Econometric Theory, **34**, 2018, 659–693.
doi:10.1017/S0266466617000251

NONPARAMETRIC IDENTIFICATION USING INSTRUMENTAL VARIABLES: SUFFICIENT CONDITIONS FOR COMPLETENESS

YINGYAO HU
Johns Hopkins University

JI-LIANG SHIU
Jinan University

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 show that (1) if the deviation of the conditional density $f(x|z_k)$ from a known complete sequence of functions is less than a sequence of values determined by the complete sequence in some distinct sequence $\{z_k : k = 1, 2, 3, \ldots\}$ converging to z_0, then $f(x|z)$ itself is complete, and (2) if the conditional density $f(x|z)$ can form a linearly independent sequence $\{f(\cdot|z_k) : k = 1, 2, \ldots\}$ in some distinct sequence $\{z_k : k = 1, 2, 3, \ldots\}$ converging to z_0 and its relative deviation from a known complete sequence of functions under some norm is finite then $f(x|z)$ itself is complete. We use these general results to provide specific sufficient conditions for completeness in three different specifications of the relationship between the endogenous regressor x and the instrumental variable z.

1. INTRODUCTION

We consider a nonparametric regression model as follows:

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

where y is an observable scalar random variable, and x is a $d_x \times 1$ vector of regressors and may be correlated with a zero mean regression error u. The parameter of interest is the nonparametric regression function $m(\cdot)$. A $d_z \times 1$ vector of

We thank Peter C.B. Phillips, two associate editors, and two anonymous referees for their time and helpful comments. We also thank the participants at the International Economic Association 16th World Congress, Beijing, China, and the 2013 Annual Meeting of American Economic Association, San Diego, USA, for various suggestions regarding this paper. The usual disclaimer applies. Address correspondence to Yingyao Hu, Department of Economics, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218, USA; e-mail: yhu@jhu.edu.

instrumental variables 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.[1] We observe a random sample of $\{y, x, z\}$, and denote the support of these random variables as $\mathcal{Y}$, $\mathcal{X}$, 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 (1) if the deviation of the conditional density $f(x|z_k)$ from a known complete sequence of functions is less than a sequence of values determined by the complete sequence in some distinct sequence $\{z_k : k = 1, 2, 3, \ldots\}$ converging to z_0, then $f(x|z)$ itself is complete, and (2) if the conditional density $f(x|z)$ can form a linearly independent sequence $\{f(\cdot|z_k) : k = 1, 2, \ldots\}$ for some distinct sequence $\{z_k : k = 1, 2, 3, \ldots\}$ converging to z_0 and its relative deviation from a known complete sequence of functions under some norm is finite then $f(x|z)$ itself is complete. 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.

The nonparametric IV regression model is applicable for a large range of empirical research. We provide a few examples in which there may exist an endogenous variable and the model is applicable. Consider an empirical study for estimating the impact of education on the female labor supply. The endogeneity may arise in the presence of an "ability bias" or the measurement error problem of education. We may use father's education as an instrumental variable for woman's education. Another example is to estimate Engel curves that describe the allocation of total nondurable consumption expenditure. In the application, total expenditure can be endogenous and we may use the gross earnings of the household head as an instrument for total expenditure.

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 $L^2(\mathcal{X}, \omega) = \{h(\cdot) : \int_{\mathcal{X}} |h(x)|^2 \omega(x) dx < \infty\}$ 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 $\int_{\mathcal{X}} \omega(x) dx < \infty$.[2] The corresponding norm is defined as: $\|f\|^2 = \langle f, f \rangle$. The space $L^2(\mathcal{X}, \omega)$ is complete under the norm $\|\cdot\|$ and 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 question of completeness with the set $\mathcal{O}$ being a distinct converging sequence $\{z_k : k = 1, 2, 3, \ldots\}$ in $\mathcal{Z}$. This case corresponds to a sequence of functions $\{f(\cdot|z_k) : k = 1,2,\ldots\}$. We start with a definition of completeness in a Hilbert space $\mathcal{H}$.

DEFINITION 1. *Denote $\mathcal{H} \equiv L^2(\mathcal{X},\omega)$ as a Hilbert space with the weight function ω. 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 all $z \in \mathcal{O}$, $\int_{\mathcal{X}} \frac{f(x|z)^2}{\omega(x)} dx < \infty$ and for any $h(\cdot) \in \mathcal{H}$*

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

implies $h(\cdot) = 0$ almost surely in $\mathcal{X}$.[3] *When it is a conditional density function defined on $\mathcal{X} \times \mathcal{Z}$, $f(x|z)$ is said to be a complete density.*[4]

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.

This definition implies that the equality in the definition can be rewritten as

$$0 = \int_{\mathcal{X}} h(x) f(x|z) dx = \int_{\mathcal{X}} h(x) \frac{f(x|z)}{\omega(x)} \omega(x) dx = \left\langle h, \frac{f(\cdot|z)}{\omega(\cdot)} \right\rangle \quad \text{for all } z \in \mathcal{O}.$$

Therefore, we use the weighted function $\frac{f(x|z)}{\omega(x)}$ instead of $f(x|z)$ in the inner product of the Hilbert space $L^2(\mathcal{X},\omega)$, when we consider a complete sequence in the Hilbert space in the appendix. The definition certainly imposes tail conditions on the conditional density function $f(x|z)$. On the other hand, because existing complete distribution functions used in this study are the exponential family and a translated density function with exponentially decaying tails, the definition is not very restrictive to the nonparametric extension of completeness from these existing complete distribution functions.

The completeness introduced in Definition 1 is close to L^2- completeness considered in Andrews (2012) with $\mathcal{H} = L^2(\mathcal{X}, f_x)$, where the density f_x may be considered as the weight function ω in $L^2(\mathcal{X},\omega)$.[5] Andrews (2012) provides broad (nonparametric) classes of L^2-complete distributions that can have

https://doi.org/10.1017/S0266466617000251 Published online by Cambridge University Press

any marginal distributions and a wide range of strengths of dependence. The L^2-complete distributions are constructed by bivariate density functions with respect to $F_x \times F_z$ which are constructed through orthonormal bases of $L^2(F_x)$ and $L^2(F_x)$. Depending on which regularity conditions are imposed on the regression function $m(\cdot)$, a different version of completeness can be also considered. For example, D'Haultfoeuillle (2011) considers three different types of completeness including (1) "standard" completeness, where h satisfies $E(|h(X)|) < \infty$, (2) P-completeness, where h is bounded by a polynomial, and (3) bounded completeness for any bounded h in nonparametric models between the two variables with an additive separability and a large support condition. D'Haultfoeuillle (2011) defines completeness in terms of dependence condition between x and z such as $x = \mu(\nu(z)+\varepsilon)$, where μ and ν are mappings and z and ε are independent. The results are useful in nonparametric regression models with a limited endogenous regressor. 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 $L^2(\mathcal{X},\omega)$ is more informative for identification than the bounded completeness because a bounded function always belongs to the weighted L^2 space $L^2(\mathcal{X},\omega)$.[6] Therefore, we consider L^2-completeness with a Hilbert space $\mathcal{H} = L^2(\mathcal{X},\omega)$ 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.[7] Our results for continuous variables extend this interpretation. Suppose that the family of conditional distributions in $\{f(\cdot|z_k) : k = 1,2,\ldots\}$ is complete in $L^2(\mathcal{X},\omega)$. As shown in the Appendix, we can extract a subfamily $\{f(\cdot|z_{r_k}) : k = 1,2,\ldots\}$ as a basis in $L^2(\mathcal{X},\omega)$. This basis interpretation implies that (1) there is no exact linear relationship among the family of the conditional distribution $\{f(\cdot|z_{r_k}) : k = 1,2,\ldots\}$ or a conditional distribution at each point z can not be expressed as a linear combination of others, and (2) every function in $L^2(\mathcal{X},\omega)$ can be approximated by linear combinations of the conditional distributions in $\{f(\cdot|z_{r_k}) : k = 1,2,\ldots\}$. In this general continuous case, a function in $L^2(\mathcal{X},\omega)$ may be expressed as an infinite sum of functions and the convergence of the infinite sum is under the norm $\|\cdot\|$.

The L^2 completeness for the nonparametric regression model (1) implies 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 $L^2(\mathbb{R},\omega)$, the regression function m can be identified within the set of functions of the form $\{\alpha + \beta x + g(x) : g \in L^2(\mathbb{R},\omega)\}$. Therefore, under our framework the functional form of the regression function m may be very flexible. Notice that the function g cannot 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 following two examples of complete $f(x|z)$ are from Newey and Powell (2003) (See their Theorems 2.2 and 2.3 for details.[8]):

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 $E[h(x)|z=z_1]=0$ for any $z_1\in\mathcal{Z}$ implies $h(\cdot)=0$ almost surely in $\mathcal{X}$; equivalently, $\{f(\cdot|z):z\in\mathcal{Z}\}$ is complete.[9]

Another case where the family $\{f(x|z):z\in\mathcal{O}\}$ 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 $s(x)>0$, the mapping from $x\rightarrow\tau(x)$ is one-to-one in x, and support of $\mu(z)$, $\mathcal{Z}$, contains an open set, then $E[h(x)|z=z_1]=0$ for any $z_1\in\mathcal{Z}$ implies $h(\cdot)=0$ almost surely in $\mathcal{X}$; equivalently, the family of conditional density functions $\{f(\cdot|z):z\in\mathcal{Z}\}$ 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), and Horowitz (2011)), semiparametric IV models (see Ai and Chen (2003) and Blundell, Chen, and Kristensen (2007)), nonparametric IV quantile models (see Chernozhukov and Hansen (2005), Chernozhukov, Imbens, and Newey (2007), and Horowitz and Lee (2007)), measurement error models (see Hu and Schennach (2008), An and Hu (2012), Carroll, Chen, and Hu (2010), and Chen and Hu (2006)), random coefficient models (see Hoderlein, Nesheim, and Simoni (2012)), and dynamic models (see Hu and Shum (2012) and Shiu and Hu (2013)), etc. We refer to D'Haultfoeuille (2011) and Andrews (2012) for more complete literature reviews. On the other hand, Canay, Santos, and Shaikh (2013) consider hypothesis testing for completeness against very general alternatives and they show that the completeness condition is, without further restrictions, untestable.

There are cases where identification and consistent estimation are relatively straightforward and not related to completeness. In the nonstationary dependent case, identification can be achieved because of the nonstationary nature of the regressor, which can act as its own instrument (see Wang and Phillips (2009), Wang and Phillips (2016), and Wang and Phillips (2007)). In the microeconometric context, identification can also be achieved by infill and spatial shifting nonstationarity (see Phillips and Su (2011)). Finally, in threshold regression cases thresholding parameters are identifiable and consistently estimable in spite of endogeneity of the regressor (see Yu and Phillips (2017)).

In this paper, we provide sufficient conditions for the completeness of a general conditional density without imposing particular functional forms. We first show the set $\mathcal{O}$ of the family $\{f(\cdot|z) : z \in \mathcal{O}\}$ in the definition of completeness can be as small as a converging sequence $\{z_k : k = 1,2,3,\ldots\}$ for some known complete families. This implies that the family $\left\{f_{x|z}(\cdot|z_k)/\omega(\cdot) : k = 1,2,\ldots\right\}$ can form a complete sequence in a Hilbert space with the weight function ω. We then use the stability properties of complete sequences in a Banach space and a Hilbert space (Sections 9 and 10 of Chapter 1 in Young (2001) and Gurarij and Meletidi (1970)) to show that (1) if the deviation of the conditional density $f(x|z_k)$ from a known complete sequence of functions is less than a sequence of values determined by the complete sequence in some distinct sequence $\{z_k : k = 1,2,3,\ldots\}$ converging to z_0, then $f(x|z)$ itself is complete, and (2) if the conditional density $f(x|z)$ can form a linearly independent sequence $\{f(\cdot|z_k) : k = 1,2,\ldots\}$ for some distinct sequence $\{z_k : k = 1,2,3,\ldots\}$ converging to z_0 and its relative deviation from a known complete sequence of functions under some norm is finite then $f(x|z)$ itself is complete.

Another observation is that this stability property of completeness is actually consistent with a result in Canay et al. (2013) that a distribution for which completeness fails can be arbitrarily close to distributions for which completeness holds. Notice that Canay et al. (2013) approach $f(X,Z)$ by a sequence of step functions (with a finite number of "steps"). The limit may be a complete function but each step function in the sequence cannot be complete. Each of the step functions in their sequence corresponds to a truncated sequence $\{f_1, f_2, \ldots, f_K\}$ with $f_k = f_{X|Z}(.|Z = z_k)$ for some finite K in our paper. The limit of this sequence $\{f_1, f_2, \ldots, f_K\}$ is $\{f_k\}_{k=1,2,\ldots}$, which may be complete. But the truncated sequence $\{f_1, f_2, \ldots, f_K\}$ is not complete with a continuous X. The key issue is approximation to $f(X,Z)$ from both dimensions of X and Z. If $f(X,Z)$ is only approximated by step functions of X, but not of Z, completeness may still hold because a sequence of step functions, such as the so-called Haar sequence, may still be complete.

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 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

https://doi.org/10.1017/S0266466617000251 Published online by Cambridge University Press

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 (1) if the deviation of the conditional density $f(x|z_k)$ from a known complete sequence of functions is less than a sequence of values determined by the complete sequence in some distinct sequence $\{z_k : k = 1,2,3,\ldots\}$ converging to z_0, then $f(x|z)$ itself is complete, and (2) if the conditional density $f(x|z)$ can form a linearly independent sequence $\{f(\cdot|z_k) : k = 1,2,\ldots\}$ for some distinct sequence $\{z_k : k = 1,2,3,\ldots\}$ converging to z_0 and its relative deviation from a known complete sequence of functions under some norm is finite then $f(x|z)$ itself is complete. 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 $\{g(\cdot|z_k) : k = 1,2,\ldots\}$ 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 forms of these. Another family of conditional distributions which implies completeness is in the form of a translated density function, i.e., $g(x|z) = g(x - z)$.[10]

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 $\mathcal{Z}$.[11]

LEMMA 1. *Suppose that $\mathcal{X}$ is a connected set. Denote $\mathcal{O}$ as an open set in $\mathcal{Z} \subset \mathbb{R}$. Let $\{z_k : k = 1,2,\ldots\}$ be a sequence of distinct $z_k \in \mathcal{O}$ converging to z_0 in the open set $\mathcal{O}$. Define*

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

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

i) $\mu(\cdot)$ is continuous differentiable with $\mu'(z_0) \neq 0$;
ii) $\tau(\cdot)$ is C^1-diffeomorphism from $\mathcal{X}$ to $\tau(\mathcal{X})$.[12]

Then, the sequence $\{g(\cdot|z_k) : k = 1,2,\ldots\}$ is complete in $L^2(\mathcal{X},\omega)$, 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. ■

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

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 implies 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.

In order to allow the space of the regression functions m to contain linear functions and polynomials, we consider $\mathcal{X} = \mathbb{R}$ with some weight function (Lemma 2). We summarize the results as follows.

LEMMA 2. *Let $\mathcal{X} = \mathbb{R}$. Denote $\mathcal{O}$ as an open set in $\mathcal{Z}$. Let $\{z_k : k = 1, 2, \ldots\}$ 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 $\mathbb{R} \times \mathcal{Z}$. If the distribution of ε is normal, then the sequence $\{g(\cdot|z_k) : k = 1, 2, \ldots\}$ is complete in $L^2(\mathbb{R}, \omega)$, where the weight function $\omega(x)$ satisfies $\omega(x) = e^{-\delta' x^2}$ for some $\delta' \in (0, \delta_3)$ and δ_3 is defined in equation (A.4).

Proof. See the appendix. ■

We need the exponentially decaying weight function to identify the possibly unbounded regression function, such as linear functions, over the whole real line in a L^2 functional space. However, this may limit the distribution of ε to be normal. Theorem 2.4 of Mattner (1993) shows that the only L^1 complete location family over a real line is the Gaussian one.[13] One remedy is to impose more restrictions on the functional space, such as bounded functions over the whole real line. As Theorem 2.1 of Mattner (1993) shows, a location family is boundedly complete if and only if its characteristic function is not zero in the whole real line. This can allow ε to have various distributions. Unfortunately, this rules out linear functions over the whole real line.

With the complete sequences explicitly specified in Lemmas 1 and 2, we are ready to extend the completeness to a more general conditional density $f(x|z)$. We will apply three different types of the stability results in Hilbert spaces to obtain the completeness. The first two are only involved with a small perturbation of a complete sequence and the third one is linked to an added structure of a Hilbert space, an orthonormal basis. The existence of such stability is based on two facts: (1) completeness is preserved by an invertible operator and (2) a bounded linear operator T on a Hilbert space is invertible whenever

$\|I - T\|_{op} < 1$, where $\|\cdot\|_{op}$ is an operator norm. Our first two sufficient conditions for completeness are summarized as follows:

THEOREM 1. *Denote $\mathcal{H} \equiv L^2(\mathcal{X}, \omega)$. 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 the Hilbert space $\mathcal{H}$ of functions defined on $\mathcal{X}$ with norm $\|\cdot\|$. Set $\mathcal{N}(z_0) = \{z \in \mathcal{Z} : \|z - z_0\| < \varepsilon$ for some small $\varepsilon > 0\} \subseteq \mathcal{Z}$ as an open neighborhood for a point z_0 such that*

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

ii) there exists a complete sequence $\{g(\cdot|z_k) : k = 1, 2, \ldots\}$ such that $f(\cdot|z)$ satisfies

$$\sum_{k=1}^{n} \|c_k(g(\cdot|z_k)/\omega(\cdot) - f(\cdot|z_k)/\omega(\cdot))\| < \lambda \sum_{k=1}^{n} \|c_k g(\cdot|z_k)/\omega(\cdot)\|$$

for some constant λ, $0 \le \lambda < 1$, and arbitrary scalars $c_1, \ldots, c_n$ $(n = 1, 2, 3, \ldots)$.

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

Proof. See the appendix. ■

THEOREM 2. *Denote $\mathcal{H} \equiv L^2(\mathcal{X}, \omega)$. 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 the Hilbert space $\mathcal{H}$ of functions defined on $\mathcal{X}$ with norm $\|\cdot\|$. Set $\mathcal{N}(z_0) = \{z \in \mathcal{Z} : \|z - z_0\| < \varepsilon$ for some small $\varepsilon > 0\} \subseteq \mathcal{Z}$ as an open neighborhood for a point z_0 such that*

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

ii) there exists a sequence of positive numbers $\{C_k : k = 1, 2, \ldots\}$ which depends on the normalized sequence $\{\frac{g(\cdot|z_k)/\omega(\cdot)}{\|g(\cdot|z_k)/\omega(\cdot)\|} : k = 1, 2, \ldots\}$ such that $\sum_{k=1}^{n} C_k \varepsilon_k < 1$ for a sequence of positive numbers $\{\varepsilon_k : k = 1, 2, \ldots\}$ and

$$\left\| \frac{g(\cdot|z_k)/\omega(\cdot)}{\|g(\cdot|z_k)/\omega(\cdot)\|} - f(\cdot|z_k)/\omega(\cdot) \right\| < \varepsilon_k.$$

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

Proof. See the appendix. ■

Theorem 2 implies that a new complete sequence always exists, although its distance from the existing complete sequence is determined by that sequence. The second stability criteria related to an orthonormal basis is the following.

https://doi.org/10.1017/S0266466617000251 Published online by Cambridge University Press

THEOREM 3. *Denote $\mathcal{H} \equiv L^2(\mathcal{X}, \omega)$. For every $z \in \mathcal{Z}$, let $f(\cdot|z)$ and $g(\cdot|z)$ be conditional densities in the Hilbert space $\mathcal{H}$ of functions defined on $\mathcal{X}$ with norm $\|\cdot\|$. Set $\mathcal{N}(z_0) = \{z \in \mathcal{Z} : \|z - z_0\| < \varepsilon$ for some small $\varepsilon > 0\} \subseteq \mathcal{Z}$ as an open neighborhood for a point z_0 such that*

i) *for every sequence $\{z_k : k = 1, 2, \ldots\}$ of distinct $z_k \in \mathcal{N}(z_0)$ converging to z_0, the corresponding sequence $\{g(\cdot|z_k) : k = 1, 2, \ldots\}$ is complete in a Hilbert space $\mathcal{H}$;*
ii) *there exists a complete sequence $\{g(\cdot|z_k) : k = 1, 2, \ldots\}$ such that the corresponding sequence $\{f(\cdot|z_k) : k = 1, 2, \ldots\}$ satisfies that*

$$\sum_{k=1}^{\infty} \frac{\left\| v_k^g - v_k^f \right\|^2}{\left\| v_k^g \right\|^2} < \infty,$$

where for $h \in \{g, f\}$, the sequence of functions v_k^h is defined as $v_1^h(\cdot) = h(\cdot|z_1)/\omega(\cdot), \ldots, v_k^h(\cdot) = h(\cdot|z_k)/\omega(\cdot) - \sum_{j=1}^{k-1} \frac{\left\langle h(\cdot|z_{r_k})/\omega(\cdot), v_j^h(\cdot)\right\rangle}{\left\langle v_j^h(\cdot), v_j^h(\cdot)\right\rangle} v_j^h$, and that for any finite subsequence $\left\{z_{k_i} : i = 1, 2, \ldots, I\right\}$ $\{f(\cdot|z_{k_i}) : i = 1, 2, \ldots, I\}$ is linearly independent, i.e.,

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

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

Proof. See the appendix. ∎

This theorem utilizes the structure of an inner product that allows length and angle in a Hilbert space. We show that if the distance between the two corresponding orthogonal sequences is finite and the new sequence is linearly independent, then the new sequence is complete.

Condition i) provides complete sequences, which may be from Lemmas 1, and 2. Condition ii) requires that the total sum of relative quadratic deviation from the orthogonal sequence $\left\{v_k^g : k = 1, 2, \ldots\right\}$ constructed by $\{g(\cdot|z_{r_k})/\omega(\cdot) : k = 1, 2, \ldots\}$ and an orthogonal sequence $\left\{v_k^f : k = 1, 2, \ldots\right\}$ constructed by $\{f(\cdot|z_{r_k})/\omega(\cdot) : k = 1, 2, \ldots\}$ is finite.

The linear independence in condition iii) imposed on $\{f(\cdot|z_k)\}$ implies that there are no redundant terms in the sequence in the sense that no term can be expressed as a linear combination of some other terms. Because a weight function is positive, the linear independence of $\{f(\cdot|z_k)\}$ is equivalent to the linear independence of $\{f(\cdot|z_k)/\omega(\cdot)\}$. For simplification, we use an ordered sequence z_k. 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

$$\lim_{x\to-\infty}\frac{f(x|z_{k+1})}{f(x|z_k)}=0 \text{ for all } k, \tag{4}$$

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 $\sum_{i=1}^{I}c_i f(x|z_{k_i})=0$ for all $x\in(-\infty,+\infty)$, we may have

$$-c_1=\sum_{i=2}^{I}c_i\frac{f(x|z_{k_i})}{f(x|z_{k_1})}.$$

The limit of the right-hand side is zero as $x\to-\infty$ so that $c_1=0$. Similarly, we may show $c_2,c_3,\ldots,c_I=0$ for all i by induction. Notice that the exponential family satisfies equation (4) for appropriate choices of μ, τ, and a sequence. When the support $\mathcal{X}$ is bounded, for example, $\mathcal{X}=[0,1]$, the condition (4) may become

$$\lim_{x\to 0}\frac{f(x|z_{k+1})}{f(x|z_k)}=0 \text{ for all } k. \tag{5}$$

For example, the Corollary (Müntz) on page 91 in Young (2001) implies that the family of function $\{x^{z_1},x^{z_2},x^{z_3},\ldots\}$ is complete in $L^2([0,1])$ if $\sum\limits_{k=1}^{\infty}\frac{1}{z_k}=\infty$. This family also satisfies the condition (5) 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 (4), i.e., $\lim_{x\to-\infty}\frac{g(x|z_{k+1})}{g(x|z_k)}=0$, then the condition $0<\left(\lim_{x\to-\infty}\frac{f(x|z_k)}{g(x|z_k)}\right)<\infty$ implies $\lim_{x\to-\infty}\frac{f(x|z_{k+1})}{f(x|z_k)}=0$ or linear independence of $\{f(\cdot|z_k)\}$. Furthermore, when $f(x|z)=h(x|z)\times g(x|z)$, the condition (4) is implied by $\lim_{x\to-\infty}\frac{g(x|z_{k+1})}{g(x|z_k)}=0$ and $\left(\lim_{x\to-\infty}\frac{h(x|z_{k+1})}{h(x|z_k)}\right)<\infty$.

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

$$W(x)=\det\begin{pmatrix} f(x|z_{k_1}) & f(x|z_{k_2}) & \ldots & f(x|z_{k_I})\\ f'(x|z_{k_1}) & f'(x|z_{k_2}) & \ldots & f'(x|z_{k_I})\\ \ldots & \ldots & \ldots & \ldots\\ \frac{d^{(I-1)}}{dx^{(I-1)}}f(x|z_{k_1}) & \frac{d^{(I-1)}}{dx^{(I-1)}}f(x|z_{k_2}) & \ldots & \frac{d^{(I-1)}}{dx^{(I-1)}}f(x|z_{k_I})\end{pmatrix} \tag{6}$$

If there exists an x_0 such that the determinant $W(x_0)\neq 0$ for every $\left\{z_{k_i}:i=1,2,\ldots,I\right\}$, 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 $\left\{z_{k_i}:i=1,2,\ldots,I\right\}$, where $G_f=\det\left(\left[\left\langle f(\cdot|z_{k_i}),\, f(\cdot|z_{k_j})\right\rangle\right]_{i,j}\right)$. This condition does not require that the function has all derivatives. We summarize these results on the linear independence as follows:

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

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) *there exists an* x_0 *such that the determinant* $W(x_0) \neq 0$ *for every* $\{z_{k_i} : i = 1, 2, \ldots, I\}$. *In particular,* $\frac{d^k}{dx^k} F_0(0) \neq 0$ *for* $k = 1, 2, \ldots$ *if* $f(x|z) = \frac{d}{dx} F_0(\mu(z)\tau(x))$ *with* $\mu'(z_0) \neq 0$, $\tau(0) = 0$.
4) *for every* $\{z_{k_i} : i = 1, 2, \ldots, I\}$, $\det\left(\left[\langle f(\cdot|z_{k_i}), f(\cdot|z_{k_j})\rangle\right]_{i,j}\right) \neq 0$.

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. Theorem 4 in Gurarij and Meletidi (1970) also shows that if a sequence $\{f_n : n = 1, 2, 3, \ldots\}$ satisfies $\|f_n - x^n\| < \epsilon_n$ for any ϵ_n such that $\lim \epsilon_n a^n = 0$ for every positive a then the sequence is also complete. Thus, we may consider $g(x|z) = x^z$ and $f(x|z) = x^z + z^{-z}x^2$ and then $\{f(\cdot|z) : z \in \mathcal{Z}\}$ is complete for $\mathcal{Z} = \mathbb{R}$. Figure 1 presents a 3D graph of $g(x|z)$ and $f(x|z)$ for (x, z) in $[0, 1] \times [1, 4]$ to illustrate the relationship between the complete sequence $\{g(\cdot|z_k)\}$ and the sequence $\{f(\cdot|z_k)\}$.

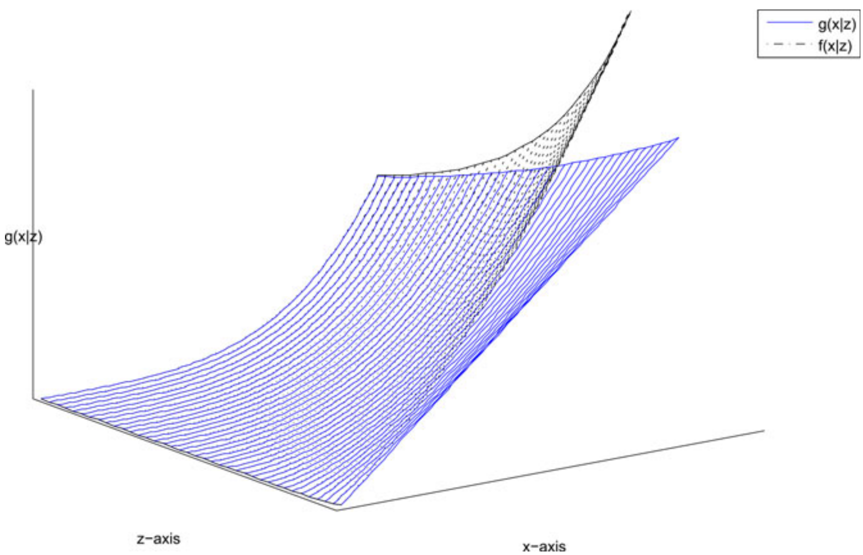

FIGURE 1. An example of $g(x|z)$ and $f(x|z)$ in Theorem 2.

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.

3.1. Extension of the Convolution Case

Lemma 2 provides a complete sequence when $x = z + \varepsilon$. Using Theorems 1, 2 and 3, 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:

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

Without loss of generality, we set

$\mu(z) = z$.

We summarize the result as follows:

LEMMA 4. *For every $z \in \mathcal{Z} \subset \mathbb{R}$, let $f(\cdot|z)$ be in $L^1(\mathbb{R})$. Suppose that there exists a point z_0 with its open neighborhood $\mathcal{N}(z_0) \subseteq \mathcal{Z}$ such that*

i) *set $f_\varepsilon(\cdot) = f(\cdot|z_0)$ and the function f_ε is normally distributed as in Lemma 2;*

ii) *there exists $\{f(\cdot|z_k) : k = 1,2,\ldots\}$ such that one of the following conditions holds:*

    1) *$\sigma(z) = 1$ for $\|z - z_0\| < \varepsilon$ for some small $\varepsilon > 0$;*

    2) *$\sigma(z)$ satisfies*

$$\sum_{k=1}^{n} \left\| c_k (f_\varepsilon(\cdot - z_k)/\omega(\cdot) - \frac{1}{\sigma(z_k)} f_\varepsilon\left(\frac{\cdot - z_k}{\sigma(z_k)}\right)/\omega(\cdot)) \right\| < \lambda \sum_{k=1}^{n} \| c_k f_\varepsilon(\cdot - z_k)/\omega(\cdot) \|$$

    for some constant λ, $0 \le \lambda < 1$, and arbitrary scalars $c_1, \ldots, c_n$ $(n = 1,2,3,\ldots)$;

    3) *there exists a sequence $\{C_k : k = 1,2,\ldots\}$ which depends on the normalized sequence $\{\frac{f_\varepsilon(\cdot - z_k)/\omega(\cdot)}{\|f_\varepsilon(\cdot - z_k)/\omega(\cdot)\|} : k = 1,2,\ldots\}$ such that $\sum_{k=1}^{n} C_k \varepsilon_k < 1$ for a sequence of positive numbers $\{\varepsilon_k : k = 1,2,\ldots\}$ and*

$$\left\| \frac{f_\varepsilon(\cdot - z_k)/\omega(\cdot)}{\|f_\varepsilon(\cdot - z_k)/\omega(\cdot)\|} - \frac{1}{\sigma(z_k)} f_\varepsilon\left(\frac{\cdot - z_k}{\sigma(z_k)}\right)/\omega(\cdot) \right\| < \varepsilon_k,$$

4) $\sigma(z)$ *satisfies*

$$\sum_{k=1}^{\infty} \frac{\left\| \upsilon_k^f - \upsilon_k^{f_\sigma} \right\|^2}{\left\| \upsilon_k^f \right\|^2} < \infty,$$

where υ_k^f *and* $\upsilon_k^{f_\sigma}$ *are defined as in Theorem 3 with* $f(x) = f_\varepsilon(x-z)/\omega(\cdot)$ *and* $f_\sigma(x) = \frac{1}{\sigma(z)} f_\varepsilon\left(\frac{x-z}{\sigma(z)}\right)/\omega(\cdot)$, *and that for any finite subsequence* $\left\{z_{k_i} : i = 1, 2, \ldots, I\right\}$, *the family of functions* $\left\{\frac{1}{\sigma\left(z_{k_i}\right)} f_\varepsilon\left(\frac{\cdot - z_{k_i}}{\sigma\left(z_{k_i}\right)}\right) : i = 1, 2, \ldots, I\right\}$ *is linearly independent.*

Then, the family $\{f(\cdot|z) : z \in \mathcal{Z}\}$ *is complete in* $L^2(\mathbb{R}, \omega)$, *where the weight function* $\omega(x)$ *satisfies* $\omega(x) = e^{-\delta' x^2}$ *for some* $\delta' \in (0, \delta_3)$.

The first part of Lemma 4 implies that one may always make a convolution sequence coincide with a complete sequence in Lemma 2 at an open neighborhood of a limit point and thereby provide more complete families. The rest of Lemma 4 is to provide sufficient conditions for the completeness under a small perturbation of the deviations defined in Theorems 1, 2, and 3. The first case immediately provides the completeness of the normal distribution with heteroskedasticity which is more flexible than the normal distribution with homoskedasticity. Suppose $\varepsilon \sim N(0,1)$ and ϕ is a standard normal PDF. Then, by Lemma 4 we have the family $\left\{f(x|z) = \frac{1}{\sigma(z)}\phi\left(\frac{x-z}{\sigma(z)}\right) : z \in \mathcal{N}(z_0)\right\}$ is complete in $L^2(\mathbb{R}, \omega)$ if $\sigma(z) = 1$ *for* $\|z - z_0\| < \varepsilon$ for some small $\varepsilon > 0$. This result is new to the literature and provides the identification for models with heteroskedasticity. Therefore, our results have shown many complete DGPs that are not previously known. Notice that a family of functions that deviate from a location family may not be in the location family anymore. To be specific, $x = z + \sigma(z)\varepsilon$ leads to $f(x|z) = \frac{1}{\sigma(z)}\phi\left(\frac{x-z}{\sigma(z)}\right)$ not in a location family. Therefore, the completeness of $f(x|z)$ does not conflict with Theorem 2.4 of Mattner (1993) after assuming that ε has a normal distribution.[14]

Another point to emphasize is that we only need the restrictions of Lemmas 2 and 3 to hold for $f_\varepsilon(\cdot) = f(\cdot|z_0)$ at an open neighborhood of 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

$$y = \alpha + \beta x + u, \quad \mathrm{E}[u|z] = 0, \tag{7}$$

with $x = z + \sigma(z)\varepsilon$ 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

$\left\{f(x|z) = \frac{1}{\sigma(z)}\phi\left(\frac{x-z}{\sigma(z)}\right) : z \in \mathcal{N}(z_0)\right\}$ is complete in $L^2(\mathbb{R},\omega)$ if $\sigma(z) = 1$ *for* $\|z - z_0\| < \varepsilon$ for some small $\varepsilon > 0$, which implies the above linear model is uniquely identified among all the functions in $L^2(\mathbb{R},\omega)$. 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 follows:[15]

$$x = h(z,\varepsilon), \text{ with } z \perp \varepsilon. \tag{8}$$

We consider the case where x and ε have the support $\mathbb{R}$. Denote CDF of ε as $F(\varepsilon)$. It is well known that the function h is related to the CDF $F_{x|z}$ as $h(z,\varepsilon) \equiv F_{x|z}^{-1}(F(\varepsilon)|z)$ 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 (8).

LEMMA 5. *Let $\mathcal{N}(z_0) \subseteq \mathcal{Z} \subset \mathbb{R}$ be an open neighborhood of some $z_0 \in \mathcal{Z}$ and equation (8) hold with $h(z_0,\varepsilon) = \varepsilon$, where the distribution function of ε, f_ε, satisfies $\int_{\mathbb{R}} |f_\varepsilon(\varepsilon)|^2 d\varepsilon < \infty$, and the conditions in Lemma 2 with a weight function ω. Suppose that*

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

ii) there exists $\{f(\cdot|z_k) \equiv \frac{\partial}{\partial x}F_\varepsilon\left(h^{-1}(z,x)\right) = \left|\frac{\partial}{\partial x}h^{-1}(z_k,\cdot)\right| f_\varepsilon\left(h^{-1}(z_k,\cdot)\right) : k = 1,2,\ldots\}$ satisfies one of the following conditions:

1) $h(z,\varepsilon) = cz + \varepsilon$, for a constant $c \neq 0$ and z satisfying $\|z - z_0\| < \varepsilon$ for some small $\varepsilon > 0$;

2) $f(\cdot|z_k)$ satisfies

$$\sum_{k=1}^{n} \|c_k(f_\varepsilon(\cdot - z_k)/\omega(\cdot) - f(\cdot|z_k)/\omega(\cdot))\| < \lambda \sum_{k=1}^{n} \|c_k f_\varepsilon(\cdot - z_k)/\omega(\cdot)\|$$

for some constant λ, $0 \le \lambda < 1$, and arbitrary scalars $c_1,\ldots,c_n$ $(n = 1,2,3,\ldots)$;

3) there exists a sequence $\{C_k : k = 1,2,\ldots\}$ which depends on the normalized sequence $\left\{\frac{f_\varepsilon(\cdot - z_k)/\omega(\cdot)}{\|f_\varepsilon(\cdot - z_k)/\omega(\cdot)\|} : k = 1,2,\ldots\right\}$ such that $\sum_{k=1}^{n} C_k\varepsilon_k < 1$ for a sequence of positive numbers $\{\varepsilon_k : k = 1,2,\ldots\}$ and

$$\left\|\frac{f_\varepsilon(\cdot - z_k)/\omega(\cdot)}{\|f_\varepsilon(\cdot - z_k)/\omega(\cdot)\|} - f(\cdot|z_k)/\omega(\cdot)\right\| < \varepsilon_k,$$

https://doi.org/10.1017/S0266466617000251 Published online by Cambridge University Press

4) $f(\cdot|z_k)$ *satisfies*

$$\sum_{k=1}^{\infty} \frac{\left\| v_k^f - v_k^{f_z} \right\|^2}{\left\| v_k^f \right\|^2} < \infty,$$

where v_k^f *and* $v_k^{f_z}$ *are defined as in Theorem 3 with* $f(\cdot) = f_\varepsilon(\cdot - z)/\omega(\cdot)$ *and* $f_z(\cdot) = f(\cdot|z)/\omega(\cdot)$, *and that for any finite subsequence* $\left\{z_{k_i} : i = 1, 2, \ldots, I\right\}$, $\left\{f(\cdot|z_{k_i}) : i = 1, 2, \ldots, I\right\}$ *is linearly independent.*

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

Proof. See the appendix. ■

Condition i) guarantees that the conditional density $f(x|z)$ is continuous in both x and z. The condition $h(z_0, \varepsilon) = \varepsilon$ is not restrictive because one may always redefine ε. Therefore, $f(x|z)$ satisfies $f(x|z_0) = f_\varepsilon(x)$. The first part of Lemma 5 implies that key sufficient assumptions for the completeness of $f(x|z)$ using the control function in equation (8) is that the control function h is locally linear in a neighborhood of a limit point in the support of z. Our results may provide sufficient conditions for completeness with a general h. For example, suppose $c \neq 0$, and small $\varepsilon > 0$, we may have

$$h(z, \varepsilon) = \begin{cases} cz + \varepsilon & \text{if } z \in (z_0 - \epsilon, z_0 + \epsilon), \\ z + e^{z - z_0}\varepsilon + \sum_{j=0}^{J} (z - z_0)^{2j} h_j(\varepsilon) & \text{else,} \end{cases}$$

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

$$h(z, \varepsilon) = \begin{cases} cz + \varepsilon & \text{if } z \in (z_0 - \epsilon, z_0 + \epsilon), \\ z + \ln\left[(z - z_0)^2 + \exp(\varepsilon)\right] & \text{else.} \end{cases}$$

3.3. Multivariate Completeness

When the endogenous variable x and the instrument z are both vectors, our main results in Theorems 1 and 3 still apply. In other words, our results can be extended to the multivariate case straightforwardly. In this section, we show that one can use Theorems 1, 2, and 3 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 $f(x_1|z_1) \times f(x_2|z_2)$. Theorems 1, 2, and 3 then imply that if conditional density $f(x_1, x_2|z_1, z_2)$ has a small deviation from $f(x_1|z_1) \times f(x_2|z_2)$ at some converging sequence in $\mathcal{Z}$ under the deviations defined in Theorems 1, 2, and 3 then $f(x_1, x_2|z_1, z_2)$ is complete. We summarize the results as follows:

LEMMA 6. *Denote $\mathcal{H} = L^2(\mathcal{X}, \omega)$ as a Hilbert space. For every $z \in \mathcal{Z} = \mathcal{Z}_1 \times \mathcal{Z}_2$, let $f_{x|z}(\cdot|z)$ be in the Hilbert space $\mathcal{H}$ of functions defined on $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2$ with norm $\|\cdot\|$. The weight function is a multiplicative product of weight functions of Hilbert spaces defined on X_1 and X_2, i.e., $\omega(x_1, x_2) = \omega(x_1)\omega(x_2)$.*[16] *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 $\{z_k : k = 1, 2, 3, \ldots\}$ of distinct $z_k \in \mathcal{N}(z_0)$ converging to z_0, the corresponding sequence $\{f_{x_1|z_1}(\cdot|z_{1k}) : k = 1, 2, 3, \ldots\}$ and $\{f_{x_2|z_2}(\cdot|z_{2k}) : k = 1, 2, 3, \ldots\}$ are complete in Hilbert spaces $\mathcal{H}$ of functions defined on $\mathcal{X}_1$ and $\mathcal{X}_2$;

ii) there exists $\{f(\cdot|z_k) : k = 1, 2, \ldots\}$ for which one of the following conditions holds:

1) *$f_{x|z}(\cdot, \cdot|z_1, z_2) = f_{x_1|z_1}(\cdot|z_1) f_{x_2|z_2}(\cdot|z_2)$ for $\|z_1 - z_{10}\| < \epsilon_1$ and $\|z_2 - z_{20}\| < \epsilon_2$ for small $\epsilon_1, \epsilon_2 > 0$;*
2) *$f_{x|z}(\cdot|z_k)$ satisfies*
$$\sum_{k=1}^{n} \left\| c_k \left(f_{x_1|z_1}(\cdot|z_{k1})/\omega(\cdot) f_{x_2|z_2}(\cdot|z_{k2})/\omega(\cdot) - f_{x|z}(\cdot,\cdot|z_{k1}, z_{k2})/\omega(\cdot,\cdot)\right)\right\| < \lambda \sum_{k=1}^{n} \left\| c_k f_{x_1|z_1}(\cdot|z_{k1})/\omega(\cdot) f_{x_2|z_2}(\cdot|z_{k2})/\omega(\cdot)\right\|$$
for some constant λ, $0 \le \lambda < 1$, and arbitrary scalars $c_1, \ldots, c_n$ $(n = 1, 2, 3, \ldots)$;
3) *there exists a sequence $\{C_k : k = 1, 2, \ldots\}$ which depends on the normalized sequence $\left\{\frac{f_{x_1|z_1}(\cdot|z_{k1})/\omega(\cdot) f_{x_2|z_2}(\cdot|z_{k2})/\omega(\cdot)}{\|f_{x_1|z_1}(\cdot|z_{k1})/\omega(\cdot) f_{x_2|z_2}(\cdot|z_{k2})/\omega(\cdot)\|} : k = 1, 2, \ldots\right\}$ such that $\sum_{k=1}^{n} C_k \varepsilon_k < 1$ for a sequence of positive numbers $\{\varepsilon_k : k = 1, 2, \ldots\}$ and*
$$\left\| \frac{f_{x_1|z_1}(\cdot|z_{k1})/\omega(\cdot) f_{x_2|z_2}(\cdot|z_{k2})/\omega(\cdot)}{\|f_{x_1|z_1}(\cdot|z_{k1})/\omega(\cdot) f_{x_2|z_2}(\cdot|z_{k2})/\omega(\cdot)\|} - f_{x|z}(\cdot,\cdot|z_{k1}, z_{k2})/\omega(\cdot,\cdot) \right\| < \varepsilon_k,$$
4) *$f_{x|z}(\cdot|z_k)$ satisfies*
$$\sum_{k=1}^{\infty} \frac{\left\| v_k^{f_{z_1 z_2}} - v_k^{f_z} \right\|^2}{\left\| v_k^{f} \right\|^2} < \infty,$$
where $v_k^{f_{z_1 z_2}}$ and $v_k^{f_z}$ are defined as in Theorem 3 with $f_{z_1 z_2}(\cdot) = f_{x_1|z_1}(\cdot|z_{k1})/\omega(\cdot) f_{x_2|z_2}(\cdot|z_{k2})/\omega(\cdot)$ and $f_z(\cdot) = f_{x|z}(\cdot|z)/\omega(\cdot,\cdot)$, and that $\left\{f(\cdot|z_{k_i}) : i = 1, 2, \ldots, I\right\}$ is linearly independent for any finite subsequence $\left\{z_{k_i} : i = 1, 2, \ldots, I\right\}$.

Then, the sequence $\{f_{x|z}(\cdot,\cdot|z_1,z_2) : z \in \mathcal{Z}\}$ *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 Theorems 1, 2, and 3 to extend the completeness for the one-dimensional sequences $\{f_{x_1|z_1}(\cdot|z_{1k}) : k = 1,2,3,\ldots\}$ and $\{f_{x_2|z_2}(\cdot|z_{2k}) : k = 1,2,3,\ldots\}$ to the multiple dimensional sequence $\{f_{x|z}(\cdot,\cdot|z_{1k},z_{2k}) : k = 1,2,3,\ldots\}$. The key assumption is that the endogenous variables are conditionally independent of each other for some value of the instruments, i.e.,

$$f_{x|z}(\cdot,\cdot|z_{10},z_{20}) = f_{x_1|z_1}(\cdot|z_{10}) f_{x_2|z_2}(\cdot|z_{20}). \tag{9}$$

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 (9) includes two steps: first, we need the independence between x_1 and x_2 only at $z = z_0$, i.e.,

$$x_1 \perp x_2 \mid z = z_0; \tag{10}$$

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 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 $\{f(x|z_k) : k = 1,2,3,\ldots\}$ as a complete sequence. At these points $z_k = (z_{1k}, z_{2k})$, an intuitive idea of Lemma 6 is the fact that the tensor product of univariate complete sequences forms a multivariate complete sequence. With the completeness of the sequence of product function $\{f_{x_1|z_1}(\cdot|z_{1k}) f_{x_2|z_2}(\cdot|z_{2k}) : k = 1,2,3,\ldots\}$, we can utilize the main perturbation results, Theorems 1, 2, and 3, to extend the result to other sequences of functions close to the sequence of the product function. At these "small" perturbation sequences, $\{f_{x|z}(\cdot,\cdot|z_{1k},z_{2k}) : k = 1,2,3,\ldots\}$ can be nonseparable and satisfies the condition (10). 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 $\{f_{x_1|z_1}(\cdot|z_{1k}) : k = 1,2,3,\ldots\}$ and

https://doi.org/10.1017/S0266466617000251 Published online by Cambridge University Press

$\{f_{x_2|z_2}(\cdot|z_{2k}) : k = 1,2,3,\ldots\}$, where z_{1k} and z_{2k} are distinct sequences converging to 1.

The family of the product function $\{f_{x_1|z_1}(\cdot|z_{1k}) f_{x_2|z_2}(\cdot|z_{2k}) : k = 1,2,3,\ldots\}$, where $f_{x_1|z_1}(x_1|z_1) = \frac{1}{z_1}e^{-x_1z_1}$ and $f_{x_2|z_2}(x_2|z_2) = \frac{1}{z_2}e^{-x_2z_2}$, is complete in $\{0\}\cup\mathbb{R}^{+2}$ because the family contains a subfamily as a basis in $\{0\}\cup\mathbb{R}^{+2}$. Then, by Theorems 1, 2, and 3 the family of multivariate density $\{f_{x|z}(\cdot,\cdot|z_{1k}, z_{2k}) : k = 1,2,3,\ldots\}$ may be complete when the family is sufficiently close to the family of product functions under the deviations defined in Theorems 1, 2, and 3. On the other hand, we can use condition ii) 1) to provide more complete families. For small $\epsilon_1, \epsilon_2 > 0$, set $\mathcal{O}_z = (z_{10} - \epsilon_1, z_{10} + \epsilon_1) \times (z_{20} - \epsilon_2, z_{10} + \epsilon_1)$. Consider the multivariate density

$$f_{x|z}(\cdot,\cdot|z_1,z_2) = \begin{cases} f_{x_1|z_1}(\cdot|z_1) f_{x_2|z_2}(\cdot|z_2) & \text{if } (z_1,z_2) \in \mathcal{O}_z, \\ \frac{c_{zk}}{z_{1k}z_{2k}} e^{-(x_1z_{1k}+x_2z_{2k}+(z_{1k}-1)^2(z_{2k}-1)^2x_1x_2)} & \text{else,} \end{cases}$$

where c_{zk} is a normalized coefficient, $z_1, z_2 > 0$, and $x_1, x_2 \in \{0\}\cup\mathbb{R}^+$. The family has an exponential decay tail over $\mathbb{R}^{+2}$ and is therefore integrable. The family at $\mathcal{O}_z$ is the same as the family of product function $\{f_{x_1|z_1}(\cdot|z_{1k}) f_{x_2|z_2}(\cdot|z_{2k}) : k = 1,2,3,\ldots\}$, where $f_{x_1|z_1}(x_1|z_1) = \frac{1}{z_1}e^{-x_1z_1}$ and $f_{x_2|z_2}(x_2|z_2) = \frac{1}{z_2}e^{-x_2z_2}$. Lemma 6 implies the sequence $\{f_{x|z}(\cdot,\cdot|z_1,z_2) : z \in \mathcal{Z}\}$ is complete.

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 complete sequences in Hilbert spaces to show that (1) if the deviation of the conditional density $f(x|z_k)$ from a known complete sequence of functions is less than a sequence of values determined by the complete sequence in some distinct sequence $\{z_k : k = 1,2,3,\ldots\}$ converging to z_0, then $f(x|z)$ itself is complete, and (2) if the conditional density $f(x|z)$ can form a linearly independent sequence $\{f(\cdot|z_k) : k = 1,2,\ldots\}$ for some distinct sequence $\{z_k : k = 1,2,3,\ldots\}$ converging to z_0 and its relative deviation from a known complete sequence of functions under some norm is finite then $f(x|z)$ itself is complete. Therefore, the regression function is nonparametrically identified.

NOTES

1. In this paper, we need to consider $E[Y|Z = z]$, i.e., conditional expectation of Y on the random variable Z taking value z in its support $\mathcal{Z}$. In other words, we need this conditional expectation to be well-defined even for the zero-probability event $\{Z = z\}$. To avoid any confusion, we assume the conditional expectation $E[Y|Z = z]$ is a continuous function of z over the support $\mathcal{Z}$. We are not the first

study to use the conditional expectation in this way for completeness. For example, Newey and Powell (2003) show that for exponential families and normal distribution, the conditional expectations in the definition of completeness can be reduced to be defined over an open subset of $\mathcal{Z}$. We follow their definition of completeness to require that conditional expectations are defined for some set $\mathcal{O} \subseteq \mathcal{Z}$ and this paper shows that $\mathcal{O}$ can be a distinct converging sequence in $\mathcal{Z}$.

2. We consider the quotient space $\in L^2(\mathbb{R}, \omega)$ where the equivalent relation $\sim$ is that $f \sim g$ if the set $\{x : f(x) \neq g(x)\}$ is a set of measure zero. If the set of elements for which a property does not hold is a set of measure zero for a probability measure, we use almost surely to indicate the property.

3. The integral in the formula makes sense because $\int_{\mathcal{X}} |h(x) f(x|z)| dx = \int_{\mathcal{X}} |h(x)| \omega(x)^{1/2} \frac{f(x|z)}{\omega(x)^{1/2}} dx \leq \left(\int_{\mathcal{X}} |h(x)|^2 \omega(x) dx\right)^{1/2} \left(\int_{\mathcal{X}} \frac{f(x|z)^2}{\omega(x)} dx\right)^{1/2} < \infty.$

4. 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 have a support like $\mathcal{X} \times \mathcal{Z}$.

5. 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).

6. 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 functions is $h(x)$, i.e., $h(x) = m(x) - \widetilde{m}(x)$, where m and $\tilde{m}$ are regression functions such that $y = m(x) + u$ or $y = \tilde{m}(x) + u$.

7. When $x, z \in \{0,1\}$, the conditional expectation $E[y|z] = \int_{\mathcal{X}} m(x) f(x|z) dx$ is equivalent to $\begin{bmatrix} E[y|z=0] \\ E[y|z=1] \end{bmatrix}^T = \begin{bmatrix} m(0) \\ m(1) \end{bmatrix}^T \begin{bmatrix} f_{x|z}(0|0) & f_{x|z}(0|1) \\ f_{x|z}(1|0) & f_{x|z}(1|1) \end{bmatrix}$. In this binary case, the regression $m(\cdot)$ may be uniquely determined from observed $E[y|z]$, and $f(x|z)$ if the last matrix is invertible, i.e., two vectors $f_{x|z}(\cdot|0)$ and $f_{x|z}(\cdot|1)$ are linearly independent

$$f_{x|z}(\cdot|0) := \begin{bmatrix} f_{x|z}(0|0) \\ f_{x|z}(1|0) \end{bmatrix} \text{ and } f_{x|z}(\cdot|1) := \begin{bmatrix} f_{x|z}(0|1) \\ f_{x|z}(1|1) \end{bmatrix}.$$

Therefore, completeness is equivalent to no-perfect collinearity among $\{f_{x|z}(\cdot|i) : i = 1,2\}$ or the rank condition on the matrix $\begin{bmatrix} f_{x|z}(0|0) & f_{x|z}(0|1) \\ f_{x|z}(1|0) & f_{x|z}(1|1) \end{bmatrix}$.

8. Theorems 2.2 and 2.3 in Newey and Powell (2003) do not specify the function space in which completeness is discussed. The definition of the completeness on page 141 of Lehmann (1986) also does not specify the function space. However, he starts to specify the property of completeness for all bounded functions and call it boundedly complete on page 144.

9. In this paper, we would use a low subscript such as z_0 to denote a point.

10. The term used here accords with the definition on page 182 of Rudin (1987), where the translate of f is defined as $f(x-z)$ for all x and a given z.

11. 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.

12. Given two open connected sets X and Y, a map f from X to Y is called a C^1-diffeomorphism if f is a bijection and both $f : X \to Y$ and its inverse $f^{-1} : Y \to X$ are continuously differentiable.

13. The family $\{p(\cdot - z) : z \in \mathcal{Z}\}$ is L^1 complete if, for all measurable real functions h such that $\mathrm{E}[|h(X)|] < \infty$,

$$\int_{\mathcal{X}} h(x) p(x-z) dx = 0 \quad \text{for all } z \in \mathcal{Z}$$

implies $h(\cdot) = 0$ almost surely in $\mathcal{X}$. We thank a referee for pointing out the result.

14. In general, we can apply the stability results to the sequence in Lemma 2. Let us take a complete sequence in Lemma 2, i.e., $\{f_\varepsilon(\cdot - z_1), f_\varepsilon(\cdot - z_2), f_\varepsilon(\cdot - z_3), f_\varepsilon(\cdot - z_4) \ldots\}$. In order to

https://doi.org/10.1017/S0266466617000251 Published online by Cambridge University Press

generate a new complete sequence, we only change the first function $f_\varepsilon(\cdot - z_1)$ with a very small deviation to $\widetilde{f}(\cdot - z_1)$. The stability result in Theorem 2 implies the new sequence $\{\widetilde{f}(\cdot - z_1), f_\varepsilon(\cdot - z_2), f_\varepsilon(\cdot - z_3), f_\varepsilon(\cdot - z_4)\ldots\}$ is also a complete sequence. (Note that this argument also applies to a sequence of basis functions.) The important observation is that the completeness of the new sequence does not conflict with Mattner's results because the new sequence is no longer in a location family.

15. 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.

16. A simple example of this type of weight function is $\omega(x_1, x_2) = e^{-(a_1 x_1^2 + a_2 x_2^2)} = e^{-a_1 x_1^2} e^{-a_2 x_2^2} = \omega(x_1)\omega(x_2)$, where $a_1, a_2 > 0$.

REFERENCES

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

An, Y. & Y. Hu (2012) Well-posedness of measurement error models for self-reported data. *Journal of Econometrics* 168(2), 259–269.

Andrews, D. (2012) Examples of L^2-complete and boundedly-complete distributions. *Cowles Foundation for Research in Economics*.

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

Canay, I., A. Santos, & A. Shaikh (2013) On the testability of identification in some nonparametric models with endogeneity. *Econometrica* 81(6), 2535–2559.

Carroll, R., X. Chen, & 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. & Y. Hu (2006) Identification and Inference of Nonlinear Models Using Two Samples with Arbitrary Measurement Errors. Cowles Foundation Discussion paper no. 1590.

Chernozhukov, V. & C. Hansen (2005) An IV model of quantile treatment effects. *Econometrica* 73(1), 245–261.

Chernozhukov, V., G. Imbens, & W. Newey (2007) Instrumental variable estimation of nonseparable models. *Journal of Econometrics* 139(1), 4–14.

Darolles, S., Y. Fan, J. Florens, & 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* 27(3), 1–12.

Dostanić, M. (1990) On the completeness of the system of functions $\{e^{-\alpha\lambda_n x} \sin \lambda_n x\}_{n=1}^{\infty}$. *Journal of Mathematical Analysis and Applications* 150(2), 519–527.

Erdös, P. & E. Straus (1953) On linear independence of sequences in a banach space. *Pacific Journal of Mathematics* 3(4), 689–694.

Florens, J., M. Mouchart, & J. Rolin (1990) *Elements of Bayesian Statistics*. Marcel Dekker.

Gurarij, V. & M. Meletidi (1970) The stability of the completeness of sequences in banach spaces. *Bulletin de l'Academie Polonaise des Sciences, Serie des Sciences, Mathematiques, Astronomiques et Physiques* 18(9), 533–536.

Hall, P. & 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, & A. Simoni (2012) Semiparametric Estimation of Random Coefficients in Structural Economic Models. Cemmap Working papers.

Horowitz, J. (2011) Applied nonparametric instrumental variables estimation. *Econometrica* 79(2), 347–394.

Horowitz, J. & S. Lee (2007) Nonparametric instrumental variables estimation of a quantile regression model. *Econometrica* 75(4), 1191–1208.

Hu, Y. & S. Schennach (2008) Instrumental variable treatment of nonclassical measurement error models. *Econometrica* 76(1), 195–216.
Hu, Y. & M. Shum (2012) Nonparametric identification of dynamic models with unobserved state variables. *Journal of Econometrics* 171(1), 32–44.
Lehmann, E. (1986) *Testing Statistical Hypotheses*, 2nd ed. Wiley.
Mattner, L. (1993) Some incomplete but boundedly complete location families. *The Annals of Statistics* 21(4), 2158–2162.
Mattner, L. (1996) Complete incomplete but boundedly complete location families. *The Annals of Statistics* 24(3), 1265–1282.
Newey, W. & J. Powell (2003) Instrumental variable estimation of nonparametric models. *Econometrica* 71(5), 1565–1578.
Phillips, P.C.B. & L. Su (2011) Nonparametric regression under location shifts. *Econometrics Journal* 14(3), 457–486.
Rudin, W. (1987) *Real and Complex Analysis*. McGraw-Hill.
San Martin, E. & M. Mouchart (2007) On joint completeness: Sampling and bayesian versions, and their connections. *The Indian Journal of Statistics* 69(4), 780–807.
Shiu, J. & Y. Hu (2013) Identification and estimation of nonlinear dynamic panel data models with unobserved covariates. *Journal of Econometrics* 175(2), 116–131.
Wang, Q. & P.C.B. Phillips (2007) Asymptotic theory for local time density estimation and nonparametric cointegrating regression. *Econometric Theory* 25(3), 710–738.
Wang, Q. & P.C.B. Phillips (2009) Structural nonparametric cointegrating regression. *Econometrica* 77(6), 1901–1948.
Wang, Q. & P.C.B. Phillips (2016) Nonparametric cointegrating regression with endogeneity and long memory. *Econometric Theory* 32(2), 359–401.
Young, R. (2001) *An Introduction to Nonharmonic Fourier Series*. Academic Press.
Yu, P. & P.C.B. Phillips (2017) Threshold regression with endogeneity. *Journal of Econometrics*, forthcoming.
Zimmer, R. (1990) *Essential Results of Functional Analysis*. University of Chicago Press.

APPENDIX A: Proofs

A.1. Preliminaries

This paper considers a weighted L^2 space $L^2(\mathcal{X},\omega) = \{h(\cdot) : \int_{\mathcal{X}} |h(x)|^2 \omega(x)dx < \infty\}$ with the inner product $\langle f, g\rangle \equiv \int_{\mathcal{X}} f(x)g(x)\omega(x)dx$. We define the corresponding norm as: $\|f\|^2 = \langle f, f\rangle$. The completion of $L^2(\mathcal{X},\omega)$ 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 $\mathcal{X}\times\mathcal{Z}$. Let ω be a weight function. If z only takes values from a countable set in $\mathcal{Z}$ then the conditional density $f(x|z)$ can be used to extend as a sequence of functions $\{f_1, f_2, f_3, \ldots\}$ in $\mathcal{H}$ with

$$f_k(\cdot) \equiv \frac{f(\cdot|z_k)}{\omega(\cdot)},$$

where $\{z_k : k = 1,2,3,\ldots\}$ 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 $\mathcal{H}$.

DEFINITION A.1. *A sequence of functions $\{f_1, f_2, f_3,\ldots\}$ in $\mathcal{H}$ is said to be a basis if for any $h \in \mathcal{H}$ there corresponds a unique sequence of scalars $\{c_1, c_2, c_3,\ldots\}$ such that*

https://doi.org/10.1017/S0266466617000251 Published online by Cambridge University Press

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

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 A.2. *A sequence of functions* $\{f_n(\cdot)\}$ *of* $\mathcal{H}$ *is said to be* $\omega-$*independent if the equality*

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

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

It is obvious that the $\omega-$independence implies 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.

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

In fact, one can show that a basis is complete. Since every element in $\mathcal{H}$ has a unique representation in terms of a basis, there is redundancy in a complete sequence. Given a complete sequence in $\mathcal{H}$, we can construct a basis from the complete sequence. One of the important properties of a complete sequence for $\mathcal{H}$ is that every element can be approximated arbitrarily close by finite combinations of the elements. We summarize these results as follows.

LEMMA A.1. (*1*) *A basis in* $\mathcal{H}$ *is also a complete sequence.*

(*2*) *Let* W *be a closed linear subspace of* $\mathcal{H}$. *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* $\{f_1, f_2, f_3, \ldots\}$ *in* $\mathcal{H}$, *we can construct an orthonormal basis* $\{g_1, g_2, g_3, \ldots\}$ *from the complete sequence for* $\mathcal{H}$.

Proof of Lemma A.1(1). Given a basis $\{f_1, f_2, f_3, \ldots\}$ in $\mathcal{H}$, applying the Gram-Schmidt process to the basis yields an orthonormal sequence $\{g_1, g_2, g_3, \ldots\}$ and $\text{span}(\{f_1, f_2, f_3, \ldots\}) = \text{span}(\{g_1, g_2, g_3, \ldots\})$. This implies that $\{g_1, g_2, g_3, \ldots\}$ is also a basis of the Hilbert space $\mathcal{H}$ and $f = \sum_{k=1}^{\infty} \langle f, g_k \rangle g_k$ for any $f \in \mathcal{H}$. Suppose that $\int f_k(x) h(x) \omega(x) dx = 0$ for all k. It follows that $\langle h, g_k \rangle = 0$ for all k. Thus, $h = \sum_{k=1}^{\infty} \langle h, g_k \rangle g_k = 0$. $\{f_1, f_2, f_3, \ldots\}$ is a complete sequence. ■

The proof of Lemma A.1(2) can be found as a corollary on page 7 in Zimmer (1990).

Proof of Lemma A.1(3). We will construct g_k using the Gram–Schmidt procedure. First, let $r_1 = f_1$ and $g_1 = \frac{r_1}{\|r_1\|}$. Then $r_2 = f_{s_2}$ where s_2 is the smallest index among

https://doi.org/10.1017/S0266466617000251 Published online by Cambridge University Press

$\{2,3,4,\ldots\}$ such that $\tilde{g}_2 \equiv f_{s_2} - \langle f_{s_2}, g_1\rangle g_1 \neq 0$. Denote $g_2 = \frac{\tilde{g}_2}{\|\tilde{g}_2\|}$. Keep the selection process going, in the k-th step, we have $r_k = f_{s_k}$ where s_k is the smallest index among $\{s_{k-1}+1, s_{k-1}+2, s_{k-1}+3,\ldots\}$ such that $\tilde{g}_k \equiv f_{s_k} - \sum_{i=1}^{k-1} \langle f_{s_k}, g_i\rangle g_i \neq 0$ and $g_k = \frac{\tilde{g}_k}{\|\tilde{g}_k\|}$. This selection procedure produces two sequences with the same span space, i.e., $\text{span}(\{f_1, f_2, f_3,\ldots\}) = \text{span}(\{g_1, g_2, g_3,\ldots\})$. In addition, $\{g_1, g_2, g_3,\ldots\}$ is an orthonormal sequence. To prove $\{g_1, g_2, g_3,\ldots\}$ is a basis, it is sufficient to show (i) the completion of $\text{span}(\{g_1, g_2, g_3,\ldots\}) = \mathcal{H}$, and (ii) $\{g_1, g_2, g_3,\ldots\}$ is $\omega-$independent. Let W be the completion of the subspace $span(\{g_1, g_2, g_3,\ldots\})$ under the norm $\|\cdot\|$. Let $W^\perp = \{h \in \mathcal{H} : \langle h, g\rangle = 0 \text{ for all } g \in W\}$. By Lemma A.1 (ii), $W \bigoplus W^\perp = \mathcal{H}$. Since the sequence $\{f_1, f_2, f_3,\ldots\}$ is complete and $\text{span}(\{f_1, f_2, f_3,\ldots\}) = \text{span}(\{g_1, g_2, g_3,\ldots\})$ then $W^\perp = \{0\}$ and $W = \mathcal{H}$. On the other hand, suppose that $\sum_{k=1}^{\infty} c_k g_k = 0$ for some scalars $c_1, c_2, c_3,\ldots$. Because $\{g_1, g_2, g_3,\ldots\}$ is an orthonormal sequence, $0 = \left\langle \sum_{k=1}^{\infty} c_k g_k, g_i \right\rangle = c_i$ for $i = 1,2,3,\ldots$. This implies that $\{g_1, g_2, g_3,\ldots\}$ is $\omega-$independent. Therefore, the sequence $\{g_1, g_2, g_3,\ldots\}$ is a basis. ■

Our proofs also need a uniqueness theorem of complex differentiable functions stated in the corollary on page 209 in Rudin (1987).

A.2. Proofs of Completeness of Existing Sequences

Proof of Lemma 1. Set $t(z) = 1$ for simplicity. In order to use the above uniqueness result of complex differentiable functions, we consider a converging sequence $\{z_k : k = 1,2,\ldots\}$ 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 subsequence $\{z_{k_i} : i = 1,2,\ldots\}$ converging to z_0 such that $\left\{\mu(z_{k_i}) : i = 1,2,\ldots\right\} \in (\mu(z_0) - \delta, \mu(z_0) + \delta) \subset \mu(\mathcal{N}(z_0))$ be a sequence of distinct numbers converging to an interior point $\mu(z_0) \in \mu(\mathcal{N}(z_0))$ and $\mu(z_{k_i})\tau(x) < \mu(z_0)\tau(x) + \delta|\tau(x)|$ for $i = 1,2,\ldots$. In addition, since $g(\cdot|z) \in L^1(\mathcal{X})$ 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 $\int_{\mathcal{X}} \frac{s(x)^2 \exp[2(\mu(z_0)\tau(x)+\delta|\tau(x)|)]}{\omega(x)} dx < \infty$ and it follows that $\int_{\mathcal{X}} g(x|z)^2/\omega(x) dx < \infty$ for $z \in \mathcal{O}$. Given $h_0 \in L^2(\mathcal{X}, \omega)$ and pick a positive constant δ_1 such that $0 < \delta_1 < \delta$. Let $w = a + ib$, where a, b are real numbers. Then, check the integrability of the function $s(x)e^{w\tau(x)}h_0(x)$ over $x \in \mathcal{X}$ for $a \in (\mu(z_0) - \delta_1, \mu(z_0) + \delta_1)$. Use Cauchy–Schwartz inequality to the function,

$$\begin{aligned}
&\left|\int_{\mathcal{X}} s(x)e^{w\tau(x)}h_0(x)dx\right|^2 \qquad \textbf{(A.1)}\\
&\leq \left(\int_{\mathcal{X}} \frac{s(x)e^{a\tau(x)}}{\omega(x)^{1/2}} |h_0(x)|\omega(x)^{1/2} dx\right)^2\\
&\leq \left(\int_{\mathcal{X}} \frac{s(x)e^{\mu(z_0)\tau(x)+\delta_1|\tau(x)|}}{\omega(x)^{1/2}} |h_0(x)|\omega(x)^{1/2} dx\right)^2
\end{aligned}$$

https://doi.org/10.1017/S0266466617000251 Published online by Cambridge University Press

$$\leq \left(\int_{\mathcal{X}} \frac{s(x)e^{\mu(z_0)\tau(x)+\delta|\tau(x)|}}{\omega(x)^{1/2}} |h_0(x)|\omega(x)^{1/2} dx \right)^2$$
$$\leq \left(\int_{\mathcal{X}} \frac{s(x)^2 \exp\left[2(\mu(z_0)\tau(x)+\delta|\tau(x)|)\right]}{\omega(x)} dx \right) \left(\int_{\mathcal{X}} |h_0(x)|^2 \omega(x) dx \right) < \infty.$$

This implies that a complex function defined as an integral of the function exists and is finite. Consider the complex function with the following form

$$f(w) = \int_{\mathcal{X}} s(x)e^{w\tau(x)} h_0(x) dx, \qquad \textbf{(A.2)}$$

where the complex variable w is in the vertical strip $R \equiv \{w : \mu(z_0) - \delta_1 < \mathrm{Re}(w) < \mu(z_0) + \delta_1\}$. (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).) 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 equation (A.2), we have

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

where we have used (1) apply the inequality $\left|\frac{e^{az}-1}{z}\right| \leq \frac{e^{\delta_3|a|}}{\delta_3}$ for $|z| \leq \delta_3$ to the factor $\frac{(e^{\eta\tau(x)}-1)}{\eta}$, (The inequality can be found on page 60 of Lehmann (1986).) and (2) $w \in R$. The right-hand side is integrable when $\delta_1 + \delta_2 < \delta$ by a similar derivation in equation (A.1). It follows from the Lebesgue dominated convergence theorem that

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

Therefore, $f'(w)$ exists and the function f defined through the integral is holomorphic.

The condition $\int_{\mathcal{X}} s(x)e^{\mu(z_{k_i})\tau(x)} h_0(x)dx = 0$ is equivalent to $f(\mu(z_{k_i})) = 0$ by equation (A.2). This implies that the complex differentiable function f is equal to zeros in the sequence $\{\mu(z_{k_1}), \mu(z_{k_2}), \mu(z_{k_3}), \ldots\}$ which has a limit point $\mu(z_0)$. Applying the uniqueness theorem 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 $L^2(\mathbb{R}, \omega)$ by

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

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 $\tilde{\tau}'(x) \neq 0$ for every x. In particular, set $w = \mu(\tilde{z}) + it$ for any real t and some $\tilde{z} \in \mathcal{O}$ such that $\mu(\tilde{z}) \in (\mu(z_0) - \delta_1, \mu(z_0) + \delta_1)$, 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. In particular, equation (A.1) implies $\hat{h}_0 \in L^1(\mathbb{R})$. (Recall that $\mu(\tilde{z}) \in (\mu(z_0) - \delta_1, \mu(z_0) + \delta_1)$, and $\hat{h}_0(x) \equiv \tilde{s}(\tau^{-1}(x)) e^{\mu(\tilde{z})x} \tilde{h}_0(\tilde{\tau}^{-1}(x)) \frac{1}{\tilde{\tau}'(x)}$. Consider

$$
\begin{aligned}
\int_{-\infty}^{\infty} |\hat{h}_0(x)| dx &\leq \int_{-\infty}^{\infty} \left| \tilde{s}(\tau^{-1}(x)) e^{\mu(\tilde{z})x} \tilde{h}_0(\tilde{\tau}^{-1}(x)) \frac{1}{\tilde{\tau}'(x)} \right| dx \\
&\leq \int_{\mathcal{X}} s(x) e^{\mu(\tilde{z})\tau(x)} |h_0(x)| dx \\
&\leq \int_{\mathcal{X}} \frac{s(x) e^{\mu(z_0)\tau(x) + \delta_1 |\tau(x)|}}{\omega(x)^{1/2}} |h_0(x)| \omega(x)^{1/2} dx \\
&\leq \int_{\mathcal{X}} \frac{s(x) e^{\mu(z_0)\tau(x) + \delta |\tau(x)|}}{\omega(x)^{1/2}} |h_0(x)| \omega(x)^{1/2} dx \\
&\leq \left(\int_{\mathcal{X}} \frac{s(x)^2 \exp\left[2(\mu(z_0)\tau(x) + \delta|\tau(x)|)\right]}{\omega(x)} dx \right)^{1/2} \left(\int_{\mathcal{X}} |h_0(x)|^2 \omega(x) dx \right)^{1/2} < \infty.
\end{aligned}
$$

This shows that $\hat{h}_0 \in L^1(\mathbb{R})$.) By the uniqueness theorem, for $\hat{h}_0 \in L^1(\mathbb{R})$ we have $\hat{h}_0 = 0$ and therefore the function $h_0 = 0$. This shows that the sequence $\{g(\cdot|z_k) = s(\cdot) t(z_k) e^{\mu(z_k)\tau(\cdot)} : k = 1, 2, \ldots\}$ is complete in $L^2(\mathcal{X}, \omega)$. ■

Proof of Lemma 2. We start with the following two inequalities. Set some positive constants c_i for $i = 1, 2, 3$, and $\delta_i > 0$ for $i = 1, 2, 3, 4, 5$. Suppose that

$$
\left| f_\varepsilon(x - z) e^{-\delta_1 z^2} \right| < c_1 e^{-\delta_2 (x - c_2 z)^2} e^{-\delta_3 x^2} \tag{A.3}
$$

and

$$
0 < \left| \int_{-\infty}^{\infty} e^{itz} f_\varepsilon(x - z) e^{-\delta_1 z^2} dz \right| < c_3 e^{-\delta_4 t^2} e^{-\delta_3 x^2} \tag{A.4}
$$

for all $t \in \mathbb{R}$. We will show Gaussian distributions satisfy equations (A.3) and (A.4) and then use these two equations to show the property of completeness. Assume $f_\varepsilon(\varepsilon) = c_\varepsilon e^{-\frac{\varepsilon^2}{2\sigma^2}}$ for some $\sigma^2 < 1$, and write $e^{-\frac{(x-z)^2}{2\sigma^2}} = e^{-\left(\frac{1}{\sigma^2} - 1\right)\frac{x^2}{2}} e^{-\frac{\left(x - \frac{z}{\sigma^2}\right)^2}{2}} e^{\frac{(1-\sigma^2)z^2}{2\sigma^4}}$. Set $\delta_1 = \frac{(1-\sigma^2)}{2\sigma^4}$. It follows that

$$
f_\varepsilon(x - z) e^{-\delta_1 z^2} = c_\varepsilon e^{\frac{-(x-z)^2}{2\sigma^2}} e^{-\delta_1 z^2} = c_\varepsilon e^{-\left(\frac{1}{\sigma^2} - 1\right)\frac{x^2}{2}} e^{-\frac{\left(x - \frac{z}{\sigma^2}\right)^2}{2}}.
$$

Let $\tilde{z} = x - \frac{z}{\sigma^2}$ and $\delta_p = \frac{1}{4}\left(\frac{1}{\sigma^2} - 1\right)$. Then

$$f_\varepsilon(x-z)e^{-\delta_1 z^2} = c_\varepsilon e^{-2\delta_p x^2} e^{-\tilde{z}^2/2},$$

and this implies that $|f_\varepsilon(x-z)e^{-\delta_1 z^2}| \le c_\varepsilon e^{-\delta_p x^2} e^{-\tilde{z}^2/4}$. Hence, the condition (A.3) is satisfied with $\delta_1 = \frac{(1-\sigma^2)}{2\sigma^4}$, $c_1 = c_\varepsilon$, $\delta_2 = \frac{1}{4}$, $c_2 = \frac{1}{\sigma^2}$, and $\delta_3 = \delta_p$.

As for equation (A.4), consider the corresponding Fourier transform in the condition (A.4):

$$\begin{aligned}
&\left|\int_{-\infty}^{\infty} e^{itz} f_\varepsilon(x-z)e^{-\delta_1 z^2} dz\right| \\
&\quad = \left|\int_{-\infty}^{\infty} e^{it(\sigma^2 x - \sigma^2\tilde{z})} c_\varepsilon e^{-2\delta_p x^2} e^{-\tilde{z}^2/2}(\sigma^2) d\tilde{z}\right| \\
&\quad = c_\varepsilon \sigma^2 \left| e^{i\sigma^2 tx} \int_{-\infty}^{\infty} e^{i(-\sigma^2 t)\tilde{z}} e^{-2\delta_p x^2} \cdot e^{-\tilde{z}^2/2} d\tilde{z}\right| \\
&\quad = c_\varepsilon \sigma^2 \cdot e^{-2\delta_p x^2} \left|\int_{-\infty}^{\infty} e^{i(-\sigma^2 t)\tilde{z}} \cdot e^{-\tilde{z}^2/2} d\tilde{z}\right| \\
&\quad = c_\varepsilon \sigma^2 e^{-2\delta_p x^2} \left| F\{e^{-\tilde{z}^2/2}\}(-\sigma^2 t)\right| \\
&\quad \le c_\varepsilon \sigma^2 e^{-2\delta_p x^2} e^{-\frac{\sigma^4 t^2}{2}} \\
&\quad \le c_\varepsilon \sigma^2 e^{-\delta_p x^2} e^{-\frac{\sigma^4 t^2}{4}},
\end{aligned}$$

where we have used the fact that $e^{-\tilde{z}^2/2}$ is an eigenfunction of the Fourier transform, $F\{e^{-\tilde{z}^2/2}\}(t) = \sqrt{2\pi}\, e^{-t^2/2}$. This implies that the Gaussian distributions satisfy the condition (A.4).

Next, choose a sequence of distinct numbers $\{z_k\}$ in the support $\mathcal{Z}$ converging to $z_0 \in \mathcal{Z}$. The inequality

$$\frac{g(x|z_k)^2}{\omega(x)} = \frac{f_\varepsilon(x-z_k)^2}{\omega(x)} < c_1^2 e^{-2\delta_2(x-c_2 z_k)^2} e^{-(2\delta_3-\delta')x^2} e^{2\delta_1 z_k^2}$$

with $\delta_2 > 0$ and $2\delta_3 > \delta'$ implies that $\int_{\mathbb{R}} \frac{g(x|z_k)^2}{\omega(x)} dx < \infty$ for all k. Suppose that for some $h_0 \in L^2(\mathbb{R}, \omega)$, $\int_{-\infty}^{\infty} h_0(x) f_\varepsilon(x-z_k)\, dx = 0$. Divide it by $e^{\delta_1 z_k^2}$ and rewrite the equation as

$$\int_{-\infty}^{\infty} h_0(x)\omega(x) \frac{f_\varepsilon(x-z_k)\, e^{-\delta_1 z_k^2}}{\omega(x)} dx = 0 \text{ for all } k. \tag{A.5}$$

Consider

$$g(z) \equiv \int_{-\infty}^{\infty} h_0(x)\omega(x) \frac{f_\varepsilon(x-z)\, e^{-\delta_1 z^2}}{\omega(x)} dx, \tag{A.6}$$

which is similar to a convolution and $g(z_k) = 0$ for all k. Because $h_0 \in L^2(\mathbb{R}, \omega)$, $h_0\omega \in L^1(\mathbb{R})$. (Suppose $h_0 \in L^2(\mathbb{R}, \omega)$. Apply Cauchy-Schwartz inequality, $\int_{\mathbb{R}} |h_0(x)\omega(x)| dx = \int_{\mathbb{R}} |h_0(x)|\omega(x)^{1/2}\omega(x)^{1/2} dx \le \left(\int_{\mathbb{R}} h_0(x)^2 \omega(x) dx\right)^{1/2} \left(\int_{\mathbb{R}} \omega(x) dx\right)^{1/2} < \infty$.

This implies $h_0\omega \in L^1(\mathbb{R})$.) Then, the condition $\delta' < \delta_3$ implies the function g is integrable because

$$\begin{aligned}
\int_{-\infty}^{\infty} |g(z)| dz &\leq \int_{-\infty}^{\infty} \int_{\mathbb{R}} |h_0(x)|\omega(x) \frac{f_\varepsilon(x-z) e^{-\delta_1 z^2}}{\omega(x)} dx dz \\
&= \int_{\mathbb{R}} |h_0(x)|\omega(x) \left(\int_{-\infty}^{\infty} \frac{f_\varepsilon(x-z) e^{-\delta_1 z^2}}{\omega(x)} dz \right) dx \\
&\leq \int_{\mathbb{R}} |h_0(x)|\omega(x) c_1 e^{-(\delta_3-\delta')x^2} \left(\int_{-\infty}^{\infty} e^{-\delta_2 (x-c_2 z)^2} dz \right) dx \\
&\leq c \int_{\mathbb{R}} |h_0(x)|\omega(x) dx \left(\int_{-\infty}^{\infty} e^{-\delta_2 z^2} dz \right) < \infty, \text{ for some } c.
\end{aligned}$$

Let $\phi_g(t) = \int_{-\infty}^{\infty} e^{itz} g(z) dz$ be the Fourier transform of g. We can derive a bound for $\phi_g(t)$ as follows:

$$\begin{aligned}
|\phi_g(t)| &= \left| \int_{-\infty}^{\infty} e^{itz} g(z) dz \right| \\
&= \left| \int_{-\infty}^{\infty} e^{itz} \int_{\mathbb{R}} h_0(x)\omega(x) \frac{f_\varepsilon(x-z) e^{-\delta_1 z^2}}{\omega(x)} dx dz \right| \\
&\leq \int_{\mathbb{R}} |h_0(x)|\, \omega(x) \left| \int_{-\infty}^{\infty} e^{itz} \frac{f_\varepsilon(x-z) e^{-\delta_1 z^2}}{\omega(x)} dz \right| dx \\
&\leq c_3 \left(\int_{\mathbb{R}} |h_0(x)|\omega(x) e^{-(\delta_3-\delta')x^2} dx \right) e^{-\delta_4 t^2} \\
&\leq c_3 \left(\int_{\mathbb{R}} |h_0(x)|\omega(x) dx \right) e^{-\delta_4 t^2} \\
&\leq c_4 e^{-\delta_4 |t|},
\end{aligned} \tag{A.7}$$

where we have used (i) an interchange of the order of integration (justified by applying Fubini's theorem to the integrable g), (ii) the inequality (A.4), and (iii) $\delta' < \delta_3$. Since $h_0\omega$ is integrable, $\phi_g(t)$ is also integrable. Both g and $\phi_g(t)$ are integrable, applying the inversion theorem to g yields that $g(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itz} \phi_g(t) dt$. Extend the function g from $\mathbb{R}$ to $\mathbb{C}$ and define

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

with

$$w = z + ib \text{ for } z, b \in \mathbb{R} \text{ with } |b| < r < \delta_4.$$

The function $f(w)$ is bounded by using equation (A.7) through

$$|f(w)| = \left| \int_{-\infty}^{\infty} e^{-itw} \phi_g(t) dt \right| \leq \int_{-\infty}^{\infty} e^{|b||t|} |\phi_g(t)| dt \leq c_4 \int_{-\infty}^{\infty} e^{-(\delta_4 - |b|)|t|} < \infty.$$

Since the right-hand side is finite, then $f(w)$ exists and is finite in $R = \{z + ib : |b| < r\}$. To prove f is analytic (complex differentiable) in R, we consider the difference quotient

https://doi.org/10.1017/S0266466617000251 Published online by Cambridge University Press

at a point $w_0 = z_0 + ib_0$ in R. For $|\eta| < r_1 < r - |b_0|$,

$$\begin{aligned}|Q(t,\eta)| &\equiv \left|\frac{e^{-it(w_0+\eta)}\phi_g(t) - e^{-itw_0}\phi_g(t)}{\eta}\right| \\ &= \left|\frac{e^{-itw_0}(e^{-it\eta}-1)}{\eta}\phi_g(t)\right| \\ &\leq \frac{\left|e^{-itw_0}\right| e^{r_1|t|}}{r_1}\left|\phi_g(t)\right| \\ &\leq \frac{e^{b_0 t}e^{r_1|t|}}{r_1}\left|\phi_g(t)\right| \\ &\leq c_4\frac{e^{-(\delta_4-|b_0|-r_1)|t|}}{r_1},\end{aligned}$$

where we have used the inequality $\left|\frac{e^{c\zeta}-1}{\zeta}\right| \leq \frac{e^{r_1|c|}}{r_1}$ for $|\zeta| \leq r_1$ and equation (A.7). The condition $|b_0| + r_1 < r < \delta_4$ makes the right-hand side integrable. Since the quotient is bounded above by an integrable function, the Lebesgue dominated convergence theorem implies

$$f'(w_0) = \lim_{\eta\to 0}\int_{-\infty}^{\infty} Q(t,\eta)dt = \int_{-\infty}^{\infty}\lim_{\eta\to 0} Q(t,\eta)dt = -it\int_{-\infty}^{\infty} e^{-itw_0}\phi_g(t)dt.$$

Because w_0 is arbitrary in R, $w \to f(w)$ is analytic (complex differentiable) in $R = \{z + ib : |b| < r\}$. Consequently, the fact that $f(z) = g(z)$ equals zero for a sequence $\{z_1, z_2, z_3, \ldots\}$ converging to z_0 in equation (A.5) implies that f is equal to zero in R by the uniqueness theorem cited in the proof of Lemma 1. This implies that $f(w)$ is equal to zero for all $w = z$ on the real line, i.e., $\int_{-\infty}^{\infty} e^{-itz}\phi_g(t)dt = 0$ for all $z \in \mathbb{R}$. Because the function $\phi_g(\cdot)$ is integrable, this yields $\phi_g(t) = 0$ for all t. (See Theorem 9.12 on page 185 in Rudin (1987).) That is $0 = \int_{\mathbb{R}} e^{itx}h_0(x)\left(\int_{-\infty}^{\infty} e^{it(z-x)} f_\varepsilon(x-z)e^{-\delta_1 z^2}dz\right)dx$ which implies $h_0(\cdot)\left(\int_{-\infty}^{\infty} e^{it(z-\cdot)} f_\varepsilon(\cdot - z)e^{-\delta_1 z^2}dz\right) = 0$ a.e. By equation (A.4), we obtain $h_0 = 0$ a.e. The family $\{g(\cdot|z) = f_\varepsilon(\cdot - z_k) : k = 1, 2, \ldots\}$ is complete in $L^2(\mathbb{R}, \omega)$. ■

A.3. Proof of Theorem 1

The prototype of the stability result in Theorem 1 comes from Problem 2 on page 41 of Young (2001) which implies the following sufficient condition for the stability of completeness in Hilbert spaces.

LEMMA A.2. *Suppose $\{g_k\}$ is a complete sequence for a Hilbert space $\mathcal{H}$. If $\{f_k\}$ is a sequence in $\mathcal{H}$ such that*

$$\sum_{k=1}^{n}\|c_k(g_k - f_k)\| \leq \lambda\sum_{k=1}^{n}\|c_k g_k\|,$$

for some constant λ, $0 \leq \lambda < 1$, and arbitrary scalars $c_1, \ldots, c_n$ $(n = 1, 2, 3, \ldots)$. Then $\{f_k\}$ is also complete in $\mathcal{H}$.

Applying Lemma A.2 to the complete sequence in condition ii) in Theorem 1, we obtain $\{f(\cdot|z_k)/\omega(\cdot) : k = 1, 2, \ldots\}$ is also a complete sequence. Because the weight function is positive, we have shown the family $\{f(\cdot|z) : z \in \mathcal{N}(z_0)\}$ is complete in $\mathcal{H}$.

A.4. Proof of Theorem 2

The stability result used for Theorem 2 comes from Theorem 1 of Dostanić (1990) which adopted the results from Gurarij and Meletidi (1970) for Banach spaces. We summarize the stability results for completeness in Hilbert spaces.

LEMMA A.3. *Let* $\{e_k : k = 1,2,\ldots\}$ *be a complete normalized sequence in a Hilbert space* $\mathcal{H}$ *with* $\|e_k\| = 1$. *Then there is a sequence* $\{C_k : k = 1,2,\ldots\}$ $(C_k > 0)$ *which depends on* $\{e_k : k = 1,2,\ldots\}$, *with the following property: for every* $\varepsilon > 0$ *and* $e \in \mathcal{H}$, $\|e\| = 1$, *there is a finite linear combination such that*

$$\left\| e - \sum_{k=1}^{N} b_k e_k \right\| < \varepsilon,$$

and $|b_k| \le C_k$ *for* $k = 1,2,\ldots,N$.

Gurarij and Meletidi (1970) used the above lemma to prove the following stability results for completeness:

LEMMA A.4. *Let* $\{e_k : k = 1,2,\ldots\}$ *be a complete, normalized sequence in a Hilbert space* $\mathcal{H}$. *If* $\{\varepsilon_k : k = 1,2,\ldots\}$ *is a sequence of positive numbers such that* $\sum_{k=1}^{\infty} C_k \varepsilon_k < 1$ (*sequence* C_k *depends on* e_k *and is given by Lemma 12*) *and for a sequence* $\{f_k : k = 1,2,\ldots\}$ *satisfies*

$$\|f_k - e_k\| < \varepsilon_k,$$

then the sequence $\{f_k : k = 1,2,\ldots\}$ *is complete in the Hilbert space* $\mathcal{H}$.

First, apply Lemma A.3 to the normalized complete sequence $\left\{\frac{g(\cdot|z_k)/\omega(\cdot)}{\|g(\cdot|z_k)/\omega(\cdot)\|} : k = 1,2,\ldots\right\}$ to ensure the existence of the sequence $\{C_k : k = 1,2,\ldots\}$ in condition ii) of Theorem 2. Then, by Lemma A.4, the sequence $\{f(\cdot|z_k)/\omega(\cdot) : k = 1,2,\ldots\}$ is a complete sequence. Because the weight function is positive, we have shown the family $\{f(\cdot|z) : z \in \mathcal{N}(z_0)\}$ is also complete in $\mathcal{H}$.

A.5. Proof of Theorem 3

We prove Theorem 3 in three steps:

1. The quadratic deviation from an orthonormal sequence $\left\{\frac{v_k^g}{\|v_k^g\|} : k = 1,2,\ldots\right\}$ to the corresponding sequence $\left\{\frac{v_k^f}{\|v_k^g\|} : k = 1,2,\ldots\right\}$ is defined as

$$\sum_{k=1}^{\infty} \frac{\left\|v_k^g - v_k^f\right\|^2}{\left\|v_k^g\right\|^2}. \tag{A.8}$$

We show that if the quadratic deviation from an orthonormal basis to an $\omega-$independent sequence is finite, then the latter sequence is also a basis. This result is summarized in Lemma A.5 which is Theorem 15 in Young (2001).

2. Condition ii) implies that the quadratic deviation in equation (A.8) is finite for an orthonormal sequence $\left\{\frac{v_k^g}{\|v_k^g\|}: k=1,2,\ldots\right\}$ constructed by $\{g(\cdot|z_{r_k})/\omega(\cdot):$ $k=1,2,\ldots\}$ and an orthogonal sequence $\left\{\frac{v_k^f}{\|v_k^g\|}: k=1,2,\ldots\right\}$ constructed by $\{f(\cdot|z_{r_k})/\omega(\cdot): k=1,2,\ldots\}$.
3. A linearly independent sequence $\left\{f(\cdot|z_{r_k})\right\}$ in a Hilbert space implies linear independence of the orthogonal sequence $\left\{\frac{v_k^f}{\|v_k^g\|}: k=1,2,\ldots\right\}$. The linearly independent sequence $\left\{\frac{v_k^f}{\|v_k^g\|}: k=1,2,\ldots\right\}$ contains an $\omega-$independent subsequence $\left\{\frac{v_{k_l}^f}{\left\|v_{k_l}^g\right\|}: l=1,2,\ldots\right\}$. Finally, for an orthonormal sequence constructed by a complete sequence $\left\{g(\cdot|z_{r_{k_l}})/\omega(\cdot)\right\}$ in a Hilbert space and the $\omega-$independent sequence $\left\{\frac{v_{k_l}^f}{\left\|v_{k_l}^g\right\|}: l=1,2,\ldots\right\}$, equation (A.8) and Lemma A.5 imply that the sequence $\left\{\frac{v_{k_l}^f}{\left\|v_{k_l}^g\right\|}: l=1,2,\ldots\right\}$ is complete in a Hilbert space, and therefore, $\{f(\cdot|z): z \in \mathcal{N}(z_0)\}$ is complete.

Step 1. We prove that if the quadratic deviation from an orthonormal basis to an $\omega-$independent sequence is finite, then the latter sequence is also a basis. This result is Theorem 15 in Young (2001) and summarized in the following lemma.

LEMMA A.5. *Suppose that*

i) *the sequence* $\{e_n(\cdot): n=1,2,\ldots\}$ *is an orthonormal basis in a Hilbert space* $\mathcal{H}$;
ii) *the sequence* $\{f_n(\cdot): n=1,2,\ldots\}$ *in* $\mathcal{H}$ *is* $\omega-$*independent*;
iii) $\sum_{n=1}^{\infty}\|f_n(\cdot)-e_n(\cdot)\|^2<\infty$.
Then, the sequence $\{f_n(\cdot): n=1,2,\ldots\}$ *is a basis in* $\mathcal{H}$.

Step 2. First, by Lemma A.1(3), we can extract a convergence subsequence $\{r_1,r_2,r_3,\ldots\}$ such that the orthogonal basis constructed by the basis $\{g(\cdot|z_{r_k})/\omega(\cdot): k=1,2,\ldots\}$ is given by

$$v_1^g(\cdot)=g(\cdot|z_{r_1})/\omega(\cdot),$$

$$v_2^g(\cdot)=g(\cdot|z_{r_2})/\omega(\cdot)-\frac{\langle g(\cdot|z_{r_2})/\omega(\cdot),v_1(\cdot)\rangle}{\langle v_1(\cdot),v_1(\cdot)\rangle}v_1^g(\cdot),$$

$$\vdots$$

$$v_k^g(\cdot)=g(\cdot|z_{r_k})/\omega(\cdot)-\sum_{j=1}^{k-1}\frac{\langle g(\cdot|z_{r_j})/\omega(\cdot),v_j(\cdot)\rangle}{\langle v_j(\cdot),v_j(\cdot)\rangle}v_j^g(\cdot),$$

$$\vdots$$

https://doi.org/10.1017/S0266466617000251 Published online by Cambridge University Press

We can normalize the orthogonal basis to obtain an orthonormal basis as $\{v_k^g(\cdot)/\|v_k^g(\cdot)\| : k = 1,2,\ldots\}$. The orthogonal basis constructed by the basis $\{f(\cdot|z_{r_k})/\omega(\cdot) : k = 1,2,\ldots\}$ is the following sequence

$$
\begin{aligned}
v_1^f(\cdot) &= f(\cdot|z_{r_1})/\omega(\cdot),\\
v_2^f(\cdot) &= f(\cdot|z_{r_2})/\omega(\cdot) - \frac{\langle g(\cdot|z_{r_2})/\omega(\cdot), v_1(\cdot)\rangle}{\langle v_1(\cdot), v_1(\cdot)\rangle} v_1^f(\cdot),\\
&\vdots\\
v_k^f(\cdot) &= f(\cdot|z_{r_k})/\omega(\cdot) - \sum_{j=1}^{k-1} \frac{\langle g(\cdot|z_{r_j})/\omega(\cdot), v_j(\cdot)\rangle}{\langle v_j(\cdot), v_j(\cdot)\rangle} v_j^f(\cdot),\\
&\vdots
\end{aligned}
$$

This implies $\sum_{k=1}^{\infty} \left\| \frac{v_k^g(\cdot)}{\|v_k^g(\cdot)\|} - \frac{v_k^f(\cdot)}{\|v_k^g(\cdot)\|} \right\|^2 < \sum_{k=1}^{\infty} \frac{\left\| v_k^g - v_k^f \right\|^2}{\|v_k^g\|^2} < \infty$ by Condition ii). This implies the sequence $\left\{ \frac{v_k^f(\cdot)}{\|v_k^g(\cdot)\|} : k = 1,2,\ldots \right\}$ is quadratically close to the orthonormal basis $\left\{ \frac{v_k^g(\cdot)}{\|v_k^g(\cdot)\|} : k = 1,2,\ldots \right\}$.

Step 3. By the construction of $\left\{ \frac{v_k^f}{\|v_k^g\|} : k = 1,2,\ldots \right\}$ and linear independence of $\{f(\cdot|z_k) : k = 1,2,\ldots\}$ in Condition iii), $\left\{ \frac{v_k^f}{\|v_k^g\|} : k = 1,2,\ldots \right\}$ is also linearly independent. 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 $\left\{ \frac{v_{k_l}^f}{\|v_{k_l}^g\|} : l = 1,2,\ldots \right\}$.

We then show that the $\omega-$independent subsequence $\left\{ \frac{v_{k_l}^f}{\|v_{k_l}^g\|} : l = 1,2,\ldots \right\}$ is complete in the Hilbert space $\mathcal{H}$. Since the sequence $\{z_{r_{k_l}}\}$ corresponding to $\{\tilde{f}(\cdot|z_{r_{k_l}})\}$ is a subsequence of $\{z_k\}$ and also converges to z_0, condition i) implies that the corresponding sequence $\{g(\cdot|z_{r_{k_l}})/w(\cdot)\}$ is complete in the Hilbert space defined on $\mathcal{X}$. This implies the orthonormal sequence constructed by a complete sequence $\{g(\cdot|z_{r_{k_l}})/\omega(\cdot)\}$ and $\left\{ \frac{v_{k_l}^f}{\|v_{k_l}^g\|} : l = 1,2,\ldots \right\}$ also satisfies equation (A.8). Lemma A.5 implies that $\left\{ \frac{v_{k_l}^f}{\|v_{k_l}^g\|} : l = 1,2,\ldots \right\}$ is a basis and thus $\left\{ \frac{v_{k_l}^f}{\|v_{k_l}^g\|} : l = 1,2,\ldots \right\}$ is complete. By the construction of $\left\{ \frac{v_{k_l}^f}{\|v_{k_l}^g\|} : l = 1,2,\ldots \right\}$, the completeness of $\left\{ \frac{v_{k_l}^f}{\|v_{k_l}^g\|} : l = 1,2,\ldots \right\}$ implies that the family $\{f(\cdot|z) : z \in \mathcal{N}(z_0)\}$ is also complete.

A.6. 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

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

According to the property of the Vandermonde matrix, the determinant $W(x)$ in equation (6) is not equal to zero when $F_0(x)$ has all the nonzero derivatives 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 the $\mu(z_k)$ are different for a distinct sequence $\{z_k\}$ around z_0, the condition that $\tau(\cdot)$ 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 function $f(x|z)$ is nonzero at $x = 0$. ■

A.7. Proof of Completeness in Applications

Proof of Lemma 4. Since f_ε is normally distributed, by Lemma 2 we may generate a complete sequence $\{g(x|z_k) = f(x - z_k|z_0) = f_\varepsilon(x - z_k) : k = 1, 2, \dots\}$ satisfying condition i) in Theorems 1, 2, and 3. Condition ii) 2), condition ii) 3), and condition ii) 4), respectively, are condition ii) of Theorems 1, 2, and 3. Thus, condition i) and condition ii) 2), condition i) and condition ii) 3), and condition i) and condition ii) 4), respectively, satisfy conditions of Theorems 1, 2, and 3 and those theorems imply $\{f(\cdot|z) : z \in \mathcal{Z}\}$ is complete in $L^2(\mathbb{R}, \omega)$.

As for the condition ii) 1), pick a distinct sequence $\{z_k : k = 1, 2, \dots\}$ such that z_k converging to z_0 and $\|z_k - z_0\| < \varepsilon$. Then, we have (1) $\{f_\varepsilon(\cdot - z_k) : k = 1, 2, \dots\}$ satisfies condition i) of Theorem 2 by condition i), and (2) condition ii) of Theorem 2 is also satisfies by

$$\left\| \frac{f_\varepsilon(\cdot - z_k)/\omega(\cdot)}{\|f_\varepsilon(\cdot - z_k)/\omega(\cdot)\|} - \frac{1}{\|f_\varepsilon(\cdot - z_k)/\omega(\cdot)\|} \frac{1}{\sigma(z_k)} f_\varepsilon\left(\frac{\cdot - z_k}{\sigma(z_k)}\right)/\omega(\cdot) \right\| = 0,$$

because $\sigma(z_k) = 1$ for $k = 1, 2, \dots$. Theorem 2 implies that $\left\{ f_\varepsilon\left(\frac{\cdot - z_k}{\sigma(z_k)}\right) : z \in \mathcal{Z} \right\}$ is complete in $L^2(\mathcal{X}, \omega)$. ■

https://doi.org/10.1017/S0266466617000251 Published online by Cambridge University Press

Proof of Lemma 5. We take distinct $z_k \to z_0$ such that $|z_k - z_0| < \epsilon$. Consider the sequence $\{g(x|z_k) = f_\varepsilon(x - z_k) : k = 1,2,\ldots\}$. This implies that $g(x|z_0) = f_\varepsilon(x) = f(x|z_0)$ because $h(z_0,\varepsilon) = \varepsilon$. In addition, the assumptions of ε imply $\{g(\cdot|z_k) : k = 1,2,\ldots\}$ is complete in $L^2(\mathbb{R},\omega)$ for the weight function ω by Lemma 2. Then the complete sequence $\{g(\cdot|z_k) : k = 1,2,\ldots\}$ satisfies the condition i) in Theorems 1, 2, and 3.

We may check that the family $\left\{f(x|z_k) = \left|\frac{\partial}{\partial x}h^{-1}(z_k,x)\right| f_\varepsilon\left(h^{-1}(z_k,x)\right) : k = 1,2,\ldots\right\}$ is in $L^2(\mathbb{R},\omega)$. Consider for some constant c_1 and $z \in \mathcal{N}(z_0)$

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

The last step holds because condition i) and the 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 $f(x|z) \in L^2(\mathbb{R})$ for $z \in \mathcal{N}(z_0)$. Since the weight function is bounded, $f(x|z) \in L^2(\mathbb{R},\omega)$ for $z \in \mathcal{N}(z_0)$.

Similar to the proof of Lemma 4, the results in condition ii) 2), condition ii) 3), and condition ii) 4) are direct applications of Theorems 1, 2, and 3, and these results imply that completeness of $\{f(\cdot|z) : z \in \mathcal{N}(z_0)\}$ in $L^2(\mathbb{R},\omega)$. As for the results in condition ii) 1), for z such that $\|z - z_0\| < \varepsilon$, $f(\cdot|z) = \left|\frac{\partial}{\partial x}h^{-1}(z,\cdot)\right| f_\varepsilon\left(h^{-1}(z,\cdot)\right) = f_\varepsilon(\cdot - cz)$. Because $c \neq 0$, cz_k is also a converging sequence. Then, the assumptions of ε also imply $\{f(\cdot|z_k) : k = 1,2,\ldots\}$ is complete in $L^2(\mathbb{R},\omega)$ by Lemma 2. We have the completeness of $\{f(\cdot|z) : z \in \mathcal{N}(z_0)\}$. ■

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 $\{f_{x_1|z_1}(\cdot|z_{1k}) : k = 1,2,3,\ldots\}$ and $\{f_{x_2|z_2}(\cdot|z_{2k}) : k = 1,2,3,\ldots\}$ are complete in their corresponding Hilbert spaces.

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

$$
\begin{aligned}
&\int\!\!\int h(x_1,x_2) f(x_1|z_1) f(x_2|z_2)\,dx_1 dx_2 \\
&\quad= \int\!\!\int h(x_1,x_2)\frac{f(x_1|z_1)f(x_2|z_2)}{\omega(x_1,x_2)}\omega(x_1,x_2)\,dx_1 dx_2 \\
&\quad= \int\left(\int h(x_1,x_2)\frac{f(x_1|z_1)}{\omega(x_1)}\omega(x_1)\,dx_1\right)\frac{f(x_2|z_2)}{\omega(x_2)}\omega(x_2)\,dx_2
\end{aligned}
$$

https://doi.org/10.1017/S0266466617000251 Published online by Cambridge University Press

$$= \int \left(\int h(x_1, x_2) f(x_1|z_1) dx_1 \right) f(x_2|z_2) dx_2$$
$$\equiv \int h'(x_2, z_1) f(x_2|z_2) dx_2.$$

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 almost sure $x_2 \in \mathcal{X}_2$ 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 almost sure $x_1 \in \mathcal{X}_1$. Therefore, the sequence $\{f_{x_1|z_1}(\cdot|z_{1k}) f_{x_2|z_2}(\cdot|z_{2k}) : k = 1, 2, 3, \ldots\}$ is complete. Thus, we have a family of functions satisfying the condition i) in Theorems 1, 2, and 3. Because its corresponding condition in the condition ii) in Theorems 1, 2, and 3 are assumed directly in condition ii) 2), condition ii) 3), and condition ii) 4), respectively, the sequence $\{f_{x_1,x_2|z_1,z_2}(\cdot,\cdot|z_{1k}, z_{2k}) : k = 1, 2, 3, \ldots\}$ is complete in these conditions. As for the first part, we can regard it as a special case of the third part with zero deviation in a converging sequence. We have reached our claim. ■

https://doi.org/10.1017/S0266466617000251 Published online by Cambridge University Press