Online Appendix to "Optimal Linear Rank Indexes for Latent Variables"*

Yingyao Hu Ji-Liang Shiu Yi Xin Jiaxiong Yao

June 11, 2024

C. Proofs

C.1. Proof of Theorem 2.1

Proof In order to apply the identification method in Hu and Schennach (2008), we first rewrite the integral equation in Equation (2.3) as follows

$$f_{X_1,X_2|X_3}(X_1,X_2|X_3) = \int_{\mathcal{X}^*} f_{X_1|X^*}(X_1|x^*) f_{X_2|X^*}(X_2|x^*) f_{X^*|X_3}(x^*|X_3) dx^*. \tag{C.1}$$

This equation is equivalent to the one in Theorem 1 of Hu and Schennach (2008) if we interpret the measurements X_2, X_1, and X_3 as the dependent variable Y, a mismeasured covariate X, and an instrumental variable Z respectively. within the context of the measurement error model.

*Hu: Department of Economics, Johns Hopkins University (email: yhu@jhu.edu); Shiu: Institute for Economic and Social Research, Jinan University (email: jishiu.econ@gmail.com); Xin: Division of the Humanities and Social Sciences, California Institute of Technology (email: yixin@caltech.edu); Yao: International Monetary Fund, 700 19th St NW, Washington, DC 20431 (email: jyao@imf.org). The usual disclaimer applies.

Denote $L^2(\mathfrak{X}) = \{h(\cdot) : \int_{\mathfrak{X}} |h(x)|^2 dx < \infty\}$. Given x_2, define operators as follows[1]:

$$\begin{aligned} &L_{f_{X_1,X_2|X_3}} : L^2(\mathscr{X}_3) \to L^2(\mathscr{X}_1) \text{ with} \\ &\qquad [L_{f_{X_1,X_2|X_3}} h](x_1) = \int f_{X_1,X_2|X_3}(x_1, x_2|x_3) h(x_3) dx_3, \end{aligned} \tag{C.2}$$

$$\begin{aligned} &L_{f_{X_1|X^*}} : L^2(\mathscr{X}^*) \to L^2(\mathscr{X}_1) \text{ with} \\ &\qquad [L_{f_{X_2|X^*}} h](x_1) = \int f_{X_1|X^*}(x_1|x^*) h(x^*) dx^*, \end{aligned} \tag{C.3}$$

$$\begin{aligned} &\Delta_{f_{X_2|X^*}} : L^2(\mathscr{X}^*) \to L^2(\mathscr{X}^*) \text{ with} \\ &\qquad [\Delta_{f_{X_2|X^*}} h](x^*) = f_{X_2|X^*}(x_2|x^*) h(x^*), \end{aligned} \tag{C.4}$$

$$\begin{aligned} &L_{f_{X^*|X_3}} : L^2(\mathscr{X}_3) \to L^2(\mathscr{X}^*) \text{ with} \\ &\qquad [L_{f_{X^*|X_3}} h](x^*) = \int f_{X^*|X_3}(x^*|x_3) h(x_3) dx_3, \end{aligned} \tag{C.5}$$

$$\begin{aligned} &L_{f_{X_1|X_3}} : L^2(\mathscr{X}_3) \to L^2(\mathscr{X}_1) \text{ with} \\ &\qquad [L_{f_{X_1|X_3}} h](x_1) = \int f_{X_1|X_3}(x_1|x_3) h(x_3) dx_3. \end{aligned} \tag{C.6}$$

For an arbitrary $h \in L^2(\mathscr{X}_3)$, using the definition of operators in Equations (C.2)–(C.5) and an interchange of integrations, we rewrite Equation (C.1) as an operator equivalence rela-

[1]The requirement for the operators defined in Equations (C.2)–(C.5) to be well-defined in these L^2-spaces is that their kernel functions are in L^2-spaces. Given x_2, assume $f(x_1, x_2|x_3) \in L^2(\mathscr{X}_1 \times \mathscr{X}_3)$, $f(x_1|x^*) \in L^2(\mathscr{X}_1 \times \mathscr{X}^*)$, $f(x_2|x^*) \in L^2(\mathscr{X}^*)$, and $f(x^*|x_3) \in L^2(\mathscr{X}^* \times \mathscr{X}_3)$. Then, for $h \in L^2(\mathscr{X}_3)$, by Cauchy-Schwarz inequality, we have

$$\begin{aligned} \int_{\mathscr{X}_1} \left| [L_{f_{X_1,X_2|X_3}} h](x_1) \right|^2 dx_1 &= \int_{\mathscr{X}_1} \left| \int_{\mathscr{X}_3} f(x_1, x_2|x_3) h(x_3) dx_3 \right|^2 dx_1 \\ &\leq \int_{\mathscr{X}_1} \left(\int_{\mathscr{X}_3} |f(x_1, x_2|x_3)|^2 dx_3 \cdot \int_{\mathscr{X}_3} |h(x_3)|^2 dx_3 \right) dx_1 \\ &= \int_{\mathscr{X}_1} \int_{\mathscr{X}_3} |f(x_1, x_2|x_3)|^2 dx_3 dx_1 \cdot \int_{\mathscr{X}_3} |h(x_3)|^2 dx_3 < \infty \end{aligned}$$

Therefore, $[L_{f_{X_1,X_2|X_3}} h](x_1) \in L^2(\mathscr{X}_1)$ and $L_{f_{X_1,X_2|X_3}}$ from $L^2(\mathscr{X}_3)$ to $L^2(\mathscr{X}_1)$ is well-defined and the support of $f(\cdot, x_1|\cdot)$ can be unbounded. We can apply similar arguments to the operators $L_{f_{X_1|X^*}}$, $\Delta_{f_{X_2|X^*}}$, and $L_{f_{X^*|X_3}}$ in Equations (C.2)–(C.5) to show the operators are well-defined in the corresponding L^2-spaces and their supports can be unbounded.

tionship as follows,

(C.7)
$$
\begin{aligned}
&\left[L_{f_{X_1,X_2|X_3}}h\right](x_1)\\
&=\int f_{X_1,X_2|X_3}(x_1,x_2|x_3)h(x_3)dx_3\\
&=\int_{\mathscr{X}_3}\left(\int_{\mathscr{X}^*} f_{X_1|X^*}(x_1|x^*)f_{X_2|X^*}(x_2|x^*)f_{X^*|X_3}(x^*|x_3)dx^*\right)h(x_3)dx_3\\
&=\int_{\mathscr{X}^*} f_{X_1|X^*}(x_1|x^*)f_{X_2|X^*}(x_2|x^*)\left(\int_{\mathscr{X}_3} f_{X^*|X_3}(x^*|x_3)h(x_3)dx_3\right)dx^*\\
&=\int_{\mathscr{X}^*} f_{X_1|X^*}(x_1|x^*)f_{X_2|X^*}(x_2|x^*)[L_{f_{X^*|X_3}}h](x^*)dx^*\\
&=\int_{\mathscr{X}^*} f_{X_1|X^*}(x_1|x^*)[\Delta_{f_{X_2|X^*}}L_{f_{X^*|X_3}}h](x^*)dx^*\\
&=[L_{f_{X_1|X^*}}\Delta_{f_{X_2|X^*}}L_{f_{X^*|X_3}}h](x_2).
\end{aligned}
$$

Thus, we express Equation (C.1) as the operator equivalence relationships:

(C.8)
$$L_{f_{X_1,X_2|X_3}}=L_{f_{X_1|X^*}}\Delta_{f_{X_2|X^*}}L_{f_{X^*|X_3}}.$$

Integrating out X_2 in Equation (C.1) yields

(C.9)
$$f_{X_1|X_3}(X_1|X_3)=\int_{\mathscr{X}^*} f_{X_1|X^*}(X_1|x^*)f_{X^*|X_3}(x^*|X_3)dx^*.$$

This is equivalent to the following operator relationship:

(C.10)
$$L_{f_{X_1|X_3}}=L_{f_{X_1|X^*}}L_{f_{X^*|X_3}}.$$

Since $L_{f_{X_1|X^*}}$ is injective by Assumption 2.3, Equation (C.10) can be written as

(C.11)
$$L_{f_{X^*|X_3}}=L^{-1}_{f_{X_1|X^*}}L_{f_{X_1|X_3}}.$$

The expression in Equation (C.11) for $L_{f_{X^*|X_3}}$ can be substituted into Equation (C.8) to yield

$$L_{f_{X_1,X_2|X_3}} = L_{f_{X_1|X^*}} \Delta_{f_{X_2|X^*}} L^{-1}_{f_{X_1|X^*}} L_{f_{X_1|X_3}}. \tag{C.12}$$

By Hu and Schennach (2008), the injectivity $L_{f_{X_3|X_1}}$ in Assumption 2.3 implies the inverse $L^{-1}_{f_{X_1|X_3}}$ exists and can be applied from the right on each side of Equation (C.12) to yield[2]

$$L_{f_{X_1,X_2|X_3}} L^{-1}_{f_{X_1|X_3}} = L_{f_{X_1|X^*}} \Delta_{f_{X_2|X^*}} L^{-1}_{f_{X_1|X^*}}. \tag{C.13}$$

Equation (C.13) implies that the known operator $L_{f_{X_1,X_2|X_3}} L^{-1}_{f_{X_1|X_3}}$ admits a spectral decomposition that takes the form of an eigenvalue-eigenfunction decomposition where the eigenvalues are given by the multiple operator $\Delta_{f_{X_2|X^*}}$ and the eigenfunctions are given by the kernel of the integral operator $L_{f_{X_1|X^*}}$. To establish uniqueness of the decomposition in Equation (C.13), we need to impose conditions on the eigenvalues $f_{X_2|X^*}$ and the eigenfunctions $f_{X_1|X^*}$ which is analogous to standard matrix diagonalization. Assumption 2.4 ensures that the eigenvalues are distinct, and Assumption 2.5 imposes a location normalization to pin down the values of the unobserved X^* relative to the observed variables. Therefore, the conditional densities $f_{X_1|X^*}$ and $f_{X_2|X^*}$ are identified.

Next, under Assumption 2.1, consider

$$f_{X_1,X_2,X_3,\cdots,X_K} = \int_{\mathscr{X}^*} f_{X_1|X^*} f_{X_2,X_3,\cdots,X_K,X^*} dx^*. \tag{C.14}$$

Following the previous derivation, we obtain the following linear operator relationship, for

[2] See Assumption 3 and Lemma 1 in Hu and Schennach (2008).

each given $(x_3,\cdots,x_K)$

(C.15)
$$L_{f_{X_1,X_2,x_3,\cdots,x_K}} = L_{f_{X_1|X^*}} L_{f_{X_2,x_3,\cdots,x_K,X^*}},$$

where

(C.16)
$$L_{f_{X_1,X_2,x_3,\cdots,x_K}} : L^2(\mathscr{X}_1) \to L^2(\mathscr{X}_2) \text{ with}$$
$$[L_{f_{X_1,X_2,x_3,\cdots,x_K}} h](x_1) = \int f_{X_1,X_2,X_3,\cdots,X_K}(x_1,x_2,x_3,\cdots,x_K) h(x_2) dx_2,$$

(C.17)
$$L_{f_{X_2,x_3,\cdots,x_K,X^*}} : L^2(\mathscr{X}_2) \to L^2(\mathscr{X}^*) \text{ with}$$
$$[L_{f_{X_2,x_3,\cdots,x_K,X^*}} h](x^*) = \int f_{X_2,X_3,\cdots,X_K,X^*}(x_2,x_3,\cdots,x_K,x^*) h(x_2) dx_2,$$

The injectivity of $L_{X_1|X^*}$ in Assumption 2.3 implies that

$$L_{f_{X_2,x_3,\cdots,x_K,X^*}} = L^{-1}_{f_{X_1|X^*}} L_{f_{X_1,X_2,x_3,\cdots,x_K}}.$$

Since $f_{X_1,X_2,x_3,\cdots,x_K}$ is observable and $f_{X_1|X^*}$ is identified from the spectral decomposition, $f_{X_2,X_3,\cdots,X_K,X^*}$ is also identified. The identification of the joint density $f_{X_2,X_3,\cdots,X_K,X^*}$ implies the identification of the conditional densities of the observable measurements given the latent variable X^*, $f_{X_k|X^*}(X_k|X^*)$, for $k = 2,\cdots,K$ and the density of the latent variable $f_{X^*}(X^*)$. $Q.E.D.$

C.2. Proof of Theorem 2.2

Proof After discussing the assumptions preceding Theorem 2.2, what remains to be shown is that Assumptions 2.7 and 2.8 imply Assumption 2.3, specifically the injectivity of $L_{f_{X_1|X^*}}$ and $L_{f_{X_3|X_1}}$. Injectivity can be expressed in terms of completeness of the kernel families.

Definition C.1. *The family* $\{f_{U|V} : v \in \mathcal{V}\}$ *is complete over* $L^2(\mathcal{V})$ *if for any function* $h \in L^2(\mathcal{V})$, $\int f_{U|V}(u|v)h(v)dv = 0$ *for all* $u \in \mathcal{U}$ *implies* $h(v) = 0$ *for almost any* $v \in \mathcal{V}$.

Hu and Shiu (2022) generalize Theorem 2.1 of Mattner (1993) to show a sufficient condition for completeness under additive independence:

Lemma C.1. *Let* f_E *be a pdf of a variable* E. *Consider* $X = D + E$, *where* $D \in \mathbb{R}$ *and* E *is independent of* D. *If the characteristic function of* X *or* E *is nonvanishing everywhere, then the nonparametric family of conditional density functions* $\left\{f(X|D) = f_E(X - D) : X \in \mathcal{X}\right\}$ *is complete in* $L^2(\mathcal{D})$.

First, we show that Assumptions 2.7 and 2.8 imply that $L_{f_{X_1|X^*}}$ is injective. That is,

$$[L_{f_{X_1|X^*}} h](x_1) \quad = \quad \int f_{X_1|X^*}(x_1|x^*)h(x^*)dx^* = 0$$

for all x_1 implies $h(x^*) = 0$ for all x^*. We have

$$\begin{aligned}
[L_{f_{X_1|X^*}} h](x_1) &= \int f_{X_1|X^*}(x_1|x^*)h(x^*)dx^* \\
&= \int f_{\varepsilon_1}(x_1 - g_1(x^*))h(x^*)dx^* \\
&= \int f_{\varepsilon_1}(x_1 - z)h(g_1^{-1}(z))dg_1^{-1}(z) \\
&= \int f_{\varepsilon_1}(x_1 - z)\left(h(g_1^{-1}(z))\frac{1}{g_1'(g_1^{-1}(z))}\right)dz \\
&= 0
\end{aligned}$$

We then have $h(g_1^{-1}(z)) = 0$ for all z, since the family $\left\{f_{\varepsilon_1}(x_1 - z) : x_1 \in \mathcal{X}\right\}$ is complete. This conclusion follows from applying Lemma C.1 under Assumption 2.7. The monotonicity of g_1 in Assumption 2.8 implies that $h(x^*) = 0$ for all x^*. Therefore, $L_{f_{X_1|X^*}}$ is injective. Similarly, $L_{f_{X_3|X^*}}$ is injective.

Next, we show that $L_{f_{X_3|X_1}}$ is injective under Assumptions 2.7 and 2.8. Specifically, we need to show that

$$[L_{f_{X_3|X_1}}h](x_3) = \int f_{X_3|X_1}(x_3|x_1)h(x_1)dx_1 = 0$$

for all x_3 implies $h(x_1) = 0$ for all x_1. Consider

$$\begin{aligned}
& \int f_{X_3|X_1}(x_3|x_1)h(x_1)dx_1 \\
= & \int\int f_{X_3|X_1X^*}(x_3|x_1,x^*)f_{X_1|X^*}(x_1|x^*)f_{X^*}(x^*)dx^*h(x_1)f_{X_1}^{-1}(x_1)dx_1 \\
= & \int\int f_{X_3|X^*}(x_3|x^*)f_{X^*}(x^*)f_{X_1|X^*}(x_1|x^*)dx^*h(x_1)f_{X_1}^{-1}(x_1)dx_1 \\
= & \int\int f_{\varepsilon_3}(x_3-g_3(x^*))f_{X^*}(x^*)f_{\varepsilon_1}(x_1-g_1(x^*))h(x_1)f_{X_1}^{-1}(x_1)dx_1dx^* \\
= & \int f_{\varepsilon_3}(x_3-g_3(x^*))\left(f_{X^*}(x^*)\int f_{\varepsilon_1}(x_1-g_1(x^*))h(x_1)f_{X_1}^{-1}(x_1)dx_1\right)dx^* \\
= & L_{f_{X_3|X^*}}\left(f_{X^*}(\cdot)\int f_{\varepsilon_1}(x_1-g_1(\cdot))h(x_1)f_{X_1}^{-1}(x_1)dx_1\right).
\end{aligned}$$

Since $L_{f_{X_3|X^*}}$ is injective under Assumptions 2.7 and 2.8, we have

$$f_{X^*}(x^*)\int f_{\varepsilon_1}(x_1-g_1(x^*))h(x_1)f_{X_1}^{-1}(x_1)dx_1 = 0 \text{ for any } x^*.$$

Because the range of g_1 is the whole real line, we have $\int f_{\varepsilon_1}(x_1-z)h(x_1)f_{X_1}^{-1}(x_1)dx_1 = 0$ for any z and this implies that

$$\int f_{-\varepsilon_1}(z-x_1)h(x_1)f_{X_1}^{-1}(x_1)dx_1 = 0 \text{ for any } z.$$

The condition that the characteristic function of ε_1 does not vanish on the real line (i.e., Assumption 2.7(iii)) implies that the characteristic function of $-\varepsilon_1$ does not vanish on the real line. Applying Lemma C.1 under Assumption 2.7, we have $\int f_{-\varepsilon_1}(z-x_1)h(x_1)f_{X_1}^{-1}(x_1)dx_1 =$

0 for all z implies that $h(x_1) = 0$ for all x_1. Therefore, $L_{f_{X_3|X_1}}$ is injective. *Q.E.D.*

C.3. Proof of Lemma B.1

Proof We will adopt the following two lemmas in Newey and Powell (2003) for our consistency results of $\widehat{\alpha}_n$.[3]

Lemma C.2. *Suppose that (i) $\mathbb{Q}(\alpha)$ has a unique minimum on $\mathscr{A}$ at α_0; (ii) $\widehat{\mathbb{Q}}_n(\alpha)$ and $\mathbb{Q}(\alpha)$ are continuous, and $\mathscr{A}$ is compact; (iii) $\max_{\alpha\in\mathscr{A}} \left|\widehat{\mathbb{Q}}_n(\alpha) - \mathbb{Q}(\alpha)\right| \xrightarrow{p} 0$; (iv) $\widehat{\mathscr{A}}$ are compact subsets of $\mathscr{A}$ such that for any $\alpha \in \mathscr{A}$ there exists $\hat{\alpha} \in \widehat{\mathscr{A}}$ such that $\hat{\alpha} \xrightarrow{p} \alpha$. Then $\hat{\alpha} = \arg\min_{\alpha\in\widehat{\mathscr{A}}} \widehat{\mathbb{Q}}_n(\alpha) \xrightarrow{p} \alpha_0$.*

Lemma C.3. *If (i) $\mathscr{A}$ is a compact subset of a space with norm $\|\alpha\|_s$; (ii) $\widehat{\mathbb{Q}}_n(\alpha) - \mathbb{Q}(\alpha) \xrightarrow{p} 0$ for all $\alpha \in \mathscr{A}$; (iii) there is $v > 0$ and $B_n = O_p(1)$ such that for all $\alpha, \widetilde{\alpha} \in \mathscr{A}$, $\left|\widehat{\mathbb{Q}}_n(\alpha) - \widehat{\mathbb{Q}}_n(\widetilde{\alpha})\right| \leq B_n \|\alpha - \widetilde{\alpha}\|_s^v$, then $\mathbb{Q}(\alpha)$ is continuous and $\sup_{\alpha\in\mathscr{A}} \left|\widehat{\mathbb{Q}}_n(\alpha) - \mathbb{Q}(\alpha)\right| \xrightarrow{p} 0$.*

We show pointwise consistency of $\widehat{\alpha}_n$ by verifying the conditions of Lemma C.2. Denote the likelihood function as follows

$$\mathscr{L}(\alpha) = E \ln \int_{\mathscr{X}^*} \prod_{k=1}^{K} f_k(X_k - g_k(x^*)) f_{K+1}(x^*) dx^*. \tag{C.18}$$

Under Assumptions 2.6–2.7, the identification result in Theorem 2.2 holds and it implies $\mathbb{Q}(\alpha) \equiv -\mathscr{L}(\alpha)$ has a unique minimum on $\mathscr{A}$ at α_0 which satisfies condition (i) of Lemma C.2. Set $\widehat{\mathbb{Q}}_n(\alpha) \equiv -\widehat{\mathscr{L}_n}(\alpha)$. While the continuity of $\widehat{\mathbb{Q}}_n(\alpha)$ and $\mathbb{Q}(\alpha)$ follow from their formulas, the compactness of $\mathscr{A}$ and $\mathscr{A}_n$ follow from Assumptions B.3 and B.4. We then apply Lemma C.3 to assure condition (iii) of Lemma C.2. The difference between $\widehat{\mathbb{Q}}_n(\alpha)$ and $\mathbb{Q}(\alpha)$ is

$$\left|\widehat{\mathbb{Q}}_n(\alpha) - \mathbb{Q}(\alpha)\right| = \left|\widehat{\mathscr{L}_n}(\alpha) - \mathscr{L}(\alpha)\right|.$$

[3]The result are Lemma A1 and Lemma A2 of Newey and Powell (2003).

By the law of large number, we have $\widehat{\mathscr{L}}_n(\alpha)-\mathscr{L}(\alpha)\xrightarrow{p}0$, which implies that $\widehat{\mathbb{Q}}_n(\alpha)-\mathbb{Q}(\alpha)\xrightarrow{p}0$ for all $\alpha\in\mathscr{A}$. We also need to verify Hölder continuity of the sample objective function $\widehat{\mathbb{Q}}_n(\alpha)$, which is equivalent to Hölder continuity of $\ell(\boldsymbol{x_i};\alpha)$. Set

$$
\begin{aligned}
\alpha &= (\sqrt{f_1(\cdot-\cdot)},\sqrt{f_2(\cdot-g_2(\cdot))},\cdots,\sqrt{f_K(\cdot-g_K(\cdot))},\sqrt{f_{K+1}(\cdot)})\\
\breve{\alpha} &= (\sqrt{\breve{f}_1(\cdot-\cdot)},\sqrt{\breve{f}_2(\cdot-\breve{g}_2(\cdot))},\cdots,\sqrt{\breve{f}_K(\cdot-\breve{g}_K(\cdot))},\sqrt{\breve{f}_{K+1}(\cdot)}).
\end{aligned}
$$

The difference of $\ell(\boldsymbol{x_i};\alpha)$ at α and $\breve{\alpha}$ is given by

$$
\ell(\boldsymbol{x_i};\alpha)-\ell(\boldsymbol{x_i};\breve{\alpha})=\frac{d}{dt}\ell(\boldsymbol{x_i};\bar{\alpha}+t(\alpha-\breve{\alpha}))\Big|_{t=0}, \tag{C.19}
$$

where $\bar{\alpha}=(\sqrt{\bar{f}_1(\cdot-\cdot)},\sqrt{\bar{f}_2(\cdot-\bar{g}_2(\cdot))},\ldots,\sqrt{\bar{f}_K(\cdot-\bar{g}_K(\cdot))},\sqrt{\bar{f}_{K+1}(\cdot)})$, a mean value between α and $\breve{\alpha}$. Denote $\bar{\varepsilon}_k=x_k-\bar{g}_k(x^*)$, $\varepsilon_k=x_k-g_k(x^*)$, and $\breve{\varepsilon}_k=x_k-\breve{g}_k(x^*)$. Equation (C.19) equals

$$
\begin{aligned}
\frac{d}{dt}\ell(\boldsymbol{x_i};\bar{\alpha}+t(\alpha-\breve{\alpha}))\Big|_{t=0} &= \frac{1}{f(\boldsymbol{x_i};\bar{\alpha})}\Big(\sum_{l=1}^{K}\int_{\mathscr{X}^*}2\sqrt{\bar{f}_l(\bar{\varepsilon}_l)}\Big(\sqrt{f_l(\varepsilon_l)}-\sqrt{\breve{f}_l(\breve{\varepsilon}_l)}\Big)\prod_{\substack{k=1\\k\neq l}}^{K}\bar{f}_k(\bar{\varepsilon}_k)\bar{f}_{K+1}(x^*)dx^*\\
&\quad+\int_{\mathscr{X}^*}\prod_{k=1}^{K}\bar{f}_k(\bar{\varepsilon}_k)2\sqrt{\bar{f}_{K+1}(x^*)}\Big(\sqrt{f_{K+1}(x^*)}-\sqrt{\breve{f}_{K+1}(x^*)}\Big)dx^*\Big).
\end{aligned}
$$

We then obtain the bound for the Hölder continuity:

$$
\begin{aligned}
&\Big|\frac{d}{dt}\ell(\boldsymbol{x_i};\bar{\alpha}+t(\alpha-\breve{\alpha}))\Big|_{t=0}\Big|\\
&\le\frac{1}{|f(\boldsymbol{x_i};\bar{\alpha})|}\Big(\sum_{l=1}^{K}\int_{\mathscr{X}^*}\Big|2\sqrt{\bar{f}_l(\bar{\varepsilon}_l)}\omega^{-1}(x_l,x^*)\Big(\Big(\sqrt{f_l(\bar{\varepsilon}_l)}-\sqrt{\breve{f}_l(\bar{\varepsilon}_l)}\Big)\omega(x_l,x^*)\Big)\times\prod_{\substack{k=1\\k\neq l}}^{K}\bar{f}_k(\bar{\varepsilon}_k)\bar{f}_{K+1}(x^*)\Big|dx^*\\
&\qquad+\int_{\mathscr{X}^*}\Big|\prod_{k=1}^{K}\bar{f}_k(\bar{\varepsilon}_k)2\sqrt{\bar{f}_{K+1}(x^*)}\omega^{-1}(x^*)\Big(\Big(\sqrt{f_{K+1}}(x^*)-\sqrt{\breve{f}_{K+1}}(x^*)\Big)\omega(x^*)\Big)\Big|dx^*\Big)\\
&\le\frac{1}{|f(\boldsymbol{x_i};\bar{\alpha})|}\Big(\sum_{l=1}^{K}\int_{\mathscr{X}^*}\Big|2\sqrt{\bar{f}_l(\bar{\varepsilon}_l)}\omega^{-1}(x_l,x^*)\prod_{\substack{k=1\\k\neq l}}^{K}\bar{f}_k(\bar{\varepsilon}_k)\bar{f}_{K+1}(x^*)\Big|dx^*
\end{aligned}
$$

$$+\int_{\mathscr{X}^*}\Big|\prod_{k=1}^{K}\bar{f}_k(\bar{\varepsilon}_k)2\sqrt{\bar{f}_{K+1}(x^*)}\omega^{-1}(x^*)\Big|dx^*\Big)\|\alpha-\tilde{\alpha}\|_s$$

$$\equiv h_1(\boldsymbol{x_i},\bar{\alpha},\bar{\omega})\|\alpha-\breve{\alpha}\|_s.$$

Assumption B.5(ii) ensures that the function h_1 is dominated by a l^2 integrable function under expectation and $\ell(\boldsymbol{x_i};\alpha)$ is Hölder continuous. Therefore, we obtain the consistency of $\hat{\alpha}_n$. $Q.E.D.$

C.4. Proof of Theorem B.1

Proof We prove the results by checking the conditions in Theorem 3.1 in Ai and Chen (2003). The assumptions in Theorem 3.1 in Ai and Chen (2003) are directly being assumed in our single-step sieve MLE. We obtain the consistency result with a $n^{-1/4}$ convergence rate. $Q.E.D.$

D. Additional Tables

Table D.1: Simulation Results for Models with Linear g_k (n=1000)

	Infeasible			ICA			PCA			Linear Rank		
DGP 1	$(\beta_{01},\beta_{02},\beta_{03})=(0.356,0.555,0.089)$											
	β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3
Mean	0.356	0.555	0.089	0.127	0.093	0.104	0.212	0.198	0.590	0.335	0.588	0.077
Median	0.356	0.556	0.089	0.213	0.144	0.152	0.212	0.198	0.590	0.337	0.583	0.077
Bias	0.000	0.000	0.000	-0.229	-0.462	0.015	-0.144	-0.357	0.501	-0.021	0.033	-0.012
Std. Dev.	0.021	0.022	0.013	0.559	0.560	0.582	0.013	0.011	0.020	0.079	0.087	0.013
DGP 2	$(\beta_{01},\beta_{02},\beta_{03})=(0.953,0.017,0.030)$											
	β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3
Mean	0.950	0.019	0.031	-0.161	0.573	-0.038	0.137	0.731	0.133	0.923	0.030	0.047
Median	0.951	0.020	0.030	-0.266	0.942	-0.054	0.147	0.716	0.138	0.927	0.026	0.044
Bias	-0.003	0.002	0.001	-1.114	0.556	-0.068	-0.816	0.714	0.103	-0.030	0.013	0.017
Std. Dev.	0.015	0.016	0.012	0.263	0.755	0.076	0.063	0.133	0.071	0.042	0.021	0.029
DGP 3	$(\beta_{01},\beta_{02},\beta_{03})=(0.481,0.039,0.480)$											
	β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3
Mean	0.481	0.039	0.480	0.072	0.032	0.090	0.217	0.566	0.217	0.487	0.031	0.482
Median	0.481	0.039	0.480	0.078	0.142	0.088	0.217	0.566	0.217	0.482	0.031	0.485
Bias	-0.003	0.002	0.001	-1.114	0.556	-0.068	-0.816	0.714	0.103	-0.030	0.013	0.017
Std. Dev.	0.021	0.009	0.021	0.404	0.814	0.400	0.008	0.014	0.008	0.114	0.004	0.114
DGP 4	$(\beta_{01},\beta_{02},\beta_{03})=(0.236,0.003,0.762)$											
	β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3
Mean	0.236	0.003	0.761	-0.032	0.641	-0.046	0.046	0.919	0.035	0.271	0.006	0.723
Median	0.236	0.008	0.760	-0.071	0.986	-0.101	0.046	0.919	0.035	0.255	0.007	0.738
Bias	-0.003	0.002	0.001	-1.114	0.556	-0.068	-0.816	0.714	0.103	-0.030	0.013	0.017
Std. Dev.	0.018	0.011	0.021	0.198	0.707	0.217	0.010	0.015	0.006	0.165	0.003	0.167
DGP 5	$(\beta_{01},\beta_{02},\beta_{03})=(0.485,0.484,0.030)$											
	β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3
Mean	0.485	0.485	0.030	0.110	0.102	0.095	0.187	0.187	0.627	0.484	0.490	0.026
Median	0.485	0.484	0.030	0.165	0.152	0.173	0.187	0.187	0.626	0.487	0.486	0.026
Bias	-0.003	0.002	0.001	-1.114	0.556	-0.068	-0.816	0.714	0.103	-0.030	0.013	0.017
Std. Dev.	0.021	0.021	0.007	0.569	0.561	0.575	0.009	0.009	0.017	0.119	0.119	0.003

Note: The population quantity $\boldsymbol{\beta}_0$ for each dgp is approximated by computing the mean of the infeasible estimator with a sample size $N = 5000$ in 1000 repetitions.

Table D.2: Percentage of Rankings Correctly Predicted for Models with Linear g_k (n=1000)

	Infeasible	ICA	PCA	Linear Rank
DGP 1	0.830	0.562	0.721	0.827
	(0.006)	(0.202)	(0.012)	(0.007)
DGP 2	0.925	0.637	0.802	0.888
	(0.003)	(0.177)	(0.016)	(0.083)
DGP 3	0.895	0.550	0.746	0.892
	(0.004)	(0.192)	(0.010)	(0.005)
DGP 4	0.846	0.605	0.676	0.831
	(0.006)	(0.117)	(0.012)	(0.047)
DGP 5	0.895	0.566	0.714	0.892
	(0.004)	(0.237)	(0.012)	(0.005)

Note: Standard deviations are reported in parenthesis, which are computed using estimates across 1000 simulations.

Table D.3: Simulation Results for Models with Non-linear g_k (n=1000)

	Infeasible			ICA			PCA			Linear Rank		
DGP 6	$(\beta_{01},\beta_{02},\beta_{03}) = (0.265, 0.639, 0.096$											
	β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3
Mean	0.264	0.641	0.096	-0.079	0.760	0.555	0.152	0.232	0.616	0.225	0.637	0.138
Median	0.264	0.640	0.096	-0.081	0.802	0.586	0.152	0.232	0.615	0.223	0.627	0.138
Bias	-0.001	0.002	0.000	-0.344	0.121	0.459	-0.113	-0.407	0.520	-0.040	-0.002	0.042
Std. Dev.	0.021	0.023	0.011	0.060	0.260	0.194	0.009	0.010	0.008	0.057	0.068	0.033
DGP 7	$(\beta_{01},\beta_{02},\beta_{03}) = (0.268, 0.596, 0.137)$											
	β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3
Mean	0.264	0.599	0.137	-0.056	0.203	0.778	0.148	0.250	0.602	0.229	0.616	0.155
Median	0.263	0.600	0.137	-0.075	0.233	0.959	0.149	0.251	0.600	0.224	0.618	0.150
Bias	-0.001	0.002	0.000	-0.344	0.121	0.459	-0.113	-0.407	0.520	-0.040	-0.002	0.042
Std. Dev.	0.022	0.022	0.013	0.145	0.191	0.541	0.010	0.008	0.015	0.048	0.061	0.033
DGP 8	$(\beta_{01},\beta_{02},\beta_{03}) = (0.083, 0.784, 0.133)$											
	β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3
Mean	0.083	0.785	0.132	0.107	0.520	0.536	0.575	0.213	0.212	0.119	0.653	0.228
Median	0.083	0.785	0.132	0.105	0.965	0.246	0.577	0.213	0.209	0.116	0.646	0.232
Bias	-0.001	0.002	0.000	-0.344	0.121	0.459	-0.113	-0.407	0.520	-0.040	-0.002	0.042
Std. Dev.	0.009	0.017	0.015	0.079	0.503	0.415	0.022	0.012	0.022	0.029	0.066	0.052
DGP 9	$(\beta_{01},\beta_{02},\beta_{03}) = (0.245, 0.650, 0.105)$											
	β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3
Mean	0.245	0.651	0.104	-0.060	0.843	0.431	0.231	0.267	0.502	0.197	0.662	0.140
Median	0.244	0.651	0.104	-0.063	0.886	0.451	0.233	0.265	0.501	0.197	0.655	0.136
Bias	0.000	0.001	-0.001	-0.305	0.193	0.326	-0.014	-0.383	0.397	-0.048	0.012	0.035
Std. Dev.	0.018	0.021	0.013	0.082	0.273	0.143	0.019	0.020	0.011	0.066	0.082	0.045

Note: The population quantity $\boldsymbol{\beta}_0$ for each dgp is approximated by computing the mean of the infeasible estimator with a sample size $N = 5000$ in 1000 repetitions.

Table D.4: Percentage of Rankings Correctly Predicted for Models with Non-linear g_k (n=1000)

	Infeasible	ICA	PCA	Linear Rank
DGP 6	0.852	0.771	0.776	0.847
	(0.006)	(0.093)	(0.008)	(0.008)
DGP 7	0.873	0.733	0.811	0.871
	(0.005)	(0.154)	(0.007)	(0.005)
DGP 8	0.880	0.809	0.793	0.872
	(0.005)	(0.082)	(0.008)	(0.008)
DGP 9	0.849	0.779	0.784	0.842
	(0.006)	(0.088)	(0.008)	(0.009)

Note: Standard deviations are reported in parenthesis, which are computed using estimates across 1000 simulations.

Table D.5: Measurements of Economic Activity and Input for Linear Rank Indexes

Measures	Frequency	Data Source	Data Construction
GDP Growth	Annual, Quarterly	WEO, Haver Analytics	
Luminosity			
DMSP-OLS	Annual	EOG	Hu and Yao (2021)
VIIRS	Monthly	EOG	Hu and Yao (2021)
CO2 Emissions	Annual	Global Carbon Project	Canadell (2003)
Google SVI	Monthly	Google Trends	Narita and Yin (2018)
Samples	**Variable Input**	**Country Coverage**	**Time Coverage**
Sample 1	GDP, DMSP/OLS luminosity, CO2	180 countries	2000-2011
Sample 2	GDP, VIIRS luminosity, CO2	180 countries	2014-2018
Sample 3	GDP, VIIRS luminosity, Google SVI	69 EMDEs	2014Q1-2020Q4

Note: (1). WEO—World Economic Outlook, International Monetary Fund; EOG—Earth Observation Group, Payne Institute, Colorado School of Mines; SVI—search volume index; EMDEs–Emerging markets and developing economies. (2). DMSP/OLS (The Defense Meteorological Satellite Program Operational Line-Scan System) and VIIRS (Visible Infrared Imaging Radiometer Suite) are two different satellite systems.

References

AI, C., AND X. CHEN (2003): "Efficient Estimation of Models with Conditional Moment Restrictions Containing Unknown Functions," *Econometrica*, 71(6), 1795–1843.

CANADELL, J. (2003): "Global Carbon Project. Science Framework and Implementation," *Earth System Science Partnership (IGBP, IHDP, WCRP, DIVERSITAS) Report No. 1; Global Carbon Project Report No. 1*.

HU, Y., AND S. SCHENNACH (2008): "Instrumental Variable Treatment of Nonclassical Measurement Error Models," *Econometrica*, 76(1), 195–216.

HU, Y., AND J.-L. SHIU (2022): "A Simple Test of Completeness in a Class of Nonparametric Specification," *Econometric Reviews*, 41(4), 373–399.

HU, Y., AND J. YAO (2021): "Illuminating Economic Growth," *Journal of Econometrics*.

MATTNER, L. (1993): "Some Incomplete but Boundedly Complete Location Families," *The Annals of Statistics*, 21(4), 2158–2162.

NARITA, M. F., AND R. YIN (2018): "In Search of Information: Use of Google Trends? Data to Narrow Information Gaps for Low-income Developing Countries," IMF Working Paper No. 18/286.

NEWEY, W., AND J. POWELL (2003): "Instrumental Variable Estimation of Nonparametric Models," *Econometrica*, 71(5), 1565–1578.