



City Research Online

City St George's, University of London

Citation: Gao, J., Kim, N. & Wongsart, P. (2020). On endogeneity and shape invariance in extended partially linear single index models. *Econometric Reviews*, 39(4), pp. 415-435. doi: 10.1080/07474938.2019.1682313

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/33937/>

Link to published version: <https://doi.org/10.1080/07474938.2019.1682313>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

1 **On Endogeneity and Shape Invariance in Extended Partially Linear**
2 **Single Index Models**

3 Jiti Gao¹
4 Monash University, Australia

5 Namhyun Kim ²
6 University of Exeter Business School, United Kingdom

7 Patrick W. Saart
8 Cardiff Business School, United Kingdom

9 **Abstract**

10 This paper elaborates the usefulness of the extended generalized partially linear single-index (EG-
11 PLSI) model introduced by Xia et al. (1999) in its ability to model a flexible shape-invariant
12 specification. More importantly, a control function approach is proposed to address endogeneity
13 in the EGPLSI model to enhance its applicability to empirical studies. Furthermore, it is shown
14 that the attractive asymptotic features of the single-index type of a semiparametric model are still
15 valid given intrinsic generated covariates. Our proposed method is then illustrated by applying to
16 address the endogeneity of expenditure in the semiparametric analysis of empirical Engel curves
17 with the British data.

18 *JEL Classification:* C14, C18, C51.

19 **Keyword:** Extended generalized partially linear single-index, control function approach, endo-
20 geneity, semiparametric regression models.

¹This author acknowledges the Australian Research council Discovery Grants Program for its support under grant numbers:DP150101012 & DP170104421.

²Namhyun, Kim, University of Exeter Business School, UK, email: n.kim@exeter.ac.uk
The authors would like to thank the editor and two anonymous referees for their kind and thoughtful suggestions and comments on the paper. All remaining errors are our own.

21 **1. Introduction**

22 Xia et al. (1999) introduced the extended generalized partially linear single-index
23 (EGPLSI) model of the form

$$Y_i = X_i' \beta_0 + g(X_i' \alpha_0) + \epsilon_i, \quad (1.1)$$

24 where (i) (X, Y) is a set of $\mathbb{R}^q \times \mathbb{R}$ -valued observable random vectors; (ii) β_0 and α_0
25 are unknown parameters vectors such that $\beta_0 \perp \alpha_0$ with $\|\alpha_0\| = 1$; (iii) $g(\cdot)$ is an
26 unknown link function such that $g(\cdot) : \mathbb{R} \rightarrow \mathbb{R}$ and $g''(\cdot) \neq 0$; and (iv) $E(\epsilon|X) = 0$
27 suggesting that $E(\epsilon|V_0) = 0$ with $V_0 = X' \alpha_0$. In fact, the EGPLSI model is the
28 extended version of the generalized partially linear single-index (GPLSI) model of
29 Carroll et al. (1997) and Xia and Härdle (2006) and hence a number of non- and
30 semiparametric models are special cases of the EGPLSI model. More importantly,
31 the EGPLSI model is useful for modelling a flexible shape-invariant specification in
32 pooling nonparametric regression curves (see Härdle and Marron (1990), and Robin-
33 son and Pinkse (1995) for examples) to model an aggregate structural relationship
34 incorporating the individual heterogeneity (see Blundell and Stoker (2007) for ex-
35 amples). The EGPLSI model allows this type of shape-invariant specification with a
36 functional flexibility because both scale and shift parameters can be incorporated in
37 the model. Therefore, the paper aims to address endogeneity in the EGPLSI model
38 causing an identification problem, to enhance its applicability to empirical studies.

39 Recently, a number of methods have been discussed in the literature on how
40 endogeneity can be best addressed in non- and semiparametric models. Among
41 these, two of the most popular alternatives are the nonparametric instrumental
42 variable estimation (NPIV) and the control function (CF) approach (see Blundell
43 and Powell (2003) for an excellent review). The NPIV approach relies on different
44 stochastic assumptions to the CF one and there are a few well-known difficulties that
45 are intrinsic to the NPIV, particularly the so-called ill-posed inverse problem (see
46 Ai and Chen (2003), and Blundell et al. (2007) for details). On the other hand, the
47 CF approach alternatively allows the specification of endogeneity, which is based on
48 an intuitive triangular structure of a model (see Blundell et al. (1998), and Blundell

49 and Powell (2003) for details).

50 This paper particularly aims to develop the CF approach. Although the gener-
51 ated covariates issue is intrinsic in the development of the CF approach, similar to
52 the study of Mammen et al. (2016), the proposed method maintains the attractive
53 features of the single-index (SI) model with relatively mild conditions in the litera-
54 ture and shows an accessible extension to strictly stationary and α -mixing process.
55 In a SI model, Härdle et al. (1993) showed that the optimal bandwidth for estimating
56 a link function can be used for the \sqrt{n} -consistent estimation of the index coefficients.
57 The current paper shows that this attractive feature is still valid with the CF ap-
58 proach and under-smoothing for estimating a first-stage reduced-form equation is
59 not required in order to archive \sqrt{n} -consistency. These results are developed in de-
60 tails with the simplest data structure, namely IID random sample, then extended to
61 a strictly stationary and α -mixing case. Furthermore, the convenient applicability
62 of our proposed CF approach is explored by analyzing the empirical Engel curves
63 based on the British data.

64 The structure of the rest of the paper is as follows. In Section 2, the usefulness
65 of the EGPLSI model for modelling a flexible shape-invariant specification is elabo-
66 rated. In addition, the development of the CF approach in the EGPLSI model and
67 a Monte Carlo exercise assessing the finite-sample performances of the proposed es-
68 timators are also presented. In Section 3, the implementation of the empirical study
69 of the cross sectional relationships between specific goods and the level of total ex-
70 penditure are investigated. Finally, Section 4 concludes the paper with a summary
71 of the main findings and the further issues to be investigated. All mathematical
72 proofs of the main theoretical results are presented in the supplemental document.

73 **2. EGPLSI Model, Shape-Invariant Specification and Endogeneity**

74 In this section, the usefulness of the EGPLSI model introduced by Xia et al.
75 (1999) is elaborated for specifying a flexible shape-invariant specification. This
76 section then introduces endogeneity into the EGPLSI model, establishes the CF
77 approach to address endogeneity and presents the asymptotic properties and finite

78 sample performances from a Monte Carlo simulation exercise for the estimators.

79 *2.1. Shape-Invariant Specification within EGPLSI Model Framework*

80 Let us discuss a flexible shape-invariant specification within the EGPLSI model
 81 framework by considering the two sets of observations. The first set of observations,
 82 $(X_1, Y_1), \dots, (X_n, Y_n)$, is assumed to follow the data generating process shown below

$$Y_i = m_1(X_i) + \varepsilon_i, \quad i = 1, \dots, n, \quad (2.1)$$

83 where ε is assumed to be independent with mean 0 and the common variance σ^2 .
 84 Suppose the second set of observations, $(X'_1, Y'_1), \dots, (X'_n, Y'_n)$, is from the following
 85 nonparametric regression model

$$Y'_i = m_2(X'_i) + \varepsilon'_i, \quad (2.2)$$

86 where ε' is independent from ε , but otherwise has the same stochastic structure as
 87 ε and has the common variance σ'^2 . The main interest here is to model the curves
 88 whose parametric nature is modelled by ³

$$m_2(X') = S_{\theta_0}^{-1}(m_1(T_{\theta_0}^{-1}(X'))), \quad (2.3)$$

89 where T_θ and S_θ are invertible transformations, particularly scalings and shifts of
 90 the axes indexed by parameters $\theta \in \Theta \subseteq R^d$, and θ_0 is the vector of true values of
 91 the parameters. A good estimate of θ_0 is provided by θ for which the curve $m_1(X)$
 92 is closely approximated by

$$m(X, \theta) = S_\theta(m_2(T_\theta(X))). \quad (2.4)$$

93 For the sake of illustration, the simple models are considered as follows

$$m_1(X) = (X - 0.4)^2 \quad \text{and} \quad m_2(X') = (X' - 0.5)^2 - 0.2, \quad (2.5)$$

94 which fit in the framework described by (2.1) to (2.4) by defining the following

$$\begin{aligned} T_\theta(X) &= \theta^{(1)}X + \theta^{(2)} \\ m_2(T_\theta(X)) &= (\theta^{(1)}X + \theta^{(2)} - 0.5)^2 - 0.2 \\ S_\theta(m_2(T_\theta(X))) &= (\theta^{(1)}X + \theta^{(2)} - 0.5)^2 - 0.2 + \theta^{(3)}X + \theta^{(4)}, \end{aligned}$$

³The case of (2.3) is available on the request from the author

95 where $\theta_0 = (\theta_0^{(1)}, \theta_0^{(2)}, \theta_0^{(3)}, \theta_0^{(4)}) = (1, 0.1, 0, 0.2)$.

96 When a curve comparison problem with a similar parametric nature to (2.3) is
 97 considered, Härdle and Marron (1990) suggested an estimation procedure by which
 98 separated kernel smoothers are used in order to compute the estimates of $m_1(\cdot)$ and
 99 $m_2(\cdot)$. The estimator of θ_0 is then found by minimizing a L^2 -norm objective function
 100 of kernel estimates of $m_1(\cdot)$ and $m_2(\cdot)$, and the approximation in (2.4). Alternatively,
 101 pooling the two sets of observations is more desirable. Modelling the data within
 102 the EGPLSI model framework enables this type of pooling nonparametric regression.
 103 The shift and scaling of the axes illustrated in the example above fit in the EGPLSI
 104 framework, shown below

$$m_3(X_1, X_2) = [\beta_{01}X_1 + \beta_{02}X_2] + \{([\alpha_{01}X_1 + \alpha_{02}X_2] - 0.5)^2 - 0.2\}, \quad (2.6)$$

105 where $X_1 = \begin{cases} X \\ X' \end{cases}$ and $X_2 = \begin{cases} 1 & \text{if } X_1 = X \\ 0 & \text{if } X_1 = X' \end{cases}$. The model examples in (2.5) can
 106 be obtained by defining

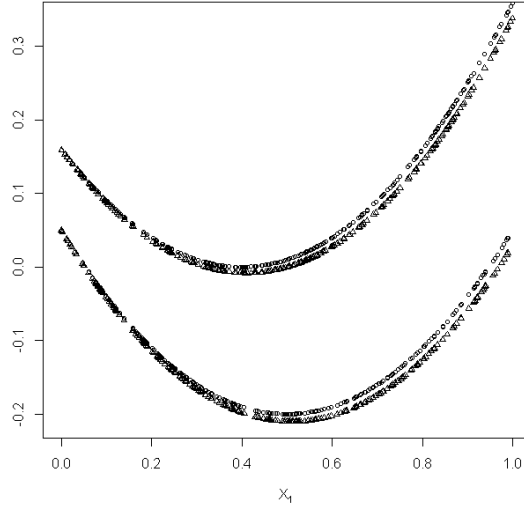
$$(\beta_{01}, \beta_{02}, \alpha_{01}, \alpha_{02}) = (0, 0.2, 1, 0.1). \quad (2.7)$$

107 Five hundred simulated observations of the model are represented by circles
 108 in Figure 2.1, where X_{1i} on the x -axis is a uniform random variable on $[0, 1]$ for
 109 $i = 1, \dots, 500$. The two sets of observations are determined by X_2 , which is a
 110 Bernoulli random variable with the parameter $p = 0.5$. It should be noted, however,
 111 that the set of values of the parameters in (2.7) do not satisfy the identification
 112 conditions which require that $\beta_0 \perp \alpha_0$ with $\|\alpha_0\| = 1$. An approximate model that
 113 satisfies these identification conditions is obtained by first setting $\beta_{02} = 0.2$ and
 114 $\alpha_{02} = 0.1$, so that $\beta_{01} = -0.02$ and $\alpha_{01} = 0.99$ can be derived. Five hundred
 115 simulated observations of this type of a model are represented by triangles in Figure
 116 2.1. In practice, when there is enough reason to believe (perhaps based on economic
 117 theory) that $\beta_{01} = 0$ and $\alpha_{01} = 1$, then such a model can be obtained by scaling and
 118 shifting, respectively, as follows

$$X_2 = v_{01} - \beta_{01}X_1 \text{ and } X_1 + \frac{\alpha_{02}}{\alpha_{01}}X_2 = \frac{v_{02}}{\alpha_{01}},$$

119 where $\beta_{01}X_1 + \beta_{02}X_2 = v_{01}$ and $\alpha_{01}X_1 + \alpha_{02}X_2 = v_{02}$. This method is illustrated in
 120 the empirical analysis in Section 3.

Figure 2.1. 500-simulated observations based on $m_3(\cdot, \cdot)$.



121
 122 *2.2. Endogeneity and Newly Proposed Estimation Procedure*

123 Despite its ability to model a flexible shape-invariant specification, the applica-
 124 bility of EGPLSI model to an empirical study is limited because of its shortfalls in
 125 addressing endogeneity. There are two potential sources of endogeneity in the model,
 126 namely endogeneity in the parametric and in the nonparametric components. If it
 127 is present, endogeneity in the parametric component is relatively easy to deal with.⁴
 128 Hence, to simplify the argument, the parametric covariates are assumed to belong to
 129 a subset $X_1 \subseteq \mathbb{R}^{q_1}$, for $q_1 < q$, of X such that $E(\epsilon|X_1) = 0$, namely the parametric
 130 covariates are exogenous, without loss of generality. In this case, endogeneity in the
 131 nonparametric component exists when $E(\epsilon|X) \neq 0$, which implies that $E(\epsilon|V_0) \neq 0$.
 132 An unanticipated property from the SI type of semiparametric models is that es-
 133 timators of the index coefficients are still \sqrt{n} -consistent even with the presence of
 134 endogeneity because of the partialling-out process in estimating the index coeffi-
 135 cients (see Ichimura (1993), Härdle et al. (1993), and Xia and Härdle (2006) for

⁴A comprehensive discussion on the presence of endogeneity in the parametric component can be found in Li and Racine (2007).

136 details). Nonetheless, the link function in the EGPLSI model is unidentifiable by
 137 using the conditional expectation relationship in the presence of endogeneity.

138 In the following, let us present the development of the CF approach in the
 139 EGPLSI model. For the sake of the notational simplicity, the simplest case is con-
 140 sidered, namely the presence of an endogenous nonparametric covariate denoted by
 141 X_2 .⁵ Hereafter, let Z denote a vector of valid instruments for X_2 as follows

$$X_{2i} = g_x(Z_i) + \eta_i, \quad (2.8)$$

142 where $E(\eta|Z) = 0$, and $E(\epsilon|X_2) = E(\epsilon|Z, \eta) = E(\epsilon|\eta) \equiv \iota(\eta)$ with (X_2, Z) is a
 143 set of $\mathbb{R} \times \mathbb{R}^{q_z}$ -valued observable random vectors, and $g_x(Z)$ and $\iota(\eta)$ are unknown
 144 real functions such that $g_x(\cdot) : \mathbb{R}^{q_z} \rightarrow \mathbb{R}$ and $\iota(\cdot) : \mathbb{R} \rightarrow \mathbb{R}$, respectively. The
 145 above stochastic assumption on ϵ is standard in the CF literature suggesting the
 146 exogeneity condition of Z , particularly $E(\epsilon|Z, \eta) = E(\epsilon|\eta)$ (see Newey et al. (1999),
 147 Blundell and Powell (2004), and Su and Ullah (2008) for examples). Furthermore,
 148 the necessary identification condition for the link function as discussed in Newey
 149 et al. (1999) is non-existence of a linear functional relationship between X_2 and η .

150 By imposing the structure of (2.8), the EGPLSI model in (1.1) in the presence
 151 of endogeneity is rewritten as

$$Y_i = X_i' \beta_0 + m(V_{0i}, \eta_i) + e_i, \quad (2.9)$$

152 where $m(v_0, \eta) \equiv g(v_0) + \iota(\eta)$ with $\iota(\eta) \neq 0$ being the endogeneity control func-
 153 tion, and $E(e|X) = 0$. The conditional expectation relationship, based on (2.9), is
 154 obtained as follows

$$m_y(v_0, \eta) \equiv m(v_0, \eta) + m_x(v_0, \eta)' \beta_0, \quad (2.10)$$

155 where $m_y(v_0, \eta) \equiv E(y|V_0, \eta)$ and $m_x(v_0, \eta) \equiv E(x|V_0, \eta)$.

156 In the following, the performance of the CF approach in the EGPLSI model
 157 based on (2.8) to (2.10) is discussed. The identification issue is first presented as

⁵The generalized version, namely more than one endogenous nonparametric covariates, is avail-
 able by a request to the author.

158 follows. Given α and β , let

$$J(\alpha, \beta) = E [Y - E(Y|V, \eta) - \{X - E(X|V, \eta)\}'\beta]^2$$

159

$$\mathcal{V} = E(\{X - E(X|V, \eta)\}\{X - E(X|V, \eta)\}'); \mathcal{W} = E(\{X - E(X|V, \eta)\}\{Y - E(Y|V, \eta)\}),$$

160 where $V = X'\alpha$. Suppose that $g(\cdot)$ is twice differentiable and that X has a positive
 161 density function on a union of a finite number of open convex subset in \mathbb{R}^q . The min-
 162 imum point of $J(\alpha, \beta)$ with $\alpha \perp \beta$ is then unique at α_0 and $\beta_{\alpha_0} = \{\mathcal{V}(\alpha_0)\}^+\mathcal{W}(\alpha_0)$,
 163 where $\{\mathcal{V}(\alpha_0)\}^+$ is the Moore-Penrose inverse.

164 Before we discuss the optimization procedure, the necessary notation is defined
 165 for the sake of convenience. We assume that the random sample $\{(X'_i, Z'_i, Y_i); i =$
 166 $1, \dots, n\}$ is IID. Let $f_x(x)$ and $f_z(z)$ denote the joint density functions of X' and
 167 Z' , respectively. Let us also denote $f_\alpha(v)$ as the density function of $V = X'\alpha$. We
 168 assume that $\mathcal{A}_j \subset \mathbb{R}^k$ is the union of a finite number of open sets such that $f_j(s) > C$
 169 on \mathcal{A}_j , where $k = q$ or q_z and $j = x$ or z for some constant $C > 0$. Hereafter, this
 170 region is considered to avoid the boundary points. Because the region is not known
 171 in practice, Xia and Härdle (2006) suggested using the weight function such that
 172 $I_n(s) = 1$ if $\frac{1}{n} \sum_{i=1}^n K_{j,i}(s) > C$ and 0 otherwise, where K_j is a corresponding kernel
 173 function. In this paper, $I_n(s)$ is omitted for the notational simplicity. In addition,
 174 C, C' and C'' denote generic constants varying from one place to another.

175 The conditional expectations, namely $E(Y|V, \eta)$ and $E(X|V, \eta)$, are then esti-
 176 mated with the leave-one-out nonparametric estimation as follows

$$\hat{E}_i(Y_i|V_i, \eta_i) = \frac{\sum_{j \neq i} L_{h_v h_\eta}(V_j - V_i, \eta_j - \eta_i) Y_j}{\sum_{j \neq i} L_{h_v h_\eta}(V_j - V_i, \eta_j - \eta_i)} \quad (2.11)$$

177

$$\hat{E}_i(X_i|V_i, \eta_i) = \frac{\sum_{j \neq i} L_{h_v h_\eta}(V_j - V_i, \eta_j - \eta_i) X_j}{\sum_{j \neq i} L_{h_v h_\eta}(V_j - V_i, \eta_j - \eta_i)}, \quad (2.12)$$

178 where $L_{h_v h_\eta}$ is a product kernel function constructed from the product of univariate
 179 kernel functions of $k_{h_v}(\cdot) \times k_{h_\eta}(\cdot)$ with the relevant bandwidth parameters, h_v and h_η .
 180 Furthermore, the first stage leave-one-out nonparametric estimation of the reduced
 181 equation in (2.8) used to estimate η_i is as follows

$$\hat{\eta}_i = X_i - \hat{g}_{x,i}(Z_i), \quad (2.13)$$

182 where $\hat{g}_{x,i}(Z_i) = \frac{\sum_{j \neq i} K_{h_z}(Z_j - Z_i) X_j}{\sum_{j \neq i} K_{h_z}(Z_j - Z_i)}$ with $K_{h_z}(\cdot)$ being the product kernel function
 183 constructed from $k_{h_{z_1}}(\cdot) \times \cdots \times k_{h_{z_{q_z}}}(\cdot)$, and h_{z_j} , for $j = 1, \dots, q_z$, is the relevant
 184 bandwidth parameter. The LS estimates of the unknown parametric coefficients are
 185 then computed, given the initial values of the index coefficients denoted by α , as
 186 follows

$$\beta = (S_{\hat{U}_2})^{-1} S_{\hat{U}_2 \hat{W}_2}, \quad (2.14)$$

187 where $S_{AB} = \frac{1}{n} \sum_{i=1}^n A_i B_i'$, $S_A = S_{AA}$, $(S_A)^{-1}$ is a generalized inverse of (S_A) ,
 188 $\hat{W}_{2i} \equiv Y_i - \hat{E}_i(Y_i | V_i, \hat{\eta}_i)$ and $\hat{U}_{2i} \equiv X_i - \hat{E}_i(X_i | V_i, \hat{\eta}_i)$. Next, based on $\beta \in B_n$, $\hat{\alpha}$, \hat{h}_v
 189 and $\hat{h}_{\hat{\eta}}$ are computed by minimizing the objective function as follows

$$\min_{\alpha \in A_n, h_v, h_{\hat{\eta}} \in \mathcal{H}_n} \hat{J}(\alpha, h_v, h_{\hat{\eta}}) \equiv \min_{\alpha \in A_n, h_v, h_{\hat{\eta}} \in \mathcal{H}_n} \frac{1}{n} \sum_{i=1}^n (\hat{W}_{2i} - \hat{U}_{2i}' \beta)^2, \quad (2.15)$$

190 where $A_n = \{\alpha : \|\alpha - \alpha_0\| \leq Cn^{-1/2}\}$, $B_n = \{\beta : \|\beta - \beta_0\| \leq Cn^{-1/2}\}$ and
 191 $\mathcal{H}_n = \{h_z, h_v, h_{\hat{\eta}} : Cn^{-1/5} \leq h_z, h_v, h_{\hat{\eta}} \leq C'n^{-1/5}\}$ for $0 < C < C' < \infty$. Finally,
 192 re-estimate β_0 by using $\hat{\alpha}$, \hat{h}_v and $\hat{h}_{\hat{\eta}}$ as follows

$$\hat{\beta} = (S_{\hat{U}_3})^{-1} S_{\hat{U}_3 \hat{W}_3}, \quad (2.16)$$

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

194 **Remark 2.1.** *The conditions for α and β below (2.15) are not as restrictive as it*
 195 *seemed because $\hat{\alpha}$ and $\hat{\beta}$ are \sqrt{n} -consistent. Furthermore, \sqrt{n} -consistency is achieved*
 196 *without under-smoothing in the first-stage of the proposed estimation procedure (i.e.*
 197 *estimation of the reduced-form equation in (2.8)). In general, under-smoothing in the*
 198 *first-stage of the estimation procedure is not required when $q_z < 3$ and $q - q_1 < 3/2$.*

199 The remaining task is then to identify the unknown link function. It is plau-
 200 sible to apply the marginal integration technique of Linton and Nielsen (1995),
 201 and Tjøstheim and Auestad (1994) to identify each of the functions because of
 202 the additive specification of the conditional expectation relation (see below (2.9)).
 203 The standard identification condition in the literature is assuming that $E(g(V_0)) =$
 204 $E(\iota(\eta)) = 0$ (see Hastie and Tibshirani (1990), Gao et al. (2006) and Gao (2007) for

205 details). Hence, the marginal integration technique identifies $g(\cdot)$ and $\iota(\cdot)$ functions
 206 up to some constant values as follows

$$m_1(V_0) \equiv \int m(V_0, \eta) dQ(\eta) = g(V_0) + C \text{ and } m_2(\eta) \equiv \int m(V_0, \eta) dQ(V_0) = \iota(\eta) + C',$$

207 where $C \equiv \int \iota(\eta) dQ(\eta)$, $C' \equiv \int g(V_0) dQ(V_0)$ and Q is a probability measure in \mathbb{R}
 208 with $\int dQ(\eta) = \int dQ(V_0) = 1$. The estimate of the link function can therefore be
 209 obtained by

$$\hat{m}_1(\hat{V}) = \frac{1}{n} \sum_{i=1}^n \hat{m}(\hat{V}, \hat{\eta}_i) \text{ and } \hat{g}(\hat{V}) = \hat{m}_1(\hat{V}) - \hat{C}, \quad (2.17)$$

210 where $\hat{m}(\hat{V}, \hat{\eta}_i) = \hat{E}(Y|\hat{V}, \hat{\eta}_i) - \hat{E}(X|\hat{V}, \hat{\eta}_i)' \hat{\beta}$, $\hat{C} = \frac{1}{n} \sum_{i=1}^n \hat{m}_1(\hat{V}_i)$, and $\hat{m}_1(\hat{V})$ is
 211 estimated by keeping \hat{V}_i at \hat{V} while taking average over $\hat{\eta}_i$.

212 Before discussing the main theoretical results of the estimators proposed above,
 213 the estimation procedure is briefly summarized as follows.

214 *Step 2.1:* Estimate the endogeneity control covariate, $\hat{\eta}$, as in (2.13).

215 *Step 2.2:* Estimate β as in (2.14) with $\hat{\eta}_i$ from Step 2.1 and α .

216 *Step 2.3:* Estimate $\hat{\alpha}$ and $\hat{\beta}$ as in (2.16) and (2.18), respectively.

217 *Step 2.4:* Estimate $\hat{m}(\hat{V}_i, \hat{\eta}_i)$ by using (2.10) with $\hat{\alpha}$ and $\hat{\beta}$ from Step 2.3, then
 218 perform the marginal integration technique to estimate $\hat{g}(\hat{V})$ as in (2.17).

219 2.3. Asymptotic Properties of Proposed Estimators

220 In this subsection, the asymptotic properties of the estimators are discussed as
 221 follows. The required necessary conditions are presented first. Given ρ , let $\mathcal{A}_{j'}^\rho$
 222 denote the set of all points in $\mathbb{R}^{k'}$, where $k' = q$ or 1 , at a distance no greater than
 223 ρ from $\mathcal{A}_{j'}$ for $j' = x, \eta$. Let $\mathcal{U} = \{(V_0, \eta) : X \in \mathcal{A}_x^\rho \text{ and } \eta \in \mathcal{A}_\eta^\rho\}$ and $f(V_0, \eta)$
 224 denote the joint density function of (V_0, η) with random arguments of X' and η .
 225 The necessary regularity conditions are then as follows.

226 **Assumption 2.1.** *The vector of instrumental variables $\{Z_i : i \geq 1\}$ satisfy (2.8).*

227 **Assumption 2.2.** *The joint density functions of $f_z(Z)$ and $f(V, \eta)$ are bounded and
 228 are bounded away from zero with bounded and continuous second derivatives on \mathcal{A}_z
 229 and \mathcal{U} for all values of $\alpha \in A_n$, respectively.*

230 **Assumption 2.3.** Assume that $g_x(Z)$, and $m(V, \eta)$, $m_y(V, \eta)$ and $m_x(V, \eta)$ have
 231 bounded and continuous second derivatives on \mathcal{A}_z and \mathcal{U} for all values of $\alpha \in A_n$.

232 **Assumption 2.4.** Assume that a univariate kernel function $k(\cdot)$ and its first deriva-
 233 tive $k^{(1)}(\cdot)$ are supported on the interval $(-1, 1)$ and $k(\cdot)$ is a symmetric density
 234 function. Furthermore, both $k(\cdot)$ and $k^{(1)}(\cdot)$ satisfy the Lipschitz conditions.

235 **Assumption 2.5.** Let $E(\eta|Z) = 0$ and $E(\eta^2|Z) = \sigma_1^2(Z)$, $E(e|X, \eta) = 0$ and
 236 $E(e^2|X, \eta) = \sigma^2(X, \eta)$, $E(u|X, \eta) = 0$ and $E(u^2|X, \eta) = \sigma_2^2(X, \eta)$, and the func-
 237 tions σ^2 , σ_1^2 and σ_2^2 are bounded and continuous. In addition, $\sup_i E\|X_i\|^l < \infty$,
 238 $\sup_i E|Y_i|^l < \infty$ and $\sup_i E\|Z_i\|^l < \infty$ for some large enough $l > 2$.

239 Assumption 2.2 is necessary to avoid the random denominator problem. As-
 240 sumptions 2.2 and 2.3 ensure that the kernel function in Assumption 2.4 leads to a
 241 second-order bias in kernel smoothing. A higher-order bias can be achieved by im-
 242 posing more restrictive conditions on the smoothness of the functions (see Robinson
 243 (1988) for details). The condition on the first derivative of the kernel function in
 244 Assumption 2.4 permits the use of the Taylor expansion argument to address the
 245 generated covariate, $\hat{\eta}_i$ (a similar condition on the derivatives of the kernel func-
 246 tion can be found in Hansen (2008)). The Lipschitz conditions for both the kernel
 247 function and its derivative provide the convenience for the proof of the uniform
 248 convergence. Finally, Assumption 2.5 grants the use of the Chebyshev inequality.

249 Now let us introduce a few necessary notations used in the main theoretic-
 250 al results below. Let $\mathcal{K}_{z,2} = \int z^2 K_{h_z}(z) dz$, $\mathcal{K}_{\eta,2} = \int \eta^2 k_{h_\eta}(\eta) d\eta$ and $\mathcal{K}_{v,2} =$
 251 $\int v_0^2 k_{h_v}(v_0) dv_0$. Furthermore, let $\mathcal{K}_z = \int k_{h_{z,j}}(z)^2 dz$ and $\mathcal{K} = \mathcal{K}_v \mathcal{K}_\eta$, where $\mathcal{K}_v =$
 252 $\int k_{h_v}(v_0)^2 dv_0$ and $\mathcal{K}_\eta = \int k_{h_\eta}(\eta)^2 d\eta$. Let $f_{z,j}^{(r)}$ be the r -th derivatives of $f_z(z)$ with
 253 respect to Z_j , for $j = 1, \dots, q_z$, and let $f_{v_0}^{(r)}(v_0, \eta)$ and $f_\eta^{(r)}(v_0, \eta)$ be the r -th partial
 254 derivatives of $f(v_0, \eta)$ with respect to V_0 and η , respectively. Moreover, let $g_{x,j}^{(r)}(z)$
 255 be the r -th partial derivatives of $g_x(z)$ with respect to Z_j , and let $m_{v_0}^{(r)}(V_0, \eta)$ and
 256 $m_\eta^{(r)}(v_0, \eta)$ be that of $m(v_0, \eta)$ with respect to V_0 and η , respectively. Then, let

$$B_z(z) \equiv \frac{\mathcal{K}_{z,2}}{2f(z)} \left\{ 2f_{z,j}^{(1)}(z)g_{x,j}^{(1)}(z) + f_z(z)g_{x,j}^{(2)}(z) \right\}$$

$$B_v(v_0, \eta) \equiv \frac{\mathcal{K}_{v,2}}{2f(v_0, \eta)} \left\{ 2f_{v_0}^{(1)}(v_0, \eta)m_{v_0}^{(1)}(v_0, \eta) + f(v_0, \eta)m_{v_0}^{(2)}(v_0, \eta) \right\}$$

$$B_\eta(v_0, \eta) \equiv \frac{\mathcal{K}_{\eta,2}}{2f(v_0, \eta)} \{2f_\eta^{(1)}(v_0, \eta)m_\eta^{(1)}(v_0, \eta) + f(v_0, \eta)m_\eta^{(2)}(v_0, \eta)\}.$$

258 In addition, let

$$\begin{aligned} IMSE_1(h_z) &\asymp \int \left\{ \left[\sum_{j=1}^{q_z} B_{z,j}(z)h_{z,j}^2 \right]^2 + \frac{\mathcal{K}_z^{q_z}}{nh_{z,1} \dots h_{z,q_2}} \frac{\sigma_1^2(z)}{f_z(z)} \right\} f(z) dz \\ IMSE_2(h_v, h_\eta) &\asymp \int \left\{ [B_v(v_0, \eta)h_v^2 + B_\eta(v_0, \eta)h_\eta^2]^2 + \frac{\mathcal{K}}{nh_v h_\eta} \frac{\sigma^2(V_0, \eta)}{f(v_0, \eta)} \right\} f(x, \eta) dx d\eta, \end{aligned}$$

259 where \asymp means that the quotient of the two sides tends to 1 as $n \rightarrow \infty$.

260 **Theorem 2.1.** *Under Assumptions 2.1 to 2.5, the minimizing objective function in*
261 *(2.15) is rewritten as follows*

$$\hat{J}(\alpha, h_v, h_\eta) = \tilde{J}(\alpha) + T_1(h_z) + T_2(h_v, h_\eta) + R_1(\alpha, h_v, h_\eta) + R_2(\alpha, h_v, h_\eta, h_z), \quad (2.18)$$

262 where $T_1(h_z) \equiv \frac{1}{n} \sum_{i=1}^n \{\hat{g}_{x,i}(Z_i) - g_x(Z_i)\}^2 = IMSE_1(h_z) + R_3(h_z)$, $T_2(h_v, h_\eta) \equiv$
263 $\frac{1}{n} \sum_{i=1}^n \{\hat{m}_i(V_{0i}, \eta_i) - m(V_{0i}, \eta_i)\}^2 = IMSE_2(h_v, h_\eta) + R_4(h_v, h_\eta)$, $\sup_{\alpha \in A_n, h_v, h_\eta \in \mathcal{H}_n} |R_1(\alpha, h_v, h_\eta)| =$
264 $o_p(n^{-1/2})$, and $\sup_{\alpha \in A_n, h_v, h_\eta, h_z \in \mathcal{H}_n} |R_2(\alpha, h_v, h_\eta, h_z)| = o_p(n^{-1/2})$ with $\hat{m}_i(\cdot)$ and $\hat{g}_{x,i}(\cdot)$ be-
265 *ing the leave-one-out local constant estimators of $m(\cdot)$ and $g_x(\cdot)$, respectively. More*
266 *importantly*

$$\tilde{J}(\alpha) = \frac{1}{n} \sum_{i=1}^n \{W_i - U_i' \beta\}^2,$$

267 where $W_i \equiv Y_i - E(Y_i | V_i, \eta_i)$ and $U_i \equiv X_i - E(X_i | V_i, \eta_i)$. Furthermore, $\sup_{h_z \in \mathcal{H}_n} |R_3(h_z)| =$
268 $o_p(n^{1/5})$ and $\sup_{h_v, h_\eta \in \mathcal{H}_n} |R_4(h_v, h_\eta)| = o_p(n^{1/5})$ because they do not depend on α .

269 The results of Theorem 2.1 show the attractive properties of our proposed CF ap-
270 proach. Similar to the results of Härdle et al. (1993) and Xia et al. (1999), Theorem
271 2.1 shows that the properties of the bandwidth parameter estimators can be studied
272 while assuming α_0 is known. Moreover, the asymptotically optimal bandwidth pa-
273 rameters for estimating $m(\cdot)$ function are assumed to be used for the \sqrt{n} -consistent
274 estimation of α_0 . In addition, under-smoothing is not required in estimating the
275 first-stage reduced-form equation, as already stated in Remark 2.1. In particular,
276 Theorem 2.1 suggests that minimizing $\hat{J}(\alpha, h_v, h_\eta)$ simultaneously with respect to α ,

277 h_v and $h_{\hat{\eta}}$, is asymptotically equivalent to separately minimizing $\tilde{J}(\alpha)$ with respect
 278 to α , $T_1(h_z)$ with respect to h_z , and $T_2(h_v, h_{\eta})$ with respect to h_v and h_{η} , assuming
 279 that α_0 and η are known. This is because the remainder terms, namely $R_1(\alpha, h_v, h_{\eta})$
 280 and $R_2(\alpha, h_z, h_v, h_{\eta})$, are shown to be asymptotically negligible.

281 Next, the asymptotic properties of $\hat{\alpha}$ and $\hat{\beta}$ are shown as a corollary of Theorem
 282 2.1 given that $\Phi_{U_0} = [\{X - E(X|V_0, \eta)\}\{X - E(X|V_0, \eta)\}']$.

283 **Corollary 2.1.** *Under the assumptions of Theorem 2.1, the asymptotic properties*
 284 *of $\hat{\alpha}$ and $\hat{\beta}$ are as follows*

$$\sqrt{n}(\hat{\beta} - \beta_0) \rightarrow_D N(0, Var_1), \quad (2.19)$$

285 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

$$\sqrt{n}(\hat{\alpha} - \alpha_0) \rightarrow_D N(0, Var_2), \quad (2.20)$$

286 where $Var_2 = \sigma^2 \left[\left\{ \left(m_0^{(1)} \right)^2 \Phi_{U_0} \right\}^- - \left\{ m_0^{(1)} \Phi_{U_0} \right\}^- \Phi_{U_0} \left\{ m_0^{(1)} \Phi_{U_0} \right\}^- \right]$.

287 Finally, the asymptotic properties of $\hat{g}(\hat{v})$ are presented in Theorem 2.2 below.

288 **Theorem 2.2.** *Under the assumptions of Theorem 2.1, and $\inf_{z \in \mathcal{A}_z} f_z(z) > 0$ and*
 289 *$\inf_{x, \eta \in \mathcal{U}} f(v_0, \eta) > 0$, the asymptotic results of $\hat{g}(\hat{v})$ are as follows*

$$\sqrt{nh_v}(\hat{g}(\hat{v}) - g(v_0) - Bias) \rightarrow_D N(0, Var),$$

290 where $Bias = h_v^2 B_v(v_0, \eta) + h_{\eta}^2 B_{\eta}(v_0, \eta)$ and $Var = f_{\alpha}(v_0) \mathcal{K}_v \int \frac{\sigma^2(V_0, \eta) f_{\eta}^2(\eta)}{f^2(v_0, \eta)} dQ(\eta)$ with
 291 $f_{\alpha}(v_0)$ and $f_{\eta}(\eta)$ denoting the density functions of V_0 and η , respectively.

292 **Remark 2.2.** *In these results, it is clear the first stage nonparametric estimation*
 293 *does not contribute to the asymptotic variance of the estimators in the final stage.*
 294 *This characteristic is common among multi-stage nonparametric estimation proce-*
 295 *dures (see Su and Ullah (2008) for an example). However, this differs from the*
 296 *work of Li and Wooldridge (2002) which considers parametrically generated covari-*
 297 *ates in a PL semiparametric regression model. Li and Wooldridge (2002) showed*
 298 *that the variance of the first stage estimation is not asymptotically negligible instead*
 299 *contributes to the variances of the estimators of the finite-dimensional parameters*
 300 *in the final stage.*

301 **Remark 2.3.** *It is also interesting to explore the case of performing the CF approach*
302 *without the presence of endogeneity. The essential stochastic assumption of the CF*
303 *approach below (2.8) implies no existence of any endogeneity control function and,*
304 *hence there is no identification problem in estimating the link function. Therefore,*
305 *performing the CF approach without the presence of endogeneity causes an unneces-*
306 *sary multi-stage nonparametric estimation and the presence of redundant covariates*
307 *in estimating the link function. However, the theoretical results of the proposed es-*
308 *timators particularly Theorems 2.1 and 2.2 and Corollary 2.1 are still valid with*
309 *minor modifications, especially in terms of $IMSE_2(h_v, h_\eta)$, Var_1 and Var_2 , and the*
310 *bias and the variance of $\hat{g}(\hat{v})$. The minor modifications are as follows*

$$IMSE_2(h_v, h_\eta)^* \asymp \int \left\{ [B_v^*(v_0, \eta) h_v^2 h_\eta^2]^2 + \frac{\mathcal{K}}{nh_v h_\eta} \frac{\sigma^{*2}(V_0, \eta)}{f(v_0, \eta)} \right\} f(x, \eta) dx d\eta$$

$$Bias^* = h_v^2 B_v^*(v_0, \eta) \text{ and } Var^* = f_\alpha(v_0) \mathcal{K}_v \int \frac{\sigma^{*2}(V_0) f_\eta^2}{f^2(v_0, \eta)} dQ(\eta),$$

311 where $B_v^*(v_0, \eta) = \frac{\mathcal{K}_{v,2}}{2f(v_0, \eta)} \left\{ 2f_{v_0}(v_0, \eta) g^{(1)}(v_0) + f(v_0, \eta) g_{v_0}^{(2)}(v_0, \eta) \right\}$ and $\sigma^{*2} = E(\epsilon^2 | X, \eta)$
312 $= E(\epsilon^2 | X)$, and Var_1^* and Var_2^* are obtained by replacing $m_0^{(1)}$ with $g_0^{(1)}$ in (2.19)
313 and (2.20) with $g_0^{(1)}$ being the first derivative of $g(v_0)$ with respect to V_0 .

314 **Remark 2.4.** *Our results can also be extended to more general data structure where*
315 *a random sample $\{(X'_t, Z'_t, Y_t); t = 1, \dots, n\}$ is a strictly stationary and α -mixing*
316 *process under Assumptions 2.6 and 2.7 below in addition to 2.1 to 2.5 above.*

317 In the rest of this section, we discuss about how to extend these established
318 theoretical results to stationary time series data as in Remark 2.4. First, let $\xi_t \equiv$
319 $(X'_t \alpha_0, \eta_t)$ and $f_\xi(\xi)$ denote the joint density function of $X' \alpha_0$ and η . The necessary
320 regularity conditions for the strictly stationary and α -mixing case are then as follows.

321 **Assumption 2.6.** (i) *The conditional densities satisfy the following conditions*

$$f_{\xi_1, \xi_l | X_1, X_l}(\xi_1, \xi_l) \leq C < \infty; f_{\xi_1, \xi_l | Y_1, Y_l}(\xi_1, \xi_l) \leq C' < \infty; f_{Z_1, Z_l | X_1, X_l}(Z_1, Z_l) \leq C'' < \infty$$

322 *for some constants $C, C', C'' > 0$ and for all $l \geq 1$. (ii) The mixing and moment*
323 *conditions are as follows*

$$\sum_l l^\alpha [\alpha(l)]^{1-2/l} < \infty, E\|X_0\|^l < \infty \text{ and } f_{\xi_1 | X_1}(\xi | X) \leq C < \infty;$$

$$\sum_l l^{a'} [\alpha(l)]^{1-2/l} < \infty, \quad E|Y_0|^l < \infty \text{ and } f_{\xi_1|Y_1}(\xi|Y) \leq C' < \infty;$$

324

$$\sum_l l^{a''} [\alpha(l)]^{1-2/l} < \infty, \quad E\|Z_0\|^l < \infty \text{ and } f_{Z_1|X_1}(z|X) \leq C'' < \infty,$$

325 where $l > 2$ and $a, a', a'' > 1 - 2/l$. (iii) There is a sequence of positive integer s_T ,
 326 which satisfies $s_T \rightarrow \infty$ and $s_T = o\{(nh_{z,T}^{q_z})^{1/2}\}$, such that $(n/h_{z,T}^{q_z})^{1/2}\alpha(s_T) \rightarrow 0$ as
 327 $T \rightarrow \infty$.

328 **Assumption 2.7.** (i) Let the density functions $f_z(z)$ and $f(v_0, \eta)$ satisfy $\inf_{z \in A_z} f_z(z) >$
 329 0 and $\inf_{x, \eta \in \mathcal{U}} f(v_0, \eta) > 0$. (ii) In addition, we require the following moments condi-
 330 tions

$$E\|X\|^s < \infty, \quad \sup_{\xi \in \mathcal{U}} \int \|X\|^s f(x, \xi) dx, \quad E|Y|^s < \infty, \quad \sup_{\xi \in \mathcal{U}} \int |Y|^s(y, \xi) dy; \quad \int_{x \in A_z} \|X\|^s f(x, z) dx,$$

331 for some $s > 2$. (iii) The bandwidth sequences, h_v , h_η and h_z , tend to zero as
 332 $T \rightarrow \infty$ and satisfy, for some $\delta > 0$,

$$T^{1-2s-1-2\delta} h_z^{q_z} \rightarrow \infty; \quad T^{1-2s-1-2\delta} h_v h_\eta \rightarrow \infty; \quad T^{1-2s-1-2\delta} (h_z^{q_z} h_v h_\eta^3)^{1/2} \rightarrow \infty.$$

333 In the proof of the \sqrt{n} -consistency of $\hat{\alpha}$ and $\hat{\beta}$ in the case of Remark 2.4, Propo-
 334 sitions A.1 to A.15 in the supplementary document encompass the extra covari-
 335 ance terms caused by the serial dependences in the sample. Under Assumptions
 336 2.1 to 2.5 and 2.6(i)(ii), those covariance terms can be shown to be $o_p(n^{-1/2})$.
 337 For instance, the extra covariance term in Proposition A.1 might be derived as
 338 $\sum_{l=1}^{n-1} (1-t/n) \text{Cov}(\hat{\varphi}_1, \hat{\varphi}_{l+1}) = o(h_v h_\eta)$. However the consistency of $\hat{g}(\hat{v})$ requires
 339 stronger conditions than the case of $\hat{\alpha}$ and $\hat{\beta}$, namely the uniform convergence of
 340 $\hat{f}(v_0, \eta)$, which requires the uniform convergences of Q_j , where $j = 1, \dots, 5$ in (B.1)
 341 in the supplementary document. Under Assumptions 2.1 to 2.5, 2.6(i)-(ii) and 2.7,
 342 Q_j are shown to be $o_p(1)$ as follows

$$\sup_{\xi \in \mathcal{U}, z \in A_z} |Q_{2i}| = \sup_{\xi \in \mathcal{U}, z \in A_z} |Q_{5i}| = O_p \left\{ \left(\frac{(\ln n)^2}{n^2 h_z^{q_z} h_v h_\eta^3} \right)^{1/2} + h_z^2 (h_v^2 + h_\eta^2) \right\}.$$

343 Furthermore, the asymptotic normality of $\hat{g}(\hat{v})$ is then obtained by applying As-
 344 sumption 2.6 (iii) for the standard nonparametric small-block and large-block ar-
 345 guments. Nonetheless, the asymptotic normalities of $\hat{\alpha}$ and $\hat{\beta}$ are obtained by ap-
 346 plying the part of Assumption 2.6 (ii), namely $\sum_l l^a [\alpha(l)]^{1-2/l} < \infty, E\|X_0\|^l < \infty,$

347 $\sum_l l^{\alpha'} [\alpha(l)]^{1-2/l} < \infty$ and $E|Y_0|^l < \infty$, to (A.6) and (A.10) for the small-block and
 348 large-block arguments of a standard strictly stationary and α -mixing process.

349 2.4. Simulation Studies

350 In this section⁶, the finite-sample performance of the proposed estimator is in-
 351 vestigated by making a comparison between the performances of the estimation
 352 method introduced in Xia et al. (1999) referred as the XTL procedure and the CF
 353 approach established in Section 2.2 as the KS procedure in the presence of endo-
 354 geneity. Throughout this section, optimization is implemented by using a limited
 355 memory Broyden-Fletcher-Goldfarb-Shanno algorithm for the bound constrained
 356 optimization of Byrd et al. (1995). All simulation exercises are conducted in R with
 357 the Gaussian kernel function and the number of replications $Q = 200$. To compare
 358 and evaluate the finite sample performances of the procedures, the mean and mean
 359 absolute errors of the estimates of both coefficients, α_0 and β_0 , across Q replications
 360 are computed in Tables 2.1 and 2.2. The averaged absolute error of the estimates
 361 of the unknown structural function is also computed as follows

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

362 where n is the number of samples.

363 In the analysis that follows, an example model of the following form is considered

$$Y_i = \beta_{01}X_{1i} + \beta_{02}X_{2i} + \beta_{03}X_{3i} + g(V_{0i}) + \epsilon_i, \quad (2.21)$$

364 where $V_0 = \alpha_{01}X_1 + \alpha_{02}X_2 + \alpha_{03}X_3$, $g(V_0) = \exp \{-2(\alpha_{01}X_1 + \alpha_{02}X_2 + \alpha_{03}X_3)^2\}$,
 365 and X_j is independently and uniformly distributed on $[-1, 1]$ for $j = 1, 2$. It is
 366 required that $\beta_0 \perp \alpha_0$ with $\|\alpha_0\| = 1$. In order for these conditions to be satisfied,
 367 define $\beta_{02} = 0.4$, $\beta_{03} = 0$, $\alpha_{01} = 0.7$, $\alpha_{02} = -0.6$, then β_{01} and α_{03} are defined as
 368 follows

$$\alpha_{03} = \sqrt{1 - \alpha_{01}^2 - \alpha_{02}^2} \quad \text{and} \quad \beta_{01} = -\frac{\beta_{02}\alpha_{02}}{\alpha_{01}}.$$

⁶The results of extensive simulation exercises for GPLSI model are available by a request to the author.

369 In this example, endogeneity is introduced by letting $X_3 = Z + \eta$, where Z and η
 370 are independently and uniformly distributed on $[-0.5, 0.5]$ and $[-1, 1]$, respectively,
 371 and $\epsilon = \eta + e$ with e is independent and standard normally distributed. Tables 2.1
 372 and 2.2 present the results based on the XTL and KS procedures, respectively.

373 The simulation results in Table 2.1 show the strong evidence against the use
 374 of XTL procedure in the presence of endogeneity. This evidence is clear when the
 375 averaged absolute errors, $ae_{\hat{g}}$, in Table 2.1 are considered. On the other hand, the
 376 simulation results in Table 2.2 suggest that the KS procedure is able to identify the
 377 link function, namely $g(\cdot)$ function, in the presence of endogeneity.

378 **Table 2.1.** *EGPLSI model with endogeneity and the XTL's procedure.*

	n	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$
379	50	0.3130	0.4332	0.8884	-0.7748	0.5597
	150	0.3088	0.4340	0.8993	-0.7671	0.5279
	300	0.3142	0.4264	0.8988	-0.7674	0.5225
	500	0.3135	0.4288	0.8960	-0.7653	0.5179

	n	$ \hat{\beta}_1 - \beta_{01} $	$ \hat{\beta}_2 - \beta_{02} $	$ \hat{\alpha}_1 - \alpha_{01} $	$ \hat{\alpha}_2 - \alpha_{02} $	$ \hat{\alpha}_3 - \alpha_{03} $	$ae_{\hat{g}}$
380	50	0.0656	0.0714	0.1691	0.1253	0.1586	0.0905
	150	0.0428	0.04572	0.0859	0.0559	0.0910	0.0891
	300	0.0331	0.03377	0.0629	0.0548	0.0426	0.0895
	500	0.0306	0.0319	0.0229	0.0156	0.0181	0.0906

381 **Table 2.2.** *EGPLSI model with endogeneity and the KS procedure.*

	n	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$
382	50	0.2645	0.4652	0.9638	-0.8249	0.5483
	150	0.3260	0.4135	0.8975	-0.7852	0.4756
	300	0.3486	0.3945	0.8090	-0.6997	0.4382
	500	0.3555	0.3891	0.7353	-0.6295	0.3992

	n	$ \hat{\beta}_1 - \beta_{01} $	$ \hat{\beta}_2 - \beta_{02} $	$ \hat{\alpha}_1 - \alpha_{01} $	$ \hat{\alpha}_2 - \alpha_{02} $	$ \hat{\alpha}_3 - \alpha_{03} $	$ae_{\hat{g}}$
383	50	0.0816	0.0684	0.1678	0.1389	0.1195	0.0632
	150	0.0307	0.0264	0.1244	0.0962	0.0769	0.0265
	300	0.0213	0.0183	0.0446	0.0327	0.0285	0.0160
	500	0.0189	0.0159	0.0416	0.0319	0.0263	0.0124

384 3. Semiparametric CF approach to Shape-Invariant Empirical Engel Curves

385 In this section, a flexible shape-invariant Engel curves system is analyzed within
 386 the framework of the EGPLSI model with the proposed CF approach above. The
 387 consumer optimization theory suggests to include a scale and a shift parameters

388 within a flexible shape-invariant empirical Engel curve for the individual household
389 heterogeneity (see Pendakur (1999), Blundell and Powell (2003) and Blundell et al.
390 (2007) for examples). In addition, the endogeneity of total expenditure is well-known
391 which is caused by the two-stage budgeting model (see Blundell et al. (1998) and
392 Blundell et al. (2007) for details). Hence, it is natural to study a shape-invariant
393 Engel curves system within the framework of the EGPLSI model with the newly
394 developed CF approach.

395 *3.1. The Empirical Model and Estimation*

396 Hereafter, let $\{Y_{il}, X_{1i}, X_{2i}\}_{i=1}^n$ represent an IID sequence of n household obser-
397 vations on the budget share Y_{il} of good $l = 1, \dots, L \geq 1$ for each household i facing
398 the same relative prices, the log of total expenditure X_{1i} , and a vector of household
399 composition variables X_{2i} . For each commodity l , budget shares and total outlay are
400 related by a general stochastic Engel curve, namely $Y_l = G_l(X_1) + \epsilon_l$, where $G_l(\cdot)$
401 is an unknown function that can be estimated by using a standard nonparametric
402 method under the exogeneity assumption of total expenditure (i.e. $E(\epsilon_l|X_1) = 0$).
403 Nonetheless, a number of previous studies have reported that household expendi-
404 tures typically display great variation with demographic composition. A simple
405 approach for estimating the model is to stratify by each distinct discrete outcome
406 of X_2 and then carry out our estimation with nonparametric smoothing within each
407 cell. At some point, however, it may be useful to pool the Engel curves across
408 different household demographic types and to allow X_1 to enter each Engel curve
409 semiparametrically. This idea leads to the specification below

$$Y_{il} = \beta'_{0l}X_{2i} + g_l(X_{1i} - \phi(\gamma'_0 X_{2i})) + \epsilon_{il}, \quad (3.1)$$

410 where $g_l(\cdot)$ is an unknown function and $\phi(\gamma'_0 X_{2i})$ is a known function up to a finite set
411 of unknown parameters γ_0 , which can be interpreted as the log of general equivalence
412 scales for household i . In the current paper, $\phi(\gamma'_0 X_{2i}) = \gamma'_0 X_{2i}$ is chosen so that (3.1)
413 is specified as follows

$$Y_{il} = \beta'_{0l}X_{2i} + g_l(X_{1i} - \gamma'_0 X_{2i}) + \epsilon_{il}. \quad (3.2)$$

414 In this application, total expenditure is allowed to be endogenous and a measure of
 415 earning of the head of each household is used as an instrument.

416 Following the CF approach discussed above, the empirical model to be estimated
 417 is the following form below

$$Y_{il} = \beta_{01,l}X_{1i} + \beta'_{0l}X_{2i} + g_l(\alpha_{01}X_{1i} + \alpha'_{02}X_{2i}) + \epsilon_{il} \quad (3.3)$$

$$X_{1i} = m_{X1}(Z_i) + \eta_i, \text{ where } E(\eta|Z) = 0 \text{ and } E(\epsilon_l|Z, \eta) = E(\epsilon_l|\eta) \neq 0, \quad (3.4)$$

418 with $m_{X1}(Z) = E(X_1|Z)$ and $\{Z_i\}_{i=1}^n$ represents an IID sequence of the measure of
 419 earning of n heads of households and (3.3) is a semiparametric model that satisfies
 420 all the identification conditions required in the construction of the EGPLSI model.
 421 The theoretically consistent model in (3.1) can then be solved based on (3.3). To
 422 this end, a similar scaling transformation to that explained in Section 2 is used. In
 423 the remainder of this section, some specific details about the estimation procedure
 424 are discussed. Rather than basing the discussion on (3.3) to (3.4), it is statistically
 425 more equivalent to do so based on as follows

$$Y_{il} = \beta'_{0l}X_{2i} + g_l(X_{1i} - \gamma'_0 X_{2i}) + \epsilon_{il} \quad (3.5)$$

$$X_{1i} = m_{X1}(Z_i) + \eta_i, \text{ where } E(\eta|Z) = 0 \text{ and } E(\epsilon_l|Z, \eta) = E(\epsilon_l|\eta) \neq 0. \quad (3.6)$$

426 These models suggest the conditional expectation relationship shown below

$$E(Y_l|(X_1 - \gamma'_0 X_2), \eta) - \beta'_{0l}E(X_2|(X_1 - \gamma'_0 X_2), \eta) = g_l(X_1 - \gamma'_0 X_2) + \iota(\eta), \quad (3.7)$$

427 where $E(\epsilon_l|(X_1 - \gamma'_0 X_2), \eta) = E(\epsilon_l|\eta) \equiv \iota(\eta) \neq 0$, which immediately leads to

$$Y_{il} = \beta'_{0l}X_{2i} + g_l(X_{1i} - \gamma'_0 X_{2i}) + \iota(\eta_i) + e_{il}, \quad (3.8)$$

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

428 where $E(e_l|X_1, X_2, \eta) = 0$. Let $m_l(\{X_{1i} - \gamma'_0 X_{2i}\}, \eta_i) = g_l(X_{1i} - \gamma'_0 X_{2i}) + \iota(\eta_i)$. In
 429 order to use (3.8), it is important to note that

$$m_{1,l}(X_1 - \gamma'_0 X_2) = \int m_l(\{X_1 - \gamma'_0 X_2\}, \eta) d\eta \text{ and } g_l(X_1 - \gamma'_0 X_2) = m_{1,l}(X_1 - \gamma'_0 X_2) - C, \quad (3.10)$$

430 where $C = \int \iota(\eta)dQ(\eta)$ and $E(g_l(\cdot)) = 0$.

431 If a linear specification is imposed on $\iota(\cdot)$, (3.8) would be similar to the extended
 432 partially linear (EPL) model discussed in Blundell et al. (1998). In this case, Blun-
 433 dell et al. (1998) showed that a test of the endogeneity null can be constructed by
 434 testing $H_0 : \iota_l = 0$, where ι_l is an unknown parameter. The current paper, however,
 435 suggests more flexible functional form for testing the endogeneity null by construct-
 436 ing the variability bands for $\iota(\cdot)$. To do so, the following procedure is employed.

437 *Step 3.1.1:* Obtain an empirical estimate of $g_l(X_1 - \gamma'_0 X_2)$ in (3.10).

438 *Step 3.1.2:* Regress (3.8) using the estimates in Step 3.1.1 to obtain the nonpara-
 439 metric estimates of $\iota_l(\cdot)$.

440 *Step 3.1.3:* Compute the bias-corrected confidence bands for the nonparametric
 441 smoothing using the procedure introduced by Xia (1998). Finally, the Bonferroni-
 442 type variability bands are obtained by using a similar procedure discussed in Eubank
 443 and Speckman (1993).

444 To perform Step 3.1.1, the estimation procedure introduced in Section 2 is used.
 445 However, some modifications are required to take the vector of index coefficients, γ_0
 446 (a general equivalence scale for household i), into account. In this case, the objec-
 447 tive function (2.15) is only used for a particular commodity l . The new objective
 448 function, $\min_{\gamma \in A_n, h_{v,l}, h_{\hat{\eta},l} \in \mathcal{H}_n} \hat{J}(\gamma, h_{v,l}, h_{\hat{\eta},l})$, is the summation of these individual functions
 449 that is minimized with respect to γ and 14 smoothing parameters, particularly two
 450 for each commodity. Finally, the estimation procedure is completed by using $\hat{\gamma}$ as
 451 well as $\hat{h}_{\hat{v},l}$ and $\hat{h}_{\hat{\eta},l}$.

452 In addition, the model in (3.8) can also be re-stated as

$$Y_{il}^* = g_l(X_{1i} - \gamma'_0 X_{2i}) + e_{il}, \quad (3.11)$$

453 where $Y_l^* \equiv Y_l - \beta'_{0l} X_2 - \iota_l(\eta)$. The use of (3.11) relies on

$$m_{2,l}(\eta) = \int m_l(v, \eta) dv = \iota_l(\eta) + C' \text{ and } \iota_l(\eta) = m_{2,l}(\eta) - C', \quad (3.12)$$

454 where $V = X_1 - \gamma' X_2$, $C' = \int g(v) dQ(v)$ and $E(\iota_l(\cdot)) = 0$, which corresponds to
 455 (3.10) above. Hence, the model in (3.11) suggests that the estimates of the shape-
 456 invariant Engel curves and the related confidence bands are obtained as follows.

457 *Step 3.2.1:* Obtain empirical estimates of $\iota_l(\eta)$ in (3.12).

458 *Step 3.2.2:* Regress (3.11) using the estimates in Step 3.2.1 to obtain the nonpara-
 459 metric estimates of $g_l(\cdot)$.

460 *Step 3.2.3:* Compute the bias-corrected confidence bands about the nonparametric
 461 estimator in Step 3.2.2 using the procedure introduced by Xia (1998).

462 *3.2. The Engel Curve Data*

463 In our application, the data set is drawn from the British Family Expenditure
 464 Survey (FES) 1995-96. The seven broad categories of goods are considered as follows:
 465 (1) fuel, light and power (fuel hereafter); (2) fares, other travel costs and running of
 466 motor vehicles (fares); (3) food; (4) alcoholic drink and tobacco (alcohol); (5) leisure
 467 goods & services (leisure goods); (6) clothing and footwear (clothing); (7) personal
 468 goods & services (personal goods).

469 **Table 3.1.** *Descriptive statistics.*

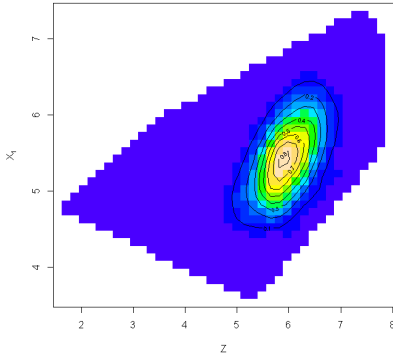
	Couples with 1 or 2 children		Couples without children	
	Mean	Std. Dev	Mean	Std. Dev
Budget shares:				
470 Fuel	0.0692	0.0011	0.0618	0.0012
Fares	0.1537	0.0025	0.1715	0.0031
Food	0.3235	0.0028	0.2768	0.0031
Alcohol	0.0844	0.0022	0.1144	0.0031
Leisure goods	0.2155	0.0038	0.2298	0.0045
Clothing	0.0926	0.0024	0.0872	0.0029
Personal goods	0.0606	0.0016	0.0581	0.0019
Expenditure and income:				
log (total expenditure)	5.4374	0.0130	5.4524	0.0161
log (income)	5.9205	0.0153	6.0397	0.0166
Sample size	1072		1278	

471 To maintain some demographic homogeneity, a subset of married or cohabiting
 472 couples are selected from the FES, particularly categories 1 and 3 of variable *ms* in
 473 table *adult*. In addition, those where the head of household is aged between 20 and
 474 55 (i.e. variable *age* in table *adult*) and in work (i.e. excluding the category 1 of the
 475 variable *fted* in the table *adult* and category 6 of the variable *a093* in the table *set8*)
 476 are considered. Finally, all households with three or more children are excluded.
 477 Our demographic variable, X_2 , is a binary dummy variable that reflects whether a
 478 couple has 1 or 2 children (where $X_2 = 1$) or no children (where $X_2 = 0$). Overall,
 479 there are 2350 observations, 1278 are couples with one or two children. Table 3.1
 480 shows larger expenditure shares for fuel, food, clothing and personal goods for the

481 households with children as expected. Also as expected, households without children
 482 are able to spend higher proportions of their total expenditure on alcohol and leisure
 483 goods. Overall, there are clear differences in the consumption patterns between the
 484 two demographic groups. The estimates of the scale and the shift coefficients are
 485 expected to reflect these differences.

486 Furthermore, the log of total expenditure on the nondurables and services is our
 487 measure of the continuous endogenous explanatory variable, X_1 . In our analysis that
 488 follows, the log of normal weekly disposable head of household income, specifically
 489 variable $p389$ of the table $set3$, is used as an instrument. The two variables show
 490 strongly-positive correlation with the correlation coefficients of 0.5660 and 0.5954
 491 for couples with and without children, respectively. Figures 3.1 and 3.2 present
 492 plots of the kernel estimates of the joint density for these variables. Finally, in the
 493 empirical application the instrument variable $Z = \Phi(\log \text{earnings})$ is taken, similar
 494 to Blundell et al. (2007).

495 **Figure 3.1.** *Kernel joint density estimates for log total expenditure and log weekly income – couples*
 496 *with 1 or 2 children.*

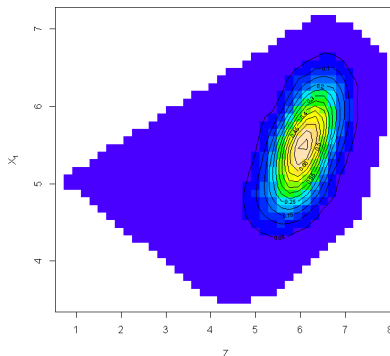


497 *3.3. Empirical Findings*

498 The important empirical findings are now presented and summarized in Table
 499 3.2. Although exact definitions of the data are not given in Blundell et al. (1998),
 500 Blundell et al. (1998) estimated the shape-invariant Engel curves for four broad
 501 categories of nondurables and services by using the FES data, namely fuel, fares,
 502 alcohol and leisure similar to this paper. The empirical estimate, $\hat{\gamma}$, of 0.36355
 503 reported in the first column is very close to 0.3698 as found in Blundell et al. (1998).
 504 Furthermore, the signs of the parameter estimates, $\hat{\beta}_i$, for the four broad categories

505 are all consistent with those of Blundell et al. (1998); specifically they are positive
 506 for food and leisure, but negative for alcohol, fares and fuel.

507 **Figure 3.2.** Kernel joint density estimates for log total expenditure and log weekly income – couples
 508 without children.



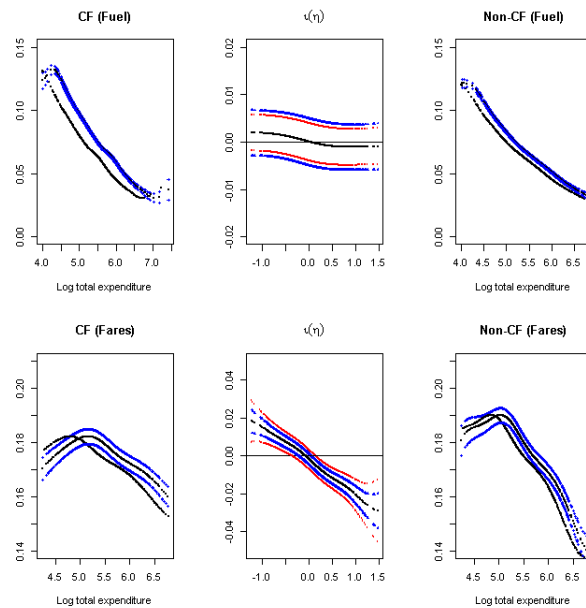
509 **Table 3.2.** Empirical results

$\hat{\gamma}$	Categories of goods	$\hat{\beta}_l$	$\hat{h}_{v,l}$	$\hat{h}_{\eta,l}$
0.36355	Fuel, light and power	-0.01401	0.14021	0.93631
	Fares, other travel costs and running of motor vehicles	-0.02027	0.19545	0.26831
	Food	0.00537	0.15120	0.25826
	Alcoholic Drink and Tobacco	-0.05205	0.30802	0.22569
	Leisure goods and services	0.05077	0.14663	0.40277
	Clothing and footwear	0.02079	0.14846	0.27234
	Personal goods and services	0.00738	0.49331	0.49335

512 The first columns of Figures 3.3 to 3.6 present the empirical estimates of the
 513 Engel curves for seven of the goods in our system based on the CF approach discussed
 514 in Section 3.1. For these plots, the smoothing parameters presented in the fourth
 515 and fifth columns of Table 3.2 are used. Furthermore, the third columns of these
 516 figures show the empirical estimates of the Engel curves computed from the Xia
 517 et al. (1999)’s procedure by which the exogeneity assumption is imposed on the
 518 total expenditure. Together with the estimated Engel curves, their 90% point-wise
 519 confidence bands are also reported. The bands are obtained by using the procedure
 520 discussed in Section 3.1. Let us now concentrate on the first columns. For fuel, food
 521 and alcohol, the Engel curves appear to demonstrate that the Working-Leser linear
 522 logarithmic formulation may provide a reasonable approximation. Nonetheless, for
 523 other shares, especially for fares, a nonlinear relationship between the shares and
 524 the log expenditure is evident. A detailed investigation of the data shows that on
 525 average, up to 70% of fares belongs to running of motor vehicles. Hence, motor

526 vehicles seemed to be a necessity good for a household for which the log of total
 527 expenditure is more than around 5.3 for those with children, for those without
 528 children, it is up to around 4.8. It seemed that motor vehicles are a superior good
 529 for those household where the log of total expenditure, is below these levels. The
 530 estimated shares for the couples with children are higher than those for couples
 531 without children, except extremely lower quantile of the log of total expenditure.
 532 This could lead to the nonlinear relationship witnessed in Figure 3.3.

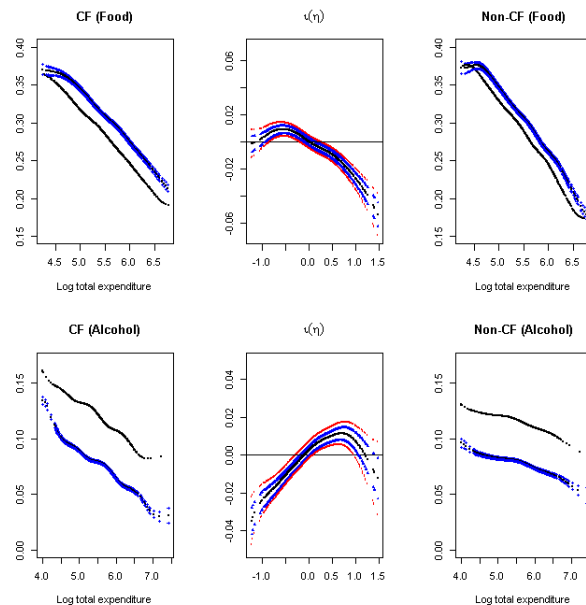
533 **Figure 3.3.** *Fuel and fares (90% confidence bands drawn for households with children)*



534 As expected, the estimated shares of fuel and food for households with children
 535 are consistently above those for households without children. Couples without chil-
 536 dren spends around 3% more of their budget on fuel and food than couples with
 537 children. In addition, the estimated shares of alcohol, leisure, clothing and per-
 538 sonal goods for households with children are consistently below those for households
 539 without children. Couples with children spend around 3%, 8% and 2% more of their
 540 budget on leisure, clothing and personal than couples with children at the same
 541 level of expenditure. In all but one case (i.e. fares), there seem to be a broadly
 542 parallel shift in the Engel curves from one demographic group to another. Our re-
 543 sults suggest that fuel, food and alcohol may be categorized as necessity goods in
 544 the sense that the demand for these goods increases proportionally less than the

545 increase in the total expenditure. These goods whose demand increases with the
 546 total expenditure are leisure, clothing and personal. The second column presents the
 547 nonparametric estimates of the control functions, $\iota_l(\cdot)$. With the estimated control
 548 functions, the two sets of bands, namely the 90% bias-corrected confidence bands
 549 for the nonparametric smoothing of Xia (1998) (blue) and the 90% Bonferroni-type
 550 variability bands of Eubank and Speckman (1993) (red) are also reported. Regarding
 551 fuel and personal, $\iota_l(\cdot)$ for these cases do not seem statistically significant. How-
 552 ever, the opposite is found for fares, food, leisure and clothing. Hence, neglecting
 553 potential endogeneity in the estimation can lead to incorrect estimates of the shape
 554 of Engel curves for these goods. This can be seen by comparing the first and the
 555 third columns of the figures. For these goods it is clear that the curvature changes
 556 significantly as the presence of endogeneity is allowed.

557 **Figure 3.4.** *Food and alcohol (90% confidence bands drawn for households with children)*

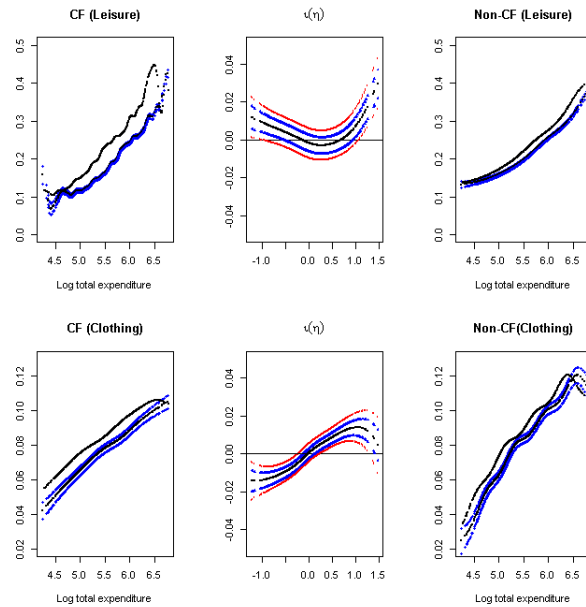


558 4. Conclusion

559 In this paper, the usefulness of the EGPLSI model in its ability to model a
 560 flexible shape-invariant specification is elaborated. A flexible shape-invariant speci-
 561 fication is easily studied within the EGPLSI framework because both scale and shift
 562 parameters are easily incorporated in the EGPLSI model. However, the applicability

563 of the EGPLSI model to an empirical study is limited because of its shortfalls in ad-
564 dressing endogeneity. Hence, the current paper develops the CF approach to address
565 endogeneity in the EGPLSI model. The proposed CF approach inherits an intrinsic
566 feature of the generated endogeneity control covariates and hence multi-stage
567 nonparametric estimation procedure. This paper establishes the theoretical validity
568 of the proposed estimation procedure and closes with the theoretical discussion by
569 providing the straightforward extension of the results to a strictly stationary and α -
570 mixing process. The paper also presents the satisfactory finite sample performance
571 of proposed estimators from a Monte Carlo simulation exercise. Finally, the semi-
572 parametric analysis of a system of shape-invariant empirical Engel curves using the
573 FES (1995-96) data set within the framework of the EGPLSI model with our pro-
574 posed CF approach is conducted. Not only are the findings interesting empirically
575 but the accessible applicability of our proposed CF approach is also explored.

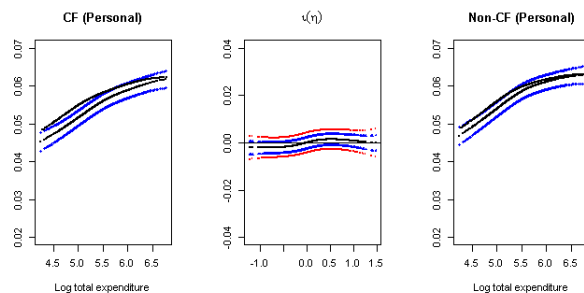
576 **Figure 3.5.** *Leisure and clothing (90% confidence bands drawn for households with children)*



577 Additionally, the development of the CF approach in this paper also provides the
578 foundation for addressing the presence of weak instruments in the EGPLSI model.
579 Han (2011) discussed how the intuitive triangular structure of the CF approach in a
580 simple nonparametric regression model translates the difficult problem (the presence
581 of weak instruments in a reduced-form equation) into a much simpler one, particu-
582 larly the multicollinearity problem in a structural equation. Hence it is plausible to

583 develop the current paper further to the presence of weak instruments case. How-
 584 ever, a thorough investigation is required to examine a number of important issues,
 585 particularly examining the \sqrt{n} -consistent estimation of α_0 and β_0 , and the proper-
 586 ties of the smoothing parameters in each stage of an estimation procedure, and how
 587 to address the presence of weak instruments in the EGPLSI model.

588 **Figure 3.6.** *Engel curves for personal (90% confidence bands drawn for households with children)*



589 References

- 590 Ai, C., Chen, X., 2003. Efficient estimation of models with conditional moment
 591 restrictions containing unknown functions. *Econometrica* 71 (6), 1795–1843.
- 592 Blundell, R., Chen, X., Kristensen, D., 2007. Semi-nonparametric iv estimation of
 593 shape-invariant engel curves. *Econometrica* 75 (6), 1613–1669.
- 594 Blundell, R., Duncan, A., Pendakur, K., 1998. Semiparametric estimation and con-
 595 sumer demand. *Journal of Applied Econometrics*, 435–461.
- 596 Blundell, R., Powell, J. L., 2003. Endogeneity in nonparametric and semiparametric
 597 regression models. *Econometric society monographs* 36, 312–357.
- 598 Blundell, R., Stoker, T. M., 2007. Models of aggregate economic relationships that
 599 account for heterogeneity. *Handbook of Econometrics* 6, 4609–4666.
- 600 Blundell, R. W., Powell, J. L., 2004. Endogeneity in semiparametric binary response
 601 models. *The Review of Economic Studies* 71 (3), 655–679.
- 602 Byrd, R. H., Lu, P., Nocedal, J., Zhu, C., 1995. A limited memory algorithm for
 603 bound constrained optimization. *SIAM Journal on Scientific Computing* 16 (5),
 604 1190–1208.

- 605 Carroll, R. J., Fan, J., Gijbels, I., Wand, M. P., 1997. Generalized partially linear
606 single-index models. *Journal of the American Statistical Association* 92 (438),
607 477–489.
- 608 Eubank, R. L., Speckman, P. L., 1993. Confidence bands in nonparametric regres-
609 sion. *Journal of the American Statistical Association* 88 (424), 1287–1301.
- 610 Gao, J., 2007. *Nonlinear Time Series: Semiparametric and Nonparametric Methods*.
611 CRC Press.
- 612 Gao, J., Lu, Z., Tjøstheim, D., et al., 2006. Estimation in semiparametric spatial
613 regression. *The Annals of Statistics* 34 (3), 1395–1435.
- 614 Han, S., 2011. Nonparametric triangular simultaneous equations models with weak
615 instruments.
- 616 Hansen, B. E., 2008. Uniform convergence rates for kernel estimation with dependent
617 data. *Econometric Theory* 24 (3), 726–748.
- 618 Härdle, W., Hall, P., Ichimura, H., 1993. Optimal smoothing in single-index models.
619 *The Annals of Statistics* 21 (1), 157–178.
- 620 Härdle, W., Marron, J. S., 1990. Semiparametric comparison of regression curves.
621 *The Annals of Statistics* 18 (1), 63–89.
- 622 Hastie, T., Tibshirani, R., 1990. *Generalized Additive Models*. John Wiley & Sons,
623 Inc.
- 624 Ichimura, H., 1993. Semiparametric least squares (sls) and weighted sls estimation
625 of single-index models. *Journal of Econometrics* 58 (1-2), 71–120.
- 626 Li, Q., Racine, J. S., 2007. *Nonparametric Econometrics: Theory and Practice*.
627 Princeton University Press.
- 628 Li, Q., Wooldridge, J. M., 2002. Semiparametric estimation of partially linear models
629 for dependent data with generated regressors. *Econometric Theory* 18 (3), 625–
630 645.

- 631 Linton, O., Nielsen, J. P., 1995. A kernel method of estimating structured nonpara-
632 metric regression based on marginal integration. *Biometrika* 82 (1), 93–100.
- 633 Mammen, E., Rothe, C., Schienle, M., 2016. Semiparametric estimation with gen-
634 erated covariates. *Econometric Theory* 32 (5), 1140–1177.
- 635 Newey, W. K., Powell, J. L., Vella, F., 1999. Nonparametric estimation of triangular
636 simultaneous equations models. *Econometrica* 67 (3), 565–603.
- 637 Pendakur, K., 1999. Semiparametric estimates and tests of base-independent equiv-
638 alence scales. *Journal of Econometrics* 88 (1), 1–40.
- 639 Robinson, P., Pinkse, C., 1995. Pooling nonparametric estimates of regression func-
640 tions with similar shape. In: *Advances in Econometrics and Quantitative Eco-*
641 *nomics*. Wiley-Blackwell, 172–195.
- 642 Robinson, P. M., 1988. Root-n-consistent semiparametric regression. *Econometrica*
643 56 (4), 931–954.
- 644 Su, L., Ullah, A., 2008. Local polynomial estimation of nonparametric simultaneous
645 equations models. *Journal of Econometrics* 144 (1), 193–218.
- 646 Tjøstheim, D., Auestad, B. H., 1994. Nonparametric identification of nonlinear time
647 series: projections. *Journal of the American Statistical Association* 89 (428), 1398–
648 1409.
- 649 Xia, Y., 1998. Bias-corrected confidence bands in nonparametric regression. *Journal*
650 *of the Royal Statistical Society: Series B (Statistical Methodology)* 60 (4), 797–
651 811.
- 652 Xia, Y., Härdle, W., 2006. Semi-parametric estimation of partially linear single-
653 index models. *Journal of Multivariate Analysis* 97 (5), 1162–1184.
- 654 Xia, Y., Tong, H., Li, W., 1999. On extended partially linear single-index models.
655 *Biometrika* 86 (4), 831–842.