ISSN 1440-771X

MONASH University

Australia

Department of Econometrics and Business Statistics

http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/

Semi-parametric Analysis of Shape-Invariant Engel Curves with Control Function Approach

Nam H. Kim, Patrick W. Saart and Jiti Gao

March 2013

Working Paper 10/13

Semi-parametric Analysis of Shape-Invariant Engel Curves with Control Function Approach

Nam. H. Kim, Patrick. W. Saart[1] and Jiti Gao[2]

University of Adelaide, University of Canterbury and Monash University

Abstract

An extended generalised partially linear single-index (EGPLSI) model provides flexibility of a partially linear model and a single-index model. Furthermore, it also allows for the analysis of the shape-invariant specification. Nonetheless, the model's practicality in the empirical studies has been hampered by lack of appropriate estimation procedure and method to deal with endogeneity. In the current paper, we establish an alternative control function approach to address the endogeneity issue in the estimation of the EGPLSI model. We also show that all attractive features of the EGPLSI model discussed in the literature are still available under the proposed estimation procedure. Economic literature suggests that semiparametric technique is an important tool for an empirical analysis of Engel curves, which often involves endogeneity of the total expenditure. We show that our newly developed method is applicable and able to address the endogeneity issue involved in semiparametric analysis of the empirical Engel curves.

JEL Classification: C14, C41, F31.

Keyword: Control function approach, endogeneity, generalised partially linear single-index, semiparametric regression.

[1]Patrick W. Saart, Department of Mathematics and Statistics, University of Canterbury, New Zealand, Email: patrick.wsaart@canterbury.ac.nz; website: https://sites.google.com/site/patrickwsaart/home

[2]The first and second authors would like to thank Dominic Lee for his suggestions and comments especially on the empirical analysis of this paper. The third author acknowledges financial support from the Australian Research Council Discovery Grants Program under Grant number: DP1096374.

1. Introduction

Since its introduction in the study by Carroll et al. (1997), the Generalized Partially Linear Single-Index (GPLSI) model has received constant attention and been studied by many researchers; see Yatchew (2003) and Gao (2007), for example. Furthermore, Xia et al. (1999) provide a useful extension to the model; in this paper, let us refer to it as the extended GPLSI (EGPLSI) model. The EGPLSI model allows for the well-known advantages of a Single-Index (SI) model and a Partially Linear (PL) model (see the discussion in Chapter 2 of Horowitz (2009) for details) and also enables the analysis of the so-called shape-invariant specification as will be illustrated in Section 3. Unlike its GPLSI counterpart, the EGPLSI model concedes instead a more extensible specification, which includes the shape-invariant one as a special case.

Recently, considerable effort has been made in studies of the shape-invariant specification in the literature. While some interesting theoretical studies can be found in Härdle and Marron (1990), and Pinkse and Robinson (1995), the best known application is in the empirical demand study literature such as Blundell et al. (1998), Blundell et al. (2003) and Blundell et al. (2007). In the context of empirical demand studies, this specification enables the analysis of both a scale coefficient and a shift coefficient of a household characteristic in the modelling specification, which is coherent with the consumer theory; see Blundell et al. (1998), Pendakur (1999), Blundell et al. (2003) and Blundell et al. (2007) for detail.

With regard to nonparametric estimation techniques employed, the study by Carroll et al. (1997) propose the local constant kernel estimation method, while Xia and Härdle (2006) consider the local polynomial estimation method of Fan and Gijbels (1996) to estimate the GPLSI model. On the other hand, Xia et al. (1999) employ the local constant kernel estimation method to estimate the EGPLSI model and to examine its identification condition. However, these methods are not directly applicable to empirical studies in various economic areas, since they do not take endogeneity into account. The so-called "endogeneity problem" is a technical name given by econometricians to a problem that is well known in developmental studies and empirical economics; see Nakamura and Nakamura (1998), and Deaton and Muellbauer (1980) for some excellent surveys. For example, the endogeneity of total expenditure is a well-known issue in the empirical demand study literature; see Blundell et al. (1998) and Blundell et al. (2007) for detail. If present, it might cause an inconsistent estimation of the model's scale coefficient and lead to nonidentification of structural Engel curves. Recently, various methods of addressing endogeneity in non and semiparametric models have been discussed in the literature. Among these, a couple of the most popular methods are the nonparametric instrumental variables (IV) estimation and the control function (CF) approaches; see Blundell and Powell (2003) for an excellent review of these methods.

In the current study, we intend to provide two main contributions to the econometric literature. Firstly, we aim to introduce a method to address endogeneity in the estimation of the above-mentioned EGPLSI model. In particular, we aim to do so by establishing a CF approach based on (i) the Robinson (1988) and Speckman (1988) type of the two-stage estimation procedure and (ii) the widely-used triangular structure of Newey et al. (1999), Pinkse (2000), Blundell and Powell (2004), and Su and Ullah (2008). The two-stage estimation procedure allows us to conveniently identify the source(s) of endogeneity and hence systematically address it in a partially linear type of semiparametric model via the partialling-out process. Furthermore, we present in detail below how imposition of the triangular structure enables us to identify the unknown structural relationship (e.g. the structural Engel curves) in a simple nonparametric additive structure which can be conveniently estimated using the marginal integration technique of Linton and Nielsen (1995), and Tjøstheim and Austad (1996). In spite of the involvement of an endogeneity control variable which is not observable in practice and hence is non-parametrically estimated for the flexibility (as in Newey et al. (1999), and Su and Ullah (2008)), we derive the asymptotic normality and the $\sqrt{n}$-consistency of parameter estimators of both the parametric coefficients and the index coefficients. More importantly, we show that the practicality of the study in Xia et al. (1999), which allows the same smoothing parameter in the estimation of the index coefficients and the unknown structural function, is still applicable to the EGPLSI model with the endogeneity control variable generated.

Secondly, we also intend to provide a further contribution to the economic literature, particularly on the cross sectional relationships between expenditure on specific goods and the level of total expenditure. To achieve this objective, we employ our newly established methods to conduct a semiparametric analysis of shape-invariant Engel curves in Australia. It should be noted that within the context of the empirical demand study, Blundell et al. (2007) address the endogeneity of the total expenditure by using the nonparametric IV method through which some regularity conditions are imposed on the inversion matrix and a constraint is placed on the space of the reduced relation to make it compact. Blundell et al. (2007) show the $\sqrt{n}$-consistency of the estimators of both the scale and the shift coefficients. On the other hand, Blundell et al. (1998) address endogeneity by using the CF approach by a parametrically generated endogeneity control variable. We will clearly explain the difference between our method and that of Blundell et al. (1998) below. Furthermore, because of the importance of this topic, even though an effective tool is lacking for testing endogeneity in semiparametrics, an additional advantage of our method is that it enables a rather simple procedure to be established for the purpose. This is brought about mainly by its ability to identify and unentangle the effect of endogeneity in the model. This simple tool relies on the variability bands being constructed over the estimates of the endogeneity measures (to be defined below) as the means of testing their statistical significance.

The structure of the rest of the paper is as follows. Section 2 immediately below discusses the first contribution in detail, i.e. introduction of an alternative method for addressing endogeneity in the estimation of the EGPLSI model. Section 3 concentrates on the second contribution, i.e. the empirical study of the cross sectional relationships between specific goods and the level of total expenditure. We conclude the paper with a summary of our results in Section 4. All mathematical proofs of the main theoretical results of the paper are presented in the Appendix.

2. EGPLSI Model with/without Endogeneity

Let us begin the current section with a brief review of the EGPLSI model and its estimation procedure as often discussed in the literature (see Xia et al. (1999) and Gao (2007), for example). We introduce endogeneity into the model and then discuss our alternative CF based estimation procedure in Section 2.2. We present the main theoretical results of this paper, which focus on the asymptotic properties of estimators of the model in Section 2.3. All mathematical proofs are discussed in the Appendix. Finally, the finite sample properties of the estimators are investigated in Section 2.4.

2.1. EGPLSI Model without Endogeneity

Generally, without the presence of endogeneity, the EGPLSI model can be defined as:

$$Y_i = X_i'\beta_0 + g(X_i'\alpha_0) + \epsilon_i, \tag{2.1}$$

where (X, Y) is a $\mathbb{R}^q \times \mathbb{R}$-valued observable random vector, β_0 and α_0 are unknown vector parameters, and $g(\cdot)$ is an unknown link function such that $g : \mathbb{R} \to \mathbb{R}$. The exogeneity assumption suggests that $E(\epsilon_i|X_i) = 0$, which implies that $E(\epsilon_i|V_{i0}) = 0$ for $V_{i0} = X_i'\alpha_0$. Throughout the rest of the paper, let us assume that the random sample $\{(X_i', Y_i); i = 1, \ldots, n\}$ is independently and identically distributed (i.i.d.). Furthermore, let $f(x)$ and $f(v_0)$ denote the density functions of x and v_0, respectively, with the random argument of X_i. We also assume that $\mathcal{A}_x \subseteq \mathbb{R}^q$ is the union of a finite number of open convex sets such that $f(x) > M_x$ on $\mathcal{A}_x$ for some constant $M_x > 0$. Finally, note the identification condition of the EGPLSI model investigated in Xia et al. (1999), the orthogonality of the two coefficients so that $\beta_0 \perp \alpha_0$ with $||\alpha_0|| = 1$.

Given α and β, we smooth the nonparametric index component out from the structural relation (2.1) to obtain the minimising objective function for both unknown coefficients as shown below:

$$\min_{\alpha,\beta} J^*(\alpha, \beta) = \min_{\alpha,\beta} E\left(W_i^* - U_i^{*\prime}\beta\right)^2, \tag{2.2}$$

where $W_i^* = Y_i - E^*(Y_i|V_i)$ and $U_i^* = X_i - E^*(X_i|V_i)$ with $V_i = X_i'\alpha$. In order to estimate those unknown parameters and functions involved in (2.1), we need to obtain a feasible version of (2.2). Firstly, consider the nonparametric kernel estimators of $E^*(Y_i|V_i)$ and $E^*(X_i|V_i)$ of the form:

$$\hat{E}^*(y|v) = \frac{\sum_{X_i \in \mathcal{A}_x} k_h(V_i - v)Y_i}{\sum_{X_i \in \mathcal{A}_x} k_h(V_i - v)} \quad \text{and} \quad \hat{E}^*(x|v) = \frac{\sum_{X_i \in \mathcal{A}_x} k_h(V_i - v)X_i}{\sum_{X_i \in \mathcal{A}_x} k_h(V_i - v)}, \tag{2.3}$$

where $k_h(\cdot) = k(\cdot/h)$, $k(\cdot)$ is a kernel function satisfying Assumption 2.4 below and h is a bandwidth parameter. Next, we turn to the corresponding estimators based on the usual cross-validation criterion. Let the estimators in (2.3) be the leave-one-out estimators by omitting (X_i, Y_i, V_i):

$$\hat{E}_i^*(y|v) = \frac{\sum_{j \neq i} k_h(V_j - v)Y_j}{\sum_{j \neq i} k_h(V_j - v)} \quad \text{and} \quad \hat{E}_i^*(x|v) = \frac{\sum_{j \neq i} k_h(V_j - v)X_j}{\sum_{j \neq i} k_h(V_j - v)}. \tag{2.4}$$

Let A_n denote the set of all unit q-vectors. Given $C > 0$ and $0 < C_1 < C_2 < \infty$, $A_n = \{\alpha \in A_n : ||\alpha - \alpha_0|| \leq Cn^{-1/2}\}$ and $\mathcal{H}_n = \left\{h : C_1 n^{-1/5} \leq h \leq C_2 n^{-1/5}\right\}$. These definitions are motivated by the fact that, since we anticipate that $\hat{\alpha}^*$ is $\sqrt{n}$-consistent and we expect $\hat{h}$ to be close to $h_0 \sim$ const $n^{1/5}$, we should look for a minimum of the feasible objective function of (2.2), i.e. $\hat{J}(\alpha, h)$, defined in Step 2.1.3 of Procedure 2.1 below. The feasible objective function involves α to be distant from α_0 by the order of $n^{-1/2}$ and h to be approximately equal to a constant multiple of $n^{-1/5}$; see Härdle et al. (1993) and Xia et al. (1999), for example. The estimation procedure of (2.1) can be summarised as follows. Hereafter, let us collectively refer to these estimation steps as *"Procedure 2.1"*.

Procedure 2.1

Step 2.1.1: Given α, obtain the feasible objective function of (2.2) by estimating $E^*(y|v)$ and $E^*(x|v)$ by $\hat{E}_i^*(y|v)$ and $\hat{E}_i^*(x|v)$ in (2.4).

Step 2.1.2: Define the feasible objective function of (2.2) as:

$$\hat{J}^*(\beta) = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{W}_i^* - \hat{U}_i^{*\prime}\beta\right)^2, \tag{2.5}$$

where $\hat{W}_i^* = Y_i - \hat{E}_i^*(Y_i|V_i)$ and $\hat{U}_i^* = X_i - \hat{E}_i^*(X_i|V_i)$. Perform the least squares (LS) estimation on (2.5) to obtain $\hat{\beta}^* = (S_{\hat{U}^*})^- S_{\hat{U}^*\hat{W}^*}$, where $S_{AB} = \frac{1}{n}\sum_{i=1}^{n} A_i B_i'$, $S_A = S_{AA}$, and $(S_{\hat{U}^*})^-$ is a generalised inverse of $(S_{\hat{U}^*})$.

Step 2.1.3: Given $\hat{\beta}^*$ from the previous step, obtain $\hat{\alpha}^*$ and $\hat{h}$ by minimising the feasible objective function:

$$\min_{\alpha\in A_n, h\in\mathcal{H}_n} \hat{J}^*(\alpha, h) = \min_{\alpha\in A_n, h\in\mathcal{H}_n} \frac{1}{n}\sum_{i=1}^{n}(\hat{W}_i^* - \hat{U}_i^{*\prime}\hat{\beta}^*)^2.$$

Step 2.1.4: Re-estimate β_0 using $\hat{\alpha}^*$ and $\hat{h}$ from Step 2.1.3 as in 2.1.2:

$$\hat{\beta}_{\hat{\alpha}}^* = \left(S_{\hat{U}_{\hat{\alpha}}^*}\right)^- S_{\hat{U}_{\hat{\alpha}}^*\hat{W}_{\hat{\alpha}}^*},$$

where $\hat{W}_{\hat{\alpha}^*,i} = Y_i - \hat{E}_i^*(Y_i|\hat{V}_i)$ and $\hat{U}_{\hat{\alpha}^*,i} = X_i - \hat{E}_i^*(X_i|\hat{V}_i)$ with $\hat{V}_i = X_i'\hat{\alpha}$, $\hat{E}_i^*(Y_i|\hat{V}_i)$ and $\hat{E}_i^*(X_i|\hat{V}_i)$ obtained by replacing α in (2.4) with $\hat{\alpha}^*$.

Step 2.1.5: Given $\hat{\alpha}^*$ and $\hat{\beta}_{\hat{\alpha}}^*$, estimate the unknown structural function $g(\cdot)$ by $\hat{g}^*(\hat{v}) = \hat{E}^*(y|\hat{v}) - \hat{E}^*(x|\hat{v})'\hat{\beta}_{\hat{\alpha}}^*$. ■

The benefits of Procedure 2.1 of Xia et al. (1999) relies on the Robinson (1988) and Speckman (1988) type of the two-stage estimation procedure and the direct extension of the study in Härdle et al. (1993) to the EGPLSI model. On the one hand, the former conveniently allows for the identification of the source(s) of endogeneity and hence a systematic way of addressing endogeneity in partially linear semiparametrics due to the *partialling-out* process as discussed above. On the other hand, the latter provides an empirical and practical way of estimating single-index semiparametrics. The study of Härdle et al. (1993) allows for the same bandwidth for the optimal estimation of $\hat{\alpha}^*$ and $\hat{g}^*(\cdot)$, and the simultaneous estimation of index coefficients and a smoothing parameter. Procedure 2.1 accommodates this practicality of Härdle et al. (1993) in the EGPLSI model. In the next section, we show that these benefits of Xia et al. (1999) can be extended to the proposed estimation procedure in the current paper to address endogeneity in the EGPLSI model.

2.2. EGPLSI Model with Endogeneity

Let us now introduce endogeneity into the EGPLSI model, (2.1). There are two potential sources of endogeneity, namely endogeneity in the parametric and the nonparametric components. Hereafter, let us refer to these as parametric-endogeneity and nonparametric-endogeneity, respectively. Clearly, these two types of endogeneity may also occur simultaneously. To simplify the argument, we assume that the parametric regressors belong to a subset of X, i.e. $X_1 \subseteq \mathbb{R}^{q_1}$ for $q_1 < q$, such that the regressors are exogenous with $E(\epsilon|x_1) = 0$. Nonparametric-endogeneity exists for the case where $E(\epsilon|x) \neq 0$, which implies that $E(\epsilon|v_0) \neq 0$. Unless the parametric regressors are endogenous, the LS estimation results in the consistent estimation of the parametric coefficients even with nonparametric-endogeneity in the model due to the partialling-out process in the two-stage estimation procedure of Robinson (1988) and Speckman (1988). Note also that, if present, parametric-endogeneity can be conveniently dealt with using the parametric IV estimation; see also the discussion in Chapter 16 of Li and Racine (2007), for example. Nonetheless, Procedure 2.1 does not take the above mentioned nonparametric-endogeneity into account and may therefore result in inconsistent estimators for the index coefficients and in nonidentification of the unknown structural function. The formal result is due to similar reasoning to that in the classical linear regression model; see also the discussion in Chapter 8 of Amemiya (1985) for details. Given β_0, reconsider the objective function of (2.2), particularly the following:

$$\begin{aligned}
J(\alpha) &= E(W_i^* - U_i^{*\prime}\beta_0)^2 \\
&= E\left[\{g(V_{0i}) - g(V_i)\} + \epsilon_i - E(\epsilon_i|V_i)\right]^2 \\
&= E\left\{g(V_{0i}) - g(V_i)\right\}^2 + E\left\{\epsilon_i - E(\epsilon_i|V_i)\right\}^2 + 2E\left[\{g(V_{0i}) - g(V_i)\}\left\{\epsilon_i - E(\epsilon_i|V_i)\right\}\right] \\
&\equiv A_{1,1,i} + A_{1,2,i} + A_{1,3,i}.
\end{aligned}$$

The feasible objective function in Step 2.1.3 of Procedure 2.1 does not converge to the function which provides consistent estimators of the index coefficients, since $A_{1,3,i}$ may not converge to 0 in probability, due to endogeneity, i.e. $E(\epsilon|x) \neq 0$; see Amemiya (1974), for example. When there is no endogeneity, the estimator of $A_{1,3,i}$ converges to 0 and the estimator of $A_{1,1,i}$ converges to the unique function providing the minimum value of the objective function with respect to the index coefficients in probability. Note that $A_{1,2,i}$ is not relevant to the index coefficients. Here more importantly, the unknown structural function is not identified. This is mainly because $E(\epsilon|x) \neq 0$, the conditional expectation of ϵ on any function of x is not 0. This leads to the conditional expectation relation $E^*(y|v) - E^*(x|v)'\beta_0 = g(v) + E(\epsilon|v)$, and $E(\epsilon|v) \neq 0$. Hence it is the case that $\hat{E}^*(y|\hat{v}) - \hat{E}^*(x|\hat{v})'\hat{\beta}^* = \hat{g}^*(\hat{v}) + \hat{E}(\epsilon|\hat{v}) \overset{p}{\nrightarrow} g(v_0)$, where $\overset{p}{\nrightarrow}$ denotes no convergence in probability.

In order to obtain consistent estimators of the index coefficients and to recover the unknown structural function when nonparametric-endogeneity is present, we propose in the current section an alternative estimation method which is based on the CF approach; see the discussions in Newey et al. (1999), Blundell and Powell (2004), and Su and Ullah (2008) for its application to the non and semiparametric models. Let Z_i denote a vector of valid instruments for X_i such that:

$$X_i = m_x(Z_i) + \eta_i, \tag{2.6}$$

where we assume the following conditions:

$$E(\eta_i|Z_i) = 0 \text{ and } E(\epsilon_i|X_i, \eta_i) = E(\epsilon_i|Z_i, \eta_i) = E(\epsilon_i|\eta_i) = \iota(\eta_i), \tag{2.7}$$

and Z is an $\mathbb{R}^{q_z}$-valued vector, $q_z \geq q_2$ with $q_2 \equiv q - q_1$, $m_x(z)$ is a vector of unknown real functions, $m_x \equiv (m_{x,l}(Z_i))'$, $\{(Z_i'); i = 1, \ldots, n\}$ is i.i.d. and $m_{x,l} : \mathbb{R}^{q_z} \to \mathbb{R}$ for $l = 1, \ldots, q_2$. Also, let $f(z)$ denote the density function of z with the random argument of Z_i. Assume that $\mathcal{A}_z \subseteq \mathbb{R}^{q_z}$ is the union of a finite number of open convex sets such that $f(z) > M_z$ on $\mathcal{A}_z$ for some constant $M_z > 0$. The conditional expectation of the disturbance term in the reduced relation of (2.6), i.e. (2.7), is the distributional exclusion restriction; see the discussion on page 658 of Blundell and Powell (2004), which leads to the following argument. Hereafter, let us define the following:

$$m_y(v_0, \eta) = E(Y_i|V_{0i} = v_0, \eta_i = \eta) \quad \text{and} \quad m_x(v_0, \eta) = E(X_i|V_{0i} = v_0, \eta_i = \eta), \tag{2.8}$$

by which:

$$Y_i = m_y(V_{0i}, \eta_i) + W_{0i} \quad \text{and} \quad X_i = m_x(V_{0i}, \eta_i) + U_{0i}, \tag{2.9}$$

where $E(W_{0i}|X_i, \eta_i) = 0$ and $E(U_{0i}|X_i, \eta_i) = 0$. We are now able to derive the conditional expectation relation which controls endogeneity by using (2.6) to (2.9):

$$m(v_0, \eta) \equiv m_y(v_0, \eta) - m_x(v_0, \eta)'\beta_0 = g(v_0) + \iota(\eta), \tag{2.10}$$

where $\iota(\eta) \neq 0$ is the endogeneity control function which controls the endogeneity in the structural relation.

By imposing the above mentioned distributional exclusion restriction (2.7), we have gained control over the endogeneity in the nonparametric regressors. As the results show, it provides the consistent estimators of the index coefficients and also a way to identify the unknown structural function. Given β_0, reconsider (2.2) so that we have:

$$\begin{aligned} J(\alpha) &= E\left(W_i - U_i'\beta_0\right)^2 \\ &= E\left[\{g(V_{0i}) - g(V_i)\} + \epsilon_i - \iota(\eta_i)\right]^2 \\ &\equiv E\left\{g(V_{0i}) - g(V_i)\right\}^2 + E(e_i)^2 - 2E\left[\{g(V_{0i}) - g(V_i)\} e_i\right] \\ &\equiv A_{2,1,i} + A_{2,2,i} + A_{2,3,i}, \end{aligned}$$

where $e_i \equiv \epsilon_i - \iota(\eta_i)$, $W_i = Y_i - E(Y_i|V_i, \eta_i)$ and $U_i = X_i - E(X_i|V_i, \eta_i)$. Note that the estimator of $A_{2,3,i}$ converges to 0 in probability, since $E(e_i|X_i, \eta_i) = 0$. Hence, the feasible objective function (2.17) defined in Step 2.2.3 of Procedure 2.2 below converges to the function which provides the local minimum value with respect to the index coefficients in probability; see Chapters 4 and 8 of Amemiya (1985) for details. Furthermore, we may now identify the unknown structural function using the marginal integration technique, since (2.10) is a simple nonparametric additive structure. The details for implementing technique are given in Step 2.2.5. of Procedure 2.2 below.

Given β and α, the minimising objective function is:

$$\min_{\alpha,\beta} J(\beta, \alpha) = \min_{\alpha,\beta} E\left(W_i - U_i'\beta\right)^2. \tag{2.11}$$

Furthermore, let:

$$\hat{E}(y|v, \eta) = \frac{\sum_{X_i \in \mathcal{A}_x, Z_i \in \mathcal{A}_z} L_{h_v,h_\eta}(V_i - v, \eta_i - \eta) Y_i}{\sum_{X_i \in \mathcal{A}_x, Z_i \in \mathcal{A}_z} L_{h_v,h_\eta}(V_i - v, \eta_i - \eta)}, \tag{2.12}$$

and:

$$\hat{E}(x|v, \eta) = \frac{\sum_{X_i \in \mathcal{A}_x, Z_i \in \mathcal{A}_z} L_{h_v,h_\eta}(V_i - v, \eta_i - \eta) X_i}{\sum_{X_i \in \mathcal{A}_x, Z_i \in \mathcal{A}_z} L_{h_v,h_\eta}(V_i - v, \eta_i - \eta)}, \tag{2.13}$$

where $L_{h_v,h_\eta}(\cdot)$ is the product kernel function constructed from the product of the univariate kernel functions of $k_{h_{\eta_1}}(\cdot) \times \cdots \times k_{h_{\eta_{q_2}}}(\cdot) \times k_{h_v}(\cdot)$, and h_v and h_{η_j} with $j = 1, \ldots, q_2$ are the relevant bandwidth parameters and are nonparametric kernel estimators of $E(y|v, \eta)$ and $E(x|v, \eta)$, respectively. Next, we turn to the corresponding leave-one-out estimators of (2.12) and (2.13) by omitting (X_i, Y_i, V_i, η_i):

$$\hat{E}_i(y|v, \eta) = \frac{\sum_{j \neq i} L_{h_v,h_\eta}(V_j - v, \eta_j - \eta) Y_j}{\sum_{j \neq i} L_{h_v,h_\eta}(V_j - v, \eta_j - \eta)} \tag{2.14}$$

and:

$$\hat{E}_i(x|v, \eta) = \frac{\sum_{j \neq i} L_{h_v,h_\eta}(V_j - v, \eta_j - \eta) X_j}{\sum_{j \neq i} L_{h_v,h_\eta}(V_j - v, \eta_j - \eta)}. \tag{2.15}$$

We redefine $\mathcal{H}_n$ in the previous section as $\mathcal{H}_n = \left\{h_v, h_\eta, h_z : C_1 n^{-1/5} \leq h_v, h_\eta, h_z \leq C_2 n^{-1/5}\right\}$. We propose the following estimation procedure. Hereafter, let us collectively refer to these estimation steps as *"Procedure 2.2"*.

Procedure 2.2

Step 2.2.0: Estimate the endogeneity control regressors from (2.6) as:

$$\hat{\eta}_i = X_i - \hat{m}_x(Z_i), \tag{2.16}$$

where $\hat{m}_x(z) = \frac{\sum_{Z_i \in \mathcal{A}_z} K_{h_z}(Z_i - z) X_i}{\sum_{Z_i \in \mathcal{A}_z} K_{h_z}(Z_i - z)}$, in which $K_{h_z}(\cdot)$ is the product kernel function constructed from the product of the univariate kernel functions of $k_{h_{z_1}}(\cdot) \times \cdots \times k_{h_{z_{q_z}}}(\cdot)$ and h_{z_j} with $j = 1, \ldots, q_z$ is the relevant bandwidth parameter. By omitting the pair (X_i, Z_i), the corresponding leave-one-out estimator is $\hat{m}_{x,i}(z) = \frac{\sum_{j \neq i} K_{h_z}(Z_j - z) X_j}{\sum_{j \neq i} K_{h_z}(Z_j - z)}$.

Step 2.2.1: Given α and the non-parametrically generated endogeneity control regressors $\hat{\eta}_i$, obtain the feasible objective function of (2.11) by the estimates of $\hat{E}_i(y|v, \hat{\eta})$ and $\hat{E}_i(x|v, \hat{\eta})$, which are the corresponding estimates of those in (2.14) and (2.15) obtained by replacing η_i with $\hat{\eta}_i$.

Step 2.2.2: Define the feasible objective function of (2.11) as given below:

$$\hat{J}(\beta) = \frac{1}{n} \sum_{i=1}^{n} \left(\hat{W}_{2i} - \hat{U}_{2i}' \beta\right)^2,$$

where $\hat{W}_{2i} = Y_i - \hat{E}_i(Y_i|V_i, \hat{\eta}_i)$ and $\hat{U}_{2i} = X_i - \hat{E}_i(X_i|V_i, \hat{\eta}_i)$. We may compute the LS estimate of the unknown parametric coefficients as:

$$\hat{\beta}_\alpha = \left(S_{\hat{U}_2}\right)^{-} S_{\hat{U}_2 \hat{W}_2}.$$

Step 2.2.3: Given $\hat{\beta}$ from the previous step, compute $\hat{\alpha}$, $\hat{h}_v$ and $\hat{h}_{\hat{\eta}}$ by minimising the feasible objective function as follows:

$$\min_{\alpha\in A_n, h_v, h_{\hat{\eta}}\in\mathcal{H}_n} \hat{J}(\alpha, h_v, h_{\hat{\eta}}) = \min_{\alpha\in A_n, h_v, h_{\hat{\eta}}\in\mathcal{H}_n} \frac{1}{n}\sum_{i=1}^{n}(\hat{W}_{2i} - \hat{U}_{2i}'\hat{\beta})^2. \tag{2.17}$$

Step 2.2.4: Re-estimate β_0 using $\hat{\alpha}$, $\hat{h}_v$ and $\hat{h}_{\hat{\eta}}$ from the previous step as follows:

$$\hat{\beta} = \left(S_{\hat{U}_3}\right)^{-} S_{\hat{U}_3\hat{W}_3},$$

where $\hat{W}_{3i} = Y_i - \hat{E}_i(Y_i|\hat{V}_i, \hat{\eta}_i)$ and $\hat{U}_{3i} = X_i - \hat{E}_i(X_i|\hat{V}_i, \hat{\eta}_i)$ with $\hat{V}_i = X_i'\hat{\alpha}$.

Step 2.2.5 below is mainly due to the involvement of the marginal integration technique in an attempt to identify the unknown structural relation in question.

Step 2.2.5: Perform the marginal integration technique of Linton and Nielsen (1995) or Tjøstheim and Austad (1996) to identify the unknown structural function. ■

In the following paragraphs, we discuss an application of the marginal integration technique in Step 2.2.5 of Procedure 2.2 in greater detail. Let us first recall from (2.10) that $m(v_0, \eta) = g(v_0) + \iota(\eta)$, which is clearly a nonparametric additive specification. Hence a standard identification condition as discussed extensively in the literature (see Gao et al. (2006) and Gao (2007), for example) assumes that $E(g(v_0)) = E(\iota(\eta)) = 0$. The implementation of the marginal integration technique identifies $g(\cdot)$ and $\iota(\cdot)$ up to some constant values as follows:

$$m(v_0) = \int m(v_0, \eta)dQ(\eta) = g(v_0) + c_1,$$

and:

$$m(\eta) = \int m(v_0, \eta)dQ(v_0) = \iota(\eta) + c_2,$$

where $c_1 = \int \iota(\eta)dQ(\eta)$, $c_2 = \int g(v_0)dQ(v_0)$ and Q is a probability measure with $\int dQ(\eta) = \int dQ(v_0) = 1$. Here, the estimate of the structural relation can therefore be obtained by the following sample version of the integration:

$$\hat{m}(v) = \frac{1}{n}\sum_{i=1}^{n} \hat{m}(v, \hat{\eta}_i), \tag{2.18}$$

and:

$$\hat{g}(v) = \hat{m}(v) - \hat{c}_1, \tag{2.19}$$

where $\hat{m}(v,\hat{\eta}_i) = \hat{E}(y|v,\hat{\eta}_i) - \hat{E}(x|v,\hat{\eta}_i)'\hat{\beta}_{\hat{\alpha}}$, and $\hat{c}_1 = \frac{1}{n}\sum_{i=1}^{n}\hat{m}(\hat{V}_i)$. Note that (2.18) is estimated by keeping $\hat{V}_i$ at v, while taking an average over the remaining regressors, $\hat{\eta}_i$. In (2.19), in order to ensure that the identification condition of a nonparametric additive model is satisfied, the constant value is estimated as $\hat{c}_1$.

An attractive feature of Procedure 2.2 is that the practicality of Xia et al. (1999), which provides a way of selecting the same smooth parameter(s) for optimal estimation of both α_0 and $g(\cdot)$ is still applicable, despite the regressors generated in order to control endogeneity in the model. The feasible objective function (2.17) can be expanded in the form of $\hat{J}(\alpha, h_v, h_{\hat{\eta}}) = \tilde{J}(\alpha) + T(h_v, h_\eta) + R_1(\alpha, h_v, h_\eta, h_z)$, where $\tilde{J}(\alpha)$ is an accurate approximation to $E(W_i - U_i'\beta_0)^2$ and does not depend on the smoothing parameters, $T(h_v, h_\eta)$ is the usual cross-validation criterion for choosing optimal bandwidths to estimate $m(x'\alpha_0, \eta)$ for known values of α_0 and true values of η, and R_1 is shown to be $o_p(n^{-1/2})$ in Theorem 2.1 below. Hence, minimising $\hat{J}(\alpha, h_v, h_{\hat{\eta}})$ simultaneously with respect to α, h_v and $h_{\hat{\eta}}$ is very much like separately minimising $\tilde{J}(\alpha)$ with respect to α and $T(h_v, h_\eta)$ with respect to h_v and h_η.

2.3. Asymptotic Properties

In this section, we present the main theoretical results of the current paper. First, we present the necessary conditions and then the main theoretical results in Theorems 2.1 and 2.2. Within the results of Theorem 2.1, the asymptotic properties of both estimators of parametric and index coefficients are presented in Corollary 2.1, particularly the fact that they are $\sqrt{n}$-consistent. The asymptotic properties of the estimator of the unknown structural function are presented in Theorem 2.2. The formal proofs of these results are presented in the Appendix.

We impose the following regularity conditions. Assume that $\mathcal{A} = \mathcal{A}_x \times \mathcal{A}_\eta \subseteq \mathcal{R}^{2q_2}$ and $\mathcal{A}_z \subseteq \mathcal{R}^{q_z}$ are the unions of a finite number of open convex sets, respectively. Given ε_x, ε_η and ε_z, let $\mathcal{A}_x^{\varepsilon_x}$, $\mathcal{A}_\eta^{\varepsilon_\eta}$ and $\mathcal{A}_z^{\varepsilon_z}$ denote the sets of all points in $\mathcal{R}^{q_2}$ and $\mathcal{R}^{q_z}$ that are no more distant than ε_x, ε_η and ε_z, respectively. Put $\mathcal{U} = \left\{(v_0 = x'\alpha_0, \eta) : x \in \mathcal{A}_x^{\varepsilon} \text{ and } \eta \in \mathcal{A}_\eta^{\varepsilon}\right\}$, where ε is the smaller value of ε_x and ε_η, and $\mathcal{U}_z = \{z : z \in \mathcal{A}_z^{\varepsilon_z}\}$. Let $f(v_0, \eta)$ denote the joint density function of $(x'\alpha_0, \eta)$ with random arguments of X_i and η_i. Assume that for some ε and ε_z, we have the assumptions below.

Assumption 2.1. *$f(x, \eta)$ and $f(z)$ are bounded away from 0 on $\mathcal{U}$ and $\mathcal{U}_z$, respectively.*

Assumption 2.2. *$f(z)$ and $m_x(z)$ have bounded and continuous second derivatives on $\mathcal{U}_z$.*

Assumption 2.3. *$m(v,\eta)$, $m_y(v,\eta)$, $m_x(v,\eta)$ and $f(v,\eta)$ have bounded and continuous second derivatives on $\mathcal{U}$ for all values of $\alpha \in A_n$.*

Assumption 2.4. *A univariate kernel function $k(\cdot)$ and its first derivative $k^{(1)}(\cdot)$ are supported on the interval $(-1,1)$ and $k(\cdot)$ is a symmetric probability density with $k^{(1)}(\cdot)$ being bounded.*

Assumption 2.5. *Let $E(e_i|X_i = x, \eta_i = \eta) = 0$ and $E(U_i|X_i = x, \eta_i = \eta) = 0$. Assume that $E(e_i^2|X_i = x, \eta_i = \eta) = \sigma^2(x,\eta)$ and $E(u_i^2|X_i = x, \eta_i = \eta) = \boldsymbol{u}^2(x,\eta)$ hold almost surely and both are continuous in (x,η). Let also $\sup_i E|Y_i|^l < \infty$ and $\sup_i E||X_i||^l < \infty$ for some $l > 2$.* ■

Assumption 2.1 is imposed to permit estimation of the functions in the regions of $\mathcal{A}^{\varepsilon}$ and $\mathcal{A}_z^{\varepsilon_z}$ in order to avoid the random denominator problem. A similar set of conditions is imposed in Härdle et al. (1993) and Xia et al. (1999). Assumptions 2.2 and 2.3 are needed to ensure that the symmetric kernel function in Assumption 2.4 leads to a second-order bias in kernel smoothing. A higher-order bias can be achieved by imposing more restrictive conditions on the smoothness of functions. For instance, Robinson (1988) reduces the bias sufficiently by employing a higher-order kernel function with strong smoothness conditions on the functions. The condition of the first derivative of the kernel function in Assumption 2.4 is required because we employ the Taylor argument to address the generated regressors, $\hat{\eta}$. A similar condition on the rth derivative of the kernel function can be found in Hansen (2008). Assumption 2.5 is imposed so that the Chebyshev inequality can be applied as in Härdle et al. (1993) and Xia et al. (1999).

Let us define the following:

$$\begin{aligned} B_v(v_0,\eta) &= \frac{K_{v,2}}{f(v_0,\eta)} \left\{ f_v^{(1)}(v_0,\eta) m_0^{(1)}(v_0) + f(v_0,\eta) m_0^{(2)}(v_0) \right\} \\ B_{\eta,j}(v_0,\eta) &= \frac{K_{\eta,2}}{f(v_0,\eta)} \left\{ f_{\eta,j}^{(1)}(v_0,\eta) m_j^{(1)}(\eta) + f(v_0,\eta) m_j^{(2)}(\eta) \right\}, \end{aligned}$$

where $\mathcal{K}_{v,2} = \int v_0^2 k_{h_v}(v_0) dv_0$, $\mathcal{K}_{\eta,2} = \int \eta^2 K_{h_\eta}(\eta) d\eta$ with $K_{h_\eta} = k_{h_{\eta_1}}(\cdot) \times \cdots \times k_{h_{\eta_{q_2}}}(\cdot)$, $f_v^{(r)}$ and $f_\eta^{(r)}$ are rth derivatives of the joint density function of $f(v_0,\eta)$ with respect to v_0 and η, respectively, and $m_0^{(r)}(v_0)$ and $m_j^{(r)}(\eta)$ are the rth partial derivatives of the function $m(v_0,\eta)$ with respect to v_0 and η_j, respectively, where $j = 1,\ldots,q_2$. Also, let $\mathcal{K} = \mathcal{K}_v \mathcal{K}_\eta^{q_2}$, where $\mathcal{K}_v = \int k_{h_v}(v_0)^2 dv_0$ and $\mathcal{K}_\eta = \int k_{\eta,j}(\eta)^2 d\eta$. In these notations, the "integrated mean squared error (IMSE)" is:

$$IMSE(h_v,h_\eta) \asymp \int \left\{ \left[B_v(v_0,\eta)h_v^2 + \sum_{j=1}^{q_2} B_\eta(v_0,\eta)h_{\eta,j}^2 \right]^2 + \frac{\mathcal{K}}{nh_v h_{\eta,1}\ldots h_{\eta,q_2}} \frac{\sigma^2(v_0,\eta)}{f(v_0,\eta)} \right\} f(x,\eta)dxd\eta,$$

where $\asymp$ means that the quotient of the two sides tends to 1 and $n \to \infty$. Now let us define the following:

$$\tilde{J}(\alpha) = \frac{1}{n}\sum_{i=1}^{n} \left\{ W_i - U_i'\hat{\beta} \right\}^2 \quad \text{and} \quad T(h_v,h_\eta) = \frac{1}{n}\sum_{i=1}^{n} \left\{ \hat{m}_i(V_{0i},\eta_i) - m(V_{0i},\eta_i) \right\}^2,$$

where $\hat{m}_i(\dot{})$ is the leave-one-out kernel estimator of $m(\cdot)$. Hence, we have the result shown in Theorem 2.1.

Theorem 2.1. *Under Assumptions 2.1 to 2.5, we can write:*

$$\hat{J}(\alpha,h_v,h_{\hat{\eta}}) = \tilde{J}(\alpha) + T(h_v,h_\eta) + R_1(\alpha,h_v,h_\eta,h_z) + R_2(\alpha,h_v,h_\eta), \tag{2.20}$$

$$T(h_v,h_\eta) = IMSE(h_v,h_\eta) + R_3(h_v,h_\eta), \tag{2.21}$$

where $R_3(h_v,h_\eta)$ does not depend on α, and:

$$\sup_{\alpha\in A_n,h_v,h_\eta,h_z\in\mathcal{H}_n} |R_1(\alpha,h_v,h_\eta,h_z)| = o_p(1), \quad \sup_{\alpha\in A_n,h_v,h_\eta\in\mathcal{H}_n} |R_2(\alpha,h_v,h_\eta)| = o_p(1),$$

and:

$$\sup_{h_v,h_\eta\in\mathcal{H}_n} |R_3(h_v,h_\eta)| = o_p(1).$$

■

The above theorem is a direct extension of the work of Xia et al. (1999) to a more complicated model associated with endogeneity. Now, let us define the following:

$$\Phi_{U_0} = X_i - E(X_i|V_{0i},\eta_i) \quad \text{and} \quad m_0^{(1)} = \partial m(v_0,\eta)/\partial v_0.$$

As the results of Theorem 2.1 show, we have the asymptotic results for the estimators of α_0 and β_0 shown in Corollary 2.1 below.

Corollary 2.1. *Under the assumptions of Theorem 2.1, we obtain the following:*

$$\sqrt{n}(\hat{\beta}-\beta_0) \to_D (0,var_1),$$

where $var_1 = \sigma^2 \left[\Phi_{U_0}^- - \left(m_0^{(1)}\Phi_{U_0}\right)^- \Phi_{U_0} \left\{m_0^{(1)}\right\}^2 \left(m_0^{(1)}\Phi_{U_0}\right)^- \right]$ *and:*

$$\sqrt{n}(\hat{\alpha}-\alpha_0) \to_D (0,var_2),$$

where $var_2 = \sigma^2 \left[\left\{(m_0^{(1)})^2\Phi_{U_0}\right\}^- - \left\{m_0^{(1)}\Phi_{U_0}\right\}^- \Phi_{U_0} \left\{m_0^{(1)}\Phi_{U_0}\right\}^- \right].$ ■

As for the estimator of the unknown structural function, i.e. $g(\cdot)$, we have the asymptotic properties shown in Theorem 2.2.

Theorem 2.2. *Under Assumptions 2.1 to 2.5, we show that:*

$$\sqrt{nh_v}\left(\hat{g}(\hat{v}) - g(v_0) - bias\right) \rightarrow_D N(0, var),$$

where $bias = h_v^2 \mathcal{B}_v(v_0, \eta) + \sum_{s=1}^{q_2} h_{\eta,s}^2 B_{\eta,s}(v_0, \eta)$ *and* $var = f(v_0)\mathcal{K}_v \int \frac{\sigma^2(v_0,\eta) f^2(\eta)}{f^2(v_0,\eta)} dQ(\eta)$. ■

The proofs of Theorems 2.1 and 2.2 as well as Corollary 2.1 are given in the Appendix below.

2.4. Simulation Studies

The purposes of the simulation exercises conducted in this section are twofold. Firstly, the section aims to investigate whether experimental evidence can be found to support the various points made in the theoretical discussion presented in the previous sections. Secondly, we aim to provide finite sample evidence for the usefulness of the newly introduced method for addressing endogeneity in the estimation of semiparametric SI models.

Remark 2.1. *The work to be presented in this section has been completed in two stages. Initially, we conduct our simulation study based on the strategy discussed in Section 2.4.1 below. The results obtained are mostly the same as what we expected, i.e. Procedure 2.2 performs superbly in the presence of nonparametric-endogeneity. On the other hand, Procedure 2.1, which was developed without an effective mechanism to deal with endogeneity, does not seem to be able to identify the unknown structural function for the models under investigation. Although this evidence alone should be more than sufficient to dismiss the use of Procedure 2.1 in the presence of endogeneity, it is surprising to see that, with a couple of exceptions, such a procedure still performs quite well overall in the estimation of the index coefficients. In order to provide further clarity, we conduct further investigations on the importance of some particular characteristics of endogeneity on the estimation outcomes. This is the work conducted in Section 2.4.2.*

2.4.1. Initial Investigation

In this section, we consider two illustrative models, namely the GPLSI-type and the EGPLSI-type, as defined in Examples 2.4.1 and 2.4.2 below. In practice, endogeneity is introduced to the models first and then Procedure 2.2 is applied. The finite sample performance of Procedure 2.2 is subsequently compared to that of Procedure 2.1.

Example 2.4.1: *GPLSI-type* The baseline model without endogeneity is:

$$Y_i = 1.2X_{1i} + g(V_{0i}) + \epsilon_i \text{ with } V_{0i} = \alpha_0 X_{2i} = \frac{1}{\sqrt{2}}(X_{2i}), \tag{2.22}$$

such that:

$$g(V_{0i}) = \frac{1}{2}\left\{\frac{\frac{1}{\sqrt{2}}(X_{2i})}{1 + \left[\frac{1}{\sqrt{2}}(X_{2i})\right]^2}\right\},$$

where X_1 and X_2 are independently and uniformly distributed on $[-1, 1]$ and $\epsilon_i \sim N(0, 1)$. Clearly, (2.22) is a GPLSI type of model such that the perpendicularity of the parameter vectors (see Xia et al. (1999), for instance) is not required. In this example, we introduce endogeneity into the nonparametric regressor by letting $X_{2i} = Z_i + \eta_i$, where Z and η are independently and uniformly distributed on $[-0.5, 0.5]$ and $[-1, 1]$, respectively, and $\epsilon_i = \eta_i + e_i$ and $e_i \sim N(0, 1)$. ■

Example 2.4.2: *EGPLSI-type* The base line model without endogeneity is:

$$Y_i = 0.3X_{1i} + 0.4X_{2i} + g(V_{0i}) + 0.1\epsilon_i \text{ with } V_{0i} = 0.8X_{1i} - 0.6X_{2i} + 0.5X_{3i}, \tag{2.23}$$

such that:

$$g(V_{0i}) = \exp\left\{-2(0.8X_{1i} - 0.6X_{2i} + 0.5X_{3i})^2\right\},$$

where for $j = 1, 2, 3$, X_j is independently and uniformly distributed on $[-1, 1]$ and $\epsilon_i \sim N(0, 1)$. Model (2.23) is an EGPLSI type of model such that the required perpendicularity of the parameter vectors is satisfied, given that their dot product is zero. In this example, we introduce endogeneity into the nonparametric regressor by letting $X_{3i} = Z_i + \eta_i$, where Z and η are independently and uniformly distributed on $[-0.5, 0.5]$ and $[-1, 1]$, respectively, and $\epsilon_i = \eta_i + e_i$ and $e_i \sim N(0, 1)$. ■

Throughout this section, optimisation is implemented using a limited memory Broyden–Fletcher–Goldfarb–Shanno algorithm for the bound constrained optimisation of Byrd et al. (1995). All simulation exercises are conducted in R with the Gaussian kernel function and the number of replications $Q = 200$. To compare and evaluate the finite sample performances of the estimation procedures introduced above, we compute the mean and mean absolute errors of the estimates of both coefficients across Q replications as tabulated in Tables 2.1

to 2.4. We also compare the averaged absolute error (ae) of the estimates the unknown structural function which is computed for Procedure 2.1 and for Procedure 2.2 using the following:

$$\text{ae}_{\hat{g}} = \frac{1}{n} \sum_{i=1}^{n} \left| \hat{g}(\hat{V}_i) - g(V_{0i}) \right|,$$

where n is the number of samples.

Table 2.1. *GPLSI-type model with nonparametric endogeneity: Procedure 2.1.*

n	$\hat{\beta}$	$\hat{\alpha}$	$\lvert\hat{\beta} - 1.2\rvert$	$\lvert\hat{\alpha} - 1/\sqrt{2}\rvert$	$\text{ae}_{\hat{g}}$
50	1.1997	0.8980	0.0060	0.1980	0.0438
150	1.1994	0.8592	0.0031	0.1592	0.0443
300	1.1999	0.7306	0.0024	0.0402	0.0443
500	1.2001	0.6523	0.0016	0.0708	0.0446

Table 2.2. *GPLSI-type model with nonparametric endogeneity: Procedure 2.2.*

n	$\hat{\beta}$	$\hat{\alpha}$	$\lvert\hat{\beta} - 1.2\rvert$	$\lvert\hat{\alpha} - 1/\sqrt{2}\rvert$	$\text{ae}_{\hat{g}}$
50	1.2000	0.8272	0.0033	0.1436	0.0266
150	1.1999	0.7784	0.0015	0.0796	0.0176
300	1.2000	0.7527	0.0082	0.0578	0.0148
500	1.9999	0.7502	0.0006	0.0580	0.0118

Table 2.3. *EGPLSI-type model with nonparametric endogeneity: Procedure 2.1.*

n	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$
50	0.2665	0.4236	0.9606	-0.7113	0.6586
150	0.2632	0.4383	0.8856	-0.6527	0.5910
300	0.2673	0.4340	0.8171	-0.6037	0.5422
500	0.2649	0.4355	0.7376	-0.5453	0.4880

n	$\lvert\hat{\beta}_1 - 0.3\rvert$	$\lvert\hat{\beta}_2 - 0.4\rvert$	$\lvert\hat{\alpha}_1 - 0.8\rvert$	$\lvert\hat{\alpha}_2 - (-0.6)\rvert$	$\lvert\hat{\alpha}_3 - 0.5\rvert$	$\text{ae}_{\hat{g}}$
50	0.0679	0.0651	0.1691	0.1253	0.1586	0.0838
150	0.0461	0.0489	0.0859	0.0559	0.0910	0.0802
300	0.0364	0.0382	0.0229	0.0156	0.0426	0.0800
500	0.0361	0.0368	0.0629	0.0548	0.0181	0.0799

Table 2.4. *EGPLSI-type model with nonparametric endogeneity: Procedure 2.2.*

n	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$
50	0.2253	0.4558	0.9649	-0.7366	0.6169
150	0.2877	0.4100	0.9226	-0.6951	0.5755
300	0.3118	0.3910	075821	-0.5670	04738
500	0.3089	0.3930	0.8068	-0.6065	0.5026

n	$\lvert\hat{\beta}_1 - 0.3\rvert$	$\lvert\hat{\beta}_2 - 0.4\rvert$	$\lvert\hat{\alpha}_1 - 0.8\rvert$	$\lvert\hat{\alpha}_2 - (-0.6)\rvert$	$\lvert\hat{\alpha}_3 - 0.5\rvert$	$\mathrm{ae}_{\hat{g}}$
50	0.0785	0.0587	0.1678	0.1389	0.1195	0.0618
150	0.0247	0.0186	0.1244	0.0962	0.0769	0.0240
300	0.0184	0.0138	0.0446	0.0327	0.0285	0.0146
500	0.0182	0.0137	0.0416	0.0319	0.0263	0.0124

Let us now present some important findings based on the results in Tables 2.1 to 2.4. Since endogeneity is introduced to the nonparametric regressor only, we expect the LS estimators of the unknown parameters in the parametric component to be consistent in all cases. Strong experimental evidence of such consistency can be clearly seen in all of the tables; see the fourth column of Tables 2.1 and 2.2, and the eighth to tenth columns of Tables 2.3 and 2.4 in particular. With the exception of some unprecedented (but not unexpected) increases in the absolute errors in the fourth column of Table 2.1 and the fifth column of Table 2.3, similar findings to that in the previous point may also be seen for the index coefficients. Despite the slightly better than expected performance, this is still strong evidence against the use of Procedure 2.1 when endogeneity is a possibility. Strong evidence against the use of Procedure 2.1 is clearly seen when the averaged absolute errors in the last columns of each table are considered. Unlike Procedure 2.2, Procedure 2.1 is clearly not able to identify the unknown structural function when endogeneity is presents.

In our view, such a conclusion should provide sufficient motivation for use of our newly established procedure in practice. However, in the next section, let us conduct a further investigation which provides more concrete evidence of the desirability of Procedure 2.2.

2.4.2. More Detailed Analysis

For the sake of clarity in illustrating the importance of some particular characteristics of endogeneity, the model used in the analysis that follows will be structurally similar to that of

Example 2.4.1. However, some modifications will be made to ensure that the experimental design is suitable to the objectives of the exercise. In this section, we will conduct two types of analysis, which are referred to hereafter as *Type A* and *Type B*, respectively.

Type A: The objective of the experimental analysis that follows is to study the importance of the conditional expectation of ϵ given η, i.e. denoted previously as $\iota(\cdot)$, for the performance of Procedure 2.1, which was originally introduced in Xia et al. (1999), in the presence of endogeneity. In such an experiment, the magnitude of endogeneity is clearly an important parameter that must be carefully controlled. In this current analysis, in order to best illustrate the impact of endogeneity, let us consider an extreme case, i.e. by defining:

$$X_{2i} = \eta_i, \tag{2.24}$$

where η_i is independently and uniformly distributed on $[-1, 1]$. Defining X_{2i} as in (2.24) enables specification of three related types of models, namely "exogeneity", "linear-endogeneity" and "nonlinear-endogeneity". In the current sections, these models can be respectively obtained by introducing the following:

$$\iota 1(\eta) = 0 \times \eta, \tag{2.25}$$

$$\iota 2(\eta) = 0.5 \times \eta, \tag{2.26}$$

$$\iota 3(\eta) = \frac{\eta}{\frac{1}{2}\left(4+\eta^2\right)}. \tag{2.27}$$

For example, (2.25) suggests that the conditional expectation of ϵ given η is zero and the model is exogenous. An example of $g(\cdot)$, $\iota 1(\cdot)$, $\iota 2(\cdot)$ and $\iota 3(\cdot)$ with $n = 500$ is presented in Figure 2.1. The simulation results in this section are presented in Tables 2.5 to 2.7.

Table 2.5. *Nonparametric-exogeneity, i.e.* $\iota 1$.

n	$\hat{\beta}$	Bias	Var	$\lvert\hat{\beta}-\beta\rvert$	$\hat{\alpha}$	Bias	Var	$\lvert\hat{\alpha}-\alpha\rvert$	$ae_{\hat{g}}$
100	1.1997	0.0002	0.0003	0.0143	0.9660	0.2660	0.0008	0.2660	0.0150
300	1.1996	0.0004	0.0001	0.0079	0.7989	0.0989	0.0012	0.0989	0.0108
500	1.2001	0.0001	0.0000	0.0055	0.7740	0.0740	0.0056	0.0740	0.0084
700	1.2005	0.0005	0.0000	0.0055	0.7330	0.0330	0.0045	0.0332	0.0073

Table 2.6. *Linear-endogeneity, i.e. $\iota 2$.*

n	$Corr_L$	$\hat{\beta}$	Bias	Var	$\lvert\hat{\beta}-\beta\rvert$	$\hat{\alpha}$	Bias	Var	$\lvert\hat{\alpha}-\alpha\rvert$	$ae_{\hat{g}}$
100	0.9852	1.1998	0.0001	0.0003	0.0145	0.9910	0.2910	0.0001	0.2910	0.2474
300	0.9852	1.1994	0.0005	0.0001	0.0079	0.8039	0.1039	0.0049	0.1039	0.2492
500	0.9853	1.2000	0.0000	0.0000	0.0057	0.8092	0.1093	0.0128	0.1092	0.2496
700	0.9853	1.2001	0.0005	0.0000	0.0056	0.7721	0.0721	0.0124	0.0898	0.2491
900	0.9853	1.1997	0.0002	0.0000	0.0043	0.8072	0.1072	0.0199	0.1341	0.2492
1,100	0.9853	1.2003	0.0003	0.0000	0.0040	0.7595	0.0595	0.0115	0.0932	0.2494
1,300	0.9853	1.1995	0.0004	0.0000	0.0035	0.7591	0.0591	0.0133	0.0982	0.2495

Table 2.7. *Nonlinear-endogeneity, i.e. $\iota 3$.*

n	$Corr_{NL}$	$\hat{\beta}$	Bias	Var	$\lvert\hat{\beta}-\beta\rvert$	$\hat{\alpha}$	Bias	Var	$\lvert\hat{\alpha}-\alpha\rvert$	$ae_{\hat{g}}$
100	0.9514	1.1998	0.0001	0.0003	0.0146	0.9852	0.2852	0.0001	0.2852	0.3748
300	0.9505	1.1995	0.0004	0.0001	0.0079	0.8573	0.1573	0.0079	0.1573	0.3771
500	0.9513	1.2000	0.0000	0.0000	0.0057	0.8882	0.1882	0.0099	0.1883	0.3777
700	0.9514	1.2005	0.0005	0.0000	0.0056	0.8592	0.1592	0.0099	0.1602	0.3771

Below, let us discuss some important findings. Note firstly that $E[\epsilon] = 0$, which implies that $E[\epsilon|\eta] = E[\epsilon] = 0$ when η and ϵ are independent. Therefore, in this case, we are able to measure the magnitude of endogeneity by simply considering the dependency between ϵ and η. The second columns of Tables 2.6 and 2.7, present averages over $Q = 200$ replications of the empirical correlation coefficients, which is a measure the linear dependence between ϵ and η. It is clear that even in such a controlled case, the functional forms of $\iota(\cdot)$ give rise to different magnitudes of endogeneity, which are measured by $Corr_L$ and $Corr_{NL}$. Since endogeneity is introduced to the nonparametric regressor only, the LS estimators of the unknown parameters in the parametric component seem to be consistent in all cases, as expected. Compared to the simulation results in Table 2.5, those in Tables 2.6 and 2.7 show clearly that Procedure 2.1 does not work well in the presence of endogeneity. Under linear-endogeneity, the procedure seems to work quite well in estimating the index coefficient up to about 700 observations. By extending the number of observations to 900, 1,100 and 1,300, it becomes clear that $|\hat{\alpha} - \alpha|$ shows no sign of converging to zero. Furthermore, the evidence suggests that the procedure is incapable of identifying the unknown structural function when (either linear or nonlinear) endogeneity is present. Overall, nonlinear-endogeneity seems to have somewhat more severe consequences when compared to its linear counterpart.

Figure 2.1. $g(\cdot)$, $\iota 1(\cdot)$, $\iota 2(\cdot)$ *and* $\iota 3(\cdot)$.

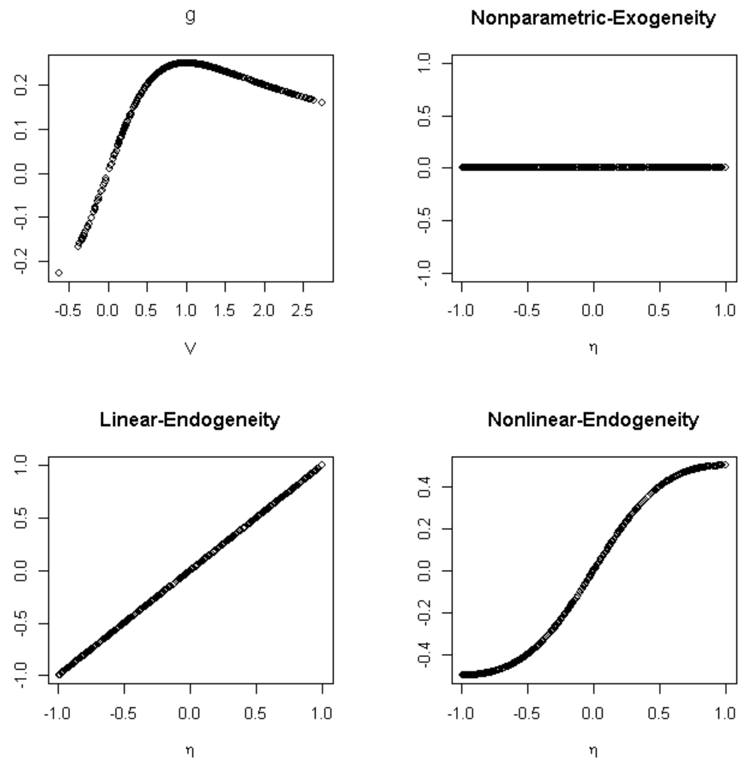


Table 2.8. $Corr_{X_{2i},Z_i}$

n	100	300	500	700
$Corr_{X_{2i},Z_i}$	0.8278	0.8302	0.8326	0.8330

Type B: The objective of the analysis that follows is to investigate the finite-sample performance of our newly introduced Procedure 2.2 in the presence of endogeneity. In practice, whether Z_i is a weak or a strong instrument may significantly affect the estimation outcomes. In order to control for such an effect, let us define the following:

$$X_{2i} = Z_i + \eta_i, \tag{2.28}$$

where Z and η are independently and uniformly distributed on $[0, 3]$ and $[-1, 1]$, respectively. Furthermore, we consider two cases of $\iota(\cdot)$, namely linear-endogeneity and nonlinear-endogeneity defined respectively as

$$\iota_2(\eta) = 1 \times \eta \quad \text{and} \quad \iota_3(\eta) = \frac{\eta}{1 + \eta^2}. \tag{2.29}$$

While Table 2.8 presents the averaged correlation coefficient of X_{2i} and Z_i at $Q = 200$ replications for $n = 100, 300, 500$ and 700, Tables 2.9 and 2.10 provide simulation results.

Table 2.9. *Linear-endogeneity, i.e. $\iota 2$.*

n	$Corr_L$	$\hat{\beta}$	Bias	Var	$\lvert\hat{\beta}-\beta\rvert$	$\hat{\alpha}$	Bias	Var	$\lvert\hat{\alpha}-\alpha\rvert$	$ae_{\hat{g}}$
100	0.9852	1.1972	0.0027	0.0008	0.0226	0.7803	0.0803	0.0076	0.0999	0.0827
300	0.9852	1.2009	0.0009	0.0001	0.0099	0.7372	0.0372	0.0009	0.0406	0.0511
500	0.9854	1.2003	0.0003	0.0001	0.0085	0.7137	0.0137	0.0005	0.0212	0.0439
700	0.9853	1.2008	0.0008	0.0000	0.0054	0.6948	0.0051	0.0002	0.0135	0.0385

Table 2.10. *Nonlinear-endogeneity, i.e. $\iota 3$.*

n	$Corr_{NL}$	$\hat{\beta}$	Bias	Var	$\lvert\hat{\beta}-\beta\rvert$	$\hat{\alpha}$	Bias	Var	$\lvert\hat{\alpha}-\alpha\rvert$	$ae_{\hat{g}}$
100	0.6744 7	1.2004	0.0004	0.0003	0.0156	0.7863	0.0863	0.0021	0.0869	0.0326
300	0.6743	1.2001	0.0002	0.0001	0.0086	0.7248	0.0248	0.0005	0.0296	0.0230
500	0.6767	1.1998	0.0002	0.0000	0.0069	0.7082	0.0082	0.0001	0.0118	0.0196
700	0.6768	1.2008	0.0008	0.0000	0.0052	0.7016	0.0016	0.0000	0.0053	0.0171

Below, let us discuss some important findings. Once again, the functional forms of $\iota(\cdot)$ seem to be important factors which determines the nature of endogeneity. With an instrument of a particular explanatory power in (2.28), linear-endogeneity tends to give a higher $Corr_L$ than $Corr_{NL}$ obtained from its nonlinear counterpart. An important observation which can be brought forward is that even for cases in which we are able to identify a strong instrument (with strong explanatory power), the impact of endogeneity is still determined by the relationship between ϵ and η, i.e. the conditional expectation of the former with respect to the latter. Furthermore, compared the results in Tables 2.9 and 2.10 to those presented in Tables 2.6 and 2.7, it is clear that our newly developed Procedure 2.2 performs much better than its Procedure 2.1 counterpart in the presence of endogeneity. Procedure 2.2 seems to be capable of obtaining consistent estimators of all the unknowns, including the parametric and index coefficients, and the unknown structural function.

3. Semi-parametric Analysis of Shape-Invariant Empirical Engel Curves

In this section, we will study the relationships between expenditure on specific goods and the level of total expenditure by using our newly established method to conduct a semiparametric analysis of shape-invariant Engel curves in the Australian context. The data used is based on the Household, Income and Labor Dynamics in Australia (HILDA) Survey, which is Australia's household-based panel study that began in 2001. The goal of such a survey is to collect information about economic and subjective well-being, the labour market and family dynamics. The survey consists of more than 7,500 households with just below 20,000 individuals. The current release, i.e. Release 8, covers the first eight waves (out of 11) of data, which has recently become publicly available. The current section consists of four subsections. In Section 3.1, we explain the empirical model which our analysis will be based on. Section 3.2 discusses the details of the relevance of endogeneity in the study at hand. In Section 3.3, we then discuss the empirical estimation of the shape-invariant Engel curves and we present a number of important findings in Section 3.4.

3.1. The Empirical Model

Hereafter, let $\{Y_{1il}, X_{1i}, X_{2i}\}_{i=1}^{n}$ represent an i.i.d. sequence of n household observations on the budget share Y_{1il} of good $l = 1, \ldots, L \geq 1$ for each household i facing the same relative prices, the log of total expenditure X_{1i}, and a vector of household composition variables X_{2i}. For each commodity l, budget shares and total outlay are related by the general stochastic Engel curve $Y_{1il} = G_l(X_{1i}) + \epsilon_{il}$, where G_l is an unknown function that can be estimated using a standard nonparametric regression method under the exogeneity assumption of X_1, i.e. $E(\epsilon_{il}|X_{1i}) = 0$. Furthermore, a number of previous studies have reported that household expenditures typically display a large variation with demographic composition. When X_2 is discrete, a simple approach for model estimation is to stratify by each distinct discrete outcome of X_2 and then estimate using nonparametric regression within each cell. At some point, however, it may be useful to pool the Engel curves across household demographic types and to allow X_1 to enter each Engel curve semiparametrically. This idea leads to the following specification:

$$Y_{1il} = g_l(X_{1i} - \phi(X_{2i}'\alpha_0)) + X_{2i}'\beta_{0l} + \epsilon_{il}, \tag{3.1}$$

where $g_l(\cdot)$ is an unknown function and $\phi(X_{2i}'\alpha_0)$ is a known function up to a finite set of unknown parameters α_0 that can be interpreted as the log of general equivalence scales for household i.

The functional form specification in (3.1) deserves a few remarks. To this end, Blundell et al. (2003) show that such the functional form specification is consistent with consumer optimisation theory; see also the discussion of Lemma 3.2 of Blundell et al. (1998). Furthermore, in the current paper, we choose $\phi(X_{2i}'\alpha_0) = X_{2i}'\alpha_0$, where X_{2i} is a vector of the demographic variables that represent different household types and α_0 is the vector of the corresponding equivalence scales. Hence we have the following EGPLSI specification:

$$Y_{1il} = g_l(X_{1i} - X_{2i}'\alpha_0) + X_{2i}'\beta_{0l} + \epsilon_{il}. \tag{3.2}$$

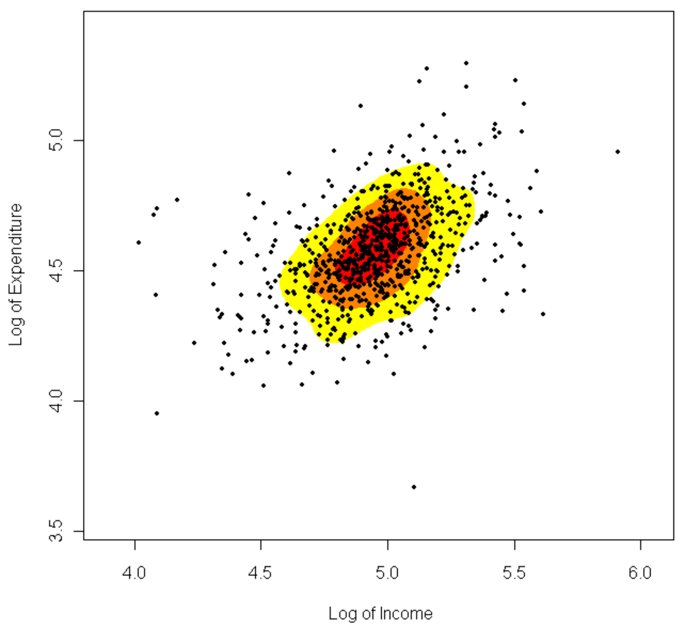
In our application, we consider six broad categories of goods, namely food, clothing, alcohol, electricity and gas, transportation and other goods. In order to preserve a degree of demographic homogeneity, we select a subset of married (or cohabiting) couples with one or two dependent children aged less than 16 years, in five Australian territory capital cities, namely Adelaide, Brisbane, Melbourne, Perth and Sydney. Therefore, our demographic variable, X_2, is simply a binary dummy variable that reflects whether the couple has one child ($X_2 = 0$) or two children ($X_2 = 1$). This leaves us with 817 observations, including 286 couples with one child.

Table 3.1. *Descriptive statistics.*

	Couples 1 child		**Couples 2 children**	
	Mean	**Std. Dev**	**Mean**	**Std. Dev**
Budget shares:				
Alcohol	0.03373	0.03608	0.02918	0.03409
Clothing	0.03060	0.02343	0.03212	0.02788
Electricity and gas	0.04077	0.16236	0.03850	0.14124
Food	0.31515	0.02600	0.31303	0.02872
Transportation	0.04076	0.00153	0.04385	0.00124
Other	0.56870	0.03060	0.57263	0.02308
Expenditure and income:				
log (total expenditure)	4.53302	0.20566	4.58983	0.17854
log (income)	4.92124	0.23414	4.96652	0.23769
Sample size	286		531	

The budget shares of these goods are presented in Table 3.1. The log of total expenditure on the these goods is our measure of the continuous endogenous explanatory variable X_1. Furthermore, Table 3.1 also presents descriptive statistics for the main variables used in this study. The table shows larger expenditure shares for alcohol, electricity and gas, and food for the couples with one child, but larger expenditure shares for clothing, transportation and other goods for the couples with two children. This indicates the differences in the consumption patterns between the two demographic groups, and we expect the estimators of the scale and shift coefficients to reflect these patterns.

Figure 3.1. *Kernel joint density estimates with a full bandwidth matrix.*



3.2. A Simple Test of Endogeneity

Regarding the empirical study in the current section, in order to see the reason why the log of total expenditure X_1 is likely to be endogenous, i.e. $E(\epsilon_l|x_1) \neq 0$, let us note firstly that the system of budget shares can be thought of as the second stage in a two-stage budgeting model (see Gorman (1959) for details), in which total expenditure and savings are first determined conditional on total expenditure, and individual commodity shares are chosen at the second stage; see Blundell (1988) for example. Hence X_1 is a variable which reflects savings and other consumption decisions made at the same time as the budget shares Y_1 are chosen. In our analysis that follows, we consider an earning variable, which is the amount that a household earned before tax in the chosen year, as an instrument.

Figures 3.1 and 3.2 present a plot of the kernel estimates for the joint density of log(total expenditure) and log(earning) and a plot for $E(\log(\text{expenditure})|\log(\text{earning}))$, respectively. The two variables show strong positive correlation such that for the sample with one child, the correlation is 0.4882 and is 0.4056 for those with two children. As seen in the figure, the joint density is also smooth and, together with the conditional mean, confirms our belief that the gross earnings variable should be a good choice for our instrumental variable. Since the kernel estimate of the density of log earnings is close to normal, we have taken the instrumental variable $Z = \Phi(\text{log earnings})$ in the empirical applications and write:

$$\eta_i = X_{1i} - m_{X1}(Z_i). \tag{3.3}$$

Our model, which consists of the index model in (3.2) and the specification of the endogeneity control regressor in (3.3), is appropriate for the application since it is coherent with the economic theory and it allows for the endogeneity of total expenditure as discussed earlier.

Figure 3.2. *Kernel estimates of conditional expectation of log(expenditure) with respect to log(income).*

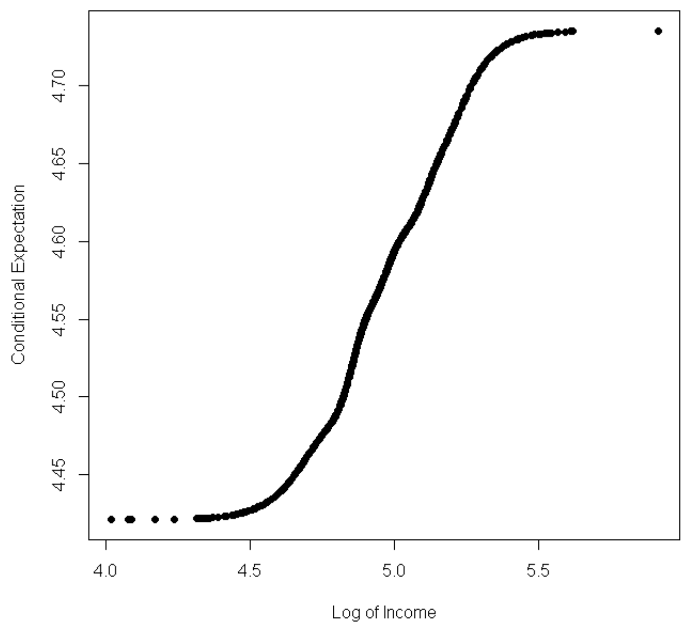
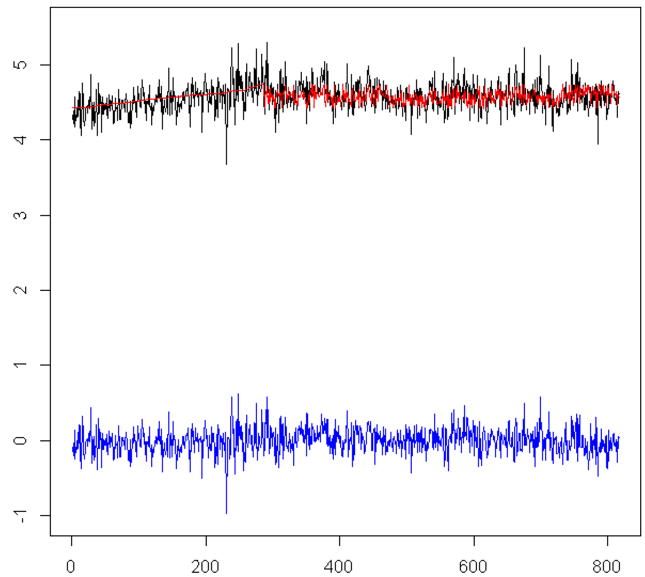


Figure 3.3 shows log(expenditure) (black line), m_{X1} (red line) and η (blue line). In the view of this triangular structure, the figure stresses that the endogenous variable, X_1, may be decomposed into the exogenous (i.e. Z) and the endogenous (i.e. η) components. An important observation to be noted is that even for cases in which we are able to identify a strong instrument (with strong explanatory power), the impact of endogeneity is still determined by the relationship between ϵ_l and η, i.e. the conditional expectation of the former with respect to the latter. We will explore this point further below.

Figure 3.3. *log(expenditure),* m_{X1} *and* η.



In the following, we discuss the construction of variability bands in our analysis and how they can be used as a preliminary test of exogeneity. For convenience, let us first restate the triangular structure as:

$$Y_{1il} = g_l(X_{1i} - X_{2i}'\alpha_0) + X_{2i}'\beta_{0l} + \epsilon_{il}, \quad (3.4)$$

$$X_{1i} = m_{X1}(Z_i) + \eta_i, \quad (3.5)$$

where $m_{X1}(z) = E(X_{1i}|Z_i = z)$, under the assumptions of the following:

$$E(\eta_i|Z_i = z) = 0 \text{ and } E(\epsilon_{il}|Z_i = z, \eta_i = \eta) = E(\epsilon_{il}|\eta_i = \eta) \neq 0. \quad (3.6)$$

The structure described in (3.4) to (3.6) suggests that we have

$$E[Y_{1il}|(X_{1i} - X_{2i}'\alpha_0), \eta_i] - E[X_{2i}|(X_{1i} - X_{2i}'\alpha_0), \eta]'\beta_{0l} = g_l(X_{1i} - X_{2i}'\alpha_0) + \iota_l(\eta_i), \quad (3.7)$$

where $E[\epsilon_{il}|(X_{1i} - X_{2i}'\alpha) = (x_1 - x_2'\alpha), \eta_i = \eta] = E[\epsilon_{il}|X_{2i} = x_2, \eta_i = \eta] = E[\epsilon_{il}|\eta_i = \eta] \equiv \iota_l(\eta) \neq 0$. Expression (3.7) then implies

$$Y_{1il} = X_{2i}'\beta_{0l} + g_l(X_{1i} - X_{2i}'\alpha_0) + \iota_l(\eta_i) + e_{il}, \quad (3.8)$$

$$X_{1i} = m_{X1}(Z_i) + \eta_i. \quad (3.9)$$

where $E(e_l|\eta) = 0$. Let $M_l[(X_{1i} - X_{2i}'\alpha_0), \eta_i] = g_l(X_{1i} - X_{2i}'\alpha_0) + \iota_l(\eta_i)$. In order to use (3.8), it is important to note that:

$$\begin{aligned} m_l(x_1 - x_2'\alpha_0) &= \int M_l(\{x_1 - x_2'\alpha_0\}, \eta)\, d\eta = g_l(x_1 - x_2'\alpha_0) + c_1 \\ g_l(x_1 - x_2'\alpha_0) &= m_l(x_1 - x_2'\alpha_0) - c_1, \end{aligned} \tag{3.10}$$

where $c_1 = \int \iota(\eta) dQ(\eta)$ and $E(g_l(\cdot)) = 0$; the estimation of which can be done based on the marginal integration technique in Step 2.2.5 of Procedure 2.2.

Now, observe that if we were to impose a linear specification on $\iota_l(\cdot)$, (3.8) would be closely similar to the extended partially linear (EPL) model discussed in Blundell et al. (1998). In this case, Blundell et al. (1998) showed that a test of the exogeneity null can be constructed by testing $H_0 : \iota_l = 0$, where ι_l is an unknown parameter. To allow for more flexibility on the functional form between the total expenditure and its instrument, as an alternative, one may apply an existing test of a parametric mean-regression model against a nonparametric alternative; see Horowitz and Spokoiny (2001), for example. However, in the current paper, we suggest that it is more convenient to simply construct the variability bands for $\iota_l(\cdot)$ since its estimate is readily available. To do so, we use the following procedure.

Procedure 3.2

Step 3.2.1: Obtain an empirical estimate of $g_l(x_1 - x_2'\alpha_0)$ in (3.10); see also Remark 3.1.

Step 3.2.2: Regress (3.9) using the estimates in Step 3.2.1 to obtain the nonparametric estimates of $\iota_l(\cdot)$.

Step 3.2.3: Compute the bias-corrected confidence bands for the nonparametric regression using the procedure introduced in Xia (1998). Finally, the above mentioned (Bonferroni-type) variability bands are obtained using a similar procedure discussed in Eubank and Speckman (1993).

Remark 3.1. *To complete Step 3.2.1, Procedure 2.2 in Section 2.2 can be useful. However, some modifications are required to take the index coefficient α_0 into account, which can be interpreted as a general equivalence scale for household i. Steps 2.2.1 and 2.2.2 are directly applicable since they are implemented using a given α across $l = 1, 2, \ldots, 6$ commodities. In this case, the objective function (2.17) in Step 2.2.3 is only used for the particular l commodity. A new objective function is the summation of these individual functions, i.e. $\min_{\alpha \in A_n, h_{v,l}, h_{\hat{\eta},l} \in \mathcal{H}_n} \hat{J}(\alpha, h_{v,l}, h_{\hat{\eta},l})$, which is minimised with respect to α and 12 bandwidth parameters, i.e. two for each commodity. Finally, Steps 2.2.4 and 2.2.5 are directly applicable using $\hat{\alpha}$ as well as $\hat{h}_{v,l}$ and $\hat{h}_{\hat{\eta},l}$.* ■

3.3. Shape-Invariant Engel Curves

First, observe that (3.8) can also be re-stated as:

$$\tilde{Y}_{1il} = g_l(X_{1i} - X'_{2i}\alpha_0) + e_{il}, \tag{3.11}$$

where $\tilde{Y}_{1il} = Y_{1il} - X'_{2i}\beta_{0l} - \iota_l(\eta_i)$. The use of (3.11) relies on the following corresponding expression of (3.10):

$$\begin{aligned} m_l(\eta) &= \int M_l(v,\eta)\ dv = \iota_l(\eta) + c_2 \\ \iota_l(\eta) &= m_l(\eta) - c_2, \end{aligned} \tag{3.12}$$

where $v = x_1 - x'_2\alpha$, $c_2 = \int g(v)dQ(v)$ and $E(\iota_l(\cdot)) = 0$. Hence (3.11) suggests that we are able to employ *Procedure 3.3* below in order to obtain the estimates of the shape-invariant Engel curves and the related confidence bands.

Procedure 3.3

Step 3.3.1: Obtain empirical estimates of $\iota_l(\eta)$ in (3.12).

Step 3.3.2: Regress (3.11) using the estimates in Step 3.3.1 to obtain the nonparametric estimates of $g_l(\cdot)$.

Step 3.3.3: Compute the bias-corrected confidence bands about the nonparametric estimator in Step 3.3.2 using the procedure introduced in Xia (1998).

3.4. Empirical Findings

Prior to presenting our empirical findings, let us recapitulate our empirical model of shape-invariant Engel curves and made a final remark on the identification of the model. The empirical model we are attempting to estimate is of the following form:

$$\begin{aligned} Y_{1il} &= g_l(X_{1i} - \alpha_0 X_{2i}) + \beta_{0l}X_{2i} + \epsilon_{il}, \\ X_{1i} &= m_{X1}(Z_i) + \eta_i. \end{aligned} \tag{3.13}$$

However, the EGPLSI structure suggests that an unrestricted version of (3.13) is, in fact, $Y_{1il} = g_l(\alpha_{01}X_{1i} - \alpha_0 X_{2i}) + \beta_{01,l}X_{1i} + \beta_{0l}X_{2i} + \epsilon_{il}$, while the restrictions $\alpha_{01} = 1$ and $\beta_{01,l} = 0$ leads to $Y_{1il} = g_l(X_{1i} - \alpha_0 X_{2i}) + \beta_{0,l}X_{2i} + \epsilon_{il}$. To ensure the model's estimability, the following assumption, which is based closely on Assumption I of Ai and Chen (2003), is required.

Assumption 3.1. *Suppose that*

$$\begin{aligned}
&E\left[\left(Y_{1il}-M_l(\alpha_{01}X_{1i}-\alpha_0X_{2i})-\beta_{01,l}X_{1i}-\beta_{0l}X_{2i}\right)|X_{2i},Z_i\right]\\
=\;&E\left[\left(Y_{1il}-M_l(\alpha_{01}X_{1i}-\alpha_0X_{2i})-\beta_{0l}X_{2i}\right)|X_{2i},Z_i\right]=0,
\end{aligned}$$

which implies

$$E[Y_{1il}|X_{2i},Z_i]=E\left[\left(M_l(\alpha_{01}X_{1i}-\alpha_0X_{2i})+\beta_{0l}X_{2i}\right)|X_{2i},Z_i\right],$$

and, therefore,

$$E[\beta_{01,l}X_{1i}|X_{2i},Z_i]=0,$$

which is what is required by the perpendicularity of Xia et al. (1999). ■

Hereafter, let us use $\hat{g}_{1,l}(\cdot)$ and $\hat{\iota}_{1,l}(\cdot)$ to denote the empirical estimates of $g_l(\cdot)$ and $\iota_l(\cdot)$ based on the marginal integration techniques, i.e. those obtained from Steps 3.2.1 and 3.3.1, respectively. Furthermore, let us use $\hat{g}_{2,l}(\cdot)$ and $\hat{\iota}_{2,l}(\cdot)$ to denote the empirical estimates of $g_l(\cdot)$ and $\iota_l(\cdot)$ which are obtained from Steps 3.2.2 and 3.3.2, respectively. Table 3.2 below presents the empirical estimates of the unknown parameters α_0 and β_{0l} (3.4). In addition, to demonstrate the validity of our Procedures 3.2 and 3.3 above, in the table we also present in the following average squared difference:

$$d_{gl}=\frac{1}{n}\sum_{i=1}^{n}\left\{\hat{g}_{1,l}(\hat{v})-\hat{g}_{2,l}(\hat{v})\right\}^2 \quad\text{and}\quad d_{\iota l}=\frac{1}{n}\sum_{i=1}^{n}\left\{\hat{\iota}_{1,l}(\hat{\eta})-\hat{\iota}_{2,l}(\hat{\eta})\right\}^2,$$

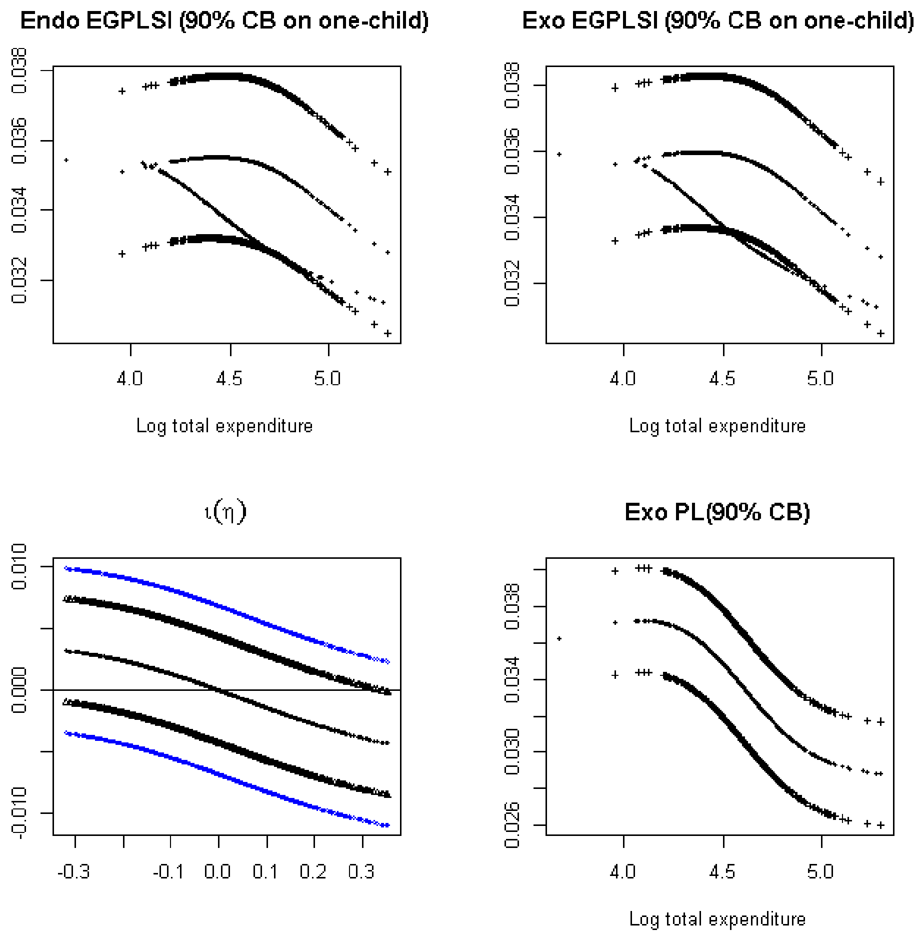
where $\hat{v}=x_1-\hat{\alpha}x_2$.

Table 3.2. *Empirical results*

$\hat{\alpha}$	Categories of goods	$\hat{\beta}_l$	d_{gl}	$d_{\iota l}$	$\hat{h}_{v,l}$	$\hat{h}_{\hat{\eta},l}$
0.5813	Alcohol	-0.0053	3.9781e-07	3.2355e-06	0.581334	0.581333
	Clothing	0.0005	7.8607e-07	6.4676e-06	0.581332	0.581330
	Food	-0.4541	3.4367e-04	1.7932e-04	0.065466	0.065465
	Electricity and Gas	0.0133	6.9226e-06	2.8772e-06	0.065465	0.065466
	Transportation	-0.0024	5.3794e-07	2.3716e-06	0.581335	0.581333
	Other	0.1245	1.6083e-04	2.8754e-04	0.065466	0.065465

We will now summarise a number of important findings based on the empirical results in Table 3.2 and Figures 3.4 to 3.9.

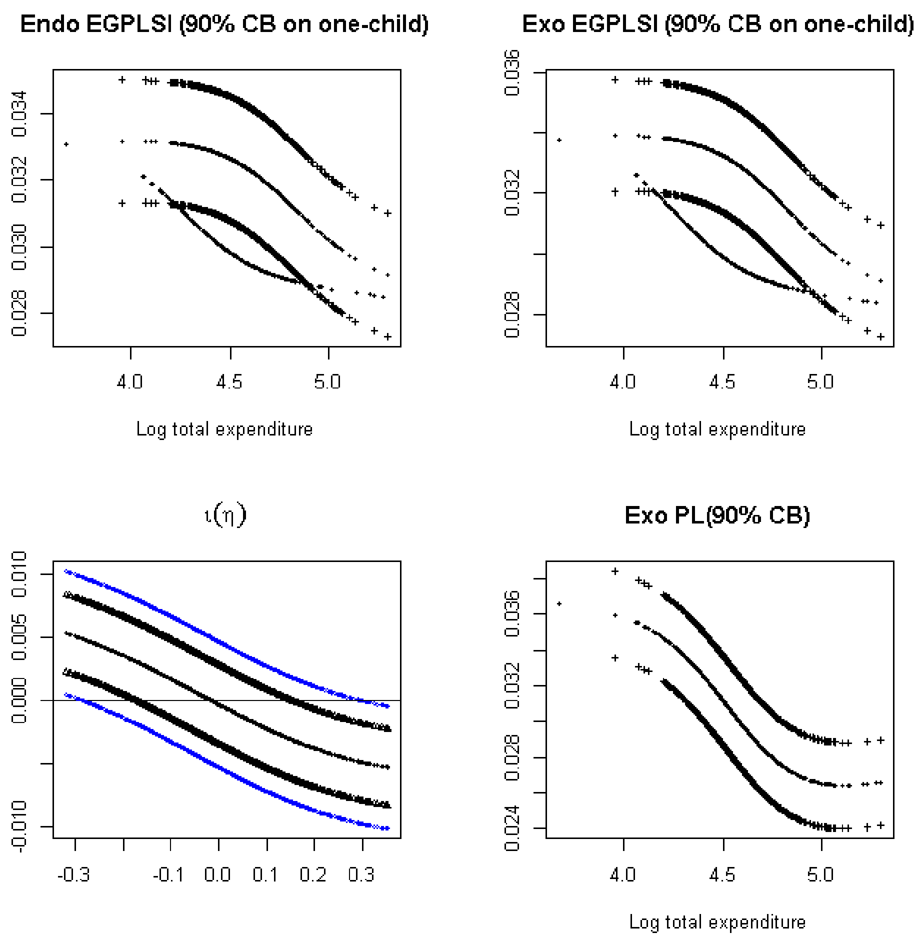
Figure 3.4. *Engel curves for alcohol*



Firstly, the average squared errors reported in the fourth and the fifth columns of Table 3.2 are virtually zero, which provides strong evidence in support of the procedures discussed in Sections 3.2 and 3.3. Secondly, the signs and magnitudes of the estimates of the parameters reported in the first and the third columns are consistent with what is reported in the existing literature; see Blundell et al. (1998) for example. Furthermore, Figures 3.4 to 3.9, present the Engel curves for the six budget shares in our HILDA sample, each of which consists of four panels. The first and second panels present estimates of the Engel curves (for couples with one child and couples with two children) based on the EGPLSI model with the endogeneity being controlled using Procedure 2.2 and the endogeneity not being controlled by Procedure 2.1 in Section 2, respectively. Xia's (1998) confidence bands are constructed for the Engel curves of couples with one child. Furthermore, the fourth panels present estimates of the

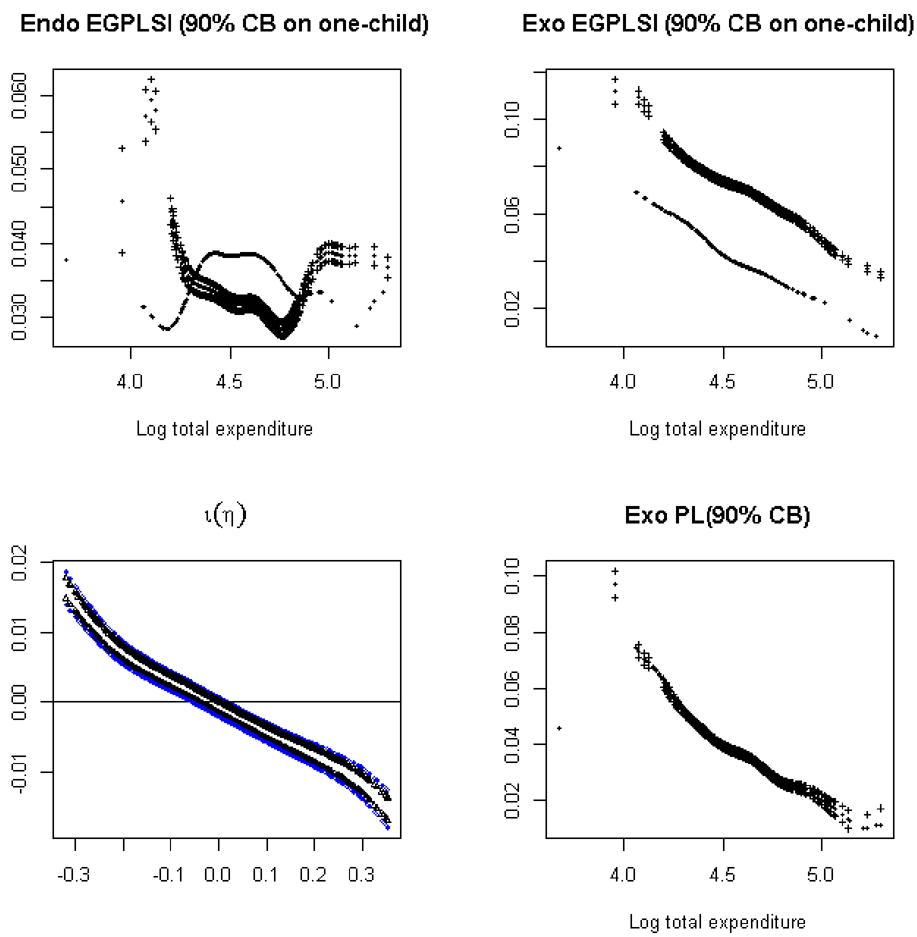
Engel curves computed using the partially linear model of Robinson (1988) for the sake of comparison with the EGPLSI model. They show clear evidence that the partially linear model restricts the empirical Engel curves to be within the same specification; see Blundell et al. (1998) and Blundell et al. (2003) for example, where all empirical Engel curves are similar to the quadratic functional form.

Figure 3.5. *Engel curves for clothing*



Finally, the third panel of each graph presents the nonparametric estimates of $\iota_l(\cdot)$ with two sets of bands, namely the bias-corrected confidence bands for the nonparametric regression of Xia (1998) (black) and the Bonferroni-type variability bands discussed in Eubank and Speckman (1993) (blue). Regarding alcohol, clothing and transportation, $\iota_l(\cdot)$ for these cases do not seem to be statistically significant. These findings can be linked to the fact that the shapes of the Engel curves presented in the top two panels are similar. In other words, we show that the seriousness of the effect of the endogeneity problem, given an instrument, depends very much on the relationship between the disturbances in the structural and the reduced relations, i.e. the relationship between ϵ and η, which, in this case, is summarised by $\iota_l(\eta)$. For a given instrument and therefore the corresponding η, $\iota_l(\cdot)$ can be a function such that the impact of endogeneity is minimal, e.g. in the case of alcohol, clothing and transportation. Otherwise, they may be functions which make the effect of the endogeneity severe, such as the case of electricity and gas.

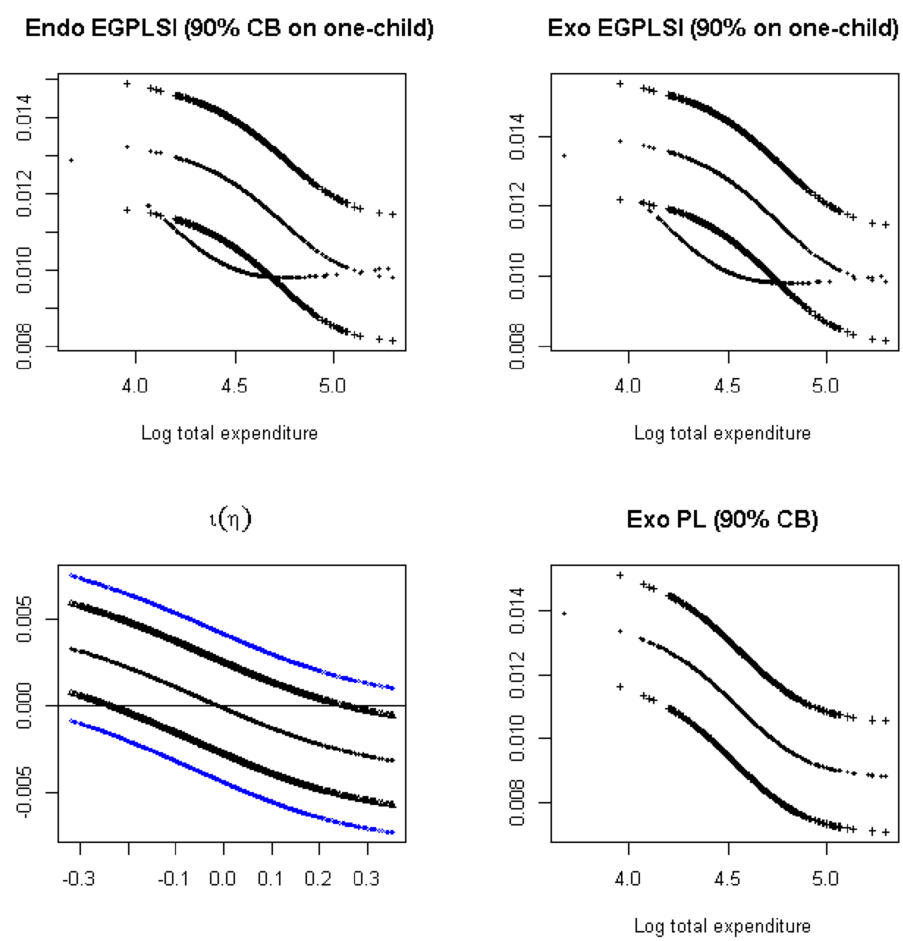
Figure 3.6. *Engel curves for electricity and gas*



Some of these Engel curves, e.g. those of alcohol, clothing and transportation, appear to demonstrate that the Working-Leser linear logarithmic (Piglog) formulation is a reasonable approximation. Nonetheless, for other shares, particularly electricity and gas, and food and other goods, a more nonlinear relationship between the shares and the log expenditure is evident. Regarding alcohol, clothing and transportation, although the Engel curves for our two demographic groups both slope downward a broadly parallel shift in the Engel curves does not seem to appear. In fact, the Engel curves of families with two children tend to decline at a much faster rate as the log total expenditure increases.

On the contrary, it is interesting to note how similar the shapes of the Engel curves are for our two demographic groups for food and other goods. In these cases, there appears to be a parallel shift in the Engel curves. A couple with one child spends around 15% more of their budget on food than a couple with two children. However, couples with two children end up spending 4% more of their budget on other goods than couples with one child at the same level of expenditure. Such outcomes seem consistent with our intuitive belief about consumption behaviour in practice, i.e. a couple with two children incurs additional costs for having an extra child which are hidden within the other goods category.

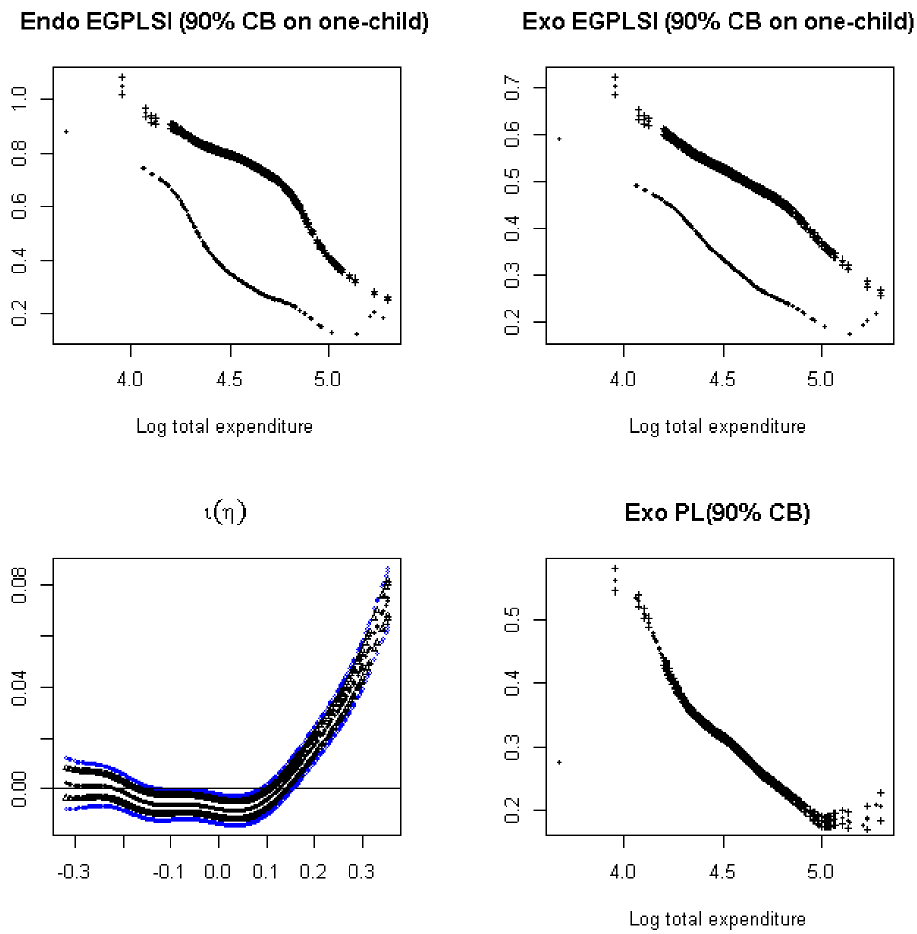
Figure 3.7. *Engel curves for transportation*



4. Conclusions

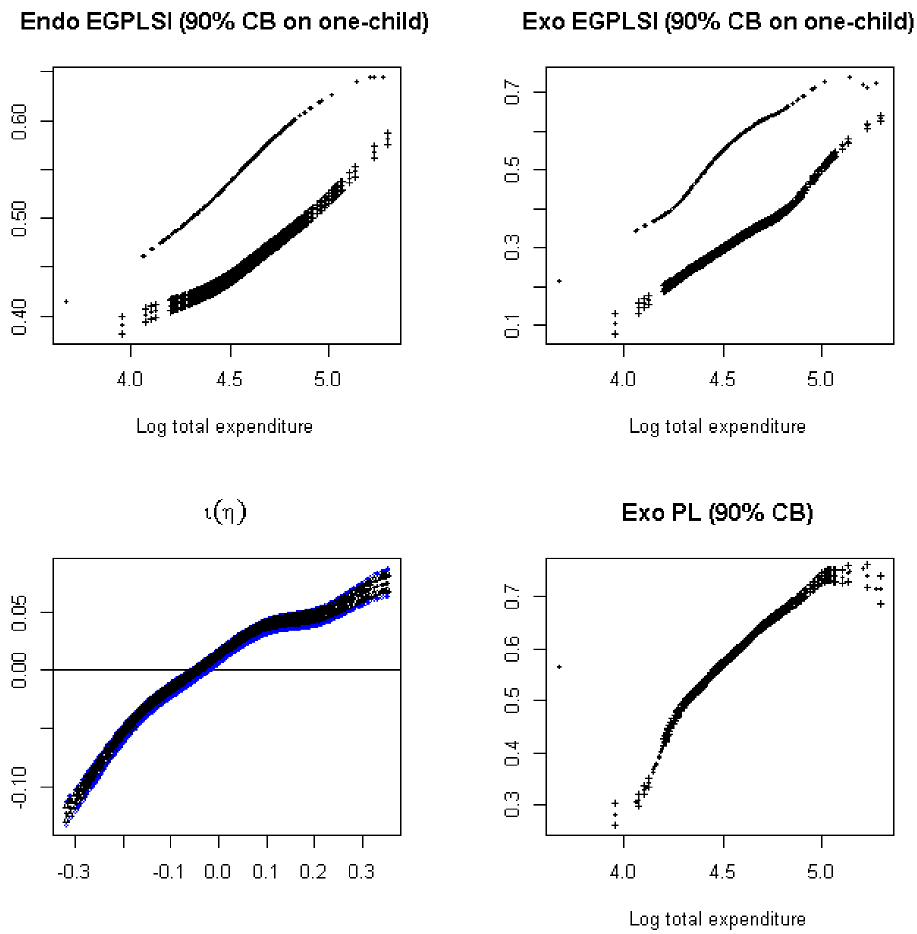
Although the GPLSI model by Carroll et al. (1997), and Xia and Härdle (2006) has great flexibility and advantages from both a PL model and a SI model perspective, it is not appropriate for modelling the shape-invariant empirical Engel curve, since it does not allow the coefficient of the household equivalence scale to be included. Hence we consider the EGPLSI model of Xia et al. (1999) and Gao (2007) in order to take the shape-invariant specification into account.

Figure 3.8. *Engel curves for food*



However, the estimation method and procedure of the existing EGPLSI model are not applicable, since the endogeneity of total expenditure is well known in the literature. Hence, we establish the CF approach in the EGPLSI model to address endogeneity instead of the nonparametric IV estimation of Ai and Chen (2003); see Blundell et al. (2007) for its application to an semiparametric analysis of empirical Engel curves. The attractive feature of the proposed estimation procedure in the current study is that the practicality of Xia et al. (1999) approach is still applicable, despite the endogeneity control variable generated. The same bandwidth parameters are used for the estimation of coefficient of an equivalence scale and a structural Engel function. In addition, we also consider the "biased-adjusted" confidence band for the nonparametric structural function since the index coefficient is estimated and the endogeneity control regressor is generated. This corrected confidence band gives us useful information such as whether the effect of endogeneity is significant by analysing whether the band is significantly different from zero.

Figure 3.9. *Engel curves for other goods*



We also provide Monte Carlo simulation studies and an application of the methodology to the Australian HILDA dataset. The simulation studies illustrate the performance of CF approach and the usefulness of the adjusted confidence band. The application illustrates that the partially linear model restricts empirical Engel curves to be within the same specification (see Blundell et al. (1998), and Blundell et al. (2003) for details), where all empirical Engel curves are similar to the quadratic functional form. However, the EGPLSI model which, coherent with consumption theory, shows different functional forms for different commodities. Also, the EGPLSI model shows that the effect of endogeneity on total expenditure is non-trivial, the magnitude of the effects can be measured by the endogeneity control functions and they are significantly different from zero.

5. Appendix

In this section, we provide the necessary mathematical proofs of the main theoretical results of the current paper. In Section 5.1, we show the proofs of Theorem 2.1 and Corollary 2.1 in two main steps. In Section 5.2, we present the proofs of Theorem 2.2.

5.1. Proofs of Theorem 2.1 and Corollary 2.1

Step 1. Proofs of Theorem 2.1: Given α, $\hat{\eta}$ and $\hat{\beta}$, the feasible objective function (2.17) is expanded as follows:

$$\begin{aligned}\hat{J}(\alpha, h_v, h_{\hat{\eta}}) &= \frac{1}{n}\sum_{i=1}^{n}\left[Y_i - \hat{Y}_i + \hat{Y}_i - \hat{Y}_{2i} - \left\{X_i - \hat{X}_i + \hat{X}_i - \hat{X}_{2i}\right\}'\hat{\beta}\right]^2 \\ &\equiv \frac{1}{n}\sum_{i=1}^{n}\left[Y_i - \hat{Y}_i - \hat{\delta}_{Y,i} - \left\{X_i - \hat{X}_i - \hat{\delta}_{X,i}\right\}'\hat{\beta}\right]^2 \\ &= \hat{J}^*(\alpha, h_v, h_\eta) + R_1(\alpha, h_v, h_\eta, h_z), \qquad (5.A.1)\end{aligned}$$

where $\hat{Y}_{2i} = \hat{m}_y(V_i, \hat{\eta}_i) + \hat{W}_{2i}$ and $\hat{X}_{2i} = \hat{m}_x(V_i, \hat{\eta}_i) + \hat{U}_{2i}$ with $\hat{W}_{2i} = \frac{\sum_{j\neq i} W_j L_{2,ij}}{\sum_{j\neq i} W_j L_{2,ij}}$, $\hat{U}_{2i} = \frac{\sum_{j\neq i} U_j L_{2,ij}}{\sum_{j\neq i} W_j L_{2,ij}}$ and $L_{2,ij} = L_{h_v,h_\eta}(V_i - V_j, \hat{\eta}_i - \hat{\eta}_j)$, and $\hat{\delta}_{Y,i} = \hat{Y}_{2i} - \hat{Y}_i$, $\hat{\delta}_{X,i} = \hat{X}_{2i} - \hat{X}_i$ with $\hat{Y}_i = \hat{m}_y(V_i, \eta_i) + \hat{W}_i$, $\hat{X}_i = \hat{m}_x(V_i, \eta_i) + \hat{U}_i$ with $\hat{W}_i = \frac{\sum_{j\neq i} W_j L_{ij}}{\sum_{j\neq i} L_{il}}$, $\hat{U}_i = \frac{\sum_{j\neq i}^{n} U_j L_{ij}}{\sum_{j\neq i}^{n} L_{il}}$ and $L_{ij} = L_{h_v,h_\eta}(V_i - V_j, \eta_i - \eta_j)$. Let us note that $m = m(v_0, \eta)$ and $\tilde{m} = E(m|\alpha)$. Note that the term in the last equation of (5.A.1), $\hat{J}^*(\alpha, h_v, h_\eta)$, is further expanded, as shown below:

$$\hat{J}^*(\alpha, h_v, h_\eta) = \frac{1}{n}\sum_{i=1}^{n}\left[Y_i - \hat{Y}_i - \left\{X_i - \hat{X}_i\right\}'\hat{\beta}_\alpha\right]^2 = \tilde{J}(\alpha) + T(h_v, h_\eta) + R_2(\alpha, h_v, h_\eta),$$

where:

$$\begin{aligned}R_2(\alpha, h_v, h_\eta) &= (\hat{\beta}-\beta_0)'S_{\tilde{m}_x-\hat{m}_x}(\beta_0-\hat{\beta}) + (\hat{\beta}-\beta_0)'S_{\hat{U}}(\beta_0-\hat{\beta}) - 2(\hat{\beta}-\beta_0)'S_{m_x-\tilde{m}_x,\tilde{m}_x-\hat{m}_x}(\beta_0-\hat{\beta}) \\ &- 2(\hat{\beta}-\beta_0)'S_{m_x-\tilde{m}_x,\hat{U}}(\beta_0-\hat{\beta}) + 2(\hat{\beta}-\beta_0)'S_{\tilde{m}_x-\hat{m}_x,U}(\beta_0-\hat{\beta}) - 2(\hat{\beta}-\beta_0)'S_{\tilde{m}_x-\hat{m}_x,\hat{U}}(\beta_0-\hat{\beta}) \\ &- 2(\hat{\beta}-\beta_0)'S_{U\hat{U}}(\beta_0-\hat{\beta}) + S_{\tilde{m}-\hat{m}} + 2S_{m-\tilde{m},\tilde{m}-\hat{m}} - 2S_{e\hat{e}} + S_{\hat{e}} - 2(\hat{\beta}-\beta_0)'S_{\tilde{m}_x-\hat{m}_x,m-\tilde{m}} \\ &+ 2(\hat{\beta}-\beta_0)'S_{\hat{U},m-\tilde{m}} - 2(\hat{\beta}-\beta_0)'S_{m_x-\tilde{m}_x,\tilde{m}-\hat{m}} - 2(\hat{\beta}-\beta_0)S_{\tilde{m}_x-\hat{m}_x,\tilde{m}-\hat{m}} - 2(\hat{\beta}-\beta_0)'S_{U,\tilde{m}-\hat{m}} \\ &+ 2(\hat{\beta}-\beta_0)'S_{\hat{U},\tilde{m}-\hat{m}} - 2(\hat{\beta}-\beta_0)'S_{\tilde{m}_x-\hat{m}_x,e} + 2(\hat{\beta}-\beta_0)'S_{\hat{U}e} + 2(\hat{\beta}-\beta_0)'S_{U\hat{e}} + 2(\hat{\beta}-\beta_0)'S_{m_x-\tilde{m}_x,\hat{e}} \\ &+ 2(\hat{\beta}-\beta_0)'S_{\tilde{m}_x-\hat{m}_x,\hat{e}} + 2(\hat{\beta}-\beta_0)'S_{U\hat{e}} - 2(\hat{\beta}-\beta_0)S_{\hat{U}\hat{e}} + 2S_{\tilde{m}-\hat{m},e} - 2S_{m-\tilde{m},\hat{e}} - 2S_{\tilde{m}-\hat{m},\hat{e}} - S_{m-\hat{m}_0}\end{aligned}$$

with $\tilde{Y}_i = \tilde{m}_{y,i}$ and $\tilde{X}_i = \tilde{m}_{x,i}$ since $E(w|x,\eta) = 0$ and $E(u|x,\eta) = 0$, and $\hat{m}_0 = \frac{\sum_{j\neq i} m_j L_{0,ij}}{\sum_{j\neq i} L_{0,ij}}$, with $L_{0,ij} = L_{h_v,h_\eta}(V_{0i} - V_{0j}, \eta_i - \eta_j)$. The results of $\sup_{\alpha\in\mathcal{A}_n, h_v, h_\eta\in\mathcal{H}_n} |R_2(\alpha, h_v, h_\eta)| = o_p(n^{-1/2})$ are easily shown using the fact that $\beta_0 - \hat{\beta} = o_p(n^{-1/2})$ shown below, and Propositions 5.A.1, 5.A.2, 5.A.3, 5.A.6, 5.A.7, 5.A.9, 5.A.12, 5.A.13, and 5.A.14. The last term in R_2 is $S_{m-\hat{m}_0} = O_p(n^{-1}h_v^{-1}h_\eta^{-q_2}) + O_p((h_v^2 + h_\eta^2)^2)$ by a simple non-parametric analysis. This is a simple extension of the results in Xia et al. (1999). Hence the objective function (5.A.1) is rewritten as:

$$\hat{J}(\alpha, h_v, h_{\hat{\eta}}) = \tilde{J}(\alpha) + T(h_v, h_\eta) + R_1(\alpha, h_v, h_\eta, h_z) + o_p(n^{-1/2}),$$

where:

$$R_1(\alpha, h_z, h_v, h_\eta) = (\hat{\beta}-\beta_0)'S_{\hat{\delta}_X}(\hat{\beta}-\beta_0) + S_{\hat{\delta}_m} + S_{\hat{\delta}_e} + 2(\hat{\beta}-\beta_0)'S_{\hat{\delta}_X\hat{\delta}_m} + 2(\hat{\beta}-\beta_0)'S_{\hat{\delta}_X\hat{\delta}_e} - 2S_{\hat{\delta}_m\hat{\delta}_e}$$
$$- 2(\hat{\beta}-\beta_0)'S_{\hat{\delta}_X, m_x-\tilde{m}_x}(\hat{\beta}-\hat{\beta}_0) - 2(\hat{\beta}-\beta_0)'S_{\hat{\delta}_X,\tilde{m}_x-\hat{m}_x}(\hat{\beta}-\beta_0) - 2(\hat{\beta}-\beta_0)'S_{\hat{\delta}_X U}(\hat{\beta}-\beta_0)$$
$$+ 2(\hat{\beta}-\beta_0)'S_{\hat{\delta}_X\hat{U}}(\hat{\beta}-\beta) + 2(\hat{\beta}-\beta_0)'S_{\hat{\delta}_X,m-\tilde{m}} + 2(\hat{\beta}-\beta_0)'S_{\hat{\delta}_X,\tilde{m}-\hat{m}} + 2(\hat{\beta}-\beta_0)'S_{\hat{\delta}_X e} - 2(\hat{\beta}-\beta_0)'S_{\hat{\delta}_X\hat{e}}$$
$$+2(\hat{\beta}-\beta_0)'S_{\hat{\delta}_m,m_x-\tilde{m}_x} + 2(\hat{\beta}-\beta_0)S_{\hat{\delta}_m,\tilde{m}_x-\hat{m}_x} + 2(\hat{\beta}-\beta_0)'S_{\hat{\delta}_m U} - 2(\hat{\beta}-\beta_0)'S_{\hat{\delta}_m\hat{U}} - 2S_{\hat{\delta}_m,m-\tilde{m}}$$
$$-2S_{\hat{\delta}_m,\tilde{m}-\hat{m}} - 2S_{\hat{\delta}_m e} + 2S_{\hat{\delta}_m\hat{e}} + 2(\hat{\beta}-\beta_0)S_{\hat{\delta}_e,m_x-\tilde{m}_x} + 2(\hat{\beta}-\beta_0)'S_{\hat{\delta}_e,\tilde{m}_x-\hat{m}_x} + 2(\hat{\beta}-\beta_0)S_{\hat{\delta}_e U}$$
$$-2(\hat{\beta}-\beta_0)'S_{\hat{\delta}_e\hat{U}} - 2S_{\hat{\delta}_e,m-\tilde{m}} - 2S_{\hat{\delta}_e,\tilde{m}-\hat{m}} - 2S_{\hat{\delta}_e e} + 2S_{\hat{\delta}_e\hat{e}}.$$

In particular, we show that $\sup_{\alpha\in A_n, h_v,h_\eta,h_z\in\mathcal{H}_n} |R_1(\alpha, h_z, h_v, h_\eta)| = o_p(n^{-1/2})$ by using the fact that $\hat{\beta} = \beta_0 + o_p(n^{-1/2})$ and Propositions 5.A.4, 5.A.5, 5.A.8, 5.A.10, 5.A.11 and 5.A.15 below. Hence we have:

$$\hat{J}(\alpha, h_v, h_{\hat{\eta}}) = \tilde{J}(\alpha) + T(h_v, h_\eta) + o_p(n^{-1/2}).$$

■

Step 2. Proofs of Corollary 2.1: We can now present the proofs of asymptotic properties of $\hat{\alpha}$ and $\hat{\beta}$. In view of the representation of $||\alpha-\alpha_0|| \le Cn^{-1/2}$, we may write, for bounded values of x:

$$m(v_0,\eta) = m(v,\eta) - x'(\alpha-\alpha_0)m_0^{(1)} + O(n^{-1}), \quad (5.A.2)$$
$$m(v_0,\eta|v,\eta) = m(v,\eta) - m_x(x|v,\eta)'(\alpha-\alpha_0)m_0^{(1)} + O(n^{-1}), \quad (5.A.3)$$

where $m_x(x|v,\eta) = E(X_{\mathcal{A}_x}|v,\eta)$. Firstly, let us consider the asymptotic properties of $\hat{\alpha}$. Using (5.A.2) and (5.A.3), we have the expansion of $\tilde{J}(\alpha)$ below:

$$\begin{aligned}
\tilde{J}(\alpha) &= \frac{1}{n}\sum_{i=1}^n \Big[m_i - \tilde{m}_i + U_i'(\beta_0-\hat{\beta}) + e_i\Big]^2 \\
&= \frac{1}{n}\sum_{i=1}^n \{m_i-\tilde{m}_i\}^2 + \frac{2}{n}\sum_{i=1}^n \{m_i-\tilde{m}_i\}e_i + \frac{2}{n}\sum_{i=1}^n \{m_i-\tilde{m}_i\}U_i'(\beta_0-\hat{\beta}) \\
&+ \text{ terms independent of } \alpha + o_p(n^{-1/2}) \\
&= (\alpha_0-\alpha)'\left[\frac{1}{n}\sum_{i=1}^n \Big\{m_0^{(1)}\Big\}^2 U_iU_i'\right](\alpha_0-\alpha) + 2\frac{1}{n}\sum_{i=1}^n e_i m_0^{(1)}U_i'(\alpha_0-\alpha) \\
&+ 2(\beta_0-\hat{\beta})'\left[\frac{1}{n}\sum_{i=1}^n m_0^{(1)}U_iU_i'\right](\alpha_0-\alpha) + o_p(n^{-1/2}) \\
&= (\alpha_0-\alpha)'\left[\frac{1}{n}\sum_{i=1}^n \Big\{m_0^{(1)}\Big\}^2 U_{0i}U_{0i}'\right](\alpha_0-\alpha) + 2\frac{1}{n}\sum_{i=1}^n e_i m_0^{(1)}U_{0i}'(\alpha_0-\alpha) \\
&+ 2(\beta_0-\hat{\beta})'\left[\frac{1}{n}\sum_{i=1}^n m_0^{(1)}U_{0i}U_{0i}'\right](\alpha_0-\alpha) + o_p(1) + O_p(n^{-1/2}), \quad (5.A.4)
\end{aligned}$$

where $U_{0i} = \{X_i - E(X_i|V_{0i},\eta_i)\}$.

Given α_0, $(\beta_0-\hat{\beta}) = -\left(\frac{1}{n}\sum_{i=1}^n U_{0i}U_{0i}'\right)^{-}\frac{1}{n}\sum_{i=1}^n U_{0i}e_i$ (see the last equation of (5.A.8) below). Hence we have:

$$\begin{aligned}\tilde{J}(\alpha) &= (\alpha_0-\alpha)'\left[\frac{1}{n}\sum_{i=1}\left\{m_0^{(1)}\right\}^2 U_{0i}U_{0i}'\right](\alpha_0-\alpha)+2\frac{1}{n}\sum_{i=1}^{n} e_i m_0^{(1)}U_{0i}'(\alpha_0-\alpha)\\ &- 2\left[\left(\frac{1}{n}\sum_{i=1}^{n}U_{0i}U_{0i}'\right)^{-}\frac{1}{n}\sum_{i=1}^{n}e_iU_{0i}'\right]\left[\frac{1}{n}\sum_{i=1}^{n}m_0^{(1)}U_{0i}U_{0i}'\right](\alpha_0-\alpha)+o_p(1)+O_p(n^{-1/2})\\ &= (\alpha_0-\alpha)'\left\{m_0^{(1)}\right\}^2 S_{U_0}(\alpha_0-\alpha)+2m_0^{(1)}S_{eU_0}(\alpha_0-\alpha)-2\left\{(S_{U_0})^{-}S_{eU_0}\right\}\left\{m_0^{(1)}S_{U_0}(\alpha_0-\alpha)\right\}\\ &+ o_p(1).\end{aligned}$$

Given $\hat{\eta}$ and α, we write the linear reduced form from Robinson (1988) as follows:

$$Y_i-\hat{Y}_{3i}=(X_i-\hat{X}_{3i})'\beta_0+(m_i-\hat{m}_{3i})+(e_i-\hat{e}_{3i}), \tag{5.A.5}$$

where $\hat{Y}_{3i}=\hat{m}_y(\hat{V}_i,\hat{\eta}_i)+\hat{W}_{3i}$, $\hat{X}_{3i}=\hat{m}_x(\hat{V}_i,\hat{\eta}_i)+\hat{U}_{3i}$, $\hat{m}_{3i}=\frac{\sum_{j=1}^{n}\tilde{m}_jL_{3,ij}}{\sum_{l=1}^{n}L_{3,il}}$, $\hat{e}_{3i}=\frac{\sum_{j=1}^{n}e_jL_{3,ij}}{\sum_{l=1}^{n}L_{3,il}}$ with $\hat{W}_{3i}=\frac{\sum_{j=1}^{n}W_jL_{3,ij}}{\sum_{l=1}^{n}L_{3,il}}$ and $\hat{U}_{3i}=\frac{\sum_{j=1}^{n}U_jL_{3,ij}}{\sum_{l=1}^{n}L_{3,il}}$ with $L_{3,ij}=L_{h_v,h_\eta}(\hat{V}_i-\hat{V}_j,\hat{\eta}_i-\hat{\eta}_j)$.

Hence from (5.A.5), we obtain:

$$\hat{\beta}-\beta_0=S^{-1}_{X-\hat{X}_3}\left(S_{X-\hat{X}_3,m-\hat{m}_3}+S_{X-\hat{X}_3,e-\hat{e}_3}\right). \tag{5.A.6}$$

We further decompose (5.A.5), as shown below:

$$\begin{aligned}Y_i-\hat{Y}_{1i}+\hat{Y}_{1i}-\hat{Y}_{3i} &= \left\{X_i-\hat{X}_{1i}+\hat{X}_{1i}-\hat{X}_{3i}\right\}'\beta_0+m_i-\hat{m}_{1i}+\hat{m}_{1i}-\hat{m}_{3i}\\ &+ e_i-\hat{e}_{1i}+\hat{e}_{1i}-\hat{e}_{3i}\\ Y_i-\hat{Y}_{1i}-\breve{\delta}_{y,i} &\equiv (X_i-\hat{X}_{1i}-\breve{\delta}_{x,i})'\beta_0+(m_i-\hat{m}_{1i}-\breve{\delta}_{m,i})+(e_i-\hat{e}_{1i}-\breve{\delta}_{e,i})\\ Y_i-\tilde{Y}_i+\tilde{Y}_i-\hat{Y}_{1i}-\breve{\delta}_{y,i} &\equiv (X_i-\tilde{X}_i+\tilde{X}_i-\hat{X}_{1i}-\breve{\delta}_{x,i})'\beta_0+(m_i-\tilde{m}_i+\tilde{m}_i-\hat{m}_{1i}-\breve{\delta}_{mi})\\ &+ (e_i-\hat{e}_{1i}-\breve{\delta}_{e,i}).\end{aligned} \tag{5.A.7}$$

The last term of the right-hand side in (5.A.7) is from $E(e|x,\eta)=0$, where $\breve{\delta}_{y,i}=\hat{Y}_{3i}-\hat{Y}_{1i}$, $\breve{\delta}_{x,i}=\hat{X}_{3i}-\hat{X}_{1i}$, $\breve{\delta}_{m,i}=\hat{m}_{3i}-\hat{m}_{1i}$, $\breve{\delta}_{e,i}=\hat{e}_{3i}-\hat{e}_{1i}$, $\hat{Y}_{1i}=\hat{m}_y(\hat{V}_i,\eta_i)+\hat{W}_{1i}$, $\hat{X}_{1i}=\hat{m}_x(\hat{V}_i,\eta_i)+\hat{U}_{1i}$, $\hat{m}_{1i}=\frac{\sum_{j=1}^{n}\tilde{m}_jL_{1,ij}}{\sum_{l=1}^{n}L_{1,il}}$, $\hat{e}_{1i}=\frac{\sum_{j=1}^{n}e_jL_{1,ij}}{\sum_{l=1}^{n}L_{il}}$, $\hat{W}_{1i}=\frac{\sum_{j=1}^{n}W_jL_{1,ij}}{\sum_{l=1}^{n}L_{1,il}}$ and $\hat{U}_{1i}=\frac{\sum_{j=1}^{n}U_jL_{1,ij}}{\sum_{l=1}^{n}L_{1,il}}$ with $L_{1,ij}=L_{h_v,h_\eta}(\hat{V}_i-\hat{V}_j,\eta_i-\eta_j)$. By utilising the decomposition in (5.A.7), we have:

$$\begin{aligned}&S_{X-\hat{X}_3}=S_{X-\tilde{X}}+S_{\tilde{X}-\hat{X}_1}+S_{\breve{\delta}_X}+2S_{X-\tilde{X},\tilde{X}-\hat{X}_1}-2S_{X-\tilde{X},\breve{\delta}_X}-2S_{\tilde{X}-\hat{X}_1,\breve{\delta}_X}\\ &= S_{m_x-\tilde{m}_x}+S_{\tilde{m}_x-\hat{m}_{x_1}}+S_U+S_{\hat{U}_1}+S_{\breve{\delta}_X}+2S_{m_x-\tilde{m}_x,\tilde{m}_x-\hat{m}_{x_1}}+2S_{m_x-\tilde{m}_x,U}-2S_{m_x-\tilde{m}_x,\hat{U}_1}\\ &- S_{m_x-\tilde{m}_x,\breve{\delta}_X}-2S_{\tilde{m}_x-\hat{m}_{x_1},U}-2S_{\tilde{m}_x-\hat{m}_{x_1},\hat{U}_1}-2S_{m_x-\tilde{m}_x,\breve{\delta}_X}-2S_{U\hat{U}_1}-2S_{U\breve{\delta}_X}+2S_{\hat{U}_1\breve{\delta}_X};\\ &S_{m-\hat{m}_3}=S_{m-\tilde{m}}+S_{\tilde{m}-\hat{m}_1}+S_{\breve{\delta}_m}+2S_{m-\tilde{m},\tilde{m}-\hat{m}_1}-2S_{m-\tilde{m},\breve{\delta}_m}-2S_{\tilde{m}-\hat{m}_1,\breve{\delta}_m};\\ &S_{e-\hat{e}_3}=S_e+S_{\hat{e}_1}+S_{\breve{\delta}_e}-2S_{e\hat{e}_1}-2S_{e\breve{\delta}_e}+2S_{\hat{e}_1\breve{\delta}_e};\\ &S_{X-\hat{X}_3,m-\hat{m}_3}=S_{m_x-\tilde{m}_x,m-\tilde{m}}+S_{m_x-\tilde{m}_x,\tilde{m}-\hat{m}_1}-S_{m_x-\tilde{m}_x,\breve{\delta}_m}+S_{\tilde{m}_x-\hat{m}_{x_1},m-\tilde{m}}+S_{\tilde{m}_x-\hat{m}_{x_1},\tilde{m}-\hat{m}_1}-S_{\tilde{m}_x-\hat{m}_{x_1},\breve{\delta}_m}\\ &+ S_{m-\tilde{m},U}+S_{\tilde{m}-\hat{m}_1,U}-S_{U\breve{\delta}_m}-S_{m-\tilde{m},\hat{U}_1}-S_{\tilde{m}-\hat{m}_1,\hat{U}_1}+S_{\hat{U}_1\breve{\delta}_m}-S_{m-\tilde{m},\breve{\delta}_X}-S_{\tilde{m}-\hat{m}_1,\breve{\delta}_X}+S_{\breve{\delta}_X\breve{\delta}_m};\\ &S_{X-\hat{X}_3,e-\hat{e}_3}=S_{m_x-\tilde{m}_x,e}-S_{m_x-\tilde{m}_x,\hat{e}_1}-S_{m_x-\tilde{m}_x,\breve{\delta}e}+S_{\tilde{m}_x-\hat{m}_{x_1},e}-S_{\tilde{m}_x-\hat{m}_{x_1},\hat{e}_1}-S_{\tilde{m}_x-\hat{m}_{x_1},\breve{\delta}_e}\\ &+ S_{Ue}-S_{U\hat{e}_1}-S_{U\breve{\delta}_e}-S_{\hat{U}_1e}+S_{\hat{U}_1\hat{e}_1}+S_{\hat{U}_1\breve{\delta}_e}-S_{\breve{\delta}_Xe}+S_{\breve{\delta}_X\hat{e}_1}+S_{\breve{\delta}X\breve{\delta}e};\\ &S_{m-\hat{m}_3,e-\hat{e}_3}=S_{m-\tilde{m},e}-S_{m-\tilde{m},\hat{e}_1}-S_{m-\tilde{m},\breve{\delta}e}+S_{\tilde{m}-\hat{m}_1,e}-S_{\tilde{m}-\hat{m}_1,\hat{e}_1}-S_{\tilde{m}-\hat{m}_1,\breve{\delta}_e}-S_{\breve{\delta}_me}+S_{\breve{\delta}_m\hat{e}_1}+S_{\breve{\delta}_e\breve{\delta}_m}.\end{aligned}$$

Note that we approximate two kernel functions to be $L_{3,ij} = L_{2,ij} + O_p(n^{-1/2}h_v^{-1})$ and $L_{1,ij} = L_{ij} + O_p(n^{-1/2}h_v^{-1})$ uniformly in i. Hence, we employ $L_{2,ij}$ and L_{ij} instead of $L_{3,ij}$ and $L_{1,ij}$, respectively, for the case of $\hat{\beta}$ in Propositions 5.A.1 to 5.A.15.

By Propositions 5.A.1 to 5.A.15, and (5.A.2) and (5.A.3), we obtain that (5.A.6) is:

$$
\begin{aligned}
(\hat{\beta}-\beta_0) &= \left(\frac{1}{n}\sum_{i=1}^{n} U_i U_i'\right)^{-} \left\{\frac{1}{n}\sum_{i=1}^{n} U_i e_i - \frac{1}{n}\sum_{i=1}^{n} U_i (m_i - \tilde{m}_i)\right\} + o_p(n^{-1/2}) \\
&= \left(\frac{1}{n}\sum_{i=1}^{n} U_i U_i'\right)^{-} \left\{\frac{1}{n}\sum_{i=1}^{n} U_i e_i - \frac{1}{n}\sum_{i=1}^{n} m_0^{(1)} U_i U_i' (\alpha_0 - \alpha)\right\} + o_p(n^{-1/2}) \\
&= \left(\frac{1}{n}\sum_{i=1}^{n} U_{0i} U_{0i}'\right)^{-} \left\{\frac{1}{n}\sum_{i=1}^{n} U_{0i} e_i - \frac{1}{n}\sum_{i=1}^{n} m_0^{(1)} U_{0i} U_{0i}' (\alpha_0 - \alpha)\right\} + o_p(1) + O_p(n^{-1/2}). \quad (5.A.8)
\end{aligned}
$$

Given β_0, $\alpha_0 - \alpha = \left(\frac{1}{n}\sum_{i=1}^{n}\left\{m_0^{(1)}\right\}^2 U_{0i}U_{0i}'\right)^{-} \frac{1}{n}\sum_{i=1}^{n} m_0^{(1)} e_i U_{0i}'$ (see the last equation in (5.A.4)). Hence we have:

$$
\begin{aligned}
(\hat{\beta}-\beta_0) &= \left(\frac{1}{n}\sum_{i=1}^{n} U_{0i} U_{0i}'\right)^{-} \left\{\frac{1}{n}\sum_{i=1}^{n} U_{0i} e_i - \frac{1}{n}\sum_{i=1}^{n} m_0^{(1)} U_{0i} U_{0i}' \left(\frac{1}{n}\sum_{i=1}^{n}\left\{m_0^{(1)}\right\}^2 U_{0i}U_{0i}'\right)^{-} \frac{1}{n}\sum_{i=1}^{n} m_0^{(1)} e_i U_{0i}'\right\} \\
&+ o_p(1) = (S_{U_0})^{-}\left\{S_{U_0 e} - m_0^{(1)} S_{U_0}\left(\left\{m_0^{(1)}\right\}^2 S_{U_0}\right)^{-} m_0^{(1)} S_{eU_0}\right\} + o_p(1).
\end{aligned}
$$

Given both $\hat{\beta}$ and $\hat{\alpha}$, the variance of e is:

$$
\begin{aligned}
\hat{\sigma}^2 &= S_{e-\hat{e}_3} + S_{m-\hat{m}_3} + (\hat{\beta}-\beta_0)' S_{(X-\hat{X}_3)} (\hat{\beta}-\beta_0) - 2(\hat{\beta}-\beta_0)' S_{(X-\hat{X}_3)', e-\hat{e}_3} \\
&- 2(\hat{\beta}-\beta_0)' S_{(X-\hat{X}_3), m-\hat{m}_3} + 2 S_{m-\hat{m}_3, e-\hat{e}_3} \qquad (5.A.9) \\
&= S_e + o_p(1) \xrightarrow{p} \sigma^2,
\end{aligned}
$$

by Propositions 5.A.1 to 5.A.15 below, the law of large numbers, and the *i.i.d.* assumption of e_i . The other nine terms are $(\hat{\beta}-\beta_0)' S_{m_x - \tilde{m}_x} (\hat{\beta}-\beta_0)$; $(\hat{\beta}-\beta_0)' S_{m_x - \tilde{m}_x, U} (\hat{\beta}-\beta_0)$; $(\hat{\beta}-\beta_0)' S_U (\hat{\beta}-\beta_0)$; $S_{m-\tilde{m}}$; $S_{m_x-\tilde{m}_x, m-\tilde{m}}$; $S_{m-\tilde{m},U}$; $S_{m_x-\tilde{m}_x,e}$; S_{Ue} and $S_{m-\tilde{m},e}$ equal to $o_p(n^{-1/2})$.

By the central limit theorem and the law of large numbers, the asymptotic normalities of $\hat{\alpha}$ and $\hat{\beta}$ are:

$$
\begin{aligned}
\sqrt{n}(\hat{\beta}-\beta_0) &= \sqrt{n}\,(S_{U_0})^{-}\left\{S_{U_0 e} - m_0^{(1)} S_{U_0}\left(\left\{m_0^{(1)}\right\}^2 S_{U_0}\right)^{-} m_0^{(1)} S_{eU_0}\right\} + o_p(1) \\
&\to_D N\left(0, \sigma^2\left[\Phi_{U_0}^{-} - \left(m_0^{(1)}\Phi_{U_0}\right)^{-}\Phi_{U_0}\left\{m_0^{(1)}\right\}^2\left(m_0^{(1)}\Phi_{U_0}\right)^{-}\right]\right) \\
\sqrt{n}(\hat{\alpha}-\alpha_0) &= \sqrt{n}\left(\left\{m_0^{(1)}\right\}^2 S_{U_0}\right)^{-}\left\{m_0^{(1)} S_{eU_0} - m_0^{(1)} S_{U_0}\,(S_{U_0})^{-} S_{eU_0}\right\} + o_p(1) \\
&\to_D N\left(0, \sigma^2\left[\left(\left\{m_0^{(1)}\right\}^2 \Phi_{U_0}\right)^{-} - \left\{m_0^{(1)}\Phi_{U_0}\right\}^{-}\Phi_{U_0}\left\{m_0^{(1)}\Phi_{U_0}\right\}^{-}\right]\right).
\end{aligned}
$$

■

Note that the stated orders of the remainder term $R_1(\alpha, h_v, h_\eta, h_z)$ are available uniformly in $\alpha \in A_n$ and $h_v, h_\eta, h_z \in \mathcal{H}_n$, using the uniform bounds in Härdle et al. (1993). Let $\varphi_n(\alpha, h_v, h_\eta, h_z)$ be a possible quantity for which we show that:

$$\sup_{\alpha\in A_n, h_v, h_\eta, h_z\in\mathcal{H}_n} |\varphi_n(\alpha, h_v, h_\eta, h_z)| = o_p(n^a), \tag{5.A.10}$$

since we have:

$$\sup_{\alpha\in A_n, h_v, h_\eta, h_z\in\mathcal{H}_n} E\left(\varphi_n(\alpha, h_v, h_\eta, h_z)/n^b\right)^{2l} = O(1), \tag{5.A.11}$$

for all integers $l \geq 1$ and where $b < a$. For details of the equations (5.A.10) and (5.A.11), see Step (ii) of the proof section 4 in Härdle et al. (1993). For proofs of Propositions 5.A.1 to 5.A.15, we assume that $h_{\eta,1} = \cdots = h_{\eta,q_2} = h_\eta$ and $h_{z,1} = \cdots = h_{z,q_2} = h_z$ for expositional simplicity.

Proposition 5.A.1.

(i) $\sqrt{n}S_{\tilde{m}_x-\hat{m}_x} = O_p(n^{-1/2}h_v^{-1}h_\eta^{-q_2}) + O_p(n^{1/2}(h_v^2+h_\eta^2)^2)$;

(ii) $\sqrt{n}S_{\tilde{m}-\hat{m}} = O_p(n^{-1/2}h_v^{-1}h_\eta^{-q_2}) + O_p(n^{1/2}(h_v^2+h_\eta^2)^2)$.

Proof: *Let $\varphi(\cdot)$ denote $m(\cdot)$ and $m_x(\cdot)$, and $\tilde{\varphi}(\cdot)$ denote $\tilde{m}(\cdot)$ and $\tilde{m}_x(\cdot)$. We deduce from (5.A.2) and (5.A.3) that, uniformly in i, we have:*

$$\begin{aligned}
\tilde{\varphi}_i - \hat{\varphi}_i &= \frac{\sum_{j\neq i}\left\{\varphi(X_i'\alpha_0, \eta_i|v,\eta) - \varphi(X_j'\alpha_0,\eta_j)\right\}L_{ij}}{\sum_{j\neq i}L_{ij}} \\
&= \frac{\sum_{j\neq i}\left\{\tilde{\varphi}_i - \tilde{\varphi}_j + U_j'(\alpha-\alpha_0)\varphi_0^{(1)}\right\}L_{ij}}{\sum_{j\neq i}L_{ij}} + O(n^{-1}) \\
&= \frac{(nh_vh_\eta^{q_2})^{-1}\sum_{j\neq i}\left\{\tilde{\varphi}_i - \tilde{\varphi}_j + U_j'(\alpha-\alpha_0)\varphi_0^{(1)}\right\}L_{ij}}{f(v,\eta)}\left[1 - \frac{\hat{f}(v,\eta)-f(v,\eta)}{\hat{f}(v,\eta)}\right] + o(1),
\end{aligned}$$

where $\varphi_0^{(1)} = \partial\varphi(v_0,\eta)/\partial v_0$. Note that since $\left(\hat{f}(v,\eta) - f(v,\eta)\right)$ is $O_p(nh_vh_\eta^{q_2})^{-1/2} + O_p(h_v^2+h_\eta^2)$ so $\left[1 - \frac{\hat{f}(v,\eta)-f(v,\eta)}{\hat{f}(v,\eta)}\right]$ can be dropped, hence we consider only the numerator terms in the rest of the section.

By identical distribution, $E(S_{\tilde{\varphi}-\hat{\varphi}}) = E\left\{(\tilde{\varphi}_i-\hat{\varphi}_i)^2\right\}$. We can easily obtain those $E(\tilde{\varphi}_i-\hat{\varphi}_i) = O(h_v^2 + h_\eta^2)$ and $Var(\tilde{\varphi}_i-\hat{\varphi}_i) = O(nh_vh_\eta^{q_2})^{-1}$, where:

$$Var(\tilde{\varphi}_i-\hat{\varphi}_i) = Var\left(\frac{1}{nh_vh_\eta^{q_2}}\sum_{j\neq i}(\tilde{\varphi}_i-\tilde{\varphi}_j)L_{ij}\right) + Var\left(\frac{1}{nh_vh_\eta^{q_2}}\sum_{j\neq i}U_j'(\alpha-\alpha_0)\varphi_0^{(1)}L_{ij}\right)$$

$$+2Cov\left(\frac{1}{nh_vh_\eta^{q_2}}\sum_{j\neq i}(\tilde{\varphi}_i-\tilde{\varphi}_j)L_{ij}, \frac{1}{nh_vh_\eta^{q_2}}\sum_{j\neq i}U_j'(\alpha-\alpha_0)\varphi_0^{(1)}L_{ij}\right),$$

$$Var\left(\frac{1}{nh_vh_\eta^{q_2}}\sum_{j\neq i}(\tilde{\varphi}_i-\tilde{\varphi}_j)L_{ij}\right) = O(nh_vh_\eta^{q_2})^{-1},$$

$$Var\left(\frac{1}{nh_vh_\eta^{q_2}}\sum_{j\neq i}U_j'(\alpha-\alpha_0)\varphi_0^{(1)}\right) = O(n^2h_vh_\eta^{q_2})^{-1},$$

$$Cov\left(\frac{1}{nh_vh_\eta^{q_2}}\sum_{j\neq i}(\tilde{\varphi}_i-\tilde{\varphi}_j)L_{ij}, \frac{1}{nh_vh_\eta^{q_2}}\sum_{j\neq i}U_j'(\alpha-\alpha_0)\varphi_0^{(1)}L_{ij}\right) = O(n^{-3/2}h_v^{-1}h_\eta^{-q_2}).$$

Hence $E(S_{\tilde{\varphi}-\hat{\varphi}}) = O(nh_vh_\eta^{q_2})^{-1} + O((h_v^2+h_\eta^2)^2)$. ■

Proposition 5.A.2.

$\sqrt{n}S_{\tilde{m}_x-\hat{m}_x,\tilde{m}-\hat{m}} = O_p(n^{-1/2}h_v^{-1}h_\eta^{-q_2}) + O_p(n^{1/2}(h_v^2+h_\eta^2)^2).$

Proof: *Proposition 5.A.1 (i) and (ii), and the Cauchy inequality provide the proof.* ■

Proposition 5.A.3.

(i) $\sqrt{n}S_{\hat{U}} = O_p(n^{-1/2}h_v^{-1}h_\eta^{-q_2})$;

(ii) $\sqrt{n}S_{\hat{e}} = O_p(n^{-1/2}h_v^{-1}h_\eta^{-q_2})$.

Proof: *Let* ϱ_i *denote* U_i *and* e_i*, and* $E(\varrho_i|\mathcal{L}) = 0$ *almost surely, where* $\mathcal{L} = (X,\eta)$*, hence* $E(S_{\hat{\varrho}}) = E(\hat{\varrho}_i^2)$*. Then we have:*

$$E(\hat{\varrho}_i)^2 = \frac{1}{n^2h_v^2h_\eta^{q_2}}E\left(\sum_{j\neq i}\varrho_j^2L_{ij}^2\right) = O(nh_vh_\eta^{q_2})^{-1}.$$

■

Proposition 5.A.4.

(i) $\sqrt{n}S_{\hat{\delta}_X} = O_p(n^{-1}h_z^{-q_z/2}h_v^{-1/2}h_\eta^{-(q_2+2)/2}) + O_p(n^{1/2}h_z^2(h_v^2+h_\eta^2))$;

(ii) $\sqrt{n}S_{\hat{\delta}_m} = O_p(n^{-1}h_z^{-q_z/2}h_v^{-1/2}h_\eta^{-(q_2+2)/2}) + O_p(n^{1/2}h_z^2(h_v^2+h_\eta^2))$;

(iii) $\sqrt{n}S_{\hat{\delta}_e} = O_p(n^{-1}h_z^{-q_z/2}h_v^{-1/2}h_\eta^{-(q_2+2)/2}) + O_p(n^{1/2}h_z^2(h_v^2+h_\eta^2))$.

Proof: *Let* δ *denote* δ_X*,* δ_m *and* δ_e*.Then we have:*

$$\hat{\delta}_i = \delta_{2,i} - \delta_{1,i} = \frac{\sum_{j\neq i}\delta_jL_{2,ij}}{\sum_{j\neq i}L_{2,ij}} - \frac{\sum_{j\neq i}\delta_jL_{ij}}{\sum_{j\neq i}L_{ij}}.$$

The Taylor expansion of the kernel function, $L_{2,ij}$*, is:*

$$L_{2,ij} = L_{ij} + L_{ij}^{(1)}\left(\frac{\triangle_{ij}}{h_\eta}\right) + L_{ij}^{(2)}(\tau)\left(\frac{\triangle_{ij}}{h_\eta}\right)^2,$$

where $L_{ij}^{(r)}$ *is the* r*th derivative of* L_{ij} *with respect to* η *with* $r = 1$ *or* 2*,* $\triangle_{ij} = \{\hat{m}_x(Z_j) - m_x(Z_j)\} - \{\hat{m}_x(Z_i) - m_x(Z_i)\}$ *and* τ *is between the segment line of* $\eta_j - \eta_i$ *and* $\hat{\eta}_j - \hat{\eta}_i$*. Hence, the denominator of* $\delta_{2,i}$ *is:*

$$\frac{1}{nh_vh_\eta^{q_2}}\sum_{j\neq i}L_{2,ij} = \frac{1}{nh_vh_\eta^{q_2}}\sum_{j\neq i}L_{ij} + \frac{1}{nh_vh_\eta^{q_2+1}}\sum_{j\neq i}L_{ij}^{(1)}\triangle_{ij} + R_{ij},$$

where R_{ij} *is the remainder term and the second term on the right-hand side is* $o_p(n^{-1/2})$*, because of the* *following:*

$$
\begin{aligned}
& E\left(\frac{1}{nh_v h_\eta^{q_2+1}}\sum_{j\neq i} L_{ij}^{(1)}\triangle_{ij}\right)^2 \\
&= \frac{1}{n^2h_v^2h_\eta^{2(q_2+1)}}E\left(\sum_{j\neq i}\left(L_{ij}^{(1)}\right)^2\triangle_{ij}^2\right) + \frac{2}{n^2h_v^2h_\eta^{2(q_2+1)}}E\left(\sum_{j\neq i}\sum_{k\neq i,j} L_{ij}^{(1)}L_{ik}^{(1)}\triangle_{ij}\triangle_{ik}\right) \\
&= \frac{1}{n^4h_z^{2q_z}h_v^2h_\eta^{2(q_2+1)}}E\left(\sum_{j\neq i}\left(L_{ij}^{(1)}\right)^2\left\{\sum_{l\neq j} C_{(l,j;K)} - \sum_{l\neq i} C_{(l,i;K)}\right\}^2\right) \\
&+ \frac{2}{n^4h_z^{2q_z}h_v^2h_\eta^{2(q_2+1)}}E\left(\sum_{j\neq i}\left(L_{ij}^{(1)}\right)^2\left\{\sum_{l\neq j} C_{(l,j;K)} - \sum_{l\neq i} C_{(l,i;K)}\right\}\left\{\sum_{m\neq j,l} C_{(m,j;K)} - \sum_{m\neq i,l} C_{(m,i;K)}\right\}\right) \\
&+ \frac{2}{n^4h_z^{2q_z}h_v^2h_\eta^{2(q_2+1)}}E\left(\sum_{j\neq i}\sum_{k\neq i,j} L_{ij}^{(1)}L_{ik}^{(1)}\left\{\sum_{l\neq j} C_{(l,j;K)} - \sum_{l\neq i} C_{(l,i;K)}\right\}\left\{\sum_{m\neq k,l} C_{(m,k;K)} - \sum_{m\neq i,l} C_{(m,i;K)}\right\}\right) \\
&= O\left(n^{-2}h_z^{-q_z}h_v^{-1}h_\eta^{-(q_2+2)}\right) + O\left(n^{-1}h_z^4h_v^{-1}h_\eta^{-(q_2+2)}\right) + O\left(h_z^4(h_v^2+h_\eta^2)^2\right),
\end{aligned}
$$

where $C_{(l,j;K)} = \{m_x(Z_l) - m_x(Z_j)\}K_{jl}$. *Hence* $\hat{\delta}_i = \frac{(nh_vh_\eta^{q_2+1})^{-1}\sum_{j\neq i}\delta_j L_{ij}^{(1)}\triangle_{ij}}{(nh_vh_\eta)^{-1}\sum_{j\neq i}L_{ij}+o_p(n^{-1/2})}$.

Now consider $E(\sqrt{n}S_{\hat{\delta}})$, *we have:*

$$
E(\sqrt{n}S_{\hat{\delta}}) = \frac{1}{n}\sum_{i=1}^{n}E(\hat{\delta}_i^2) + \frac{2}{n}\sum_{i=1}^{n}\sum_{j=1,\neq i}^{n}E(\hat{\delta}_i\hat{\delta}_j). \tag{5.A.12}
$$

Using a similar argument to the above, the two terms in the right-hand-side of (5.A.12) are:

$$
\begin{aligned}
& E\left(\hat{\delta}_i^2\right) = \frac{1}{n^4h_z^{2q_z}h_v^2h_\eta^{2(q_2+1)}}E\left(\sum_{j\neq i}\delta_j L_{ij}^{(1)}\left\{\sum_{l\neq j} C_{(l,j;K)} - \sum_{l\neq i} C_{(l,i;K)}\right\}\right)^2 \\
&= \frac{1}{n^4h_z^{2q_z}h_v^2h_\eta^{2(q_2+1)}}E\left(\sum_{j\neq i}\delta_j^2\left(L_{ij}^{(1)}\right)^2\left\{\sum_{l\neq j} C_{(l,j;K)} - \sum_{l\neq i} C_{(l,i;K)}\right\}^2\right) \\
&+ \frac{2}{n^4h_z^{2q_z}h_v^2h_\eta^{2(q_2+1)}}E\left(\sum_{j\neq i}\delta_j^2\left(L_{ij}^{(1)}\right)^2\left\{\sum_{l\neq j} C_{(l,j;K)} - \sum_{l\neq i} C_{(l,i;K)}\right\}\left\{\sum_{m\neq j,l} C_{(m,j;K)} - \sum_{m\neq i,l} C_{(m,i;K)}\right\}\right) \\
&= O\left(n^{-2}h_z^{-q_z}h_v^{-1}h_\eta^{-(q_2+2)}\right) + O\left(n^{-1}h_z^4h_v^{-1}h_\eta^{-(q_2+2)}\right),
\end{aligned}
$$

and

$$
\begin{aligned}
& E\left(\hat{\delta}_i\hat{\delta}_j\right) = \frac{2}{n^4h_z^{2q_z}h_v^2h_\eta^{2(q_2+1)}}\sum_{j\neq i}\sum_{k\neq i,j} \\
&\times E\left(\delta_j L_{ij}^{(1)}\delta_k L_{ik}^{(1)}\left\{\sum_{l\neq j} C_{(l,j;K)} - \sum_{l\neq i} C_{(l,i;K)}\right\}\left\{\sum_{m\neq k} C_{(m,k;K)} - \sum_{m\neq i,l} C_{(m,i;K)}\right\}\right) \\
&= O\left(h_z^4(h_v^2+h_\eta^2)^2\right).
\end{aligned}
$$

■

Proposition 5.A.5.

(i) $\sqrt{n}S_{\hat{\delta}_X\hat{\delta}_m} = O_p(n^{-1}h_z^{-q_z/2}h_v^{-1/2}h_\eta^{-(q_2+2)/2}) + O_p(n^{1/2}h_z^2(h_v^2 + h_\eta^2))$;

(ii) $\sqrt{n}S_{\hat{\delta}_X\hat{\delta}_e} = O_p(n^{-1}h_z^{-q_z/2}h_v^{-1/2}h_\eta^{-(q_2+2)/2}) + O_p(n^{1/2}h_z^2(h_v^2 + h_\eta^2))$;

(iii) $\sqrt{n}S_{\hat{\delta}_m\hat{\delta}_e} = O_p(n^{-1}h_z^{-q_z/2}h_v^{-1/2}h_\eta^{-(q_2+2)/2}) + O_p(n^{1/2}h_z^2(h_v^2 + h_\eta^2))$.

Proof: *Proposition 5.A.4 (i), (ii) and (iii), and the Cauchy inequality provide the proof.* ∎

Proposition 5.A.6.

(i) $\sqrt{n}S_{U\hat{U}} = O_p(n^{-1/2}h_v^{-1/2}h_\eta^{-q_2/2})$;

(ii) $\sqrt{n}S_{\hat{U}e} = O_p(n^{-1/2}h_v^{-1/2}h_\eta^{-q_2/2})$;

(iii) $\sqrt{n}S_{e\hat{e}} = O_p(n^{-1/2}h_v^{-1/2}h_\eta^{-q_2/2})$;

(iv) $\sqrt{n}S_{U\hat{e}} = O_p(n^{-1/2}h_v^{-1/2}h_\eta^{-q_2/2})$.

Proof: *Since $E(\varrho_i|\mathcal{L}) = 0$, we have:*

$$E\left(\sqrt{n}S_{\varrho\hat{\varrho}}\right)^2 = \frac{1}{n}\sum_{i=1}^{n} E\left(\varrho_i^2\hat{\varrho}_i^2\right),$$

where:

$$E(\varrho_i^2\hat{\varrho}_i^2) = \frac{1}{n^2h_v^2h_\eta^{2q_2}}E\left(\varrho_i^2\sum_{j\neq i}\varrho_j^2L_{ij}^2\right) = O(nh_vh_\eta^{q_2})^{-1}.$$

∎

Proposition 5.A.7.

(i) $\sqrt{n}S_{\tilde{m}_x-\hat{m}_x,U} = O_p(n^{-1/2}h_v^{-1/2}h_\eta^{-q_2/2})$;

(ii) $\sqrt{n}S_{\tilde{m}-\hat{m},U} = O_p(n^{-1/2}h_v^{-1/2}h_\eta^{-q_2/2})$;

(iii) $\sqrt{n}S_{\tilde{m}_x-\hat{m}_x,e} = O_p(n^{-1/2}h_v^{-1/2}h_\eta^{-q_2/2})$;

(iv) $\sqrt{n}S_{\tilde{m}-\hat{m},e} = O_p(n^{-1/2}h_v^{-1/2}h_\eta^{-q_2/2})$.

Proof: *Since $E(\varrho_i|\mathcal{L}) = 0$, we have:*

$$E(\sqrt{n}S_{\varrho,\tilde{\varphi}-\hat{\varphi}})^2 = \frac{1}{n}\sum_{i=1}^{n} E\left\{\varrho_i^2(\tilde{\varphi}_i - \hat{\varphi}_i)^2\right\},$$

where

$$\begin{aligned} E\left\{\varrho_i^2(\tilde{\varphi}_i - \hat{\varphi}_i)^2\right\} &= \frac{1}{n^2h_v^2h_\eta^{2q_2}}E\left(\varrho_i^2\sum_{j\neq i}\left(C^*_{(i,j;L)}\right)^2\right) + \frac{2}{n^2h_v^2h_\eta^{2q_2}}E\left(\varrho_i^2\sum_{j\neq i}\sum_{l\neq i,j}C^*_{(i,j;L)}C^*_{(i,l;L)}\right) \\ &= O(n^{-1}h_v^{-1}h_\eta^{-q_2}) + O((h_v^2 + q_2h_\eta^2)^2) \end{aligned}$$

with $C^*_{(i,j;L)} = \left\{\tilde{\varphi}_i - \tilde{\varphi}_j + U_j'(\alpha-\alpha_0)\varphi_0^{(1)}\right\}L_{ij}$. ∎

Proposition 5.A.8.

(i) $\sqrt{n}S_{U\hat{\delta}_X} = O_p(n^{-1}h_z^{-q_z/2}h_v^{-1/2}h_\eta^{-(q_2+2)/2}) + O_p(n^{-1/2}h_z^2h_v^{-1/2}h_\eta^{-(q_2+2)/2})$;

(ii) $\sqrt{n}S_{e\hat{\delta}_e} = O_p(n^{-1}h_z^{-q_z/2}h_v^{-1/2}h_\eta^{-(q_2+2)/2}) + O_p(n^{-1/2}h_z^2h_v^{-1/2}h_\eta^{-(q_2+2)/2})$;

(iii) $\sqrt{n}S_{e\hat{\delta}_m} = O_p(n^{-1}h_z^{-q_z/2}h_v^{-1/2}h_\eta^{-(q_2+2)/2}) + O_p(n^{-1/2}h_z^2h_v^{-1/2}h_\eta^{-(q_2+2)/2})$;

(iv) $\sqrt{n}S_{e\hat{\delta}_X} = O_p(n^{-1}h_z^{-q_z/2}h_v^{-1/2}h_\eta^{-(q_2+2)/2}) + O_p(n^{-1/2}h_z^2h_v^{-1/2}h_\eta^{-(q_2+2)/2})$;

(v) $\sqrt{n}S_{U\hat{\delta}_m} = O_p(n^{-1}h_z^{-q_z/2}h_v^{-1/2}h_\eta^{-(q_2+2)/2}) + O_p(n^{-1/2}h_z^2h_v^{-1/2}h_\eta^{-(q_2+2)/2})$;

(vi) $\sqrt{n}S_{U\hat{\delta}_e} = O_p(n^{-1}h_z^{-q_z/2}h_v^{-1/2}h_\eta^{-(q_2+2)/2}) + O_p(n^{-1/2}h_z^2h_v^{-1/2}h_\eta^{-(q_2+2)/2})$.

Proof: *Since $E(\varrho_i|\mathcal{L}) = 0$, we have*

$$E(\sqrt{n}S_{\varrho\hat{\delta}})^2 = \frac{1}{n}\sum_{i=1}^{n} E\left(\varrho_i^2\hat{\delta}_i^2\right),$$

where

$$\begin{aligned}
E(\varrho_i^2\hat{\delta}_i^2) = {} & \frac{1}{n^4h_z^{2q_z}h_v^2h_\eta^{2(q_2+1)}}E\left(\varrho_i^2\sum_{j\neq i}\delta_j^2\left(L_{ij}^{(1)}\right)^2\left\{\sum_{l\neq j}C_{(l,j;K)} - \sum_{l\neq i}C_{(l,i;K)}\right\}^2\right) \\
& + \frac{2}{n^4h_z^{2q_z}h_v^2h_\eta^{2(q_2+1)}}E\left(\varrho_i^2\sum_{j\neq i}\delta_j^2\left(L_{ij}^{(1)}\right)^2\left\{\sum_{l\neq j}C_{(l,j;K)} - \sum_{l\neq i}C_{(l,i;K)}\right\}\right. \\
& \left.\times\left\{\sum_{m\neq j,l}C_{(m,j;K)} - \sum_{m\neq i,l}C_{(m,i;K)}\right\}\right) \\
= {} & O\left(n^{-2}h_z^{-q_z}h_v^{-1}h_\eta^{-(q_2+2)}\right) + O\left(n^{-1}h_z^4h_v^{-1}h_\eta^{-(q_2+2)}\right),
\end{aligned}$$

using similar arguments to those in Proposition 5.A.4. ■

Proposition 5.A.9.

(i) $\sqrt{n}S_{\tilde{m}_x-\hat{m}_x,\hat{U}} = O_p(nh_vh_\eta^{q_2})^{-1} + O_p(n^{1/2}(h_v^2+h_\eta^2)^2)$;

(ii) $\sqrt{n}S_{\tilde{m}-\hat{m},\hat{U}} = O_p(nh_vh_\eta^{q_2})^{-1} + +O_p(n^{1/2}(h_v^2+h_\eta^2)^2)$;

(iii) $\sqrt{n}S_{\tilde{m}_x-\hat{m}_x,\hat{e}} = O_p(nh_vh_\eta^{q_2})^{-1} + O_p(n^{1/2}(h_v^2+h_\eta^2)^2)$;

(iv) $\sqrt{n}S_{\tilde{m}-\hat{m},\hat{e}} = O_p(nh_vh_\eta^{q_2})^{-1} + O_p(n^{1/2}(h_v^2+h_\eta^2)^2)$.

Proof*:*

$$E\left(\sqrt{n}S_{\tilde{\varphi}-\hat{\varphi},\hat{\varrho}}\right)^2 = \frac{1}{n}\sum_{i=1}^{n}E\left\{\hat{\varrho}_i^2(\tilde{\varphi}_i-\hat{\varphi}_i)^2\right\} + \frac{2}{n}\sum_{i=1}^{n}\sum_{j=1,\neq i}^{n}E\left\{\hat{\varrho}_i\hat{\varrho}_j(\tilde{\varphi}_i-\hat{\varphi}_i)(\tilde{\varphi}_j-\hat{\varphi}_j)\right\},$$

where

$$E\left\{\hat{\varrho}_i^2(\tilde{\varphi}_i-\hat{\varphi}_i)^2\right\}=\frac{1}{n^4h_v^4h_\eta^{4q_2}}E\left(\sum_{j\neq i}\sum_{l\neq i}\varrho_j^2L_{ij}^2\left(C^*_{(i,l:L)}\right)^2\right)$$
$$+\ \frac{2}{n^4h_v^4h_\eta^{4q_2}}E\left(\sum_{j\neq i}\sum_{l\neq i}\sum_{k\neq i,l}\varrho_j^2L_{ij}^2C^*_{(i,l;L)}C^*_{(i,k,L)}\right)$$
$$=\ O(n^{-2}h_v^{-2}h_\eta^{-2q_2})+O(n^{-1}h_v^{-1}h_\eta^{-q_2}(h_v^2+h_\eta^2)^2),$$
$$E\left\{\hat{\varrho}_i\hat{\varrho}_j(\tilde{\varphi}_i-\hat{\varphi}_i)(\tilde{\varphi}_j-\hat{\varphi}_j)\right\}=\frac{1}{n^4h_v^4h_\eta^{4q_2}}E\left(\sum_{l\neq i}\sum_{s\neq j}\sum_{k\neq i}\sum_{m\neq j}\varrho_l\varrho_sL_{il}L_{js}C^*_{(i,k;L)}C^*_{(j,m;L)}\right)$$
$$=\ O((h_v^2+h_\eta^2)^4).$$

■

Proposition 5.A.10.

(i) $\sqrt{n}S_{\hat{U}\hat{\delta}_X}=O_p\left(n^{-1}h_z^{-q_z/2}h_v^{-1}h_\eta^{-(q_2+1)}\right)+O_p\left(n^{1/2}h_z^2(h_v^2+h_\eta^2)\right)$;

(ii) $\sqrt{n}S_{\hat{e}\hat{\delta}_e}=O_p\left(n^{-1}h_z^{-q_z/2}h_v^{-1}h_\eta^{-(q_2+1)}\right)+O_p\left(n^{1/2}h_z^2(h_v^2+h_\eta^2)\right)$;

(iii) $\sqrt{n}S_{\hat{U}\hat{\delta}_m}=O_p\left(n^{-1}h_z^{-q_z/2}h_v^{-1}h_\eta^{-(q_2+1)}\right)+O_p\left(n^{1/2}h_z^2(h_v^2+h_\eta^2)\right)$;

(iv) $\sqrt{n}S_{\hat{e}\hat{\delta}_X}=O_p\left(n^{-1}h_z^{-q_z/2}h_v^{-1}h_\eta^{-(q_2+1)}\right)+O_p\left(n^{1/2}h_z^2(h_v^2+h_\eta^2)\right)$;

(v) $\sqrt{n}S_{\hat{e}\hat{\delta}_m}=O_p\left(n^{-1}h_z^{-q_z/2}h_v^{-1}h_\eta^{-(q_2+1)}\right)+O_p\left(n^{1/2}h_z^2(h_v^2+h_\eta^2)\right)$;

(vi) $\sqrt{n}S_{\hat{U}\hat{\delta}_e}=O_p\left(n^{-1}h_z^{-q_z/2}h_v^{-1}h_\eta^{-(q_2+1)}\right)+O_p\left(n^{1/2}h_z^2(h_v^2+h_\eta^2)\right)$.

Proof:

$$E\left(\sqrt{n}S_{\hat{\varrho}\hat{\delta}}\right)^2=\frac{1}{n}\sum_{i=1}^{n}E\left(\hat{\varrho}_i^2\hat{\delta}_i^2\right)+\frac{2}{n}\sum_{i=1}^{n}\sum_{j=1,\neq i}^{n}E\left(\hat{\varrho}_i\hat{\delta}_i\hat{\varrho}_j\hat{\delta}_j\right),$$

where

$$E\left(\hat{\varrho}_i^2\hat{\delta}_i^2\right)=\frac{1}{n^6h_z^{2q_z}h_v^4h_\eta^{2(2q_2+1)}}E\left(\sum_{j\neq i}\sum_{l\neq i}\varrho_jL_{ij}\delta_l\left(L_{il}^{(1)}\right)\left\{\sum_{k\neq j}C_{(k,j;K)}-\sum_{k\neq i}C_{(k,i;K)}\right\}\right)^2$$
$$=\frac{1}{n^6h_z^{2q_2}h_v^4h_\eta^{2(2q_2+1)}}E\left(\sum_{j\neq i}\sum_{l\neq i}\varrho_j^2L_{ij}^2\delta_l^2\left\{L_{il}^{(1)}\right\}^2\left\{\sum_{k\neq l}C_{(k,l;K)}-\sum_{k\neq i}C_{(k,i;K)}\right\}^2\right)$$
$$+\frac{2}{n^6h_z^{2q_z}h_v^4h_\eta^{2(2q_2+1)}}E\left(\sum_{j\neq i}\sum_{l\neq i}\varrho_j^2L_{ij}^2\delta_l^2\left(L_{il}^{(1)}\right)^2\left\{\sum_{k\neq l}C_{(k,l;K)}-\sum_{k\neq i}C_{(k,i,;K)}\right\}\right.$$
$$\left.\times\left\{\sum_{m\neq l,k}C_{(m,l;K)}-\sum_{m\neq i,k}C_{(m,i;K)}\right\}\right)=O\left(n^{-2}h_z^{-q_z}h_v^{-2}h_\eta^{-(2q_2+2)}\right)+O\left(n^{-2}h_z^4h_v^{-2}h_\eta^{(-2q_2+2)}\right),$$

and the cross product term, $E\left(\hat{\varrho}_i\hat{\delta}_i\hat{\varrho}_j\hat{\delta}_j\right)$, *is* $(n^6h_z^{2q_z}h_v^4h_\eta^{2(2q_2+1)})^{-1}$ *times:*

$$E\left(\sum_{j\neq i}\sum_{s\neq j}\sum_{l\neq i}\sum_{t\neq j}\varrho_j\varrho_sL_{ij}L_{js}\delta_l\delta_tL_{il}^{(1)}L_{jt}^{(1)}\left\{\sum_{k\neq l}C_{(k,l;K)}-\sum_{k\neq i}C_{(k,i;K)}\right\}\left\{\sum_{m\neq t}C_{(m,t;K)}-\sum_{m\neq j}C_{(m,j;K)}\right\}\right).$$

Hence the cross product term is $O\left(h_z^4(h_v^2+h_\eta^2)^2\right)$.

■

Proposition 5.A.11.

(i) $\sqrt{n}S_{\tilde{m}_x-\hat{m}_x,\hat{\delta}_X}=O_p\left(n^{-1}h_z^{-q_z/2}h_v^{-1}h_\eta^{-(q_2+1)}\right)+O_p\left(n^{1/2}h_z^2(h_v^2+h_\eta^2)^2\right)$;

(ii) $\sqrt{n}S_{\tilde{m}_x-\hat{m}_x,\hat{\delta}_m}=O_p\left(n^{-1}h_z^{-q_z/2}h_v^{-1}h_\eta^{-(q_2+1)}\right)+O_p\left(n^{1/2}h_z^2(h_v^2+h_\eta^2)^2\right)$;

(iii) $\sqrt{n}S_{\tilde{m}_x-\hat{m}_x,\hat{\delta}_e}=O_p\left(n^{-1}h_z^{-q_z/2}h_v^{-1}h_\eta^{-(q_2+1)}\right)+O_p\left(n^{1/2}h_z^2(h_v^2+h_\eta^2)^2\right)$;

(iv) $\sqrt{n}S_{\tilde{m}-\hat{m},\hat{\delta}_m}=O_p\left(n^{-1}h_z^{-q_z/2}h_v^{-1}h_\eta^{-(q_2+1)}\right)+O_p\left(n^{1/2}h_z^2(h_v^2+h_\eta^2)^2\right)$;

(v) $\sqrt{n}S_{\tilde{m}-\hat{m},\hat{\delta}_X}=O_p\left(n^{-1}h_z^{-q_z/2}h_v^{-1}h_\eta^{-(q_2+1)}\right)+O_p\left(n^{1/2}h_z^2(h_v^2+h_\eta^2)^2\right)$;

(vi) $\sqrt{n}S_{\tilde{m}-\hat{m},\hat{\delta}_e}=O_p\left(n^{-1}h_z^{-q_z/2}h_v^{-1}h_\eta^{-(q_2+1)}\right)+O_p\left(n^{1/2}h_z^2(h_v^2+h_\eta^2)^2\right)$.

Proof:

$$E\left(\sqrt{n}S_{\tilde{\varphi}-\hat{\varphi},\hat{\delta}}\right)^2=\frac{1}{n}\sum_{i=1}^{n}E\left((\tilde{\varphi}_i-\hat{\varphi}_i)^2\hat{\delta}_i^2\right)+\frac{2}{n}\sum_{i=1}^{n}\sum_{j=1,\neq i}^{n}E\left((\tilde{\varphi}_i-\hat{\varphi}_i)\hat{\delta}_i(\tilde{\varphi}_j-\hat{\varphi}_j)\hat{\delta}_j\right),$$

where:

$$\begin{aligned}&E\left((\tilde{\varphi}_i-\hat{\varphi}_i)^2\hat{\delta}_i\right)^2\\&=\frac{1}{n^6h_z^{2q_z}h_v^4h_\eta^{2(2q_2+1)}}E\left(\sum_{j\neq i}\left(C^*_{(i,j;L)}\right)^2\sum_{l\neq j}\delta_l^2\left(L_{il}^{(1)}\right)^2\left\{\sum_{k\neq l}C_{(k,l:K)}-\sum_{k\neq i}C_{(k,i;K)}\right\}^2\right)\\&+\frac{2}{n^6h_z^{2q_z}h_v^4h_\eta^{2(2q_2+1)}}E\left(\sum_{j\neq i}\left(C^*_{(i,j;L)}\right)^2\sum_{l\neq j}\delta_l^2\left(L_{il}^{(1)}\right)^2\left\{\sum_{k\neq l}C_{(k,l:K)}-\sum_{k\neq i}C_{(k,i;K)}\right\}\right.\\&\left.\times\left\{\sum_{m\neq l,k}C_{(m,l:K)}-\sum_{m\neq i,k}C_{(m,i;K)}\right\}\right)=O\left(n^{-2}h_z^{-q_z}h_v^{-2}h_\eta^{-(2q_2+2)}\right)+O\left(n^{-2}h_z^4h_v^{-2}h_\eta^{-(2q_2+2)}\right),\end{aligned}$$

and the cross product term, $E\left(\hat{\varrho}_i\hat{\delta}_i\hat{\varrho}_j\hat{\delta}_j\right)$ *is* $(n^6h_z^{2q_z}h_v^2h_\eta^{2(2q_2+1)})^{-1}$ *times:*

$$E\left(\sum_{j\neq i}\sum_{ks\neq j}\sum_{l\neq i}\sum_{t\neq j}C^*_{(i,j;L)}C^*_{(j,s;L)}\delta_l\delta_tL_{il}^{(1)}L_{jt}^{(1)}\left\{\sum_{k\neq l}C_{(l,k:K)}-\sum_{k\neq i}C_{(k,i;K)}\right\}\left\{\sum_{m\neq t}C_{(m,t:K)}-\sum_{m\neq j}C_{(m,j;K)}\right\}\right).$$

Hence the cross product term is $O\left(h_z^4(h_v^2+h_\eta^2)^4\right)$. ■

Proposition 5.A.12.

$\sqrt{n}S_{\hat{U}\hat{e}}=O_p(nh_vh_\eta^{q_2})^{-1}+O_p(n^{1/2}(h_v^2+h_\eta^2)^2)$.

Proof:

$$E\left(\sqrt{n}S_{\hat{U}\hat{e}}\right)^2=\frac{1}{n}\sum_{i=1}^{n}E\left\{\hat{U}_i^2\hat{e}_i^2\right\}+\frac{2}{n}\sum_{i=1}^{n}\sum_{j=1,\neq i}^{n}E\left\{\hat{U}_i\hat{U}_j'\hat{e}_i\hat{e}_j\right\},$$

where:

$$E\left\{\hat{U}_i^2\hat{e}^2\right\} = \frac{1}{n^4h_v^4h_\eta^{4q_2}}E\left\{U_jU_j'\sum_{j\neq i}L_{ij}^2e_l^2\sum_{l\neq i}L_{il}^2\right\} = O(n^{-2}h_v^{-2}h_\eta^{-2q_2}),$$

and:

$$E\left\{\hat{U}_i\hat{U}_j'\hat{e}_i\hat{e}_j\right\} = \frac{1}{n^4h^4h_\eta^{4q_2}}E\left\{U_lU_l'\sum_{l\neq i}\sum_{l\neq j}L_{il}L_{jl}e_l^2\sum_{k\neq i}\sum_{k\neq j}L_{ik}L_{jk}\right\} = O((h_v^2+h_\eta^2)^4).$$

■

Proposition 5.A.13.

(i) $\sqrt{n}S_{m_x-\tilde{m}_x,\tilde{m}_x-\hat{m}_x} = O_p\left(n^{-1}h_v^{-1/2}h_\eta^{-q_2/2}\right)$;

(ii) $\sqrt{n}S_{m-\tilde{m},\tilde{m}-\hat{m}} = O_p\left(n^{-1}h_v^{-1/2}h_\eta^{-q_2/2}\right)$;

(iii) $\sqrt{n}S_{m_x-\tilde{m}_x,\tilde{m}-\hat{m}} = O_p\left(n^{-1}h_v^{-1/2}h_\eta^{-q_2/2}\right)$;

(iv) $\sqrt{n}S_{m-\tilde{m},\tilde{m}_x-\hat{m}_x} = O_p\left(n^{-1}h_v^{-1/2}h_\eta^{-q_2/2}\right)$.

Proof: *By (5.A.2) and (5.A.3) we deduce that, uniformly in i, we have:*

$$\varphi_i - \tilde{\varphi}_i = U_i'(\alpha_0-\alpha)\varphi_0^{(1)}(X_i'\alpha_0,\eta_i) + O(n^{-1}). \tag{5.A.13}$$

Hence we have:

$$(\varphi_i-\tilde{\varphi}_i)(\tilde{\varphi}_i-\hat{\varphi}_i) = \frac{1}{nh_vh_\eta^{q_2}}\sum_{j\neq i}t_i\left\{\tilde{\varphi}_i-\tilde{\varphi}_j+U_j'(\alpha-\alpha_0)\varphi_0^{(1)}\right\}L_{ij},$$

where $t_i = U_i'(\alpha_0-\alpha)\varphi_0^{(1)}$.

For the rest of proofs, we use similar arguments to those in Proposition 5.A.7 because $E(U_i|\mathcal{L}) = 0$. Hence we have:

$$E\left(\sqrt{n}S_{\varphi-\tilde{\varphi},\tilde{\varphi}-\hat{\varphi}}\right)^2 = \frac{1}{n}\sum_{i=1}^{n}E\left(t_i^2(\tilde{\varphi}_i-\hat{\varphi}_i)^2\right),$$

where:

$$\begin{aligned} E\left(t_i^2(\tilde{\varphi}_i-\hat{\varphi}_i)^2\right) &= \frac{1}{n^2h_v^2h_\eta^2}E\left\{\sum_{j\neq i}t_i^2\left(C^*_{(i,j;L)}\right)^2L_{ij}^2\right\} + \frac{2}{n^2h_v^2h_\eta^2}E\left\{\sum_{j\neq i}\sum_{l\neq i,j}t_i^2C^*_{(i,j;L)}L_{ij}C^*_{(i,l;L)}L_{il}\right\} \\ &= O\left(n^{-2}h_v^{-1}h_\eta^{-q_2}\right) + O\left(n^{-1}(h_v^2+h_\eta^2)^2\right). \end{aligned}$$

■

Proposition 5.A.14.

(i) $\sqrt{n}S_{m_x-\tilde{m}_x,\hat{U}} = O_p\left(n^{-1}h_v^{-1/2}h_\eta^{-q_2/2}\right)$;

(ii) $\sqrt{n}S_{m_x-\tilde{m}_x,\hat{e}} = O_p\left(n^{-1}h_v^{-1/2}h_\eta^{-q_2/2}\right)$;

(iii) $\sqrt{n}S_{m-\tilde{m},\hat{U}} = O_p\left(n^{-1}h_v^{-1/2}h_\eta^{-q_2/2}\right)$;

(iv) $\sqrt{n}S_{m-\tilde{m},\hat{e}} = O_p\left(n^{-1}h_v^{-1/2}h_\eta^{-q_2/2}\right)$.

Proof: *By (5.A.13) and $E(\varrho_i|\mathcal{L}) = 0$, we use similar arguments to those in Proposition 5.A.6 for the rest of the proofs.*

$$E(\sqrt{n}S_{\varphi-\tilde{\varphi},\hat{\varrho}})^2 = \frac{1}{n}\sum_{i=1}^{n} E\left(t_i^2\hat{\varrho}_i^2\right),$$

where:

$$E\left(t_i^2\hat{\varrho}_i^2\right) = \frac{1}{n^2h_v^2h_\eta^{2q_2}}E\left(\sum_{j\neq i}t_i^2\varrho_j^2L_{ij}^2\right) = O\left(n^{-2}h_v^{-1}h_\eta^{-q_2}\right).$$

■

Proposition 5.A.15.

(i) $\sqrt{n}S_{m_x-\tilde{m}_x,\hat{\delta}_X} = O_p\left(n^{-2}h_z^{-q_z/2}h_v^{-1/2}h_\eta^{-(q_2+2)/2}\right)$;

(ii) $\sqrt{n}S_{m-\tilde{m},\hat{\delta}_m} = O_p\left(n^{-2}h_z^{-q_z/2}h_v^{-1/2}h_\eta^{-(q_2+2)/2}\right)$;

(iii) $\sqrt{n}S_{m-\tilde{m},\hat{\delta}_X} = O_p\left(n^{-2}h_z^{-q_z/2}h_v^{-1/2}h_\eta^{-(q_2+2)/2}\right)$;

(iv) $\sqrt{n}S_{m_x-\tilde{m}_x,\hat{\delta}_m} = O_p\left(n^{-2}h_z^{-q_z/2}h_v^{-1/2}h_\eta^{-(q_2+2)/2}\right)$;

(v) $\sqrt{n}S_{m_x-\tilde{m}_x,\hat{\delta}_e} = O_p\left(n^{-2}h_z^{-q_z/2}h_v^{-1/2}h_\eta^{-(q_2+2)/2}\right)$;

(vi) $\sqrt{n}S_{m-\tilde{m},\hat{\delta}_e} = O_p\left(n^{-2}h_z^{-q_z/2}h_v^{-1/2}h_\eta^{-(q_2+2)/2}\right)$.

Proof: *By (5.A.13) and $E(U_i|\mathcal{L}) = 0$, the rest of proofs is similar to that of Proposition 5.A.8.*

$$E(\sqrt{n}S_{\varphi-\tilde{\varphi},\hat{\delta}})^2 = \frac{1}{n}\sum_{i=1}^{n} E\left(t_i^2\hat{\delta}_i^2\right),$$

where

$$\begin{aligned}
E\left(t_i^2\hat{\delta}_i^2\right) &= \frac{1}{n^6h_z^{2q_z}h_v^2h_\eta^{2(q_2+1)}}E\left(\sum_{j\neq i}t_i^2\delta_j^2\left\{L_{ij}^{(1)}\right\}^2\left\{\sum_{l\neq j}C_{(l,j;K)} - \sum_{l\neq i}C_{(l,i;K)}\right\}^2\right) \\
&+ \frac{2}{n^6h_z^{2q_z}h_v^2h_\eta^{2(q_2+1)}}E\left(\sum_{j\neq i}t_i^2\delta_j^2\left(L_{ij}^{(1)}\right)^2\left\{\sum_{l\neq j}C_{(l,j;K)} - \sum_{l\neq i}C_{(l,i;K)}\right\}\right. \\
&\times \left.\left\{\sum_{k\neq j,l}C_{(k,j;K)} - \sum_{k\neq i,l}C_{(k,i;K)}\right\}\right) = O(n^{-4}h_z^{-q_z}h_v^{-1}h_\eta^{-(q_2+2)}) + O(n^{-4}h_z^4h_v^{-1}h_\eta^{-(q_2+2)}).
\end{aligned}$$

■

5.2. Proof of Theorem 2.2

Given $\hat{\beta}$ and $\hat{\alpha}$, we have:

$$\begin{aligned}
\hat{m}(\hat{v},\hat{\eta}_i) - m(v_0,\eta_i) \quad &= \quad \left\{\hat{m}_{y**}(\hat{v},\eta_i) - \widetilde{m}_{y**}(\hat{v},\eta_i) + \widetilde{m}_{y**}(\hat{v},\eta_i) - m_{y**}(v_0,\eta_i) + \breve{\delta}_{y**,i}\right\} \\
&- \quad \left\{\hat{m}_x(\hat{v},\eta_i) - \widetilde{m}_x(\hat{v},\eta_i) + \widetilde{m}_x(\hat{v},\eta_i) - m_x(v_0,\eta_i) + \breve{\delta}_{x,i}\right\}'(\hat{\beta} - \beta_0), \quad (5.B.1)
\end{aligned}$$

where $Y_i^{**} = Y_i - X_i'\beta_0$, $\breve{\delta}_{y^{**}} = \breve{\delta}_y - \breve{\delta}_x'\beta_0$ and $\widetilde{m}(\hat{v},\eta) = E(m|\hat{\alpha})$. As the results of Section 5.1, the second term in the right-hand side of (5.B.1) is $o_p(n^{-1/2})$, uniformly in i, by applying (5.A.10) and (5.A.11) as $\sup_{X_i,\eta_i\in\mathcal{A},Z_i\in\mathcal{A}_z} |\varphi_i| = o_p(n^a)$ since $\sup_{X_i,\eta_i\in\mathcal{A},Z_i\in\mathcal{A}_z} E|\varphi_i/n^b|^{2l} = O(1)$. Hence (5.B.1) is:

$$\hat{m}(\hat{v},\hat{\eta}_i) - m(v_0,\eta_i) = \left\{\hat{m}_{y**}(\hat{v},\eta_i) - \widetilde{m}_{y**}(\hat{v},\eta_i) + \widetilde{m}_{y**}(\hat{v},\eta_i) - m_{y**}(v_0,\eta_i) + \breve{\delta}_{y**,i}\right\} + o_p(1), \quad (5.B.2)$$

where $\breve{\delta}_{y^{**}} = o_p(n^{-1/2})$ by similar arguments to those in Proposition 5.A.4 and $\widetilde{m}_{y^{**}}(\hat{v},\eta_i) - m_{y^{**}}(v_0,\eta_i) = O_p(n^{-1/2})$ by (5.A.2) and (5.A.3), uniformly in i. Hence (5.B.2) is:

$$\begin{aligned} \hat{m}(\hat{v},\hat{\eta}_i) - m(v_0,\eta_i) &= \hat{m}_{y^{**}}(\hat{v},\eta_i) - \widetilde{m}_{y^{**}}(\hat{v},\eta_i) + o_p(1) \\ &\equiv \hat{m}(\hat{v},\eta_i) - \widetilde{m}(\hat{v},\eta_i) + o_p(1), \end{aligned} \quad (5.B.3)$$

where:

$$\begin{aligned} \hat{m}(\hat{v},\eta_i) - \widetilde{m}(\hat{v},\eta_i) &= \frac{\sum_{j\neq i}\{m(v_0,\eta_j) - \widetilde{m}(\hat{v},\eta_i)\}L_{1,ij}}{\sum_{j\neq i}L_{1,ij}} \\ &= \frac{\sum_{j\neq i}\{m(v_0,\eta_j) - m(v_0,\eta_i)\}L_{1,ij}}{\sum_{j\neq i}L_{1,ij}} + U_i'(\hat{\alpha}-\alpha_0)m_0^{(1)} + O(n^{-1}) \\ &= \frac{\sum_{j\neq i}\{m(v_0,\eta_j) - m(v_0,\eta_i)\}\{L_{0,ij} + O(n^{-1/2}h_v^{-1})\}}{\sum_{j\neq i}L_{0,ij} + o(1)} + O_p(n^{-1/2}). \end{aligned}$$

Hence (5.B.3) is:

$$\hat{m}(\hat{v},\hat{\eta}_i) - m(v_0,\eta_i) = \hat{m}(v_0,\eta_i) - m(v_0,\eta_i) + o_p(1). \quad (5.B.4)$$

Let us define $\check{m}(v_0,\eta_i) = \hat{m}(v_0,\eta_i)\hat{f}(v_0,\eta_i)$. Then we can rewrite the term in the right-hand side of (5.B.4) as follows:

$$\begin{aligned} \hat{m}(v_0,\eta_i) - m(v_0,\eta_i) &= \frac{\check{m}(v_0,\eta_i) - m(v_0,\eta_i)\hat{f}(v_0,\eta_i)}{\hat{f}(v_0,\eta_i)} \\ &= \frac{\check{m}(v_0,\eta_i) - m(v_0,\eta_i)\hat{f}(v_0,\eta_i)}{f(v_0,\eta_i)}\left[1 - \frac{\hat{f}(v_0,\eta_i) - f(v_0,\eta_i)}{f(v_0,\eta_i)}\right]. \end{aligned} \quad (5.B.5)$$

First, we consider the bias term $E(\hat{m}(v_0,\eta_i) - m(v_0,\eta_i)) = f^{-1}(v_0,\eta_i)\left(E\check{m}(v_0,\eta_i) - m(v_0,\eta_i)E(\hat{f}(v_0,\eta_i))\right)$, where

$$\begin{aligned} E\check{m}(v_0,\eta_i) &= E\left[\frac{1}{nh_vh_\eta^{q_2}}\sum_{j=1}^n K_v\left(\frac{V_{0,j}-v_0}{h_v}\right)K_\eta\left(\frac{\eta_j-\eta_i}{h_\eta}\right)Y_j^{**}\right] \\ &= E\left[E_{v_0,\eta_i}\left\{\frac{1}{nh_vh_\eta^{q_2}}\sum_{j=1}^n K_v\left(\frac{V_{0,j}-v_0}{h_v}\right)K_\eta\left(\frac{\eta_j-\eta_i}{h_\eta}\right)Y_j^{**}\right\}\right] \\ &= E\left[\frac{1}{nh_vh_\eta^{q_2}}\sum_{j=1}^n K_v\left(\frac{V_{0,j}-v_0}{h_v}\right)K_\eta\left(\frac{\eta_j-\eta_i}{h_\eta}\right)m(V_{0,j},\eta_j)\right] \\ &= f(v_0,\eta_i)m(v_0,\eta_i) + \mathcal{K}_{v,2}h_v^2\left\{f_v^{(1)}(v_0,\eta_i)m_0^{(1)}(v_0) + f(v_0,\eta_i)m_0^{(2)}(v_0)\right\} \\ &+ \mathcal{K}_{\eta,2}\sum_{s=1}^{q_2}h_{\eta,s}^2\left\{f_{\eta,s}^{(1)}(v_0,\eta_i)m^{(1)}(\eta_{s,i}) + f(v_0,\eta_i)m_{\eta,s}^{(2)}(\eta_i)\right\} + O(h_v^3) + O\left(\sum_s h_{\eta,s}^3\right). \end{aligned}$$

In the expression above, E_{v_0,η_i} is the expectation conditional on v_0 and η_i. Hence

$$E(\hat{m}(v_0,\eta_i) - m(v_0,\eta_i)) = \left\{ h_v^2 B_v(v_0,\eta_i) + \sum_{s=1}^{q_2} h_{\eta,s}^2 B_{\eta,s}(v_0,\eta_i) \right\} + o(1). \tag{5.B.6}$$

The single sum of (5.B.6) converges to its population mean by the Chebyshev's law of large numbers (see Linton and Härdle (1996)).

Now let us consider the variance term. Note that $f(v_0,\eta_i) = f(v_0,\eta) + O_p(n^{1/2})$ and $m(v_0,\eta_i) = m(v_0,\eta) + O_p(n^{-1/2})$ by the law of large numbers since both functions satisfy the bounded moment conditions. Hence we have:

$$\begin{aligned}
V\left(\frac{1}{n}\sum_{i=1}^{n}\hat{m}(v_0,\eta_i)\right) &= f(v,\eta)^{-2} V\left(\frac{1}{n}\sum_{i=1}^{n}\left\{\check{m}(v_0,\eta_i) - m(v_0,\eta_i)\hat{f}(v_0,\eta_i)\right\}\right) \\
&= f(v_0,\eta)^{-2} V\left(\frac{1}{n}\sum_{i=1}^{n}\check{m}(v_0,\eta_i)\right) + f(v_0,\eta)^{-2} m(v_0,\eta)^2 V\left(\frac{1}{n}\sum_{i=1}^{n}\hat{f}(v_0,\eta_i)\right) \\
&- f(v_0,\eta)^{-2} 2m(v_0,\eta) Cov\left(\frac{1}{n}\sum_{i=1}^{n}\check{m}(v_0,\eta_i), \frac{1}{n}\sum_{i=1}^{n}\hat{f}(v_0,\eta)\right),
\end{aligned}$$

where $V(\cdot)$ and $Cov(\cdot)$ denote variance and covariance, respectively, and:

$$\begin{aligned}
& V\left(\frac{1}{n}\sum_{i=1}^{n}\check{m}(v_0,\eta_i)\right) = E\left(V_{v_0,\eta_i}\left\{\frac{1}{n}\sum_{i=1}^{n}\check{m}(v_0,\eta_i)\right\}\right) + V\left(E_{v_0,\eta_i}\left\{\frac{1}{n}\sum_{i=1}^{n}\check{m}(v_0,\eta_i)\right\}\right) \\
&= \sigma^2 f(\eta)^2 E\left[\frac{1}{nh_v^q}\sum_{j=1}^{n} K_v\left(\frac{V_{j,0} - v_0}{h_v}\right)\right]^2 + f(\eta)^2 V\left[\frac{1}{nh_v^q}\sum_{j=1}^{n} K_v\left(\frac{V_{0,j} - v_0}{h_v}\right) m(V_{0,j},\eta_j)\right] \\
&= \frac{\sigma^2 f(\eta)^2}{nh_v^q}\mathcal{K}_v + \frac{m(v_0,\eta)^2 f(\eta)^2 f(v_0)}{nh_v^q}\mathcal{K}_v + O(n^{-1}), \\
& V\left(\frac{1}{n}\sum_{i=1}^{n}\hat{f}(v_0,\eta_i)\right) = \frac{f(\eta)^2 f(v)\mathcal{K}_v}{nh_v^q} + O(n^{-1}) \\
& \text{Cov}\left(\frac{1}{n}\sum_{i=1}^{n}\check{m}(v_0,\eta_i), \frac{1}{n}\sum_{i=1}^{n}\hat{f}(v_0,\eta_i)\right) = E\left\{\frac{1}{n}\sum_{i=1}^{n}\check{m}(v_0,\eta_i)\frac{1}{n}\sum_{i=1}^{n}\hat{f}(v_0,\eta_i)\right\} \\
&- E\left\{\frac{1}{n}\sum_{i=1}^{n}\check{m}(v_0,\eta_i)\right\} E\left\{\frac{1}{n}\sum_{i=1}^{n}\hat{f}(v_0,\eta_i)\right\} = \frac{m(v_0,\eta) f(\eta)^2 f(v)\mathcal{K}_v}{nh_v^q} + O(n^{-1}),
\end{aligned}$$

where V_{v_0,η_i} denotes the variance conditional on v_0 and η_i. Hence we have:

$$\sqrt{nh_v^q}(\hat{m}(\hat{v}) - m(v_0) - bias) \to_D N(0, var).$$

The consistency of $\hat{g}(\hat{v})$ and its asymptotic normality are argued in the same way as above, since $m(v_0) = g(v_0) + c_1$. ■

References

Ai, C., Chen, X., 2003. Efficient estimation of models with conditional moment restrictions containing unknown functions. Econometrica 71 (6), 1795–1843.

Amemiya, T., 1974. The non-linear two-stage least-squares estimator. Journal of Econometrics 2, 105–110.

Amemiya, T., 1985. Advanced Econometrics. Harvard University Press, Cambridge, MA.

Blundell, R., 1988. Consumer behaviour: theory and empirical evidence. Economic Journal 98 (389), 16–65.

Blundell, R., Browning, M., Crawford, I., 2003. Nonparametric Engel curves and revealed preference. Econometrica 71 (1), 205–240.

Blundell, R., Chen, X., Kristensen, D., 2007. Semi-nonparametric IV estimation of shape-invariant Engel curves. Econometrica 75 (6), 1613–1669.

Blundell, R., Duncan, A., Pendakur, K., 1998. Semiparametric estimation and consumer demand. Journal of Applied Econometrics 13 (5), 435–461.

Blundell, R., Powell, J., 2003. Endogeneity in Nonparametric and Semiparametric Regression Models: Advances in Economics and Econometrics, Eighth World Congress. Vol. 2. Cambridge University Press, Cambridge.

Blundell, R., Powell, J., 2004. Endogeneity in semiparametric binary response models. Review of Economics Studies 71 (3), 655–679.

Byrd, R., Lu, P., Nocedal, J., Zhu, C., 1995. A limited memory algorithm for bound constrained optimization. SIAM Journal on Scientific Computing 16 (5), 1190–1208.

Carroll, R., Fan, J., Gijbels, I., Wand, M., 1997. Generalized partially linear single-index models. Journal of the American Statistical Association 92 (438), 477–489.

Deaton, A., Muellbauer, J., 1980. An almost ideal demand system. The American economic review (70), 312–326.

Eubank, R., Speckman, P., 1993. Confidence bands in nonparametric regression. Journal of the American Statistical Association 88 (424), 1287–1301.

Fan, J., Gijbels, I., 1996. Local Polynomial Modeling and its Applications. Chapman & Hall, London.

Gao, J., 2007. Nonlinear Time Series. Vol. 108 of Monographs on Statistics and Applied Probability. Chapman & Hall/CRC, Boca.

Gao, J., Lu, Z., Tjøstheim, D., 2006. Estimation in semiparametric spatial regression. The Annals of Statistics 34 (3), 1395–1435.

Gorman, W., 1959. Separable utility and aggregation. Econometrica, 469–481.

Hansen, B., 2008. Uniform convergence rates for kernel estimation with dependent data. Econometric Theory 24 (3), 726–748.

Härdle, W., Hall, P., Ichimura, H., 1993. Optimal smoothing in single-index models. The Annals of Statistics 21 (1), 157–178.

Härdle, W., Marron, J., 1990. Semiparametric comparison of regression curves. The Annals of Statistics 18 (1), 63–89.

Hastie, T., Tibishirani, R., 1991. Generalized Additive Models. Chapman and Hall.

Horowitz, J., 2009. Semiparametric and Nonparametric Methods in Econometrics. Springer Science and Business Media, New York.

Horowitz, J., Spokoiny, V., 2001. An adaptive, rate-optimal test of a parametric mean-regression model against a nonparametric alternative. Econometrica 69 (3), 599–631.

Li, Q., Racine, J., 2007. Nonparametric Econometrics: Theory and Practice. Princeton University Press, Princeton.

Linton, A., Härdle, W., 1996. Estimation of additive regression models with known links. Biometrika 83, 529–540.

Linton, A., Nielsen, J., 1995. A kernel method of estimating structural nonparametric regression based on marginal integration. Biometrika 82 (1), 93–100.

Nakamura, A., Nakamura, M., 1998. Model specification and endogeneity. Journal of Econometrics 83 (1-2), 213–237.

Newey, W., Powell, J., Vella, F., 1999. Nonparametric estimation of triangular simultaneous equations models. Econometrica 67 (3), 565–603.

Pendakur, P., 1999. Semiparametric estimates and tests of base-independant equivalence scales. Journal of Econometrics 88, 1–40.

Pinkse, C., Robinson, P., 1995. Pooling nonparametric estimates of regression functions with a similar shape. Advances in Econometrics and Quantitative Economics, eds. G. Maddala, P. Phillips and TN Srinvinsan, 172–195.

Pinkse, J., 2000. Nonparametric two-step regression estimation when regressors and error are dependent. The Canadian Journal of Statistics 28 (2), 289–300.

Robinson, P., 1988. Root-N-consistent semiparametric regression. Econometrica 56 (4), 931–954.

Speckman, P., 1988. Kernel smoothing in partial linear models. Journal of the Royal Statistical Society. Series B (Methodological) 50 (3), 413–436.

Su, L., Ullah, A., 2008. Local polynomial estimation of nonparametric simultaneous equations models. Journal of Econometrics 144 (1), 193–218.

Tjøstheim, D., Austad, R., 1996. Nonparametric identification of nonlinear time series: projection. Journal of American Statistical Associtation 89 (428), 1398–1409.

Xia, Y., 1998. Biased-corrected confidence bands in nonparametric regression. Journal of the Royal Statistical Society, Series B 60 (4), 797–811.

Xia, Y., Härdle, W., 2006. Semi-parametric estimation of partially linear singel-index models. Journal of Multivariate Analysis 97, 1162–1184.

Xia, Y., Tong, H., Li, W., 1999. On extended partially linear single-index models. Biometrika 86 (4), 831–842.

Yatchew, A., 2003. Semiparametric Regression for the Applied Econometrician. Cambridge University Press, Cambridge.