Nonparametric Identification Using Instrumental Variables: Sufficient Conditions For Completeness*

Yingyao Hu† Ji-Liang Shiu‡

Johns Hopkins University National Chung-Cheng University

October 18, 2012

Abstract

This paper provides sufficient conditions for the nonparametric identification of the regression function $m(\cdot)$ in a regression model with an endogenous regressor x and an instrumental variable z. It has been shown that the identification of the regression function from the conditional expectation of the dependent variable on the instrument relies on the completeness of the distribution of the endogenous regressor conditional on the instrument, i.e., $f(x|z)$. We provide sufficient conditions to extend known complete distribution functions to complete nonparametric families without imposing a specific functional form. We show that if the conditional density $f(x|z)$ can form a linearly independent sequence and coincides with an known complete density at a limit point in the support of z, then $f(x|z)$ itself is complete, and therefore, the regression function $m(\cdot)$ is nonparametrically identified. We use this general result to provide specific sufficient conditions for completeness in three different specifications of the relationship between the endogenous regressor x and the instrumental variable z.

Keywords: *nonparametric identification, instrumental variable, completeness, endogeneity.*

---

*We thank the participants at the International Economic Association 16^{th} World Congress, Bejing, China, for various suggestions regarding this paper. The usual disclaimer applies.

†Department of Economics, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218. Email: yhu@jhu.edu.

‡Department of Economics, National Chung-Cheng University, 168 University Rd. Min-Hsiung Chia-Yi, Taiwan. Email: jishiu@ccu.edu.tw.

1. Introduction

We consider a nonparametric regression model as follows:

$$y = m(x) + u, \tag{1}$$

where the regressor x may be correlated with a zero mean regression error u. The parameter of interest is the nonparametric regression function $m(\cdot)$. An instrumental variable z is conditional mean independent of the regression error u, i.e., $E(u|z) = 0$, which implies

$$E[y|z] = \int_{-\infty}^{+\infty} m(x) f(x|z) dx, \tag{2}$$

where the probability measure of x conditional on z is absolutely continuous w.r.t. the Lebesgue measure. We observe a random sample of $\{y, x, z\}$, which take values in the state spaces $\mathcal{Y}$, $\mathcal{X}$ and $\mathcal{Z}$, respectively. This paper provides sufficient conditions on the conditional density $f(x|z)$ under which the regression function $m(\cdot)$ is nonparametrically identified from, i.e., uniquely determined by, the observed conditional mean $E[y|z]$. We show that if the conditional density $f(x|z)$ can form a linearly independent sequence and coincides with an existing complete density at a limit point in the support of z under appropriate assumptions, then $f(x|z)$ itself is complete, and consequently, the regression function $m(\cdot)$ is nonparametrically identified. Our sufficient conditions for completeness impose no specific functional form on $f(x|z)$, such as the exponential family.

We assume the regression function $m(\cdot)$ is in a Hilbert space $\mathcal{H}$ of functions defined on $\mathcal{X}$ the support of regressor x. This paper considers a weighted L^2 space $\mathcal{L}^2(\mathcal{X}, \omega) = \{h(\cdot) : \int_{\mathcal{X}} |h(x)|^2 \omega(x) dx < \infty\}$ with the inner product $\langle f, g\rangle \equiv \int_{\mathcal{X}} f(x)g(x)\omega(x)dx$, where the positive weight function $\omega(x)$ is bounded almost everywhere and $\int_{\mathcal{X}} \omega(x)dx < \infty$. The corresponding norm is defined as: $\|f\|^2 = \langle f, f\rangle$. The completion of $\mathcal{L}^2(\mathcal{X}, \omega)$ under the norm $\|\cdot\|$ is a Hilbert space.

One may show that the uniqueness of the regression function $m(\cdot)$ is implied by the completeness of the family $\{f(\cdot|z) : z \in \mathcal{O}\}$ in $\mathcal{H}$, where $\mathcal{O} \subseteq \mathcal{Z}$ is a subset of $\mathcal{Z}$ the support of z. The set $\mathcal{O}$ may be $\mathcal{Z}$ itself or some subset of $\mathcal{Z}$. In particular, this paper considers the completeness with the set $\mathcal{O}$ being a sequence $\{z_k : k = 1, 2, 3, ...\}$ in $\mathcal{Z}$. This case corresponds to a sequence of functions $\{f(\cdot|z_k) : k = 1, 2, ...\}$. We start with the definition of the completeness in a Hilbert space $\mathcal{H}$.

Definition 1. *Denote $\mathcal{H}$ as a Hilbert space. The family $\{f(\cdot|z) \in \mathcal{H} : z \in \mathcal{O}\}$ for some set $\mathcal{O} \subseteq \mathcal{Z}$ is said to be complete in $\mathcal{H}$ if for any $h(\cdot) \in \mathcal{H}$*

$$\int_{\mathcal{X}} h(x)f(x|z)dx = 0 \quad \text{for all } z \in \mathcal{O}$$

implies $h(\cdot) = 0$ almost everywhere in $\mathcal{X}$. When it is a conditional density function defined on $\mathcal{X} \times \mathcal{Z}$, $f(x|z)$ is said to be a complete density.[1]

The completeness introduced in Definition 1 is close to L^2- completeness considered in Andrews (2011) with $\mathcal{H} = \mathcal{L}^2(\mathcal{X}, f_x)$, where the density f_x may be considered as the weight function ω in $\mathcal{L}^2(\mathcal{X}, \omega)$[2]. Andrews (2011) constructs broad (nonparametric) classes of L^2-complete distributions that can have any marginal distributions and a wide range of strengths of dependence. Depending on which regularity conditions are imposed on the regression function $m(\cdot)$, a different versions of completeness can be also considered. For example, D'Haultfoeuille (2011) considers bounded completeness in a nonparametric model between the two variables with an additive separability and a large support condition. Regardless of whether the support $\mathcal{X}$ is bounded or unbounded, such as the unit interval $[0, 1]$ or the real line $\mathbb{R}$, respectively, the completeness in $\mathcal{L}^2(\mathcal{X}, \omega)$ is more informative for identification than the bounded completeness because a bounded function always belongs to the weighted L^2 space $\mathcal{L}^2(\mathcal{X}, \omega)$.[3] Therefore, we consider L^2-completeness with a Hilbert space $\mathcal{H} = \mathcal{L}^2(\mathcal{X}, \omega)$ in this paper.

In the extreme case where x and z are discrete, completeness is the same as a no-perfect-collinearity or a full rank condition on a finite number of distributions of x conditional on different values of z. Our results for continuous variables extend this interpretation. Suppose that the family of conditional distributions in $\{f(\cdot|z_k) : k = 1, 2, ...\}$ is complete in $\mathcal{L}^2(\mathcal{X}, \omega)$. Our linear independence interpretation implies that (1) there is no exact linear relationship

[1]The conditional density function $f(x|z)$ has a two dimensional variation from x and z and we treat it as a special class of the function form $f(x, z)$ which can has a support like $\mathcal{X} \times \mathcal{Z}$.

[2]This is under the assumption that the density function f_x exists. Closely related definitions of L^2-completeness can also be found in Florens, Mouchart, and Rolin (1990), Mattner (1996), and San Martin and Mouchart (2007).

[3]In a bounded domain, bounded completeness may also be less informative than L^2-completeness. For instance, consider a function $h(x) = x^{-1/4}$ over (0,1). Bounded completeness can not distinguish the case that the difference of two regression function is $h(x)$, i.e., $h(x) = m(x) - \widetilde{m}(x)$.

among the family of the conditional distribution $\{f(\cdot|z_k) : k = 1, 2, ...\}$ or a conditional distribution at each point z can not be expressed as a linear combination of others, and (2) every function in $\mathcal{L}^2(\mathcal{X}, \omega)$ can be expressed in terms of linear combinations of the conditional distributions in $\{f(\cdot|z_k) : k = 1, 2, ...\}$. In this general continuous case, the linear combination may be a sum of an infinite number of functions.

The $\mathcal{L}^2$ completeness for the nonparametric regression model (1) suggests that identification is achieved among functions whose difference with the true one is square integrable w.r.t. the weighted Lebesgue measure. As an illustration, suppose that $m(x) = \alpha + \beta x$. With completeness in $\mathcal{L}^2(\mathbb{R}, \omega)$, the regression function m can be identified within the set of functions of the form $\{\alpha + \beta x + g(x) : g \in \mathcal{L}^2(\mathbb{R}, \omega)\}$ Therefore, under our framework the functional form of the regression function m may be very flexible. Notice that the function g can't be linear over $\mathbb{R}$ under bounded completeness, which implies that bounded completeness is not enough to distinguish the true linear regression function $m(x) = \alpha + \beta x$ from another linear function $\widetilde{m}(x) = \widetilde{\alpha} + \widetilde{\beta} x$.

The uniqueness (identification) of the regression function $m(\cdot)$ is implied by the completeness of the family $\{f(\cdot|z) : z \in \mathcal{O}\}$ in $\mathcal{H}$ for some set $\mathcal{O} \subseteq \mathcal{Z}$. This sufficient condition may be shown as follows. Suppose that $m(\cdot)$ is not identified so that there are two different functions $m(\cdot)$ and $\widetilde{m}(\cdot)$ in $\mathcal{H}$ which are observationally equivalent, i.e., for any $z \in \mathcal{Z}$

$$E[y|z] = \int_{\mathcal{X}} m(x) f(x|z) dx = \int_{\mathcal{X}} \widetilde{m}(x) f(x|z) dx. \tag{3}$$

We then have for some $h(x) = m(x) - \widetilde{m}(x) \neq 0$

$$\int_{\mathcal{X}} h(x) f(x|z) dx = 0 \text{ for any } z \in \mathcal{Z}$$

which implies that $\{f(\cdot|z) : z \in \mathcal{O}\}$ for any $\mathcal{O} \subseteq \mathcal{Z}$ is not complete in $\mathcal{H}$. Therefore, if $\{f(\cdot|z) : z \in \mathcal{O}\}$ for some $\mathcal{O} \subseteq \mathcal{Z}$ is complete in $\mathcal{H}$, then $m(\cdot)$ is uniquely determined by $E[y|z]$ and $f(x|z)$, and therefore, is nonparametrically identified.

The following two examples of complete $f(x|z)$ are from Newey and Powell (2003) (See their Theorem 2.2 and 2.3 for details.[4]):

[4]Theorem 2.2 and 2.3 in Newey and Powell (2003) do not specify which functional space the completeness is discussed. The definition of the completeness in page 141 of Lehmann (1986) also does not specify the functional space. However, he starts to specify the property of completeness for all bounded functions and call

Example 1: *Suppose that the distribution of x conditional on z is $N(a+bz,\sigma^2)$ for $\sigma^2>0$ and the support of z contains an open set, then $E[h(\cdot)|z]=0$ for any $z\in\mathcal{Z}$ implies $h(x)=0$ almost everywhere in $\mathcal{X}$; equivalently, $\{f(x|z):z\in\mathcal{Z}\}$ is complete .*

Another case where the $\{f(x|z):z\in\mathcal{O}\}$ is complete in $\mathcal{H}$ is that $f(x|z)$ belongs to an exponential family as follows:

Example 2: *Let $f(x|z)=s(x)t(z)\exp[\mu(z)\tau(x)]$, where $s(x)>0$, $\tau(x)$ is one-to-one in x, and support of $\mu(z)$, $\mathcal{Z}$, contains an open set, then $E[h(\cdot)|z]=0$ for any $z\in\mathcal{Z}$ implies $h(x)=0$ almost everywhere in $\mathcal{X}$; equivalently, the family of conditional density functions $\{f(x|z):z\in\mathcal{Z}\}$ is complete.*

These two examples show the completeness of a family $\{f(x|z):z\in\mathcal{O}\}$, where $\mathcal{O}$ is an open set. In order to extend the completeness to general density functions, we further reduce the set $\mathcal{O}$ from an open set to a countable set with a limit point, i.e. a converging sequence in the support $\mathcal{Z}$.

This paper focuses on the sufficient conditions for completeness of a conditional density. These conditions can be used to obtain global or local identification in a variety of models including the nonparametric IV regression model (see Newey and Powell (2003); Darolles, Fan, Florens, and Renault (2011); Hall and Horowitz (2005); Horowitz (2011)), semiparametric IV models (see Ai and Chen (2003); Blundell, Chen, and Kristensen (2007)), nonparametric IV quantile models (see Chernozhukov and Hansen (2005); Chernozhukov, Imbens, and Newey (2007); Horowitz and Lee (2007)), measurement error models (see Hu and Schennach (2008); An and Hu (2009); Carroll, Chen, and Hu (2010); Chen and Hu (2006)), random coefficient models (see Hoderlein, Nesheim, and Simoni (2010)), and dynamic models (see Hu and Shum (2009); Shiu and Hu (2010)), etc. We refer to D'Haultfoeuille (2011) and Andrews (2011) for more complete literature reviews. On the other hand, Canay, Santos, and Shaikh (2011) consider hypothesis testing problem concerns completeness and they show that the completeness condition is, without further restrictions, untestable.

In this paper, we provide sufficient conditions for the completeness of a general conditional density without imposing particular functional forms. If we consider the set $\mathcal{O}$ in Definition 1 as a sequence $\{z_k:k=1,2,3,...\}$ in $\mathcal{Z}$, then the completeness is determined by a sequence of functions $\{f(\cdot|z_k):k=1,2,...\}$. In a Hilbert space, a complete sequence contains a

it boundedly complete in page 144.

subsequence as a basis. Then, we utilize the results related to the stability of bases in Hilbert space (section 10 of chapter 1 in Young (1980)) to show that a linearly independent sequence is complete if its relative deviation from a complete sequence of function is finite. We then show that two sequences of density functions have a finite deviation when they have the same limit. Based on this observation, we may deviate from the existing complete density function without losing the completeness.

We apply the general results to show the completeness in three scenarios. First, we extend Example 1 to a general setting. In particular, we show the completeness of $f(x|z)$ when x and z satisfy for some function $\mu(\cdot)$ and $\sigma(\cdot)$

$$x = \mu(z) + \sigma(z)\varepsilon \text{ with } z \perp \varepsilon.$$

Second, we consider a general control function

$$x = h(z, \varepsilon) \text{ with } z \perp \varepsilon,$$

and provide conditions for completeness of $f(x|z)$ in this case. Third, our results imply that the completeness of a multidimensional conditional density, e.g.,

$$f(x_1, x_2|z_1, z_2),$$

may be reached by the completeness of two conditional densities of lower dimension, i.e.,$f(x_1|z_1)$ and $f(x_2|z_2)$.

This paper is organized as follows: section 2 provides sufficient conditions for completeness; section 3 applies the main results to the three cases with different specifications of the relationship between the endogenous variable and the instrument; section 4 concludes the paper and all the proofs are in the appendix.

2. Sufficient Conditions for Completeness

In this section, we show that a sequence $\{f(\cdot|z_k)\}$ is complete if it can form a linearly independent sequence and coincides with a complete sequence $\{g(\cdot|z_k)\}$ at a limit point z_0. We start with the introduction of two well-known complete families in Examples 1 and 2. Notice that these completeness results are established on an open set $\mathcal{O}$ instead of a countable set with a limit point, i.e., a converging sequence. In order to extend the completeness to a new function $f(x|z)$, we first establish the completeness on a sequence of z_k.

As we will show below, the completeness of an existing sequence $\{g(x|z_k) : k = 1, 2, ...\}$ is essential to show the completeness for a new function $f(x|z)$. An important family of conditional distributions which admit completeness is the exponential family. Many distributions encountered in practice can be put into the form of exponential families, including Gaussian, Poisson, Binomial, and certain multivariate form of these. Another family of conditional distribution which implies completeness is in the form of a translated density function, i.e., $g(x|z) = g(x - z)$.[5]

Based on the existing results, such as in Examples 1 and 2 in the introduction, we may generate complete sequences from the exponential family or a translated density function. We start with an introduction of a complete sequence in the exponential family. Example 2 shows the completeness of the family $\{g(\cdot|z) : z \in \mathcal{O}\}$, where $\mathcal{O}$ is an open set in $\mathcal{Z}$. In the next lemma, we reduce the set $\mathcal{O}$ from an open set to a countable set with a limit point, i.e. a converging sequence in $\mathcal{Z}$.[6]

Lemma 1. *Denote $\mathcal{O}$ as an open set in $\mathcal{Z}$. Let $\{z_k : k = 1, 2, ...\}$ be a sequence of distinct $z_k \in \mathcal{O}$ converging to z_0 in the open set $\mathcal{O}$. Define*

$$g(x|z) = s(x)t(z) \exp\left[\mu(z)\tau(x)\right]$$

on $\mathcal{X} \times \mathcal{Z}$ with $s(\cdot) > 0$ and $t(\cdot) > 0$. Suppose that $g(\cdot|z) \in \mathcal{L}^1(\mathcal{X})$ for $z \in \mathcal{O}$ and

i) $\mu(\cdot)$ is continuous with $\mu'(z_0) \neq 0$;

ii) $\tau(\cdot)$ is monotonic over $\mathcal{X}$.

Then, the sequence $\{g(\cdot|z_k) : k = 1, 2, ...\}$ is complete in $\mathcal{L}^2(\mathcal{X}, \omega)$, where the weight function $\omega(x)$ satisfies $\int_{\mathcal{X}} \frac{s(x)^2 \exp[2(\mu(z_0)\tau(x)+\delta|\tau(x)|)]}{\omega(x)} dx < \infty$ for some $\delta > 0$.

Proof: See the appendix.

[5]The term used here is according to a definition in page 182 of Rudin (1987), where the translate of f is defined as $f(x - z)$ for all x and a given z.

[6]It is important to show the completeness of a family defined on a countable set because all the statistical asymptotics are based on an infinitely countable number of observations, i.e., the sample size approaching infinity, instead of a continuum of observations, for example, all the possible values in an open set.

The restrictions on the weight function is mild and there are many potential candidates. For example, since $\mathcal{O}$ is open and $\mu(\cdot)$ is continuous with $\mu'(z_0) \neq 0$, there exists some $\tilde{z} \in \mathcal{O}$ such that $\mu(z_0)\tau(x) + \delta|\tau(x)| < \mu(\tilde{z})\tau(x)$ for $\delta > 0$. One particular choice of the weight function is $\omega(x) = s(x)^2 \exp[2\mu(\tilde{z})\tau(x)]$.

Another case where the completeness of $g(x|z)$ is well studied is when $g(x|z) = f_\varepsilon(x - z)$, which is usually due to a translation between the endogenous variable x and instrument z as follows

$$x = z + \varepsilon \text{ with } z \perp \varepsilon.$$

Example 1 suggests that the family $\{g(\cdot|z) \in \mathcal{H} : z \in \mathcal{O}\}$ is complete if $\mathcal{O}$ is an open set in $\mathcal{Z}$ and ε is normal. Again, we show the completeness still holds when the set $\mathcal{O}$ is a converging sequence. We summarize the results as follows.

Lemma 2. *Denote $\mathcal{O}$ as an open set in $\mathcal{Z}$. Let $\{z_k : k = 1, 2, ...\}$ be a sequence of distinct $z_k \in \mathcal{O}$ converging to z_0 in the open set $\mathcal{O}$. Define*

$$g(x|z) = f_\varepsilon(x - z)$$

on $\mathbb{R} \times \mathcal{Z}$. Suppose that $g(\cdot|z) \in \mathcal{L}^1(\mathbb{R})$ for $z \in \mathcal{O}$ and the Fourier transform ϕ_ε of f_ε satisfies

$$0 < |\phi_\varepsilon(t)| < Ce^{-\delta|t|} \tag{4}$$

for all $t \in \mathbb{R}$ and some constants C, $\delta > 0$.

Then, the sequence $\{g(\cdot|z_k) : k = 1, 2, ...\}$ is complete in $\mathcal{L}^2(\mathbb{R}, \omega)$, where the weight function $\omega(x)$ satisfies $\int_{\mathcal{X}} \frac{\exp(-2\delta'|x|)}{\omega(x)} dx < \infty$ for some $\delta' \in (0, \delta)$.

Proof: See the appendix.

Equation (4) implies that the characteristic function is not equal to zero on the real line and that the characteristic function has exponentially decaying tails. For example, the distribution of ε may be normal, Cauchy, or their convolutions with other distributions. The restriction on the weight function is also mild. For example, we may pick $\omega(x) = 1/(1 + x^2)$. With the complete sequences explicitly specified in Lemma 1 and 2, we are ready to extend the completeness to a more general conditional density $f(x|z)$. Our sufficient conditions for completeness are summarized as follows:

Theorem 1. *Suppose $f(\cdot|z)$ and $g(\cdot|z)$ are conditional densities. For every $z \in \mathcal{Z}$, let $f(\cdot|z)$ and $g(\cdot|z)$ be in a Hilbert space $\mathcal{H}$ of functions defined on $\mathcal{X}$ with norm $\|\cdot\|$. Suppose that there exists a point z_0 with its open neighborhood $\mathcal{N}(z_0) \equiv \{z \in \mathcal{Z} : \|z - z_0\| < \epsilon$ for some small $\epsilon > 0\} \subseteq \mathcal{Z}$ such that*

i) for every sequence $\{z_k : k = 1, 2, ...\}$ of distinct $z_k \in \mathcal{N}(z_0)$ converging to z_0, the corresponding sequence $\{g(\cdot|z_k) : k = 1, 2, ...\}$ is complete in a Hilbert space $\mathcal{H}$;

ii) $g(\cdot|z)$ is continuous at z_0 and $\|g(\cdot|z_0)\| > 0$ such that the relative deviation $D(z) = \frac{\|f(\cdot|z)-g(\cdot|z)\|}{\|g(\cdot|z)\|}$ is well-defined and Lipschitz continuous in z on $\mathcal{N}(z_0)$ and $f(\cdot|z)$ coincides with $g(\cdot|z)$ at z_0 in $\mathcal{H}$, i.e.,

$$\|f(\cdot|z_0) - g(\cdot|z_0)\| = 0;$$

iii) there exists a sequence $\{z_k : k = 1, 2, ...\}$ of distinct $z_k \in \mathcal{N}(z_0)$ converging to z_0 such that the sequence $\{f(\cdot|z_k) : k = 1, 2, ...\}$ is linearly independent, i.e.,

$$\sum_{i=1}^{I} c_i f(x|z_{k_i}) = 0 \text{ for all } x \in \mathcal{X} \text{ implies } c_i = 0 \text{ for all } I.$$

Then, the family $\{f(\cdot|z) : z \in \mathcal{N}(z_0)\}$ is complete in $\mathcal{H}$.

Proof: See the appendix.

Condition i) provides complete sequences, which may be from Lemma 1 and 2. Condition ii) requires that the relative deviation $D(z) = \frac{\|f(\cdot|z)-g(\cdot|z)\|}{\|g(\cdot|z)\|}$ is Lipschitz continuous in z on $\mathcal{N}(z_0)$. When the Hilbert space $\mathcal{H}$ is the $\mathcal{L}^2(\mathcal{X}, \omega)$, the relative deviation $D(z)$ is Lipschitz continuous if $\|g(\cdot|z_0)\| > 0$ and the first-order derivatives $\frac{\partial}{\partial z} f(\cdot|z)$ and $\frac{\partial}{\partial z} g(\cdot|z)$ are also in $\mathcal{L}^2(\mathcal{X}, \omega)$ for $z \in \mathcal{N}(z_0)$.[7] In the proof of Theorem 1, we show that the Lipschitz continuity of $D(z)$ implies that there exists a sequence $\{z_k\}$ converging to z_0 such that the total deviation from the sequence $\{g(\cdot|z_k)\}$ to $\{f(\cdot|z_k)\}$ is finite, i.e.,

$$\sum_{k=1}^{\infty} \frac{\|f(\cdot|z_k) - g(\cdot|z_k)\|}{\|g(\cdot|z_k)\|} < \infty. \tag{5}$$

Intuitively, this condition implies that the sequence $\{f(\cdot|z_k)\}$ is close to a complete sequence $\{g(\cdot|z_k)\}$ so that the former sequence may also be complete.

The linear independence in condition iii) imposed on $\{f(\cdot|z_k)\}$ implies that there are

[7] The proof of this claim is provided in the appendix.

no redundant terms in the sequence in the sense that no term can be expressed as a linear combination of some other terms. This condition imposes a mild restriction on $f(x|z)$ because Equation $\sum_{i=1}^{I} c_i f(x|z_{k_i}) = 0$ for all $x \in \mathcal{X}$, which implies an infinite number of restrictions on a finite number of constants c_i. When the support of $f(\cdot|z_k)$ is the whole real line for all z_k, a sufficient condition for the linear independence is that

$$\lim_{x \to -\infty} \frac{f(x|z_{k+1})}{f(x|z_k)} = 0 \text{ for all } k, \tag{6}$$

which implies $\lim_{x \to -\infty} \frac{f(x|z_{k+m})}{f(x|z_k)} = 0$ for any $m \geq 1$ and for all x. If $\sum_{i=1}^{I} c_i f(\cdot|z_{k_i}) = 0$ for all $x \in (-\infty, +\infty)$, we may have

$$-c_1 = \sum_{i=1}^{I} c_i \frac{f(x|z_{k_i})}{f(x|z_{k_1})}.$$

The limit of the right-hand side is zero as $x \to -\infty$ so that $c_1 = 0$. Similarly, we may show $c_{2,} c_3, ..., c_I = 0$ for all i by induction. Notice that the exponential family satisfies Equation (6) for appropriate choices of μ, τ and a sequence. When the support $\mathcal{X}$ is bounded, for example, $\mathcal{X} = [0, 1]$, the condition (6) may become

$$\lim_{x \to 0} \frac{f(x|z_{k+1})}{f(x|z_k)} = 0 \text{ for all } k. \tag{7}$$

For example, the Corollary (Müntz) on page 91 in Young (1980) implies that the family of function $\{x^{z_1}, x^{z_2}, x^{z_3}, ...\}$ is complete in $\mathcal{L}^2([0, 1])$ if $\sum\limits_{k=1}^{\infty} \frac{1}{z_k} = \infty$. This family also satisfies the condition (7) for a strictly increasing $\{z_k\}$. For an existing function $g(x|z) > 0$, we may always have $f(x|z) = \frac{f(x|z)}{g(x|z)} \times g(x|z)$. If the existing sequence $\{g(\cdot|z_k)\}$ satisfies Equation (6), i.e., $\lim_{x \to -\infty} \frac{g(x|z_{k+1})}{g(x|z_k)} = 0$, then it is enough to have $0 < \left(\lim_{x \to -\infty} \frac{f(x|z_k)}{g(x|z_k)}\right) < \infty$.

Furthermore, when $f(x|z) = h(x|z) \times g(x|z)$, the condition (6) is implied by $\lim_{x \to -\infty} \frac{g(x|z_{k+1})}{g(x|z_k)} = 0$ and $\left(\lim_{x \to -\infty} \frac{h(x|z_{k+1})}{h(x|z_k)}\right) < \infty$. We may also consider

$$f(x|z) = \lambda(z) h(x|z) + [1 - \lambda(z)] g(x|z). \tag{8}$$

In this case, the conditional density $f(\cdot|z)$ is a mixture of two continuous conditional densities h, g and the weight λ in the mixture depends on z. At the limit point z_0, $f(\cdot|z_0)$ coincides with $g(\cdot|z_0)$ if $\lim_{z_k \to z_0} \lambda(z) = 0$. The linear independence condition in Equation (6) holds when $h(x|z)$ and $g(x|z)$ satisfy

$$\lim_{x \to -\infty} \frac{g(x|z_{k+1})}{g(x|z_k)} = 0 \text{ and } \lim_{x \to -\infty} \frac{h(x|z_k)}{g(x|z_k)} < \infty. \tag{9}$$

The advantage of this condition is that there are only mild restrictions imposed on the functional form of $h(x|z)$ and $\lambda(z)$.

Suppose the function $f(x|z)$ is differentiable with respect to the variable x up to any finite order. We may consider the so-called Wronskian determinant as follows:

$$W(x) = \det \begin{pmatrix} f(x|z_{k_1}) & f(x|z_{k_2}) & ... & f(x|z_{k_I}) \\ f'(x|z_{k_1}) & f'(x|z_{k_2}) & ... & f'(x|z_{k_I}) \\ ... & ... & ... & ... \\ \frac{d^{(I-1)}}{dx^{(I-1)}} f(x|z_{k_1}) & \frac{d^{(I-1)}}{dx^{(I-1)}} f(x|z_{k_2}) & ... & \frac{d^{(I-1)}}{dx^{(I-1)}} f(x|z_{k_I}) \end{pmatrix} \tag{10}$$

If there exists an x_0 such that the determinant $W(x_0) \neq 0$ for every $\{z_{k_i} : i = 1, 2, ..., I\}$, then $\{f(\cdot|z_k)\}$ is linear independent.

Another sufficient condition for the linear independence is that the so-called Gram determinant G_f is not equal to zero for every $\{z_{k_i} : i = 1, 2, ..., I\}$, where $G_f = \det \left(\left[\langle f(\cdot|z_{k_i}), f(\cdot|z_{k_j}) \rangle \right]_{i,j} \right)$. This condition does not require the function has all the derivatives.

We summarize these results on the linear independence as follows:

Lemma 3. *The sequence $\{f(\cdot|z_k)\}$ corresponding to a sequence $\{z_k : k = 1, 2, ...\}$ of distinct $z_k \in \mathcal{N}(z_0)$ converging to z_0 is linearly independent if one of the following conditions hold:*

1) $\sum_{i=1}^{I} c_i f(x|z_{k_i}) = 0$ for all $x \in \mathcal{X}$ implies $c_i = 0$ for all I.

2) for all k, $\lim_{x \to -\infty} \frac{f(x|z_{k+1})}{f(x|z_k)} = 0$ or $\lim_{x \to x_0} \frac{f(x|z_{k+1})}{f(x|z_k)} = 0$ for some x_0;

3) $f(x|z) = \frac{d}{dx} F_0(\mu(z)\tau(x))$ with $\mu'(z_0) \neq 0$, $\tau(0) = 0$, and $\frac{d^k}{dx^k} F_0(0) \neq 0$ for $k = 1, 2, ...$;

4) for every $\{z_{k_i} : i = 1, 2, ..., I\}$, $\det \left(\left[\langle f(\cdot|z_{k_i}), f(\cdot|z_{k_j}) \rangle \right]_{i,j} \right) \neq 0$.

Proof: See the appendix.

In order to illustrate the relationship between the complete sequence $\{g(\cdot|z_k)\}$ and the sequence $\{f(\cdot|z_k)\}$, we present numerical examples of these two functions as follows. Consider $g(x|z) = x^z$ over $\mathcal{L}^2([0, 0.8])$ for $\mathcal{Z} = (\frac{2}{5}, \frac{3}{5})$. We pick $z_k = \frac{1}{2} - \frac{2}{(k+1)^2}$ with $z_k \to z_0 = \frac{1}{2}$. Since

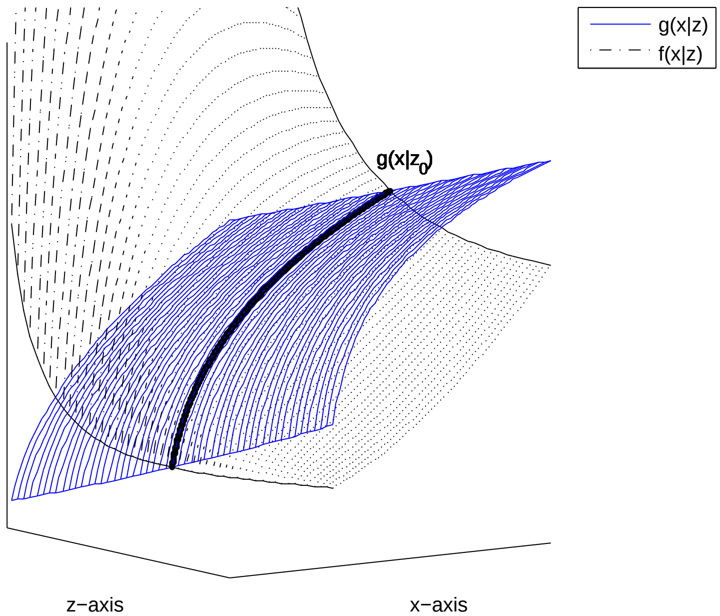

Figure 1: An example of $g(x|z)$ and $f(x|z)$ in Theorem 1.

$\sum_{k=1}^{\infty} \frac{1}{z_k} = \infty$, by the Corollary (Müntz) on page 91 in Young (1980) the family of function $\{x^{z_1}, x^{z_2}, x^{z_3}, ...\}$ is complete in $\mathcal{L}^2([0, 0.8])$. Let $f(x|z) = \left(1 - \frac{2(z-\frac{1}{2})}{(z-0.62)}(x-1)\right) x^z$.[8] Since $\lim_{x \to 0} \frac{f(x|z_{k+1})}{f(x|z_k)} = 0$, our Theorem 1 implies that $\{f(\cdot|z_k)\}$ is also complete in $\mathcal{L}^2([0, 0.8])$ with $g(x|z_0) = f(x|z_0) = \sqrt{x}$. Figure 1 presents a 3D graph of $g(x|z)$ and $f(x|z)$ for (x, z) in $[0, 0.8] \times (\frac{2}{5}, \frac{3}{5})$ to illustrate the relationship between the complete sequence $\{g(\cdot|z_k)\}$ and the sequence $\{f(\cdot|z_k)\}$.

3. Applications

We consider three applications of our main results: first, we show the sufficient conditions for the completeness of $f(x|z)$ when $x = \mu(z) + \sigma(z)\varepsilon$ with $z \perp \varepsilon$; second, we consider the completeness with a general control function $x = h(z, \varepsilon)$; finally, we show how to use our results to transform a multivariate completeness problem to a single variable one.

[8]Choosing such a particular function is only for a suitable illustration in Figure 1.

3.1. Extension of the convolution case

Lemma 2 provides a complete sequence when $x = z + \varepsilon$. Using Theorem 1, we may provide sufficient conditions for the completeness of $f(x|z)$ when the endogenous variable x and the instrument z satisfy a general heterogeneous structure as follows:

$$x = \mu(z) + \sigma(z)\,\varepsilon \text{ with } z \perp \varepsilon.$$

We summarize the result as follows:

Lemma 4. *For every $z \in \mathcal{Z}$, let $f(\cdot|z)$ be in $\mathcal{L}^1(\mathbb{R})$. Let ω be a weight function satisfying the restriction in Lemma 2. Suppose that there exists a point z_0 with its open neighborhood $\mathcal{N}(z_0) \subseteq \mathcal{Z}$ such that*

i) the characteristic function $\phi_{z_0}(t)$ of $f(\cdot|z_0)$ satisfies $0 < |\phi_{z_0}(t)| < Ce^{-\delta|t|}$ for all $t \in \mathbb{R}$ and some constants C, $\delta > 0$;

ii) $\frac{\partial}{\partial z} f(\cdot|z)$ for $z \in \mathcal{N}(z_0)$ and $\frac{\partial}{\partial x} f(\cdot|z_0)$ are in $\mathcal{L}^2(\mathbb{R}, \omega)$;

iii) the function $f(\cdot|z)$ satisfies conditions iii) in Theorem 1, i.e., there exists a sequence $\{z_k : k = 1, 2, ...\}$ of distinct $z_k \in \mathcal{N}(z_0)$ converging to z_0 such that the sequence $\{f(\cdot|z_k) : k = 1, 2, ...\}$ is linearly independent.

Then, the family $\{f(\cdot|z) : z \in \mathcal{N}(z_0)\}$ is complete in $\mathcal{L}^2(\mathbb{R}, \omega)$.

In particular, when

$$f(x|z) = \frac{1}{\sigma(z)} f_\varepsilon\left(\frac{x - \mu(z)}{\sigma(z)}\right)$$

on $\mathbb{R} \times \mathcal{Z}$, we assume

i') the characteristic function $\phi_\varepsilon(t)$ of f_ε satisfies $0 < |\phi_\varepsilon(t)| < Ce^{-\delta|t|}$;

ii') $\mu(\cdot)$, $\sigma(\cdot)$, and $f_\varepsilon(\cdot)$ are continuously differentiable with $\mu'(z_0) \neq 0$, $\sigma(z_0) \neq 0$ and $\int_{-\infty}^{+\infty} |x f_\varepsilon'(x)|^2 \, dx < \infty$;

iii') $\lim_{x \to -\infty} \frac{f_\varepsilon(x-c)}{f_\varepsilon(x)} = 0$ for any constant $c > 0$.

Then, the family $\{f(\cdot|z) : z \in \mathcal{N}(z_0)\}$ is complete in $\mathcal{L}^2(\mathbb{R}, \omega)$.

Proof: See the appendix.

The first part of Lemma 4 implies that one may always make a sequence coincide with a convolution sequence. Consider a sequence $\{f(\cdot|z_k) : k = 1, 2, ...\}$ with a sequence $\{z_k : k = 1, 2, ...\}$

of distinct $z_k \in \mathcal{N}(z_0)$ converging to z_0. We may always generate a convolution sequence $\{g(\cdot|z_k) : k = 1, 2, ...\}$ where

$$g(x|z_k) = f\left(x - \mu\left(z_k\right)|z_0\right) \text{ with } \mu\left(z_0\right) = 0 \text{ and } \mu'\left(z_0\right) \neq 0.$$

Condition i) implies that the sequence satisfies the conditions in Lemma 2 and is complete. Condition ii) guarantees that the first-order derivatives $\frac{\partial}{\partial z} f\left(\cdot|z\right)$ and $\frac{\partial}{\partial z} g\left(\cdot|z\right)$ are in $\mathcal{L}^2\left(\mathcal{X}, \omega\right)$ so that the relative deviation $D(z) = \frac{\|f(\cdot|z) - g(\cdot|z)\|}{\|g(\cdot|z)\|}$ is Lipschitz continuous in z. Since $D(z_0) = 0$ by construction, the condition ii) in Theorem 1 holds. Thus, the completeness holds for $\{f(\cdot|z) : z \in \mathcal{N}(z_0)\}$. In the heterogeneous case where $x = \mu(z) + \sigma(z)\varepsilon$, a primitive condition for the linear independence is that $\lim_{x \to -\infty} \frac{f_\varepsilon(x-c)}{f_\varepsilon(x)} = 0$.

The main purpose of Lemma 4 is to provide sufficient conditions that are easier to check and more accessible than conditions in Theorem 1. While the part (i) of Theorem 1 is a generalized version of the condition (1), the part (iii) of Theorem 1 is equivalent to the condition (2) by Lemma 3. This particular case immediately provide the completeness of the normal distribution with heterogeneity which is complement to the normal distribution with homoskedasticity. Suppose $\varepsilon \sim N(0, 1)$. Then, by Lemma 4 we have the family $\left\{f(x|z) = \frac{1}{\sigma(z)}\phi\left(\frac{x-\mu(z)}{\sigma(z)}\right) : z \in \mathcal{N}(z_0)\right\}$ is complete in $\mathcal{L}^2(\mathbb{R}, \omega)$. This result is new to the literature and provides the identification for models with heterogeneity. On the other hand, consider a conditional density function $f(x|z) = c_z \left(1 - \frac{2(z-\frac{1}{2})}{(z-0.62)}(x-1)\right) x^z$ where c_z is a normalized coefficient. We have shown $\{f(\cdot|z_k) : k = 1, 2, ...\}$ is complete in $\mathcal{L}^2([0, 0.8])$ in the end of Section 2. Therefore, our results have shown many complete DGPs that are not previously known. Another point to emphasize is that we only need this $0 < |\phi_\varepsilon(t)| < Ce^{-\delta|t|}$ restriction at the limit point z_0 not over all z. Any distribution containing a normal factor, say a convolution of normal and another distribution, satisfies this tail restriction.

We may then consider the nonparametric identification of a regression model

$$y = \alpha + \beta x + u, \quad \mathrm{E}[u|z] = 0, \tag{11}$$

with $x = \mu(z) + \sigma(z)\varepsilon$ and $\varepsilon \sim N(0, 1)$. Here the true regression function $m(x)$ is linear, which is unknown to researchers. We have shown that the family $\left\{f(x|z) = \frac{1}{\sigma(z)}\phi\left(\frac{x-\mu(z)}{\sigma(z)}\right) : z \in \mathcal{N}(z_0)\right\}$ is complete in $\mathcal{L}^2(\mathbb{R}, \omega)$, which implies the above linear model is uniquely identified among all the functions in $\mathcal{L}^2(\mathbb{R},\omega)$. Notice that the bounded completeness is not enough for such an identification.

3.2. Completeness with a control function

We then consider a general expression of the relationship between the endogenous variable x and the instrument z. Let a control function describe the relationship between an endogenous variable x and an instrument z as follow:[9]

$$x = h(z,\varepsilon), \text{ with } z \perp \varepsilon. \tag{12}$$

We consider the case where x and ε have the support $\mathbb{R}$. Without loss of generality, we assume ε has a standard normal distribution with the cdf Φ. It is well known that the function h is related to the cdf $F_{x|z}$ as $h(z,\varepsilon) \equiv F_{x|z}^{-1}\left(\Phi\left(\varepsilon\right)|z\right)$ when the inverse of $F_{x|z}$ exists and h is strictly increasing in ε. Given the function h, we are interested in what restrictions on h are sufficient for the completeness of the conditional density $f(x|z)$ implied by Equation (12).

Lemma 5. *Let $\mathcal{N}\left(z_0\right) \subseteq \mathcal{Z}$ be an open neighborhood of some $z_0 \in \mathcal{Z}$ and ω be a weight function satisfying the restriction in Lemma 2. Let Equation (12) hold with $h(z_0,\varepsilon) = \varepsilon$, where ε has a standard normal distribution with the support $\mathbb{R}$. Suppose that*

i) for $z \in \mathcal{N}\left(z_0\right)$, the function $h(z,\varepsilon)$ is strictly increasing in ε and twice differentiable in z and ε;

ii) for $z \in \mathcal{N}\left(z_0\right)$, the functions $f(\cdot|z)$ and $\frac{\partial}{\partial z}f\left(\cdot|z\right)$ are in $\mathcal{L}^2(\mathbb{R},\omega)$, where $f(x|z) = \frac{\partial}{\partial x}F_{\varepsilon}\left(h^{-1}(z,x)\right)$;

iii) there exists a sequence $\{z_k : k = 1,2,...\}$ of distinct $z_k \in \mathcal{N}\left(z_0\right)$ converging to z_0 such that the sequence $\{f(\cdot|z_k) : k = 1,2,...\}$ is linearly independent; in particular, such a sequence exists if for any $\widetilde{z} \neq \widehat{z}$ in $\mathcal{N}\left(z_0\right)$, $\lim_{\varepsilon\to-\infty}\left[h(\widetilde{z},\varepsilon) - h(\widehat{z},\varepsilon)\right] \neq 0$, and $\frac{\partial}{\partial z}h(z_0,\varepsilon) \neq 0$ as $\varepsilon \to -\infty$.

Then, the family $\{f(\cdot|z) : z \in \mathcal{N}\left(z_0\right)\}$ is complete in $\mathcal{L}^2(\mathbb{R},\omega)$.

Proof: See the appendix.

[9] Here we call h the control function without assuming that the IV z is independent of (u,ε) as in the usual control function approach.

Condition i) guarantees that the conditional density $f(x|z)$ is continuous in both x and z around z_0. The condition $h(z_0, \varepsilon) = \varepsilon$ is not restrictive because one may always redefine ε. Therefore, $f(x|z)$ satisfies $f(x|z_0) = f_\varepsilon(x)$, which may be considered as a limit point in the translated family such as $\{g(x|z) = f_\varepsilon(x - \mu(z_k)) : k = 1, 2, ...\}$ with $\mu(z_0) = 0$, i.e. $f(x|z_0) = g(x|z_0)$. We may then use Theorem 1 to show $f(x|z)$ is complete. Condition ii) implies that the deviation $D(z) = \frac{\|f(\cdot|z) - g(\cdot|z)\|}{\|g(\cdot|z)\|}$ is Lipschitz continuous in z. Condition iii) guarantees the linear independence of the sequence $\{f(\cdot|z_k)\}$.

Lemma 5 implies that key sufficient assumptions for the completeness of $f(x|z)$ implied by the control function in Equation (12) is that the control function h is invertible with respect to ε around a limit point in the support of z and linearly independence of the sequence $\{f(\cdot|z_k) : k = 1, 2, ...\}$. Our results may provide sufficient conditions for completeness with a general h. For example, suppose $\frac{\partial}{\partial z}\mu(z_0) < 0$, we may have

$$h(z, \varepsilon) = \mu(z) + e^{z - z_0}\varepsilon + \sum_{j=0}^{J} (z - z_0)^{2j} h_j(\varepsilon),$$

where $h_j(\cdot)$ are increasing functions. The function h may also have a nonseparable form such as

$$h(z, \varepsilon) = \mu(z) + \ln\left[(z - z_0)^2 + \exp(\varepsilon)\right].$$

3.3. Multivariate completeness

When the endogenous variable x and the instrument z are both vectors, our main results in Theorem 1 still applies. In other words, our results can be extended to the multivariate case straightforwardly. In this section, we show that one can use Theorem 1 to reduce a multivariate completeness problem to a single variate one. Without loss of generality, we consider $x = (x_1, x_2)$, $z = (z_1, z_2)$, $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2$, and $\mathcal{Z} = \mathcal{Z}_1 \times \mathcal{Z}_2$. One may show that the completeness of $f(x_1|z_1)$ and $f(x_2|z_2)$ implies that of $f(x_1|z_1) \times f(x_2|z_2)$. Theorem 1 then implies that if conditional density $f(x_1, x_2|z_1, z_2)$ coincides with $f(x_1|z_1) \times f(x_2|z_2)$ at a limit point in $\mathcal{Z}$ then $f(x_1, x_2|z_1, z_2)$ is complete. We summarize the results as follows:

Lemma 6. *For every $z \in \mathcal{Z} = \mathcal{Z}_1 \times \mathcal{Z}_2$, let $f(\cdot|z)$ and $g(\cdot|z)$ be in a Hilbert space $\mathcal{H}$ of functions defined on $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2$ with norm $\|\cdot\|$. Suppose that there exists a point $z_0 = (z_{10}, z_{20})$ with its open neighborhood $\mathcal{N}(z_0) \subseteq \mathcal{Z}$ such that* *i) for every sequence $\{z_k : k = 1, 2, 3, ...\}$ of distinct $z_k \in \mathcal{N}(z_0)$ converging to z_0, the corresponding sequence $\{f_{x_1|z_1}(\cdot|z_{1k}) : k = 1, 2, 3, ...\}$ and $\{f_{x_2|z_2}(\cdot|z_{2k}) : k = 1, 2, 3, ...\}$ are complete in Hilbert spaces $\mathcal{H}$ of functions defined on $\mathcal{X}_1$ and $\mathcal{X}_2$;*

ii) the relative deviation $D(z_1, z_2) = \frac{\left\|f_{x|z}(\cdot,\cdot|z_1,z_2) - f_{x_1|z_1}(\cdot|z_1) f_{x_2|z_2}(\cdot|z_2)\right\|}{\left\|f_{x_1|z_1}(\cdot|z_1) f_{x_2|z_2}(\cdot|z_2)\right\|}$ is continuous in $z = (z_1, z_2)$ on $\mathcal{N}(z_0)$ with $f_{x|z}(\cdot,\cdot|z_{10}, z_{20}) = f_{x_1|z_1}(\cdot|z_{10}) f_{x_2|z_2}(\cdot|z_{20})$.

iii) there exists a sequence $\{z_k : k = 1, 2, 3, ...\}$ of distinct $z_k \in \mathcal{O}$ converging to z_0 such that the sequence $\{f_{x|z}(\cdot,\cdot|z_{1k}, z_{2k}) : k = 1, 2, 3, ...\}$ is linearly independent,

Then, the sequence $\{f_{x|z}(\cdot,\cdot|z_{1k}, z_{2k}) : k = 1, 2, 3, ...\}$ is complete in the Hilbert space $\mathcal{H}$ of functions defined on $\mathcal{X}_1 \times \mathcal{X}_2$.

Proof: See the appendix.

In many applications, it is difficult to show the completeness for a multivariate conditional density. The results above use Theorem 1 to extend the completeness for the one-dimensional sequences $\{f_{x_1|z_1}(\cdot|z_{1k}) : k = 1, 2, 3, ...\}$ and $\{f_{x_2|z_2}(\cdot|z_{2k}) : k = 1, 2, 3, ...\}$ to the multiple dimensional sequence $\{f_{x|z}(\cdot,\cdot|z_{1k}, z_{2k}) : k = 1, 2, 3, ...\}$. The key assumption is that the endogenous variables are conditionally independent of each other for some value of the instruments, i.e.

$$f_{x|z}(\cdot,\cdot|z_{10}, z_{20}) = f_{x_1|z_1}(\cdot|z_{10}) f_{x_2|z_2}(\cdot|z_{20}). \tag{13}$$

We may then use the completeness of one-dimensional conditional densities $f_{x_1|z_1}(\cdot|z_{1k})$ and $f_{x_2|z_2}(\cdot|z_{2k})$ to show the completeness of a multi-dimensional density $f_{x|z}(\cdot,\cdot|z_{1k}, z_{2k})$. Therefore, Lemma 6 may reduce the dimension as well as the difficulty of the problem.

What we need for the multivariate case (Lemma 6) in equation (13) includes two steps: first, we need the independence between x_1 and x_2 only at $z = z_0$, i.e.,

$$x_1 \perp x_2 \mid z = z_0; \tag{14}$$

The second step requires with $z_0 = (z_{10}, z_{20})$

$$f_{x_1|z}(\cdot|z_0) = f_{x_1|z_1}(\cdot|z_{10}) \text{ and } f_{x_2|z}(\cdot|z_0) = f_{x_2|z_2}(\cdot|z_{20}).$$

This step is for simplicity and convenience because $f_{x_1|z}(\cdot|z_{10}, z_{20})$ and $f_{x_2|z}(\cdot|z_{10}, z_{20})$ are already one-dimensional densities and we may re-define the two sequences in condition i) in

Lemma 6 corresponding to $f_{x_1|z}(\cdot|z_0)$ and $f_{x_2|z}(\cdot|z_0)$. Such a simplification is particularly useful when one can find an instrument corresponding to each endogenous variable.

The completeness of a conditional density function $f(x|z)$ implies there exists a sequence of conditional density function $\{f(x|z_k) : k = 1, 2, 3, ...\}$ as a basis. At these points $z_k = (z_{1k}, z_{2k})$, an intuitive idea of Lemma 6 is the fact that the tensor product of univariate basis are multivariate basis. With the completeness of the sequence of product function $\{f_{x_1|z_1}(\cdot|z_{1k}) f_{x_2|z_2}(\cdot|z_{2k}) : k = 1, 2, 3, ...\}$, we can utilize the main perturbation result, Theorem 1, to extend the result to other sequence of function close to the sequence of the product function. At these "small" perturbation sequences, $\{f_{x|z}(\cdot, \cdot|z_{1k}, z_{2k}) : k = 1, 2, 3, ...\}$ can be nonseparable and satisfy the condition 14. For example, set $f_{x_1|z_1}(x_1|z_1) = \frac{1}{z_1} e^{-x_1 z_1}$ and $f_{x_2|z_2}(x_2|z_2) = \frac{1}{z_2} e^{-x_2 z_2}$ where $z_1, z_2 > 0$ and $x_1, x_2 \in \{0\} \cup \mathbb{R}^+$. Applying the results of Lemma 1 (a generalized version of Example 2) to these two density functions, we can obtain the completeness of the two families $\{f_{x_1|z_1}(\cdot|z_{1k}) : k = 1, 2, 3, ...\}$ and $\{f_{x_2|z_2}(\cdot|z_{2k}) : k = 1, 2, 3, ...\}$ where z_{1k} and z_{2k} are distinct sequences converging to 1. Then, consider $\{f_{x|z}(\cdot, \cdot|z_{1k}, z_{2k}) = \frac{c_{zk}}{z_{1k} z_{2k}} e^{-(x_1 z_{1k} + x_2 z_{2k} + (z_{1k}-1)(z_{2k}-1) x_1 x_2)} : k = 1, 2, 3, ...\}$ where c_{zk} is a normalized coefficient. This family has satisfied the conditions (i) and (ii) of Lemma 6 with $z_0 = (1, 1)$. If the family is linear independent then Lemma 6 ensures its completeness. Suppose that there exists $c_1, ..., c_m$ such that

$$c_1 e^{-(x_1 z_{1m} + x_2 z_{2m} + (z_{1m}-1)(z_{2m}-1) x_1 x_2)} + ... + c_m e^{-(x_1 z_{1m} + x_2 z_{2m} + (z_{1m}-1)(z_{2m}-1) x_1 x_2)} = 0$$

Differentiating the equation $m - 1$ times and plugging $(x_1, x_2) = (0, 0)$, we obtain

$$\begin{cases} c_1 + \ldots + c_m & = & 0 \\ c_1 z_{k_1} + \ldots + c_m z_{k_m} & = & 0 \\ \vdots & \ldots & \vdots \\ c_1 z_{k_1}^{m-1} + ... + c_m z_{k_m}^{m-1} & = & 0 \end{cases}.$$

Rewrite the above equation as

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ z_{k_1} & z_{k_2} & \dots & z_{k_m} \\ \vdots & \vdots & \ddots & \vdots \\ z_{k_1}^{m-1} & z_{k_2}^{m-1} & \dots & z_{k_m}^{m-1} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

The first matrix is the Vandermonde matrix and its property implies the linear independence of the family. Therefore, the multivariate result provides more results than the product case at z_0 and may not be too restrictive.

4. Conclusion

We provide sufficient conditions for the nonparametric identification of the regression function in a regression model with an endogenous regressor x and an instrumental variable z. The identification of the regression function from the conditional expectation of the dependent variable is implied by the completeness of the distribution of the endogenous regressor conditional on the instrument, i.e., $f(x|z)$. Sufficient conditions are then provided for the completeness of $f(x|z)$ without imposing a specific functional form, such as the exponential family. We use the results in the stability of bases in Hilbert space to show that if the relative deviation from a complete sequence of function is finite then $f(x|z)$ itself is complete, and therefore, the regression function is nonparametrically identified.

5. Appendix: Proofs

5.1. Preliminaries

Let $\mathcal{L}^2(\mathcal{X}) = \{h(\cdot) : \int_{\mathcal{X}} |h(x)|^2 dx < \infty,\}$ be a L^2 space with the following inner product $\langle f, g \rangle \equiv \int_{\mathcal{X}} f(x)g(x)dx$. We define the corresponding norm as: $\|f\|^2 = \langle f, f \rangle$. The completion of $\mathcal{L}^2(\mathcal{X})$ under the norm $\|\cdot\|$ is a Hilbert space, which may be denoted as $\mathcal{H}$. The conditional density of interest $f(x|z)$ is defined over $\mathcal{X} \times \mathcal{Z}$. If z only takes values from a countable set in $\mathcal{Z}$ then $f(x|z)$ can be considered as a sequence of functions $\{f_1, f_2, f_3, ...\}$ in $\mathcal{H}$ with

$$f_k \equiv f(\cdot|z_k),$$

where $\{z_k : k = 1, 2, 3, ...\}$ is a sequence in $\mathcal{Z}$. The property of the sequence $\{f_k\}$ determines the identification of the regression function in (2).

We then introduce the definition of a basis in a Hilbert space.

Definition 2. *A sequence of functions $\{f_1, f_2, f_3, ...\}$ in a Hilbert space $\mathcal{H}$ is said to be a basis if for any $h \in \mathcal{H}$ there corresponds a unique sequence of scalars $\{c_1, c_2, c_3, ...\}$ such that*

$$h = \sum_{k=1}^{\infty} c_k f_k.$$

The identification of a regression function in Equation (2) actually only requires a sequence $\{f_1, f_2, f_3, ...\}$ containing a basis, instead of a basis itself. Therefore, we consider a complete sequence of functions $\{f_1, f_2, f_3, ...\}$ which satisfies that $\langle g, f_k \rangle = 0$ for $k = 1, 2, 3...$ implies $g = 0$.

In fact, one can show that a basis is complete and that a complete sequence contains a basis. Since every element in a Hilbert space has a unique representation in terms of a basis, there is redundancy in a complete sequence. Given a complete sequence in a Hilbert space, we can extract a basis from the complete sequence. One of the important properties of a complete sequence for a Hilbert space is that every element can be approximated arbitrarily close by finite combinations of the elements. We summarize these results as follows.

Lemma 7. *(1) A basis in the Hilbert space $\mathcal{H}$ is also a complete sequence.*

(2) Let W be a closed linear subspace of a Hilbert space. Set $W^{\perp} = \{h \in \mathcal{H} : \langle h, g \rangle = 0$ for all $g \in W\}$. Then $W^{\perp}$ is a closed linear subspace such that, $W \bigoplus W^{\perp} = \mathcal{H}$.

(3) Given a complete sequence of functions $\{f_1, f_2, f_3, ...\}$ in a Hilbert space $\mathcal{H}$, there exists a subsequence $\{r_1, r_2, r_3, ...\}$ which is a basis in the Hilbert space $\mathcal{H}$.

Any function f in a Hilbert space can be expressed as a linear combination of the basis function with a unique sequence of scalars $\{c_1, c_2, c_3, ...\}$. Therefore, we can consider c_n as a function of f. In fact, $c_n(\cdot)$ is the so-called coefficient functional.

Definition 3. *If $\{f_1, f_2, f_3, ...\}$ is a basis in a Hilbert space $\mathcal{H}$, then every function f in $\mathcal{H}$ has a unique series $\{c_1, c_2, c_3, ...\}$ such that*

$$f = \sum_{n=1}^{\infty} c_n(f) f_n.$$

Each c_n is a function of f. The functionals c_n $(n = 1, 2, 3, ...)$ are called the coefficient functionals associated with the basis $\{f_1, f_2, f_3, ...\}$.

It is clear that c_n is a linear function of f. The following results regarding the coefficient functionals are from Theorem 3 in section 6 in Young (1980).

Lemma 8. *If $\{f_1, f_2, f_3, ...\}$ is a basis in a Hilbert space $\mathcal{H}$. Define c_n as coefficient functionals associated with the basis. Then, there exists a constant M such that*

$$1 \leq \|f_n\| \cdot \|c_n\| \leq M, \tag{15}$$

for all n.

In our proofs, we limit our attention to linearly independent sequences when providing sufficient conditions for completeness. The linear independence of an infinite sequence is considered as follows.

Definition 4. *A sequence of functions $\{f_n(\cdot)\}$ of a Hilbert space $\mathcal{H}$ is said to be $\omega-$independent if the equality*

$$\sum_{n=1}^{\infty} c_n f_n(x) = 0 \text{ for all } x \in \mathcal{X}$$

is possible only for $c_n = 0$, $(n = 1, 2, 3, ...)$.

It is obvious that the $\omega-$independence implies that linear independence. But the converse argument does not hold. A complete sequence may not be $\omega-$independent, but it contains a basis, and therefore, contains an $\omega-$independent subsequence.

Our proofs also need a uniqueness theorem of complex differentiable functions. Let $w = a + ib$, where a, b are real number and $i = \sqrt{-1}$. Define $\mathbb{C} = \{w = a + ib : a, b \in \mathbb{R}\}$ and it is called a complex plane. The complex differentiable function is defined as follows.

Definition 5. *Denote Ω as an open set in $\mathbb{C}$. Suppose f is a complex function defined in Ω. If $z_0 \in \Omega$ and*

$$\lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

exists, we denote this limit by $f'(z_0)$ and call it the derivative of f at z_0. If $f'(z_0)$ exists for every $z_0 \in \Omega$, f is called a complex differentiable function in Ω.

To be more precise, $f'(z_0)$ exists if for every $\varepsilon > 0$ there corresponds a $\delta > 0$ such that

$$\left| \frac{f(z) - f(z_0)}{z - z_0} - f'(z_0) \right| < \varepsilon \text{ for all } 0 < |z - z_0| < \delta.$$

A complex differentiable function has a large number of interesting properties which are different from a real differentiable function. One of them is the following uniqueness theorem, as stated in a corollary on page 209 in Rudin (1987).

Lemma 9. *If g and f are complex differentiable functions in an open connected set Ω and if $f(z) = g(z)$ for all z in some set which has a limit point in Ω, then $f(z) = g(z)$ for all $z \in \Omega$.*

5.2. Proofs of completeness of existing sequences

Proof of Lemma 1: Set $t(z) = 1$ is for simplicity. In order to use the above uniqueness result of complex differentiable functions, we consider a converging sequence $\{z_k : k = 1, 2, ...\}$ in $\mathcal{Z}$ as the set with a limit point. Since $\mu(\cdot)$ is continuous with $\mu'(z_0) \neq 0$ for some limit point $z_0 \in \mathcal{Z}$, there exists $\delta > 0$ and a sequence $\{z_k : k = 1, 2, ...\}$ converging to z_0 such that $\{\mu(z_k) : k = 1, 2, ...\} \in (\mu(z_0) - \delta, \mu(z_0) + \delta) \subset \mu(\mathcal{N}(z_0))$ be a sequence of distinct numbers converging to an interior point $\mu(z_0) \in \mu(\mathcal{N}(z_0))$. In addition, since $g(x|z) \in \mathcal{L}^1(\mathcal{X})$ for $z \in \mathcal{O}$,

$$\int_{\mathcal{X}} s(x) \exp\left[\mu(z_0)\tau(x) + \delta|\tau(x)|\right] dx < \infty.$$

Choose a weight function $\omega(x)$ satisfying $\int_{\mathcal{X}} \frac{s(x)^2 \exp[2(\mu(z_0)\tau(x) + \delta|\tau(x)|)]}{\omega(x)} dx < \infty$.

Given $h_0 \in \mathcal{L}^2(\mathcal{X}, \omega)$ and $\delta_1 < \delta$. Consider a complex function with the following form

$$f(w) = \int_{\mathcal{X}} s(x) e^{w\tau(x)} h_0(x) dx, \tag{16}$$

where the complex variable w is in the vertical strip $R \equiv \{w : \mu(z_0) - \delta_1 < \mathrm{Re}\,(w) < \mu(z_0) + \delta_1\}$.[10] Let $w = a + ib$, where a, b are real numbers. Applying Cauchy-Schwarz inequality to Eq. (16), we have for $a \in (\mu(z_0) - \delta_1, \mu(z_0) + \delta_1)$

$$\begin{aligned} \big|f(w)\big|^2 &\leq \Big|\int_{\mathcal{X}} s(x) e^{w\tau(x)} h_0(x) dx\Big|^2 \\ &\leq \Big(\int_{\mathcal{X}} \frac{s(x) e^{a\tau(x)}}{\omega(x)^{1/2}} |h_0(x)| \omega(x)^{1/2} dx\Big)^2 \\ &\leq \Big(\int_{\mathcal{X}} \frac{s(x) e^{\mu(z_0)\tau(x) + \delta_1|\tau(x)|}}{\omega(x)^{1/2}} |h_0(x)| \omega(x)^{1/2} dx\Big)^2 \\ &\leq \Big(\int_{\mathcal{X}} \frac{s(x) e^{\mu(z_0)\tau(x) + \delta|\tau(x)|}}{\omega(x)^{1/2}} |h_0(x)| \omega(x)^{1/2} dx\Big)^2 \\ &\leq \Big(\int_{\mathcal{X}} \frac{s(x)^2 \exp\left[2(\mu(z_0)\tau(x) + \delta|\tau(x)|)\right]}{\omega(x)} dx\Big)\Big(\int_{\mathcal{X}} |h_0(x)|^2 \omega(x) dx\Big) < \infty. \end{aligned} \tag{17}$$

This suggests that $f(w)$ defined in Eq. (16) exists and is finite. Suppose $\eta \in \mathbb{C}$ such that $|\eta| \leq \delta_2$ and $\delta_1 + \delta_2 < \delta$. Given $w \in R$. Consider the difference quotient of the integrand in Eq. (16), we have

$$\begin{aligned} |Q(x,\eta)| &\equiv \Big|\frac{s(x) e^{(w+\eta)\tau(x)} h_0(x) - s(x) e^{w\tau(x)} h_0(x)}{\eta}\Big| \\ &= \Big|s(x) \frac{e^{w\tau(x)} \left(e^{\eta\tau(x)} - 1\right)}{\eta} h_0(x)\Big| \\ &\leq s(x) \left|\frac{e^{w\tau(x) + \delta_2|\tau(x)|}}{\delta_2}\right| \Big|h_0(x)\Big| \\ &\leq s(x) \left|\frac{e^{(w+\delta_2)\tau(x)} + e^{(w-\delta_2)\tau(x)}}{\delta_2}\right| \Big|h_0(x)\Big| \\ &\leq 2s(x) \frac{e^{\mu(z_0)\tau(x) + (\delta_1+\delta_2)|\tau(x)|}}{\delta_2} \Big|h_0(x)\Big|, \end{aligned}$$

where we have used (i) apply the inequality $|\frac{e^{az}-1}{z}| \leq \frac{e^{\delta_2|a|}}{\delta_2}$ for $|z| \leq \delta_2$ to the factor $\frac{\left(e^{\eta\tau(x)}-1\right)}{\eta}$, and (ii) $w \in R$. The right-hand side is integrable when $\delta_1 + \delta_2 < \delta$ by a similar derivation in Eq. (17). It follows from the Lebesgue dominated convergence theorem that

$$f'(w) = \lim_{\eta \to 0} \int_{\mathcal{X}} Q(x,\eta) dx = \int_{\mathcal{X}} \lim_{\eta \to 0} Q(x,\eta) dx = \int_{\mathcal{X}} s(x)\tau(x) e^{w\tau(x)} h_0(x) dx.$$

[10] A holomorphic (or analytic) function defined with a similar function form in a strip is also discussed in the proof of Theorem 1 in Section 4.3 of Lehmann (1986). The proof provided here is close to the proof of Theorem 9 in Section 2.7 of Lehmann (1986).

Therefore, the function f defined through the integral is holomorphic.

The condition $\int_{\mathcal{X}} s(x)e^{\mu(z_k)\tau(x)}h_0(x)dx = 0$ is equivalent to $f(\mu(z_k)) = 0$ by Equation (16). This implies that the complex differentiable function f is equal to zeros in the sequence $\{\mu(z_1), \mu(z_2), \mu(z_3), ...\}$ which has a limit point $\mu(z_0)$. Applying the uniqueness theorem (Lemma 9) quoted above to f results in $f(w) = 0$ on $\{w : \mu(z_0) - \delta_1 < \mathrm{Re}\,(w) < \mu(z_0) + \delta_1\}$. If $\mathcal{X}$ is a bounded domain, we extend h_0 to a function in $\mathcal{L}^2(\mathbb{R}, \omega)$ by

$$\tilde{h}_0(x) = \begin{cases} h_0 & \text{if } x \in \mathcal{X}, \\ 0 & \text{otherwise.} \end{cases}$$

We also extend $s(x)$ and $\tau(x)$ to functions in $\mathbb{R}$, $\tilde{s}(x)$ and $\tilde{\tau}(x)$ respectively with the following properties, $\tilde{s}(x) > 0$ and $\tilde{\tau}'(x) \neq 0$. In particular, choose $w = \mu(\tilde{z}) + it$ for any real t, we have

$$\begin{aligned} f(w) &= \int_{\mathcal{X}} s(x)e^{\mu(\tilde{z})\tau(x)}e^{it\tau(x)}h_0(x)dx = 0 \\ &= \int_{-\infty}^{\infty} \tilde{s}(\tau^{-1}(x))e^{\mu(\tilde{z})x}e^{itx}\tilde{h}_0(\tilde{\tau}^{-1}(x))\frac{1}{\tilde{\tau}'(x)}dx \\ &\equiv \int_{-\infty}^{\infty} e^{itx}\hat{h}_0(x)dx. \end{aligned}$$

The last step implies that the Fourier transform of $\hat{h}_0(x)$ is zero on the whole real line. And Eq. (17) implies $\hat{h}_0 \in \mathcal{L}^1(\mathbb{R})$. By the uniqueness Theorem 9.12 in Rudin (1987) for $\hat{h}_0 \in \mathcal{L}^1(\mathbb{R})$, we have $\hat{h}_0 = 0$ and therefore the function $h_0 = 0$. This shows that the sequence $\{g(\cdot|z_k) = s(\cdot)t(z_k)e^{\mu(z_k)\tau(\cdot)} : k = 1, 2, ...\}$ is complete in $\mathcal{L}^2(\mathcal{X}, \omega)$. QED.

Proof of Lemma 2: Choose a sequence of distinct numbers $\{z_k\}$ in the support $\mathcal{Z}$ converging to $z_0 \in \mathcal{Z}$. Pick $0 < \delta' < \delta$ and choose a weight function $\omega(x)$ satisfying $\int \frac{\exp(-2\delta'|x|)}{\omega(x)}dx < \infty$.[11] Suppose that $\int_{-\infty}^{\infty} f_\epsilon\,(x - z_k)\,h_0(x)dx = 0$ for some $h_0 \in \mathcal{L}^2(\mathbb{R}, \omega)$. Consider

$$g\,(z) \equiv \int_{\mathcal{X}} h_0(x)f_\epsilon\,(x - z)\,dx,$$

[11] A specific choice of the weight function is $\omega(x) \equiv \exp\left(-2\delta_\omega|x|\right)$ for $0 < \delta_\omega < \delta'$.

which is a convolution. Let ϕ_g stands for the Fourier transform of g as follows:

$$\begin{aligned}
\phi_g(t) &= \int_{-\infty}^{\infty} e^{itz} g(z)\, dz \\
&= \int_{-\infty}^{\infty} e^{it(x-(x-z))} \int_{\mathcal{X}} h_0(x) f_\epsilon(x-z)\, dxdz \\
&= \int_{\mathcal{X}} e^{itx} h_0(x) \left(\int_{-\infty}^{\infty} e^{it(z-x)} f_\epsilon(-(z-x))\, dz \right) dx \\
&= \phi_{h_0}(t)\phi_{-\epsilon}(t) \\
&= \phi_{h_0}(t)\phi_\epsilon(-t).
\end{aligned}$$

We have $g(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itz} \phi_{h_0}(t)\phi_\epsilon(-t)dt$. We define

$$f(w) = \int_{-\infty}^{\infty} e^{-itw} \phi_{h_0}(t)\phi_\epsilon(-t)dt,$$

for

$$w = z + ib \text{ for } z, b \in \mathbb{R} \text{ and } \ |b| < \delta - \delta',$$

where δ is in Equation (4). Consider

$$\begin{aligned}
\big|f(w)\big|^2 &= \Big| \int_{-\infty}^{\infty} e^{-itw} \phi_{h_0}(t)\phi_\epsilon(-t)dt \Big|^2 \\
&\leq \Big(\int_{-\infty}^{\infty} |\phi_\epsilon(-t)| \frac{e^{bt}}{\omega(t)^{1/2}} |\phi_{h_0}(t)| \omega(t)^{1/2} dt \Big)^2 \\
&\leq \Big(\int_{-\infty}^{\infty} |\phi_\epsilon(-t)|^2 \frac{e^{2bt}}{\omega(t)} dt \Big) \Big(\int_{-\infty}^{\infty} |\phi_{h_0}(t)|^2 \omega(t) dt \Big) \\
&\leq C^2 \Big(\int_{-\infty}^{\infty} \frac{e^{-2(\delta-b)|t|}}{\omega(t)} dt \Big) \Big(\int_{-\infty}^{\infty} |\phi_{h_0}(t)|^2 \omega(t) dt \Big) \\
&\leq C^2 \Big(\int_{-\infty}^{\infty} \frac{e^{-2\delta'|t|}}{\omega(t)} dt \Big) \Big(\int_{-\infty}^{\infty} |\phi_{h_0}(t)|^2 \omega(t) dt \Big) < \infty
\end{aligned}$$

by $\phi_{h_0}(t) \in \mathcal{L}^2(\mathbb{R}, \omega)$ since $h_0 \in \mathcal{L}^2(\mathbb{R}, \omega)$.[12] Since the right-hand side is finite, then $f(w)$ exists and is finite in $R = \{z + ib : |b| < \delta - \delta'\}$. To prove $f(w)$ is analytic (complex differentiable) in

[12] See Theorem 9.13 on page 186 in Rudin (1987).

R, we consider the difference quotient at a point $w_0 = z_0 + ib_0$ in R. For $|\eta| < \delta_1 < \delta - \delta' - b_0$,

$$
\begin{aligned}
|Q(t,\eta)| &\equiv \Big|\frac{e^{-it(w_0+\eta)}\phi_{h_0}(t)\phi_\epsilon(-t) - e^{-itw_0}\phi_{h_0}(t)\phi_\epsilon(-t)}{\eta}\Big| \\
&= \Big|\frac{e^{-itw_0}(e^{-it\eta}-1)}{\eta}\phi_{h_0}(t)\phi_\epsilon(-t)\Big| \\
&\le \Big|\frac{e^{-itw_0}e^{\delta_1|t|}}{\delta_1}\Big|\Big|\phi_{h_0}(t)\Big|\Big|\phi_\epsilon(-t)\Big| \\
&\le \Big|\frac{e^{b_0 t}e^{\delta_1|t|}}{\delta_1}\Big|\Big|\phi_{h_0}(t)\Big|\Big|\phi_\epsilon(-t)\Big| \\
&\le 2C\Big|\frac{e^{(b_0+\delta_1)|t|}}{\delta_1\omega(t)^{1/2}}\Big|e^{-\delta|t|}\Big|\phi_{h_0}(t)\Big|\omega(t)^{1/2} \\
&\le \frac{2C}{\delta_1}\Big|\frac{e^{(b_0+\delta_1-\delta)|t|}}{\omega(t)^{1/2}}\Big|\Big|\phi_{h_0}(t)\Big|\omega(t)^{1/2},
\end{aligned}
$$

where we have used the inequalities $|\frac{e^{az}-1}{z}| \le \frac{e^{\delta_1|a|}}{\delta_1}$ for $|z| \le \delta_1$, and $|\phi_\varepsilon(t)| < Ce^{-\delta|t|}$. The condition $b_0 + \delta_1 - \delta < -\delta'$ and $\phi_{h_0}(t) \in \mathcal{L}^2(\mathbb{R}, \omega)$ make the right-hand side integrable. Since the quotient is bounded above by an integrable function, the Lebesgue dominated convergence theorem implies

$$
f'(w) = \lim_{\eta\to 0}\int_{-\infty}^{\infty} Q(t,\eta)dt = \int_{-\infty}^{\infty}\lim_{\eta\to 0} Q(t,\eta)dt = -it\int_{-\infty}^{\infty} e^{-itw}\phi_{h_0}(t)\phi_\epsilon(-t)dt.
$$

Thus, $f(w)$ is analytic (complex differentiable) in $R = \{z + ib : |b| < \delta - \delta'\}$. Consequently, the fact that $f(w)$ equals zero for a sequence $\{z_1, z_2, z_3, ...\}$ converging to z_0 implies that $f(w)$ is equal to zero in R by the uniqueness theorem cited in the proof of Lemma 1. This suggests that $f(w)$ is equal to zero for all $w = z$ on the real line, i.e., $\int_{-\infty}^{\infty} e^{-itz}\phi_{h_0}(t)\phi_\epsilon(-t)dt = 0$ for all $z \in \mathbb{R}$. Since $\int_{-\infty}^{\infty} |\phi_{h_0}(t)\phi_\epsilon(-t)|dt \le (\int_{-\infty}^{\infty} |\phi_{h_0}(t)|^2 e^{-2\delta_1|t|}dt)^{1/2}(\int_{-\infty}^{\infty} |\phi_\epsilon(-t)|^2 e^{2\delta_1|t|}dt)^{1/2} < \infty$, $\phi_{h_0}(t)\phi_\epsilon(-t) \in \mathcal{L}^1(\mathbb{R})$. Thus, the characteristic function $\phi_{h_0}(t)\phi_\epsilon(-t) = 0$ for all t.[13] By Eq. (4), i.e., $\phi_\epsilon(t) \ne 0$, we have $\phi_{h_0}(t) = 0$ for all $t \in \mathbb{R}$ so $h_0 = 0$. The family $\{g(x|z) = f_\varepsilon(x - z_k) : k = 1, 2, ...\}$ is complete in $\mathcal{L}^2(\mathcal{X}, \omega)$. QED.

5.3. Proof of Theorem 1

We prove Theorem 1 in three steps:

1. The total deviation from a complete sequence $\{g(\cdot|z_k)\}$ to the corresponding sequence

[13] See Theorem 9.12 on page 185 in Rudin (1987).

$\{f(\cdot|z_k)\}$ is

$$\sum_{k=1}^{\infty} \frac{\|f(\cdot|z_k) - g(\cdot|z_k)\|}{\|g(\cdot|z_k)\|}. \tag{18}$$

We prove that if the total deviation from a basis to an $\omega-$ independent sequence is finite, then the latter sequence is also a basis. This result is summarized in Lemma 10 as the cornerstone of the proof of Theorem 1.

2. Condition ii) implies that the total deviation in Eq. (18) is finite.

3. A linearly independent sequence $\{f(\cdot|z_k)\}$ in a normed space contains an $\omega-$ independent subsequence $\{f(\cdot|z_{k_l})\}$. Finally, for a complete sequence $\{g(\cdot|z_{k_l})\}$ and the $\omega-$ independent sequence $\{f(\cdot|z_{k_l})\}$, Equation (18) and Lemma 10 imply that the sequence $\{f(\cdot|z_{k_l})\}$ is complete, and therefore, $\{f(\cdot|z) : z \in \mathcal{N}(z_0)\}$ is complete.

Step 1: We prove that if the total deviation from a basis to an $\omega-$ independent sequence is finite, then the latter sequence is also a basis. This result is summarized in the following lemma as the cornerstone of the proof of Theorem 1.

Lemma 10. *Suppose that*

i) the sequence $\{e_n(\cdot) : n = 1, 2, ...\}$ *is a basis in a Hilbert space* $\mathcal{H}$*;*

ii) the sequence $\{f_n(\cdot) : n = 1, 2, ...\}$ *in* $\mathcal{H}$ *is* $\omega-$*independent;*

iii) $\sum_{n=1}^{\infty} \frac{\|f_n(\cdot) - e_n(\cdot)\|}{\|e_n(\cdot)\|} < \infty$.

Then, the sequence $\{f_n(\cdot) : n = 1, 2, ...\}$ *is a basis in* $\mathcal{H}$*.*

Proof of Lemma 10: As in the proof of Theorem 15 on page 45 of Young (1980), we consider for any function $f \in \mathcal{H}$

$$f = \sum_{n=1}^{\infty} c_n(f) e_n,$$

where $c_n(f)$ is the so-called coefficient functional corresponding to the basis $\{e_n\}$. It is clear that $c_n(f)$ is a linear function of f. Define an operator $T : \mathcal{H} \to \mathcal{H}$ as

$$Tf = \sum_{n=1}^{\infty} c_n(f) \left(e_n - f_n\right).$$

It is clear that T is linear. Since $c_n(e_n) = 1$ and $c_k(e_n) = 0$ for $k \neq n$, we have

$$Te_n = \sum_{n=1}^{\infty} c_n(e_n)\,(e_n - f_n) = e_n - f_n.$$

By using the triangle inequality and the definition of functional, we have

$$\begin{aligned} \|Tf\| &= \left\| \sum_{n=1}^{\infty} c_n(f)\,(e_n - f_n) \right\| \\ &\leq \sum_{n=1}^{\infty} \|c_n(f)\,(e_n - f_n)\| \\ &\leq \left(\sum_{n=1}^{\infty} \|e_n - f_n\|\,\|c_n\| \right) \|f\|\,. \end{aligned}$$

Lemma 8 suggests that

$$1 \leq \|e_n\|\,\|c_n\| \leq M.$$

Therefore, we have

$$\begin{aligned} \|Tf\| &\leq \left(\sum_{n=1}^{\infty} \frac{\|e_n - f_n\|}{\|e_n\|}\,\|e_n\|\,\|c_n\| \right) \|f\| \\ &\leq M \left(\sum_{n=1}^{\infty} \frac{\|e_n - f_n\|}{\|e_n\|} \right) \|f\| \end{aligned}$$

The relationship above implies that the linear operator T is bounded if

$$\sum_{n=1}^{\infty} \frac{\|e_n - f_n\|}{\|e_n\|} < \infty,$$

which will be shown in the next step to be implied by condition (iii). We then show that T is a compact operator. Set

$$T_N f = \sum_{n=1}^{N} c_n(f)\,(e_n - f_n)\,.$$

Consider

$$\begin{aligned}
\|(T-T_n)f\| &= \left\|\sum_{n=N+1}^{\infty} c_n(f)\,(e_n-f_n)\right\| \\
&\le \sum_{n=N+1}^{\infty} \|c_n(f)\,(e_n-f_n)\| \\
&\le \left(\sum_{n=N+1}^{\infty} \|e_n-f_n\|\,\|c_n\|\right)\|f\|\,.
\end{aligned}$$

Follow the previous derivation of $\|Tf\|$, we can obtain

$$\|(T-T_n)f\| \le M\left(\sum_{n=N+1}^{\infty} \frac{\|e_n-f_n\|}{\|e_n\|}\right)\|f\|\,.$$

This implies that $\|(T-T_n)\| \le M\left(\sum_{n=N+1}^{\infty} \frac{\|e_n-f_n\|}{\|e_n\|}\right)$. Assumption iii) of Lemma 10 suggests that $\|T-T_n\| \to 0$. Since each T_N has finite dimensional range and $\|T-T_N\| \to 0$ as $N\to\infty$, T is an compact operator.[14]

Next, we show that $Ker(I-T)=\{0\}$, i.e., $(I-T)$ is invertible. Consider

$$\begin{aligned}
0 &= (I-T)\,f \\
&= f-\sum_{n=1}^{\infty} c_n(f)\,(e_n-f_n) \\
&= \sum_{n=1}^{\infty} c_n(f)e_n-\sum_{n=1}^{\infty} c_n(f)e_n+\sum_{n=1}^{\infty} c_n(f)f_n \\
&= \sum_{n=1}^{\infty} c_n(f)f_n
\end{aligned}$$

Since $\{f_n(\cdot)\}$ is an $\omega-$independent sequence, we have $c_n(f)=0$ for all n, and therefore, $0=(I-T)\,f$ implies $f=0$.

Therefore, T is a compact operator defined in a Hilbert space $\mathcal{H}$ with $Ker(I-T)=\{0\}$. By the Fredholm alternative, this shows that $(I-T)$ is a bounded invertible operator.[15]

Since T is bounded, $(I-T)$ is also bounded. Therefore, we have shown that $(I-T)$ is a bounded invertible operator. Clearly, we have $(I-T)\,e_n=f_n$. Consider any $h\in\mathcal{H}$. Then,

[14] If an bounded linear operator T is the limit of operators of finite rank, then T is compact. See Exercise 13 on page 112 in Rudin (1991).

[15] See the Fredholm alternative in Rudin (1991), Exercise 13 on page 112.

$(I-T)^{-1}h$ has an unique series expression $(I-T)^{-1}h = \sum_{n=1}^{\infty} c_n e_n$ since $\{e_n(\cdot)\}$ is a basis. Since $(I-T)$ is bounded, applying $(I-T)$ to the expression above results in $h = \sum_{n=1}^{\infty} c_n f_n$. This series expansion is unique because $\{f_n(\cdot)\}$ is ω-independent. The argument above shows that every element $h \in \mathcal{H}$ has a unique series expansion in terms of f_n. Thus, $\{f_n(\cdot)\}$ is also a basis for $\mathcal{H}$. QED

Step 2: We show condition ii) implies Equation (18), i.e.,

$$\sum_{k=1}^{\infty} \frac{\|f(\cdot|z_k) - g(\cdot|z_k)\|}{\|g(\cdot|z_k)\|} < \infty. \tag{19}$$

We choose a sequence $\{z_k : k = 1, 2, ...\} \subset \mathcal{N}(z_0)$ converging to $z_0 \in \mathcal{N}(z_0)$. In other words, z_0 is a limit point in $\mathcal{N}(z_0)$. Define

$$D(z) \equiv \frac{\|f(\cdot|z) - g(\cdot|z)\|}{\|g(\cdot|z)\|}$$

for z close to z_0. Condition ii) imply that $D(z)$ is Lipschitz continuous at z_0 with $D(z_0) = 0$. Then, we have for some constant C and z close to z_0

$$\frac{\|f(\cdot|z) - g(\cdot|z)\|}{\|g(\cdot|z)\|} = D(z) - D(z_0) \leq C\,|z - z_0|\,.$$

Therefore, we may choose $|z_k - z_0| = O(k^{-p})$ for $p > 1$ so that Equation (18) holds with $\sum_{k=1}^{\infty} D(z_k) = O(\sum_{k=1}^{\infty} k^{-p}) < \infty$. Thus, there exists a sequence $\{z_k : k = 1, 2, ...\}$ converging to z_0 such that Equation (18) holds.

Step 3: Condition iii) implies that there exists a linearly independent sequence in $\{f(\cdot|z_k)\}$. According to the second Theorem in Erdös and Straus (1953), any linearly independent sequence in a normed space contains an $\omega-$ independent subsequence. We obtain an $\omega-$ independent subsequence $\{f(\cdot|z_{k_l})\}$ in $\{f(\cdot|z_k)\}$.

We then show that the $\omega-$ independent subsequence $\{f(\cdot|z_{k_l})\}$ is complete in the Hilbert space $\mathcal{H}$. Since the sequence $\{z_{k_l}\}$ corresponding to $\{f(\cdot|z_{k_l})\}$ is a subsequence of $\{z_k\}$ and also converges to z_0, condition i) implies that the corresponding sequence $\{g(\cdot|z_{k_l})\}$ is complete in the Hilbert space defined on $\mathcal{X}$. The two sequences also satisfies Equation (18) i.e.,

$$\sum_{l=1}^{\infty} \frac{\|f(\cdot|z_{k_l}) - g(\cdot|z_{k_l})\|}{\|g(\cdot|z_{k_l})\|} < \infty. \tag{20}$$

Let $\{e_n\}$ denote a basis contained in the complete sequence $\{g(\cdot|z_{k_l})\}$ and $\{f_n\}$ be the corresponding subsequence in $\{f(\cdot|z_{k_l})\}$, which is also $\omega-$ independent. Then $\{e_n\}$ and $\{f_n\}$ also satisfies $\sum_{n=1}^{\infty} \frac{\|f_n(\cdot)-e_n(\cdot)\|}{\|e_n(\cdot)\|} < \infty$. Lemma 10 implies that $\{f_n\}$ is a basis and therefore $\{f(\cdot|z_{k_l})\}$ is complete in the Hilbert space $\mathcal{H}$. Since the sequence $\{z_k\}$ is in $\mathcal{N}(z_0)$, the family $\{f(\cdot|z) : z \in \mathcal{N}(z_0)\}$ is complete in the Hilbert space $\mathcal{H}$. QED.

5.4. Proof of a Sufficient Condition for Lipschitz Continuity

Proof: For $z \in \mathcal{N}(z_0)$,

$$\begin{aligned}
|D(z) - D(z_0)| &\leq c_1 |D(z) - D(z_0)||D(z) + D(z_0)| \\
&= c_1 |D(z)^2 - D(z_0)^2| \\
&= c_1 \frac{\left| \|f(\cdot|z) - g(\cdot|z)\|^2 \|g(\cdot|z_0)\|^2 - \|f(\cdot|z_0) - g(\cdot|z_0)\|^2 \|g(\cdot|z)\|^2 \right|}{\|g(\cdot|z)\|^2 \|g(\cdot|z_0)\|^2} \\
&\leq c_2 \frac{\left| d(z) - d(z_0) \right|}{\|g(\cdot|z_0)\|^4} \\
&\leq c_2 \frac{\left| \frac{\partial}{\partial z} d(\tilde{z}) \right| |z - z_0|}{\|g(\cdot|z_0)\|^4},
\end{aligned}$$

where $d(z) \equiv \|f(\cdot|z) - g(\cdot|z)\|^2 \|g(\cdot|z_0)\|^2 - \|f(\cdot|z_0) - g(\cdot|z_0)\|^2 \|g(\cdot|z)\|^2$, and $\tilde{z}$ is some point in $\mathcal{N}(z_0)$. $D(z)$ is Lipschitz continuous if the function $d(z)$ has a bounded derivative at $\tilde{z}$. This holds because for some function $h(\cdot|z) \in \mathcal{L}^2(\mathcal{X}, \omega)$ with $\frac{\partial}{\partial z} h(\cdot|z) \in \mathcal{L}^2(\mathcal{X}, \omega)$ the derivative of $\|h(\cdot|z)\|^2$ w.r.t. z is finite due to the Cauchy-Schwarz inequality as follows:

$$\left| \frac{\partial}{\partial z} \left(\|h(\cdot|z)\|^2 \right) \right| \leq 2 \|h(\cdot|z)\| \left\| \frac{\partial}{\partial z} h(\cdot|z) \right\|.$$

Furthermore, if $\mathcal{X}$ is bounded, we only need $\frac{\partial}{\partial z} h(\cdot|z)$ to be bounded. QED.

5.5. Proof of the linear independence

Proof of Lemma 3(3): We have for $z > 0$ and $0 \in \mathcal{X}$

$$f(x|z) = \frac{d}{dx} F_0(z \times x)$$

with

$$W(0) = \Pi_{i=1}^{I} \left(z_{k_i} \frac{d^{(i)} F_0(0)}{dx^{(i)}} \right) \times \det \begin{pmatrix} 1 & 1 & ... & 1 \\ z_{k_1} & z_{k_2} & ... & z_{k_I} \\ ... & ... & ... & ... \\ (z_{k_1})^{I-1} & (z_{k_2})^{I-1} & ... & (z_{k_I})^{I-1} \end{pmatrix}.$$

According to the property of the Vandermonde matrix, the determinant $W(x)$ is not equal to zero when $F_0(x)$ has all the nonzero derivative at $x = 0$ and z_k are nonzero and distinctive. We may also generalize the above argument to show $\{f(\cdot|z_k)\}$ is linear independent with

$$f(x|z) = \frac{d}{dx} F_0(\mu(z)\tau(x))$$

where $\mu'(z) \neq 0$ and $\tau(\cdot)$ is monotonic with $\tau(0) \equiv 0$. While the restriction $\mu'(z) \neq 0$ guarantees that $\mu(z_k)$ are different for a distinct sequence $\{z_k\}$ around z_0, the condition $\tau(\cdot)$ is monotonic ensures that the linear independence for any x is the same as that for any $\tau(x)$. If $\sum_{i=1}^{I} c_i f(\cdot|z_{k_i}) = 0$, then it is equivalent to $\sum_{i=1}^{I} c_i \frac{d}{dx} F_0(\mu(z_{k_i})\tau(\cdot)) = 0$. This implies $\sum_{i=1}^{I} c_i \frac{d}{dx} F_0(\mu(z_{k_i})\tau) = 0$ for all $\tau \in \tau(\mathcal{X})$. Thus, we may show the determinant of $W(x)$ of the the function $f(x|z)$ is nonzero at $x = 0$. QED.

5.6. Proof of completeness in applications

Proof of Lemma 4: Let $\mathcal{N}(z_0)$ be an open neighborhood of z_0. Since the characteristic function $\phi_{z_0}(t)$ of $f(\cdot|z_0)$ satisfies Equation (4) in Lemma 2, we may generate a complete sequence $\{g(x|z_k) = f(x - \mu(z_k)|z_0) : k = 1, 2, ...\}$ satisfying condition i) in Theorem 1 with $\mu(z_0) = 0$ and $\mu'(z_0) \neq 0$. We have $f(\cdot|z_0) = g(\cdot|z_0)$ and $\|g(\cdot|z_0)\| > 0$ due to $|\phi_{z_0}(t)| > 0$.

As discussed below Theorem 1, when the Hilbert space $\mathcal{H}$ is the $\mathcal{L}^2(\mathcal{X},\omega)$, the relative deviation $D(z)$ is Lipschitz continuous if $\|g(\cdot|z_0)\|>0$ and the first-order derivatives $\frac{\partial}{\partial z}f(\cdot|z)$ and $\frac{\partial}{\partial z}g(\cdot|z)$ are also in $\mathcal{L}^2(\mathcal{X},\omega)$ for $z\in\mathcal{N}(z_0)$. This is because the derivative of $\|f(\cdot|z)\|^2$ w.r.t. z is bounded by the Cauchy-Schwarz inequality as follows:

$$\begin{aligned}\left|\frac{\partial}{\partial z}\left(\|f(\cdot|z)\|^2\right)\right| &= \left|\frac{\partial}{\partial z}\int f(x|z)^2\omega(x)dx\right| \\ &= \left|\int 2f(x|z)\omega(x)^{1/2}\frac{\partial}{\partial z}f(x|z)\omega(x)^{1/2}dx\right| \\ &\leq 2\|f(\cdot|z)\|\left\|\frac{\partial}{\partial z}f(\cdot|z)\right\|.\end{aligned}$$

For $g(x|z)=f(x-\mu(z)|z_0)$, we have

$$\begin{aligned}\frac{\partial}{\partial z}g(x|z) &= \frac{\partial}{\partial z}f(x-\mu(z)|z_0) \\ &= \frac{\partial}{\partial x}f(x-\mu(z)|z_0)\left(-\mu'(z)\right).\end{aligned}$$

The condition ii) of Lemma 4 implies that $\frac{\partial}{\partial z}g(\cdot|z)$ is in $\mathcal{L}^2(\mathcal{X},\omega)$ so that the relative deviation $D(z)=\frac{\|f(\cdot|z)-g(\cdot|z)\|}{\|g(\cdot|z)\|}$ is Lipschitz continuous in z. Since $D(z_0)=0$ by definition, the condition ii) in Theorem 1 holds. Thus, the completeness holds for $\{f(\cdot|z):z\in\mathcal{N}(z_0)\}$.

We then consider the special case $f(x|z)=\frac{1}{\sigma(z)}f_\varepsilon\left(\frac{x-\mu(z)}{\sigma(z)}\right)$. Without loss of generality, we set $\sigma(z_0)=1$ because we may always redefine $\frac{1}{\sigma(z_0)}f_\varepsilon\left(\frac{x}{\sigma(z_0)}\right)$ as $f_\varepsilon(x)$. Since $\mu(\cdot)$ is continuous with $\mu'(z_0)\neq 0$, the sequence $\{\mu(z_k):k=1,2,3,...\}\subset\mu(\mathcal{N}(z_0))$ may be a distinct sequence converging to $\mu(z_0)\in\mu(\mathcal{N}(z_0))$. Applying the results in Lemma 2 with the sequence $\{\mu(z_k):k=1,2,3,...\}$, we may show that $\{g(x|z_k)=f_\varepsilon(x-\mu(z_k)):k=1,2,...\}$ is complete. We then extend the completeness of $\{g(x|z_k)=f_\varepsilon(x-\mu(z_k)):k=1,2,...\}$ to $\{f(x|z_k)=\frac{1}{\sigma(z_k)}f_\varepsilon\left(\frac{x-\mu(z_k)}{\sigma(z_k)}\right):k=1,2,...\}$. Since $\sigma(z_0)=1$, we have $f(x|z_0)=g(x|z_0)$.

We then check $f(\cdot|z) \in \mathcal{L}^2(\mathbb{R})$ for any $z \in \mathcal{N}(z_0)$. We have for some constant C_1, C_2

$$\begin{aligned}
\|f(\cdot|z)\| &= \int_{\mathcal{X}} \left| \frac{1}{\sigma(z)} f_\varepsilon\left(\frac{x-\mu(z)}{\sigma(z)}\right)\right|^2 \omega(x)dx \\
&\leq C_1 \int_{\mathcal{X}} \left| \frac{1}{\sigma(z)} f_\varepsilon\left(\frac{x-\mu(z)}{\sigma(z)}\right)\right|^2 dx \\
&= C_1 \int_{\mathcal{X}} \left| \frac{1}{\sigma(z)} f_\varepsilon(\varepsilon)\right|^2 \sigma(z)\, d\varepsilon \\
&\leq \frac{C_2}{\sigma(z_0)} \int_{\mathbb{R}} |f_\varepsilon(\varepsilon)|^2 \, d\varepsilon.
\end{aligned}$$

The last step is due to the continuity of $\sigma(\cdot)$ and $\sigma(z_0) > 0$. Since $|\phi_\varepsilon(t)| < Ce^{-\delta|t|}$, we have $\int_{\mathbb{R}} |\phi_\varepsilon(t)|^2 \, dt < \infty$, the last expression is finite, and therefore, $f(\cdot|z)$ is in $\mathcal{L}^2(\mathbb{R}, \omega)$ for $z \in \mathcal{N}(z_0)$.

In order to show the Lipschitz continuity of $D(z)$, we show $\frac{\partial}{\partial z} f(\cdot|z)$ is also in $\mathcal{L}^2(\mathbb{R}, \omega)$. We have

$$\begin{aligned}
\frac{\partial}{\partial z} f(x|z) &= \frac{-\sigma'(z)}{\sigma^2(z)} f_\varepsilon\left(\frac{x-\mu(z)}{\sigma(z)}\right) + f'_\varepsilon\left(\frac{x-\mu(z)}{\sigma(z)}\right)\left(\frac{-\mu'(z)}{\sigma^2(z)}\right) \\
&\quad + \frac{x-\mu(z)}{\sigma(z)} f'_\varepsilon\left(\frac{x-\mu(z)}{\sigma(z)}\right)\left(\frac{-\sigma'(z)}{\sigma^2(z)}\right).
\end{aligned}$$

The function $\frac{\partial}{\partial z} f(\cdot|z)$ for $z \in \mathcal{N}(z_0)$ is in $\mathcal{L}^2(\mathbb{R})$ because of condition ii'). It follows that these $\frac{\partial}{\partial z} f(\cdot|z)$ are in $\mathcal{L}^2(\mathbb{R}, \omega)$. Therefore, the total deviation

$$D(z) = \frac{\left\| \frac{1}{\sigma(z)} f_\varepsilon\left(\frac{x-\mu(z)}{\sigma(z)}\right) - f_\varepsilon(x-\mu(z)) \right\|}{\| f_\varepsilon(x-\mu(z)) \|}$$

is Lipschitz continuous in z.

We show the linear independence of $\{f(\cdot|z_k)\}$ as follows:

$$\lim_{x \to -\infty} \frac{f(x|z_{k+1})}{f(x|z_k)} = \lim_{x \to -\infty} \frac{\left|\frac{1}{\sigma(z_{k+1})}\right| f_\varepsilon\left(\frac{x-\mu(z_{k+1})}{\sigma(z_{k+1})}\right)}{\left|\frac{1}{\sigma(z_k)}\right| f_\varepsilon\left(\frac{x-\mu(z_k)}{\sigma(z_k)}\right)},$$

where

$$
\begin{aligned}
\frac{f_\varepsilon\left(\frac{x-\mu(z_{k+1})}{\sigma(z_{k+1})}\right)}{f_\varepsilon\left(\frac{x-\mu(z_k)}{\sigma(z_k)}\right)} &= \frac{f_\varepsilon\left(\frac{x-\mu(z_k)}{\sigma(z_k)}-\left(\frac{x-\mu(z_k)}{\sigma(z_k)}-\frac{x-\mu(z_{k+1})}{\sigma(z_{k+1})}\right)\right)}{f_\varepsilon\left(\frac{x-\mu(z_k)}{\sigma(z_k)}\right)} \\
&= \frac{f_\varepsilon\left(\frac{x-\mu(z_k)}{\sigma(z_k)}-\left(\frac{[\sigma(z_{k+1})-\sigma(z_k)]x-\sigma(z_{k+1})\mu(z_k)+\sigma(z_k)\mu(z_{k+1})}{\sigma(z_k)\sigma(z_{k+1})}\right)\right)}{f_\varepsilon\left(\frac{x-\mu(z_k)}{\sigma(z_k)}\right)} \\
&< \frac{f_\varepsilon\left(\frac{x-\mu(z_k)}{\sigma(z_k)}-c\right)}{f_\varepsilon\left(\frac{x-\mu(z_k)}{\sigma(z_k)}\right)}.
\end{aligned}
$$

If $\sigma'(z)=0$, i.e., $\sigma(z_{k+1})=\sigma(z_k)$, we may pick z_k such that $\mu(z_{k+1})>\mu(z_k)$ so that the last inequality holds because $f_\varepsilon(x)$ decreases as $x\to-\infty$. If $\sigma'(z)\neq 0$, we may pick z_k such that $\sigma(z_{k+1})<\sigma(z_k)$ and therefore $\left(\frac{[\sigma(z_{k+1})-\sigma(z_k)]x-\sigma(z_{k+1})\mu(z_k)+\sigma(z_k)\mu(z_{k+1})}{\sigma(z_k)\sigma(z_{k+1})}\right)>c>0$ for some constant c as $x\to-\infty$. Therefore, condition iii') implies that condition (2) in Lemma 3 holds.

Finally, Theorem 1 implies that the family $\{f(\cdot|z): z\in\mathcal{N}(z_0)\}$ is complete in $\mathcal{L}^2(\mathbb{R},\omega)$. In fact, the proof of Theorem 1 suggests that the sequence $\{f(\cdot|z_k): k=1,2,...\}$ is also complete. QED.

Proof of Lemma 5: Without loss of generality, we assume $\frac{\partial}{\partial z}h(z_0,\varepsilon)<0$ as $\varepsilon\to-\infty$. Then we choose distinct $z_k\uparrow z_0$ such that $|z_k-z_0|<\frac{1}{k^p}$ for some $p>2$. We use the complete sequence $\{g(x|z_k)=f_\varepsilon(x-\mu(z_k)): k=1,2,...\}$ with $\mu(z_0)=0$ and $\mu'(z_0)\neq 0$ from Lemma 4. This implies that $g(x|z_0)=f_\varepsilon(x)=f(x|z_0)$ because $h(z_0,\varepsilon)=\varepsilon$. In addition, the normality assumption of ε suggests $\{g(\cdot|z_k): k=1,2,...\}$ is complete in $\mathcal{L}^2(\mathbb{R},\omega)$ for a weight function ω by Lemma 2. Then the complete sequence $\{g(\cdot|z_k): k=1,2,...\}$ satisfies the condition i) in Theorem 1.

We may check that the family $\{f(x|z_k)=\left|\frac{\partial}{\partial x}h^{-1}(z_k,x)\right|f_\varepsilon\left(h^{-1}(z_k,x)\right): k=1,2,...\}$ is in

$\mathcal{L}^2(\mathbb{R},\omega)$. Consider for some constant c_1 and $z \in \mathcal{N}(z_0)$

$$\begin{aligned}\int_{\mathbb{R}} |f(x|z)|^2 \, dx &= \int_{\mathbb{R}} \left| \frac{\partial h^{-1}(z,x)}{\partial x} f_\varepsilon \left(h^{-1}(z,x)\right) \right|^2 dx \\ &= \int_{\mathbb{R}} \left| \left(\frac{\partial h(z,\varepsilon)}{\partial \varepsilon} \right)^{-1} f_\varepsilon(\varepsilon) \right|^2 \frac{\partial h(z,\varepsilon)}{\partial \varepsilon} d\varepsilon \\ &= \int_{\mathbb{R}} \left| \frac{\partial h(z,\varepsilon)}{\partial \varepsilon} \right|^{-1} |f_\varepsilon(\varepsilon)|^2 \, d\varepsilon \\ &\leq c_1 \int_{\mathbb{R}} \left| \frac{\partial h(z_0,\varepsilon)}{\partial \varepsilon} \right|^{-1} |f_\varepsilon(\varepsilon)|^2 \, d\varepsilon \\ &= \frac{c_1}{C} \int_{\mathbb{R}} |f_\varepsilon(\varepsilon)|^2 \, d\varepsilon < \infty\end{aligned}$$

The last step is because conditions i) and the normality assumption of ε imply $\left|\frac{\partial h(z_0,\varepsilon)}{\partial \varepsilon}\right| > C > 0$ and $\int_{\mathbb{R}} |f_\varepsilon(\varepsilon)|^2 \, d\varepsilon < \infty$. That means $f(x|z) \in \mathcal{L}^2(\mathbb{R})$ for $z \in \mathcal{N}(z_0)$. Since a weight function is bounded, $f(x|z) \in \mathcal{L}^2(\mathbb{R},\omega)$ for $z \in \mathcal{N}(z_0)$. The condition ii) of Lemma 5 implies that $\frac{\partial}{\partial z} f(\cdot|z)$ and $\frac{\partial}{\partial z} g(\cdot|z)$ are in $\mathcal{L}^2(\mathbb{R},\omega)$ so that the relative deviation $D(z) = \frac{\|f(\cdot|z)-g(\cdot|z)\|}{\|g(\cdot|z)\|}$ is Lipschitz continuous in z. Since $D(z_0) = 0$ by definition, the condition ii) in Theorem 1 holds.

We show the linear independence of $\{f(\cdot|z_k)\}$ and the corresponding CDF sequence $\{F(\cdot|z_k)\}$ as follows. We consider

$$\begin{aligned}\lim_{x\to-\infty} \frac{F(x|z_{k+1})}{F(x|z_k)} &= \lim_{x\to-\infty} \frac{F_\varepsilon(h^{-1}(z_{k+1},x))}{F_\varepsilon(h^{-1}(z_k,x))} \\ &= \lim_{x\to-\infty} \frac{F_\varepsilon(h^{-1}(z_k,x) - (h^{-1}(z_k,x) - h^{-1}(z_{k+1},x)))}{F_\varepsilon(h^{-1}(z_k,x))}.\end{aligned}$$

Since the function $h(z,\varepsilon)$ is strictly increasing in ε for $z \in \mathcal{N}(z_0)$, this implies that

$$\begin{aligned}&h^{-1}(z_k,x) - h^{-1}(z_{k+1},x) \\ \equiv\ &\varepsilon_k - h^{-1}(z_{k+1}, h(z_k,\varepsilon_k)) \\ =\ &\varepsilon_k - h^{-1}(z_{k+1}, h(z_{k+1},\varepsilon_k) + [h(z_k,\varepsilon_k) - h(z_{k+1},\varepsilon_k)]).\end{aligned}$$

Since $\frac{\partial}{\partial z} h(z_0,\varepsilon) < 0$ and $z_k \uparrow z_0$, we have $h(z_k,\varepsilon_k) - h(z_{k+1},\varepsilon_k) = c' > 0$ for $\varepsilon_k \to -\infty$.

Using $h(z,\varepsilon)$ is strictly increasing in ε, we have

$$\begin{aligned} & h^{-1}(z_k,x)-h^{-1}(z_{k+1},x) \\ = \;& \varepsilon_k - h^{-1}(z_{k+1},h(z_{k+1},\varepsilon_k)+c') \\ = \;& \varepsilon_k - h^{-1}(z_{k+1},h(z_{k+1},\varepsilon_k))+c \\ = \;& \varepsilon_k-\varepsilon_k+c\neq 0 \end{aligned}$$

for some constant $c>0$ as $\varepsilon_k\to-\infty$. Given F_ε is increasing, we may pick z_k such that $c>0$ to have

$$\lim_{x\to-\infty}\frac{F(x|z_{k+1})}{F(x|z_k)}=\lim_{x\to-\infty}\frac{F_\varepsilon(h^{-1}(z_k,x)-c)}{F_\varepsilon(h^{-1}(z_k,x))}=0.$$

The last step is because F_ε satisfies $\lim_{x\to-\infty}\frac{F_\varepsilon(x-c)}{F_\varepsilon(x)}=\lim_{x\to-\infty}\frac{f_\varepsilon(x-c)}{f_\varepsilon(x)}=0$ by L'Hôpital's rule and ε is normally distributed. Therefore, condition (2) in Lemma 3 holds. Theorem 1 then implies that completeness of $\{f(\cdot|z):z\in\mathcal{N}(z_0)\}$ in $\mathcal{L}^2(\mathbb{R},\omega)$. QED.

Proof of Lemma 6: Without loss of generality, we consider $x=(x_1,x_2)$, $z=(z_1,z_2)$, $\mathcal{X}=\mathcal{X}_1\times\mathcal{X}_2$, and $\mathcal{Z}=\mathcal{Z}_1\times\mathcal{Z}_2$. Condition i) implies that $\{f_{x_1|z_1}(\cdot|z_{1k}):k=1,2,3,...\}$ and $\{f_{x_2|z_2}(\cdot|z_{2k}):k=1,2,3,...\}$ are complete in their corresponding Hilbert spaces.

We then show the sequence $\{f_{x_1|z_1}(\cdot|z_{1k})f_{x_2|z_2}(\cdot|z_{2k}):k=1,2,3,...\}$ is complete because $\{f_{x_1|z_1}(\cdot|z_{1k}):k=1,2,3,...\}$ and $\{f_{x_2|z_2}(\cdot|z_{2k}):k=1,2,3,...\}$ are complete in corresponding Hilbert spaces. Consider

$$\begin{aligned}\int\int h(x_1,x_2)f(x_1|z_1)f(x_2|z_2)dx_1dx_2 &= \int\left(\int h(x_1,x_2)f(x_1|z_1)dx_1\right)f(x_2|z_2)dx_2 \\ &\equiv \int h'(x_2,z_1)\,f(x_2|z_2)dx_2.\end{aligned}$$

If the LHS is equal to zero for any $(z_1,z_2)\in\mathcal{Z}_1\times\mathcal{Z}_2$, then for any given z_1 $\int h'(x_2,z_1)\,f(x_2|z_2)dx_2$ equals to zero for any z_2. Since $f(x_2|z_2)$ is complete, we have $h'(x_2,z_1)=0$ for any $x_2\in\mathcal{X}_2$ and any given $z_1\in\mathcal{Z}_1$. Furthermore, for any given $x_2\in\mathcal{X}_2$, $h'(x_2,z_1)=0$ for any $z_1\in\mathcal{Z}_1$ implies $h(x_1,x_2)=0$ for any $x_1\in\mathcal{X}_1$. Therefore, the sequence $\{f_{x_1|z_1}(\cdot|z_{1k})f_{x_2|z_2}(\cdot|z_{2k}):k=1,2,3,...\}$ is complete. We then apply Theorem 1 to show that the sequence $\{f_{x_1,x_2|z_1,z_2}(\cdot,\cdot|z_{1k},z_{2k}):k=1,2,3,...\}$ is complete because it is close to a complete sequence $\{f_{x_1|z_1}(\cdot|z_{1k})f_{x_2|z_2}(\cdot|z_{2k}):k=1,2,3,...\}$. QED.

References

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

An, Y., and Y. Hu (2009): "Well-posedness of Measurement Error Models for Self-reported Data," *CeMMAP working papers*.

Andrews, D. (2011): "Examples of L^2-Complete and Boundedly-Complete Distributions," *Cowles Foundation for Research in Economics*.

Blundell, R., X. Chen, and D. Kristensen (2007): "Semi-nonparametric IV Estimation of Shape-Invariant Engel Curves," *Econometrica*, 75(6), 1613.

Canay, I., A. Santos, and A. Shaikh (2011): "On the Testability of Identification in Some Nonparametric Models with Endogeneity," *Working Paper*.

Carroll, R., X. Chen, and Y. Hu (2010): "Identification and Estimation of Nonlinear Models Using Two Samples with Nonclassical Measurement Errors," *Journal of Nonparametric Satistics*, 22(4), 379–399.

Chen, X., and Y. Hu (2006): "Identification and Inference of Nonlinear Models Using Two Samples With Arbitrary Measurement Errors," *Cowles Foundation Discussion Paper No. 1590*.

Chernozhukov, V., and C. Hansen (2005): "An IV Model of Quantile Treatment Effects," *Econometrica*, 73(1), 245–261.

Chernozhukov, V., G. Imbens, and W. Newey (2007): "Instrumental Variable Estimation of Nonseparable Models," *Journal of Econometrics*, 139(1), 4–14.

Darolles, S., Y. Fan, J. Florens, and E. Renault (2011): "Nonparametric Instrumental Regression," *Econometrica*, 79(5), 1541–1565.

D'Haultfoeuille, X. (2011): "On the Completeness Condition in Nonparametric Instrumental Problems," *Econometric Theory*, 1, 1–12.

Erdös, P., and E. Straus (1953): "On Linear Independence of Sequences in a Banach Space," *Pacific J. Math*, 3, 689–694.

FLORENS, J., M. MOUCHART, AND J. ROLIN (1990): *Elements of Bayesian Statistics.* New York: Marcel Dekker.

HALL, P., AND J. HOROWITZ (2005): "Nonparametric Methods for Inference in the Presence of Instrumental Variables," *The Annals of Statistics*, 33(6), 2904–2929.

HODERLEIN, S., L. NESHEIM, AND A. SIMONI (2010): "Semiparametric Estimation of Random Coefficients in Structural Economic Models," *Working Paper.*

HOROWITZ, J. (2011): "Applied Nonparametric Instrumental Variables Estimation," *Econometrica*, 79(2), 347–394.

HOROWITZ, J., AND S. LEE (2007): "Nonparametric Instrumental Variables Estimation of a Quantile Regression Model," *Econometrica*, 75(4), 1191–1208.

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

HU, Y., AND M. SHUM (2009): "Nonparametric Identification of Dynamic Models with Unobserved State Variables," *Jonhs Hopkins University, Dept. of Economics Working Paper*, 543.

LEHMANN, E. (1986): *Testing Statistical Hypotheses.* 2nd ed. New York: Wiley.

MATTNER, L. (1996): "Complete Order Statistics in Parametric Models," *The Annals of Statistics*, 24(3), 1265–1282.

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

RUDIN, W. (1987): *Real and Complex Analysis.* McGraw-Hill.

——— (1991): *Functional Analysis.* McGraw-Hill, Inc., New York.

SAN MARTIN, E., AND M. MOUCHART (2007): "On Joint Completeness: Sampling and Bayesian Versions, and Their Connections," *The Indian Journal of Statistics*, 69(4), 780–807.

Shiu, J., and Y. Hu (2010): "Identification and Estimation of Nonlinear Dynamic Panel Data Models with Unobserved Covariates," *Economics Working Paper Archive.*

Young, R. (1980): *An Introduction to Nonharmonic Fourier Series.* Academic Press.