Errors-in-variables models: a generalized functions approach

Victoria Zinde-Walsh*

McGill University and CIREQ

November 5, 2007

Abstract

Generalized functions are a powerful tool for examining errors-in-variables models, since they extend consideration to wide model classes. Schennach (Econometrica, 2007) - (S) applies this approach to prove identification in a general class of models. Here the problems addressed in (S) are revisited because various features of the generalized functions approach need to be clarified. The nonparametric identification theorem in (S) applies less generally than claimed (e.g.

*The support of the Social Sciences and Humanities Research Council of Canada (SSHRC), the *Fonds québecois de la recherche sur la société et la culture* (FRQSC) is gratefully acknowledged. The author thanks Peter C.B.Phillips for very illuminating discussions.

disallowing functions with fractional power growth) by relying on decomposition of generalized functions into ordinary and singular parts which may not hold. This paper highlights the issues of importance in applying generalized functions and provides the general nonparametric identification result relating it to possibility of estimation.

Keywords: errors-in-variables model, generalized functions

1 Introduction

Identification and estimation in nonlinear regression with measurement error is a difficult problem; solutions are known in various cases such as for polynomial regression functions (Hausman, Newey, Ichimura and Powell (1991) or for regression function in the L_1 space (Wang and Hsiao (1995). Schennach (2007) - hereafter (S) - proposes a solution for nonparametric identification in a wide class of models by utilizing generalized functions and their Fourier transforms; the assumptions allow for polynomial growth in the regression function (without requiring a parametric form) and make no requirements on the underlying distribution. The approach relies on the use of generalized functions in the form of tempered distributions. Results on semiparametric identification and estimation in (S) also make use of generalized functions.

Generalized functions offer many advantages as tools to deal with lack of differentiability and to regularize divergent integrals. Generalized functions are often applied in physics to solve equations where singularities may arise; in econometrics Phillips (1991,1995) worked with tempered distributions to derive asymptotics for LAD estimators, and used tempered distributions in the 1980s to characterize stable distributions (1985); generalized functions were employed by Zinde-Walsh and Phillips (2003) and Zinde-Walsh (2008) to deal with singularities. Often generalized functions are needed in intermediate steps only and lead to results that can be expressed via ordinary functions. The proofs that involve generalized functions are mostly based on verifying the definitions and known properties, but since some of the results have no ordinary analogues care must be taken that all the details are fully explored; some of the examples in this paper demonstrate how intuition based on pointwise approach to functions fails. For example, one of the difficulties is that non-linear operations such as multiplication cannot be generally defined in spaces of generalized functions and while they can be performed on some pairs of functions, it is important to verify that the resulting product is a legitimate generalized function (even when the product involves pointwise multiplication of ordinary functions it may not exist as a tempered distribution). Despite the technical difficulties presented by employing generalized functions, they represent a useful and appropriate tool for expanding the results in the errors-in-variables literature to wide classes of regression functions and distributions.

The model considered in (S) is an instrumental variables model with an unknown regression function g; it is represented as

$$\begin{aligned} Y &= g(X^*) + \Delta Y; E[\Delta Y\, |Z, U] = 0; \qquad (1) \\ X &= X^* + \Delta X; E[\Delta X\, |Z, U, \Delta Y] = 0; \\ X^* &= Z - U; U \text{ is independent of } Z; E(U) = 0 \end{aligned}$$

with scalar variables X, Y and Z which are observable (Z represents a transformation of the instruments via an identified projection of X on the instruments), while X^* and the errors $\Delta X, \Delta Y, U$ are not observable. As suggested in Newey (2001) two conditions are considered for identification of the unknown function g, with F denoting the distribution of U :

$$E[Y\,|Z=z] = \int g(z-u)dF(u); \tag{2}$$

$$E[ZY\,|Z=z] = \int (z-u)g(z-u)dF(u), \tag{3}$$

The convention is that all real integrals are over $[-\infty, \infty]$; overdot $\dot{f}$ represents derivative of f; $(i)^2 = -1$.

Consider Fourier Transforms:

$$\begin{aligned} \varepsilon_y(\zeta) &= \int E[Y\,|Z=z]e^{i\zeta z}dz; \qquad (4) \\ \varepsilon_{xy}(\zeta) &= \int E[XY\,|Z=z]e^{i\zeta z}dz; \\ \gamma(\zeta) &= \int g(x^*)e^{i\zeta x^*}dx^*; \\ \phi(\zeta) &= \int e^{i\zeta u}dF(u). \end{aligned}$$

If the functions are not integrable their Fourier Transforms may be singular. To resolve problems arising from possible singularity *tempered distributions* (see definitions below in Section 2) can be used since in the space of tempered distributions Fourier Transform as well as its inverse is always defined. The main source for generalized functions as tempered distributions is Lighthill

(1959) (L); other classes of generalized functions can also be useful and, as will be shown here, are necessary to get identification. Monographs by Gel'fand and Shilov (1964) -(GS), Schwartz (1950) and Sobolev (1992)(Sob) discuss other spaces of generalized functions, in (Sob) convenient diagrams of mappings of these spaces are provided. Using Fourier Transform requires considering complex-valued generalized functions, however as (L, p.2) and (GS, v.1, pp.15-16) point out, no special treatment is necessary and the properties of generalized functions extend to complex-valued cases.

It is known that generally on spaces of integrable functions Fourier Transform transforms convolutions when they exist (as in (2,3)) into products: solving (2,3) for g is equivalent (S, Lemma 1) to solving in the space of tempered distributions

$$\varepsilon_y(\zeta) = \gamma(\zeta)\phi(\zeta); \tag{5}$$

$$i\varepsilon_{xy}(\zeta) = \dot{\gamma}(\zeta)\phi(\zeta), \tag{6}$$

with two unknown functions given by tempered distributions, γ and ϕ. Tempered distributions are differentiable, thus $\dot{\gamma}(\zeta)$ is defined as a generalized derivative in the space of tempered distributions.

The equations (5,6) are employed to provide a general nonparametric identification result for g in (S). As is discussed in section 2 here the proof of Theorem 1 in (S) relies on the decomposition of a generalized function into an ordinary and a purely singular part; this decomposition does not generally hold, must be assumed and restricts application of the result to function classes that e.g. exclude some functions g with fractional power growth. A proof of the nonparametric identification result that does not rely on the decomposition and thus applies to the general function class is provided in section 2. The solution also requires multiplication by ϕ^{-1} possibly on all of R; the product may not exist in the space of tempered distributions and thus working in the space of tempered distributions is insufficient for the proof, a wider space of generalized functions has to be used. This is not purely a technical point since when an operation cannot be performed in the space of tempered distributions identification cannot be achieved as a result of a continuous mapping; this has implications for consistent estimation and inference, possibly turning them into ill-posed problems. While identification is necessary for the possibility of consistent estimation and for inference it does not guarantee the validity of estimation and testing, see, e.g. Dufour (2003) for a discussion and examples.

Parametric identification and a semiparametric estimation procedures are proposed in (S) for a class of models with the regression function given by a sum of a polynomial and a function whose Fourier Transform is an ordinary function. It fact, the result does not hold for some of the models in the class as is demonstrated in Section 3; the proof may require multiplication of a discontinuous function with a singular $\delta-$function - such a product is undefined. An example given here illustrates that the estimation procedure may not work; the assumption that the ordinary part of the Fourier Transform be differentiable is required.

The proofs are in section 4.1 of the Appendix ; discussion of multiplication for generalized functions with examples is in 4.2 of the Appendix.

2 Non-parametric identification

The problem of identification of g can be stated as follows. Define the functions W_1, W_2;

$$\begin{aligned} W_1(z) &= E(Y|Z=z); \\ W_2(z) &= E(Y(Z-X)|Z=z). \end{aligned}$$

Following (S) make the following assumptions.

Assumption 1. $|g(x^*)|, |W_1(z)| \equiv |E[Y|Z=z]|, |W_2(z)| \equiv |E[XY|Z=z]|$ are defined and bounded by polynomials for $x^*, z \in R$.

For the distribution F the assumption is in terms of implications for Fourier Transform, ϕ.

Assumption 2. (i) Absolute moment of U exists and (ii) $\phi(\zeta) \neq 0$ everywhere on R.

Differentiability for ϕ follows from 2(i).

There is a restriction on the Fourier Transform of g.

Assumption 3. There exists a positive finite or infinite $\bar{\zeta}$ such that $\gamma(\zeta) \neq 0$ almost everywhere in $[-\bar{\zeta}, \bar{\zeta}]$ and (ii) $\gamma(\zeta) = 0$ for all $|\zeta| > \bar{\zeta}$.

Denote by A the class of functions (g, F), by A^* the class of functions (g), by B the class of functions (W_1, W_2) where assumptions 1-3 are assumed to hold.

Then (2,3) can be written as a mapping

$$M : A \to B,$$

such that $M(g, F) = (W_1, W_2)$. If for any $g' \neq g$ (on a set of positive Lebesgue measure) and any F' such that $(g', F') \in A$ we have $M(g, F) \neq M(g', F')$ then the function g is identified. In such a case there exists a mapping

$$M^* : B \to A^* \tag{7}$$

that provides a unique function g for any (W_1, W_2). This could be a basis for estimation if a computational algorithm to implement the mapping can be found; consistent estimation would require that the mapping be continuous.

In Theorem 1 of (S) g is obtained as the inverse Fourier Transform of $\gamma(\zeta)$ defined by

$$\gamma(\zeta) = \begin{cases} 0 & \text{if } \varepsilon_y(\zeta) = 0; \\ \varepsilon_y(\zeta)/\phi(\zeta) & \text{if } \varepsilon_y(\zeta) \neq 0; \end{cases}$$

(note that $\varepsilon_y(\zeta) \neq 0$ almost everywhere in $[-\bar{\zeta}, \bar{\zeta}]$). The notation used in Theorem 1 of (S) is kept here; note that here ϕ differs from the Fourier transform in (4): the function $\phi(\zeta)$ is restricted to the set where $\varepsilon_y(\zeta) \neq 0$

and defined there by

$$\phi(\zeta) = \exp(\int_0^\zeta \frac{i\varepsilon_{(z-x)y,o}(\xi)}{\varepsilon_{y,o}(\xi)} d\xi). \tag{8}$$

The functions subscripted by o in (8) are defined in (S) as ordinary functions representing the ordinary parts of the Fourier transforms of known functions (based on W_1, W_2). It is claimed that they can be obtained via the decomposition of the Fourier Transforms of the generalized functions into the sum of an ordinary and purely singular components: $\varepsilon = \varepsilon_o + \varepsilon_s$ (the subscript o refers to ordinary and s to purely singular parts).

Thus the result in (S) relies on two operations in the space of tempered distributions: decomposition into ordinary and purely singular components and multiplication (e.g. by ϕ^{-1}). An example given here in 2.1.2 demonstrates that the required decomposition may not hold (e.g. for a function with fractional power growth) and needs to be assumed resulting in a more restricted class of functions for which the identification result in (S) holds; to provide a general result (Theorem 1 here) a proof that does not rely on the decomposition is given.

The advantage of Fourier Transform is that convolutions are transformed into products. Products of tempered distributions cannot be generally defined (Schwartz's impossibility result, 1954) although there are cases when specific products can be defined (see L, p. 18). It is necessary to verify that the product exists, since even for two ordinary functions, each of which is a tempered distribution, the product may not be defined as a tempered distribution (see, e.g. Kaminski and Rudnicki, 1991 and example 2 in the Appendix). The proof of the identification result relies on existence of products such as

$$\varepsilon_y(\zeta)\phi^{-1}(\zeta) = [\gamma\,(\zeta)\,\phi(\zeta)]\phi^{-1}(\zeta)$$

to obtain γ, the Fourier Transform of the unknown function. The "apparent" cancellation on the r.h.s. cannot be taken for granted to produce γ; the products that are needed do not always exist in the space of tempered distributions (and multiplication does not have to have the associative property). The problem is that tempered distributions are sensitive to behavior in the tails of the function and the assumptions 1-3 do not place sufficient restrictions on the characteristic function of the error density to permit multiplication by ϕ^{-1}.

After clarifying the issues about decomposition and multiplication of generalized functions in the next subsection 2.1 a general identification theorem is proved in 2.2.

2.1 Spaces of generalized functions and required properties

2.1.1 Tempered distributions, other generalized functions and Fourier Transform

Denote by T the space of test functions (Lighthill's "good functions"):

$$T = \left\{ s \in C_\infty(R) : \left| \frac{d^k s(t)}{dt^k} \right| = O(|t|^{-l}) \text{ as } t \to \infty, \text{ for all } k, l \in N^+ \right\}$$

where $C_\infty(R)$ is the space of all infinitely differentiable functions; $N^+ = \{0, 1, ...\}$; $k = 0$ corresponds to the function itself; $|\cdot|$ is the absolute value. Fourier Transform can be defined on T and maps T onto itself (L,2.1). A generalized function, b, on T, also called a tempered distribution, is defined via equivalence classes of converging sequences of test functions in T :

$$b = \left\{ \{b_n\} : b_n \in T, \text{ such that for any } s \in T, \lim_{n\to\infty} \int b_n(t)s(t)dt = (b, s) < \infty \right\}.$$

where (b, s) is a constant independent of the sequence $\{b_n\}$; denote (b, s) by $\int b(t)s(t)dt$. It is shown (L, GS) that tempered distributions form a linear space T' which is the space of all linear continuous functionals on T. Sums, derivatives, linear transformations of arguments, multiplication by a function from T and Fourier Transform are consistently defined as continuous operations for generalized functions by applying these operations to the test functions in each converging sequence of the equivalence class (L, 2.2). All functions from T are in T', as well as all locally summable functions with no more than power growth, but T' also contains singular elements, such as the Dirac's delta-function, $\delta : \int \delta(t)s(t)dt = s(0)$. More complicated singular generalized functions may have non-isolated singularities; even a non-countable set of singularities as in the derivative of the continuous Cantor function. A purely singular function b_s has support on a set of Lebesgue measure zero, meaning that for some set $S \subset R$ of Lebesgue measure zero and any

$$\bar{s}(t) \in T : \bar{s}(t) = 0 \text{ for } t \in S$$

$\int b_s(t)\bar{s}(t)dt = 0$ (e.g. $t = 0$ is the only point of support for $\delta-$ function).

Consider also $D \subset T$ containing all the functions in T with finite support, and the corresponding space of generalized functions, D', that is also often used to permit enlargement of spaces of functions to include singular generalized functions; D' is studied in detail in (GS) and (Sob). The space D is not closed under Fourier Transform: Fourier transform of a function in D may not belong to D (see GS, p.167), thus D' cannot provide the same advantages as T' here. Note, however, that $D' \supset T'$ so that T' is a linear subspace but it is not a closed subspace of D' : some classes of sequences that diverge in T' converge in D'.

There is a difference between what is understood as an "ordinary function" depending on the space of test functions considered, thus (GS) call any locally integrable real (or complex) function an ordinary function; any such function $\in D'$, but not necessarily to T'; for an ordinary function to be a tempered distribution it cannot grow excessively fast at infinity, so that by (L) an ordinary function $b(t)$ satisfies

$$\int (1+t^2)^{-l} |b(t)| \, dt < \infty \text{ for some } l \geq 0.$$

Other useful spaces of generalized functions are $D'(U)$, similarly defined but for test functions with support restricted to U; $D(U) \subset D$. Sometimes spaces D_m where test functions are only $m = 0, 1, ..$ times continuously differentiable and corresponding spaces of functionals D'_m are used (also possibly restricted to some U).

2.1.2 Decomposition of generalized functions and products of generalized functions

A crucial point for the proof of (S, Theorem 1) is the existence of a decomposition of a tempered distribution into a sum of an ordinary function and a purely singular part. The pointwise argument for the existence of such a decomposition in (S) (see also supplementary material to (S)) is not correct.

Suppose that the function g has fractional power growth, to simplify the discussion, say grows as $|x|^{\frac{1}{2}}$.

$$g(x) = 2^{-\frac{1}{4}} \frac{|x|^{\frac{1}{2}}}{\Gamma(\frac{2}{4})};$$

then its Fourier Transform, $\gamma,$ is (see. GS, v. 1 p.173, function $f_\lambda(x)$)

$$\gamma(\zeta) = c\,|\zeta|^{-\frac{3}{2}}\ ,\ \text{where } c = \frac{2^{\frac{3}{4}}}{\Gamma(-\frac{1}{4})}$$

The function $|\zeta|^{-\frac{3}{2}}$ has a singularity at $\zeta = 0$; by (GS, v.1, p.51) $\gamma(\xi)$ as a generalized function applied to $\psi \in T$ provides

$$\int \gamma(\xi)\psi(\xi)d\xi = \int c\,|\zeta|^{-\frac{3}{2}}\,\psi(\zeta)d\zeta = c\int_0^\infty \zeta^{-\frac{3}{2}}\{\psi(\zeta)+\psi(-\zeta)-2\psi(0)\}d\zeta.$$

This generalized function γ cannot be represented as a sum of an ordinary function and a purely singular function. Indeed it clearly has a singularity at zero only; by definition of local properties (see, eg, GS, v.1, p.14) it can be said that it coincides (pointwise) with $c\,|\zeta|^{-\frac{3}{2}}$ for any $\zeta > 0$ or $\zeta < 0$.At the same time for any open $U = (0, a)$ or $U = (-a, 0)$ $(a > 0)$ the function $c\,|\zeta|^{-\frac{3}{2}}$ is not an ordinary function since it is not integrable there.

The function g above satisfies assumptions 1-3, but an attempt to use (8) for this case as required by (S,Theorem 1) will not provide $\phi(\xi)$ since the needed integral diverges. Thus Theorem 1 in (S) will apply only to classes of generalized functions where such a decomposition holds. The following assumption needs to be added to 1-3 for (S, theorem 1) to hold.

Assumption D. Decomposition into ordinary and singular parts holds for the Fourier Transform γ of g.

Some classes of functions that satisfy Assumption D can be identified.

For example, Theorem 19 of (L) provides a class of functions $\{g\}$ where both the functions and the Fourier Transforms not only satisfy Assumption D but also have easily expressible Fourier Transforms: the functions are defined explicitly as sums of derivatives of $\delta-$functions and some ordinary functions in T'. Of course, Assumption D is satisfied in all cases where Fourier Transform of g, γ, is an ordinary function. If an addition to Assumptions 1-3 Assumption D also holds (S, Theorem 1) holds[1].

For the proof of identification one needs to verify the validity of products between some generalized functions and some continuous functions. Consider a continuous function a and a generalized function b that is defined in T' or in D'. For the time being denote by G' the space of generalized functions (G the corresponding space of test functions); G' is either T' or D' here. The definition of a product of a tempered distribution with a continuous function in (L) specifies that a should be a "fairly good function" (L, Def.2): an infinitely differentiable function such that it and all its derivatives are bounded by a polynomial function at infinity; the proof of existence of the product (L., p.18) requires verifying that the same limit obtains no matter which sequence is used for the generalized function. Here to extend the definition of product to functions that may not be "good functions" we need to make the requirement that the product does not depend on the sequence that defines the generalized function. We thus say that ab is defined in G' if

[1] The proof in (S) however will require a further modification since generally multiplication that is used to obtain identification cannot be performed in the space of tempered distributions.

for any sequence b_n from the equivalence class that defines b there exists a sequence $(ab)_n$ in G such that for any $\psi \in G$

$$\lim \int a(x)b_n(x)\psi(x)dx \text{ exists and equals } \lim \int (ab)_n (x)\psi(x)dx. \qquad (9)$$

The following proposition gives necessary and sufficient conditions for the product ab to be defined. Denote by 0_n a zero-convergent sequence that belongs to the equivalence class defining the function that is identically zero in G'.

Proposition *For the product ab to be defined it is necessary and sufficient that (i)* (9) *hold for some sequence $\tilde{b}_n$ in the class that defines b and (ii) for any zero-convergent sequence, $0_n(x)$,*

$$\lim \int a(x)0_n(x)\psi(x)dx = 0. \qquad (10)$$

Proof.

Any sequence b_n differs from a specific $\tilde{b}_n$ by a zero-convergent sequence.$\blacksquare$

Denote $I(A) = 1$ if A is true, 0 otherwise.

Lemma 1 *Under Assumptions 1-3 the products $\gamma\phi$ and $\gamma\dot{\phi}$ are defined in T' and in D'; for $\tilde{\phi}^{-1} = \phi^{-1}(\zeta)I(|\zeta| < \bar{\zeta})$ the product $(\gamma\phi) \cdot \tilde{\phi}^{-1}$ is defined in D'; the product $(\gamma\phi) \cdot \tilde{\phi}^{-1}$ is defined in T' if either $\bar{\zeta} < \infty$* or *ϕ^{-1} is such that*

$$\int (1 + x^2)^{-l} \left|\phi(x)^{-1}\right| dx < \infty \text{ for some } l \geq 0. \qquad (11)$$

Proof. See Appendix.

2.2 The nonparametric identification theorem

The theorem here differs from the result in (S) in two ways. Firstly, it does not rely on the decomposition of generalized functions into ordinary and singular parts; this provides the identification result under the less restrictive Assumptions 1-3 without the additional Assumption D. Secondly, the condition under which the identification mapping is continuous is given here (the same condition applies to the identification result in (S) when it holds).

Theorem 1 *For functions satisfying Assumptions 1-3 the mapping* M^* *in (7) exists and provides identification for* g*; if (11) holds or if* $\bar{\zeta} < \infty$ *the mapping is continuous; it can be discontinuous when (11) does not hold.*

Proof. See Appendix.

The implication of this Theorem is that the identification result holds under the general assumptions 1-3. If ϕ is too thin-tailed, however, the mapping whereby the identification is achieved may not be continuous: this point is illustrated by the example in the proof of Theorem 1 where high frequency components b_n are magnified by multiplication with ϕ^{-1} from a thin-tailed distribution; this produces inverse Fourier Transforms that diverge.

Assuming that ϕ satisfies *(11)* it is possible to sketch a procedure for consistent nonparametric estimation of g. Suppose that W_{1n}, W_{2n} are some estimates of conditional moments functions, W_1 and W_2;

$$W_{1n} = g_n * f_n;\ W_{2n} = (zg_n) * f_n,$$

where the (generalized) functions g_n and f_n are not known but satisfy Assumptions 1-3. Denote by ε_{1n} the Fourier Transform of W_{1n} and by ε_{2n} the Fourier Transform of W_{2n}. Suppose that these Fourier Transforms are consistent estimators of $\varepsilon_1 \equiv \varepsilon_y, \varepsilon_2 \equiv \varepsilon_{xy}$ in (4) that converge to those limits as $n \to \infty$ in the space of tempered distributions; assume additionally that they are given by continuous functions that are non-zero a.e.. Consider functions with support on all of R (it may be possible to restrict attention to $|\zeta| < \bar{\zeta}$ if this bound were known). From arguments similar to those in the proof of Theorem 1 in the Appendix a continuous function $f_n(\zeta)$, that satisfies the equation

$$f_n(\zeta)\varepsilon_{1n}(\zeta) + (i\varepsilon_{2n}(\zeta) - \dot{\varepsilon}_{1n}(\zeta)) = 0 \tag{12}$$

in generalized functions, exists and is unique. Since all the functions that determine f_n are continuous (12) holds for ordinary continuous functions and since ε_{1n} is non-zero a.e. we have

$$f_n(\zeta) = (i\varepsilon_{2n}(\zeta) - \dot{\varepsilon}_{1n}(\zeta))(\varepsilon_{1n}(\zeta))^{-1}.$$

The generalized functions

$$f_n\varepsilon_{1n} - f\varepsilon_1 = i(\varepsilon_{2n} - \varepsilon_2) + (\dot{\varepsilon}_{in} - \dot{\varepsilon}_1) \text{ and } f(\varepsilon_1 - \varepsilon_{1n})$$

converge to zero as generalized functions; as a result, so does $(f_n - f)\varepsilon_{1n}$, but since this is a continuous function this implies pointwise convergence and it follows from $\varepsilon_{1n} \neq 0$ that $f_n \to f$. From the differential equation $\phi_n^{-1}\dot{\phi}_n = f_n$ with the condition $\phi_n(0) = 1$ the function ϕ_n is uniquely determined; and $\phi_n \to \phi$ where $\phi^{-1}\dot{\phi} = f, \phi(0) = 1$. Then also since ϕ is non-zero, $\phi_n^{-1} \to \phi^{-1}$ pointwise; if ϕ satisfies *(11)* so does ϕ_n so that $\varepsilon_{1n}\phi_n^{-1}$ can be defined as a tempered distribution. Finally consider

$$\varepsilon_{1n}\phi_n^{-1} - \varepsilon_1\phi^{-1} = \varepsilon_{1n}(\phi_n^{-1} - \phi^{-1}) + (\varepsilon_{1n} - \varepsilon_1)\phi^{-1};$$

since the continuous function $\varepsilon_{1n}(\phi_n^{-1} - \phi^{-1}) \to 0$ and $(\varepsilon_{1n} - \varepsilon_1)\phi^{-1} \to 0$ in T' (as a tempered distribution) $\varepsilon_{1n}\phi_n^{-1}$ converges to γ in T', and its inverse Fourier Transform converges to g as a tempered distribution (by continuity of inverse Fourier Transform in T').

3 Parametric specification and semiparametric estimation

Here we examine the semiparametric estimation proposed in (S). The nonparametric identification result of required that the characteristic function $\phi(\zeta)$ be positive for all $\zeta \in R$. If $g(x^*)$ is parametrically specified as $g(x^*, \theta)$ this requirement can be replaced by the less restrictive assumption 4 in (S) that demands only that $\phi \neq 0$ on a possibly finite interval around zero that is sufficient for identification of the true θ^*. The method for semiparametric estimation in (S) is applied to a function $g(x, \theta)$ such that it can be represented as a sum of a polynomial with a function that has an ordinary function as its Fourier transform. It turns out that this class is too wide for the proposed method to work; it cannot be applied to some regression functions with ordinary Fourier transforms, specifically to functions such that their generalized derivatives have singularities.

Semiparametric estimation results in (S) are provided for two classes of cases. The first is when all of the Fourier Transforms in (4) are ordinary functions. Note that if $\dot{\gamma}$ is an ordinary function then γ is differentiable a.e.. The second applies to functions with singularities of γ given by a linear combination of derivatives of $\delta-$functions at zero, specifically,

$$\gamma(\zeta, \theta) = \gamma_o(\zeta, \theta) + 2\pi \sum_{k=0}^{\bar{k}} \gamma_k(\theta)(-i)^k \delta^{(k)}(\zeta), \tag{13}$$

where $\gamma_o(\zeta,\theta)$ is an ordinary function, $\bar{k} \in N$, and the $\gamma_k(\theta)$ for $k = 0, ...\bar{k}$ are $\theta-$dependent scalar parameters; $\gamma_{\bar{k}}(\theta) \neq 0$. ((S), Assumption 5). The assumption does not require differentiability of $\gamma_o(\zeta,\theta)$ and in this respect differs from the case treated in (S, 3.1.1)). The proof in (S) does not make this distinction; the possibility of an ordinary part γ_o with $\dot{\gamma}_o$with singularity was overlooked.

Consider a specific example with $g(x^*,\theta) = \theta sinc(\frac{\theta x^*}{2})$ where

$$sinc(x) = \frac{\sin x}{x};$$

to satisfy assumption 5 in (S) a polynomial could be added, but this does not alleviate the problem. The *sinc* function is widely used in signal processing (it is the so-called low-pass filter); it is also known as the zeroth order Bessel function of the first kind. Its Fourier Transform is the so-called *rect* function:

$$\begin{aligned} \gamma(\zeta,\theta) &= \pi rect\left(\frac{\zeta}{\theta}\right) \\ &= \begin{cases} 0 & |\zeta| > \frac{1}{2}\theta; \\ \frac{1}{2}\pi & |\zeta| = \frac{1}{2}\theta; \\ \pi & |\zeta| < \frac{1}{2}\theta. \end{cases} \end{aligned} \tag{14}$$

Note that then

$$\dot{\gamma}(\zeta,\theta) = \pi\left[\delta(\zeta + \frac{1}{2}\theta) - \delta(\zeta - \frac{1}{2}\theta)\right]. \tag{15}$$

We see that while γ is an ordinary function its derivative exists only as a generalized function and has singularities at points $\pm\frac{\theta}{2}$.

To construct the functions for the moment conditions in proof of Theorem 2 (S) equates the ordinary parts of (5,6), multiplies the left-hand side of (5) by the right-hand-side of (6) and thus eliminating ϕ obtains the equation (S,(68))

$$\varepsilon_{y,o}(\zeta)\dot{\gamma}_o(\zeta,\theta) = i\varepsilon_{xy,o}(\zeta)\gamma_o(\zeta,\theta); \tag{16}$$

this is used to find the weighting functions for the GMM conditions. Clearly, this does not hold if $\gamma_o(\zeta,\theta)$ is given by (14); then (16) would involve the product of the function $\varepsilon_{y,o}(\zeta)$, which has a point of discontinuity (recall its definition in (5)), with a $\delta-$ function shifted to the same point; this product does not exist as a generalized function. Thus the proof does not apply.

The problem for estimation is caused by the dependence of the location of the singularity in the derivative of the Fourier Transform of the regression function on the unknown parameters; the moment conditions in (S) cannot apply to such a case and thus do not apply to some ordinary functions.

For Theorems 2 and 3 in (S) to hold it is necessary to strengthen Assumption 5 in (S) to require differentiability of the ordinary part of $\gamma(\zeta,\theta)$:

$$\gamma(\zeta,\theta) = \gamma_o(\zeta,\theta) + 2\pi\sum_{k=0}^{\bar{k}}\gamma_k(\theta)(-i)^k\delta^{(k)}(\zeta), \tag{17}$$

where $\gamma_o(\zeta,\theta)$ and its generalized derivative $\dot{\gamma}_o(\zeta,\theta)$ are ordinary functions.

For consistent semiparametric estimation of functions $g(x^*,\theta)$ with Fourier Transform given by the more general assumption (13) a method that can deal with parameters that are associated with singular points of the derivative of the ordinary part still needs to be developed.

4 Appendix

4.1 Proofs

Proof of Lemma 1.

By Assumption 1, γ and $\phi \in T' \subset D'$, by (5) $\gamma\phi \in T'$, and additionally (by applying the product rule to (5,6)) $\gamma\dot{\phi} \in T'$. Since $T' \subset D'$, the products are defined in D' as well. Now consider a sequence $(\gamma\phi)_n$ defined as follows: select some sequence $\tilde{\gamma}_n$ for γ from D; then each γ_n has finite support; for a sequence of numbers $\varepsilon_n \to 0$ select $\tilde{\phi}_n$ in D such that $\left|\tilde{\phi}_n - \phi\right| < \frac{\varepsilon_n}{\sup\left|\gamma_n^{-1}\phi\right|}$ on compact support of γ_n. Then for $(\gamma\phi)_n = \tilde{\gamma}_n\tilde{\phi}_n$

$$\int \tilde{\gamma}_n\tilde{\phi}_n\phi^{-1}\psi = \int \tilde{\gamma}_n\psi + \int \tilde{\gamma}_n(\tilde{\phi}_n - \phi)\phi^{-1}\psi \to \int \gamma\psi.$$

Since $T \supset D$ this is valid in both spaces. The same applies to multiplication by $\dot{\phi}$. Now we check that (10) holds for $a = \phi^{-1}$. In D support of any ψ is bounded, on that compact set ϕ^{-1} is bounded thus (10) will hold and the product is defined in D'. If $\bar{\zeta} < \infty$ the product with $\phi^{-1}(\zeta)I(|\zeta| < \bar{\zeta})$ is similarly defined in T'. Under condition (11) consider $\tilde{\phi}(\zeta) = (1+\zeta^2)^l\phi(\zeta) \in$

T. Then in T'

$$\begin{aligned} & \lim \left| \int^{-1} \phi^{-1}(\zeta) 0_n(\zeta) \psi(\zeta) d\zeta \right| \\ \leq \ & \sup (1+\zeta^2)^{-l} \left| \phi(\zeta)^{-1} \right| \lim \int \left| 0_n(\zeta) \tilde{\psi}(\zeta) \right| d\zeta = 0, \end{aligned}$$

thus the product with a zero-convergent sequence is zero and the product exists in T'■

Proof of Theorem 1.

The proof makes use of different spaces of generalized functions and exploits relations between them. It proceeds in two parts.

First in part one, it is shown that from equations (5,6) the continuous function $f = \dot{\phi}\phi^{-1}$ can be uniquely determined on the interval $[-\bar{\zeta}, \bar{\zeta}]$ (where γ and consequently ε_1 differ from zero a.e. as generalized functions); this requires additionally considering the generalized functions spaces, D' and $D_0(U)'$. The function $\tilde{\phi}$ is uniquely defined on the interval $[-\bar{\zeta}, \bar{\zeta}]$ as the solution of the corresponding differential equation that satisfies the condition $\phi(0) = 1$; define $\tilde{\phi} = \phi I(|\zeta| < \bar{\zeta})$; define $\tilde{\phi}^{-1}$ to equal $\phi^{-1} I(|\zeta| < \bar{\zeta})$. Of course, when $\bar{\zeta} = \infty$, $\tilde{\phi} = \phi$ and $\tilde{\phi}^{-1} = \phi^{-1}$ on R.

Next in part two, γ is defined as $\varepsilon_1 \tilde{\phi}^{-1}$. By Lemma 1 this product can always be uniquely defined as a generalized function in D'; by construction $\gamma \in T' \subset D'$; this provides the required mapping M^* by applying inverse Fourier Transform to γ. If either $\bar{\zeta} < \infty$ or ϕ satisfies *(11)* $\tilde{\phi}^{-1}$ satisfies *(11)* and the product $\varepsilon_1 \tilde{\phi}^{-1}$ is defined in T'; in this case the mapping M^*

is continuous. The proof concludes with an example that demonstrates that the mapping can be discontinuous if *(11)* does not hold.

Part one. Consider the space of generalized functions D'. By Assumption 2 ϕ is non-zero and continuously differentiable, then by differentiating (5), substituting (6) and making use of Lemma 1 to multiply by ϕ^{-1} in D' we get that the generalized function

$$\varepsilon_1 \phi^{-1} \dot{\phi} - (\dot{\varepsilon}_1 - i\varepsilon_2)$$

equals zero as a generalized function in D'. Denoting $f = \dot{\phi}\phi^{-1}$ we can characterize f as a continuous function in D' that satisfies the equation

$$\varepsilon_1 f - (\dot{\varepsilon}_1 - i\varepsilon_2) = 0. \tag{18}$$

If (18) holds in D', it holds also for any test functions with support limited to $U : \psi \in D(U) \subset D$, and thus holds in any $D(U)'$.

We show that the function f is uniquely determined on $[-\bar{\zeta}, \bar{\zeta}]$ by (18) holding in $D(U)'$ for any interval $U \subset [-\bar{\zeta}, \bar{\zeta}]$. Proof is by contradiction. Suppose that there are two distinct continuous functions $f_1 \neq f_2$ that satisfy (18), then $f_1(\bar{x}) \neq f_2(\bar{x})$ for some $\bar{x} \in [-\bar{\zeta}, \bar{\zeta}]$; by continuity $f_1 \neq f_2$ everywhere for some interval $U \in [-\bar{\zeta}, \bar{\zeta}]$. Consider now $D(U)'$; we can write

$$\int \varepsilon_1 (f_1 - f_2)\psi = 0$$

for any $\psi \in D(U)$. A generalized function that is zero for all $\psi \in D(U)$ coincides with the ordinary zero function on U and is also zero for all $\psi \in D_0(U)$, where D_0 denotes the space of continuous test functions. For the space of test function $D_0(U)$ multiplication by continuous $(f_1 - f_2) \neq 0$ is an isomorphism. Then from (18) we can write

$$0 = \int [\varepsilon_1(f_1 - f_2)]\,\psi = \int \varepsilon_1\,[(f_1 - f_2)\psi]$$

implying that $\varepsilon_1$is defined and is a zero generalized function in $D_0(U)'$. If that were so ε_1 would be a zero generalized function in $D(U)'$ since $D(U) \subset D_0(U)$; this contradicts Assumption 2. This concludes the first part of the proof since from f the function

$$\phi(\zeta) = \exp\int_0^\zeta (f(\xi))d\xi$$

that solves on $[-\bar{\zeta}, \bar{\zeta}]$

$$\dot{\phi}\phi^{-1} = f; \phi(0) = 1$$

is uniquely determined on $[-\bar{\zeta}, \bar{\zeta}]$ and $\tilde{\phi}$ (and $\tilde{\phi}^{-1}$) defined above are uniquely determined.

Part two.

Consider two cases.

Case 1. Either $\bar{\zeta} < \infty$ or the condition (11) holds in T' in which case (11) holds for $\tilde{\phi}^{-1}$. Multiplying $\varepsilon_1(= \gamma\tilde{\phi})$ by $\tilde{\phi}^{-1}$ provides a tempered distribution by Lemma 1 here; it is equal to γ. The inverse Fourier Transform provides

g. The theorem holds and moreover, since all the operations by which the solution was obtained were continuous in T', the function g is recovered by a continuous mapping M^*.

Case 2. The condition (11) does not hold and multiplication by $\tilde{\phi}^{-1} = \phi^{-1}$ may not lead to a tempered distribution. Consider now D'; $T' \subset D'$. Multiplication by ϕ^{-1} is a continuous operation in D'; define the same differential equations, solve to obtain ϕ and get via multiplication $(\gamma\phi) \cdot \phi^{-1}$ in D' the function $\gamma \in D'$. Since γ is the Fourier transform of g (a tempered distribution) it also belongs to T', and it is possible to recover g by an inverse Fourier Transform.

In the following example the mapping $M^*(7)$ is not continuous. Define β_n as a function with Fourier Transform equal to b_n of Example 2 in Appendix (4.2). Suppose that $W_{1n} = W_1 + \beta_n$; from $b_n \to 0$ in T' and the continuity of the Fourier Transform mapping in T', it follows that $\beta_n \to 0$ and $W_{1n} \to W_1$ in T'. Then $\varepsilon_{yn} = \varepsilon_y + b_n$. Suppose that ϕ was identified exactly and is proportionate to $\tilde{\phi}$ in Example 2. Then each $\gamma_n = \varepsilon_{yn}\phi^{-1}$ is in T', the inverse Fourier transform, $\tilde{g}_n$, exists, but $\tilde{g}_n$ does not converge to g in T'. Indeed, if it did so converge, then that would imply convergence $\gamma_n \to \gamma$ in T', but by Example 2 the sequence γ_n diverges in T' ■

4.2 Products involving ordinary functions, $\delta-$functions and zero functions in the space of tempered distributions

Generally products are undefined in spaces of generalized functions (Schwartz, 1954). However, some functions are "mutual multipliers" (Sob), but it is necessary to establish this by showing that the product is a valid generalized function.

Lighthill (L. p.18) defines a product between some $b \in T'$ and a real continuous function F by the class of sequences $\{b_n F\}$; he actually only considers his "fairly good" functions (that do not diverge too fast) as F and shows that the product always exists. A very common misconception about tempered distributions is possibly fueled by a misreading of Lighthill's introduction of the delta-function (L, 1.5) where it is stated that

$$\int \delta(x)F(x)dx = F(0) \tag{19}$$

for any "suitably continuous function $F(x)$". "Suitably" continuous cannot be taken to mean any continuous at zero function, since the statement is proved only for $F \in T$. Moreover, a seeming representation of a $\delta-$function as an ordinary function everywhere outside of $x = 0$ in (L,2.4) can also easily lead to a misreading (not intended in the book), say, along the following lines: "If the $\delta-$function is distinct from 0 only at 0, and behaves as a function identically equal to 0 elsewhere it can be multiplied by any ordinary function F as long as F is continuous at zero, regardless of how it behaves elsewhere; it follows that for any ϕ that is continuous at zero, $\delta(x)\phi(x)$ is defined as a generalized function that equals $\phi(0)\delta(x)$, since for any $\psi \in T$ the function $F(x) = \phi(x)\psi(x)$ is continuos at zero and (19) applies"[3]. This is incorrect as the following example 1 shows.

To maintain relevance to (S) suppose that $\tilde{\phi}(x) = e^{-x^2}$. Then the product of generalized functions $\delta(x)\tilde{\phi}(x)$ exists, since $\tilde{\phi}$ is a good function, (L.,p.19) and provides

$$\delta(x)\tilde{\phi}(x) = \delta(x).$$

Consider the function $\tilde{\phi}(x)^{-1}$; it is continuous at 0; note that due to its fast divergence it is not an ordinary function in T'.

Example 1 *The product $\delta(x)\tilde{\phi}(x)^{-1}$ does not exist as a tempered distribution.*

To define $\delta(x)\tilde{\phi}(x)^{-1}$ as a tempered distribution we need to verify that

$$\lim_{n\to\infty} \int \delta_n(x)\tilde{\phi}(x)^{-1}\psi(x)dx$$

exists for any delta-convergent (in T') $\delta_n(x) \in T$ and arbitrary $\psi(x) \in T$.

[3]See, e.g., proof of Lemma 2 in (S) that assumes this.

Consider a sequence of infinitely differentiable functions such that:

$$\tilde{\delta}_n(x) = \begin{cases} n & \text{if } -\frac{1}{n} < x < \frac{1}{n}; \\ 0 \leq \tilde{\delta}_n(x) \leq n & \text{if } -\frac{2}{n} < x < \frac{2}{n}; \\ e^{-n} & \text{if } n - \frac{1}{n} < x < n + \frac{1}{n}; \\ 0 \leq \tilde{\delta}_n(x) \leq e^{-n} & \text{if } n - \frac{2}{n} < x < n + \frac{2}{n}; \\ 0 & \text{otherwise.} \end{cases}$$

Each $\tilde{\delta}_n(x)$ is in T (even has finite support).

$$2 + 2\frac{e^{-n}}{n} \leq \int \tilde{\delta}_n(x)dx \leq 4 + 4\frac{e^{-n}}{n}$$

and $\int \tilde{\delta}_n(x)dx = \int_{-2/n}^{2/n} \tilde{\delta}_n(x)dx + O(e^{-n}n)$. Define $\delta_n(x) = \frac{\tilde{\delta}_n(x)}{\int \tilde{\delta}_n(x)dx}$.

This $\delta_n(x)$ is a $\delta-$convergent sequence in T'. Indeed for any $\psi \in T$

$$\begin{aligned} \int \delta_n(x)\psi(x)dx &= \int_{-2/n}^{2/n} \delta_n(x)\psi(x)dx + \int_{n-2/n}^{n+2/n} \delta_n(x)\psi(x)dx \\ &= \psi(0) + O(e^{-n}n) + o(1) \to \psi(0). \end{aligned}$$

Each product $\delta_n(x)\tilde{\phi}(x)^{-1}$ is a tempered distribution since it has finite support. However, the sequence of tempered distributions $\delta_n(x)\phi(x)^{-1}$ does not converge in the space T' of tempered distributions. Write

$$\int \delta_n(x)e^{x^2}\psi(x)dx = \int_{-2/n}^{2/n} \delta_n(x)e^{x^2}\psi(x)dx + \int_{n-2/n}^{n+2/n} \delta_n(x)e^{x^2}\psi(x)dx$$
$$= \psi(0) + O(e^{-n}n) + \int_{n-2/n}^{n+2/n} \delta_n(x)e^{x^2}\psi(x)dx.$$

Now, for $\psi \in T$ such that $\psi(x) = \exp(-|x|)$ for, e.g. $|x| > 1$

$$\int_{n-2/n}^{n+2/n} \delta_n(x)e^{x^2}\psi(x)dx \geq e^{-n} \int_{n-1/n}^{n+1/n} e^{x^2-x}dx \geq \frac{2}{n}e^{-2n+(n-1/n)^2}.$$

This sequence diverges.■

This example illustrates that the product $\varepsilon_1\phi^{-1}$ may not exist in the space of tempered distributions, e.g. in the case where $E[Y\,|Z = z]$ is a constant function plus a function with some ordinary Fourier transform on R, and u has a Gaussian distribution, since then the product $\delta_n(x)\tilde{\phi}(x)^{-1}$ of Example 1 appears in the equation.

But it is not only the product with a singular function that may not exist for an ordinary function; as the next example shows the product of a function that equals zero everywhere with a continuous function divergent at infinity such as $\tilde{\phi}(x)^{-1}$ of Example 1 may not exist as a tempered distribution. Recall that any tempered distribution is actually an equivalence class of functions, and thus a function that equals zero considered as a tempered distribution has to be identified with such a class leading to a possible complication as in the example below.

Example 2 *Product of function $b(x) \equiv 0$ and $\tilde{\phi}(x)^{-1}$ does not exist as a tempered distribution.*

Consider

$$b_n(x) = \begin{cases} e^{-n} & \text{if } n - \frac{1}{n} < x < n + \frac{1}{n}; \\ 0 \le b_n(x) \le e^{-n} & \text{if } n - \frac{2}{n} < x < n + \frac{2}{n}; \\ 0 & \text{otherwise.} \end{cases}$$

This $b_n(x)$ converges to $b(x) \equiv 0$ in T'. Indeed for any $\psi \in T$

$$\int b_n(x)\psi(x)dx = \int_{n-2/n}^{n+2/n} b_n(x)\psi(x)dx \to 0.$$

But $b_n(x)\tilde{\phi}(x)^{-1}$ does not converge in the space T' of tempered distributions.For $\psi \in T$ such that $\psi(x) = \exp(-|x|)$ for, e.g. $|x| > 1$

$$\int_{n-2/n}^{n+2/n} b_n(x)e^{x^2}\psi(x)dx \ge e^{-n}\int_{n-1/n}^{n+1/n} e^{x^2-x}dx \ge \frac{2}{n}e^{-2n+(n-1/n)^2}.$$

This diverges.■

References

[1] Dufour, J-M(2003) Identification, weak instruments, and statistical inference in econometrics, Canadian Journal of Economics, v. 36, 767-808.

[2] Gel'fand, I.M. and G.E.Shilov (1964) Generalized Functions, Vol.1, Properties and Operations, Academic Press, San Diego.

[3] Gel'fand, I.M. and G.E.Shilov (1964) Generalized Functions, Vol.2, Spaces of Test functions and Generalized Functions, Academic Press, San Diego.

[4] Hausman J., W. Newey, H. Ichimura and J. Powell (1991), Measurement error in Polynomial Regression Models, Journal of Econometrics, v. 50, 273-295.

[5] Kaminski, A and R. Rudnicki (1991), A note on the convolution and the product D' and S', Internat. J. Math&Math Sci., v.14, 275-282.

[6] Lighthill, M.J. (1959) Introduction to Fourier Analysis and Generalized Functions, Cambridge University Press.

[7] Newey, W. (2001), Flexible Simulated Moment Estimation of Nonlinear Errors-in-Variables Models, Review of Economics and Statistics, v. 83, 616-627.

[8] Phillips, P.C.B. (1985) A theorem on the tail behavior of probability distributions with an application to the stable family, Canadian Journal of Economics, XVIII, 58-65.

[9] Phillips, P.C.B. (1991) A Shortcut to LAD Estimator Asymptotics, Econometric Theory, 7,450-463

[10] Phillips, P.C.B. (1995) Robust Non-Stationary Regression, Econometric Theory, 11, 912-951.

[11] Schwartz, L. (1950) Théorie des Distributions, v.1,2.. Hermann, Paris.

[12] Schwartz, L. (1954) "Sur l'impossibilité de la multiplication des distributions"C. R. Acad. Sci.,239, No. 15, 847–848.

[13] Schennach, S. (2007) Instrumental variable estimation in nonlinear errors-in-variables models, Econometrica, v.75, pp. 201-239.

[14] Sobolev, S. (1992) Cubature formulas and Modern Analysis, Gordon and Breach Science Publishers, S.A..

[15] Wang, L. and C. Hsiao (1995), Simulation-Based Semiparametric Estimation of Nonlinear Errors-in-Variables Models, working paper, University of Southern California.

[16] Zinde-Walsh, V. (2008) Kernel estimation when density may not exist, Econometric Theory, 24, forthcoming.

[17] Zinde-Walsh, V. and P. C. B. Phillips (2003) Fractional Brownian motion as a differentiable generalized Gaussian process, in K. Athreya, M. Majumdar, M. Puri and W. Waymire (eds.) Probability, Statistics and their Applications: Papers in Honor of Rabi Bhattacharya. Institute of Mathematical Statistics Lecture Notes-Monograph Series, Beachwood, Ohio, v. 41, 285-292. .