DEPARTMENT OF ECONOMICS
UNIVERSITY OF CYPRUS

DOUBLE KERNEL NON-PARAMETRIC ESTIMATION IN SEMIPARAMETRIC ECONOMETRIC MODELS

**Thanasis Stengos and
Boping Yan**

Discussion Paper 99-02

P.O. Box 20537, 1678 Nicosia, CYPRUS Tel.: ++357-2-892101, Fax: ++357-2-750310
Web site: http://www.econ.ucy.ac.cy

Double Kernel Non-parametric Estimation in Semiparametric Econometric Models.

Thanasis Stengos and Boping Yan
Department of Economics
University of Guelph
Guelph, Ontario N1G 2W1
Canada

December 1998.

Abstract

In this paper we estimate the coe±cients of a generated regressor in the context of a partially linear semiparametric regression model. The generated regressor is part of the linear part of the model and the estimator is obtained by double kernel estimation. It is established that the double kernel estimator is root-n-consistent and asymptotically normal. Monte Carlo results suggest that it has satisfactory small samples properies. The usefulness of the proposed method is illustrated in an application to a model of wage determination.

Key Words: Nonparametric estimation; Correlation; Double kernel estimation.

1 Introduction

Following the work of Pagan (1984, 1986) generated regressors have introduced a number of interesting econometric problems in the context of fully parametric regression models. Models of this type may naturally arise when modelling individual behavior under uncertainty, where actions depend upon predictions (conditional extectations) of unobserved outcomes or errors that arise from these predictions. Many macroeconomic models with rational expectations fall in this category, where the generated regressors enter as "surprise" variables or simply as conditional expectations, see for example Bean's (1986) model of the consumption function. Also in econometric models of labor markets many models that analyze contarct duration or labor supply responses make use of wage estimates as right hand variables. Another labor market model where generated regressors appear is when a variable, say length of stay at the current job (job tenure) appears on the right hand side in a wage equation. This variable may be viewed as endogenous to the wage formation process and hence it may enter as a predicted variable from another ¯rst stage regression.

This paper investigates in the context of a nonparametric model, the estimation of the parameter of a variable that has been generated by a preliminary nonparametric ¯lter. The model becomes a partially linear semiparametric regression model, see Robinson (1988), in which some explanatory variables in the linear part (the part of interest to the researcher) are the unknown conditional means of certain observables given other observable regressors. The researcher is primarily interested in estimating the impact of these unknown conditional means on the dependent variable. The framework assumes that the conditional mean of the variable(s) that enters the linear part of the partial linear semiparametric model is a smooth unknown function of other independent explanatory variables formulated in an auxiliary equation. Furthermore, there is a correlated error structure between the equation of interest and the auxiliary equation. We estimate the parameter(s) of interest using double kernel nonparametric estimation. Double kernel estimation was also applied by Delgado, Li and Stengos (1995) in the context of two non-nested nonparametric regression models to derive a J-type test statistic of one model against the other, see Davdson and Mackinnon (1981).

Other approaches, especially the two-step approach method in estimation of a semiparametric model should be mentioned. Andrews (1991, 1994) and Newey (1994) provide results for situations in which the generated regressors are estimated nonparametrically, but the equation of interest is parametric which is different from our nonparametric case. However, the parametric specification used in the second step may be too strong and may lead to inconsistent estimates. The present analysis does not require the parametric assumption in the second step. Ahn (1995) and Rilstone (1996) have focused on a situation in which both the equation of interest and the auxiliary equation are estimated nonparametrically. Ahn (1997) established that the two-step estimator is $\sqrt{n}$-consistent and asymptotically normal provided that the kernel estimates of both steps converge uniformly at a sufficiently rapid rate.

For our specific model, allowing for correlation of the error structures between the equation of interest and the auxiliary equation, we establish that the double kernel estimator is $\sqrt{n}$-consistent and asymptotically normal. Our Monte Carlo study shows that the double kernel estimator behaves quite satisfactorily in medium to larger samples in terms of mean absolute bias and mean squared error.

Section 2 derives the estimator in a special case of a parametric equation of interest, except for the unknown conditional mean. Section 3 discusses the semiparametric formulation of the model. We present in detail the proposed double kernel estimator and we derive its asymptotic distribution. Section 4 provides Monte Carlo simulations results. Section 5 looks at an empirical example where the proposed estimator is applied to a labor supply model using Canadian data from the Labor and Manpower Activity Survey, LMAS89. Finally, section 6 offers concluding remarks. The Appendices A and B collect lengthy proofs.

2 The Case of a Parametric Regression Function.

Data consists of independent observations $\{(x_i; z_i; y_i; s_i); i = 1; \ldots; n\}$ identically distributed as the $R^p \times R^q \times R \times R$-valued multivariate random variable $(x; z; y; s)$. We consider the impact of the conditional mean $E(s|z)$; which is the unknown function of z, on the dependent variable y: We allow that the variables z may be correlated with the x's in the equation of interest. The model we propose is

$$y = \mu(x) + E(s|z)\alpha + u \tag{1}$$

The auxilliary regression is written as

$$s = E(s|z) + \varepsilon \tag{2}$$

The errors u and the ε are correlated: Let $g_i = g(z_i) = E(s_i|z_i)$, and use the notation $\tilde{A}_i = \tilde{E}(A_i|z_i) = \left[\frac{1}{nb^q}\sum_{j \neq i} A_j \bar{K}_{ij}\right]/\tilde{f}_i$, where $\bar{K}_{ij} = \bar{K}(\frac{z_i - z_j}{b})$ is the kernel function associated with z and b is the corersponding smoothing parameter. (In this context $\bar{K}(\cdot)$ is the product kernel). The equation of interest becomes

$$y = \theta(x) + \tilde{E}(s|z)\alpha + \eta \tag{3}$$

where

$$\eta = u + [E(s|z) - \tilde{E}(s|z)]\alpha$$

We consider below the following special model in which the function of the explanatory variables in the equation of interest is parametric, that is, $\theta(x) = x'\gamma$:

$$y = x'\gamma + E(s|z)\alpha + u \tag{4}$$

$$y_i = x_i'\gamma + E(s_i|z_i)\alpha + u_i; \qquad E(u_i|x_i; z_i) = 0; \tag{5}$$

For example, y_i might be the wage rate, x_i refers to certain variables that affect wages, such as level of education, age and gender, whereas s_i represents length of time at the present job. This variable itself may be affected by other variables z_i, such as age, marital status, number of children and other demographic characteristics which are all assumed to be exogenous. Hence, $g(z_i) = E(s_i|z_i)$ is the expected job tenure, while the functional form of $g(\cdot)$ is not specified.

We are interested in estimating γ and α.

$$y = x'\gamma + \tilde{E}(s|z)\alpha + \eta \tag{6}$$

where

$$\eta = u + [E(s|z) - \tilde{E}(s|z)]\alpha$$

Then,

$$y_i = x_i'\gamma + \tilde{s}_i\alpha + (g_i - \tilde{s}_i)\alpha + u_i \tag{7}$$

From (7), we have

$$\begin{aligned} y_i &= (x_i'; \tilde{s}_i)(\gamma'; \alpha)' + (g_i - \tilde{s}_i)\alpha + u_i & (8) \\ &= \hat{X}_i\delta + (g_i - \tilde{s}_i)\alpha + u_i; & (9) \end{aligned}$$

where $\delta = (\gamma', \alpha)'$ and $\hat{X}_i = (x_i', \hat{s}_i)$. We estimate δ by regressing y on $\hat{X}$:

$$\begin{aligned}\hat{\delta} &= (\hat{X}'\hat{X})^{-1}\hat{X}'y \\ &= \delta + (\hat{X}'\hat{X})^{-1}\hat{X}'[(g - \hat{s})\alpha + u];\end{aligned}$$

where $\hat{X} = (\hat{X}_1', \hat{X}_2', ..., \hat{X}_n')'$ and $\hat{s}_i = \hat{E}(s_i|z_i) \xrightarrow{p} g_i = g(z_i)$.
To derive the asymptotic distribution of $\sqrt{n}(\hat{\delta} - \delta)$, (and consequently that of $\sqrt{n}(\hat{\alpha} - \alpha)$), the following definitions and assumptions will be used.

Let G_l^α denote the class of functions such that if $g \in G_l^\alpha$ ($\alpha > 0$ and $l \geq 1$ is an integer), then g is l times differentiable, g and its derivatives (up to order l) are all bounded by some function that has αth order finite moments. Also K_l, $l \geq 1$, denote the class of even functions $k : R \to R$ satisfying $\int k(u)u^m du = \delta_{0m}$ for $m = 0, 1, ..., l-1$ and $k(u) = O((1 + |u|^{l+1+\delta})^{-1})$, some $\delta > 0$. Denote $g(z) = E(s|z)$.

(A1) (y_i, x_i, z_i, s_i) are independently distributed as (y, x, z, s), x admits a pdf $f_x = f(x) \in G_\nu^\infty$, also $\theta(x) \in G_\nu^4$ and $h(x) \in G_\nu^4$, where $\nu \geq 2$ is a positive integer. z admits a pdf $f_z = f(z) \in G_\mu^\infty$, $g(z) \in G_\mu^4$ and $E(g(z)|x) \in G_\mu^4$, where $\mu \geq 2$ is a positive integer. Moreover, (x, z) admits a joint pdf $\psi(x, z) \in G_\mu^\infty$. $\sigma^2(x)$, f_x, f_z and ψ are uniformly bounded, where $\sigma^2(x) = E(u^2|x)$.

(A2) $k \in K_\nu$. As $n \to \infty$, $na^{2p} \to \infty$, $na^{4\nu} \to 0$; $\bar{k} \in K_\mu$. As $n \to \infty$, $nb^{2q} \to \infty$ and $nb^{4\mu} \to 0$.

(A3) (i) (y_i, x_i, z_i, s_i) is a strictly stationary absolutely regular process with $E(r_i|x_i, z_i) = 0$, $r_i = u_i$ or w_i. (ii) $f_z(\cdot)$, $f(x, \cdot)$ and $g(\cdot)$ all satisfy some (global) Lipschitz-type conditions: $|r(u) - r(v)| \leq D_4(v)||u - v||$ for all $u, v \in R^q$ ($||\cdot||$ is the Euclidean norm), where $D_4(\cdot)$ has finite 4th moments, $r(\cdot) = f_z(\cdot)$ or $g(\cdot)$, and in the case of $f(x, \cdot)$, $r(\cdot) = f(x, \cdot)$. (iii) Both u_i and w_i have finite $4 + \epsilon$ moments for some small $\epsilon > 0$.

Assumption (A1) presents some smoothness and moments conditions. (A2) is similar to the conditions used by Robinson (1988) or Fan, Li and Stengos (1992). It requires a higher order kernel to be used for k ($\bar{k}$) if $p \geq 4$ ($q \geq 4$). Then we have,

Theorem 1 Define $X_i = (x_i', g_i)$ and $\eta_i = 2\alpha f(z_i)E[X_j|z_j = z_i]$. Then under assumptions (A1), (A2) and (A3), we have

$$\sqrt{n}(\hat{\delta} - \delta) = \sqrt{n}S_{\hat{X}}^{-1}S_{\hat{X}, u+(g-\hat{s})\alpha} \xrightarrow{d} N(0, \Sigma_1);$$

where $\Sigma_1 = \Phi_1^{-1}(\Omega_1 + \Omega_2 - 2\Omega_{12})\Phi_1^{-1}$, $\Phi_1 = E[X_i'X_i]$, $\Omega_1 = E[X_i'X_i\sigma^2]$, $\Omega_2 = E[w_i^2\eta_i'\eta_i]$, and $\Omega_{12} = E[w_iu_iX_i'\eta_i]$.

The proofs are presented in Appendix A.

3 The General Case of a Semiparametric Regression Model.

As in the previous section the data consists of independent observations $\{(x_i; z_i; y_i; s_i); i = 1; \ldots; n\}$ identically distributed as the $R^p \times R^q \times R \times R$-valued multivariate random variable $(x; z; y; s)$. The model is given by

$$y = \theta(x) + E(s|z)\alpha + u \tag{10}$$

The auxilliary regression is written as

$$s = E(s|z) + \epsilon \tag{11}$$

The errors u and the ϵ may be correlated. Hence s and u may be correlated, but we assume that $(x; z)$ is uncorrelated with u. The regression function of interest is written as

$$y = \theta(x) + \bar{E}(s|z)\alpha + \eta; \tag{12}$$

where $\eta = u + [E(s|z) - \bar{E}(s|z)]\alpha$.

Following Robinson's (1988) semi-parametric estimation approach, α in (12) could be estimated by

$$\hat{\alpha} = \frac{\sum_i(\bar{E}(s_i|z_i) - \hat{E}[\bar{E}(s_i|z_i)|x_i])(y_i - \hat{E}(y_i|x_i))}{\sum_i(\bar{E}(s_i|z_i) - \hat{E}[\bar{E}(s_i|z_i)|x_i])^2} \tag{13}$$

where $\hat{E}(\cdot|x)$ is a nonparametric estimate of $E(\cdot|x)$, and $\bar{E}(\cdot|z)$ is a nonparametric estimate of $E(\cdot|z)$. A direct application of Robinson's (1988) method in (12) requires two trimming parameters (in addition to the two smoothing parameters) to overcome the random denominator problem that arises in kernel estimation. However, the technical difficulties of using a trimming method in the context of double kernel estimation prove difficult to overcome. Therefore, we choose to assume a bounded density function to avoid the random denominator problem. We estimate $E(s_i|x_i)$ and $E(y_i|x_i)$ respectively by

$$\hat{s}_i = \frac{\frac{1}{na^p}\sum_{j\neq i} s_jK_{ij}}{\hat{f}_x(x_i)}; \tag{14}$$

$$\hat{y}_i = \frac{\frac{1}{na^p}\sum_{j\neq i} y_j K_{ij}}{\hat{f}_x(x_i)}, \tag{15}$$

and $f_x(x_i)$, the probability density function (p.d.f) of x_i, by $\hat{f}_x(x_i) = \frac{1}{na^p}\sum_{j\neq i} K_{ij}$, where $K_{ij} = K(\frac{x_i - x_j}{a})$ is the kernel function and a is the smoothing parameter. We use a product kernel, $K(u) = \prod_{l=1}^{p} k(u_l)$; u_l is the lth component of u.

We estimate $E(s_i|z_i)$ by

$$\tilde{s}_i = \frac{\frac{1}{nb^q}\sum_{j\neq i} s_j \tilde{K}_{ij}}{\tilde{f}_z(z_i)}, \tag{16}$$

and $f_z(z_i)$ is estimated by $\tilde{f}_z(z_i) = \frac{1}{nb^q}\sum_{j\neq i}\tilde{K}_{ij}$, where $\tilde{K}_{ij} = \tilde{K}(\frac{z_i - z_j}{b})$ is the kernel function associated with z and b is the corresponding smoothing parameter ($\tilde{K}(\cdot)$ is also a product kernel). We also need to estimate $E\{E(s_i|z_i)|x_i\}$. Its kernel estimate is given by

$$\hat{\tilde{s}}_i = \frac{\frac{1}{na^p}\sum_{j\neq i}\tilde{s}_j K_{ij}}{\hat{f}_x(x_i)} \equiv \frac{1}{na^p\hat{f}_x(x_i)}\sum_{j\neq i}\Big[\frac{1}{nb^q\tilde{f}_z(z_j)}\sum_{l\neq j} s_l\tilde{K}_{lj}\Big]K_{ij}. \tag{17}$$

Let $g_i = g(z_i) = E(s_i|z_i)$, and using the same notation introduced above, where "~" denotes kernel estimator conditional on z, e.g. $\tilde{A}_i = \tilde{E}(A_i|z_i) = \Big[\frac{1}{nb^q}\sum_{j\neq i} A_j\tilde{K}_{ij}\Big]/\tilde{f}_i$; we have that (12) becomes

$$y_i = \theta_i + \tilde{s}_i\alpha + (g_i - \tilde{s}_i)\alpha + u_i \tag{18}$$

From (18), we have

$$\hat{y}_i = \hat{\theta}_i + \hat{\tilde{s}}_i\alpha + (\hat{g}_i - \hat{\tilde{s}}_i)\alpha + \hat{u}_i \tag{19}$$

where

$$\hat{A}_i = \hat{E}(A_i|x_i) = \frac{\frac{1}{na^p}\sum_{j\neq i} A_j K_{ij}}{\hat{f}_i}$$

where as before "^" denotes estimation conditional on x. Subtracting (19) from (18) yields,

$$y_i - \hat{y}_i = \theta_i - \hat{\theta}_i + (\tilde{s}_i - \hat{\tilde{s}}_i)\alpha + (g_i - \tilde{s}_i)\alpha - (\hat{g}_i - \hat{\tilde{s}}_i)\alpha + u_i - \hat{u}_i \tag{20}$$

We estimate α by regressing $y - \hat{y}$ on $\tilde{s} - \hat{\tilde{s}}$. Denotes $S_{A,B} = \frac{1}{n}\sum_i A_i B_i'$ and $S_A = S_{A,A}$, we have

$$\hat{\alpha} = S^{-1}_{\tilde{s}-\hat{\tilde{s}}} S_{\tilde{s}-\hat{\tilde{s}},\, y-\hat{y}}.$$

Using the assumptions that were given in the previous section we are now presenting the main results in the form of two theorems. Theorem 2 collects some intermediate results that are important in the derivation of the asymptotic distribution of $\hat{\alpha}$ given in Theorem 3.

Theorem 2 Under assumptions (A1) and (A2), as $n \to \infty$,

$$\text{(i)} \qquad S_{\bar{s}-\hat{\bar{s}}} \xrightarrow{p} E[(g_1 - h_1)'(g_1 - h_1)]$$

$$\text{(ii)} \qquad \sqrt{n}\, S_{\bar{s}-\hat{\bar{s}},u} \xrightarrow{d} N(0, \Sigma)$$

$$\text{(iii)} \qquad \sqrt{n}\left(S_{\bar{s}-\hat{\bar{s}},\theta-\hat{\theta}} + S_{\bar{s}-\hat{\bar{s}},(g-\bar{s})\alpha} - S_{\bar{s}-\hat{\bar{s}},(\hat{g}-\hat{\bar{s}})\alpha} - S_{\bar{s}-\hat{\bar{s}},\hat{u}} \right) = o_p(1)$$

where $\Sigma = \sigma^2 E[(g_1 - h_1)'(g_1 - h_1)]$.

Theorem 3 Under assumptions (A1) and (A2), as $n \to \infty$,

$$\sqrt{n}(\hat{\alpha} - \alpha) \xrightarrow{d} N\left(0, \sigma^2 (E[(g_1 - h_1)'(g_1 - h_1)])^{-1}\right) \tag{21}$$

The proofs are presented in the appendix B.

In the next section we will analyze the properties of the above estimator by means of a Monte Carlo invetsigation.

4 The Results of Monte Carlo Study

Based on the model of the previous section our design has

$$\theta(x) = (\beta_1 x_1 + \beta_2 x_2)^2; \qquad (\beta_1 = \beta_2 = 1)$$

$$s = \left[(z_1 + z_2)^2;\ z_2\right]\gamma + e; \qquad (\gamma = (1; 1)')$$

where e is normal with $N(0; \sigma_e^2)$, x_i (i=1, 2) are generated from a uniform distribution on [1, 2] and z_i (i=1, 2) are generated according to

$$z_i = x_i \delta_i + v_i; \ i = 1; 2;$$

where v_i is normal with $N(0; \sigma_v^2)$, $\delta_i = \mathrm{cov}(x_i; z_i)/\sigma_{x_i}^2$. We assume that x_i and z_i are correlated with $\rho_i = \frac{\mathrm{cov}(x_i; z_i)}{\sigma_{x_i}\sigma_{z_i}}$. We choose the coefficients of $\delta_i (= \delta)$ (i=1, 2) by using $\delta = \frac{\sqrt{12}\sigma\rho}{\sqrt{1-\rho^2}}$. For instance, δ = 0.35, 2 and 7.15 if $\sigma = 1$ such that ρ = 0.1, 0.5 and 0.9 respectively, the correlation coefficients between x and z. We set $\sigma_v^2 = 1$. Then y is generated by

$$y = \theta(x) + E(s|z)\alpha + u;$$

where $E(s|z) = [(z_1 + z_2)^2; z_2]\gamma$ and $u \sim N(0; \sigma_u^2)$.

We use the following three methods to estimate α. The we proceed to compare the different estimates in terms of their Mean Absolute Bias and Mean Squared Error performance.

(i) Truel Model Estimation: We use the true $E(s|z)$ as a regressor and estimate the model to obtain an estimate of its coefficient α. This is the case of an unattainable benchmark.

(ii) Misspecified Linear Estimation: We treat as if $E(s|z)$ were a linear function of z and use the estimate of $E(s|z)$ which comes from a linear regression to estimate α. This is the case of a misspecified benchmark where we treat the generating regression (auxilliary regression) as a linear one.

(iii) Double Kernel Estimation: We deal with an unknown function of the conditional mean $E(s|z)$, and use the double normal kernel to estimate $E[E(s|z)|x]$: Then obtain we obtain the estimate of α.

Application of the non-parametric estimation requires that kernels and bandwidths be chosen properly. In addition to choose the normal kernels, we choose the same bandwidth for estimation, which $h = cn^{-\frac{1}{4+p}}$ where $c = 1$, $p = 2$ and n is the corresponding sample size.

Table 1 and Table 2 report the results of mean absolute bias and mean square error (MSE) for α in the case of $\sigma = 1$ and $\sigma = 2$ respectively. By varying σ we can control the noise in the data generating process. There were 4000, 2000 and 1000 replications for sample sizes of $n = 100; 200$ and 400 respectively. We consider the different cases in the correlation of $\rho = 0.1$, 0.5 and 0.9 between x and z, the explanatory variables.

From Table 1 and Table 2, we see that the bias and MSE both decrease in the true model and the double kernel estimations as the sample size n increases, but the misspecified model (linear) shows a small increase. That is, the misspecified estimate will get worse when the sample size is large, which is to be expected because the estimates will be inconsistent in this case. Also the bias and MSE in the double kernel estimation method decrease significantly as the correlation between the z's and the x's increases, since that is when the estimator is designed to be at its best. However, as expected for very low correlations (low values of ρ) the double kernel estimates are not as good, especially when σ is small. However, even then they show and improvement asymptotically. In all cases as expected the estimator of α from the true model dominates the others. However this is an unrealistic benchmark, since a researcher can hardly be expected to know the data generating mechanism of the auxilliary regression. The case of the misspecified linear is of interest, since it demonstrates the danger of assuming a linear expecations formation mechanism as it is done routinely in the literature, if in fact this is an incorrect speci¯cation. In that case the parameter estimates will be severely biased and inconsistent.

The performance of the double kernel nonparametric estimator is encouraging as it behaves fairly well in moderate samples and shows an improvement with sample size. The payo® from its use is most noticeable when the explanatory variables are correlated.

Table 1: ¾ = 1

CASE	½ = 0:1		½ = 0:5		½ = 0:9	
	BIAS	MSE	BIAS	MSE	BIAS	MSE
n = 100						
Truel Model	0:0423	0:0028	0:0434	0:0020	0:0144	0:0002
Linear Model	0:5885	0:3500	0:8469	0:7173	0:9545	0:9111
Double Kernel	0:6359	0:4337	0:4067	0:1706	0:0981	0:0101
n = 200						
Truel Model	0:0311	0:0014	0:0399	0:0016	0:0140	0:0002
Linear Model	0:5947	0:3553	0:8483	0:7197	0:9547	0:9115
Double Kernel	0:5020	0:2614	0:3226	0:1055	0:0816	0:0068
n = 400						
Truel Model	0:0239	0:0008	0:0359	0:0013	0:0135	0:0002
Linear Model	0:5990	0:3595	0:8493	0:7213	0:9549	0:9117
Double Kernel	0:3977	0:1610	0:2577	0:0669	0:0679	0:0046

Table 2: ¾ = 2

CASE	½ = 0:1		½ = 0:5		½ = 0:9	
	BIAS	MSE	BIAS	MSE	BIAS	MSE
n = 100						
Truel Model	0:0138	0:0003	0:0110	0:0001	0:0036	0:0000
Linear Model	0:7867	0:6201	0:9254	0:8564	0:9775	0:9555
Double Kernel	0:1994	0:0494	0:1207	0:0159	0:0371	0:0018
n = 200						
Truel Model	0:0099	0:0001	0:0102	0:0001	0:0035	0:0000
Linear Model	0:7905	0:6255	0:9260	0:8575	0:9776	0:9557
Double Kernel	0:1546	0:0270	0:0929	0:0090	0:0299	0:0010
n = 400						
Truel Model	0:0072	0:0001	0:0092	0:0001	0:0034	0:0000
Linear Model	0:7925	0:6283	0:9263	0:8580	0:9777	0:9558
Double Kernel	0:1206	0:0156	0:0726	0:0054	0:0236	0:0006

5 An Application to Labor Survey Data

In this section, the estimator ® is calculated for a sample from the Labor and Manpower Activity Survey of 1989 in Canada (LMAS89). The data set consists of 8254 observations in Ontario on the wages and various demographic characteristics which are education, gender, job length, age, marital status, children and birth place (born in Canada or not). We want to examine the extent to which expected job length a®ects wage earnings. Thus, we use the logaritm of annual wage earnings (y) as the dependent variable of interest, we let x denote education and gender, s denote the job length (job tenure) and z denote age, marital status, number of children and place of birth. Then E(sjz) is the expected job length conditional on the variables of z. This decomposition of the conditioning variables is a somewhat arbitrary, but it can be justi¯ed on the grounds that the z variables a®ect salary only to the extent that they determine expected job length.

Table 3 lists the overall information for this data set of total 8254 observations. We can see that the ln(wage) values are in the range between 2.4849 and 13.7441, that is, the annual wage in falls in the range [$12; $931; 079]. The education variable takes values from 1 to 7 which denote the following categories; 1: 0 to 8 years; 2: some secondary education; 3: graduated from high school; 4: some post-secondary; 5: post-secondary certi¯cate or diploma; 6: university degree; 7: trades certi¯cate or diploma. The index of 1 in Gender stands for Female, 0 for Male. The job length is in the range between 0 and 44 years. The age is grouped as 8 groups which take values 1: 16 years; 2: 17-19 years; 3: 20-24; 4: 25-34; 5: 35-44; 6: 45-54; 7: 55-64; 8: 65-69 years.

Table 4 reports the results of estimator ®, standard deviation and its t-statistic. It shows that the impact of wage on the expected job length is signi¯cant. The above example illustrates the usefulness of the method which avoids imposing a linear speci¯cation of the auxilliary equation. As it was seen by the limited Monte Carlo results reported in the previous section this can have serious consequences for the quality of the estimates if the linearity assumption is incorrect.

Table 3: A Summary of Labor Survey Data

Variable	Size(N)	Mean	StdDev	Minimum	Maximum
Ln(Wage)	8254	9:7122	1:1276	2:4849	13:7441
Education	8254	3:7894	1:6321	1	7
Gender	8254	0:4850	0:4998	0	1
Job Length(year)	8254	6:7001	7:9036	0	44
Age	8254	4:5394	1:4569	1	8
Marital Status	8254	0:6461	0:4782	0	1
Kids	8254	0:7122	0:9903	0	6
Birth Place	8254	0:8190	0:3850	0	1

Table 4: Double Kernel Estimate Results

	EstimatorValue	StdDev	MSE	T ¡ Statistic
®	0:10008	0:003182		31:4453
$\overset{1}{E}$(sjz)	6:74977	3:540385	46:0167	

6 Conclusion

In this paper we derive the asymptotic distribution of a double kernel nonparametric estimator for a partially linear semiparametric model with generated regressors among the variables of the linear part. The generated regressor is in this case expressed as the conditional mean of certain exogenous variables (the auxilliary regression). This conditional mean is left to be of an unknown function form. We assume a correlated error structure between the partial linear equation of interest and the auxilliary equation. Monte Carlo evidence suggests that the proposed estimation behaves quite well in samples of moderate size. The usefulness of the estimator is illustrated using a sample of Canadian data.

Appendix A

First, we define $g(z_i) = E(s_i|z_i)$, $w_i = s_i - E(s_i|z_i) \equiv s_i - g(z_i)$, $\bar{s}_i = \bar{E}(s_i|z_i)$, $\hat{\bar{s}}_i = \hat{\bar{E}}(s_i|z_i)$, and $v_i = g(z_i) - h(x_i)$ with $E(v_i|x_i) = 0$. Note that $\bar{s}_i$ estimates $g(z_i)$ and $\hat{\bar{s}}_i$ estimates $E[E(s_i|z_i)|x_i] = E[g(z_i)|x_i] \equiv h(x_i)$.

Lemma 1 Let $X_i = (x_i', g_i)$ and $\hat{X} = (\hat{X}_1', \hat{X}_2', \ldots, \hat{X}_n')'$. Also $w_i = s_i - g(z_i)$ and $E(w_i|z_i) = 0$.

(i) $S_{\hat{X}} = \Sigma_1 + o_p(1)$;

(ii) $S_{\hat{X},(g-\bar{s})\alpha} = -n^{-1}\sum w_i \eta_i + o_p(n^{-1/2})$;

(iii) $S_{\hat{X},u} = n^{-1}\sum X_i u_i + o_p(n^{-1/2})$,

where $\eta_i = 2\alpha f(z_i) E(X_j|z_j = z_i)$.

Proof of (i): Notice that $S_{\hat{X}} = S_X + S_{\hat{X}-X} + 2S_{X,\hat{X}-X}$ and $\hat{X}_i - X_i = (0', \bar{s} - g)$. Then $S_{\hat{X}-X} \sim S_{\bar{s}-g} = S_{\bar{g}-g+\bar{w}} = S_{\bar{g}-g} + S_{\bar{w}} + 2S_{\bar{g}-g,\bar{w}} = o_p(n^{-1/2})$ by Proposition 3, 4 of Appendix B and Cauchy Inequality. Also $S_X = \Sigma_1 + o_p(1)$ by exactly the same arguments as in the proof of lemma A.6 of Fan and Li (1996), and $S_{X,\hat{X}-X} \le \{S_X S_{\hat{X}-X}\}^{1/2} = \{O_p(1)o_p(1)\}^{1/2} = o_p(1)$. Therefore $S_{\hat{X}} = \Sigma_1 + o_p(1)$.

Proof of (ii): $S_{\hat{X},(g-\bar{s})\alpha} = S_{X,(g-\bar{s})\alpha} + S_{\hat{X}-X,(g-\bar{s})\alpha} = S_{X,(g-\bar{g})\alpha-\bar{w}\alpha} + S_{\hat{X}-X,(g-\bar{s})\alpha} \sim S_{X,(g-\bar{g})\alpha} - S_{X,\bar{w}\alpha} + S_{g-\bar{s},(g-\bar{s})\alpha} \sim -S_{X,\bar{w}\alpha} + o_p(n^{-1/2})$ by the same proof in (i) above and in Proposition 9 of Appendix B. Then we show that $S_{X,\bar{w}\alpha} = n^{-1}\sum w_i \eta_i$ as follows.

$$S_{X,\bar{w}\alpha} = n^{-1}\sum_i X_i \bar{w}_i \alpha = \frac{\alpha}{n^2 b^q}\sum_i X_i \sum_{j\neq i} w_j \bar{K}_{ji} = \frac{1}{n^2}\sum_i \sum_{j>i} H(D_i, D_j)$$

where $H(D_i, D_j) = b^{-q}(X_i w_j + X_j w_i)\alpha \bar{K}_{ij}$ and $D_i = (x_i, z_i, w_i)$.

Note that $E(w_i|z_i) = 0$ and $H(D_i, D_j)$ is symmetric. By H-decomposition of U-statistic, we have $S_{X,\bar{w}\alpha} = \frac{2}{n}\sum_i H(D_i) + o(1)$, where

$$\begin{aligned}
H(D_i) &= \int H(D_i, D_j)\,dF(D_j) \\
&= \int b^{-q}(X_i w_j + X_j w_i)\alpha \bar{K}_{ij}\,dF(D_j) \\
&= \alpha b^{-q} X_i \int w_j \bar{K}_{ij}\,dF(D_j) + \alpha b^{-q} w_i \int X_j \bar{K}_{ij}\,dF(D_j) \\
&= \alpha b^{-q} X_i E_j[w_j \bar{K}_{ij}] + \alpha b^{-q} w_i \int X_j \bar{K}_{ij}\,dF(D_j) \\
&= \alpha b^{-q}\left\{ X_i E_j[\bar{K}_{ij} E(w_j|z_j)] + w_i \int X_j \bar{K}_{ij} f(X_j, z_j)\,dX_j\,dz_j \right\} \\
&= \alpha b^{-q} w_i \int X_j \bar{K}\left(\frac{z_j - z_i}{b}\right) f(X_j|z_j) f(z_j)\,dX_j\,dz_j
\end{aligned}$$

$$= \alpha w_i \int X_j K^1(t) f(X_j|z_j = z_i + bt) f(z_i + bt) dX_j dt$$
$$= \alpha w_i \int X_j f(X_j|z_j = z_i) f(z_i) dX_j \int K^1(t) dt + o(1)$$
$$= \alpha w_i f(z_i) E[X_j|z_j = z_i] + o(1)$$

Therefore, $S_{\hat{X};(g-\tilde{s})\alpha} = -\frac{2}{n}\sum_i \alpha w_i f(z_i) E[X_j|z_j = z_i] + o_p(n^{-1/2}) = -\frac{1}{n}\sum_i w_i \eta_i + o_p(n^{-1/2})$.

Proof of (iii): $S_{\hat{X};u} = S_{X;u} + S_{\hat{X}-X;u} \sim S_{X;u} + S_{\tilde{s}-g;u}$. Notice that $S_{\tilde{s}-g;u} = S_{\tilde{g}-g+\tilde{w};u} = S_{\tilde{g}-g;u} + S_{\tilde{w};u} = o_p(n^{-1/2})$ by Proposition 14 and 17 of Appendix B. So $S_{\hat{X};u} = n^{-1}\sum X_i u_i + o_p(n^{-1/2})$.

Proof of Theorem 1.

By the proof of Lemma 1, we have $S_{\hat{X};(g-\tilde{s})\alpha+u} = n^{-1}\sum_i (X_i u_i - w_i \eta_i) + o_p(n^{-1/2})$. Since $E(X_i u_i - w_i \eta_i) = 0$ and the variance of $(X_i u_i - w_i \eta_i)$ is equal to $E[(X_i u_i - w_i \eta_i)'(X_i u_i - w_i \eta_i)] = E[X_i' X_i \sigma^2] + E[w_i^2 \eta_i' \eta_i] - 2E[u_i w_i X_i' \eta_i] = \Omega_1 + \Omega_2 - 2\Omega_{12}$, by Lindeberg Central Limit Theorem, we have $\sqrt{n}(\hat{\delta} - \delta) = \sqrt{n} S_{\hat{X}}^{-1} S_{\hat{X};u+(g-\tilde{s})\alpha} \xrightarrow{d} N(0; \Sigma_1)$.

Appendix B

As defined in Appendix A, we have $w_i = s_i - E(s_i|z_i) \equiv s_i - g(z_i)$, $\tilde{s}_i = \tilde{E}(s_i|z_i) \xrightarrow{p} g(z_i)$, $\hat{s}_i = \hat{E}(s_i|z_i) \xrightarrow{p} E[E(s_i|z_i)|x_i] = E[g(z_i)|x_i] \equiv h(x_i)$, and $v_i = g(z_i) - h(x_i)$ with $E(v_i|x_i) = 0$. Therefore we have

$$s_i = g_i + w_i \qquad \text{where} \quad E(w_i|z_i) = 0$$

$$\tilde{s}_i = \tilde{g}_i + \tilde{w}_i = g_i + (\tilde{g}_i - g_i) + \tilde{w}_i$$

$$\hat{s}_i = \hat{g}_i + \hat{w}_i = h_i + (\hat{h}_i - h_i) + (\hat{g}_i - h_i) + \hat{w}_i$$

$$\tilde{s}_i - \hat{s}_i = (g_i - h_i) + (\tilde{g}_i - g_i) + \tilde{w}_i - (\hat{h}_i - h_i) - (\hat{g}_i - \hat{h}_i) - \hat{w}_i$$

Proposition 1 $S_{\theta-\hat{\theta}} = O_p\left(a^{2\epsilon} + \frac{1}{na^p}\right) = o_p(n^{-1/2})$, that is, $E[(\hat{\theta}_1 - \theta_1)^2] = O_p\left(a^{2\epsilon} + \frac{1}{na^p}\right)$

$$\begin{aligned}
E[(\hat{\theta}_1 - \theta_1)^2] &= E\left[\frac{\frac{1}{na^p}\sum_{i\neq 1}(\theta_i - \theta_1)K_{i1}}{\hat{f}_{x_1}}\right]^2 \\
&= E\left[\frac{1}{na^p}\sum_{i\neq 1}(\theta_i - \theta_1)K_{i1}\left(\frac{1}{f_{x_1}} + \left(\frac{1}{\hat{f}_{x_1}} - \frac{1}{f_{x_1}}\right)\right)\right]^2 \\
&= E\left[\frac{1}{na^p}\sum_{i\neq 1}(\theta_i - \theta_1)K_{i1}\left(\frac{1}{f_{x_1}} + \frac{f_{x_1} - \hat{f}_{x_1}}{f_{x_1}^2} + \frac{(f_{x_1} - \hat{f}_{x_1})^2}{f_{x_1}^2}\frac{1}{\hat{f}_{x_1}}\right)\right]^2 \\
&= E\left[\frac{\frac{1}{na^p}\sum_{i\neq 1}(\theta_i - \theta_1)K_{i1}}{f_{x_1}}\right]^2 + (s.o.) \\
&\leq \left(\frac{1}{na^p}\right)^2 c^2 \sum_{i\neq 1}\sum_{j\neq 1} E[(\theta_i - \theta_1)(\theta_j - \theta_1)K_{i1}K_{j1}] \\
&= \left(\frac{1}{na^p}\right)^2 c^2 \sum_{i=j\neq 1} E[(\theta_i - \theta_1)^2 K_{i1}^2] \\
&\quad + \left(\frac{1}{na^p}\right)^2 c^2 \sum_{i\neq j; i\neq 1}\sum_{j\neq 1} E[(\theta_i - \theta_1)(\theta_j - \theta_1)K_{i1}K_{j1}] \\
&\leq \frac{c^2}{(na^p)^2}(na^p + n^2 a^{p+\epsilon} a^{p+\epsilon}) = O\left(a^{2\epsilon} + \frac{1}{na^p}\right)
\end{aligned}$$

where (s.o.) is a smaller order than the fpreceeding item. This is because it is true that $\hat{f}(x) \xrightarrow{p} f(x)$, i.e., $\hat{f}(x) - f(x) = O((na^p)^{-1/2})$, and the density function of $f(x)$ is bounded by assumption. So, $\frac{(f_x - \hat{f}_x)}{f_x} \xrightarrow{p} 0$ as $n \to \infty$, $a \to 0$ and $na^p \to \infty$. Hence, $\frac{(f_x - \hat{f}_x)^k}{f_x^k} \xrightarrow{p} 0$ for all k. It is obvious that the series of $\sum_{k=1}^{\infty} \frac{(f_x - \hat{f}_x)^k}{f_x^k}$ is a convergent series, and goes to zero if as $n \to \infty$, $a \to 0$ and $na^p \to \infty$. Therefore, the small order of (s.o) comes from

$$
\begin{aligned}
& E\left[\frac{1}{na^p}\sum_{i\neq 1}(\theta_i - \theta_1)K_{i1}\left(\frac{1}{f_{x_1}} + \frac{f_{x_1} - \hat{f}_{x_1}}{f_{x_1}^2} + \frac{(f_{x_1} - \hat{f}_{x_1})^2}{f_{x_1}^2}\frac{1}{\hat{f}_{x_1}}\right)\right]^2 \\
&= E\left[\frac{1}{na^p}\sum_{i\neq 1}(\theta_i - \theta_1)K_{i1}\frac{1}{f_{x_1}}\left(1 + \sum_{k=1}^{\infty}\frac{(f_{x_1} - \hat{f}_{x_1})^k}{f_{x_1}^k}\right)\right]^2 \\
&= E\left[\frac{1}{na^p}\sum_{i\neq 1}(\theta_i - \theta_1)K_{i1}\frac{1}{f_{x_1}}(1 + o(1))\right]^2 \\
&= E\left[\frac{\frac{1}{na^p}\sum_{i\neq 1}(\theta_i - \theta_1)K_{i1}}{f_{x_1}}\right]^2 + (s.o.)
\end{aligned}
$$

It is the same argument for (s.o.) in the following discussions.

Similar to proof of Proposition 1, we have Proposition 2 and 3 as follows:

Proposition 2 $S_{h-\hat{h}} = O_p\left(a^{2\nu} + \frac{1}{na^p}\right) = o_p(n^{-1/2})$.

Proposition 3 $S_{g-\hat{g}} = O_p\left(b^{2\mu} + \frac{1}{nb^q}\right) = o_p(n^{-1/2})$.

Proposition 4 $S_{\tilde{r}} = o_p(n^{-1/2})$, ($r = w$, or u).

We first consider the case of $r = w$.

$$
\begin{aligned}
E|S_{\tilde{w}}| &= \frac{1}{n}\sum_i E[\tilde{w}_i^2] = E[\tilde{w}_1^2] = E\{\frac{1}{f_{z_1}^2}[(nb^q)^{-1}\sum_{i\neq 1} w_i \tilde{K}_{1i}][(nb^q)^{-1}\sum_{j\neq 1} w_j \tilde{K}_{1j}]\} \\
&\leq M_z E\{[(nb^q)^{-2}\sum_{i\neq 1} w_i^2 \tilde{K}_{1i}^2]\} + s.o. \\
&= E\{[n^{-1}b^{-2q}E_1(w_2^2\tilde{K}_{12}^2)]\} + s.o. = O((nb^q)^{-1}); \qquad (B.1)
\end{aligned}
$$

by Lemma 1 of Robinson (1998). We also used the fact that $E(w_i|Z, X_{-i}) = 0$, where $X_{-i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$. Obviously (B.1) also proves $S_{\tilde{w}} = O_p((nb^q)^{-1})$ (by bounded f_1^2).

The proof is identical to the case of $r = w$, simply replacing w by u in the above proof. Similarly we have the following proposition:

Proposition 5 $S_{\hat{v}} = o_p(n^{-1/2})$.

Proposition 6 $S_{\tilde r} = o_p(n^{-1/2})$, $(r = w \text{ or } u)$.

Proof: For the case of $r = w$,
$E|S_{\hat w}| = \frac{1}{n}\sum_i E[\hat w_i^2] = E[\hat w_1^2] = [n^4a^{2p}b^{2q}]^{-1}\sum_{i\neq 1}\sum_{j\neq i}\sum_{i'\neq 1}\sum_{j'\neq i'} E[w_j w_{j'}\tilde K_{ij}\tilde K_{i'j'}K_{1i}K_{1i'}\frac{1}{\hat f_{x_1}^2\tilde f_{z_i}\tilde f_{z_{i'}}}] = [n^4a^{2p}b^{2q}]^{-1}\sum_{i\neq 1}\sum_{j\neq i,i'}\sum_{i'\neq 1} E[w_j^2\tilde K_{ij}\tilde K_{i'j}K_{1i}K_{1i'}\frac{1}{\hat f_{x_1}^2\tilde f_{z_i}\tilde f_{z_{i'}}}] = [n^4a^{2p}b^{2q}]^{-1}\{\sum_{i\neq 1}\sum_{j\neq i,i'}\sum_{i'\neq 1}(n-1)(n-1)(n-3)E[w_3^2\tilde K_{23}\tilde K_{43}K_{12}K_{14}] + (n-1)(n-2)E[w_3^2\tilde K_{23}^2K_{12}^2] + s.o. = O(n^{-1} + (n^2a^pb^q)^{-1}) = O(n^{-1})$. The same proof leads to $E|S_{\hat u}| = O(n^{-1})$.

Proposition 7 $S_{\hat g - \hat h} = o_p(n^{-1/2})$.

By the definition of $g_i = h_i + v_i$, where $E[v_i|x_i] = 0$, we have $\hat h_i = \tilde g_i - \tilde v_i$. Then $\hat g_i - \hat h_i = \hat g_i - \tilde g_i + \tilde v_i$.

$$\begin{aligned}
\hat g_i - \hat h_i &= \hat g_i - \tilde g_i + \tilde v_i = \frac{1}{na^p}\sum_{j\neq i}(\tilde g_j - g_j)K_{ji}\frac{1}{\hat f_{x_i}} + \tilde v_i \\
&= \frac{1}{na^p}\frac{1}{nb^q}\sum_{j\neq i}\sum_{l\neq j}(g_l - g_j)\tilde K_{lj}K_{ji}\frac{1}{\hat f_{x_i}\tilde f_{z_j}} + \tilde v_i \\
&= T_i + \tilde v_i
\end{aligned}$$

where $T_i = \frac{1}{na^p}\frac{1}{nb^q}\sum_{j\neq i}\sum_{l\neq j}(g_l - g_j)\tilde K_{lj}K_{ji}\frac{1}{\hat f_{x_i}\tilde f_{z_j}}$.
Hence $S_{\hat g - \hat h} = S_{T+\tilde v}$. By the Cauchy inequality, we only need to show that $S_T = o_p(n^{-1/2})$ and $S_{\tilde v} = o_p(n^{-1/2})$. Obviously Proposition 5 gives the result of the latter. So we only prove the former below.

Proof for $S_T = O((na^pnb^q)^{-1} + b^{2\mu})$:
Using $T_1 = [(na^p)^{-1}(nb^q)^{-1}\sum_{i\neq 1}\sum_{j\neq i}K_{1i}(g(z_j) - g(z_i))\tilde K_{ij}\frac{1}{\hat f_{x_1}\tilde f_{z_i}}]$, we have

$$\begin{aligned}
E|S_T| &= E(T_1^2) \\
&= E\{\frac{1}{n^4a^{2p}b^{2q}}\sum_{i\neq 1}\sum_{j\neq i}\sum_{i'\neq 1}\sum_{j'\neq i'}\frac{1}{\hat f_{x_1}^2\tilde f_{z_i}\tilde f_{z_{i'}}}K_{1i}(g_j - g_i)\tilde K_{ij}K_{1i'}(g_{j'} - g_{i'})\tilde K_{i'j'}\} \\
&= E\{\frac{1}{n^4a^{2p}b^{2q}}\sum_{i\neq 1}\sum_{j\neq i}\sum_{i'\neq 1}\sum_{j'\neq i'}\frac{1}{f_{x_1}^2 f_{z_i} f_{z_{i'}}}K_{1i}(g_j - g_i)\tilde K_{ij}K_{1i'}(g_{j'} - g_{i'})\tilde K_{i'j'}\} + (s.o.) \\
&\le \frac{c}{n^4a^{2p}b^{2q}}\sum_{i\neq 1}\sum_{j\neq i}\sum_{i'\neq 1}\sum_{j'\neq i'}E[K_{1i}(g_j - g_i)\tilde K_{ij}K_{1i'}(g_{j'} - g_{i'})\tilde K_{i'j'}] + (s.o.) \\
&= I_g + (s.o.)
\end{aligned}$$

Case (1), all are different for $i; j; i'; j'$:

$$I_g = \frac{cn^4}{n^4a^{2p}b^{2q}}E[K_{12}(g_4 - g_2)\tilde K_{24}K_{13}(g_5 - g_3)\tilde K_{35}]$$

$$
\begin{aligned}
&= \frac{c}{a^{2p}b^{2q}} E[E_2[K_{12}](g_4 - g_2)\bar{K}_{24} E_3[K_{13}](g_5 - g_3)\bar{K}_{35}] \\
&= \frac{c}{a^{2p}b^{2q}} O(a^{2p}) E[E_2[(g_4 - g_2)\bar{K}_{24}] E_3[(g_5 - g_3)\bar{K}_{35}]] \quad \text{by Lemma 2; Robinson (1988)} \\
&= \frac{c}{a^{2p}b^{2q}} O(a^{2p}) O(b^{2p+2\epsilon}) \quad \text{by Lemma 1; Li (1996)} \\
&= O(b^{2\epsilon})
\end{aligned}
$$

Case (2), $j \neq i$ but $i' \neq i$, $j' = j$ or $i' \neq j$, $j' = i$. We only prove the case of $i' \neq i$, $j' = j$:

$$
\begin{aligned}
I_g &= \frac{cn^3}{n^4 a^{2p}b^{2q}} E[K_{12}(g_4 - g_2)\bar{K}_{24} K_{13}(g_4 - g_3)\bar{K}_{34}] \\
&= \frac{c}{na^{2p}b^{2q}} E[E_2[K_{12}](g_4 - g_2)\bar{K}_{24} E_3[K_{13}](g_4 - g_3)\bar{K}_{34}] \\
&= \frac{c}{na^{2p}b^{2q}} O(a^{2p}) E[E_4[(g_4 - g_2)\bar{K}_{24}] E_4[(g_4 - g_3)\bar{K}_{34}]] \quad \text{by Lemma 2; Robinson (1988)} \\
&= \frac{c}{na^{2p}b^{2q}} O(a^{2p}) O(b^{2p+2\epsilon}) \quad \text{by Lemma 1; Li (1996)} \\
&= O\left(\frac{b^{2\epsilon}}{n}\right)
\end{aligned}
$$

Case (3), $j \neq i$ but $i' = i$, $j' = j$ or $i' = j$, $j' = i$. We only prove the case of $i' = i$, $j' = j$:

$$
\begin{aligned}
I_g &= \frac{cn^2}{n^4 a^{2p}b^{2q}} E[K_{12}^2 (g_3 - g_2)^2 \bar{K}_{32}^2] = \frac{c}{n^2 a^{2p}b^{2q}} E[E_2[K_{12}^2](g_3 - g_2)^2 \bar{K}_{32}^2] \\
&= \frac{c}{n^2 a^{2p}b^{2q}} O(a^p) E[(g_3 - g_2)^2 \bar{K}_{32}^2] = O\left(\frac{1}{n^2 a^p b^q}\right)
\end{aligned}
$$

In summary, we have shown that $E|S_T| = E(T_1^2) = O(b^{2\epsilon}) + O(\frac{b^{2\epsilon}}{n}) + O(\frac{1}{n^2 a^p b^q})$. Therefore, $S_T = O_p(b^{2\epsilon} + \frac{1}{n^2 a^p b^q}) = o_p(n^{-1/2})$.

Proposition 8 $S_{g-h;\theta-\hat{\theta}} = o_p(n^{-1/2})$

$$
\begin{aligned}
E(S^2_{g-h;\theta-\hat{\theta}}) &= \frac{1}{n^2} \sum_i \sum_j E[(g_i - h_i)(g_j - h_j)(\theta_i - \hat{\theta}_i)(\theta_j - \hat{\theta}_j)] \\
&= \frac{1}{n^2} \sum_{i=1} E[(g_i - h_i)^2 (\theta_i - \hat{\theta}_i)^2] \\
&\quad + \frac{1}{n^2} \sum_i \sum_{j \neq i} E[(g_i - h_i)(g_j - h_j)(\theta_i - \hat{\theta}_i)(\theta_j - \hat{\theta}_j)] \\
&\leq \frac{1}{n^2} nM E[(\theta_1 - \hat{\theta}_1)^2] \\
&\quad + \frac{1}{n^2} \sum_i \sum_{j \neq i} E\{(\theta_i - \hat{\theta}_i)(\theta_j - \hat{\theta}_j) E[(g_i - h_i)(g_j - h_j)|X]\} \\
&= O\left(\frac{1}{n}\right) O\left(a^{2\gamma} + \frac{1}{na^p}\right) = O\left(\frac{a^{2\gamma}}{n} + \frac{1}{n^2 a^p}\right)
\end{aligned}
$$

where $X = (x_1; x_2; \ldots; x_n)$, and using Proposition 1 and $E[(g_i - h_i)(g_j - h_j)|X] = E[(g_i - h_i)|X]E[(g_j - h_j)|X] = E[v_i|X]E[v_j|X] = 0$ by $i \neq j$ independence and $E[v_i|x_i] = 0$. It follows that $S_{g-h;\theta-\hat\theta} = O\left(\frac{\gamma}{\sqrt{n}} + \frac{1}{n\sqrt{a}}\right) = o_p(n^{-1/2})$.

Proposition 9 $S_{g-h;g-\bar g} = o_p(n^{-1/2})$.

$$
\begin{aligned}
E[S^2_{g-h;g-\bar g}] &= \frac{1}{n^2}\sum_i E[(g_i - h_i)^2(g_i - \bar g_i)^2] \\
&\quad + \frac{1}{n^2}\sum_i\sum_{j\neq i} E[(g_i - h_i)(g_j - h_j)(g_i - \bar g_i)(g_j - \bar g_j)] \\
&\leq \frac{c}{n}E[(g_1 - \bar g_1)^2] \\
&\quad + \frac{1}{n^2}\sum_i\sum_{j\neq i} E[v_i v_j \frac{1}{nb^q}\sum_{i'\neq i}(g_i - g_{i'})\frac{\bar K_{ii'}}{\tilde f_{z_i}}\frac{1}{nb^q}\sum_{j'\neq j}(g_j - g_{j'})\frac{\bar K_{jj'}}{\tilde f_{z_j}}] \\
&= O\left(\frac{1}{n}(b^{2\mu} + \frac{1}{nb^q})\right) \\
&\quad + \frac{1}{n^4 b^{2q}}\sum_i\sum_{j\neq i}\sum_{i'\neq i}\sum_{j'\neq j} E[v_i v_j (g_i - g_{i'})\bar K_{ii'}(g_j - g_{j'})\bar K_{jj'}\frac{1}{f_{z_i} f_{z_j}}] + (s.o.) \\
&= O\left(\frac{b^{2\mu}}{n} + \frac{1}{n^2 b^q}\right) + I_{vg1} + (s.o.)
\end{aligned}
$$

Case (1), all are different for $i \neq j \neq i' \neq j'$:

$$
\begin{aligned}
I_{vg1} &= \frac{1}{n^4 b^{2q}} n^4 E[v_1 v_2 \frac{1}{f_{z_1} f_{z_2}} E[(g_1 - g_3)\bar K_{13}|x_1; z_1]E[(g_2 - g_4)\bar K_{24}|x_2; z_2]] \\
&\leq \frac{M_z}{b^{2q}} E[|v_1 v_2|]O(b^{q+\mu})O(b^{q+\mu}) = O(b^{2\mu}) \quad \text{by Lemma 1; Li (1996)}
\end{aligned}
$$

Case (2), for taking three of $i; j; i'; j'$ different, there are two of them: $i' = j$ or $j' = i$. We only prove the case of $i' = j$ because of symetric.

$$
\begin{aligned}
I_{vg1} &= \frac{1}{n^4 b^{2q}} n^3 E[v_1 v_2 \frac{1}{f_{z_1} f_{z_2}}(g_1 - g_2)\bar K_{12} E[(g_2 - g_4)\bar K_{24}|x_2; z_2]] \\
&\leq \frac{M_z}{n b^{2q}} O(b^q)O(b^{q+\mu}) = O(\frac{b^\mu}{n})
\end{aligned}
$$

Case (3), for taking two of $i; j; i'; j'$ different, there is a case where $i \neq j$: $i' = j$, $j' = i$.

$$
\begin{aligned}
I_{vg1} &= \frac{1}{n^4 b^{2q}} n^2 E[v_1 v_2 \frac{1}{f_{z_1} f_{z_2}}(g_1 - g_2)\bar K_{12}(g_2 - g_1)\bar K_{21}] \\
&\leq \frac{M_z}{n^2 b^{2q}} E[v_1 v_2 (g_1 - g_2)^2 \bar K_{12}^2] \\
&= O\left(\frac{1}{n^2 b^{2q}}\right) O(b^q) = O(\frac{1}{n^2 b^q})
\end{aligned}
$$

In summary, we have $E[S^2_{g-h;g-\tilde g}] \le O\left(\frac{b^{2\nu}}{n} + \frac{1}{n^2b^q}\right) + I_n = O\left(\frac{b^{2\nu}}{n} + \frac{1}{n^2b^q}\right) + O(b^{2\nu}) + O(\frac{b^\nu}{n}) + O(\frac{1}{n^2b^q})$. Then $S_{g-h;g-\tilde g} \le O\left(\frac{b^\nu}{\sqrt{n}} + \frac{1}{n\sqrt{b^q}}\right) + O(b^\nu) + O(\frac{\sqrt{b^\nu}}{\sqrt{n}}) + O(\frac{1}{n\sqrt{b^q}})$. So, $S_{g-h;g-\tilde g} = o_p(n^{-1/2})$.

Proposition 10 $S_{g-h;\tilde w} = o_p(n^{-1/2})$.

$$
\begin{aligned}
E(S^2_{g-h;\tilde w}) &= \frac{1}{n^2}\sum_i\sum_j E[(g_i - h_i)(g_j - h_j)\tilde w_i \tilde w_j] \\
&= \frac{1}{n^2}\sum_{i=j} E[(g_i - h_i)^2 \tilde w_i^2] + \frac{1}{n^2}\sum_{i\neq j}\sum_j E[(g_i - h_i)(g_j - h_j)\tilde w_i \tilde w_j] \\
&\le \frac{c}{n}E[S_{\tilde w}] + \frac{1}{n^4b^{2q}}\sum_{i\neq j}\sum_j\sum_{i'\neq i}\sum_{j'\neq j} E[\frac{1}{\tilde f_{z_i}\tilde f_{z_j}} v_i v_j w_{i'} \tilde K_{ii'} w_{j'} \tilde K_{jj'}] \\
&= \frac{c}{n}E[S_{\tilde w}] + \frac{1}{n^4b^{2q}}\sum_{i\neq j}\sum_j\sum_{i'\neq i}\sum_{j'\neq j} E[\frac{1}{f_{z_i} f_{z_j}} v_i v_j w_{i'} \tilde K_{ii'} w_{j'} \tilde K_{jj'}] + (s.o.) \\
&\le \frac{c}{n}O(\frac{1}{nb^q}) + \frac{c}{n^4b^{2q}}\sum_{i\neq j}\sum_j\sum_{i'\neq i}\sum_{j'\neq j} E[v_i v_j w_{i'} \tilde K_{ii'} w_{j'} \tilde K_{jj'}] \\
&= O(\frac{1}{n^2b^q}) + \frac{c}{n^4b^{2q}}\sum_{i\neq j}\sum_j\sum_{i'\neq i}\sum_{j'\neq j} E[E[v_i v_j|X] w_{i'} \tilde K_{ii'} w_{j'} \tilde K_{jj'}] \\
&= O(\frac{1}{n^2b^q})
\end{aligned}
$$

where $X = (x_1, x_2, ..., x_n)$, and using proposition 4 and $E[(g_i - h_i)(g_j - h_j)|X] = E[(g_i - h_i)|X]E[(g_j - h_j)|X] = 0$ by $i \neq j$ independence. It follows that $S_{g-h;\tilde w} = O_p\left(\frac{1}{n\sqrt{b^q}}\right) = o_p(n^{-1/2})$.

Proposition 11 $S_{g-h;\hat r} = o_p(n^{-1/2})$, where $r = v$ or $r = u$.

We only prove the case of $r = v$, the same as the case $r = u$.

$$
\begin{aligned}
E(S^2_{g-h;\hat v}) &= \frac{1}{n^2}\sum_i\sum_j E[(g_i - h_i)(g_j - h_j)\hat v_i \hat v_j] \\
&= \frac{1}{n^2}\sum_{i=j} E[(g_i - h_i)^2 \hat v_i^2] + \frac{1}{n^2}\sum_{i\neq j}\sum_j E[(g_i - h_i)(g_j - h_j)\hat v_i \hat v_j] \\
&\le \frac{c}{n}E[S_{\hat v}] + E[(g_1 - h_1)(g_2 - h_2)\hat v_1 \hat v_2] \\
&= \frac{c}{n}E[S_{\hat v}] + \frac{1}{n^2a^{2p}}\sum_{i\neq 1}\sum_{j\neq 2} E[\frac{1}{\hat f_{x_1}\hat f_{x_2}} v_1 v_2 K_{i1} K_{j2} v_i v_j]
\end{aligned}
$$

$$
\begin{aligned}
&\le \frac{c}{n}E[S_{\hat v}] + \frac{c_1}{n^2a^{2p}}E[v_1v_2K_{i1}K_{j2}v_iv_j] + (s.o.) \\
&= O(n^{-2}(a^p)^{-1}) + I_v \\
I_v &= O(\frac{1}{n^2a^{2p}})\sum_{i\ne 1}\sum_{j\ne 2} E[v_1v_2K_{i1}K_{j2}v_iv_j]
\end{aligned}
$$

It is clear to see that $I_v = 0$ because for $i \ne j$, $E[v_1v_2K_{i1}K_{j2}v_iv_j] = E\{E[v_1K_{i1}v_i|i]E[v_2K_{j2}v_j|j]\}$ $= E\{(E[K_{i1}v_iE[v_1|x_1,i])|i]E[(K_{j2}v_jE[v_2|x_2,j])|j]\} = 0$ by $E[v_1|x_1,i] = E[v_1|x_1] = 0$ or $E[v_2|x_2,j] = E[v_2|x_2] = 0$. If $i = j$ (but $i,j \ne 1,2$), then $E[v_1v_2K_{i1}K_{i2}v_i^2] = E\{E[v_1v_2K_{i1}K_{i2}|i]v_i^2\}$ $= E\{v_i^2E[v_1K_{i1}|i]E[v_2K_{i2}|i]\} = E\{v_i^2E[(K_{i1}E[v_1|x_1,i])|i]E[(K_{i2}E[v_2|x_2,i])|i]\} = 0$ by $E(v_i|x_i) = 0$. Thus, $S_{g-h,\hat v} = o_p(n^{-1/2})$.

Proposition 12 $S_{g-h,\hat w} = o_p(n^{-1/2})$.

$$
\begin{aligned}
E(S^2_{g-h,\hat w}) &= \frac{1}{n^2}\sum_i\sum_j E[(g_i-h_i)(g_j-h_j)\hat w_i\hat w_j] \\
&= \frac{1}{n^2}\sum_{i=j} E[(g_i-h_i)^2\hat w_i^{\,2}] + \frac{1}{n^2}\sum_{i\ne j}\sum_j E[(g_i-h_i)(g_j-h_j)\hat w_i\hat w_j] \\
&\le \frac{c}{n}E[S_{\hat w}] + E[(g_1-h_1)(g_2-h_2)\hat w_1\hat w_2] \\
&= \frac{c}{n}E[S_{\hat w}] + \frac{1}{n^4a^{2p}b^{2q}}\sum_{i\ne 1}\sum_{j\ne 2}\sum_{i'\ne i}\sum_{j'\ne j} E[\frac{1}{\hat f_{x_1}\hat f_{x_2}\tilde f_{z_i}\tilde f_{z_j}}v_1v_2K_{i1}K_{j2}w_{i'}\bar K_{ii'}w_{j'}\bar K_{jj'}] \\
&\le \frac{c}{n}E[S_{\hat w}] + \frac{c_1}{n^4a^{2p}b^{2q}}E[v_1v_2K_{i1}K_{j2}w_{i'}\bar K_{ii'}w_{j'}\bar K_{jj'}] + (s.o.) \\
&= O(n^{-2}) + I_{vw} \\
I_{vw} &= O(\frac{1}{n^4a^{2p}b^{2q}})\sum_{i\ne 1}\sum_{j\ne 2}\sum_{i'\ne i}\sum_{j'\ne j} E[v_1v_2K_{i1}K_{j2}w_{i'}\bar K_{ii'}w_{j'}\bar K_{jj'}]
\end{aligned}
$$

Case (1), all are different for $1,2,i,j,i',j'$. Then $I_{vw} = O(\frac{1}{a^{2p}b^{2q}})E[v_1v_2K_{i1}K_{j2}w_{i'}\bar K_{ii'}w_{j'}\bar K_{jj'}] =$ $O(\frac{1}{a^{2p}b^{2q}})E\{E[v_1v_2K_{i1}K_{j2}|i,j]E[w_{i'}\bar K_{ii'}w_{j'}\bar K_{jj'}|i,j]\} = 0$, where $E[v_1v_2K_{i1}K_{j2}|i,j] =$ $E[v_1K_{i1}|i]E[v_2K_{j2}|j] = E\{K_{i1}E[v_1|1,i]|i\}E\{K_{j2}E[v_2|2,j]|j\} = 0$ by $E[v_i|x_i] = 0$.

Case (2), all are different for $1,2,i,j,i',j'$ except for one pair. We only prove the case of $i' = 1$ because of the same argument: $I_{vw} = O(\frac{1}{na^{2p}b^{2q}})E[v_1v_2K_{i1}K_{j2}w_1\bar K_{i1}w_{j'}\bar K_{jj'}] =$ $O(\frac{1}{na^{2p}b^{2q}})E\{E[v_1K_{i1}w_1\bar K_{i1}v_2K_{j2}|i,j]E[w_{j'}\bar K_{jj'}|i,j]\} = 0$, where $E[v_1K_{i1}w_1\bar K_{i1}v_2K_{j2}|i,j] =$ $E[v_1K_{i1}w_1\bar K_{i1}|i]E[v_2K_{j2}|j] = E[v_1K_{i1}w_1\bar K_{i1}|i]E\{K_{j2}E[v_2|2,j]|j\} = 0$ by $E[v_2|x_2] = 0$.

Case (3), all are different for $1; 2; i; j; i'; j'$ except for two pairs. We only prove the case of $i' = 1; j' = 2$ because of the same argument: $I_{vw} = O(\frac{1}{n^2a^{2p}b^{2q}})E[v_1v_2K_{i1}K_{j2}w_1\bar{K}_{i1}w_2\bar{K}_{j2}] = O(\frac{1}{n^2a^{2p}b^{2q}})E\{v_1w_1E[K_{i1}\bar{K}_{i1}|1]v_2w_2E[K_{j2}\bar{K}_{j2}|2]\} = O(\frac{1}{n^2a^{2p}b^{2q}})O(a^{2p}b^{2q}) = O(n^{-2})$.

Case (4), all are different for $1; 2; i; j; i'; j'$ except for three pairs. We only prove the case of $i' = 1; j' = 1; i' = j'$ because of the same argument: $I_{vw} = O(\frac{1}{n^3a^{2p}b^{2q}})E[v_1v_2K_{i1}K_{j2}w_1^2\bar{K}_{i1}\bar{K}_{j1}] = O(\frac{1}{n^3a^{2p}b^{2q}})E\{v_1w_1^2E[K_{i1}\bar{K}_{i1}\bar{K}_{j1}|1]v_2E[K_{j2}|2]\} = O(\frac{1}{n^3a^{2p}b^{2q}})O(a^pb^{2q}) = O(\frac{1}{n^3a^p}) = O(n^{-2})$.
Thus, $E(S^2_{g-h;\hat{w}}) = O(n^{-2})$. That is, $S_{g-h;\hat{w}} = o_p(n^{-1/2})$.

Proposition 13 $S_{g-h;\hat{g}-\hat{h}} = o_p(n^{-1/2})$.

$$E(S^2_{g-h;\hat{g}-\hat{h}}) = \frac{1}{n^2}\sum_i\sum_j E[(g_i - h_i)(g_j - h_j)(\hat{g}_i - \hat{h}_i)(\hat{g}_j - \hat{h}_j)]$$
$$= \frac{1}{n^2}\sum_{i=j} E[(g_i - h_i)^2(\hat{g}_i - \hat{h}_i)^2] + \frac{1}{n^2}\sum_{i\neq j}\sum_j E[(g_i - h_i)(g_j - h_j)(\hat{g}_i - \hat{h}_i)(\hat{g}_j - \hat{h}_j)]$$
$$\leq \frac{c}{n}E[S_{\hat{g}-\hat{h}}] + E[(g_1 - h_1)(g_2 - h_2)(\hat{g}_1 - \hat{h}_1)(\hat{g}_2 - \hat{h}_2)]$$
$$= \frac{c}{n}E[S_{\hat{g}-\hat{h}}] + \frac{1}{n^4a^{2p}b^{2q}}\sum_{i\neq 1}\sum_{j\neq 2}\sum_{i'\neq i}\sum_{j'\neq j} E[\frac{v_1v_2}{\hat{f}_{x_1}\hat{f}_{x_2}\tilde{f}_{z_i}\tilde{f}_{z_j}}K_{i1}(g_{i'} - g_i)\bar{K}_{ii'}K_{j2}(g_{j'} - g_j)\bar{K}_{jj'}]$$
$$\leq \frac{c}{n}E[S_{\hat{g}-\hat{h}}] + \frac{c_1}{n^4a^{2p}b^{2q}}\sum_{i\neq 1}\sum_{j\neq 2}\sum_{i'\neq i}\sum_{j'\neq j} E[v_1v_2K_{i1}(g_{i'} - g_i)\bar{K}_{ii'}K_{j2}(g_{j'} - g_j)\bar{K}_{jj'}] + (s.o.)$$
$$= O(n^{-1})O(b^{2\gamma} + (n^2a^pb^q)^{-1}) + I_{vg}$$
$$I_{vg} = O(\frac{1}{n^4a^{2p}b^{2q}})\sum_{i\neq 1}\sum_{j\neq 2}\sum_{i'\neq i}\sum_{j'\neq j} E[v_1v_2K_{i1}(g_{i'} - g_i)\bar{K}_{ii'}K_{j2}(g_{j'} - g_j)\bar{K}_{jj'}]$$

Case (1), all are different for $1; 2; i; j; i'; j'$. Then $I_{vg} = O(\frac{1}{a^{2p}b^{2q}})E[v_1v_2K_{i1}K_{j2}(g_{i'} - g_i)\bar{K}_{ii'}(g_{j'} - g_j)\bar{K}_{jj'}] = O(\frac{1}{a^{2p}b^{2q}})E\{E[v_1v_2K_{i1}K_{j2}|i; j]E[(g_{i'} - g_i)\bar{K}_{ii'}(g_{j'} - g_j)\bar{K}_{jj'}|i; j]\} = 0$
where $E[v_1v_2K_{i1}K_{j2}|i; j] = E[v_1K_{i1}|i]E[v_2K_{j2}|j] = E\{K_{i1}E[v_1|1; i]|i\}E\{K_{j2}E[v_2|2; j]|j\} = 0$ by $E[v_i|x_i] = 0$.

Case (2), all are different for $1; 2; i; j; i'; j'$ except for one pair. We only prove the case of $i' = 1$ because of the same argument: $I_{vg} = O(\frac{1}{na^{2p}b^{2q}})E[v_1v_2K_{i1}K_{j2}(g_1 - g_i)\bar{K}_{i1}(g_{j'} - g_j)\bar{K}_{jj'}] = O(\frac{1}{na^{2p}b^{2q}})E\{E[v_1K_{i1}(g_1 - g_i)\bar{K}_{i1}v_2K_{j2}|i; j]E[(g_{j'} - g_j)\bar{K}_{jj'}|i; j]\} = 0$, where $E[v_1K_{i1}(g_1 - g_i)\bar{K}_{i1}v_2K_{j2}|i; j] = E[v_1K_{i1}(g_1 - g_i)\bar{K}_{i1}|i]E[v_2K_{j2}|j] = E[v_1K_{i1}(g_1 - g_i)\bar{K}_{i1}|i]E\{K_{j2}E[v_2|2; j]|j\} =$

0 by $E[v_2|x_2] = 0$.

Case (3), all are different for $1, 2, i, j, i', j'$ except for two pairs. We only prove the case of $i' = 1, j' = 2$ because of the same argument: $I_{vg} = O(\frac{1}{n^2a^{2p}b^{2q}})E[v_1v_2K_{i1}K_{j2}(g_1 - g_i)\bar{K}_{i1}(g_2 - g_i)\bar{K}_{j2}] = O(\frac{1}{n^2a^{2p}b^{2q}})E\{v_1(g_1 - g_i)E[K_{i1}\bar{K}_{i1}|1]v_2(g_2 - g_i)E[K_{j2}\bar{K}_{j2}|2]\} = O(\frac{1}{n^2a^{2p}b^{2q}})O(a^{2p}b^{2q}) = O(n^{-2})$.

Case (4), all are different for $1, 2, i, j, i', j'$ except for three pairs. We only prove the case of $i' = 1, j' = 1, i' = j'$ because of the same argument: $I_{vg} = O(\frac{1}{n^3a^{2p}b^{2q}})E[v_1v_2K_{i1}K_{j2}(g_1 - g_i)^2\bar{K}_{i1}\bar{K}_{j1}] = O(\frac{1}{n^3a^{2p}b^{2q}})E\{v_1(g_1 - g_i)^2E[K_{i1}\bar{K}_{i1}\bar{K}_{j1}|1]v_2E[K_{j2}|2]\} = O(\frac{1}{n^3a^{2p}b^{2q}})O(a^pb^{2q}) = O(\frac{1}{n^3a^p}) = O(n^{-2})$. Thus, $E(S^2_{g-h,\hat{g}-\hat{h}}) = O(n^{-1})O(b^{2\gamma} + (n^2a^pb^q)^{-1}) + O(n^{-2})$. That is, $S_{g-h,\hat{g}-\hat{h}} = o_p(n^{-1/2})$.

Proposition 14 $S_{\bar{g}-g,u} = o_p(n^{-1/2})$.

Using the independence of $\{u_i\}$ and $\{x_i, z_i\}$, and $E(u_i|x_i, z_i) = 0$ as well, $E[u_i(g_i - \bar{g}_i)u_j(g_j - \bar{g}_j)] = 0$ for $i \neq j$. Then we have

$$
\begin{aligned}
E[S^2_{g-\bar{g},u}] &= \frac{1}{n^2}\sum_i E[u_i^2(g_i - \bar{g}_i)^2] + \frac{1}{n^2}\sum_i\sum_{j\neq i} E[u_i(g_i - \bar{g}_i)u_j(g_j - \bar{g}_j)] \\
&= \frac{1}{n}E[u_1^2(g_1 - \bar{g}_1)^2] \\
&= \frac{1}{n}E[u_1^2\frac{1}{n^2b^{2q}\tilde{f}^2_{z_1}}\sum_{i\neq 1}\sum_{j\neq 1}(g_i - g_1)\bar{K}_{i1}(g_j - g_1)\bar{K}_{j1}] \\
&= \frac{1}{n}E[u_1^2\frac{1}{n^2b^{2q}f^2_{z_1}}\sum_{i\neq 1}\sum_{j\neq 1}(g_i - g_1)\bar{K}_{i1}(g_j - g_1)\bar{K}_{j1}] + (s.o.) \\
&\leq \frac{c}{n^3b^{2q}}\sum_{i\neq 1}\sum_{j\neq 1}E[u_1^2(g_i - g_1)\bar{K}_{i1}(g_j - g_1)\bar{K}_{j1}] + (s.o.) \\
&= \frac{c}{n^2b^{2q}}E[u_1^2(g_2 - g_1)^2\bar{K}_{21}] + \frac{c}{nb^{2q}}E[u_1^2(g_2 - g_1)\bar{K}_{21}(g_3 - g_1)\bar{K}_{31}] + (s.o.) \\
&= O(\frac{1}{n^2b^q}) + \frac{c}{nb^{2q}}E\{u_1^2E_1[(g_2 - g_1)\bar{K}_{21}]E_1[(g_3 - g_1)\bar{K}_{31}]\} \\
&= O(\frac{1}{n^2b^q}) + O(\frac{b^{2\mu}}{n}) \\
S_{g-\bar{g},u} &= O(\frac{1}{n\sqrt{b^q}}) + O(\frac{b^{\mu}}{\sqrt{n}}) = o_p(n^{-1/2}).
\end{aligned}
$$

Proposition 15 $S_{\hat{h}-h,u} = o_p(n^{-1/2})$.

Similar arguments to Proposition 14, we can show that $S_{\hat{h}-h;u} = O(\frac{1}{n\sqrt{a^p}}) + O(\frac{a^{\gamma}}{\sqrt{n}}) = o_p(n^{-1/2})$.

Proposition 16 $S_{\hat{g}-\hat{h};u} = o_p(n^{-1/2})$.

As we used in proof of Proposition 14 that the independence of $\{u_i\}$ and $\{x_i; z_i\}$, and $E(u_i|x_i; z_i) = 0$, we have

$$
\begin{aligned}
E(S^2_{\hat{g}-\hat{h};u}) &= \frac{1}{n}E[u_1^2(\hat{g}_1 - \hat{h}_1)^2] \\
&= \frac{1}{n^3a^{2p}\hat{f}^2_{x_1}}\sum_{i\neq 1}\sum_{j\neq 1} E[u_1^2(\tilde{g}_i - h_i)K_{i1}(\tilde{g}_j - g_j)K_{j1}] \\
&= \frac{1}{n^5a^{2p}b^{2q}\hat{f}^2_{x_1}}\sum_{i\neq 1}\sum_{j\neq 1}\sum_{i'\neq i}\sum_{j'\neq j}\frac{1}{\tilde{f}_{z_i}\tilde{f}_{z_j}}E[u_1^2K_{i1}(g_{i'} - g_i)\bar{K}_{ii'}K_{j1}(g_{j'} - g_j)\bar{K}_{jj'}] \\
&= \frac{1}{n^5a^{2p}b^{2q}f^2_{x_1}}\sum_{i\neq 1}\sum_{j\neq 1}\sum_{i'\neq i}\sum_{j'\neq j}\frac{1}{f_{z_i}f_{z_j}}E[u_1^2K_{i1}(g_{i'} - g_i)\bar{K}_{ii'}K_{j1}(g_{j'} - g_j)\bar{K}_{jj'}] + (s.o.) \\
&\leq \frac{c}{n^5a^{2p}b^{2q}}\sum_{i\neq 1}\sum_{j\neq 1}\sum_{i'\neq i}\sum_{j'\neq j}E[u_1^2K_{i1}(g_{i'} - g_i)\bar{K}_{ii'}K_{j1}(g_{j'} - g_j)\bar{K}_{jj'}] + (s.o.) \\
&= \frac{c}{n^4a^{2p}b^{2q}}\sum_{i'\neq 2}\sum_{j'\neq 2}E[u_1^2K^2_{21}(g_{i'} - g_2)\bar{K}_{2i'}(g_{j'} - g_2)\bar{K}_{2j'}] \\
&+ \frac{c}{n^3a^{2p}b^{2q}}\sum_{i'\neq 2}\sum_{j'\neq 3}E[u_1^2K_{21}K_{31}(g_{i'} - g_2)\bar{K}_{2i'}(g_{j'} - g_3)\bar{K}_{3j'}] + (s.o.) \\
&= I_{i=j} + I_{i\neq j}
\end{aligned}
$$

Case (1), all are different for $i'; j'; 1; 2; 3$:

$I_{i=j} = O(\frac{1}{n^4a^{2p}b^{2q}})\sum_{i'\neq 2}\sum_{j'\neq 2}E[u_1^2K^2_{21}(g_{i'} - g_2)\bar{K}_{2i'}(g_{j'} - g_2)\bar{K}_{2j'}] = O(\frac{1}{n^4a^{2p}b^{2q}})O(n^2)$ $E\{u_1^2E_1[K^2_{21}]E_2[(g_{i'} - g_2)\bar{K}_{2i'}]E_2[(g_{j'} - g_2)\bar{K}_{2j'}]\} = O(\frac{1}{n^4a^{2p}b^{2q}})O(n^2)O(a^p)O(b^{q+\mu})O(b^{q+\mu})E[u_1^2] = O(\frac{b^{2\mu}}{n^2a^p})$.

$I_{i\neq j} = O(\frac{1}{n^3a^{2p}b^{2q}})\sum_{i'\neq 2}\sum_{j'\neq 3}E[u_1^2K_{21}K_{31}(g_{i'} - g_2)\bar{K}_{2i'}(g_{j'} - g_3)\bar{K}_{3j'}] = O(\frac{1}{na^{2p}b^{2q}})$ $E\{u_1^2E_1[K_{21}]E_1[K_{31}]E_2[(g_{i'} - g_2)\bar{K}_{2i'}]E_3[(g_{j'} - g_3)\bar{K}_{3j'}]\} = O(\frac{b^{2\mu}}{n})$.

Case (2), all are different except for one pairs: $i' = j'$ or $i' = 1$ or $j' = 1$ or $i' = 3$ or $j' = 2$ (in $I_{i\neq j}$). We only prove the cases of $i' = j'$ and $i' = 1$ for the reason of similarity.

For $i' = j'$: $I_{i=j} = O(\frac{1}{n^3a^{2p}b^{2q}})E[u_1^2K^2_{21}(g_3 - g_2)^2\bar{K}^2_{23}] = O(\frac{1}{n^3a^pb^q})$.

$I_{i\neq j} = O(\frac{1}{n^2a^{2p}b^{2q}})E[u_1^2K_{21}K_{31}(g_4 - g_2)\bar{K}_{24}(g_4 - g_3)\bar{K}_{34}] = O(\frac{b^{2\mu}}{n^2})$.

For $i' = 1$: $I_{i=j} = O(\frac{1}{n^3a^{2p}b^{2q}})E\{u_1^2K^2_{21}(g_1 - g_2)\bar{K}_{21}E_{21}[(g_3 - g_2)\bar{K}_{21}]\} = O(\frac{b^{\mu}}{n^3a^{2p}b^q})E\{u_1^2K^2_{21}(g_1 - g_2)\bar{K}_{21}\} = O(\frac{b^{\mu}}{n^3a^p})$.

$I_{i\neq j} = O(\frac{1}{n^2a^{2p}b^{2q}})E[u_1^2K_{21}K_{31}(g_1 - g_2)\tilde{K}_{21}(g_4 - g_3)\tilde{K}_{34}] = O(\frac{b^{\epsilon}}{n^2a^pb^q})E[u_1^2K_{21}(g_1 - g_2)\tilde{K}_{21}] = O(\frac{b^{\epsilon}}{n^2})$.

Case (3), only two pairs are equal: $i' = j' = 1$ or $i' = j; j' = i$.

For case $i' = j' = 1$: $I_{i=j} = O(\frac{1}{n^3a^{2p}b^{2q}})E[u_1^2K_{21}^2(g_1 - g_2)^2\tilde{K}_{21}^2] = O(\frac{1}{n^3a^pb^q})$.

$I_{i\neq j} = O(\frac{1}{n^2a^{2p}b^{2q}})E\{u_1^2E_1[K_{21}(g_1 - g_2)\tilde{K}_{21}]E_1[K_{31}(g_1 - g_3)\tilde{K}_{31}]\} = O(\frac{1}{n^2})$.

For case $i' = j; j' = 1$: $I_{i=j}$ not applicable because $i = j$ and $i' \neq i$.

$I_{i\neq j} = O(\frac{1}{na^{2p}b^{2q}})E\{u_1^2E[K_{21}K_{31}|x_1;Z](g_3 - g_2)\tilde{K}_{23}(g_1 - g_3)\tilde{K}_{31}\} = O(\frac{1}{nb^{2q}})$

$E\{u_1^2(g_1 - g_3)\tilde{K}_{31}E[(g_3 - g_2)\tilde{K}_{23}|1;3]\} = O(\frac{b^{\epsilon}}{nb^q})E\{u_1^2(g_1 - g_3)\tilde{K}_{31}\} = O(\frac{b^{2\epsilon}}{n})$.

From the cases above, we can say that $I_{i=j} + I_{i\neq j} = o_p(n^{-1/2})$. Thus, $S_{\hat{g}-\hat{h};u} = o_p(n^{-1/2})$.

Proposition 17 $S_{\bar{w};u} = o_p(n^{-1/2})$.

$$\begin{aligned}
E(S_{\bar{w};u}^2) &= \frac{1}{n^2}\sum_i\sum_j E[u_iu_j)\bar{w}_i\bar{w}_j] = \frac{1}{n}E[u_1^2\bar{w}_1^2] \\
&= \frac{1}{n^3b^{2q}}\sum_{i\neq 1}\sum_{j\neq 1} E[\frac{1}{\tilde{f}_{z_1}\tilde{f}_{z_1}}u_1^2w_i\tilde{K}_{i1}w_j\tilde{K}_{j1}] \\
&= \frac{1}{n^2b^{2q}}E[\frac{1}{f_{z_1}^2}u_1^2w_2^2\tilde{K}_{21}] + \frac{1}{nb^{2q}}E[\frac{1}{f_{z_1}^2}u_1^2w_2\tilde{K}_{21}w_3\tilde{K}_{31}] + (s.o.) \\
&\leq \frac{c}{n^2b^{2q}}E[u_1^2w_2^2\tilde{K}_{21}] + \frac{c}{nb^{2q}}E[u_1^2w_2\tilde{K}_{21}w_3\tilde{K}_{31}] + (s.o.) \\
&= O(\frac{1}{n^2b^{2q}})E[u_1^2w_2^2\tilde{K}_{21}] = O(\frac{1}{n^2b^q});
\end{aligned}$$

where we used $E[u_1^2w_2\tilde{K}_{21}w_3\tilde{K}_{31}] = E\{u_1^2\tilde{K}_{21}\tilde{K}_{31}E[w_2|1;2;3]E[w_3|1;2;3]\} = 0$ by $E[w_i|z_i] = 0$. Hence, $S_{\bar{w};u} = O(\frac{1}{n\sqrt{b^q}}) = o_p(n^{-1/2})$.

Proposition 18 $S_{\hat{\bar{w}};u} = o_p(n^{-1/2})$.

$$\begin{aligned}
E(S_{\hat{\bar{w}};u}^2) &= \frac{1}{n^2}\sum_i\sum_j E[u_iu_j)\hat{\bar{w}}_i\hat{\bar{w}}_j] = \frac{1}{n}E[u_1^2\hat{\bar{w}}_1^{\,2}] \\
&= \frac{1}{n^5a^{2p}b^{2q}}\sum_{i\neq 1}\sum_{j\neq 1}\sum_{i'\neq i}\sum_{j'\neq j} E[\frac{1}{\hat{f}_{x_1}^2\tilde{f}_{z_i}\tilde{f}_{z_j}}u_1^2K_{i1}K_{j1}w_{i'}\tilde{K}_{ii'}w_{j'}\tilde{K}_{jj'}] \\
&= \frac{1}{n^5a^{2p}b^{2q}}\sum_{i\neq 1}\sum_{j\neq 1}\sum_{i'\neq i}\sum_{j'\neq j} E[\frac{1}{f_{x_1}^2f_{z_i}f_{z_j}}u_1^2K_{i1}K_{j1}w_{i'}\tilde{K}_{ii'}w_{j'}\tilde{K}_{jj'}] + (s.o.)
\end{aligned}$$

$$
\begin{aligned}
&\le \frac{c}{n^5a^{2p}b^{2q}} \sum_{i\neq 1}\sum_{j\neq 1}\sum_{i'\neq i}\sum_{j'\neq j} E[u_1^2K_{i1}K_{j1}w_{i'}\bar{K}_{ii'}w_{j'}\bar{K}_{jj'}] + (s.o.) \\
&= \frac{c}{n^4a^{2p}b^{2q}} \sum_{i'\neq 2}\sum_{j'\neq 2} E[u_1^2K_{21}^2w_{i'}\bar{K}_{2i'}w_{j'}\bar{K}_{2j'}] \\
&+ \frac{c}{n^3a^{2p}b^{2q}} \sum_{i'\neq 2}\sum_{j'\neq 3} E[u_1^2K_{21}K_{31}w_{i'}\bar{K}_{2i'}w_{j'}\bar{K}_{3j'}] + (s.o.) \\
&= A_{i=j} + A_{i\neq j} + (s.o.)
\end{aligned}
$$

Case (1), all are different for $i'; j'; 1; 2; 3$. Then

$A_{i=j} = \frac{c}{n^2a^{2p}b^{2q}} E[u_1^2K_{21}^2w_3\bar{K}_{23}w_4\bar{K}_{24}] = \frac{c}{n^2a^{2p}b^{2q}} E\{u_1^2K_{21}^2E[w_3\bar{K}_{23}w_4\bar{K}_{24}|1;2]\} = \frac{c}{n^2a^{2p}b^{2q}}$
$E\{u_1^2K_{21}^2E[w_3\bar{K}_{23}|1;2]E[w_4\bar{K}_{24}|1;2]\} = \frac{c}{n^2a^{2p}b^{2q}} E\{u_1^2K_{21}^2E[(E[w_3|3]\bar{K}_{23})|1;2]E[(E[w_4|4]\bar{K}_{24})|1;2]\}$
$= 0$ by $E[w_i|z_i] = 0$.

$A_{i\neq j} = \frac{c}{na^{2p}b^{2q}} E[u_1^2K_{21}K_{31}w_4\bar{K}_{24}w_5\bar{K}_{35}] = 0$ by the same arguments above.

Case (2), all are different except for one pairs: $i' = j'$ or $i' = 1$ or $j' = 1$ or $i' = 3$ or $j' = 2$ (in $A_{i\neq j}$). We only prove the cases of $i' = j'$ and $i' = 1$ for the reason of similarity.

For $i' = j'$: $A_{i=j} = O(\frac{1}{n^3a^{2p}b^{2q}})E[u_1^2K_{21}^2w_3^2\bar{K}_{23}^2] = O(\frac{1}{n^3a^pb^q})$.

$A_{i\neq j} = O(\frac{1}{n^2a^{2p}b^{2q}})E[u_1^2K_{21}K_{31}w_4^2\bar{K}_{24}\bar{K}_{34}] = O(\frac{1}{n^2})$.

For $i' = 1$: $A_{i=j} = O(\frac{1}{n^3a^{2p}b^{2q}})E\{u_1^2K_{21}^2w_1\bar{K}_{21}w_3\bar{K}_{32}\} = 0$ by $E[w_3\bar{K}_{32}|1;2]$
$= E\{(E[w_3|1;2;z_3]\bar{K}_{32})|1;2\} = 0$.

Case (3), only two pairs are equal: $i' = j' = 1$ or $i' = j; j' = i$.

For case $i' = j' = 1$: $A_{i=j} = O(\frac{1}{n^3a^{2p}b^{2q}})E[u_1^2K_{21}^2w_1^2\bar{K}_{21}^2] = O(\frac{1}{n^3a^pb^q})$.

$A_{i\neq j} = O(\frac{1}{n^2a^{2p}b^{2q}})E\{u_1^2E[K_{21}K_{31}|1]w_1^2E[\bar{K}_{21}\bar{K}_{31}|1]\} = O(\frac{1}{n^2})$.

For case $i' = j; j' = 1$: $A_{i=j}$ not applicable because $i = j$ and $i' \neq i$.

$A_{i\neq j} = O(\frac{1}{na^{2p}b^{2q}})E\{u_1^2K_{21}K_{31}w_3\bar{K}_{23}w_1\bar{K}_{31}\} = 0$ by $E[w_i|z_i] = 0$.

From the cases above, we can say that $A_{i=j} + A_{i\neq j} = o_p(n^{-1/2})$. So, $S_{\hat{w};u} = o_p(n^{-1/2})$.

Proof of Theorem 2

First we prove Theorem 2 (i). Notice that $S_{\tilde{s}-\hat{s}} = S_{(g-h)+(\tilde{g}-g)+w-(\hat{h}-h)-(\hat{\tilde{g}}-\hat{h})-\hat{w}} =$
$S_{v+(\tilde{g}-g)+w-(\hat{h}-h)-(\hat{\tilde{g}}-\hat{h})-\hat{w};v+(\tilde{g}-g)+w-(\hat{h}-h)-(\hat{\tilde{g}}-\hat{h})-\hat{w}} = S_v + 2S_{v;(\tilde{g}-g)+w-(\hat{h}-h)-(\hat{\tilde{g}}-\hat{h})-\hat{w}} = S_v + o_p(1)$
by Cauchy Inequality and privious proofs of Propositions. Therefore, $S_v \xrightarrow{p} E[(g_1 - h_1)'(g_1 - h_1)]$ by Proposition 1 in Li (1996). Thus, $S_{\tilde{s}-\hat{s}} \xrightarrow{p} E[(g_1 - h_1)'(g_1 - h_1)]$.

Then we prove Theorem 2 (ii). Notice that

$$
\begin{aligned}
\sqrt{n}S_{\bar{s}-\hat{\bar{s}};u} &= \sqrt{n}S_{(g-h)+(\bar{g}-g)+w-(\hat{h}-h)-(\hat{\bar{g}}-\hat{h})-\hat{w};u} \\
&= \sqrt{n}S_{v;u}+\sqrt{n}S_{\bar{g}-g;u}-\sqrt{n}S_{\hat{h}-h;u}-\sqrt{n}S_{\hat{\bar{g}}-\hat{h};u}+\sqrt{n}S_{w;u}-\sqrt{n}S_{\hat{w};u} \\
&= \sqrt{n}S_{v;u}+o_p(1) \quad \text{by Propositions } 14-18. \\
&\stackrel{d}{\to} N(0;\Sigma). \quad \text{by Levy CLT under (A1) (A2).}
\end{aligned}
$$

where the last equality is obtained by using Proposition 15 in Robinson (1988). $\Sigma = \sigma^2 E[(g_1-h_1)'(g_1-h_1)]$, where $g_1 = g(z_1) = E(s_1|z_1)$ and $h_1 = h(x_1) = E[g(z_1)|x_1] = E\{E[s_1|z_1]|x_1\}$.

Finally in order to prove Theorem 2 (iii), we need to show that

(a) $I = \sqrt{n}S_{\bar{s}-\hat{\bar{s}};\theta-\hat{\theta}} = o_p(1)$;

(b) $II = \sqrt{n}S_{\bar{s}-\hat{\bar{s}};g-\bar{s}} = o_p(1)$;

(c) $III = \sqrt{n}S_{\bar{s}-\hat{\bar{s}};\hat{g}-\hat{\bar{s}}} = o_p(1)$;

(d) $IV = \sqrt{n}S_{\bar{s}-\hat{\bar{s}};\hat{u}} = o_p(1)$.

Then $\sqrt{n}S_{\bar{s}-\hat{\bar{s}};(g-\bar{s})\alpha} = o_p(1)$, and $\sqrt{n}S_{\bar{s}-\hat{\bar{s}};(\hat{g}-\hat{\bar{s}})\alpha} = o_p(1)$. So, Theorem 2 (iii) holds.

(a). Proof of $I = \sqrt{n}S_{\bar{s}-\hat{\bar{s}};(\theta-\hat{\theta})} = o_p(1)$.

$$
\begin{aligned}
I &= \sqrt{n}S_{\bar{s}-\hat{\bar{s}};\theta-\hat{\theta}} = \sqrt{n}S_{(g-h)+(\bar{g}-g)+w-(\hat{h}-h)-(\hat{\bar{g}}-\hat{h})-\hat{w};\theta-\hat{\theta}} \\
&= \sqrt{n}S_{g-h;\theta-\hat{\theta}}+\sqrt{n}S_{\bar{g}-g;\theta-\hat{\theta}}-\sqrt{n}S_{\hat{h}-h;\theta-\hat{\theta}}-\sqrt{n}S_{\hat{\bar{g}}-\hat{h};\theta-\hat{\theta}}+\sqrt{n}S_{w;\theta-\hat{\theta}}-\sqrt{n}S_{\hat{w};\theta-\hat{\theta}} \\
&= I_1+I_2+I_3+I_4+I_5+I_6
\end{aligned}
$$

Proof of $I_1 = \sqrt{n}S_{g-h;\theta-\hat{\theta}} = o_p(1)$.
By proposition 8, it follows that

$$
I_1 = \sqrt{n}S_{g-h;\theta-\hat{\theta}} = \sqrt{n}O_p\left(\frac{a^\gamma}{\sqrt{n}}+\frac{1}{n\sqrt{a^p}}\right) = O_p\left(a^\gamma+\frac{1}{\sqrt{na^p}}\right) = o_p(1).
$$

Proof of $I_2 = \sqrt{n}S_{\bar{g}-g;\theta-\hat{\theta}} = o_p(1)$.
Notice that $S_{\bar{g}-g;\theta-\hat{\theta}} \le 2(S_{\bar{g}-g}+S_{\theta-\hat{\theta}})$ by Cauchy Inequality. Then we have $E[S_{\bar{g}-g;\theta-\hat{\theta}}] \le 2\{E[S_{\bar{g}-g}]+E[S_{\theta-\hat{\theta}}]\} = 2\{E[\frac{1}{n}\sum_i(\bar{g}_i-g_i)^2]+E[\frac{1}{n}\sum_i(\theta_i-\hat{\theta}_i)^2]\} = 2\{E[(\bar{g}_1-g_1)^2]+E[(\theta_1-\hat{\theta}_1)^2]\}$. Therefore it is true that $E[S_{\bar{g}-g;\theta-\hat{\theta}}] = O\left(b^{2\mu}+\frac{1}{nb^q}\right)+O\left(a^{2\gamma}+\frac{1}{na^p}\right)$ by Proposition 1 and 3. That is, $I_2 = o_p(1)$.

The proof of $I_3 = o_p(1)$ is the same as the proof of I_2 by using Proposition 1 and 2.

Proof of $I_5 = \sqrt{n} S_{\bar w, \mu - \hat\mu} = o_p(1)$.

By Cauchy Inequality, and using Proposition 1 and 4, it is clear that

$$S_{\bar w, \mu - \hat\mu} \le 2\{S_{\bar w} + S_{\mu - \hat\mu}\} = o_p(n^{-1/2}) + O\left(a^{2\gamma} + \frac{1}{na^p}\right)$$

That is, $I_5 = \sqrt{n} S_{\bar w, \mu - \hat\mu} = o_p(1)$.

Proof of $I_6 = \sqrt{n} S_{\hat{\bar w}, \mu - \hat\mu} = o_p(1)$.

By Cauchy Inequality, and using Proposition 1 and 6, it is proved.

Finally, we use Proposition 7 to have $I_4 = \sqrt{n} S_{\hat{\bar g} - \hat h, \mu - \hat\mu} = o_p(1)$.

(b) Proof of $II = \sqrt{n} S_{\bar s - \hat{\bar s}, \bar g - \bar s} = o_p(1)$.

$$\begin{aligned} II &= \sqrt{n} S_{\bar s - \hat{\bar s}, \bar g - \bar s} = \sqrt{n} S_{(g-h)+(\bar g - g)+\bar w - (\hat h - h) - (\hat{\bar g} - \hat h) - \hat{\bar w}, (g - \bar g) - \bar w} \\ &= \sqrt{n}\{S_{g-h, g-\bar g} - S_{\bar g - g} - S_{\hat h - h, g - \bar g} - S_{\hat{\bar g} - \hat h, g - \bar g} - S_{\hat{\bar w}, g - \bar g} \\ &\quad - S_{g-h, \bar w} - 2S_{\bar g - g, \bar w} - S_{\bar w} + S_{\hat h - h, \bar w} + S_{\hat{\bar g} - \hat h, \bar w} + S_{\hat{\bar w}, \bar w}\} \end{aligned}$$

For $\sqrt{n} S_{g-h, g-\bar g}$ and $\sqrt{n} S_{g-h, \bar w}$, we can show that they both equal to $o_p(1)$ by Proposition 9 and 10. For the rest of them, by using Cauchy Inequality and Propositions of 1-7, all are proved.

(c) Proof of $III = \sqrt{n} S_{\bar s - \hat{\bar s}, \hat{\bar g} - \hat{\bar s}} = o_p(1)$.

$$\begin{aligned} III &= \sqrt{n} S_{\bar s - \hat{\bar s}, \hat{\bar g} - \hat{\bar s}} = \sqrt{n} S_{(g-h)+(\bar g - g)+\bar w - (\hat h - h) - (\hat{\bar g} - \hat h) - \hat{\bar w}, -(\hat{\bar g} - \hat h) + \bar v - \hat{\bar w}} \\ &= \sqrt{n} S_{g-h, \hat h - \hat{\bar g}} + \sqrt{n} S_{\bar g - g, \hat h - \hat{\bar g}} - \sqrt{n} S_{\hat h - h, \hat h - \hat{\bar g}} + \sqrt{n} S_{\hat{\bar g} - \hat h} + \sqrt{n} S_{\bar w, \hat h - \hat{\bar g}} - \sqrt{n} S_{\hat{\bar w}, \hat h - \hat{\bar g}} \\ &\quad + \sqrt{n} S_{g-h, \bar v} + \sqrt{n} S_{\bar g - g, \bar v} - \sqrt{n} S_{\hat h - h, \bar v} - \sqrt{n} S_{\hat{\bar g} - \hat h, \bar v} + \sqrt{n} S_{\bar w, \bar v} - \sqrt{n} S_{\hat{\bar w}, \bar v} \\ &\quad - \sqrt{n} S_{g-h, \hat{\bar w}} - \sqrt{n} S_{\bar g - g, \hat{\bar w}} + \sqrt{n} S_{\hat h - h, \hat{\bar w}} + \sqrt{n} S_{\hat{\bar g} - \hat h, \hat{\bar w}} - \sqrt{n} S_{\bar w, \hat{\bar w}} + \sqrt{n} S_{\hat{\bar w}} \end{aligned}$$

Using Propositions of 11-13, it is proved that $\sqrt{n} S_{g-h, \bar v} = o_p(1)$, $\sqrt{n} S_{g-h, \hat{\bar w}} = o_p(1)$, and $\sqrt{n} S_{g-h, \hat h - \hat{\bar g}} = o_p(1)$. By using Cauchy Inequality and Propositions of 1-7, the rest all are proved.

(d). Proof of $IV = \sqrt{n} S_{\bar s - \hat{\bar s}, \bar u} = o_p(1)$.

$$\begin{aligned} IV &= \sqrt{n} S_{\bar s - \hat{\bar s}, \bar u} \\ &= \sqrt{n} S_{(g-h)+(\bar g - g)+\bar w - (\hat h - h) - (\hat{\bar g} - \hat h) - \hat{\bar w}, \bar u} \\ &= \sqrt{n} S_{g-h, \bar u} + \sqrt{n} S_{\bar g - g, \bar u} - \sqrt{n} S_{\hat h - h, \bar u} - \sqrt{n} S_{\hat{\bar g} - \hat h, \bar u} + \sqrt{n} S_{\bar w, \bar u} - \sqrt{n} S_{\hat{\bar w}, \bar u} \end{aligned}$$

Proposition 11 gives us that $\sqrt{n}S_{g-h;\hat{u}} = o_p(1)$. By using Cauchy Inequality and Propositions of 1-7, the rest all are proved.

Proof of Theorem 3

We now prove Theorem 3. By Theorem 2, it is clear to have

$$
\begin{aligned}
\sqrt{n}(\hat{\alpha} - \alpha) &= \sqrt{n}S^{-1}_{\bar{s}-\hat{s}} S_{\bar{s}-\hat{s};u+(\theta-\hat{\theta})+(g-\bar{s})\alpha-(\hat{g}-\hat{s})\alpha-\hat{u}} \\
&= S^{-1}_{\bar{s}-\hat{s}}\sqrt{n}S_{\bar{s}-\hat{s};u} + S^{-1}_{\bar{s}-\hat{s}}\sqrt{n}\left(S_{\bar{s}-\hat{s};\theta-\hat{\theta}} + S_{\bar{s}-\hat{s};(g-\bar{s})\alpha} - S_{\bar{s}-\hat{s};(\hat{g}-\hat{s})\alpha} - S_{\bar{s}-\hat{s};\hat{u}}\right) \\
&= (S_{g-h})^{-1}\sqrt{n}S_{g-h;u} + o_p(1) \\
&\xrightarrow{d} (E[(g_1-h_1)'(g_1-h_1)])^{-1} N\left(0; \sigma^2 E[(g_1-h_1)'(g_1-h_1)]\right) \\
&= N\left(0; \sigma^2 (E[(g_1-h_1)'(g_1-h_1)])^{-1}\right).
\end{aligned}
$$

References

Ahn, H. and Powell, J.L. (1993) \Semiparametric Estimation of Censored Selection Models with a Nonparametric Selection Mechanism" Journal of Econometrics 58, 3-29.

Ahn, H. (1995) \Nonparametric Two-Stage Estimation of Conditional Choice Probabilities in Binary Choice Model Under Uncertainty" Journal of Econometrics 67, 337-378.

Ahn, H. (1997) \Semiparametric Estimation of a Single-Index Model with Nonparametrically Generated Regressors" Econometric Theory 13, 3-31.

Andrews, D.W.K. (1991) \Asymptotic Normality of Series Estimators for Nonparametric and Semiparametric Regression Models" Econometrica 59, 307-345.

Andrews, D.W.K. (1994) \Asymptotics for Semiparametric Econometric Models via Stochastic Equicontinuity" Econometrica, forthcoming.

Bean C.R. (1986) \The Estimation of "Surprise" Models and the "Surprise" consumption Function" Review of Economic Studies, 53, 497-516.

Davidson R. and J. Mackinnon (1981) " Several Tests for Model Speci¯cation in the Presence of of Alternative Hypotheses", Econometrica, 49, 781-793.

Delgado, M.A., Li Q. and T. Stengos (1995) \Nonparametric Speci¯cation Testing of Non-Nested Econometric Models" manuscript, Depertment of Economics, University of Guelph.

Fan, Y., Q. Li and T. Stengos, (1995) \Root-N-consistent semi-parametric regression with conditionally heteroskedastic disturbances" Journal of Quantitative Economics, Vol. 11, pp. 229-240.

Fan, Y., Q. Li, (1996) \Consistent Model Speci¯cation Tests: Omitted Variables and Semiparametric Functional Forms." Econometrica 64, 865-890.

Li, Q. (1996) \On The Ratio-N-Consistent Semiparametric Estimation of Partially Linear Models" Economics Letters, Vol. 51, 277-285.

Newey, W. K., (1994) \Kernel Estimation of Partial Means and a General Variance Estimator" Econometric Theory 10, 233-253.

Pagan A. (1984) \Econometric Issues in the Analysis of Regressions with Generated Regressors" International Economic Review 25, 221-247.

Pagan A. (1986) \Two Stage and Related Estimators and their Applications" Review of Economic Studies, 53, 517-538.

Powell J.L "Estimation of Semiparametric Models" in Handbook of Econometrics, Volume IV, edited by R. Engle and D.L McFadden, Elsevier Science, the Netherlands.

Powell, J.L., J.H. Stock, and T.M. Stocker, (1989), \Semiparametric estimation of the index coe±cients" Econometrica 57, 1043-1430.

Rilstone, P. (1996) \Nonparametric Estimation of Model with Generated Regressors" International Economic Review, Vol. 37, No. 2, 299-313.

Robinson, P., (1988) \Root-N consistent semi-parametric regression" Econometrica 156, 931-954.

SELECTED RECENT PUBLICATIONS

Bertaut, C. and Haliassos, M,. "Precautionary Portfolio Behavior from a Life - Cycle Perspective", *Journal of Economic Dynamics and Control,* 21, 1511-1542, 1997).

Caporale, G. and Pittis, N. "Causality and Forecasting in Incomplete Systems", *Journal of Forecasting*, 1997, 16, 6, 425-437.

Caporale, G. and Pittis, N. "Efficient estimation of cointegrated vectors and testing for causality in vector autoregressions: A survey of the theoretical literature", *Journal of Economic Surveys*, forthcoming.

Caporale, G. and Pittis, N. "Unit root testing using covariates: Some theory and evidence", *Oxford Bulletin of Economics and Statistics*, forthcoming.

Caporale, W., Hassapis, C. and Pittis, N. "Unit Roots and Long Run Causality: Investigating the Relationship between Output, Money and Interest Rates". *Economic Modeling*, 15(1), 91-112, January 1998.

Clerides, K., S. "Is Learning-by-Exporting Important? Micro-Dynamic Evidence from Colombia, Morocco, and Mexico." *Quarterly Journal of Economics* 113(3), pp. 903-947, August 1998, (with Lach and J.R. Tybout).

Cukierman, A., Kalaitzidakis, P., Summers, L. and Webb, S. "Central Bank Independence, Growth, Investment, and Real Rates", Reprinted in Sylvester Eijffinger (ed), *Independent Central Banks and Economic Performance,* Edward Elgar, 1997, 416-461.

Eicher, Th. and Kalaitzidakis, P. "The Human Capital Dimension to Foreign Direct Investment: Training, Adverse Selection and Firm Location". In Bjarne Jensen and Kar-yiu Wong (eds), *Dynamics,Economic Growth, and International Trade,* The University of Michigan Press, 1997, 337-364.

Gatsios, K., Hatzipanayotou, P. and Michael, M. S. "International Migration, the Provision of Public Good and Welfare", *Journal of Development Economics*, 60/2, 561-577, 1999.

Haliassos, M. "On Perfect Foresight Models of a Stochastic World", *Economic Journal*, 104, 477-491, 1994.

Haliassos, M. and Bertaut, C., "Why Do So Few Hold Stocks?", *The Economic Journal,* 105, 1110- 1129, 1995.

Haliassos, M. and Tobin, J. "The Macroeconomics of Government Finance", reprinted in J. Tobin, *Essays in Economics,* Vol. 4, Cambridge: MIT Press, 1996.

Hassapis, C., Pittis, N. and Prodromidis, K. "Unit Roots and Granger Causality in the EMS Interest Rates: The German Dominance Hypothesis Revisited", *Journal of International Money and Finance*, pp. 47-73, 1999.

Hassapis, C., Kalyvitis, S., and Pittis, N. "Cointegration and Joint Efficiency of International Commodity Markets", *The Quarterly Review of Economics and Finance,* Vol 39, pp. 213-231, 1999.

Hassapis, C., Pittis, N., and Prodromides, K. "EMS Interest Rates: The German Dominance Hypothesis or Else?" in *European Union at the Crossroads: A Critical Analysis of Monetary Union and Enlargement*, Aldershot, UK., Chapter 3, pp. 32-54, 1998. Edward Elgar Publishing Limited.

Hatzipanayotou, P. and Michael, M. S. "General Equilibrium Effects of Import Constraints Under Variable Labor Supply, Public Goods and Income Taxes", *Economica*, 66, 389-401, 1999.

Hatzipanayotou, P. and Michael, M. S. "Public Good Production, Nontraded Goods and Trade Restriction", *Southern Economic Journal,* 63, 4, 1100-1107, 1997.

Hatzipanayotou, P. and Michael, M. S. "Real Exchange Rate Effects of Fiscal Expansion Under Trade Restrictions", *Canadian Journal of Economics,* 30-1, 42-56, 1997.

Kalaitzidakis, P. "On-the-job Training Under Firm-Specific Innovations and Worker Heterogeneity", *Industrial Relations,* 36, 371-390, July 1997.

Kalaitzidakis, P., Mamuneas, Th. and Stengos, Th. "European Economics: An Analysis Based on Publications in Core Journals*." European Economic Review*,1999.

Lyssiotou, P., Pashardes, P. and Stengos, Th. "Testing the Rank of Engel Curves with Endogenous Expenditure", *Economics Letters,* 64, 1999, 61-65.

Lyssiotou, P., Pashardes, P. and Stengos, Th. "Preference Heterogeneity and the Rank of Demand Systems", *Journal of Business and Economic Statistics*, Vol 17, No 2, April 1999, 248-252.

Lyssiotou, Panayiota, "Comparison of Alternative Tax and Transfer Treatment of Children using Adult Equivalence Scales", *Review of Income and Wealth*, Series 43, No. 1 March 1997, 105-117.

Mamuneas, T.P. (with Demetriades P.). "Intertemporal Output and Employment Effects of Public Infrastructure Capital: Evidence from 12 OECD Economies", *Economic Journal,* forthcoming.

Mamuneas T.P. (with Kalaitzidakis P. and Stengos T.). "A Nonlinear Sensitivity Analysis of Cross-Country Growth Regressions", *Canadian Journal of Economics,* forthcoming.

Mamouneas T.P. (with Bougheas S. and Demetriades P.)."I΄nfrastructure, Specialization and Economic Growth", *Canadian Journal of Economics*, forthcoming.

Mamuneas, Theofanis P. "Spillovers from Publicly – Financed R&D Capital in High-Tech Industries", *International Journal of Industrial Organization,* 17(2), 215-239, 1999.

Mamuneas, T. P. (with Nadiri, M.I.). "R&D Tax Incentives and Manufacturing-Sector R&D Expenditures", in *Borderline Case: Interntational Tax Policy, Corporate Research and Development, and Investment*, James Poterba (ed.), National Academy Press, Washington D.C., 1997. Reprinted in *Chemtech*, 28(9), 1998.

Michaelides, A. and Ng, S. “Estimating the Rational Expectations Model of Speculative Storage: A Monte Carlo Comparison of three Simulation Estimators”, *Journal of Econometrics,* forthcoming.

Pashardes, Panos. “Equivalence Scales in a Rank-3 Demand System”, *Journal of Public Economics*, 58, 143-158, 1995.

Pashardes, Panos. “Bias in Estimating Equivalence Scales from Grouped Data”, *Journal of Income Distribution*, Special Issue: Symposium on Equivalence Scales, 4, 253-264, 1995.

Pashardes, Panos. “Abstention and Aggregation in Consumer Demand”, *Oxford Economic Papers*, 46, 502-518, 1994 (with V. Fry).

Pashardes, Panos. “Bias in Estimation of the Almost Ideal Demand System with the Stone Index
Approximation”, *Economic Journal*, 103, 908-916, 1993.

Pashardes, Panos. “What Do We Learn About Consumer Demand Patterns From Micro-Data?”, *American Economic Review*, 83, 570-597, 1993 (with R. Blundell and G. Weber).

Pashardes, Panos. “Non-Linearities and Equivalence Scales”, *The Economic Journal*, 103, 359-368, 1993 (with R. Dickens and V. Fry).

Spanos, Aris " Revisiting Date Mining: ´hunting´ with or without a license”, forthcoming *Journal of Methodology,* July 2000.

Spanos, Aris. "On Modeling Heteroscedasticity: The Student's *t* and Elliptical Linear Regression Models", *Econometric Theory*, 10, 286-315, 1994.

Spanos, Aris. "On Normality and the Linear Regression Model", *Econometric Reviews,* 14, 195-203, 1995.

Spanos, Aris. "On Theory Testing in Econometrics: Modeling with nonexperimental Data", *Journal of Econometrics,* 67, 189-226, 1995.