Then, the local linear regression estimator performs a weighted regression of $y_t$ against $z_t^{'}$ using weights $[K_h(x_t - x)]^{1/2}$, and hence the local linear estimator of $m(x)$ in model (2) is given by

$$\hat{m}_{LL}(x) = e_1^{'}\left(\sum_{t=1}^{n} z_t K_h(x_t - x) z_t^{'}\right)^{-1}\left(\sum_{t=1}^{n} z_t K_h(x_t - x) y_t\right), \tag{3}$$

where $e_1^{'} = (1,\ 0)$.

It is well established that the Nadaraya-Watson estimator and the local linear estimator converge at rate $\sqrt{nh}$, where $h \to 0$ in the traditional nonparametric asymptotics for regressions with stationary time series, whereas the Nadaraya-Watson estimator in nonparametric cointegrating regressions converges at rate $\left(nh^2\right)^{1/4}$ as shown in KMT and WP. To develop the asymptotic theory of the local linear estimator, let $\{\epsilon_j, -\infty < j < \infty\}$ be a sequence of iid random variables with $\mathbb{E}\epsilon_0 = 0$, $\mathbb{E}\epsilon_0^2 = 1$ and characteristic function $\varphi(t)$ of $\epsilon_0$ satisfying $\int_{-\infty}^{\infty} |\varphi(t)| dt < \infty$. In addition, we use the following assumptions in the asymptotic development.

**Assumption 1.** *The kernel $K(\cdot)$ is a symmetric and nonnegative density with integrable characteristic function $\varphi_K$ and satisfies the following conditions:*

$$\int_{-\infty}^{\infty} K(s)\, ds = 1,\ \int_{-\infty}^{\infty} K^2(s)\, ds < \infty,\ \int_{-\infty}^{\infty} s^2 K(s)\, ds < \infty,\ \sup_s K(s) < \infty.$$

**Assumption 2.** *For given $x$, $m(x)$ has a continuous, bounded second derivative in a neighborhood of $x$.*

**Assumption 3.** *$(u_t, \mathcal{F}_t, 1 \le l \le n)$ is a martingale difference with $\mathbb{E}\left(u_t^2|\mathcal{F}_{t-1}\right) \to_{a.s.} \sigma^2 > 0$ as $t \to \infty$ and $\sup_{1\le t\le n} \mathbb{E}\left(|u_t|^q|\mathcal{F}_{t-1}\right) < \infty$ a.s. for some $q > 2$.*

**Assumption 4.** *$x_t = \sum_{j=1}^{t} \nu_j$ where $\nu_j = \sum_{k=0}^{\infty} \phi_k \epsilon_{j-k}$ with $\sum_{k=0}^{\infty} |\phi_k| < \infty$ and $\phi(1) \equiv \sum_{k=0}^{\infty} \phi_k \neq 0$.*

The conditions in Assumption 1 are quite weak and simply verified for various kernels $K(x)$, for instance, if $K(x)$ is a standard normal kernel or has a compact support as in KMT. Assumption 2 includes a wide range of regression functions $m(x)$. Thus, commonly occurring functions like $m(x) = x^{\alpha}$ and $m(x) = 1/(1 + x^{\alpha})$ for some $\alpha > 0$ satisfy the Assumption 2. Assumption 3 is a standard condition for error processes in time series regression models. In the asymptotic development below, $x_t$ is a partial sum of linear process innovations that satisfy a simple summability condition with long run moving average coefficient $\phi(1) \neq 0$ as stated in Assumption 4.

## 3. Local Linear Estimation

In this section, we present that the limit theory for the local linear estimator $\hat{m}_{LL}(x)$ in the nonparametric cointegrating regression.

**Theorem 1.** *Let Assumptions 1–4 hold. Suppose that $(\nu_t)_1^n$ is independent of $(u_t)_1^n$. Then, for any $h$ satisfying $nh^2 \to \infty$ and $nh^{10} \to 0$,*

$$\left(h \sum_{t=1}^{n} K_h(x_t - x)\right)^{1/2} (\hat{m}_{LL}(x) - m(x)) \to_d N\left(0, \sigma_1^2\right), \tag{4}$$

*where $\sigma_1^2 = \sigma^2 \int_{-\infty}^{\infty} K^2(s)\, ds$.*

Our next theorem considers the effect of some relaxation of the restriction on the independence between $\nu_t$ and $u_t$. To do so, denote the stochastic processes $U_n$ and $V_n$ on $D[0,1]$ by

$$U_n(r) = x_{[n,r],n} \quad \text{and} \quad V_n(r) = \frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} u_t,$$

where $x_{k,n} = x_k/\sqrt{n}$, $0 \leq k \leq n, n \geq 1$ (define $x_{0,n} \equiv 0$) is a random triangular array and $[a]$ denotes the integer part of $a$.

**Theorem 2.** *Suppose Assumptions 1-4 hold. Suppose that, for each $n \geq 2$, $x_{i,n}$ is adapted to $\mathcal{F}_{i-1}$, $2 \leq i \leq n$, and $(U_n, V_n) \to_d (U, V)$ on $D[0,1]^2$ as $n \to 0$, where $(U, V)$ is a standard vector Brownian motion. Then, (4) still holds true for any $h$ satisfying $nh^2 \to \infty$ and $nh^{10} \to 0$.*

Theorems 1 and 2 show that the asymptotic distribution of the local linear estimator is normal and the rate of convergence is same as the Nadaraya-Watson estimator of WP. Also, it is clear from the proof of theorems that $\hat{m}_{LL}(x) \to_p m(x)$ for any $h$ satisfying $h \to 0$ and $nh^2 \to \infty$. The bandwidth $h$ has to satisfy a certain rate of conditions to ensure the stated asymptotic normality. In case of stationary nonparametric models, the convergence rate of a local linear regression estimator is $\sqrt{nh}$ requiring that $nh \to \infty$. Undersmoothing in stationary models to remove the asymptotic bias requires that $nh^5 \to 0$. In the nonstationary model, the corresponding rate in such regressions is now $\sqrt[4]{nh^2}$, which requires that $nh^2 \to \infty$. Undersmoothing to avoid the asymptotic bias in this nonstationary model requires that $nh^{10} \to 0$. The conditions $nh^2 \to \infty$ and $nh^{10} \to 0$ in the theorem require that $h$ should converge to zero faster than $n^{-1/10}$ but not as fast as $n^{-1/2}$. This is mainly because an integrated regressor has a stochastic trend and hence is observed less around any fixed point for a certain class of function. Therefore, it is naturally expected that a smaller rate should be required for enough observations to remove the asymptotic bias in nonstationary cointegrating regression.

Interestingly, the bandwidth of the local linear estimator needs a certain rate of conditions to remove the asymptotic bias that are different from the rate of bandwidth for the Nadaraya-Watson estimator. Undersmoothing to remove the asymptotic bias of the Nadaraya-Watson estimator of WP requires that $nh^6 \to 0$. However, the range for the bandwidth $h$ is improved by employing the local linear estimator. This can be explained as follows. Bias of the local linear estimator involves only the quadratic term of the Taylor expansion, so that the corresponding rate is $nh^{10} \to 0$. On the other hand, the bias of the Nadaraya-Watson estimator has an additional term which is $O_p(h)$ when function $m$ satisfy certain conditions as in WP. Wang and Phillips (2009c) provides related results for such a model with the Nadaraya-Watson estimator as well as local linear estimator. Specifically, Wang and Phillips (2009c) improves the range for the bandwidth by adding an explicit bias term in the limit theory and finds that the rate $h = o_p\left(n^{-1/14}\right)$ is sufficient for undersmoothing. The discussion regarding bias stated above is investigated more clearly in the following Theorem 3.

Theorem 2 shows that the local linear estimator can be used to estimate a nonparametric unit root autoregressive model

$$y_t = m\left(y_{t-1}\right) + u_t, \;\; m\left(y_{t-1}\right) = \alpha y_{t-1} \; a.s. \quad t = 1, 2, ..., n$$

with $\alpha = 1$ and $y_0 = 0$. In this model, $x_t = y_{t-1} = \sum_{i=1}^{t-1} u_i$. Then, the local linear estimator of an autoregressive parameter in case of a unit root is asymptotically normal upon the usual standardization. As demonstrated in Phillips and Park (1998), Guerre (2004) and WP, but the proper convergence rate is much slower than the one needed in nonparametric autoregression with stationary process.

In the following theorem, we derive the explicit orders for the bias terms of the local linear estimator ($\hat{m}_{LL}$) and the Nadaraya-Watson estimator ($\hat{m}_{NW}$).

**Theorem 3.** *Under Assumptions 1–4, we have*

$$Bias\left(\hat{m}_{LL}\right) = \frac{h^2}{2}\mu_2\left(K\right) m''\left(x^*\right) \tag{5}$$

$$Bias\left(\hat{m}_{NW}\right) = O_p\left(n^{-1/4}h^{1/2}\right) + \frac{h^2}{2}\mu_2\left(K\right) m''\left(x^*\right) \tag{6}$$

*where* $\mu_2\left(K\right) = \int_{-\infty}^{\infty} s^2 K\left(s\right) ds$.

Theorem 3 shows that two estimators have different biases. Specifically, the bias of the Nadaraya-Watson estimator has an additional term compared to that of the local linear estimator. The first bias term of the Nadaraya-Watson estimator is dominated by the second term as $n \to \infty$ provided $nh^6 \to \infty$ since we can write the bias of the Nadaraya-Watson estimator as $O_p\left(n^{-1/4}h^{1/2}\right) + O_p\left(h^2\right) = h^2\left(\frac{1}{\sqrt[4]{nh^6}}O_p\left(1\right) + O_p\left(1\right)\right)$. Theorem 3 shows that the main difference of our paper to WP. WP shows that both estimators have the same limit distributions whereas Theorem 3 explicitly shows that both estimators have different rates of bias reductions with different choices of bandwidth parameters. This theoretical result can be confirmed by simulation in the following chapter.

It is interesting to notice that the relationship between the local linear estimator and the Nadaraya-Watson estimator in a nonparametric cointegrating regression has a similar aspect in stationary regression models in the sense that those two estimators in the stationary models have the same variance, convergence rate and limit distribution but different bias as summarized in Table 1. However, both the linear term and the quadratic term in the bias of the stationary models have the same asymptotic orders whereas the first term in the bias of the Nadaraya-Watson estimator becomes negligible provided $nh^6 \to \infty$. Table 2 shows the summary of these results.

## 4. Simulations

In this section, we investigate the finite sample properties of the local linear estimator and the Nadaraya-Watson estimator. The data generating process follows (1) and has the form

$$y_t = m(x_t) + u_t, \quad \triangle x_t = \nu_t,$$

where $(\nu_t, u_t)$ are $iidN\left(0, \sigma^2 I_2\right)$. The following regression function was used in our simulations:

$$m(x) = \sum_{j=1}^{4} \frac{(-1)^{j+1} \sin(j\pi x)}{j^2}, \text{ for } [0,1] \text{ and } [-2,2].$$

The function corresponds to the function that has been commonly used in the literature, for example, Hall and Horowitz (2005) and Wang and Phillips (2009b). Simulations were performed for the sample size $n = 500$, with $\sigma = 0.1$ and 10,000 replications. A Gaussian kernel was used with bandwidths $h = n^{-1/2}$, $n^{-1/3}$, $n^{-1/5}$. Bias, variance and mean squared error for the estimates of $m(x)$ were computed on the grid of values $\{x = 0.01k \mid k = 0, 1, ..., 100\}$ for $[0, 1]$ and $\{x = -2 + 0.04k \mid k = 0, 1, ..., 100\}$ for $[-2, 2]$.[†]

Table 3 shows the finite sample performances of both the Nadaraya-Watson estimator $(\hat{m}_{NW}(x))$ and the local linear estimator $(\hat{m}_{LL}(x))$ computed over the intervals $[0, 1]$ and $[-2, 2]$. The bias is reduced for smaller bandwidths at the cost of some increase in dispersion as theory predicts. Finite sample performance in terms of MSE seems to be optimized at $n^{-1/2}$ although this choice violates the condition $nh^2 \to \infty$ of Theorem 1. More importantly, it is clear that the local linear estimator has smaller bias and MSE than the Nadaraya-Watson estimator in every case. We also computed the integrated mean squared error to investigate the finite sample performance of two estimators to limit the boundary effects of nonparametric estimations. The IMSE is computed by

$$IMSE = \int_{-\infty}^{\infty} \mathbb{E}\left((\hat{m}(x) - m(x))^2\right) w(x)\, dx$$

[†]The simluation scheme is almost the same as Wang and Phillips (2008b). Other regression functions like power functions were also used for the simulations, but results are very similar. The finite sample performances for other cases were skipped in the paper to save space.

where $w(x) = I_{[0.1,0.9]}(x), I_{[-1.6,1.6]}(x)$ with $I_A(x)$ denoting an indicator function. According to Table 3, the local linear estimator is superior to the Nadaraya-Watson estimator not only in terms of MSE, but also IMSE.

Figures 1-2 show the results for the Monte Carlo approximations to $\mathbb{E}(\hat{m}_{NW}(x))$ (dotted) and $\mathbb{E}(\hat{m}_{LL}(x))$ (dashes) for $m(x)$ (solid) over the various intervals corresponding to bandwidth $h = n^{-1/5}$. Figures show that the local linear estimator is fitted better than the Nadaraya-Watson estimator particularly for larger bandwidth or near the boundary of the interval. The difference between local linear estimator and the Nadaraya-Watson estimator can be clearly seen in the figures especially near the boundary of the interval. They have different rate of bias reduction in finite sample, which can be naturally expected as the theory predicts.

## 5. Concluding Remarks

In this paper, we provide an asymptotic theory of the local linear estimator in a nonparametric cointegrating regression model. The three main results in this paper have meaningful implications. First, the local linear estimator of a nonparametric cointegrating model is straightforward in practice just like the Nadaraya-Watson estimator. The local linear estimator is consistent and has a asymptotic normal distribution, thereby validating conventional methods of inference in the nonstationary nonparametric setting. Second, the local linear estimator has the same rate of convergence and variance as the Nadaraya-Watson estimator as shown in WP. Finally, the bias of the Nadaraya-Watson estimator has two terms whereas the local linear estimator has only one bias term. However, the bias of the Nadaraya-Watson estimator has an additional term which vanishes as $n \to \infty$ provided $nh^6 \to \infty$, which implies the linear term is eliminated from the asymptotic bias. Therefore, the local linear nonparametric estimator shares the same limit distribution with the Nadaraya-Watson estimator as long as bandwidth restrictions are met.

To demonstrate the finite sample performance, we examine the Monte Carlo simulations in various settings. We evaluate the local linear estimator and the Nadaraya-Watson over two different domains with three different bandwidths and compare the finite sample properties of the two estimators. The simulation results demonstrate that the finite sample performance of the local linear estimator is quite better than the Nadaraya-Watson estimator in terms of bias, MSE and IMSE.

## References


Engle, R.F. and Granger, C.W.J. (1987). Cointegration and Error Correction: Representation, Estimation and Testing. *Econometrica*, 55, 251-276.

Granger, C.W.J. (1987). Cointegrated variables and error-correcting models. *UCSD discussion paper*, Univ. California, San Diego.

Guerre, E. (2004). Design-Adaptive pointwise nonparametric regression estimation for recurrent Markov time series. *Unpublished manuscript.*

Hall, P.; Heyde, C. C. (1980). *Martingale limit theory and its application*, Probability and Mathematical Statistics. Academic Press, Inc.

Hall, P. and J. L. Horowitz (2005). Nonparametric methods for inference in the presence of instrumental variables. *Annals of Statistics*, 33, 2904-2929.

Karlsen, H. A., Myklebust, T. and Tjøstheim, D. (2007). Nonparametric estimation in a nonlinear cointegration model, *Annals of Statistics*, 35, 252-299.

Li, Q., Racine, J. (2006). *Nonparametric Econometrics: Theory and Practice.* Princeton Univ. Press.

Nadaraya, E.A. (1964). On estimating regression. *Theory of Probability and its Applications*, 9, 141-142.

Pagan, A. and Ullah, A. (1999). *Nonparametric Econometrics* Cambridge Univ. Press.

Park, J. Y. and P. C. B. Phillips (2001). Nonlinear regressions with integrated time series, *Econometrica*, 69, 117-161.

Phillips, P. C. B. and J. Y. Park (1998). Nonstationary Density Estimation and Kernel Autoregression. *Cowles Foundation discuss paper No. 1181.*

Stone, C.J. (1977). Consistent nonparametric regression. *Annals of Statistics* 5. 595-620.

Wang, Q. and P. C. B. Phillips (2009a). Asymptotic theory for local time density estimation and nonparametric cointegrating regression. *Econometric Theory*, 25,1-29.

Wang, Q. and P. C. B. Phillips (2009b). Structural Nonparametric Cointegrating Regression. *Econometrica* (forthcoming).

Wang, Q. and P. C. B. Phillips (2009c). Asymptotic Theory for Zero Energy Density Estimation with Nonparametric Regression Applications. *Cowles foundation discussion paper* NO. 1687. 143.164.

Watson, G.S. (1964). Smooth regression analysis. *Sankhyā*. 26. 101-116.

## Appendix: Proofs

**Proof of Theorem 1.** In order to prove (4), we split $\hat{m}_{LL}(x) - m(x)$ as

$$\begin{aligned}
&\hat{m}_{LL}(x) - m(x) \\
=&e_1^{'}\left(\sum_{t=1}^{n} z_t K_h(x_t - x) z_t^{'}\right)^{-1}\left(\sum_{t=1}^{n} z_t K_h(x_t - x) u_t\right) \\
&+ e_1^{'}\left(\sum_{t=1}^{n} z_t K_h(x_t - x) z_t^{'}\right)^{-1}\left(\sum_{t=1}^{n} z_t K_h(x_t - x)\left(\frac{\partial^2 m}{\partial x^2}(x^*)\frac{(x_t - x)^2}{2}\right)\right). \qquad (7)
\end{aligned}$$

It is readily seen that

$$\left(h\sum_{t=1}^{n} K_h(x_t - x)\right)^{1/2}\left(\hat{m}_{LL}(x) - m(x)\right) = \sum_{t=1}^{n} Y_{nt}u_t + \Theta_{2n}\Theta_{1n}, \qquad (8)$$

where $Y_{nt} = \Theta_{2n}\left(\frac{\sqrt{n}}{nh}\right)^{\frac{1}{2}} z_{nt}\frac{1}{h}K\left(\frac{\sqrt{n}x_{t,n}-x}{h}\right)$, $\Theta_{1n} = \left(\frac{\sqrt{n}}{nh}\right)^{\frac{1}{2}}\sum_{t=1}^{n} z_{nt}\frac{1}{h}K\left(\frac{\sqrt{n}x_{t,n}-x}{h}\right)\left(\frac{\left(\sqrt{n}x_{t,n}-x\right)^2 m''(x^*)}{2}\right)$ and $\Theta_{2n} = e_1^{'}\left(\frac{\sqrt{n}}{nh}\sum_{t=1}^{n} K\left(\frac{\sqrt{n}x_{t,n}-x}{h}\right)\right)^{1/2}\left(\frac{\sqrt{n}}{nh}\sum_{t=1}^{n} z_{nt}\frac{1}{h}K\left(\frac{\sqrt{n}x_{t,n}-x}{h}\right)z_{nt}^{'}\right)^{-1}$ with $z_{nt}^{'} = (1, \sqrt{n}x_{t,n} - x)$.

Now, to prove (4), it suffices to show that, for any $h$ satisfying $nh^{10} \to 0$ and $nh^2 \to \infty$,

$$\Theta_{2n}\Theta_{1n} \to_p 0, \qquad (9)$$

and

$$V_n \equiv \frac{1}{\Lambda_n}\sum_{t=1}^{n} u_t Y_{nt} \to_d N\left(0, \sigma^2\right), \qquad (10)$$

where $\Lambda_n^2 = \sum_{t=1}^n Y_{nt}Y_{nt}^{'}$. Indeed it follows from Theorem 2.1 of Wang and Phillips (2009a) and hence we have

$$\begin{aligned}
\Lambda_n^2 &= \sum_{t=1}^n Y_{nt}Y_{nt}^{'} \\
&= e_1^{'} A_0^{\frac{1}{2}} \begin{pmatrix} \frac{1}{h}A_0 & A_1 \\ A_1 & hA_2 \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{h^2}B_0 & \frac{1}{h}B_1 \\ \frac{1}{h}B_1 & B_2 \end{pmatrix} \begin{pmatrix} \frac{1}{h}A_0 & A_1 \\ A_1 & hA_2 \end{pmatrix}^{-1} A_0^{\frac{1}{2}} e_1 \\
&= \frac{A_0A_2^2B_0 - 2A_0A_1A_2B_1 + A_0A_1^2B_2}{\left(A_0A_2 - A_1^2\right)^2} \to_p \int_{-\infty}^{\infty} K^2\left(s\right) ds
\end{aligned}$$

and

$$\begin{aligned}
\Theta_{2n}\Theta_{1n} &= e_1^{'} \left(\frac{nh}{\sqrt{n}}\right)^{\frac{1}{2}} A_0^{\frac{1}{2}} \begin{pmatrix} \frac{1}{h}A_0 & A_1 \\ A_1 & hA_2 \end{pmatrix}^{-1} \begin{pmatrix} hA_2 \\ h^2A_3 \end{pmatrix} \frac{m^{''}\left(x^*\right)}{2} \\
&= h^2 \left(\frac{nh}{\sqrt{n}}\right)^{\frac{1}{2}} A_0^{\frac{1}{2}} \frac{A_2^2 - A_1A_3}{A_0A_2 - A_1^2} \frac{m^{''}\left(x^*\right)}{2} \to_p 0
\end{aligned}$$

where

$$A_i = \frac{\sqrt{n}}{nh} \sum_{t=1}^n \left(\frac{\sqrt{n}x_{t,n} - x}{h}\right)^i K\left(\frac{\sqrt{n}x_{t,n} - x}{h}\right) \to_p \begin{cases} L_G\left(1,0\right) \int_{-\infty}^{\infty} s^i K\left(s\right), & i \text{ is even;} \\ 0, & i \text{ is odd,} \end{cases}$$

and

$$B_i = \frac{\sqrt{n}}{nh} \sum_{t=1}^n \left(\frac{\sqrt{n}x_{t,n} - x}{h}\right)^i K^2\left(\frac{\sqrt{n}x_{t,n} - x}{h}\right) \to_p L_G\left(1,0\right) \int_{-\infty}^{\infty} s^i K^2\left(s\right)$$

with $L_G\left(1,0\right)$ is the local time at the origin of a continuous limiting Gaussian process $G$ over the time interval $[0,1]$ such that $\sup_{0\leq t\leq 1} |x_{[nt],n} - G\left(t\right)| = o_p\left(1\right)$.

As for (10), by noting that, given $\nu_1, ...\nu_n, (Y_{nt}u_t,\ t = 1, ..., n)$ is a martingale difference since $\nu_t$ is independent of $u_t$, it follows from Theorem 3.9 [(3.75) there] in Hall and Heyde (1980) with $\delta = \frac{q}{2} - 1$ that

$$\sup_x |P\left(V_n \leq x\sigma | \epsilon_1, ..., \epsilon_n\right) - \Phi\left(x\right)| \leq M\left(\delta\right) \mathcal{L}_n^{1/(1+q)}, \ \ a.s.,$$

where $M(\delta)$ is a constant depending only on $\delta$ and

$$\begin{aligned}
\mathcal{L}_n &= \frac{1}{\sigma^q \Lambda_n^q} \sum_{k=1}^{n} |Y_{nk}|^q \mathbb{E}|u_k|^q + \mathbb{E}|\frac{1}{\sigma^2 \Lambda_n^2} \sum_{k=1}^{n} Y_{nk}^2 \left[\mathbb{E}\left(u_k^2|\mathcal{F}_{k-1}\right) - \sigma^2\right]|^{q/2} \\
&< \frac{\sup_t \mathbb{E}|u_t|^q}{\sigma^q \Lambda_n^q} \sum_{k=1}^{n} |Y_{nk}|^q + o_p(1) \\
&= \frac{\sup_t \mathbb{E}|u_t|^q}{\sigma^q \Lambda_n^q} \sum_{k=1}^{n} |\frac{A_0^{\frac{1}{2}} \left(\frac{\sqrt{n}}{nh}\right)^{\frac{1}{2}} \left(K\left(\frac{\sqrt{n}x_{k,n}-x}{h}\right) A_2 - \left(\frac{\sqrt{n}x_{k,n}-x}{h}\right) K\left(\frac{\sqrt{n}x_{k,n}-x}{h}\right) A_1\right)}{A_0 A_2 - A_1^2}|^q + o_p(1) \\
&= \frac{\sup_t \mathbb{E}|u_t|^q}{\sigma^q \Lambda_n^q} \left(\frac{\sqrt{n}}{nh}\right)^{\frac{q}{2}} \sum_{k=1}^{n} |\frac{A_0^{\frac{1}{2}} \left(K\left(\frac{\sqrt{n}x_{k,n}-x}{h}\right) A_2 - \left(\frac{\sqrt{n}x_{k,n}-x}{h}\right) K\left(\frac{\sqrt{n}x_{k,n}-x}{h}\right) A_1\right)}{A_0 A_2 - A_1^2}|^q + o_p(1) \\
&= \frac{M}{\sigma^q \Lambda_n^q} \left(\frac{\sqrt{n}}{nh}\right)^{\frac{q-2}{2}} + o_p(1) = o_p(1)
\end{aligned}$$

since assumption 3, $\Lambda_n^2 \to_p \int_{-\infty}^{\infty} K^2(s)\, ds$, $q > 2$ and $nh^2 \to \infty$. Therefore, we obtain

$$\sup_x |P(V_n \leq x\sigma) - \Phi(x)| \leq \mathbb{E}\left[\sup_x |P(V_n \leq x\sigma | x_1, ..., x_n) - \Phi(x)|\right] \to 0.$$

This proves (10) and also completes the proof of Theorem 1.

**Proof of Theorem 2.** Under the assumption $(U_n, V_n) \to_d (U, V)$, it follows from the so-called Skorohod-Dudley-Wichura representation theorem that there is a common probability space $(\Omega, \mathcal{F}, P)$ supporting $\left(U_n^0, V_n^0\right)$ and $(U, V)$ such that

$$(U_n, V_n) =_d \left(U_n^0, V_n^0\right) \quad \text{and} \quad \left(U_n^0, V_n^0\right) \to_{a.s.} (U, V) \tag{11}$$

in $D[0,1]^2$ with the uniform topology. Moreover, as in Lemma 2.1 in Park and Phillips (2001), $\left(U_n^0, V_n^0\right)$ can be chosen such that, for each $n \geq 1$

$$U_n^0\left(\frac{k}{n}\right) =_d U_n\left(\frac{\tau_{nk}}{n}\right) \quad \text{and} \quad V_n^0\left(\frac{k}{n}\right) =_d V_n\left(\frac{\tau_{nk}}{n}\right), \qquad k = 1, 2, ..., n, \tag{12}$$

where $\tau_{nt}$, $0 \leq t \leq 1$, are stopping times in $(\Omega, \mathcal{F}, P)$ with $\tau_{no} = 0$ satisfying

$$\sup_{0 \leq t \leq 1} |\frac{\tau_{nt} - t}{n^{\delta}}| \to_{a.s.} 0 \tag{13}$$

as $n \to \infty$ for any $\delta > \max\left(\frac{1}{2}, \frac{2}{q}\right)$. These facts, together with (8), yield that, under the extended probability space,

$$\left(h\sum_{t=1}^{n}K_h\left(x_t-x\right)\right)^{1/2}\left(\hat{m}_{LL}\left(x\right)-m\left(x\right)\right)=_d\sum_{t=1}^{n}Y_{nt}^{*}\sqrt{n}\left[V^0\left(\frac{\tau_{n,t}}{n}\right)-V^0\left(\frac{\tau_{n,t-1}}{n}\right)\right]+\Theta_{2n}^{*}\Theta_{1n}^{*},\tag{14}$$

where $Y_{nt}^{*}=\Theta_{2n}^{*}\left(\frac{\sqrt{n}}{nh}\right)^{\frac{1}{2}}z_{nt}^{*}\frac{1}{h}K\left[\frac{\sqrt{n}}{h}U_n^0\left(\frac{t}{n}\right)-\frac{x}{h}\right]$,
$\Theta_{1n}^{*}=\left(\frac{\sqrt{n}}{nh}\right)^{\frac{1}{2}}\sum_{t=1}^{n}z_{nt}^{*}\frac{1}{h}K\left[\frac{\sqrt{n}}{h}U_n^0\left(\frac{t}{n}\right)-\frac{x}{h}\right]\left(\frac{\left(\sqrt{n}U_n^0\left(\frac{t}{n}\right)-x\right)^2m''\left(x^*\right)}{2}\right)$ and
$\Theta_{2n}^{*}=e_1'\left(\frac{\sqrt{n}}{nh}\sum_{t=1}^{n}K\left[\frac{\sqrt{n}}{h}U_n^0\left(\frac{t}{n}\right)-\frac{x}{h}\right]\right)^{1/2}\left(\frac{\sqrt{n}}{nh}\sum_{t=1}^{n}z_{nt}^{*}\frac{1}{h}K\left[\frac{\sqrt{n}}{h}U_n^0\left(\frac{t}{n}\right)-\frac{x}{h}\right]z_{nt}^{*\prime}\right)^{-1}$ with
$z_{nt}^{*\prime}=\left(1,\sqrt{n}U_n^0\left(\frac{t}{n}\right)-x\right)$.

Since (11) implies that Assumption 2.2* in WP holds true for $U_n^0\left(\frac{t}{n}\right)$ with $G\left(t\right)$ being a Brownian motion, it follows from a similar argument to the proofs of Theorem 1 that

$$A_i^{*}=\frac{\sqrt{n}}{nh}\sum_{t=1}^{[nr]}\left(\frac{\sqrt{n}}{h}U_n^0\left(\frac{t}{n}\right)-\frac{x}{h}\right)^iK\left(\frac{\sqrt{n}}{h}U_n^0\left(\frac{t}{n}\right)-\frac{x}{h}\right)\to_p\begin{cases}L_U\left(r,0\right)\int_{-\infty}^{\infty}s^iK\left(s\right), & i\text{ is even;}\\ 0, & i\text{ is odd,}\end{cases}$$

and

$$B_i^{*}=\frac{\sqrt{n}}{nh}\sum_{t=1}^{[nr]}\left(\frac{\sqrt{n}}{h}U_n^0\left(\frac{t}{n}\right)-\frac{x}{h}\right)^iK^2\left(\frac{\sqrt{n}}{h}U_n^0\left(\frac{t}{n}\right)-\frac{x}{h}\right)\to_p L_U\left(r,0\right)\int_{-\infty}^{\infty}s^iK^2\left(s\right)$$

uniformly in $r\in[0,1]$. Therefore, we have

$$\Theta_{2n}^{*}\Theta_{1n}^{*}=h^2\left(\frac{nh}{\sqrt{n}}\right)^{\frac{1}{2}}A_0^{*\frac{1}{2}}\frac{A_2^{*2}-A_1^{*}A_3^{*}}{A_0^{*}A_2^{*}-A_1^{*2}}\frac{m''\left(x^*\right)}{2}\to_p 0,$$

whenever $nh^{10}\to 0$ and $nh^2\to\infty$.

These facts, together with (14) and an argument similar to that in the proof of Theorem 3.3 in Hall and Heyde (1980) imply that (4) will follow if we prove

$$\sum_{t=1}^{n}Y_{nt}^{*}\sqrt{n}\left[V^0\left(\frac{\tau_{n,t}}{n}\right)-V^0\left(\frac{\tau_{n,t-1}}{n}\right)\right]\to_d\int_{-\infty}^{\infty}K^2\left(s\right)dsN\left(0,1\right).\tag{15}$$

In order to prove (15), write

$$M_n\left(r\right)=\sqrt{n}\sum_{t=1}^{j-1}Y_{nt}^{*}\left[V^0\left(\frac{\tau_{n,t}}{n}\right)-V^0\left(\frac{\tau_{n,t-1}}{n}\right)\right]+\sqrt{n}Y_{n,j-1}^{*}\left[V^0\left(\frac{r}{n}\right)-V^0\left(\frac{\tau_{n,j-1}}{n}\right)\right],\tag{16}$$

for $\frac{\tau_{n,j-1}}{n} < r \leq \frac{\tau_{n,j}}{n}$, $j = 1, 2, ..., k$. It is readily seen that $M_n$ is a continuous martingale with the quadratic variation process $[M_n]$ given by

$$[M_n]_r = n \sum_{t=1}^{j-1} Y_{nt}^{*2} \left( \frac{\tau_{n,t}}{n} - \frac{\tau_{n,t-1}}{n} \right) + nY_{n,j-1}^{*2} \left( \frac{r}{n} - \frac{\tau_{n,j-1}}{n} \right) \rightarrow_p \int_{-\infty}^{\infty} K^2(s)\, ds,$$

uniformly in $r \in [0, 1]$, since (13) and

$$\begin{aligned}
&\frac{\sqrt{n}}{h} \int_0^r \left( \frac{\sqrt{n}}{h} U_n^0(s) - \frac{x}{h} \right)^{\kappa} K^{\lambda} \left( \frac{\sqrt{n}}{h} U_n^0(s) - \frac{x}{h} \right) ds \\
&= \frac{\sqrt{n}}{nh} \sum_{t=1}^{[nr]} \left( \frac{\sqrt{n}}{h} U_n^0\left(\frac{t}{n}\right) - \frac{x}{h} \right)^{\kappa} K^{\lambda} \left( \frac{\sqrt{n}}{h} U_n^0\left(\frac{t}{n}\right) - \frac{x}{h} \right) \\
&+ \frac{\sqrt{n}}{nh} \left( -\frac{x}{h} \right)^{\kappa} K^{\lambda} \left( -\frac{x}{h} \right) \\
&+ \frac{\sqrt{n}}{nh} \left( \frac{\sqrt{n}}{h} U_n^0\left(\frac{[nr]}{n}\right) - \frac{x}{h} \right)^{\kappa} K^{\lambda} \left( \frac{\sqrt{n}}{h} U_n^0\left(\frac{[nr]}{n}\right) - \frac{x}{h} \right) (nr - [nr]) \\
&\rightarrow_p L_U(r, 0) \int_{-\infty}^{\infty} s^{\kappa} K^{\lambda} ds
\end{aligned}$$

with $\kappa = 0, 2$ and $\lambda = 2$. For the covariance process $[M_n, U]$ of $M_n$ and $U$, we also have

$$[M_n, U]_r = \sqrt{n} \sum_{t=1}^{j-1} Y_{nt}^{*} \left( \frac{\tau_{n,t}}{n} - \frac{\tau_{n,t-1}}{n} \right) \sigma_{uv} + \sqrt{n} Y_{n,j-1}^{*} \left( \frac{r}{n} - \frac{\tau_{n,j-1}}{n} \right) \sigma_{uv} \rightarrow_p 0$$

where $\sigma_{uv} = cov(U, V)$. Now, following the proof of Theorem 3.2 in Park and Phillips (2001), we obtain that

$$\sum_{t=1}^{n} Y_{nt}^{*} \sqrt{n} \left[ V^0\left( \frac{\tau_{n,t}}{n} \right) - V^0\left( \frac{\tau_{n,t-1}}{n} \right) \right] = M_n\left( \frac{\tau_{n,n}}{n} \right) \rightarrow_d \int_{-\infty}^{\infty} K^2(s)\, ds N(0, 1). \tag{17}$$

which yields (4). This completes the Theorem 2.

**Proof of Theorem 3.** In order to prove (6), we split $\hat{m}_{NW}(x) - m(x)$ as

$$\mathbb{E}\left[ \hat{m}_{NW}(x) - m(x) | x_t \right] = \frac{m'(x) I_a}{\sum_{t=1}^{n} K\left( \frac{x_t - x}{h} \right)} + \frac{m''(x^*) I_b}{\sum_{t=1}^{n} K\left( \frac{x_t - x}{h} \right)}. \tag{18}$$

where $I_a = h \sum_{t=1}^{n} \left( \frac{x_t - x}{h} \right) K\left( \frac{x_t - x}{h} \right)$ and $I_b = \frac{h^2}{2} \sum_{t=1}^{n} \left( \frac{x_t - x}{h} \right)^2 K\left( \frac{x_t - x}{h} \right)$.

It is readily seen from Theorem 2.1 in Wang and Phillips (2008a, 2009) that

$$\frac{\left( nh^{-2} \right)^{1/4} m'(x) I_a}{\sum_{t=1}^{n} K\left( \frac{x_t - x}{h} \right)} \rightarrow_d \tau L_W^{-1/2}(1, 0) N(0, 1), \tag{19}$$

where $\tau^2 = \left[m'(x)\right]^2 |\phi|^{-1} \int s^2 K^2(s)\, ds$, and

$$\frac{h^{-2} m''(x^*) I_b}{\sum_{t=1}^{n} K\left(\frac{x_t - x}{h}\right)} \to_p \frac{m''(x^*)}{2} \mu_2(K). \tag{20}$$

This gives the stated result (6).

In case of the local linear estimator, we have

$$\begin{aligned}
\mathbb{E}\left[\hat{m}_{LL}(x) - m(x) | x_t\right] &= e_1' \begin{pmatrix} \frac{1}{h} A_0 & A_1 \\ A_1 & h A_2 \end{pmatrix}^{-1} \begin{pmatrix} h A_2 \\ h^2 A_3 \end{pmatrix} \frac{m''(x^*)}{2} \\
&= e_1' \frac{1}{A_0 A_2 - A_1^2} \begin{pmatrix} h A_2 & -A_1 \\ -A_1 & \frac{1}{h} A_0 \end{pmatrix} \begin{pmatrix} h A_2 \\ h^2 A_3 \end{pmatrix} \frac{m''(x^*)}{2} \\
&= h^2 \frac{m''(x^*)}{2} \frac{A_2^2 - A_1 A_3}{A_0 A_2 - A_1^2} \\
&\to_p h^2 \frac{m''(x^*)}{2} \int_{-\infty}^{\infty} s^2 K(s)\, ds.
\end{aligned}$$

This gives the desired result (5).

Table 1: Asymptotics for nonparametric estimators with stationary time series

| | Nadaraya-Watson | Local linear |
|---|---|---|
| Order | $\sqrt{nh}$ | $\sqrt{nh}$ |
| Bias | $\frac{h^2}{2}\mu_2(K)\{m''(x)+2m'(x)\frac{f'}{f}(x)\}$ | $\frac{h^2}{2}\mu_2(K)m''(x)$ |
| Variance | $\sigma^2\int_{-\infty}^{\infty}K^2(s)\,ds/f(x)$ | $\sigma^2\int_{-\infty}^{\infty}K^2(s)\,ds/f(x)$ |

where $\mu_2(K)=\int_{-\infty}^{\infty}s^2K(s)\,ds$ and $f(x)$ is the density function of random variable $X$ at a point $x$.

Table 2: Asymptotics for nonparametric estimators with nonstationary time series

| | Nadaraya-Watson | Local linear |
|---|---|---|
| Order | $\left(h\sum_{t=1}^{n}K_h(x_t-x)\right)^{1/2}$ | $\left(h\sum_{t=1}^{n}K_h(x_t-x)\right)^{1/2}$ |
| Bias | $O_p\left(n^{-1/4}h^{1/2}\right)+\frac{h^2}{2}\mu_2(K)m''(x)$ | $\frac{h^2}{2}\mu_2(K)m''(x)$ |
| Variance | $\sigma^2\int_{-\infty}^{\infty}K^2(s)\,ds$ | $\sigma^2\int_{-\infty}^{\infty}K^2(s)\,ds$ |

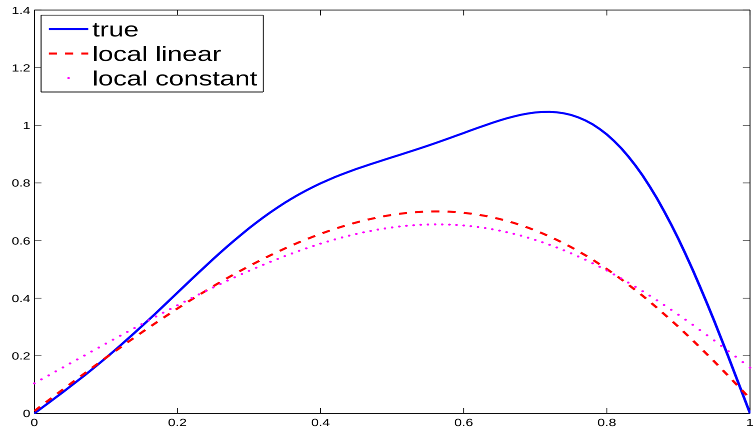
Note that $\left(h\sum_{t=1}^{n}K_h(x_t-x)\right)^{1/2}=O_p\left(\sqrt{\sqrt{n}h}\right)$.

Table 3: $m(x)$ on $[0, 1]$

| | | NW | LL |
|---|---|---|---|
| $h = n^{-1/5}$ | bias | 0.2658 | 0.2121 |
| | SD | 0.0085 | 0.0095 |
| | MSE | 0.1070 | 0.0762 |
| | IMSE | 0.0917 | 0.0708 |
| $h = n^{-1/3}$ | bias | 0.0874 | 0.0597 |
| | SD | 0.0133 | 0.0142 |
| | MSE | 0.0152 | 0.0077 |
| | IMSE | 0.0103 | 0.0068 |
| $h = n^{-1/2}$ | bias | 0.0210 | 0.0094 |
| | SD | 0.0236 | 0.0248 |
| | MSE | 0.0020 | 0.0011 |
| | IMSE | 0.0012 | 0.0009 |

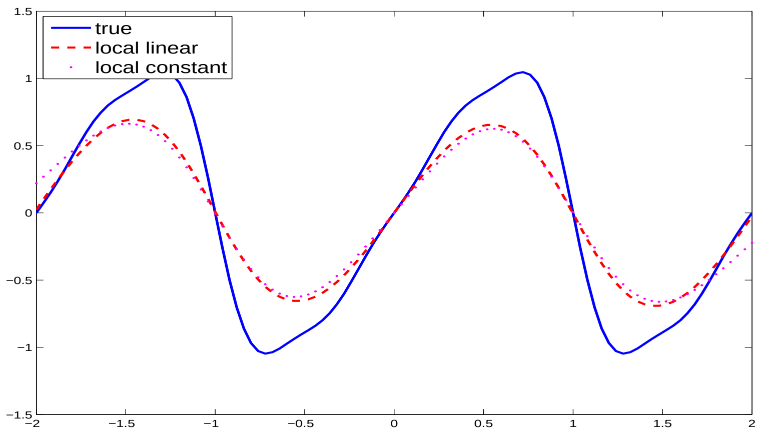
Table 4: $m(x)$ on $[-2, 2]$

| | | NW | LL |
|---|---|---|---|
| $h = n^{-1/5}$ | bias | 0.2927 | 0.2332 |
| | SD | 0.0106 | 0.0116 |
| | MSE | 0.1285 | 0.0905 |
| | IMSE | 0.4799 | 0.3581 |
| $h = n^{-1/3}$ | bias | 0.0953 | 0.0617 |
| | SD | 0.0169 | 0.0181 |
| | MSE | 0.0179 | 0.0085 |
| | IMSE | 0.0653 | 0.0335 |
| $h = n^{-1/2}$ | bias | 0.0248 | 0.0097 |
| | SD | 0.0302 | 0.0322 |
| | MSE | 0.0028 | 0.0017 |
| | IMSE | 0.0096 | 0.0055 |




(a) $h = n^{-\frac{1}{5}}$

Figure 1: Graphs over the interval $[0, 1]$ of $m(x)$ (solid) and Monte Carlo estimates of $\mathbb{E}(\hat{m}_{NW}(x))$ (dotted) and $\mathbb{E}(\hat{m}_{LL}(x))$ (dashes) with $\sigma = 0.1$




(a) $h = n^{-\frac{1}{5}}$

Figure 2: Graphs over the interval $[-2, 2]$ of $m(x)$ (solid) and Monte Carlo estimates of $\mathbb{E}(\hat{m}_{NW}(x))$ (dotted) and $\mathbb{E}(\hat{m}_{LL}(x))$ (dashes) with $\sigma = 0.1$