INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES
Série des Documents de Travail du CREST
(Centre de Recherche en Economie et Statistique)

n° 2006-32

On the Completeness Condition in Nonparametric Instrumental Problems

X. D'HAULTFOEUILLE [1]

[1] ENSAE, CREST – INSEE and University Paris I, Panthéon Sorbonne.
Email : xavier.d'haultfoeuille@ensae.fr

On the Completeness Condition in Nonparametric Instrumental Problems

Xavier d'Haultfoeuille*
ENSAE, CREST - INSEE and
Université de Paris I, Panthéon-Sorbonne

December 2006

Abstract

The notion of completeness between two random variables has been considered recently to provide identification in nonparametric instrumental problems. This condition is quite abstract, however, and characterizations have been obtained only in special cases. The aim of this paper is to provide general sufficient conditions to achieve completeness or bounded completeness. The difference between these two notions is stressed, and implications for the nonparametric instrumental regression are given.

Keywords: completeness, nonparametric identification, instrumental variables.

Résumé

La notion de complétude entre variables aléatoires a été considérée récemment pour obtenir l'identification dans les problèmes non-paramétriques instrumentaux. Cette condition est assez abstraite, cependant, et des caractérisations n'ont été obtenues que dans quelques cas particuliers. L'objectif de ce papier est de présenter des conditions suffisantes générales permettant d'obtenir la complétude ou la « complétude bornée ». La différence entre ces deux notions est soulignée, et les conséquences de ces résultats sur la régression non-paramétrique instrumentale sont examinées.

Mots-clés : complétude, identification non-paramétrique, variables instrumentales.

*ENSAE, 18 boulevard Adolphe Pinard, 75 675 Paris Cedex, France. E-mail: xavier.d'haultfoeuille@ensae.fr. I am very grateful to Philippe Février, Jean-Pierre Florens, Xavier Mary and Jean-Marc Robin for helpful discussions and comments.

1 Introduction

Let X and Z denote two random variables. Analogously to the completeness of a statistic for a parameter, X is said to be complete for Z if, for all measurable real functions h such as $\mathbb{E}[|h(X)|] < +\infty$,

$$\Big(\mathbb{E}[h(X)|Z] = 0 \quad \text{a.s.}\Big) \Longrightarrow \Big(h(X) = 0 \quad \text{a.s.}\Big). \tag{1.1}$$

X will be bounded complete if the same holds for any bounded h. Note that completeness is equivalent to the injectivity of the conditional expectation operator. Thus, not surprisingly, it has appeared to be a key identifying condition in nonparametric instrumental problems. Darolles et al. (2002) and Newey and Powell (2003) used it in the study of nonparametric instrumental regression, Florens et al. (2003) in the theory of local instrumental variables, Blundell et al. (2003) in the estimation of Engle curves with endogenous total expenditures, and Hu and Schennach (2006) in the case of nonclassical measurement error problems.[1]

This dependence condition is quite abstract though, and a characterization or at least sufficient conditions on the joint distribution of (X, Z) are desirable. Newey and Powell (2003) address the finite support and exponential families cases, Darolles et al. (2002) the normal distributions and Blundell et al. (2003) focus on the analysis of a particular semiparametric model. However, results are still lacking in the general case to properly define completeness in terms of dependence between the two variables. The aim of this paper is to go one step in this direction by considering a nonparametric model on (X, Z). In this framework, bounded completeness is implied by a large support condition and weak technical assumptions on the distribution of the error of the model. Completeness, on the other hand, requires stringent conditions on this distribution. This difference between completeness and bounded completeness is in line with previous results of the statistical literature, see e.g. Lehmann (1986, p. 173), Hoeffding (1977) and Mattner (1993) (see also Blundell et al., 2003, for a discussion on this point).

[1]Indeed, their assumption 2 is equivalent under technical conditions to a completeness condition.

Besides, the paper shows that despite the seemingly asymmetry between X and Z, their roles can be, up to a certain extent, exchanged in our nonparametric model. This result can be useful when the causality between X and Z is reversed, as it happens for instance in the measurement error model of Hu and Schennach (2006).

Implications for the nonparametric instrumental regression are also examined. We consider a nonparametric system which can be seen as generalization of the linear model with instruments. In this setting, the rank condition can be formulated roughly as a large support assumption. If one is reluctant to impose additional structure on the errors, only local nonparametric identification will be achieved in general. However, global identification can be recovered under the extra assumption that the derivative of the regression function is bounded.

The paper is organized as follows. The main results are given in section two. Section three examines the consequence of these results on the identification of nonparametric instrumental regression. The proofs are deferred to section four.

2 Main results

In the sequel, X and Z belong to $\mathbb{R}^p$ and $\mathbb{R}^q$ respectively, with $p \leq q$. X and Z may share elements in common, and we let W denote these common elements, $W \in \mathbb{R}^r$. For instance, in an instrumental nonparametric regression (see e.g. Newey and Powell, 2003), W corresponds to the exogenous components of X. The remaining elements of X and Z are called respectively X_0 and Z_0, so that $X = (X_0, W)$ and $Z = (Z_0, W)$. In this framework, we will say that X is complete (resp. bounded complete) for W if (1.1) holds for all h such as, for $\mathbb{P}^W-$ almost all w, $h(., w)$ is integrable with respect to $\mathbb{P}^{X_0}$ (resp. bounded). In the following, we first recall the existing result when X_0 has a finite support, and then we address the infinite support case.

2.1 Finite support case

When X_0 and Z_0 have finite supports, which are denoted respectively $(x_1, ..., x_s)$ and $(z_1, ..., z_t)$, completeness can be easily characterized.

Proposition 2.1 *(Newey and Powell, 2003) X is complete for Z if and only if*

$$\mathbb{P}(rank(P(W)) = s) = 1,$$

where $P(W)$ is the matrix of typical elements $\mathbb{P}(X_0 = x_i | Z_0 = z_j, W)$.

In the finite support case, completeness has two implications. First the support of Z_0 should be at least as rich as the one of X_0 ($t \geq s$). Second, the dependence between X_0 and Z_0 should be strong enough for s distinct conditional distributions $P(X_0 = .|Z_0 = z_j, W)$ to exist. Note also that in this case, completeness and bounded completeness are equivalent.

2.2 Infinite support case

The situation is more involved when the support of X_0 is infinite. By a direct adaptation of classical statistical results, completeness can be obtained when the conditional distribution of X belongs to an exponential family. Unfortunately, condition (2.1) is unnatural in terms of dependence between X and Z, except for special cases such as normal variables.

Proposition 2.2 *(Newey and Powell, 2003) Suppose that with probability one, the distribution of X_0 conditionally on Z is absolutely continuous with respect to the Lebesgue measure with density*

$$f_{X_0|Z_0,W}(x, z, w) = s(x, w)t(z, w) \exp(\mu(z, w)'\tau(x, w)), \tag{2.1}$$

where, almost surely, $\tau(., W)$ is one-to-one and the support of $\mu(Z)$ given W is an open set. Then X is complete for Z.

We now explore a more general situation by considering a nonparametric link between X and Z. More precisely, suppose that there exists maps μ_1 and ν_1, from respectively $\mathbb{R}^p$ and $\mathbb{R}^q$ to $\mathbb{R}^{p-r}$, such that

$$\mu_1(X) = \nu_1(Z) + \varepsilon_1. \tag{2.2}$$

Furthermore, we consider the following assumptions.

A1. For $\mathbb{P}^W$-almost all w, $\mu_1(., w)$ is one-to-one.

A2. $Z_0 \perp\!\!\!\perp \varepsilon_1 \,|\, W$.

A3. For $\mathbb{P}^W$-almost all w, the measure of $\nu_1(Z_0, w)$ is continuous with respect to the Lebesgue measure and its support is $\mathbb{R}^{p-r}$.

A4. For $\mathbb{P}^W$-almost all w, ε_1 admits a density $f_{\varepsilon_1|W}(., w)$.

When X_0 is real, A1 is satisfied if for instance $\mu_1(X_0, W) = g(X_0)$ for any monotonous function g. A2 states that $\mu_1(X)$ can be decomposed into two (conditionally) independent parts. Because mean-independence can always be achieved by setting $\nu_1(Z) = \mathbb{E}(\mu_1(X)|Z)$,[2] A2 means in fact that there exists μ_1 such that mean-independence can be strengthen into independence.

A3 is a continuity and large support condition. Note that it may hold without the distribution of Z being continuous. Only one continuous component is required. The large support condition is restrictive but widespread in the literature (see e.g. Manski (1988) or Lewbel (2000)). Moreover, only $\nu_1(Z)$, not necessarily Z, should satisfy this condition. This means that $p - r$ regressors with large support may be sufficient. Lastly, A4 restricts the analysis to the case of continuous residuals.

To achieve completeness or bounded completeness, further restrictions on ε_1 are required. The assumptions below, which are adapted from Mattner (1993), enable to underline the difference between the two notions.

[2]Then, indeed, $\mathbb{E}(\varepsilon_1|Z_0, W) = 0 = \mathbb{E}(\varepsilon_1|W)$.

A5. $\mathbb{P}^W-$almost surely, the characteristic function $\psi_{\varepsilon_1|W}$ of ε_1 conditionally on W vanishes nowhere.

A6. ε_1 is gaussian or satisfies, $\mathbb{P}^W$-almost surely on w and for all $x, y \in \mathbb{R}^{p-r}$, there exists $C(.)$ and $k(.)$ such as

$$f_{\varepsilon_1|W}(x+y, w) \leq C(w)(1+||x||^2)^{k(w)} f_{\varepsilon_1|W}(y, w),$$

where $||.||$ is the euclidian norm. Moreover $\psi_{\varepsilon_1|W}(., w)$ is infinitely often differentiable in $\mathbb{R}^p \backslash A$ for some finite set A.

A7. One of the following statement holds:
i) For $\mathbb{P}^W-$almost all w, there exists $s_0 \in \mathbb{R}^{p-r}$ such as $\mathbb{E}(\exp(-s_0' X_0)|W = w) < \infty$ and $t \in \mathbb{C}^{p-r}$, $|\mathfrak{Re}(t)| < |s_0|$, such as $\psi_{\varepsilon_1|W}(t, w) = 0$.
ii) For $\mathbb{P}^W-$almost all w, ε_1 is not normal conditionally on $W = w$ and there exists $\delta_1, \delta_2 > 0$ such as $\mathbb{E}\left(\exp\left(\delta_1 ||\varepsilon_1||^{1+\delta_2}\right)|W = w\right) < +\infty$.

A5 is often assumed in deconvolution problems (see e.g. Devroye (1989), Liu and Taylor (1989), Fan (1991), Fan and Truong (1993) or Li and Vuong (1998)) and is satisfied, among others, by gaussian, Student, Laplace and $\alpha-$stable distributions. The only common continuous distributions that fail to satisfy it are the uniform and triangular ones. A6, on the other hand, is restrictive. It imposes in particular that $f_{\varepsilon_1|W}(., w)$ is either gaussian or has heavy tails.[3] The condition holds for instance for Student and $\alpha-$ stable distributions (see Mattner, 1992). Lastly, the second part of assumption A7 i) should not be seen as the opposite of assumption A5, because the zeros of the characteristic function are allowed to be complex and not only real (see Mattner, 1993, for examples of distributions satisfying this condition). Assumption A7 ii) is satisfied for instance when the support of ε_1 is compact.

Theorem 2.3 *Suppose that (2.2) and A1-A4 hold. Then*
1) if A5 holds, X is bounded complete for Z.

[3]Put $x = -y$ to see that $1/f_{\varepsilon_1|W}$ must be at most of polynomial order. It can also be shown (see Mattner, 1992) that A7 is implied by the condition $0 < c(w) \leq f_{\varepsilon_1|W}(x, w)(1 + ||x||)^{\gamma(w)} \leq C(w) < \infty$ for all $x \in \mathbb{R}^{p-r}$ and some real $c(w), C(w)$ and $\gamma(w) > 0$.

2) If A5 and A6 hold, X is complete for Z.

3) If A7 holds, X is not complete for Z.

The proposition is based on the results of Gosh and Singh (1966) and Mattner (1992, 1993) on the completeness of location families. It emphasizes the difference between completeness and bounded completeness. Whereas the first is satisfied for many densities, completeness imposes stringent restrictions on $f_{\varepsilon_1|W}$. $f_{\varepsilon_1|W}$ cannot have a bounded support for instance. On the other hand, because A5 and A7 may hold together (see Mattner, 1993, for examples), X can still be bounded complete for Z in such situations.

The easiest way to interpret (2.2) is that Z causes X. However, it may be convenient sometimes to suppose instead that X causes Z. In the measurement error model of Hu and Schennach (2006) for instance, the unobserved variable must be complete for the measured variable. In this case, the model (2.2) is unnatural because we would prefer to write the measure as a function of the unobserved variable and an independent error rather than the contrary. Consider now the following model:

$$\mu_2(Z) = \nu_2(X) + \varepsilon_2, \tag{2.3}$$

where μ_2 and ν_2 are maps from $\mathbb{R}^q$ (resp. $\mathbb{R}^p$) to $\mathbb{R}^{q-r}$. We assume the following hypotheses, which are close to A1-A3.

A8. $X_0 \perp\!\!\!\perp \varepsilon_2 \,|\, W$.

A9. For $\mathbb{P}^W$-almost all w, $\nu_2(., w)$ is a one-to-one diffeomorphism on $\mathbb{R}^{q-r}$. Moreover, the conditional distribution of $\nu_2(X_0, w)$ admits a bounded positive density on $\mathbb{R}^{q-r}$ with respect to the Lebesgue measure.

Assumption A8 is exactly equivalent to A2. Assumption A9 is similar but stronger than A3. In particular, $\nu_2(., w)$ being one-to-one implies $q = p$. Moreover, restrictions are imposed on the conditional density of $\nu_2(X_0, w)$. On the other hand, no conditions like A1 are set on $\mu_2(., w)$.

Theorem 2.4 *Suppose that (2.3), A4-A5 (for ε_2) and A8-A9 hold. Then X is bounded complete for Z.*

Thus, even if completeness seems asymmetric in X and Z, to a certain extent the roles of X and Z can be exchanged. In this sense, the condition is similar to the rank condition in linear instrumental models.

3 Implications for the nonparametric instrumental regression

The previous result sheds light on the nonparametric instrumental regression. Indeed, suppose that

$$\begin{cases} Y &= \varphi(X) + \eta \\ \mu_1(X) &= \nu_1(Z) + \varepsilon_1 \end{cases} \tag{3.1}$$

with $\mathbb{E}(\eta|Z) = 0$. Such a system may be seen as the nonparametric version of the linear model with instruments. In this framework, the rank condition corresponds to the large support assumption A3 and the technical conditions A5 or A6 on ε_1. If one is reluctant to impose A6, bounded completeness and thus local identification of φ will be obtained solely. Identification is achieved globally only when φ is known to be bounded. This can happen for theoretical reasons, as for instance, in Blundell et al. (2003), or when X_0 has a finite support, as in Florens et al. (2003).

Proposition 3.1 *Suppose that (3.1) and A1-A5 holds. Then φ is identified on*

$$\mathcal{B} = \{h/(h-\varphi)(.,w) \text{ is bounded for } \mathbb{P}^W\text{-almost all } w\}.$$

If $\varphi(.,w)$ is bounded for $\mathbb{P}^W$-almost all w, then φ is identified globally.

This result is rather negative, because if $\varphi(.,w)$ is not bounded (e.g., $\varphi(.,w)$ is a linear form), only local nonparametric identification is achieved. However, global identification can be recovered if $\varphi(.,w)$ has a bounded derivative under the system defined by (3.1), with $\mu_1(X_0, W) = X_0$.[4]

[4] This particular case is considered for the sake of simplicity. The result could be extended to any known diffeomorphism $\mu_1(.,w)$ whose inverse has a bounded derivative.

A10. For $\mathbb{P}^W$-almost all w, $x \mapsto \varphi(x, w)$ is differentiable and $\frac{\partial \varphi}{\partial x}(., w)$ is bounded.

Proposition 3.2 *Suppose that (3.1) with $\mu_1(X_0, W) = X_0$, A2-A5 and A10 hold. Then φ is globally identified.*

Roughly speaking, the result stems from the fact that under (3.1), the order of the derivative and the expectation can be exchanged. This enables to identify $\mathbb{E}\left(\frac{\partial \varphi}{\partial x}(X_0, W)\middle|Z\right)$ and thus $\frac{\partial \varphi}{\partial x}(X_0, W)$ by bounded completeness.

4 Proofs

4.1 Theorem 2.3

Let $T_0 = \mu_1(X)$ and $T = (T_0, W)$. First, completeness (resp. bounded completeness) of X for Z can be deduced from completeness (resp. bounded completeness) of T for Z. Indeed, suppose that for all g such that $\mathbb{E}[|g(X_0, w)|] < +\infty$ (for $\mathbb{P}^W$-almost all w),

$$(\mathbb{E}[g(T)|Z] = 0 \text{ a.s.}) \Rightarrow (g(T) = 0 \;\text{ a.s.}) \,.$$

Then, by A1 there exists ψ_1 such that $X = \psi_1(T)$ almost surely. Moreover, if $\mathbb{E}[|h(X_0, w)|] < +\infty$, then $\mathbb{E}[|h \circ \psi_1(T_0, w)|] < +\infty$ and

$$(\mathbb{E}[h(X)|Z] = 0 \text{ a.s.}) \Rightarrow (h \circ \psi_1(T) = 0 \text{ a.s.}) \Rightarrow (h(X) = 0 \text{ a.s.}) \,.$$

The same holds with bounded functions $h(., w)$, because $h(., w)$ bounded implies that $h \circ \psi_1(., w)$ is bounded.

Now, let us rewrite the completeness statement. First, by conditional independence of Z_0 and ε_1,

$$\begin{aligned}
\mathbb{E}[h(T)|Z] &= \mathbb{E}[h(\nu_1(Z) + \varepsilon_1, W)|Z] \\
&= \int h(\nu_1(Z) + u, W) f_{\varepsilon_1|W}(u, W) du \quad \text{a.s.} \\
&= \int h(t, W) f_{\varepsilon_1|W}(t - \nu_1(Z), W) dt \qquad \text{a.s.}
\end{aligned}$$

Conditionally on W, the support of $\nu_1(Z)$ is $\mathbb{R}^{p-r}$ and its distribution is continuous. Moreover, T_0 is continuous conditionally on W. Hence, for any h such that $h(.,w)$ is integrable (resp. is bounded), completeness (resp. bounded completeness) of T for Z is equivalent to, $\mathbb{P}^W-$almost surely in w,

$$\left(\int h(t,w)f_{\varepsilon_1|W}(t-u,w)dt = 0 \quad \text{almost everywhere in } u\right) \Rightarrow (h(t,w)=0 \quad \text{a.e. in } t) \tag{4.1}$$

This statement corresponds to the completeness of the location family with density $f_{\varepsilon_1|W}$, except that the left part of (4.1) holds almost everywhere and not everywhere. But in theorem 1.3 of Mattner (1992) (and hence in his theorem 1.1), the statement also holds almost everywhere, so that we can apply it to obtain part 2 of the theorem. Moreover, a quick inspection of lemma 2.3 and theorem 2.4 of Mattner (1993) shows that it also holds almost everywhere. Part 3 is then a consequence of these two results.

To obtain part 1, we adapt the proof of theorem 2.4 of Gosh and Singh (1966). Let L^1 (resp. L^∞) denote the space of equivalent classes of integrable (resp. essentially bounded) functions with respect to the Lebesgue measure. Let w be such as $h(.,w)\in L^\infty$, $\Psi_{\varepsilon|W}(.,w)$ does not vanish anywhere and the left part of (4.1) holds (the set of such w being of probability one). Let $f_{w,u}(x) = f_{\varepsilon_1|W}(x-u,w)$, $\mathcal{P}_w = \text{span}\left\{f_{w,u},\ u\in\mathbb{R}^{p-r}\ /\ \int h(t,w)f_{w,u}(t)dt = 0\right\}$ and $\mathcal{Q}_w = \{f_{w,u}\ /\ u\in\mathbb{R}^{p-r}\}$.

Let $A_w = \{u\ /\ f_{w,u}\in\mathcal{P}_w\}$. Because the Lebesgue measure of ${}^cA_w$ is zero, there exists a sequence u_n of elements of A_w such as $u_n\to u$ for all $u\in {}^cA_w$. By Scheffe's theorem (see e.g. van der Vaart (1998), p. 22), $||f_{w,u_n}-f_{w,u}||_{L_1}\to 0$. Thus $\mathcal{Q}_w$ is included in the closure of $\mathcal{P}_w$.

Now, by Wiener's tauberian theorem (see e.g. Yoshida (1974), p. 357), $\mathcal{Q}_w$ is dense in L^1. Thus, $\mathcal{P}_z$ is dense in L^1. By continuity of the linear form $\phi\mapsto\int h(t,w)\phi(t)dt$ and the Riesz theorem (see e.g. Rudin (1998), p. 158), $h(t,w)=0$ for almost every t and almost all w.

4.2 Theorem 2.4

$\mathbb{E}[h(X)|Z] = 0$ implies that for any bounded measurable g, $\mathbb{E}[h(X)g(\mu_2(Z))|W] = 0$. In other terms,

$$\mathbb{E}[h(X)\mathbb{E}(g(\mu_2(Z))|X)|W] = 0.$$

By A8 and A9, the distribution of $\mu_2(Z)$ is continuous conditionally on X. Thus,

$$\int h(x,w)\left[\int g(u) f_{\mu_2(Z)|X_0,W}(u,x,w)du\right] d\mathbb{P}^{X_0|W=w}(x) = 0$$

for $\mathbb{P}^W-$almost all w. By Fubini's theorem,

$$\int g(u)\left[\int h(x,w) f_{\mu_2(Z)|X_0,W}(u,x,w) d\mathbb{P}^{X_0|W=w}(x)\right] du = 0.$$

Because this holds for every bounded measurable g, we get, for almost all w and almost everywhere in u,

$$\int h(x,w) f_{\mu_2(Z)|X_0,W}(u,x,w) d\mathbb{P}^{X_0|W=w}(x) = 0.$$

Hence, using (2.3), A8 and a change of variable $t = \nu_2(x,w)$,

$$\int h(\psi_2(t,w),w) f_{\nu_2(X_0,w)|W}(t,w) f_{-\varepsilon_2|W}(t-u,w)dt = 0$$

for almost everywhere in u and where $\psi_2(.,w)$ denotes the inverse map of $\nu_2(.,w)$. Now, because $f_{\nu_2(X_0,w)|W}(.,w)$ is bounded by A9, $h(\psi_2(.,w),w) \times f_{\nu_2(X_0,w)|W}(.,w)$ is bounded for every bounded h. Hence, we can apply the same device than in the previous proof, and $h(\psi_2(.,w),w) \times f_{\nu_2(X_0,z)|W}(.,w) = 0$ almost everywhere. The result follows because $f_{\nu_2(X_0,z)|W}(.,w) \neq 0$ and $\psi_2(.,w)$ is one-to-one.

4.3 Proposition 3.1

$\mathbb{E}(Y|Z) = \mathbb{E}(\varphi(X)|Z)$, so that any candidate φ' of φ satisfies

$$\mathbb{E}[(\varphi' - \varphi)(X)|Z] = 0.$$

If $\varphi' \in \mathcal{B}$, we can apply theorem 2.3, so that $\varphi'(X) = \varphi(X)$ almost surely. If $\varphi(., W)$ is known to be bounded, any candidate must be also bounded so that $(\varphi' - \varphi')(., W)$ is bounded. Thus, in this case $\varphi(., W)$ is globally identified.

4.4 Proposition 3.2

Because $\mathbb{E}(Y|Z) = \mathbb{E}(\varphi(X)|Z)$ and $f_{X_0|Z_0,W}(x, z, w) = f_{\varepsilon_1|W}(x - \nu_1(z, w), w)$, we get, almost surely,

$$\begin{aligned} \mathbb{E}(Y|Z_0 = z, W = w) &= \int \varphi(u + \nu_1(z, w), w) f_{\varepsilon_1|W}(u, w) du \\ &= \mathbb{E}(Y|\nu_1(Z_0, W) = \nu_1(z, w), W = w). \end{aligned}$$

In other terms, for almost all t and w,

$$\mathbb{E}(Y|\nu_1(Z) = t, W = w) = \int \varphi(u + t, w) f_{\varepsilon_1|W}(u, w) du.$$

Let w be such as $\varphi(., w)$ is differentiable. Because $\frac{\partial \varphi}{\partial x}(., w)$ is bounded, there exists $M(w)$ such as for every t,

$$\left| \frac{\partial}{\partial t} \left[\varphi(u + t, w) f_{\varepsilon_1|W}(u, w) \right] \right| \leq M(w) f_{\varepsilon_1|W}(u, w)$$

where the inequality must be understood termwise. Because the right term is integrable, $t \mapsto \mathbb{E}(Y|\nu_1(Z) = t, W = w)$ is differentiable and, almost everywhere,

$$\frac{\partial \mathbb{E}(Y|\nu_1(Z) = t, W = w)}{\partial t} = \int \frac{\partial \varphi}{\partial x}(u + t, w) f_{\varepsilon_1|W}(u, w) du.$$

Upon adding $\mathbb{E}(\varepsilon_1|W)$ to ν_1, we can always suppose that $\mathbb{E}(\varepsilon_1|W) = 0$. This normalization makes $\nu_1(.)$ identifiable because then, by A2, $\nu_1(Z) = \mathbb{E}(X|Z)$. Hence $\mathbb{E}(Y|\nu_1(Z) = t, W = w)$, and thus its derivative, are identifiable (almost surely).

Let φ' denote a candidate for φ and $g = \varphi' - \varphi$. Then, for almost every t,

$$\int \frac{\partial g}{\partial x}(x, w) f_{\varepsilon_1|W}(x - t, w) dx = 0.$$

In other terms,

$$\mathbb{E}\left(\frac{\partial g}{\partial x}(X_0, W)\big|Z\right) = 0 \quad \text{a.s.}$$

Because g has a bounded derivative and X is bounded complete for Z by theorem 2.3, $\frac{\partial g}{\partial x}(X_0, W) = 0$ almost surely. Hence $\frac{\partial \varphi}{\partial x}(X_0, W)$ is identified almost surely. Now, there exists φ_0 such as

$$\varphi(x, w) = \int_0^x \frac{\partial \varphi}{\partial x}(u, w)du + \varphi_0(w).$$

In other terms,

$$Y = \int_0^{X_0} \frac{\partial \varphi}{\partial x}(u, W)du + \varphi_0(W) + \eta.$$

Thus,

$$\mathbb{E}\left(Y - \int_0^{X_0} \frac{\partial \varphi}{\partial x}(u, W)du\big|W\right) = \varphi_0(W)$$

Hence φ_0 is identified and the result follows.

References

Blundell, R., Chen, X. and Kristensen, D. (2003), Nonparametric iv estimation of shape-invariant engel curves. cemmap Working Papers, CWP15/03.

Darolles, S., Florens, J. P. and Renault, E. (2002), Nonparametric instrumental regression. Working paper 05-2002 CRDE.

Devroye, L. (1989), 'Consistent deconvolution in density estimation', *Canadian Journal of Statistics* **17**, 235–239.

Fan, J. (1991), 'On the optimal rates of convergence for nonparametric deconvolution problems', *The Annals of Statistics* **19**, 1257–1272.

Fan, J. and Truong, Y. K. (1993), 'Nonparametric regression with errors in variables', *The Annals of Statistics* **21**, 1900–1925.

Florens, J. P., Heckman, J. J., Meghir, C. and Vytlacil, E. (2003), Instrumental variables, local instrumental variables and control functions. IDEI Working Paper 249.

Gosh, J. K. and Singh, R. (1966), 'Unbiased estimation of location and scale parameters', *The Annals of Statistics* **37**, 1671–1675.

Hoeffding, W. (1977), 'Some incomplete and boundedly complete families of distributions', *The Annals of Statistics* **5**, 278–291.

Hu, Y. and Schennach, S. (2006), Identification and estimation of nonclassical errors-in-variable models with continuous distributions using instruments. Working paper.

Lehmann, E. L. (1986), *Testing Statistical Hypothesis*, 2nd ed. Wiley: New-York.

Lewbel, A. (2000), 'Semiparametric qualitative response model estimation with unknown heteroscedasticity or instrumental variables', *Journal of Econometrics* **97**, 145–177.

Li, T. and Vuong, Q. (1998), 'Nonparametric estimation of the measurement error model using multiple indicators', *Journal of Multivariate Analysis* **65**, 139–165.

Liu, M. C. and Taylor, R. (1989), 'A consistent non-parametric density estimator for the deconvolution problem', *Canadian Journal of Statistics* **17**, 427–438.

Manski, C. (1988), 'Identification of binary response models', *Journal of the American Statistical Association* **83**, 729–738.

Mattner, L. (1992), 'Completeness of location families, translated moments, and uniqueness of charges', *Probability Theory and Related Fields* **92**, 137–149.

Mattner, L. (1993), 'Some incomplete but boundedly complete location families', *The Annals of Statistics* **21**, 2158–2162.

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

Rudin, W. (1998), *Analyse réelle et complexe, 3ème édition*, Dunod.

van der Vaart, A. W. (1998), *Asymptotic Statistics*, Cambridge Series in Statistical and Probabilistic Mathematics.

Yoshida, K. (1974), *Functional Analysis, fourth edition*, Springer Verlag.