Journal of Econometrics 200 (2017) 48–58

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Injectivity of a class of integral operators with compactly supported kernels

Yingyao Hu [a], Susanne M. Schennach [b], Ji-Liang Shiu [c,*]

[a] Department of Economics, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218, USA
[b] Department of Economics, Brown University, 270 Bay State Road, Boston, MA 02215-1403, USA
[c] Institute for Economic and Social Research, Jinan University, China

ARTICLE INFO

Article history:
Received 6 June 2015
Received in revised form 13 January 2017
Accepted 11 May 2017
Available online 12 June 2017

Keywords:
Nonparametric identification
Instrumental variable
Completeness
Endogeneity

ABSTRACT

Injectivity of integral operators is related to completeness conditions of their corresponding kernel functions. Completeness provides a useful way of obtaining nonparametric identification in various models including nonparametric regression models with instrumental variables, nonclassical measurement error models, and auction models, etc. However, the condition is quite abstract for empirical work and lacks a proper economic interpretation. We rely on known results regarding the Volterra equation to provide sufficient conditions for completeness conditions for densities with compact support. Our conditions include various smoothness assumptions and monotonously moving support assumptions on the kernel function of the operator. We apply our results to establish nonparametric identification in nonparametric IV regression models, nonclassical measurement error models, and auction models with an accessible interpretation and without specific functional form restrictions.

1. Introduction

Rank conditions are widely used to identify linear and/or parametric economic models. However, nonparametric models have been receiving considerable and increasing attention in recent years, and the intuitive notion of "full rank" in that context must be generalized to the more abstract notions of injectivity or completeness. The latter, unfortunately, do not admit simple, intuitive, and general sufficient conditions. In this paper, we use known conditions for the unique solution of the so-called Volterra equation to provide simple sufficient conditions for injectivity of a class of integral operators with compactly supported kernel functions. Consider the following integral equation

$$h(y) = \int_{\mathcal{X}_y} K(y, x)\, g(x)\, dx \text{ for any } y \in \mathcal{Y}, \tag{1}$$

where $h(y)$, $K(y, x)$ are given functions and $g(x)$ is an unknown function to be determined. Here, the support of $K(y, x)$ for each $y \in \mathcal{Y}$ is $\mathcal{X}_y$. If $\mathcal{X}_y = [a, y]$, Eq. (1) is a linear Volterra integral equation of the first kind. In this case, if the kernel function $K(y, x)$ is nonzero on the "diagonal", i.e., $K(y, x) \neq 0$ for $y = x$, and if $K(y, x)$ and $\frac{\partial}{\partial y} K(y, x)$ are differentiable and square integrable, then there exists a unique square-integrable solution of the Volterra equation.

Define a $\mathcal{L}^2$ space $\mathcal{L}^2(\mathcal{X}) = \{h(\cdot) : \int_{\mathcal{X}} |h(x)|^2 dx < \infty\}$, where $\mathcal{X}$ is a closed interval in $\mathbb{R}$. We can rewrite Eq. (1) as an integral operator relationship through the following:

$$(T_K g)(y) = \int_{\mathcal{X}_y} K(y, x)\, g(x)\, dx = h(y), \tag{2}$$

where T_K is an integral operator from $\mathcal{L}^2(\mathcal{X})$ to $\mathcal{L}^2(\mathcal{Y})$ with the kernel function $K(y, x)$. The unique solution of the Volterra equation implies that the corresponding integral operator is injective.

The injective property of an integral operator is related to the concept of completeness of its corresponding kernel function. Let X and Z be two random variables. For $m(\cdot) \in \mathcal{L}^2(\mathcal{X})$,

$$E[m(X)|Z = z] = \int m(x) f(x|z) dx, \tag{3}$$

where the probability measure of X conditional on Z is absolutely continuous w.r.t. the Lebesgue measure. The conditional distribution $f(X|Z)$ of random elements X and Z is $\mathcal{L}^2$-complete w.r.t. X if $E[m(X)|Z] = 0$ a.s. implies $m(X) = 0$ a.s. The completeness equals the injectivity of the conditional expectation operator using the conditional distribution $f(X|Z)$ as a kernel function. Completeness has been used to obtain identification conditions for various nonparametric and semiparametric econometric models, including nonparametric models with instrumental variables (see Newey and Powell (2003); Ai and Chen (2003); Chernozhukov and Hansen (2005); Hall and Horowitz (2005); Blundell et al. (2007); Chernozhukov et al. (2007); Horowitz and

* Corresponding author.
E-mail addresses: yhu@jhu.edu (Y. Hu), smschenn@brown.edu (S.M. Schennach), jishiu.econ@gmail.com (J.-L. Shiu).

http://dx.doi.org/10.1016/j.jeconom.2017.05.013
0304-4076/

Lee (2007); Darolles et al. (2011); Horowitz (2011)), measurement error models (see Hu and Schennach (2008); An and Hu (2012); Carroll et al. (2010); Chen and Hu (2006)), and dynamic models (see Hu and Shum (2012); Shiu and Hu (2013)), etc.

The definition of completeness may not relate to any economic meaning. Therefore, the condition is quite abstract for empirical work and is treated as a high-level condition. A number of papers provided sufficient conditions for different versions of completeness. Newey and Powell (2003) give sufficient conditions for completeness of distributions with discrete finite support and exponential parametric families. Andrews (2011) considers $\mathcal{L}^2$-completeness and provides nonparametric classes of $\mathcal{L}^2$-complete distributions for bivariate density functions as for $F_x \times F_z$ through orthonormal bases of $\mathcal{L}^2$ spaces of x and z. D'Haultfoeuille (2011) considers different versions of completeness in three different spaces including (1) $\mathcal{L}^1$ spaces, (2) P-completeness, where a function is bounded by a polynomial, and (3) bounded completeness for any bounded function in nonparametric models. Hu and Shiu (2016) use stability results in Banach or Hilbert spaces to provide sufficient conditions to extend known complete distribution functions to complete nonparametric families without imposing a specific functional form. D'Haultfoeuille and Février (2010) consider the issue of identifying nonparametric mixture models where all observed variables depend on a common and unobserved component. They show that these models are identified nonparametrically if support of the observed variables move with the true value of the unobserved component.

This paper develops additional examples of sufficient conditions for completeness conditions in a compact support by applying the result of the Volterra equation. For nonparametric regression models with a limited endogenous regressor, we consider that the support $\mathcal{X}$ is bounded, e.g., equal to $[0, 1]$. We show that if the conditional distribution $f(x|z)$ satisfies some regularity conditions and the support of the function $f(\cdot|z)$ is moving with the variable z like $[0, z]$ for $z \in [0, 1]$ and the support is degenerate at $z = 0$ then the density function is complete and the regression function is identified. The conditions are particularly well suited to empirical applications, because they are easily interpreted and imposing such conditions is not very restrictive. The moving support condition is closely related to the condition in D'Haultfoeuille and Février (2010), although the aims differ from those of the present paper. They seek to identify the distribution of the conditioning variable from repeated conditionally independent measurements and propose a specific estimator that uses the moving support condition to exploit extreme values of the observed variables to infer the conditional distribution of the unobservable variables. Our proposed approach is more general in that it focuses on the concept of injectivity and can be applied in any approach that relies on this concept, independent of repeated measurement availability. The methods of proof are also completely different, as our approach does not rely on rare extreme events.

The use of our sufficient conditions for completeness in the econometrics literature is potentially vast because a number of econometric models can be written as an integral equation, such as nonparametric regression models with instrumental variables, nonclassical measurement error models, and auction models, etc. Our results allow one to establish nonparametric identification in these models with an accessible interpretation and without specific functional form restrictions.

The plan of the paper is as follows. In Section 2, we describe sufficient conditions for injectivity of integral operators with compactly supported kernels. In Section 3, we apply the conditions in Section 2 to several models and discuss identification including nonparametric regression models with instrumental variables, nonclassical measurement error models, and auction models. In Section 4 we present our conclusions while in Appendix we provide proof.

2. Integral operators with compactly supported kernels

Consider the equation

$$h(y) = \int_{\underline{b}(y)}^{\bar{b}(y)} K(y, x)\, g(x)\, dx \text{ for any } y \in \mathcal{Y} \tag{4}$$

where we let $\mathcal{Y} = [\underline{y}, \bar{y}]$ and $\mathcal{X} \equiv [\inf_{y\in\mathcal{Y}} \underline{b}(y), \sup_{y\in\mathcal{Y}} \bar{b}(y)]$ and where $\underline{b} : \mathcal{Y} \mapsto \mathcal{X}$, $\bar{b} : \mathcal{Y} \mapsto \mathcal{X}$, $h : \mathcal{Y} \mapsto \mathbb{R}$, $g : \mathcal{X} \mapsto \mathbb{R}$, $K : \mathcal{Y} \times \mathcal{X} \mapsto \mathbb{R}$.

Let parenthesized superscripts denote the number of derivatives for each argument of a function (e.g. $K^{(1,2)}(y, x) \equiv \partial^3 K(y, x)/\partial y \partial x^2$). Suppose a is a boundary point of the set S and denote $\text{int}(S)$ as the interior of S. The derivative of a function f at the point a is defined as follows:

$$f^{(1)}(a) = \lim_{\substack{s\to a\\ s\in \text{int}(S)}} \frac{f(s) - f(a)}{s - a}.$$

Denote $\mathbb{N}^*$ as the set of all strictly positive integers and let the following assumptions hold for some $L \in \mathbb{N}^*$:

Assumption 2.1. $h^{(L)}(y)$ exists and is continuous for all $y \in \mathcal{Y}$.

Assumption 2.2. (i) $\underline{B} \le \underline{b}(y) \le \bar{b}(y) \le \bar{B}$ for all $y \in \mathcal{Y}$ for some $\underline{B}, \bar{B} \in \mathbb{R}$ and (ii) $\underline{b}(\underline{y}) = \bar{b}(\underline{y}) = \underline{B}$.

Assumption 2.3. $\underline{b}^{(1)}(y)$ and $\bar{b}^{(1)}(y)$ exist and are continuous for all $y \in \mathcal{Y}$.

Assumption 2.4. $K^{(L,0)}(y, x)$ exists and is continuous on $\cup_{y\in\mathcal{Y}} \{y\} \times [\underline{b}(y), \bar{b}(y)]$.

Assumption 2.5. At each $y \in \mathcal{Y}$, either (i) $\underline{b}^{(1)}(y) = 0$ or (ii) $K^{(\ell,0)}(y, \underline{b}(y)) = 0$ for $\ell = 0, \ldots, L-1$.

Assumption 2.6. $K^{(L,0)}(y, \underline{b}(y)) = 0$ for all $y \in \mathcal{Y}$.

Assumption 2.7. (i) $\bar{b}^{(1)}(y) > 0$ for all $y \in \mathcal{Y}$, (ii) $K^{(L-1,0)}(y, \bar{b}(y)) \neq 0$ for all $y \in \mathcal{Y}$ and (iii) if $L > 1$, $K^{(\ell,0)}(y, \bar{b}(y)) = 0$ for $\ell = 0, \ldots, L-2$ and for all $y \in \mathcal{Y}$.

Theorem 2.1. *Let Assumptions 2.1–2.7 hold and let h and K be given. Then, Eq. (4) admits a unique solution $g(x)$.*

Proof. See Appendix.

The proof of Theorem 2.1 is to treat Eq. (4) as a Volterra equation of the first kind and then transform the equation into a Volterra equation of the second kind, through iterated differentiations (to handle our more general differentiability conditions) and suitable changes of variables (to handle our more general integral bounds).

Assumptions 2.1 and 2.4 impose regularity conditions on h and K. As we show in the proof of Theorem 2.1, these regularity conditions allow us to differentiate Eq. (4) with respect to y and then transform the differentiated equation into a Volterra equation of the second kind. Assumption 2.2 imposes uniform bound conditions for $\underline{b}(\cdot)$ and $\bar{b}(\cdot)$ and ensures that the two endpoints of the support attain the lowest value $\underline{B}$ in the lowest value of y. This implies we have a degenerate support at $\underline{y}$, $\mathcal{X}_{\underline{y}} = \underline{B}$ and we can refer to Assumption 2.2(ii) as a *degenerate support* condition. Assumption 2.3 imposes regularity conditions on the two endpoints of the support. Assumptions 2.5 and 2.7(iii) are used to show some boundary terms vanish and we can obtain a Volterra equation of the second kind. This implies there is a constant lower endpoint or we need to require that $K(y, \cdot)$ has vanishing derivatives at the lower endpoints for each y. Assumption 2.6 further assumes $K(y, \cdot)$

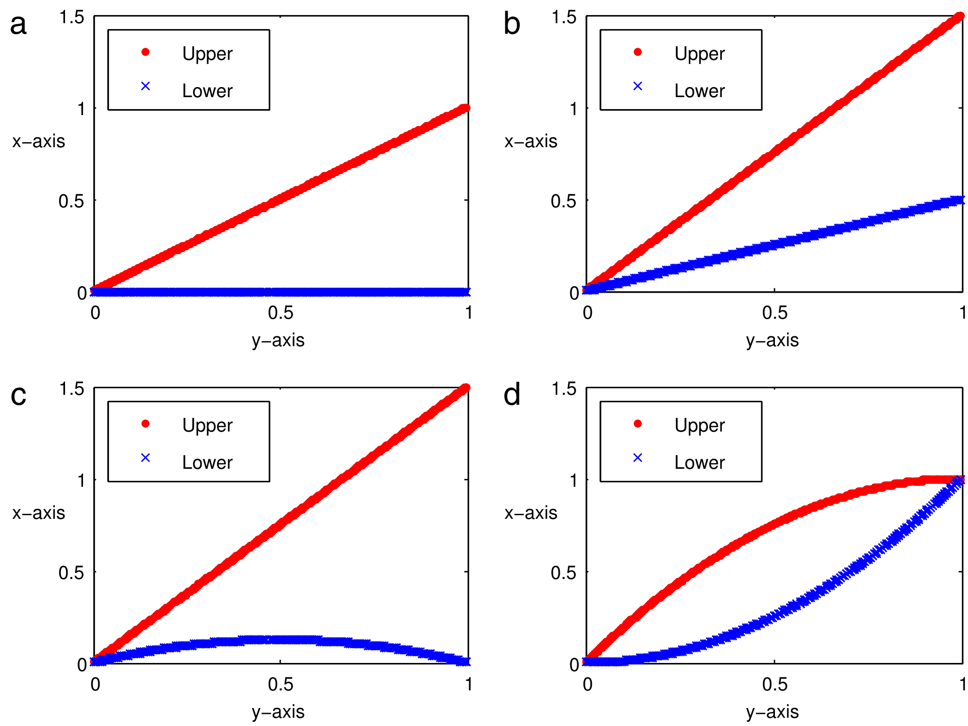

Fig. 1. The two endpoints satisfy the assumptions.

has a vanishing higher order derivative at the lower endpoints. Assumption 2.7(i) & (ii) ensures functions in the derived Volterra equation of the second kind are well defined so we can apply results on Volterra equation. We also consider the discrete case at the end of this section. A number of remarks are in order.

Remark 2.1. Consider the baseline case

$$h(y) = \int_{\underline{y}}^{y} K(y, x)\, g(x)\, dx \text{ for any } y \in \mathcal{Y}. \tag{5}$$

In this case, the support of $g(x)$ becomes $\cup_{y\in\mathcal{Y}} \left[\underline{y}, y\right] = \left[\underline{y}, \bar{y}\right]$ so x and y have to share the same support $\left[\underline{y}, \bar{y}\right]$. That means $g(\cdot)$ is only uniquely identified over this support.

Remark 2.2. When we allow the upper bound and lower bound of the integral equation to change with y, the support of the function $g(\cdot)$ is not over the interval $\left[\underline{y}, \bar{y}\right]$ instead is over $\cup_{y\in\mathcal{Y}} \left[\underline{b}(y), \bar{b}(y)\right]$.

Remark 2.3. The fact that the conditions on the derivatives are slightly asymmetric on the two endpoints of the support is important to obtain injectivity. It is more likely that the two endpoints have different smoothness rather than exactly the same level of smoothness, so the condition is not unnatural. The examples in which the support $\left[\underline{b}(y), \bar{b}(y)\right]$ for each $y \in [0, 1]$ satisfy or fail our assumptions are plotted in Figs. 1–2. In Fig. 1, the upper endpoints of (a), (b), (c), and (d) are strictly increasing and then satisfy Assumption 2.7(i) and the starting points of $\underline{b}(\cdot)$ and $\bar{b}(\cdot)$ both attain the lowest value $\underline{B}$, which satisfies Assumption 2.2(ii). Only Fig. 1(a) has a constant lower endpoint and this satisfies Assumption 2.5(i). As for Fig. 1(b),(c),(d), an additional derivative condition for $K(y, x)$ in Assumption 2.5(ii) is required. In Fig. 2, (a),(b), and (c) fails Assumption 2.2(ii). In Fig. 2(d), the upper endpoint is not strictly increasing and this implies Assumption 2.7(i) fails.

Remark 2.4. For a single endogenous regressor, including multi-dimensional instruments is trivial because if injectivity is proven for one instrument (or for one linear combination of many instruments), injectivity will automatically hold if additional instruments are included. An extension to multivariate endogenous regressor is more challenging and would deserve a separate paper. An appealing aspect of our present approach is that, as the boundary of the support moves, the boundary terms that arise when we differentiate our integral operator are functions of the unknown function at a single point, which enables a fairly elegant solution. In the multivariate case, the boundary terms would be integrals themselves (over a lower dimensional manifold) and a similar solution would not be as straightforward.

Remark 2.5. For random variables with compact support, it would seem more likely that the support of one variable conditional on another would change as the conditioning variable changes, rather than remaining *exactly* the same. In fact, other researchers have argued in favor of the plausibility of support variations (see, e.g. D'Haultfoeuille and Février (2010)). If one firmly believes that all variables must necessarily have infinite support, then, indeed, our conditions would not hold. However, it would be hard to argue that all empirically available quantities necessarily have this extreme property. Most empirical researchers would certainly agree that all measurable quantities have some finite (albeit large) bound.

The results presented in Theorem 2.1 can be extended to the following case:

$$h(y_1, y_2) = \int_{\underline{b}(y_1,y_2)}^{\bar{b}(y_1,y_2)} K(y_1, y_2, x)\, g(x, y_1)\, dx \tag{6}$$

where x and y_2 are scalars, and $y_1 \in \mathcal{Y}_1 \subset \mathbb{R}^{d_1}$. For a fixed y_1, Eq. (6) has the same form as Eq. (4). Thus, for each y_1 we can apply Theorem 2.1 to $h(y_1, \cdot)$ and $K(y_1, \cdot, \cdot)$ to obtain a unique $g(\cdot, y_1)$ over the support $\cup_{y_2\in\mathcal{Y}_2}[\underline{b}(y_1, y_2), \bar{b}(y_1, y_2)]$ where $\mathcal{Y}_2 \equiv \left[\underline{y}_2, \bar{y}_2\right]$ is an interval containing y_2. Let $y \equiv (y_1, y_2) \in \mathcal{Y} \equiv \mathcal{Y}_1 \times \mathcal{Y}_2 \subset \mathbb{R}^{d_1+1}$ and the support of $K(y_1, y_2, \cdot)$ for each $y_1 \in \mathcal{Y}_1$ be $\mathcal{X}_{y_2;y_1}$.

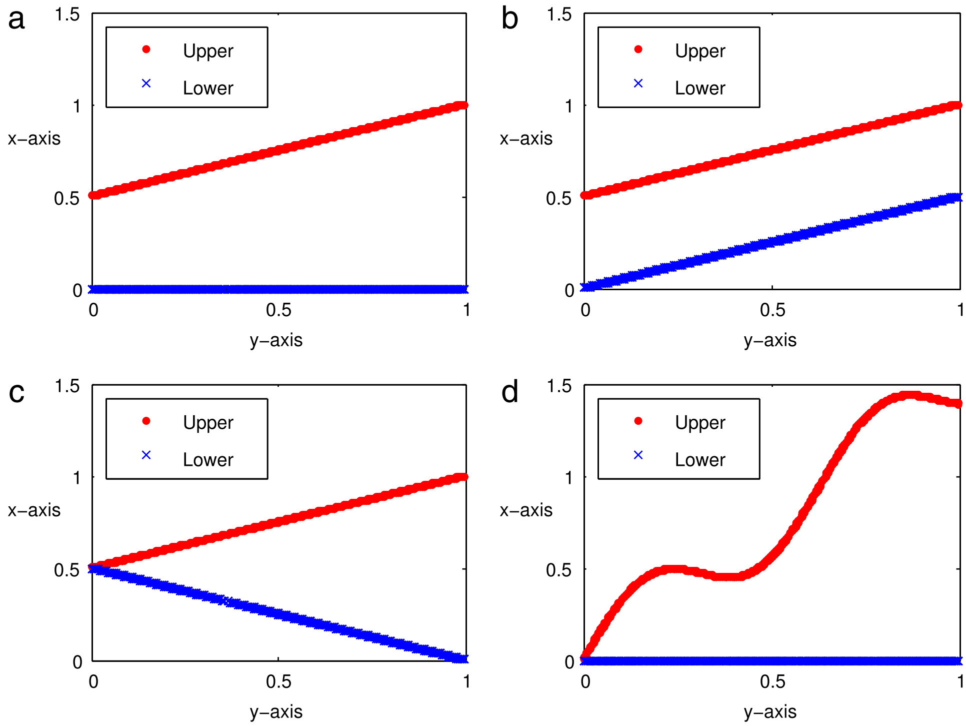

Fig. 2. The two endpoints fail the assumptions.

Assumption 2.8. $h^{(0,L)}(y_1, y_2)$ exists and is continuous for all $y \in \mathcal{Y}$.

Assumption 2.9. (i) $\underline{B} \le \underline{b}(y) \le \bar{b}(y) \le \bar{B}$ for all $y \in \mathcal{Y}$ for some $\underline{B}, \bar{B} \in \mathbb{R}$ and (ii) $\underline{b}(y_1, \underline{y}_2) = \bar{b}(y_1, \underline{y}_2) = \underline{B}$ for each y_1.

Assumption 2.10. $\underline{b}^{(0,1)}(y_1, y_2)$ and $\bar{b}^{(0,1)}(y_1, y_2)$ exist and are continuous for all $y \in \mathcal{Y}$.

Assumption 2.11. $K^{(0,L,0)}(y_1, y_2, x)$ exists and is continuous on $\cup_{y\in\mathcal{Y}}\{y\} \times \left[\underline{b}(y), \bar{b}(y)\right]$.

Assumption 2.12. At each $y \in \mathcal{Y}$, either (i) $\underline{b}^{(0,1)}(y_1, y_2) = 0$ or (ii) $K^{(0,\ell,0)}\left(y_1, y_2, \underline{b}(y)\right) = 0$ for $\ell = 0, \ldots, L-1$.

Assumption 2.13. $K^{(0,L,0)}\left(y_1, y_2, \underline{b}(y)\right) = 0$ for all $y \in \mathcal{Y}$.

Assumption 2.14. (i) $\bar{b}^{(0,1)}(y_1, y_2) > 0$ for all $y \in \mathcal{Y}$, (ii) $K^{(0,L-1,0)}\left(y_1, y_2, \bar{b}(y)\right) \neq 0$ for all $y \in \mathcal{Y}$ and (iii) if $L > 1$, $K^{(0,\ell,0)}\left(y_1, y_2, \bar{b}(y)\right) = 0$ for $\ell = 0, \ldots, L-2$ and for all $y \in \mathcal{Y}$.

The following result is a direct application of Theorem 2.1; hence, we omit its proof.

Corollary 2.1. *Let Assumptions 2.8–2.14 hold and let $h(y_1, \cdot)$ and $K(y_1, \cdot, \cdot)$ be given for each $y_1 \in \mathcal{Y}_1$. Then, Eq. (6) admits a unique solution $g(\cdot, y_1)$ for each y_1.*

For a given function $h(y_1, \cdot)$ for each y_1 satisfying Eq. (6) and Assumption 2.8, Corollary 2.1 shows how to identify the function $g(\cdot, y_1)$ for each y_1 under suitable assumptions. This suggests that we can rewrite Eq. (6) as an integral operator relationship. Let $\mathcal{C}^L(\mathcal{U})$ stand for the space of a function $f(\cdot)$ whose L−th derivative is continuous over the support $\mathcal{U}$ with a sup norm, $\|f\| = \sup_{x\in\mathcal{U}} |f(x)| < \infty$. For each y_1, define an integral operator as follows:

$$T_{K_{y_1}} : \mathcal{C}(\mathcal{X}_{y_2;y_1} \times \{y_1\}) \to \mathcal{C}^L(\{y_1\} \times \mathcal{Y}_2)$$

$$(T_{K_{y_1}} g)(y_1, y_2) = \int_{\underline{b}(y)}^{\bar{b}(y)} K(y_1, y_2, x)\, g(x, y_1)dx,$$

where $K(y_1, y_2, x)$ is the kernel function of the operator $T_{K_{y_1}}$. Using the notation, Corollary 2.1 implies there exists a unique solution $g(\cdot, y_1)$ such that $T_{K_{y_1}} g = h$ for a given $h \in \mathcal{C}^L(\{y_1\} \times \mathcal{Y}_2)$ for each y_1.

Corollary 2.2. *Under Assumption 2.8–2.14, the integral operator $T_{K_{y_1}}$ is injective for each y_1.*

The injectivity of an integral operator is connected to the completeness of the family of functions related to its kernel function.

Definition 2.1. Set $\mathcal{X}_{y_2;y_1} = \left[\underline{b}(y_1, y_2), \bar{b}(y_1, y_2)\right]$ and then $\mathcal{X} = \cup_{y\in\mathcal{Y}} \mathcal{X}_{y_2;y_1}$. The family $\{K(y, \cdot) \in \mathcal{C}(\mathcal{X}) : y \in \mathcal{Y}\}$ is said to be complete in $\mathcal{C}(\mathcal{X}_{y_2;y_1} \times \{y_1\})$ for each y_1 if for each y_1 and for any $g(\cdot, y_1) \in \mathcal{C}(\mathcal{X}_{y_2;y_1} \times \{y_1\})$

$$\int_{\mathcal{X}_{y_2;y_1}} K(y_1, y_2, x)\, g(x, y_1)dx = 0 \quad \textit{for all } y_2 \in \mathcal{Y}_2$$

implies $g(\cdot, y_1) = 0$ almost surely in $\mathcal{X}_{y_2;y_1}$ for each y_1.

In this case, the support of the kernel function $K(y, \cdot)$ is $\mathcal{X}_{y_2;y_1}$, which changes with y. The key assumption of Theorem 2.1 and Corollary 2.1 is that the upper bound of the support increases with y (by Assumption 2.7(i) and Assumption 2.14(i) respectively) and the support is degenerated at $\underline{y}$ (by Assumption 2.2(ii) and Assumption 2.9(ii), respectively). Therefore, under some regularity conditions, the monotonously moving support with a degenerated condition is a sufficient condition for completeness. We summarize this as follows:

Corollary 2.3. *Under Assumption 2.8–2.14, the family $\{K(y, \cdot) \in \mathcal{C}(\mathcal{X}) : y \in \mathcal{Y}\}$ is complete in $\mathcal{C}(\mathcal{X}_{y_2;y_1} \times \{y_1\})$ for each y_1.*

In the rest of the section, we will present a discrete case to illustrate some key aspects of the identification proof. The discrete case refers to the supports $\mathcal{Y}$ and $\mathcal{X}$ being discrete sets:

$y \in \mathcal{Y} \equiv \{y_1, y_1, \ldots, y_{J_1}\}$ and $x \in \mathcal{X} \equiv \{x_1, x_2, \ldots, x_{J_2}\}$.

In Eq. (4), the upper bound $\bar{b}(y)$ is one-to-one and this implies that the number of elements in $\mathcal{Y}$ should be greater than the number of elements in $\mathcal{X}$, i.e. $J_1 \geq J_2$. We start with a simple case, $J_1 = J_2 = J$ and $y_j = x_j$ for $j = 1, \ldots, J$. Consider the following integral equation

$$h(y) = \int_{y_1}^{y} K(y, x)\, g(x)\, dx \text{ for any } y \in \mathcal{Y}. \tag{7}$$

Because dx is a discrete measure concentrated on $\mathcal{X}$, we can write the above integral equation as follows:

$$\begin{aligned}
h(y_1) &= K(y_1, x_1)\, g(x_1) \\
h(y_2) &= K(y_2, x_1)\, g(x_1) + K(y_2, x_2)\, g(x_2) \\
\vdots &= \vdots \\
h(y_J) &= K(y_J, x_1)\, g(x_1) + K(y_J, x_2)\, g(x_2) \\
&\quad + \cdots + K(y_J, x_J)\, g(x_J).
\end{aligned}$$

This can be expressed in terms of matrices as

$$\underbrace{M_h}_{Given} = \underbrace{M_K}_{Given} \cdot \underbrace{M_g}_{Unknown}$$

where

$$M_h = \begin{bmatrix} h(y_1) \\ \vdots \\ h(y_J) \end{bmatrix}_{J\times 1}, \quad M_g = \begin{bmatrix} g(y_1) \\ \vdots \\ g(y_J) \end{bmatrix}_{J\times 1},$$

and

$$M_K = \begin{bmatrix} K(y_1, x_1) & 0 & \cdots & 0 \\ K(y_2, x_1) & K(y_2, x_2) & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ K(y_J, x_1) & K(y_J, x_2) & \cdots & K(y_J, x_J) \end{bmatrix}_{J\times J}.$$

When $L = 1$, Assumption 2.7(ii), suggests that $K(y_j, x_j) \neq 0$ for $j = 1, \ldots, J$, which implies the diagonal elements of the triangular matrix M_K are all nonzero and M_K is invertible. The invertibility permits the identification of the matrix $M_g = M_K^{-1} M_h$.

As for the case, $J_1 > J_2$, the matrix M_K corresponding to $K(y, x)$ is not a square matrix. We can pick J_2 distinct elements from $\mathcal{Y}$ and relabel them as the first J_2 distinct elements such that $K(y_j, x_j) \neq 0$ for $j = 1, \ldots, J_2$. It is straightforward to see that we can write the relationship in terms of matrix decomposition in Eq. (7) with J_2 replacing J. Because $K(y_j, x_j) \neq 0$ for $j = 1, \ldots, J_2$, we also obtain the identification of the matrix M_g.

The discrete case is useful to illustrate the need for an operator kernel that is, in some sense, "nonzero on the diagonal", as required in Assumption 2.7(ii). However, the continuous case is more complex, because smoothness constraints play an important role.

3. Identification of several models

In this section, we will apply Section 2's results to provide simple sufficient conditions for the injectivity that is used for identifying several models, including nonparametric regression models with instrumental variables, nonclassical measurement error models, and auction models. Each of the models we consider is connected to an integral equation or an integral operator relationship relating observable functions to the unobservable functions of interest. Injectivity of the integral operators can help us recover unobservable functions of interest from observable functions. We will give a separate description of each model because their structures differ.

3.1. Nonparametric regression models with instrumental variables

Consider a nonparametric regression model as follows:

$$y = m(x, z_1) + \varepsilon, \text{ with } E(\varepsilon|z) = 0, \quad z = (z_1, z_2) \tag{8}$$

where y is an observable scalar random variable, x is a one-dimensional endogenous regressor and may be correlated with a zero mean disturbance ε, and z_1, z_2 are d_1-dimensional and one-dimensional instrumental variables, respectively.

The conditional expectation of Eq. (8) yields the integral equation

$$E[y|z] = \int m(x, z_1) f(x|z_1, z_2) dx \text{ for all } z \in \mathcal{Z}, \tag{9}$$

where $\mathcal{Z} = \mathcal{Z}_1 \times \mathcal{Z}_2 \subset \mathbb{R}^{d_1+1}$ is the support of z. We can consider the conditional density function $f(\cdot|z)$ in Eq. (9) over a support moving with the instrument variable z_2. Define $\mathcal{X}_{z_2;z_1} \equiv [\underline{b}(z), \bar{b}(z)]$ for each $z \in \mathcal{Z}$, where $\underline{b} : \mathcal{Z} \mapsto \mathcal{X}$, $\bar{b} : \mathcal{Z} \mapsto \mathcal{X}$, and $\mathcal{X} = \cup_{z\in\mathcal{Z}} \mathcal{X}_{z_2;z_1}$. For each $z_1 \in \mathcal{Z}_1$, define an operator as follows:

$$T_{X|Z_2;z_1} : \mathcal{C}(\mathcal{X}_{z_2;z_1} \times \{z_1\}) \to \mathcal{C}^L(\{z_1\} \times \mathcal{Z}_2)$$

$$(T_{X|Z_2;z_1} g)(z) = \int_{\mathcal{X}_{z_2;z_1}} g(x, z_2) f(x|z_1, z_2) dx.$$

From Eq. (9), the observed conditional expectation $E[y|z]$ is in the range of the above integral operator. Injectivity of the operator $T_{X|Z_2;z_1}$ implies that the regression function can be identified as

$$m(x, z_1) = T_{X|Z_2;z_1}^{-1}(E[y|z]) \quad \text{for each } z_1.$$

We can use the conditions in Corollary 2.1 to provide a sufficient condition for the operator's injectivity $T_{X|Z_2;z_1}$ for each z_1.

Let L be a positive integer:

Assumption 3.1.1. The Lth derivative of the conditional mean function $\frac{\partial^L E[y|z_1,z_2]}{\partial z_2}$ exists for each z_1 and is continuous for all $z \in \mathcal{Z}$.

Assumption 3.1.2. (i) Uniform bound of $\mathcal{X}_{z_2;z_1}$, $\underline{B} \leq \underline{b}(z) \leq \bar{b}(z) \leq \bar{B}$ *for all* $z \in \mathcal{Z}$ for some $\underline{B}, \bar{B} \in \mathbb{R}$ and (ii) $\underline{b}(z_1, \underline{z_2}) = \bar{b}(z_1, \underline{z_2}) = \underline{B}$ for each z_1 where $\underline{z_2}$ is the lower endpoint of $\mathcal{Z}_2$.

Assumption 3.1.3. $\underline{b}^{(0,1)}(z_1, z_2)$ and $\bar{b}^{(0,1)}(z_1, z_2)$ exist and are continuous for all $z \in \mathcal{Z}$.

Assumption 3.1.4. The Lth derivative of the conditional density $\frac{\partial^L f(x|z_1,z_2)}{\partial z_2}$ exists for each z_1 and is continuous on $\cup_{z\in\mathcal{Z}} [\underline{b}(z), \bar{b}(z)] \times \{z\}$.

Assumption 3.1.5. At each $z \in \mathcal{Z}$, either (i) $\underline{b}^{(0,1)}(z_1, z_2) = 0$ or (ii) $\left.\frac{\partial^\ell f(x|z_1,z_2)}{\partial z_2}\right|_{x=\underline{b}(z)} = 0$ for $\ell = 0, \ldots, L-1$.

Assumption 3.1.6. $\left.\frac{\partial^L f(x|z_1,z_2)}{\partial z_2}\right|_{x=\underline{b}(z)} = 0$ for all $z \in \mathcal{Z}$.

Assumption 3.1.7. (i) $\bar{b}^{(0,1)}(z_1, z_2) > 0$ for all $z \in \mathcal{Z}$, (ii) $\left.\frac{\partial^{L-1} f(x|z_1,z_2)}{\partial z_2}\right|_{x=\bar{b}(z)} \neq 0$ for all $z \in \mathcal{Z}$ and (iii) if $L > 1$, $\left.\frac{\partial^\ell f(x|z_1,z_2)}{\partial z_2}\right|_{x=\bar{b}(z)} = 0$ for $\ell = 0, \ldots, L-2$ and for all $z \in \mathcal{Z}$.

Theorem 3.1.1. *If Assumption 3.1.1–3.1.7 hold. Then, the integral operator $T_{X|Z_2;z_1}$ is injective for each z_1 and then the regression function $m(\cdot, z_1)$ is identified for each z_1.*

The above result is a direct application of Corollary 2.1; hence, we omit its proof. This result should be contrasted with other popular sufficient conditions for injectivity, namely, (i) the completeness property of the exponential families and normal distributions Newey and Powell (2003) and (ii) the injectivity of convolution operations under assumptions regarding the characteristic function Mattner (1993) or under compact support assumptions (Hu and Ridder, 2010). Our result holds for a nonparametric family of distributions (unlike Newey and Powell (2003)) and holds for operators that are more general than convolutions (unlike Mattner (1993)). Interestingly, our result can be see as a generalization of Hu and Ridder (2010) that also exploits boundary conditions on a compact support to establish injectivity but in a context far more general than merely convolutions. Going beyond convolutions is essential to credibly cover nonparametric instrumental variables (whereas convolutions may be sufficient in many measurement error problems).

Providing novel assumptions that guarantee injectivity, as we do here, is a different problem than showing the impossibility of testing for injectivity against very general alternatives (as in Canay et al. (2013)). The two findings can easily be seen to be compatible: If one took one joint density that satisfies our assumptions (and, thus, has a jump in some derivative at a boundary point),[1] one can simply smear that function with infinitely many times differentiable function whose support has an arbitrarily small width, to obtain a function that violates our assumptions but that can be made arbitrarily close to the original function.

Assumption 3.1.1–3.1.7 can be divided into three parts, properties of the conditional expectation of y given z, the conditional c.d.f. of x given z, and the two endpoints, $\underline{b}(\cdot)$ and $\bar{b}(\cdot)$. Assumption 3.1.1 imposes regularity conditions on the conditional expectation of y given z. By constraining the derivatives of $E(y|z)$ in Eq. (9) this assumption implicitly imposes some continuity restrictions on the regression function $m(\cdot)$. Assumptions 3.1.4 and 3.1.5(ii), 3.1.6 and 3.1.7(ii) &(iii) assume regularity conditions on $f(x|z_1, z_2)$ and vanishing partial derivatives of $f(x|z_1, z_2)$ at the lower endpoints. Assumptions 3.1.2, 3.1.3 and 3.1.5(i), and 3.1.7(i) impose restrictions on the endpoints of the support and some discussion has been provided in Remark 2.3. Here, we provide a bit of intuition for the support condition in practice. If there is a positive correlation between x and z_2 and the variance of x increases with z_2, we would expect the distribution of x to accumulate over z_2 and the support of x would be a function of z_2. Thus, the moving support condition is consistent with a requirement of an instrumental variable, a nonzero correlation between an endogenous variable and an instrumental variable.

The support conditions should be plausible in several empirical applications. In an application to determine the effect of alcohol consumption on the wage rate, we can use the negative value of the local market price of alcohol z_2 as an instrumental variable for individual's alcohol consumption x if there is a negative correlation between alcohol consumption and the local alcohol price, and the variance of individual's alcohol consumption decreases with the local alcohol price. The degeneracy support condition holds, in this case, because there would be no alcohol consumption for extremely high alcohol prices.

The monotonously moving support with a degenerated condition, along with some regularity conditions on the conditional mean function and the conditional density, secures the identification of nonparametric regression models and allows for fairly general functional forms of the regression function in the investigation of the returns to schooling.

3.2. Nonclassical measurement error models with instrumental variable

Consider nonclassical measurement error models with the following conditional density:

$$f_{Y|X^*W}(y|x^*, w)$$

where y is the endogenous variables, x^* is the one-dimensional unobserved error-free explanatory variable, and w is an additional d_1-dimensional vector of observed error-free covariates. Suppose that, in addition to the mismeasured observed variable x, we have access to a one-dimensional instrumental variable z satisfying (1) $f_{Y|XX^*WZ}(y|x, x^*, w, z) = f_{Y|X^*W}(y|x^*, w)$, and (2) $f_{X|X^*WZ}(x|x^*, w, z) = f_{X|X^*W}(x|x^*, w)$.[2] Using the law of the total probability and these assumptions, following the derivation in Hu and Schennach (2008), we can write the observed conditional density $f_{YX|WZ}(y, x|w, z)$ as:

$$f_{YX|WZ}(y, x|z, w) = \int f_{Y|X^*W}(y|x^*, w) f_{X|X^*W}(x|x^*, w) f_{X^*|WZ}(x^*|w, z) dx^*.$$

Identifying this measurement error model then consists in showing that this integral equation has a unique solution $(f_{Y|X^*W}, f_{X|X^*W}, f_{X^*|WZ})$ for a given observable conditional density $f_{YX|WZ}(y, x|w, z)$. To this effect, for each w, define the following integral operators:

$$T_{X|X^*;w} : \mathcal{C}(\mathcal{X}^*_{x;w} \times \{w\}) \to \mathcal{C}^L(\mathcal{X} \times \{w\})$$
$$(T_{X|X^*;w} g)(x) = \int g(x^*, w) f_{X|X^*W}(x|x^*, w) dx^*,$$
$$T_{Z|X;w} : \mathcal{C}(\mathcal{X}_{z;w} \times \{w\}) \to \mathcal{C}^L(\{w\} \times \mathcal{Z})$$
$$(T_{Z|X;w} g)(z) = \int g(x, w) f_{Z|XW}(z|x, w) dx.$$

Hu and Schennach (2008) assume injectivity of the above operators, $T_{X|X^*;w}$ and $T_{Z|X;w}$ along with other location assumptions to obtain uniqueness of eigenvalue–eigenfunction decomposition of an integral operator and provide the identification of $(f_{Y|X^*W}, f_{X|X^*W}, f_{X^*|WZ})$.[3] Denote the changing supports as $\mathcal{X}^*_{x;w} = \left[\underline{b}_1(x, w), \bar{b}_1(x, w)\right]$ and $\mathcal{X}_{z;w} = \left[\underline{b}_2(w, z), \bar{b}_2(w, z)\right]$, where $\underline{b}_1$ and $\bar{b}_1$ are from $\mathcal{X} \times \{w\}$ into $\mathcal{X}^*$ for each w, and $\underline{b}_2$ and $\bar{b}_2$ are from $\{w\} \times \mathcal{Z}$ into $\mathcal{X}$ for each w. It follows that $\mathcal{X}^* = \cup_{(x,w)\in\mathcal{X}\times\mathcal{W}} \mathcal{X}^*_{x;w} = \cup_{(x,w)\in\mathcal{X}\times\mathcal{W}} \left[\underline{b}_1(x, w), \bar{b}_1(x, w)\right]$ and $\mathcal{X} = \cup_{(w,z)\in\mathcal{W}\times\mathcal{Z}} \mathcal{X}_{z;w} = \cup_{(w,z)\in\mathcal{W}\times\mathcal{Z}} \left[\underline{b}_2(w, z), \bar{b}_2(w, z)\right]$.

For an operator T maps $\mathcal{C}(\mathfrak{X})$ into $\mathcal{C}^L(\mathfrak{Y})$, where $\mathfrak{X}$ and $\mathfrak{Y}$ denote subsets of $\mathbb{R}$, the range of T is denoted by Range(T):

$$\text{Range}(T) = \{f \in \mathcal{C}^L(\mathfrak{Y}) : Tg = f \text{ for some } g \in \mathcal{C}(\mathfrak{X})\}.$$

Let L be a positive integer:

[1] The condition that a jump in a derivative at a boundary point can happen at the lowest endpoint $\underline{b}(\underline{y}) = \bar{b}(\underline{y}) = \underline{B}$ (Assumption 2.2. (ii)). This implies that $\underline{b}^{(1)}(\underline{y}) > 0$. Because $\underline{b}(y) \le \bar{b}(y)$ for all $y \in \mathcal{Y}$, we have $\underline{b}^{(1)}(\underline{y}) \le \bar{b}^{(1)}(\underline{y})$. Combining these observations yields $0 < \underline{b}^{(1)}(\underline{y}) \le \bar{b}^{(1)}(\underline{y})$. Thus, the curve describing the two endpoints of the support has a "kink" at the boundary point $\underline{y}$. The curves in Fig. 1 provide some illustration of the situation. Thus, our conditions require a jump in some derivative at the boundary point $\underline{y}$.

[2] The assumption is Assumption 2 in Hu and Schennach (2008). This indicates that the latent true value x^* has already provided enough information for y and x than other observable variables.

[3] As discussed in Hu and Schennach (2008) the injectivity assumptions can also be imposed on the operators $T_{X|X^*;w}$ and $T_{Z|X^*;w}$, where $(T_{Z|X^*;w} g)(z) = \int g(x^*, w) f_{Z|X^*W}(z|x^*, w) dx^*$.

Assumption 3.2.1. For each w, the ranges of the operators $T_{X|X^*;w}$ *and* $T_{Z|X;w}$ contain functions other than the zero function. In other words, Range$(T_{X|X^*;w}) \neq \{0\}$ and Range$(T_{Z|X;w}) \neq \{0\}$ for each w.

Assumption 3.2.2. (i) Uniform bounds of $\mathcal{X}^*_{x;w}$ and $\mathcal{X}_{z;w}$, for some $\underline{B}_i, \bar{B}_i \in \mathbb{R}$, we have $\underline{B}_i \leq \underline{b}_i(\cdot,\cdot) \leq \bar{b}_i(\cdot,\cdot) \leq \bar{B}_i$ for all points in the domains of $\underline{b}_i$ and $\bar{b}_i$ for $i = 1, 2$, and (ii) Suppose that for each w, $\underline{b}_1(\underline{x}, w) = \bar{b}_1(\underline{x}, w) = \underline{B}_1$. $\underline{b}_2(w, \underline{z}) = \bar{b}_2(w, \underline{z}) = \underline{B}_2$, where $\underline{x} = \inf_{x\in\mathcal{X}} x$ and $\underline{z} = \inf_{z\in\mathcal{Z}} z$.

Assumption 3.2.3. For each w, $\underline{b}_1^{(1,0)}(x, w)$, $\bar{b}_1^{(1,0)}(x, w)$ and $\underline{b}_2^{(0,1)}(w, z)$ and $\bar{b}_2^{(0,1)}(w, z)$ exist and are continuous in its domain.

Assumption 3.2.4. The Lth derivatives of the conditional densities $\frac{\partial^L f_{X|X^*W}(x|x^*,w)}{\partial x}$ and $\frac{\partial^L f_{Z|XW}(z|x,w)}{\partial z}$ exist and are continuous on the domains $\cup_{(x,w)\in\mathcal{X}\times\mathcal{W}} \{x\} \times \left[\underline{b}_1(x, w), \bar{b}_1(x, w)\right] \times \{w\}$ and $\cup_{(w,z)\in\mathcal{W}\times\mathcal{Z}} \{z\} \times \left[\underline{b}_2(w, z), \bar{b}_2(w, z)\right] \times \{w\}$, respectively.

Assumption 3.2.5. For each w, either (i) $\underline{b}_1^{(1,0)}(x, w) = 0$ at each $x \in \mathcal{X}$, and $\underline{b}_2^{(0,1)}(w, z) = 0$ at each $z \in \mathcal{Z}$ or (ii) $\left.\frac{\partial^\ell f_{X|X^*W}(x|x^*,w)}{\partial x}\right|_{x^*=\underline{b}_1(x,w)} = 0$ at each $x \in \mathcal{X}$ and $\left.\frac{\partial^\ell f_{Z|XW}(z|x,w)}{\partial z}\right|_{x=\underline{b}_2(w,z)} = 0$ at each $z \in \mathcal{Z}$ for $\ell = 0, \ldots, L-1$.

Assumption 3.2.6. For each w, the Lth derivatives $\left.\frac{\partial^L f_{X|X^*W}(x|x^*,w)}{\partial x}\right|_{x^*=\underline{b}_1(x,w)} = 0$ for all $x \in \mathcal{X}$ and $\left.\frac{\partial^L f_{Z|XW}(z|x,w)}{\partial z}\right|_{x=\underline{b}_2(w,z)} = 0$ for all $z \in \mathcal{Z}$.

Assumption 3.2.7. For each w, *(i)* $\bar{b}_1^{(1,0)}(x, w) > 0$ for all $x \in \mathcal{X}$ and $\bar{b}_2^{(0,1)}(w, z) > 0$ for all $z \in \mathcal{Z}$, (ii) $\left.\frac{\partial^{L-1} f_{X|X^*W}(x|x^*,w)}{\partial x}\right|_{x^*=\bar{b}_1(x,w)} \neq 0$ for all $x \in \mathcal{X}$ and $\left.\frac{\partial^{L-1} f_{Z|XW}(z|x,w)}{\partial z}\right|_{x=\bar{b}_2(w,z)} \neq 0$ for all $z \in \mathcal{Z}$ and (iii) if $L > 1$, $\left.\frac{\partial^\ell f_{X|X^*W}(x|x^*,w)}{\partial x}\right|_{x^*=\bar{b}_1(x,w)} = 0$ for all $x \in \mathcal{X}$ and $\left.\frac{\partial^\ell f_{Z|XW}(z|x,w)}{\partial z}\right|_{x=\bar{b}_2(w,z)} = 0$ for all $z \in \mathcal{Z}$ for $\ell = 0, \ldots, L-2$.

Theorem 3.2.1. *If Assumption 3.2.1–3.2.7 hold. Then, the integral operators $T_{X|X^*;w}$ and $T_{Z|X;w}$ are injective for each w.*

The proof of Theorem 3.2.1 is to apply Corollary 2.1 to each operator of $T_{X|X^*;w}$ and $T_{Z|X;w}$; this is involved but straightforward and we, thus, omit its proof. Assumption 3.2.1 essentially imposes smoothness conditions on functions in the ranges of the operators. Assumptions 3.2.4 and 3.2.5(ii), 3.2.6 and 3.2.7(ii) &(iii) are smoothness and boundary conditions imposed for the densities $f_{X|X^*W}$ and $f_{Z|XW}$. Assumptions 3.2.2, 3.2.3 and 3.2.5(i), and 3.2.7(i) are conditions for the endpoints of the supports, $\underline{b}_1(x, w)$, $\bar{b}_1(x, w)$, $\underline{b}_2(w, z)$, and $\bar{b}_2(w, z)$ for each w. Assumption 3.2.7(i) requires for each w the upper endpoints of the supports of the densities $f_{X|X^*W}(x|\cdot, w)$ and $f_{Z|XW}(z|\cdot, w)$ to increase with x and z, respectively.

As discussed in Hu and Schennach (2008), when $f_{X|X^*W}(x|\cdot, w)$ and $f_{Z|XW}(z|\cdot, w)$ can be written in the forms $f_{X|X^*W}(x|x^*, w) = f_{\varepsilon_1,w}(x - x^*)$ and $f_{Z|XW}(z|x, w) = f_{\varepsilon_2,w}(z - x)$ respectively for each w, then $f_{X|X^*W}(x|\cdot, w)$ and $f_{Z|XW}(z|\cdot, w)$ are injective if and only if the Fourier transform of $f_{\varepsilon_2,w}$ and $f_{\varepsilon_2,w}$ are everywhere nonvanishing.[4]

The moving support conditions for injectivity of the operator $T_{Z|X;w}$ for each w are consistent with the moving support condition in nonparametric regression models with instrumental variables in Section 3.1, where we need injectivity of the operator $T_{X|Z_2;z_1}$ for each z_1. While the support in the nonparametric regression models is $\mathcal{X}_{z_2;z_1} = \left[\underline{b}(z_1, z_2), \bar{b}(z_1, z_2)\right]$ for each z_1, the support in the measurement error models is $\mathcal{X}_{z;w} = \left[\underline{b}_2(w, z), \bar{b}_2(w, z)\right]$ for each w. In discussing the moving support conditions for the nonparametric regression models in Section 3.1, we can provide sufficient conditions to ensure injectivity of the operator $T_{Z|X;w}$ for each w. However, the situation is more complicated in the measurement error models because the assumptions ensure injectivity of two operators not just one.

Similar to the discussion in Section 3.1, if (1) there exists a positive correlation between x and z and the variance of x increases with z and (2) the measurement error is correlated with the true regressor and its variance increases with the true regressor then the moving support conditions in the assumptions are more likely to hold and the integral operators $T_{X|X^*;w}$ and $T_{Z|X;w}$ are injective for each w.

In applying the measurement error model to determine the effect of alcohol consumption on the wage rate, the moving support conditions in the assumptions are more likely to hold if (1) there is a negative correlation between alcohol consumption and the local alcohol price, the variance of individual's alcohol consumption decreases with the local alcohol price, and (2) the measurement error between individual's observed alcohol consumption and true alcohol consumption is correlated with the true alcohol consumption and the variance of the measurement error increases with the true alcohol consumption. The degenerate support conditions in this case are that there is no alcohol consumption for extremely high alcohol prices and when the true alcohol consumption is zero, the observed alcohol consumption is also zero.

3.3. *Auction models with unobserved heterogeneity*

Suppose w is a d_1-dimensional vector of observed state variable which captures auction characteristics. Consider the following simple auction model with unobserved heterogeneity satisfying

$$g(b, w) = \int g_{B|VW}(b|v, w) f(v, w)\, dv \text{ for any } b \in \mathcal{B}$$

where b stands for bids, with observed density $g(b, w)$, v is the unobserved heterogeneity (such as a common value) with unobserved density $f(v, w)$, and $g_{B|VW}(b|v, w)$ is the bid distribution conditional on the unobserved heterogeneity. The model is closely related to the models in An et al. (2010); Hu et al. (2013). Let $D = D(\mathcal{B})$ be the discretized bid for some discretization $D(\cdot)$. A key assumption for the identification of $g_{B|VW}$ and f is the injectivity of the linear operator $T_{B|V;w}$ for each w, where

$$(T_{B|V;w} f)(b, w) = \int g_{B|VW}(b|v, w) f(v, w)\, dv$$

where the unobserved heterogeneity V is discrete, we can choose some discretization $D(\cdot)$ such that the linear operator $T_{B|V;w}$ is expressed in terms of a square matrix as $T_{B|V;w} = [g_{B|VW}(b|v, w)]$ for each w and our changing support assumption implies that the matrix $T_{B|V;w}$ is a lower or upper triangular form for each w. When the probability on the upper (or lower) boundary of the support is nonzero, i.e., the diagonal elements of the triangular matrix are nonzero, the moving support assumption implies the completeness condition.

In the continuous case where the unobserved heterogeneity is continuous, the result remains with matrices replaced by integral operators. In this case, a triangular matrix is extended to an integral operator on a function space, e.g.,

$$(T_{B|V;w} f)(b) = \int_{\mathcal{V}_{b;w}} g_{B|VW}(b|v, w) f(v, w)\, dv,$$

where $\mathcal{V}_{b;w} = \left[\underline{b}(b, w), \bar{b}(b, w)\right]$.

[4] The result is an application of Theorem 2.1 in Mattner (1993).

Corollary 2.2 shows such integral operator is injective in a continuous space if for each w, the kernel function $g_{B|VW}(b|v, w)$ is nonzero on the "diagonal", and the changing support $\mathcal{V}_{b;w}$ satisfies some regularity conditions. Therefore, the changing support assumption can be used to imply that the operator corresponding to the conditional bid distribution $g_{B|VW}$ is injective and the distribution of the unobserved heterogeneity is identified. We will state the conditions in Corollary 2.2 for these models.

Let L be some positive integer:

Assumption 3.3.1. The Lth derivative of the density of bids $\frac{\partial^L g(b,w)}{\partial b}$ exists and is continuous for all $b \in \mathcal{B}$.

Assumption 3.3.2. (i) Uniform bound of $\mathcal{V}_{b;w}$, $\underline{B} \le \underline{b}(b, w) \le \bar{b}(b, w) \le \bar{B}$ for all $b \in \mathcal{B}$ for some $\underline{B}, \bar{B} \in \mathbb{R}$ and (ii) for each w $\underline{b}\left(\underline{b}_l, w\right) = \bar{b}\left(\underline{b}_l, w\right) = \underline{B}$, where $\underline{b}_l$ is the lower endpoint of $\mathcal{B}$.

Assumption 3.3.3. $\underline{b}^{(1,0)}(b, w)$ and $\bar{b}^{(1,0)}(b, w)$ exist and are continuous for all $(b, w) \in \mathcal{B} \times \mathcal{W}$.

Assumption 3.3.4. The Lth derivative of the conditional density $\frac{\partial^L g_{B|VW}(b|v,w)}{\partial b}$ exists and is continuous on $\cup_{(b,w)\in\mathcal{B}\times\mathcal{W}} \{b\} \times \left[\underline{b}(b, w), \bar{b}(b, w)\right] \times \{w\}$.

Assumption 3.3.5. At each $(b, w) \in \mathcal{B}\times\mathcal{W}$, either (i) $\underline{b}^{(1,0)}(b, w) = 0$ or (ii) the derivatives $\left.\frac{\partial^\ell g_{B|VW}(b|v,w)}{\partial b}\right|_{v=\underline{b}(b,w)} = 0$ for $\ell = 0, \dots, L - 1$.

Assumption 3.3.6. $\left.\frac{\partial^L g_{B|VW}(b|v,w)}{\partial b}\right|_{v=\underline{b}(b,w)} = 0$ for all $(b, w) \in \mathcal{B}\times\mathcal{W}$.

Assumption 3.3.7. (i) $\bar{b}^{(1,0)}(b, w) > 0$ for all $(b, w) \in \mathcal{B} \times \mathcal{W}$, (ii) $\left.\frac{\partial^{L-1} g_{B|VW}(b|v,w)}{\partial b}\right|_{v=\bar{b}(b,w)} \ne 0$ for all $(b, w) \in \mathcal{B} \times \mathcal{W}$ and (iii) if $L > 1$, $\left.\frac{\partial^\ell g_{B|VW}(b|v,w)}{\partial b}\right|_{v=\bar{b}(b,w)} = 0$ for $\ell = 0, \dots, L - 2$ and for all $(b, w) \in \mathcal{B} \times \mathcal{W}$.

Theorem 3.3.1. *Assume the conditional bid distribution $g_{B|VW}$ exists. If Assumption 3.3.1– 3.3.7 hold, the integral operator $T_{B|V;w}$ is injective for each w and the distribution of the unobserved heterogeneity $f(v, w)$ is identified for each w.*[5]

In this model, the moving support conditions of Theorem 3.3.1 require that the support of valuations changes with the bids. The moving support assumption has a strong and natural economic basis in this example, as it asserts that well-trained bidders can effectively rule out an exceptionally large or small deviation of individual-specific valuations from a common value. This should be contrasted with Fourier-based approaches (e.g., Li et al. (2000)), where injectivity is established based on the condition non-vanishing characteristic functions, which has no clear connection to the economic model.

Assumption 3.3.1 imposes smoothness conditions on the distribution of bids. Assumptions 3.3.4 and 3.3.5(ii), 3.3.6 and 3.3.7(ii) &(iii) collect all the regularity conditions for the bid distribution conditional on the unobserved heterogeneity. Assumptions 3.3.2, 3.3.3 and 3.3.5(i), and 3.3.7(i) are the moving support conditions. D'Haultfoeuille and Février (2010) use at least three repeated conditionally independent measurements and some moving support assumptions to identify an auction model with unobserved heterogeneity nonparametrically. The moving support assumption requires that the upper endpoints of the supports increase with the true value of the unobserved variable, which is close to Assumption 3.3.7(i). Because no bidder would rationally bid higher than their own valuation, the values of the bids should be lower than the valuation and this implies that a zero valuation induces a zero value of the bids. The degenerate support condition is, thus, plausible in an auction model.

[5] An et al. (2010), and Hu et al. (2013) have shown the identification of the conditional bid distribution $g_{B|VW}$ from observable information.

4. Conclusion

This paper presents sufficient conditions for injectivity of integral operators with compactly supported kernels. The injectivity is related to completeness, which is used increasingly in various nonparametric models to obtain identification including nonparametric regression models with instrumental variables, nonclassical measurement error models, and auction models, etc. Our results show that if compactly supported kernels satisfy certain regularity conditions, monotonously moving support assumptions and degeneracy of the support assumptions then its corresponding integral operator is injective or, equivalently, its kernels are complete.

Two possible extensions are left for future research. First, one may wonder whether the similar assumption can be made for kernels with infinite support because a compact support assumption may not hold in some applications. Second, the random variables in this paper are restricted to be scalars. An extension to vectors is natural but more involved.

Appendix. Proofs

A.1. Existing results on Volterra equations

For easy reference, we first summarize existing results on Volterra equations from Bôcher (1909). We define a function $u(x)$ on

$\mathbb{I} = [a, b]$

and a function of two variables $K(x, \xi)$ on the triangle

$\mathbb{T} = \{(x, \xi) : a \le \xi \le x \le b\}$.

The Volterra equations of the first kind and of the second kind are respectively defined as

$$f(x) = \int_a^x K(x, \xi) u(\xi)\, d\xi \tag{10}$$

$$u(x) = f(x) + \int_a^x K(x, \xi) u(\xi)\, d\xi \tag{11}$$

where $f(x)$ and $K(x, \xi)$ are known and $u(x)$ is usually the function of interest.

Definition A.1.1. The discontinuities of a function of (x, ξ) are said to be regularly distributed if they all lie on a finite number of curves with continuously turning tangents no one of which is met by a line parallel to the axis of x or of ξ in more than a finite number of points.

The uniqueness of the solution of Eq. (11) is summarized in Theorem 1 on p. 16 of Bôcher (1909) as follows:

Theorem A.1.1. *If $K(x, \xi)$ is uniformly bounded on $\mathbb{T}$ and its discontinuities, if it has any, are regularly distributed, a necessary and sufficient condition that* Eq. (11) *has a solution continuous throughout $\mathbb{I}$ is that $f(x)$ be continuous throughout $\mathbb{I}$, and if this condition is fulfilled, Eq. (11) has only one continuous solution, which is given by the absolutely and uniformly convergent series in* Eq. (13).

Proof. First, we provide a proof of the sufficient condition. Successively substituting for $u(\cdot)$ in Eq. (11) yields that

$$\begin{aligned} u(x) &= f(x) + \int_a^x K(x,\xi) f(\xi) d\xi \\ &\quad + \int_a^x K(x,\xi) \int_a^\xi K(\xi,\xi_1) u(\xi_1)\, d\xi_1 d\xi \\ &= \cdots \\ &= S_n(x) + R_n(x) \end{aligned} \tag{12}$$

where

$$\begin{aligned} S_n(x) &= f(x) + \int_a^x K(x,\xi) f(\xi) d\xi \\ &\quad + \int_a^x K(x,\xi) \int_a^\xi K(\xi,\xi_1) f(\xi_1)\, d\xi_1 d\xi \\ &\quad + \cdots + \int_a^x K(x,\xi) \int_a^\xi K(\xi,\xi_1) \cdots \\ &\quad \times \int_a^{\xi_{n-1}} K(\xi_{n-1},\xi_n) f(\xi_n)\, d\xi_n \cdots d\xi_1 d\xi, \end{aligned}$$

$$\begin{aligned} R_n(x) &= \int_a^x K(x,\xi) \int_a^\xi K(\xi,\xi_1) \cdots \\ &\quad \times \int_a^{\xi_{n-1}} K(\xi_{n-1},\xi_n) u(\xi_n)\, d\xi_n \cdots d\xi_1 d\xi. \end{aligned}$$

Because $K(x,\xi)$ is uniformly bounded on $\mathbb{T}$ and $f(x)$ is continuous throughout $\mathbb{I}$, we have

$$|K(x,\xi)| \le M \text{ and } |f(x)| \le N.$$

Define

$$\begin{aligned} F_n(x) &= \int_a^x K(x,\xi) \int_a^\xi K(\xi,\xi_1) \cdots \\ &\quad \times \int_a^{\xi_{n-1}} K(\xi_{n-1},\xi_n) f(\xi_n)\, d\xi_n \cdots d\xi_1 d\xi. \end{aligned}$$

Then, we have $S_n(x) = S_{n-1}(x) + F_n(x)$, and

$$\begin{aligned} |F_n(x)| &\le NM^{n+1} \int_a^x \int_a^\xi \cdots \int_a^{\xi_{n-1}} d\xi_n \cdots d\xi_1 d\xi \\ &= NM^{n+1} \frac{(x-a)^n}{(n+1)!} \le NM^{n+1} \frac{(b-a)^n}{(n+1)!}. \end{aligned}$$

This implies that $S_n(x)$ converges absolutely and uniformly over I. The continuity of $u(x)$ over I implies that $|u(x)| \le N' < \infty$. We then have

$$|R_n(x)| \le N'M^{n+1} \frac{(x-a)^n}{(n+1)!} \le N'M^{n+1} \frac{(b-a)^n}{(n+1)!}$$

and therefore, $R_n(x)$ converges absolutely and uniformly over I with

$$\lim_{n\to\infty} R_n(x) = 0.$$

Thus, we have an infinite series which converges absolutely and uniformly over I and

$$\begin{aligned} s(x) &\equiv \lim_{n\to\infty} S_n(x) \\ &= f(x) + \int_a^x K(x,\xi) f(\xi) d\xi \\ &\quad + \int_a^x K(x,\xi) \int_a^\xi K(\xi,\xi_1) f(\xi_1)\, d\xi_1 d\xi + \cdots. \end{aligned} \tag{13}$$

It is straightforward to check that $s(\cdot)$ is a solution of Eq. (11) by

$$\begin{aligned} \int_a^x K(x,\xi) s(\xi)\, d\xi &= \int_a^x K(x,\xi) f(\xi) d\xi \\ &\quad + \int_a^x K(x,\xi) \int_a^\xi K(\xi,\xi_1) f(\xi_1)\, d\xi_1 d\xi + \cdots \\ &= s(x) - f(x). \end{aligned}$$

Because the finite series $S_n(x)$ is continuous and $S_n(x)$ converges to $S(x)$ absolutely and uniformly over I, $S(x)$ is continuous throughout $\mathbb{I}$.[6] On the other hand, any continuous solution of Eq. (11) must satisfy $u(x) = S_n(x) + R_n(x)$ in Eq. (12). This implies that there exists only one continuous solution given by

$$u(x) = \lim_{n\to\infty} S_n(x) = s(x).$$

As for a proof for the necessary condition, define $H(x) = \int_a^x K(x,\xi) u(\xi)\, d\xi$. Because the discontinuities of $K(x,\xi)$ are regularly distributed and $u(\cdot)$ is continuous, the function $H(x)$ is continuous throughout $\mathbb{I}$.[7] Then, by the continuity of $u(\cdot)$ and $H(\cdot)$, given $\varepsilon > 0$, there exists $\delta > 0$ such that $|x - x'| < \delta$ and

$$\begin{aligned} |f(x) - f(x')| &\le |u(x) - u(x')| \\ &\quad + \left| \int_a^x K(x,\xi) u(\xi)\, d\xi - \int_a^{x'} K(x',\xi) u(\xi)\, d\xi \right| \\ &= |u(x) - u(x')| + |H(x) - H(x')| < \varepsilon. \end{aligned}$$

This shows that $f(x)$ must be continuous throughout $\mathbb{I}$. □

We may go through the same steps again to get

$$\begin{aligned} F_n(x) &\equiv \int_{A(x)}^x K(x,\xi) \int_{A(\xi)}^\xi K(\xi,\xi_1) \cdots \\ &\quad \times \int_{A(\xi_{n-1})}^{\xi_{n-1}} K(\xi_{n-1},\xi_n) f(\xi_n)\, d\xi_n \cdots d\xi_1 d\xi. \end{aligned}$$

This term is also bounded by

$$\begin{aligned} |F_n(x)| &\le NM^{n+1} \int_{A(x)}^x \int_{A(\xi)}^\xi \cdots \int_{A(\xi_{n-1})}^{\xi_{n-1}} d\xi_n \cdots d\xi_1 d\xi \\ &\le NM^{n+1} \int_a^x \int_a^\xi \cdots \int_a^{\xi_{n-1}} d\xi_n \cdots d\xi_1 d\xi \\ &= NM^{n+1} \frac{(x-a)^n}{(n+1)!}. \end{aligned}$$

In the similar manner as the proof of Theorem A.1.1, the unique solution exists and is given by an absolutely and uniform convergent series.

Eq. (10) leads to the form of Eq. (11) by taking the first derivative as follows:

$$f^{(1)}(x) = K(x,x)u(x) + \int_a^x K^{(1,0)}(x,\xi) u(\xi)\, d\xi. \tag{14}$$

Therefore, the results for Eq. (11) also apply. Bôcher (1909) further shows, in his Theorem 1*, the following result.

[6] Because the discontinuities of $K(x,\xi)$ are regularly distributed and $f(\cdot)$ is continuous, the function $F_n(x)$ is continuous throughout $\mathbb{I}$ for each n. The result is in a corollary on p. 3 of Bôcher (1909) and the statement is the following: *If $\phi(x,y)$ and $\psi(x,y)$ are finite in $\mathbb{T}$ and their discontinuities, if they have any, are regularly distributed, the function*

$$H(x,y) = \int_y^x \phi(x,\xi) \psi(\xi,y)\, d\xi$$

is continuous throughout $\mathbb{T}$.

[7] The result adopted here is the same as the previous footnote and is in a corollary on p. 3 of Bôcher (1909).

Theorem A.1.2. *If $K(x,\xi)$ is continuous in $\mathbb{T}$ and has a derivative $K^{(1,0)}(x,\xi)$ finite in $\mathbb{T}$ and whose discontinuities are regularly distributed, and if $K(x,x)$ does not vanish at any point of $\mathbb{I}$, a necessary and sufficient condition that* Eq. (10) *has a continuous solution is that $f(x)$ and its derivative $f'(x)$ be continuous in $\mathbb{I}$ and $f(a)=0$. If these conditions are fulfilled, Eq. (10) has only one continuous solution, namely the continuous solution of Eq. ((14)).*

This proof implies that we may allow both boundaries to vary as long as they are bounded. Consider a varying lower bound as

$$u(x)=f(x)+\int_{A(x)}^{x}K(x,\xi)u(\xi)\,d\xi,$$

where the key is

$$a\le A(x)<x.$$

This condition guarantees there is a triangular operator. For example, a lower triangular matrix can have zero near the lower left corner, which does not affect the invertibility.

A.2. Proof of Theorem 2.1

The proof proceeds by first differentiating (4) until it can be cast into the form of Volterra Equation of the second kind for which Theorem A.1.1 applies.

Step 1: Differentiating (4) with respect to y using the Leibniz integral rule yields:

$$\begin{aligned}h^{(1)}(y)&=\bar b^{(1)}(y)K\left(y,\bar b(y)\right)g\left(\bar b(y)\right)-\underline{b}^{(1)}(y)K\left(y,\underline{b}(y)\right)g\left(\underline{b}(y)\right)\\&\quad+\int_{\underline{b}(y)}^{\bar b(y)}K^{(1,0)}(y,x)g(x)\,dx \qquad (15)\\&=\bar b^{(1)}(y)K\left(y,\bar b(y)\right)g\left(\bar b(y)\right)+\int_{\underline{b}(y)}^{\bar b(y)}K^{(1,0)}(y,x)g(x)\,dx\end{aligned}$$

where the second term in (15) vanishes by Assumption 2.5(i) or (ii). For the $L=1$ case, we stop at this stage and go to step 2 below.

If $L>1$, the first term of (15) also vanishes by Assumption 2.7(iii) and we obtain

$$h^{(1)}(y)=\int_{\underline{b}(y)}^{\bar b(y)}K^{(1,0)}(y,x)g(x)\,dx.$$

Repeating the differentiation process yields:

$$\begin{aligned}h^{(2)}(y)&=\bar b^{(1)}(y)K^{(1,0)}\left(y,\bar b(y)\right)g\left(\bar b(y)\right)\\&\quad-\underline{b}^{(1)}(y)K^{(1,0)}\left(y,\underline{b}(y)\right)g\left(\underline{b}(y)\right)\\&\quad+\int_{\underline{b}(y)}^{\bar b(y)}K^{(2,0)}(y,x)g(x)\,dx\\&=\bar b^{(1)}(y)K^{(1,0)}\left(y,\bar b(y)\right)g\left(\bar b(y)\right)\\&\quad+\int_{\underline{b}(y)}^{\bar b(y)}K^{(2,0)}(y,x)g(x)\,dx\end{aligned}$$

where the second term again vanishes by Assumption 2.5(i) or (ii).

Step 2: In general, repeating the differentiation process L times (for $L\ge1$) thus yields:

$$\begin{aligned}h^{(L)}(y)&=\bar b^{(1)}(y)K^{(L-1,0)}\left(y,\bar b(y)\right)g\left(\bar b(y)\right)\\&\quad+\int_{\underline{b}(y)}^{\bar b(y)}K^{(L,0)}(y,x)g(x)\,dx.\end{aligned}$$

Step 3: Making the change of variable $x=\bar b(u)$ in the integral yields:

$$\begin{aligned}h^{(L)}(y)&=\bar b^{(1)}(y)K^{(L-1,0)}\left(y,\bar b(y)\right)g\left(\bar b(y)\right)\\&\quad+\int_{\bar b^{-1}(\underline{b}(y))}^{y}K^{(L,0)}\left(y,\bar b(u)\right)g\left(\bar b(u)\right)\bar b^{(1)}(u)\,du\\&=\bar b^{(1)}(y)K^{(L-1,0)}\left(y,\bar b(y)\right)\tilde g(y)\\&\quad+\int_{\bar b^{-1}(\underline{b}(y))}^{y}K^{(L,0)}\left(y,\bar b(u)\right)\bar b^{(1)}(u)\tilde g(u)\,du\end{aligned}$$

where the inverse of $\bar b$ exists and is one-to-one by Assumptions 2.3 and 2.7(i) and where we have introduced the function $\tilde g(u)\equiv g\left(\bar b(u)\right)$. Rearranging the equation yields:

$$\tilde g(y)=\tilde h(y)-\int_{\underline{y}}^{y}\tilde K(y,u)\tilde g(u)\,du \qquad (16)$$

where

$$\tilde h(y)=\frac{h^{(L)}(y)}{\bar b^{(1)}(y)K^{(L-1,0)}\left(y,\bar b(y)\right)}$$

$$\tilde K(y,u)=\begin{cases}\dfrac{K^{(L,0)}\left(y,\bar b(u)\right)\bar b^{(1)}(u)}{\bar b^{(1)}(y)K^{(L-1,0)}\left(y,\bar b(y)\right)} & \text{if } u\ge\bar b^{-1}\left(\underline{b}(y)\right)\\ 0 & \text{if } u\in\left[\underline{y},\bar b^{-1}\left(\underline{b}(y)\right)\right)\end{cases}.$$

The constant lower integration bound in (16) has been obtained by replacing $\bar b^{-1}\left(\underline{b}(y)\right)$ by its minimum value (over $y\in\mathcal{Y}$) and by padding $\tilde K(y,u)$ with zeros accordingly. This minimum is found through $\bar b^{-1}\left(\underline{b}(y)\right)\ge\bar b^{-1}\left(\underline{b}\left(\underline{y}\right)\right)=\bar b^{-1}\left(\bar b\left(\underline{y}\right)\right)=\underline{y}$, where we have used the monotonicity of $\bar b(\cdot)$ (and thus $\bar b^{-1}(\cdot)$) by Assumption 2.7(i), the fact that $\underline{b}(y)\ge\underline{B}=\underline{b}\left(\underline{y}\right)$ and that $\underline{b}\left(\underline{y}\right)=\bar b\left(\underline{y}\right)$ by Assumption 2.2.

Step 4: Eq. (16) has the form of a Volterra Integral Equation of the second kind. Moreover, $\tilde h(y)$ is continuous because so are $h^{(L)}(y)$, $\bar b^{(1)}(y)$ and $K^{(L-1,0)}\left(y,\bar b(y)\right)$ and since $\bar b^{(1)}(y)$ and $K^{(L-1,0)}\left(y,\bar b(y)\right)$ are nonvanishing by Assumption 2.7. Similarly, $\tilde K(y,u)$ is continuous because so are $K^{(L,0)}\left(y,\bar b(u)\right)$, $\bar b^{(1)}(\cdot)$ and $K^{(L-1,0)}\left(y,\bar b(y)\right)$ (with $\bar b^{(1)}(\cdot)$ and $K^{(L-1,0)}\left(y,\bar b(y)\right)$ nonvanishing) and because, at the junction point $u=\bar b^{-1}\left(\underline{b}(y)\right)$, we have $K^{(L,0)}\left(y,\bar b(u)\right)=K^{(L,0)}\left(y,\bar b\left(\bar b^{-1}\left(\underline{b}(y)\right)\right)\right)=K^{(L,0)}\left(y,\underline{b}(y)\right)=0$ by Assumption 2.6.

It follows by Theorem A.1.1 that the solution $\tilde g(u)$ to Eq. (16) is unique. The solution $g(x)$ can then be recovered from $g(x)=\tilde g\left(\bar b^{-1}(x)\right)$, since $\bar b(y)$ is one-to-one by Assumptions 2.3 and 2.7(i). □

References

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

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

An, Y., Hu, Y., Shum, M., 2010. Estimating first-price auctions with an unknown number of bidders: a misclassification approach. J. Econometrics 157 (2), 328–341.

Andrews, D., 2011. Examples of L^2-complete and boundedly-complete distributions. Cowles Foundation for Research in Economics.

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

Bôcher, M., 1909. An introduction to the study of integral equations. Cambridge University Press.

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

Carroll, R., Chen, X., Hu, Y., 2010. Identification and estimation of nonlinear models using two samples with nonclassical measurement errors. J. Nonparam. Stat. (ISSN: 1048-5252) 22 (4), 379–399.

Chen, X., Hu, Y., 2006. Identification and inference of nonlinear models using two samples with arbitrary measurement errors. Cowles Foundation Discussion Paper No. 1590.

Chernozhukov, V., Hansen, C., 2005. An IV model of quantile treatment effects. Econometrica (ISSN: 1468-0262) 73 (1), 245–261.

Chernozhukov, V., Imbens, G., Newey, W., 2007. Instrumental variable estimation of nonseparable models. J. Econometrics (ISSN: 0304-4076) 139 (1), 4–14.

Darolles, S., Fan, Y., Florens, J., Renault, E., 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.

D'Haultfoeuille, X., Février, P., (2010) 2010. Identification of Mixture Models Using Support Variations. Working Paper.

Hall, P., Horowitz, J., 2005. Nonparametric methods for inference in the presence of instrumental variables. Ann. Statist. (ISSN: 0090-5364) 33 (6), 2904–2929.

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

Horowitz, J., Lee, S., 2007. Nonparametric instrumental variables estimation of a quantile regression model. Econometrica (ISSN: 1468-0262) 75 (4), 1191–1208.

Hu, Y., McAdams, D., Shum, M., 2013. Identification of first-price auctions with non-separable unobserved heterogeneity. J. Econometrics 174 (2), 186–193.

Hu, Y., Ridder, G., 2010. On deconvolution as a first stage nonparametric estimator. Econometric Rev. 29 (4), 1–32.

Hu, Y., Schennach, S., 2008. Instrumental variable treatment of nonclassical measurement error models. Econometrica 76 (1), 195–216.

Hu, Y., Shiu, J., 2016. Nonparametric Identification Using Instrumental Variables: Sufficient Conditions for Completeness. Working paper.

Hu, Y., Shum, M., 2012. Nonparametric identification of dynamic models with unobserved state variables. J. Econometrics 171 (1), 32–44.

Li, T., Perrigne, I., Vuong, Q, 2000. Conditionally independent private information in OCS wildcat auctions. J. Econometrics 98 (1), 129.

Mattner, L., 1993. Some incomplete but boundedly complete location families. Ann. Statist. 21 (4), 2158–2162.

Newey, W., Powell, J., 2003. Instrumental variable estimation of nonparametric models. Econometrica 71 (5), 1565–1578.

Shiu, J., Hu, Y., 2013. Identification and estimation of nonlinear dynamic panel data models with unobserved covariates. J. Econometrics (ISSN: 0304-4076) 175 (2), 116–131.