KOÇ UNIVERSITY-TÜSİAD ECONOMIC RESEARCH FORUM
WORKING PAPER SERIES

IDENTIFICATION VIA COMPLETENESS FOR DISCRETE COVARIATES AND ORTHOGONAL POLYNOMIALS

Yevgeniy Kovchegov
Nese Yıldız

Working Paper 1203
January 2012

KOÇ UNIVERSITY-TÜSİAD ECONOMIC RESEARCH FORUM
Rumelifeneri Yolu 34450 Sarıyer/Istanbul

Identification via completeness for discrete covariates and orthogonal polynomials *

Yevgeniy Kovchegov[†] Neşe Yıldız[‡]

First version: June 2009. This version: January 8, 2012.

Abstract

We solve a class of identification problems for nonparametric and semiparametric models when the endogenous covariate is discrete with unbounded support. Then we proceed with an approach that resolves a polynomial basis problem for the above class of discrete distributions, and for the distributions given in the sufficient condition for completeness in Newey and Powell (2003). Thus, in addition to extending the set of econometric models for which nonparametric or semiparametric identification of structural functions is guaranteed to hold, our approach provides a natural way of estimating these functions. Finally, we extend our polynomial basis approach to Pearson-like and Ord-like families of distributions.

KEYWORDS: nonparametric methods, identification, instrumental variables.

*Many of the results of this paper were presented as part of a larger project at University of Chicago and Cowles Foundation, Yale University, econometrics research seminar in the spring of 2010, as well as the 2010 World Congress of the Econometric Society in Shanghai. We would like to thank participants of those seminars for valuable comments and questions.

[†]Department of Mathematics, Oregon State University, Kidder Hall 368C, Corvallis, OR 97331; Email: kovchegy@math.oregonstate.edu; Phone: 541-737-1379; Fax: 541-737-0517.

[‡]Corresponding author: Department of Economics, University of Rochester, 231 Harkness Hall, Rochester, NY 14627; Email: nese.yildiz@rochester.edu; Phone: 585-275-5782; Fax: 585-256-2309.

1 Introduction

In this paper we extend the set of econometric models for which nonparametric or semiparametric identification of structural functions is guaranteed to hold by showing completeness when the endogenous covariate is discrete with unbounded support. Note that the case of discrete endogenous covariate X with unbounded support is not covered by the sufficiency condition given in [6]. In addition, using the theory of differential equations we develop a novel orthogonal polynomial basis approach for a large class of the distributions given in Theorem 2.2 in [6], and in the case of discrete endogenous covariate X for which the identification problem is solved in this paper. Our approach is new in economics and provides a natural link between identification and estimation of structural functions. We also discuss how our polynomial basis results can be extended to the case when the conditional distribution of $X|Z$ belongs to either the modified Pearson or modified Ord family. Finally, as will be remarked following formula (3) in the beginning of Section 3, our orthogonal polynomial approach in many cases provides a certain L^2-completeness condition that itself can be used as an alternative approach for obtaining identification results. Thus, our orthogonal polynomial approach extends the class of problems where identification can be solved via the L^2-completeness concept we introduce.

Our approach is motivated by the form of the conditional density (either with respect to Lebesgue or counting measure) of covariates given instruments. Using the form of this density function we can define an operator (a type of Stein operator) whose eigenfunctions are orthogonal polynomials (in covariates) under certain sufficient conditions. One could use the eigenfunctions of the Stein operator to approximate the structural functions of interest in such models. Since the conditional expectations of these orthogonal basis functions given instruments are known up to a certain function of the instruments (namely, they are polynomials in $\mu(Z)$, which will be defined below), this approach is likely to simplify estimation.

Completeness, L^2-completeness, and bounded completeness[1] are often used in nonparametric and semiparametric models to obtain identification of structural parameters of interest. A related but different L^2-completeness condition is studied in [1] and the references therein. In their L^2-completeness condition the functions considered are square integrable with respect to the distribution of covariates and included instruments whereas in our condition the functions considered have to be square integrable with respect to a different weight function that depends on covariates and included instruments. Also the methods and goals of [1] and our paper are different.

Here is the main identification result of this paper. We let X denote the endogenous random variable and $Z = (Z_1, Z_2)$ denote the instrumental variables.

Theorem 1. *Let X be a discrete random variable, with conditional distribution of $X|Z$ given by*

$$P(X = x|Z = z) := p(x|z) = t(z)s(x, z_1)\prod_{j=1}^{d}[\mu_j(z) - m_j]^{\tau_j(x,z_1)} \qquad \tau(x, z_1) \in \mathbb{Z}_+^d,$$

[1]See for example [2] or [3] for a discussion on bounded completeness.

where $t(z) > 0$, $s(x, z_1) > 0$, $\tau(x, z_1) = (\tau_1(x, z_1), \ldots, \tau_d(x, z_1))$ *is one-to-one in* x, *and the support of* $\mu(Z) = (\mu_1(Z), \ldots, \mu_d(Z))$ *given* Z_1 *contains a non-trivial open set in* $\mathbb{R}^d$, *and* $\mu_j(Z) > m_j$ $(Z - a.s.)$ *for each* $j = 1, \ldots, d$. *Then*

$$E[g(X, Z_1)|Z_1, Z_2] = 0 \quad Z - a.s. \quad \textit{implies} \quad g(X, Z_1) = 0 \quad (X, Z_1) - a.s.$$

The above theorem is a discrete counterpart to Theorem 2.2 in [6], where it was shown that if with probability one conditional on Z, the distribution of X is absolutely continuous w.r.t. Lebesgue measure, and its conditional density is given by

$$f_{X|Z}(x|z) = t(z)s(x, z_1)\exp[\mu(z) \cdot \tau(x, z_1)], \tag{1}$$

where $t(z) > 0$, $s(x, z_1) > 0$, $\tau(x, z_1)$ is one-to-one in x, and the support of $\mu(Z)$ given Z_1 contains a non-trivial open set, then for each $g(x, z_1)$ with finite expectation $E[g(X, Z_1)|Z] = 0$ $(Z - a.s.)$ implies that $g(X, Z_1) = 0 \quad (X, Z_1) - a.s.$

The condition requiring the support of $\mu(Z)$ given Z_1 to contain a nontrivial open set in $\mathbb{R}^d$ in both our Theorem 1 and Theorem 2.2 in [6] can be weakened to requiring that the support of $\mu(Z)$ given Z_1 be a countable set that is dense in a nontrivial open set in $\mathbb{R}^d$.

The paper is organized as follows. The identification result of Theorem 1 is proved in Section 2. Section 3 contains the orthogonal polynomial approach for the basis problem. Finally, Section 4 contains the concluding remarks.

2 Identification results for a class of discrete endogenous covariates X

In order to prove the identification result in Theorem 1, we need the following lemma.

Lemma 1. *Let* $T = (T_1, \ldots, T_d)$ *be a* d*-dimensional random vector such that its support contains an open ball* $\mathcal{B}_\rho(\zeta)$ *in* $\mathbb{R}^d$ *of some positive radius* ρ. *Then*

$$\sum_{j \in \mathbb{Z}_+^d} a_j T^j = 0 \quad (T - a.s.) \quad \textit{implies} \;\; a_j = 0 \;\; \textit{for all } j \in \mathbb{Z}_+^d,$$

where for $t = (t_1, \ldots, t_d) \in \mathbb{R}^d$ *and* $j = (j_1, \ldots, j_d) \in \mathbb{Z}_+^d$, *we define* $t^j := \prod_{k=1}^{d} t_k^{j_k}$.

Proof. Let us first consider the case of $d = 1$. Since $\sum_{j=0}^{\infty} a_j T^j = 0 \quad (T - a.s.)$, the real-valued interval $\mathcal{B}_\rho(\zeta) = \{\Im(z) = 0,\ \zeta - \rho < \Re(z) < \zeta + \rho\}$, with $\Im(z)$ and $\Re(z)$ denoting the imaginary and real parts of z, respectively, is contained inside the ball $B_R(0) = \{|z| < R\}$ in $\mathbb{C}$, where R is the radius of convergence of $\sum_{j=0}^{\infty} a_j z^j$. Hence there is a unique analytic function $f(z) = \sum_{j=0}^{\infty} a_j z^j$ over $B_R(0) = \{|z| < R\}$. Now since $\sum_{j=0}^{\infty} a_j T^j = 0 \quad (T - a.s.)$ and the support of T contains $\mathcal{B}_\rho(\zeta) = \{\Im(z) = 0,\ \zeta - \rho < \Re(z) < \zeta + \rho\}$, the function $f(z)$ must be equal to zero on a dense set of points in $\{\Im(z) = 0,\ \zeta - \rho < \Re(z) < \zeta + \rho\}$. Hence the function equals zero over whole $\{\Im(z) = 0,\ \zeta - \rho < \Re(z) < \zeta + \rho\}$, and therefore

$$f(z) \equiv 0 \ \text{ for all } \ |z| < R$$

We conclude that $a_j = 0$ for all $j \in \mathbb{Z}_+$.

For the general d-dimensional case, we use Lemma 2.4.1 in [4] in order to conclude that $\mathcal{B}_\rho(\zeta)$ in $\mathbb{R}^d \subset \mathbb{C}^d$ is inside the *domain of convergence* $D \subset \mathbb{C}^d$ for the series $\sum_{j \in \mathbb{Z}_+^d} a_j z^j$ over $z \in \mathbb{C}^d$, and $\sum_{j \in \mathbb{Z}_+^d} a_j z^j$ is analytic in D. Now, Theorem 2.4.5 in [4] implies the uniqueness of the analytic function $f(z) = \sum_{j \in \mathbb{Z}_+^d} a_j z^j$ over the following *Reinhardt domain* containing the origin

$$\{z = (z_1, \dots, z_d) \in \mathbb{C}^d \ : \ \exists t \in \mathcal{B}_\rho(\zeta) \subset \mathbb{R}^d \ \ s.t. \ \ |z_1| \le |t_1|, \dots, |z_d| \le |t_d|\}$$

It is important to observe that $f(z)$ is analytic in each single variable. Thus we can iterate the above one dimensional argument since $\mathcal{B}_\rho(\zeta)$ belongs to the support of T and $\sum_{j \in \mathbb{Z}_+^d} a_j T^j = 0 \quad (T - a.s.)$. Thus obtaining $f \equiv 0$ on the above Reinhardt domain, and $a_j = 0$ for all $j \in \mathbb{Z}_+^d$. □

The above lemma enables us to prove the identification Theorem 1.

Proof of Theorem 1. Suppose

$$E[g(X, Z_1)|Z_1, Z_2] = \sum_x g(x, Z_1)t(Z)s(x, Z_1)\prod_{j=1}^d [\mu_j(Z) - m_j]^{\tau_j(x, Z_1)} = 0 \qquad (Z - a.s.)$$

Thus, for a.e. $Z_1 = z_1$, letting $T = \mu(z_1, Z_2) - (m_1, \dots, m_d)$ we have

$$\sum_x g(x, z_1)s(x, z_1)\prod_{j=1}^d T_j^{\tau_j(x, z_1)} = 0 \qquad (T - a.s.)$$

as $t(z_1, Z_2) > 0$. Since $\tau(x, z_1)$ is one-to-one in x, we can re-index the above series by letting $y = \tau(x, z_1)$ as follows.

$$\sum_{y \in \mathbb{Z}_+^d} a_y T^y = 0 \qquad (T - a.s.),$$

where $a_y = g(x, z_1)s(x, z_1)$ for each given $y = \tau(x, z_1)$ in $\mathbb{Z}_+^d$.

Now, since the support of $\mu(Z)$ given $Z_1 = z_1$ contains a non-trivial open set in $\mathbb{R}^d$, the conditions of Lemma 1 are satisfied, and $a_y = 0$ for all $y \in \mathbb{Z}_+^d$. Therefore

$$g(x, z_1) = 0$$

for almost every $(X, Z_1) = (x, z_1)$ as $s(x, z_1) > 0$. □

3 Polynomial basis results

Consider the continuous conditional distributions $f_{X|Z}(x,z)$ from Theorem 2.2 in [6] with $\tau(x,z_1)=x\in\mathbb{R}^d$ for simplicity. Assume that for *a.e.* $Z_1=z_1$, $s(x,z_1)$ is differentiable in x, so that $f_{X|Z}(x,z)=s(x,z_1)t(z)e^{\mu(z)^Tx}$ for $x=(x_1,\dots,x_d)^T$ and $\mu(z)=(\mu_1(z),\dots,\mu_d(z))^T$. Also denote $\nabla_x:=\left(\frac{\partial}{\partial x_1},\dots,\frac{\partial}{\partial x_d}\right)^T$ to be the gradient operator w.r.t. x variable, and therefore $\nabla_x^T\nabla_x=\triangle_x$ to be the corresponding Laplacian operator. We differentiate $f_{X|Z}(x|z)$ to obtain

$$\nabla_x^T f_{X|Z}(x|z)=\frac{\nabla_x^T s(x,z_1)}{s(x,z_1)}f_{X|Z}(x|z)+\mu^T(z)f_{X|Z}(x|z).$$

Let $\Omega(z)$ denote the support of X given $Z=z$. For a set B, let ∂B denote the boundary of B. The following statement holds for almost every $Z=z$. For a function Q such that it is differentiable in x, and $Q(x,z_1)f_{X|Z}(x|z)=0$ for each $x\in\partial\Omega(z)$,[2] the integration by parts implies

$$E[AQ(X,Z_1)|Z]=-\mu(Z)^TE[Q(X,Z_1)|Z], \tag{2}$$

where

$$AQ(x,z_1)=\frac{1}{s(x,z_1)}\nabla^T[s(x,z_1)Q(x,z_1)]=\frac{\nabla_x^T s(x,z_1)Q(x,z_1)}{s(x,z_1)}+\nabla_x^TQ(x,z_1)$$

for Q such that $E\left[\frac{\partial}{\partial x}Q(X,Z_1)|Z\right]<\infty$.

Now, let $L^2(\mathbb{R}^d,s(x,z_1))$ denote the set of measurable g such that $\int g^2(x,z_1)s(x,z_1)\mathrm{d}x<\infty$ for given z_1. Also define the following operator[3] :

$$\mathcal{A}Q=A(\nabla_xQ)=\frac{1}{s(x,z_1)}\nabla_x^T\Big[s(x,z_1)\nabla_xQ(x,z_1))\Big]=\frac{\nabla_xs(x,z_1)\cdot\nabla_xQ(x,z_1)}{s(x,z_1)}+\triangle_xQ(x,z_1)$$

Then with standard boundary conditions (i.e. the boundary integral in the integration by parts calculation must be zero), the Sturm-Liouville operator $\mathcal{A}$ is self-adjoint with respect to the Hilbert space $L^2(\mathbb{R}^d,s(x,z_1))$. Thus its eigenvalues, λ_j, are real, and its eigenfunctions, $Q_j(x,z_1)$, solve the following Sturm-Liouville differential equation

$$\sum_{i=1}^d\frac{\partial^2Q_j(x,z_1)}{\partial x_i^2}+\frac{1}{s(x,z_1)}\sum_{i=1}^d\frac{\partial s(x,z_1)}{\partial x_i}\frac{\partial Q_j(x,z_1)}{\partial x_i}-\lambda_jQ_j(x,z_1)=0 \tag{3}$$

with the corresponding boundary conditions. See equation (6).

Assume that for *a.e.* $Z_1=z_1$, $s(x,z_1)\in C^\infty(\mathbb{R}^d)$ w.r.t. variable x, for each nonnegative integer $j=(j_1,\dots,j_d)$. If $Q_j(x,z_1)=\frac{(-1)^{j_1+\dots+j_d}}{s(x,z_1)}\frac{\partial^{j_1+\dots+j_d}}{\partial x_1^{j_1}\dots\partial x_d^{j_d}}s(x,z_1)$ are the orthogonal

[2] If $\partial\Omega(z)$ contains a point at infinity, this statement should be taken to hold in the limit. Also, as will be seen below the Q_j's, which are the eigenfunctions of the Stein operator satisfy this condition.

[3] A is the Stein-Markov operator for the distribution that has Lebesgue density equal to $\frac{s(x,z_1)}{\int s(x,z_1)\mathrm{d}x}$, and $\mathcal{A}$ is the corresponding Stein operator.

eigenfunctions in $L^2(\mathbb{R}^d, s(x, z_1))$, then their projections

$$P_j(Z) := E[Q_j(X)|Z] = \prod_{k=1}^{d} \mu_k(Z)^{j_k} = \mu(Z)^j$$

due to integration by parts under the boundary conditions requiring the corresponding boundary integral to be zero. If in addition $Q_j(x)$ are an orthogonal basis for $L^2(\mathbb{R}^d, s(x, z_1))$, then the identification problem follows immediately from Lemma 1 giving us a somewhat alternative approach to the identification result of Theorem 2.2 in [6] that works for all $g \in L^2(\mathbb{R}^d, s(x, z_1))$. In particular, using the Rodrigues' formula for the Sturm-Liouville boundary value problem, we can show that when

$$s(x, z_1) = \gamma(z_1) \exp\left[\alpha(z_1)\frac{x^T x}{2} + \beta(z_1)\right],$$

with $\alpha(z_1) < 0$ for each z_1, there is a series of eigenvalues $\lambda_0, \lambda_1, \lambda_2, ...$ that lead to solutions $\{Q_j(x, z_1)\}_{j=0}^{\infty}$, where each $Q_j(x, z_1) = \frac{(-1)^{j_1+\cdots+j_d}}{s(x,z_1)} \frac{\partial^{j_1+\cdots+j_d}}{\partial x_1^{j_1} \ldots \partial x_d^{j_d}} s(x, z_1)$ is a multidimensional Hermite-type orthogonal polynomial basis for $L^2(\mathbb{R}^d, s(x, z_1))$.[4]

3.1 The orthogonal polynomial basis results for continuous X

First we show how orthogonal polynomial approach works when $\tau(x, z_1)$ is a one-to-one function of x for *a.e.* $Z_1 = z_1$. To keep the notation simple, we are going to assume that $d = 1$ in the rest of the paper (with the exception of Example 2 below), but all the results go through when d is a finite, positive integer. When $\tau(x, z_1)$ is a one-to-one function of x,

$$\frac{\partial f_{X|Z}(x|z)}{\partial x} = \frac{\frac{\partial s(x,z_1)}{\partial x}}{s(x)} f_{X|Z}(x|z) + \mu(z)\frac{\partial \tau(x, z_1)}{\partial x} f_{X|Z}(x|z).$$

Let

$$AQ(x, z_1) = \frac{\partial}{\partial x}\left(\frac{s(x, z_1)Q(x, z_1)}{\frac{\partial \tau(x,z_1)}{\partial x}}\right)\frac{1}{s(x, z_1)} = \frac{\frac{\partial Q(x,z_1)}{\partial x}}{\frac{\partial \tau(x,z_1)}{\partial x}} + \frac{\frac{\partial s(x,z_1)}{\partial x}}{s(x, z_1)}\frac{Q(x, z_1)}{\frac{\partial \tau(x,z_1)}{\partial x}} - \frac{Q(x, z_1)\frac{\partial^2 \tau(x,z_1)}{\partial x^2}}{\left[\frac{\partial \tau(x,z_1)}{\partial x}\right]^2}.$$

Then, assuming $\frac{Q(x,z_1) f_{X|Z}(x|z)}{\frac{\partial \tau(x,z_1)}{\partial x}}$ equals 0 on $\partial\Omega(z)$ for *a.e.* $Z = z$,[5] we arrive at (2) using integration by parts. This operator A applied to the derivative of Q with respect to x gives us the following operator

$$\mathcal{A}Q(x, z_1) := \frac{\frac{\partial^2 Q(x,z_1)}{\partial x^2}}{\frac{\partial \tau(x,z_1)}{\partial x}} + \left(\frac{\frac{\partial s(x,z_1)}{\partial x}}{s(x, z_1)}\frac{1}{\frac{\partial \tau(x,z_1)}{\partial x}} - \frac{\frac{\partial^2 \tau(x,z_1)}{\partial x^2}}{\left[\frac{\partial \tau(x,z_1)}{\partial x}\right]^2}\right)\frac{\partial Q(x, z_1)}{\partial x}.$$

[4] When $s(x, z_1)$ is of this form $Q_j(x, z_1)$ are polynomials. In general equation (3) may have solutions for other $s(x, z_1)$ that are not necessarily polynomials.

[5] As before we mean this equality in the limit if $\partial\Omega(z)$ contains a point at infinity, and again, as will be shown below, the eigenfunctions of the Stein operator will satisfy this condition.

We would like to find functions Q_j and numbers λ_j such that

$$\mathcal{A}Q_j = \lambda_j Q_j. \tag{4}$$

Defining $\phi(x,z_1) := -\frac{1}{\frac{\partial\tau(x,z_1)}{\partial x}}$ and $\psi(x,z_1) := -\frac{1}{\frac{\partial\tau(x,z_1)}{\partial x}}\left[\frac{\frac{\partial s(x,z_1)}{\partial x}}{s(x,z_1)} - \frac{\frac{\partial^2\tau(x,z_1)}{\partial x^2}}{\frac{\partial\tau(x,z_1)}{\partial x}}\right]$, then each solution of the eigenvalue problem (4) satisfies

$$\phi(x,z_1)\frac{\partial^2 Q(x,z_1)}{\partial x^2} + \psi(x,z_1)\frac{\partial Q(x,z_1)}{\partial x} + \lambda Q(x,z_1) = 0. \tag{5}$$

with the following boundary conditions

$$\begin{aligned} c_1 Q(\alpha_1(z_1),z_1) + c_2\frac{\partial Q(\alpha_1(z_1),z_1)}{\partial x} &= 0 \qquad c_1^2 + c_2^2 > 0,\\ d_1 Q(\alpha_2(z_1),z_1) + d_2\frac{\partial Q(\alpha_2(z_1),z_1)}{\partial x} &= 0 \qquad d_1^2 + d_2^2 > 0, \end{aligned} \tag{6}$$

where $[\alpha_1(z_1),\alpha_2(z_1)]$ denotes the support of X conditioned on $Z_1 = z_1$. The differential equation (5) with the boundary conditions (6) is a Sturm-Liouville type problem. The solution to this problem exists when one of the three sufficient conditions listed below is satisfied. See [8] and [7].[6] Moreover, in that case, the solutions are orthogonal polynomials with respect to the weight function

$$W(x,z_1) := \frac{R(x,z_1)}{\phi(x,z_1)},$$

where $R(x,z_1) := \exp\left(\int \frac{\psi(x,z_1)}{\phi(x,z_1)}\mathrm{d}x\right) = \exp\left(\int \left[\frac{\frac{\partial s(x,z_1)}{\partial x}}{s(x,z_1)} - \frac{\frac{\partial^2\tau(x)}{\partial x^2}}{\frac{\partial\tau(x,z_1)}{\partial x}}\right]\mathrm{d}x\right) = \frac{|s(x,z_1)|}{|\partial\tau(x,z_1)/\partial x|}$, and for each j, $Q_j(x,z_1)$ is proportional to

$$\frac{1}{W(x,z_1)}\frac{\partial^j}{\partial x^j}\left(W(x,z_1)[\phi(x,z_1)]^j\right).$$

Note that this formula works for any W that is a scalar multiple of the form above. *Finally, equation (2) implies that $P_j(Z) = E[Q_j(X,Z_1)|Z]$ are* **polynomials** *of order j in $\mu(Z)$.*

Here we list sufficient conditions for eigenfunctions $\{Q_j(x,z_1)\}_{j=0}^{\infty}$ to be orthogonal polynomials in x together with the corresponding examples of continuous conditional densities $f_{X|Z}(x|z)$.

1. **Hermite-like polynomials:** ϕ is a non-zero constant, ψ is linear and the leading term of ψ has the opposite sign of ϕ. In this case, let $\phi(x,z_1) = c(z_1) \neq 0$, then

[6] [7] and [8] give results for Hermite, Laguerre and Jacobi polynomials, the other cases are obtained by defining $\tilde{x} = ax + b$ and applying the results in [7] and [8]. Also note that these conditions are sufficient for the solutions to be polynomials. Solutions that are not polynomials, but nevertheless form an orthogonal basis might exist under less restrictive conditions.

$\tau(x,z_1) = -\frac{1}{c(z_1)}x + d(z_1)$. Then, $\psi(x,z_1) = c(z_1)\frac{\frac{\partial s(x,z_1)}{\partial x}}{s(x,z_1)} = a(z_1)x + b(z_1)$. Thus, we have $\frac{\frac{\partial s(x,z_1)}{\partial x}}{s(x,z_1)} = \frac{a(z_1)}{c(z_1}x + \frac{b(z_1)}{c(z_1)}$. Let $\alpha(z_1) := a(z_1)/c(z_1)$ and $\beta(z_1) := b(z_1)/c(z_1)$, where $\alpha(z_1) < 0\ \forall z_1$, since $a(z_1)$ and $c(z_1)$ always have opposite signs. Solving for $s(x,z_1)$ we get $s(x,z_1) = \gamma(z_1)\exp\left(\alpha(z_1)x^2/2 + \beta(z_1)x\right)$.

Example 1: Suppose there is no z_1 so that $z = z_2$, and $d = 1$. Consider $f_{X|Z}(x|z) = \frac{1}{2\sqrt{\pi\sigma^2}}\exp\left(-\frac{(x-z)^2}{2\sigma^2}\right)$. Then $t(z) = \frac{1}{2\sqrt{\pi\sigma^2}}\exp\left(-z^2/2\right)$, $s(x) = \exp\left[-x^2/(2\sigma^2)\right]$, $\mu(z) = z$, $\tau(x) = x/\sigma^2$, and $W(x) = s(x)$. The orthogonal polynomials Q_j are

$$Q_j(x) = (-1)^j e^{\frac{x^2}{(2\sigma^2)}} \frac{\mathrm{d}^j}{\mathrm{d}x^j} e^{-\frac{x^2}{(2\sigma^2)}},$$

$P_j(z) = \frac{z^j}{\sigma^j}$ and $\lambda_j = -j$ for each $j > 1$.

Example 2: Suppose there is no z_1 so that $z = z_2$, and $d > 1$. For $x = (x_1,\ldots,x_d)^T$ and $z = (z_1,\ldots,z_d)^T$, let $f_{X|Z}(x|z) = \frac{\sqrt{\det M}}{(2\pi)^{\frac{d}{2}}} e^{-\frac{(x-z)^T M(x-z)}{2}}$, where M is the inverse of the variance-covariance matrix. Then $t(z) = \frac{\sqrt{\det M}}{(2\pi)^{\frac{d}{2}}} e^{-\frac{z^T Mz}{2}}$, $s(x) = e^{-\frac{x^T Mx}{2}}$, $\mu(z) = z$, $\tau(x) = Mx$, and $W(x) = s(x)$. For each nonnegative integer $j = (j_1,\ldots,j_d)$, the polynomial Q_j is given by

$$Q_j(x) = (-1)^{j_1+\cdots+j_d} e^{\frac{x^T Mx}{2}} \frac{\partial^{j_1+\cdots+j_d}}{\partial^j x_1^{j_1}\ldots\partial x_d^{j_d}} e^{-\frac{x^T Mx}{2}}.$$

Then

$$P_j(Z) = E[Q_j(X)|Z] = (m_1 Z)^{j_1}\ldots(m_d Z)^{j_d},$$

where $m_1 = e_1 M,\ldots,m_d = e_d M$ are the row vectors of M.

2. **Laguerre-like polynomials:** ϕ and ψ are both linear, the roots of ϕ and ψ are different, and the leading terms of ϕ and ψ have the same sign if the root of ψ is less than the root of ϕ or vice versa.
   Suppose $\phi(x,z_1) = a(z_1)x + b(z_1)$ and $\psi(x,z_1) = c(z_1)x + d(z_1)$ with $b(z_1)/a(z_1) \neq d(z_1)/c(z_1)$. Then

$$\frac{\partial\tau(x,z_1)}{\partial x)} = \frac{1}{-a(z_1)x - b(z_1)},$$

so

$$\tau(x,z_1) = \frac{1}{a(z_1)}\log[a(z_1)x + b(z_1)| + C(z_1).$$

Moreover,

$$\psi(x,z_1) = [a(z_1)x+b(z_1)]\frac{\frac{\partial s(x,z_1)}{\partial x}}{s(x,z_1)} + a(z_1) = c(z_1)x + d(z_1) \Leftrightarrow \frac{\frac{\partial s(x,z_1)}{\partial x}}{s(x,z_1)} = \frac{c(z_1)x + d^*(z_1)}{a(z_1)x + b(z_1)},$$

where $d^*(z_1) = d(z_1) - a(z_1)$. This means that

$$s(x, z_1) = \rho(z_1) \exp\left\{ \int \frac{c(z_1)x + d^*(z_1)}{a(z_1)x + b(z_1)} \mathrm{d}x \right\}.$$

Example: Suppose there is no z_1 so that $z = z_2$, and $d_2 = 1$. Let $\delta, r > 0$. Consider

$$f_{X|Z}(x, z) = \frac{1}{\Gamma(r+z)} \delta^{r+z} x^{r+z-1} e^{-\delta x} \quad (x > 0),$$

where $Z > -r$. Then $t(z) = \frac{1}{\Gamma(r+z)} \delta^{r+z}$, $s(x) = x^{r-1} e^{-\delta x}$, $\mu(z) = z$, and $\tau(x) = \log x$, since $x^z = e^{z \log x}$. In this case, $\phi(x) = -x$, $\psi(x) = \delta x - r$, and $W(x) = s(x)$. The orthogonal polynomials Q_j are

$$Q_j(x) = \frac{x^{-(r-1)} e^{\delta x}}{j!} \frac{\mathrm{d}^j}{\mathrm{d}x^j} \left(x^{j+r-1} e^{-\delta x} \right),$$

for $j > 1$, $P_j(z) = z(z-1) \cdots (z-n+1)$, and $\lambda_j = -\delta j$.

3. **Jacobi-like polynomials:** ϕ is quadratic, ψ is linear, ϕ has two distinct real roots, the root of ψ lies between the two roots of ϕ, and the leading terms of ϕ and ψ have the same sign.

   In this case,

$$\frac{\partial \tau(x, z_1)}{\partial x} = -\frac{1}{(x - r_1(z_1))(x - r_2(z_1))},$$

   with $r_1 \neq r_2$ and x not equal to either one of them. In this case, however, τ is not one-to-one on x, and the condition given in Theorem 2.2 of Newey and Powell does not hold unless specific support conditions are met.

   Solving the last differential equation we get

$$\tau(x, z_1) = \frac{1}{r_1(z_1) - r_2(z_1)} \left[\log|x - r_2(z_1)| - \log|x - r_1(z_1)| \right] + c(z_1).$$

   Plugging this into the formula for ψ yields

$$\psi(x, z_1) = (x - r_1(z_1))(x - r_2(z_1)) \left[\frac{\frac{\partial s(x,z_1)}{\partial x}}{s(x, z_1)} + \frac{2x - r_1(z_1) - r_2(z_1)}{(x - r_1(z_1))(x - r_2(z_1))} \right] = a(z_1)x + b(z_1).$$

   Rearranging terms gives us

$$\begin{aligned} \frac{\frac{\partial s(x,z_1)}{\partial x}}{s(x, z_1)} &= -\frac{2x - r_1(z_1) - r_2(z_1)}{(x - r_1(z_1))(x - r_2(z_1))} \\ &+ \frac{1}{r_1(z_1) - r_2(z_1)} \left[\frac{a(z_1)r_1(z_1) + b(z_1)}{x - r_1(z_1)} - \frac{a(z_1)r_2(z_1) + b(z_1)}{x - r_2(z_1)} \right] \\ &=: \kappa(x, z_1). \end{aligned}$$

Let $\alpha(x, z_1) := \int \kappa(x, z_1)\mathrm{d}x$. Then

$$\begin{aligned} \alpha(x, z_1) &= -\log|(x - r_1(z_1))(x - r_2(z_2))| \\ &+ \frac{a(z_1)r_1(z_1) + b(z_1)}{r_1(z_1) - r_2(z_1)} \log|x - r_1(z_1)| \\ &- \frac{a(z_1)r_2(z_1) + b(z_1)}{r_1(z_1) - r_2(z_1)} \log|x - r_2(z_1)|, \end{aligned}$$

and

$$s(x, z_1) = \rho(z_1) \exp[\alpha(x, z_1)].$$

Example: Suppose there is no z_1, and $f_{X|Z}(x, z) = \frac{1}{\mathcal{B}(a+z,b-z)} x^{a+z-1}(1-x)^{b-z-1}$, for $x \in (0, 1)$, where $\mathcal{B}(\cdot, \cdot)$ denotes the beta function. Suppose the following condition is satisfied:

$$\lim_{x\to 0+} x^{a+Z}Q(x) = \lim_{x\to 1-}(1 - x)^{b-Z}Q(x) = 0 \qquad Z - a.s. \tag{7}$$

We also assume the support of Z is in $(-a, b)$. Then $t(z) = \frac{\mathcal{B}(a,b)}{\mathcal{B}(a+z,b-z)}$, $s(x) = \frac{1}{\mathcal{B}(a,b)} x^{a-1}(1-x)^{b-1}$, $\mu(z) = z$ and $\tau(x) = \log\left(\frac{x}{1-x}\right)$, since $\left(\frac{x}{1-x}\right)^z = \exp\left[z \log\left(\frac{x}{1-x}\right)\right]$. Then $\phi(x) = -x(1 - x)$, $\psi(x) = (a - b)x - a$, and $W(x) = s(x)$. The orthogonal polynomial Q_j are the scaled Jacobi polynomials and satisfy the following hypergeometric differential equations of Gauss:

$$x(1 - x)Q_j'' + (a - (a + b)x)Q_j' + j(j + a + b - 1)Q_j = 0$$

for each degree $j = 0, 1, \ldots$. See Section 4.21 of [8] and [9]. These scaled Jacobi polynomials can be expressed with the hypergeometric functions

$$Q_j(x) := P_j^{(a-1,b-1)}(1 - 2x) = \frac{(\alpha)_j}{j!} \cdot {}_2F_1(-j, j + a + b - 1; a; x) ,$$

where $(\alpha)_j := \alpha(\alpha + 1)\cdots(\alpha + j - 1)$, and for $c \notin \mathbb{Z}_-$, ${}_2F_1(a, b; c; x) := \sum_{j=0}^{\infty} \frac{(a)_j(b)_j x^j}{(c)_j j!}$. Note that these Q_j's satisfy equation (7). Moreover, the eigenvalues are $\lambda_j = -j(j + a + b - 1)$ and for $j > 1$,

$$P_j(Z) = E[Q_j(X)|Z] = -\frac{Z}{\lambda_j} E[Q_j'(X)|Z].$$

3.2 The orthogonal polynomial basis results for discrete X

Here we show that the orthogonal polynomial basis results of the previous section go through when X is discrete and satisfies the conditions in Theorem 1. Suppose for simlicity X is one-dimensional with its conditional distribution given by

$$P(X = x|Z = z) := p(x|z) = t(z)s(x, z_1)[\mu(z) - m]^x \tag{8}$$

for

$$x \in a + \mathbb{Z}_+ = \{a, a+1, a+2, \dots\},$$

where $\mu(Z) > m$ *a.s.*, and a given $-\infty \le a < \infty$.

For a function h, define respectively the backwards and forwards difference operators as

$$\begin{aligned} \nabla h(x) &:= h(x) - h(x-1), \\ \Delta h(x) &:= h(x+1) - h(x). \end{aligned}$$

Let $Ah(x, z_1) := \frac{s(x-1,z_1)}{s(x,z_1)} \nabla h(x, z_1) - \left[m + \frac{s(x-1,z_1)}{s(x,z_1)} \right] h(x, z_1)$, and let $s(a-1, z_1) = 0$ for almost every $Z = z$.

Lemma 2. *Suppose g is such that $Eg(X, Z_1) < \infty$. Then*

$$E[Ag(X, Z_1)|Z] = -\mu(Z) E[g(X, Z_1)|Z] \qquad (Z - a.s.)$$

Proof.

$$\begin{aligned} E[Ag(X, Z_1)|Z] &= \sum_{x \in a+\mathbb{Z}_+} \frac{s(x-1, Z_1)}{s(x, Z_1)} [g(x, Z_1) - g(x-1, Z_1)] t(z) s(x, Z_1) [\mu(Z) - m]^x \\ &- \sum_{x \in a+\mathbb{Z}_+} \left[m + \frac{s(x-1, Z_1)}{s(x, Z_1)} \right] g(x, Z_1) t(z) s(x, Z_1) [\mu(Z) - m]^x \\ &= [m - \mu(Z)] \sum_{x \in a+\mathbb{Z}_+} g(x-1, Z_1) t(z) s(x-1, Z_1) [\mu(Z) - m]^{x-1} \\ &- m \sum_{x \in a+\mathbb{Z}_+} g(x, Z_1) t(z) s(x, Z_1) [\mu(Z) - m]^x = -\mu(Z) E[g(X, Z_1|Z]. \end{aligned}$$

□

Note that the result holds when the support of $p(x|z) = P(x = x|Z = z)$ is

$$a - \mathbb{Z}_+ = \{\dots, a-2, a-1, a\}$$

with $-\infty < a < \infty$, $Ah(x, z_1) := \frac{s(x+1,z_1)}{s(x,z_1)} \Delta h(x, z_1) - \left[m + \frac{s(x+1,z_1)}{s(x,z_1)} \right] h(x, z_1)$, and $s(a+1, z_1) = 0$ for almost every $Z = z$.

From the above lemma we see that equation (2) holds, and iterating on that equation yields

$$E[A^k g(X)|Z] = (-\mu(Z))^k E[g(X)|Z]. \tag{9}$$

The corresponding Stein operator $\mathcal{A}$ is defined as $\mathcal{A}h = A\Delta h$. The eigenfunctions of $\mathcal{A}$ are orthogonal polynomials Q_j such that

$$\mathcal{A} Q_j(x) = \lambda_j Q_j(x).$$

See [8], [7]. Then by (2) and (9) we have

$$\lambda_j E[Q_j(X)|Z] = E[A\Delta Q_j(X)|Z] = -\mu(Z)E[\Delta Q_j(X)|Z],$$

so that

$$E[Q_j(X)|Z] = \frac{-\mu(Z)}{\lambda_j} E[\Delta Q_j(X)|Z]$$

for $j > 1$. Thus, we know recursively that $P_j(Z) := E[Q_j(X)|Z]$ is an j-th degree polynomial in $\mu(Z)$.

We now present the following specific examples.

1. **Charlier polynomials:** Suppose there is no Z_1, and $X|Z$ has a Poisson distribution with density $p(x|z) = \frac{e^{-(\tilde{m}_0+z)}[\tilde{m}_0+z]^x}{x!} = e^{-z}\frac{e^{-\tilde{m}_0}\tilde{m}_0^x}{x!}\left[1+\frac{z}{\tilde{m}_0}\right]^x$, for $x \in \mathbb{N}$, so that $t(z) = e^{-z}$, $s(x) = \frac{e^{-\tilde{m}_0}\tilde{m}_0^x}{x!}$, $m_0 = 1$, and $\mu(z) = \frac{z}{\tilde{m}_o}$. Then $Ah(x) = h(x) - \frac{x}{\tilde{m}_0}h(x-1)$ is the Stein-Markov operator. The eigenfunctions of the Stein operator are the Charlier polynomials $Q_j(x) = C_j(x;\tilde{m}_0)(x) = \sum_{r=0}^{j}\binom{j}{r}(-1)^{j-r}\tilde{m}_0^{-r}x(x-1)\ldots(x-r+1)$ which are orthogonal w.r.t. Poisson-Charlier weight measure $\rho(x) := \frac{e^{-\tilde{m}_0}\tilde{m}_0^x}{x!}\sum_{k=0}^{\infty}\delta_k(x)$, where $\delta_k(x)$ equals 1 if $k = x$, and 0 otherwise. See [7]. Finally,
$P_j(Z) = E[Q_j(X)|Z] = \sum_{r=0}^{j}\sum_{x=r}^{\infty} e^{-(\tilde{m}_0+Z)}\frac{(\tilde{m}_0+Z)^x}{(x-r)!}\binom{j}{r}(-1)^{j-r}\tilde{m}_0^{-r} = \frac{Z^j}{\tilde{m}_0^j}$.

2. **Meixner polynomials:** Suppose there is no Z_1, and for $x \in \mathbb{N}$ and α an integer greater than or equal to 1, $p(x|z) = \binom{x+\alpha-1}{x}p^\alpha[1-p+\mu(z)]^x t(z)$, where $t(z) = \left[\sum_{x=0}^{\infty}\frac{\Gamma(x+\alpha)}{x!\Gamma(\alpha)}p^\alpha[1-p+\mu(z)]^x\right]^{-1}$. The above lemma applies with $s(x) = \binom{x+\alpha-1}{x}p^\alpha$, $m_0 = 1-p$. Then $Ah(x) = (1-p)h(x) - \frac{x}{x+\alpha}h(x-1)$ is the Stein-Markov operator. The eigenfunctions of the Stein operator are the Meixner polynomials $Q_j(x) = M_j(x;\alpha,p)(x) = \sum_{k=0}^{j}(-1)^k\binom{j}{k}\binom{x}{k}k!(x-\alpha)_{j-k}p^{-k}$, where $(a)_j := a(a+1)\ldots(a+j-1)$. which are orthogonal w.r.t. weight measure $\rho(x) := s(x)\sum_{k=0}^{\infty}\delta_k(x)$.

3.3 Extension to Pearson-like and Ord-like Families:

Suppose X is a random variable with either Lebesgue density or density with respect to counting measure, f, that satisfies

$$D[\phi(x)f(x)] = \psi(x)f(x),$$

where D denotes derivative when X is continuous, and the forward difference operator Δ when X is discrete, $\phi(x)$ is a polynomial of degree at most two and $\psi(x)$ is a decreasing linear function, $\phi(x) > 0$, $a < x < b$, $\phi(a), \phi(b) = 0$ if a, b is finite. This relation describes the Pearson family when X is continuous and Ord family, when X is discrete. Many continuous distributions fall into the Pearson family, and many discrete ones fall into Ord's family. See [7] and the references therein.

Suppose X is a random variable in either Pearson or Ord family. Following [7], define the Stein-Markov operator for F_X as $AQ(x) = \phi(x)D^*Q(x) + \psi(x)Q(x)$, where D^* denotes the derivative when X is continuous and the backwards difference operator, ∇, when X is discrete, and for Q such that $EQ(X), ED^*Q(X) < \infty$. Then $E[AQ(X)] = 0$. Let the Stein operator, $\mathcal{A}$, be defined as $\mathcal{A}Q := ADQ$. As before, let Q_j denote the eigenfunctions of $\mathcal{A}$. Next, suppose that the conditional distribution of X given Z is such that the Stein-Markov operator of $F_{X|Z}$ equals $A_\mu Q = \phi D^* Q + (\psi + c\mu(Z))Q$, where c is a constant. Then $E[A_\mu Q(X)|Z] = 0$. Since Q_j are eigenfunctions of $\mathcal{A}$

$$\begin{aligned} \lambda_j E[Q_j(X)|Z] &= E[\mathcal{A}Q_j(X)|Z] = E[ADQ_j(X)|Z] = E[(A - A_\mu)DQ_j(X)|Z] \\ &= -c\mu(Z)E[DQ_j(X)|Z]. \end{aligned}$$

Letting $P_j(Z) := E[Q_j(X)|Z]$ we see that P_j's are j^{th}-order polynomials in $\mu(Z)$. Thus, our main theorem applies whenever the Stein-Markov operator of $F_{X|Z}$ equals $A_\mu Q = \phi D^* Q + (\psi + c\mu(Z))Q$. The question then arises for which, if any, conditional distributions of $X|Z$ the Stein-Markov operator is of this form. It should be pointed out that this current approach extends to multidimensional discrete $X|Z$, and other types of distributions with well defined Stein-Markov operators. We now give some examples for such discrete distributions.

Examples:

1. Binomial distribution: Suppose $X|Z \sim Bin(N + \mu(Z), p)$, with $\mu(Z) \in \mathbb{Z}_+$. In this case, $\phi(x) = (1-p)x$, $\psi(x) = pN - x$, and

   $$A_\mu Q(x) = (1-p)x\nabla Q(x) + [pN + p\mu(Z) - x]Q(x)$$

   Let $Q_{-1}(x) := 0$, $Q_1(x) = 0$, and
   $Q_j(x) = K_j(x, N, p) = \sum_{l=0}^{j} (-1)^{j-l} \binom{N-x}{j-l} \binom{x}{l} p^{j-l}(1-p)^l$, the Krawtchouk polynomials, are orthogonal with respect to the binomial $Bin(N, p)$ distribution.

2. Pascal/Negative binomial distribution: Suppose

   $$P(X = x|Z = z) = p(x|z) = \binom{x + \alpha + \mu(z) - 1}{x} p^{\alpha+\mu(z)}(1-p)^x,$$

   for $x \in \mathbb{N}_+$. Then $\phi(x) = x$, $\psi(x) = (1-p)\alpha - px$, and

   $$A_\mu Q(x) = x\nabla Q(x) + [(1-p)\alpha + (1-p)\mu(Z) - px]Q(x).$$

   In this case, $Q_j = M_j(x; \alpha, p)$, where $M_j(x; \alpha, p)$ denote Meixner polynomials which were defined in the previous section and are orthogonal with respect to the Pascal distribution with parameter vector (α, p).

4 Conclusion

In this paper we solve an identification problem for nonparametric and semiparametric models in the case when the conditional distribution of X given Z belongs to the generalized power series distributions family.[7] Using an approach based on differential equations, Sturm-Liouville theory specifically, we solved orthogonal polynomial basis problem for the conditional expectation transformation , $E[g(X)|Z]$. Finally, we discussed how our polynomial basis results can be extended to the case when the conditional distribution of $X|Z$ belongs to either the modified Pearson or modified Ord family.

In deriving our results we encountered a second order differential (or difference, in the case of discrete X) equation with boundary values, which is a Sturm-Luiouville type equation. In this paper we focused on cases in which the solutions to the Sturm-Liuouville problem, which are the eigenfunctions of the operator $\mathcal{A}$, are an orthogonal polynomial basis. Our approach is more general than this. In particular, one might question for what conditional distributions the eigenfunctions of the Stein operator $\mathcal{A}$ are orthogonal basis functions, but not necessarily orthogonal polynomials. Our paper does not address this question. Addressing this question is left for future research. Finally, the work of applying the orthogonal polynomial basis approach for estimating structural functions is nearing completion.

References

[1] Andrews, D. W. K. (2011): "Examples of L^2-Complete and Boundedly-Complete Distributions," *Cowles Foundation Discussion Paper No.* 1801.

[2] Blundell, R., X. Chen, and D. Kristensen (2007): "Semi-Nonparametric IV Estimation of Shape-Invariant Engel Curves," *Econometrica*, 75, 1613-1669.

[3] d'Haultfoeuille, X. (2011): "On the Completeness Condition in Nonparameteric Instrumental Problems," *Econometric Theory*, 27, 460-471.

[4] Hörmander, L. (1973): *An Introduction to Complex Analysis in Several Variables.* Second edition. North-Holland Mathematical Library, Vol. **7**.

[5] Johnson, N. L., S. Kotz and A. W. Kemp (1992): *Univariate Discrete Distributions* Second edition. (Wiley Series in Probability and Statistics)

[6] Newey, W. K. and J. L. Powell (2003): "Instrumental Variable Estimation of Nonparametric Models," *Econometrica*, 71, 1565-1578.

[7] Schoutens, W. (2000): *Stochastic Processes and Orthogonal Polynomials.* Lecture notes in statistics (Springer-Verlag), Vol. **146**.

[8] Szegö, G. (1975): *Orthogonal Polynomials.* Fourth edition. AMS Colloquium Publications, Vol. **23**.

[9] Whittaker, E. T. and G. N. Watson (1935): *A Course of Modern Analysis.* Fourth edition. Cambridge Mathematical Library.

[7] We borrow this term from [5]