Journal of Econometrics 196 (2017) 156–179

Contents lists available at ScienceDirect

Journal of Econometrics

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

Inference in semiparametric conditional moment models with partial identification☆

Shengjie Hong *

Tsinghua University, Department of Economics, WeiLun 543, China
School of Economics and Management, Tsinghua University, Beijing, 100084, China

ARTICLE INFO

Article history:
Received 30 November 2012
Received in revised form
30 September 2016
Accepted 30 September 2016
Available online 13 October 2016

JEL classification:
C12
C14

Keywords:
Conditional moment equalities
Identified set
Sieve space
Bootstrap

ABSTRACT

This paper develops inference methods for conditional moment models in which the unknown parameter is possibly partially identified and may contain infinite-dimensional components. For a conjectured restriction on the parameter, we consider testing the hypothesis that the restriction is satisfied by at least one element of the identified set. We propose using the sieve minimum of a Kolmogorov–Smirnov type statistic as the test statistic, derive its asymptotic distribution, and provide consistent bootstrap critical values. In this way a broad family of restrictions can be consistently tested, making the proposed procedure applicable to testing the model specification and constructing confidence set for any given component or some feature of the parameter. Our methods are robust to partial identification, and allow for the moment functions to be nonsmooth. As an illustration, we apply the proposed inference methods to study the quantile instrumental variable Engel curves for gasoline in Brazil. A Monte Carlo study demonstrates finite sample performance.

1. Introduction

Partial identification analysis, pioneered by Manski (1990, 2003), has become a rapidly developing field in econometrics. Within this field, most analysis so far has been kept parametric, in the sense that (the component of) the parameter that is allowed to be unidentified must be finite dimensional. This requirement rules out potential applications involving unknown functions that may not be identified. In an effort to remove such limitation, attention very recently is turned to analysis of semi/non-parametric partial identification. An important contribution in this direction is by Santos (2012), who proposed inference methods for nonparametric instrumental variable (NPIV) models $E[Y-\theta_0(X)|Z]=0$ in which the regression function $\theta(\cdot)$ need not to be identified.[1]

In this paper, we consider models taking the form

$$E[m(Y,\ \theta_0(X))|Z]=0 \tag{1}$$

where the residual function $m(\cdot)$ is known up to an unknown function $\theta(\cdot)$, and $\theta\in\Theta$ need not to be identified. We extend Santos (2012)'s analysis to allow for $m(\cdot)$ to be non-linear or even non-smooth in θ. In addition, $m(\cdot)$ can be vector-valued in our analysis. This extension is motivated by numerous problems of partial identification, the underlying structural relations of which imply conditional moment restrictions with more complex residual functions than linear ones. For example, in nonparametric quantile IV (NPQIV) regressions, the residual function is an indicator function, which is discontinuous. This model is treated as our main motivating example. It is introduced in the following section, along with two other motivating examples that are related to binary choice dynamic panel data models and inference in incomplete information entry games, respectively. We return to all three

☆ This is a revised version of Chapter 1 in my Ph.D. thesis at the University of WI-Madison, entitled "Essays on Semiparametric Models with Partial Identification". I acknowledge continuous guidance and encouragement from my advisors Bruce Hansen and Jack Porter. I thank Andres Santos for useful comments and for sharing the Pesquisa de Orcamentos Familiares 2002–2003 data set. I also benefit from helpful discussions with Andres Aradillas-Lopez, Ivan Canay, Jianhua Huang, Qi Li, Whitney Newey, Joris Pinkse, Demian Pouzo, Xiaoxia Shi, Ya Su, Yu-Chin Hsu, Ping Yu, and participants of the 2011 Meeting of the Midwest Econometrics Group, 2012 Midwest Economics Association Annual Meeting, 2012 North American Summer Meeting of the Econometric Society, and seminars at Vanderbilt, CUHK, UW-Madison, Academia Sinica, and Tsinghua. Any errors are my own responsibility.

* Correspondence to: School of Economics and Management, Tsinghua University, Beijing, 100084, China.
E-mail address: hongshj@sem.tsinghua.edu.cn.

[1] Although Santos (2012) is focused on the NPIV, his method is applicable to conditional moment models in which the residual function is linear in $h(\cdot)$, such as the partially linear model $E\left[Y-X_1'\beta_0-h_0(X_2)|Z\right]=0$.

http://dx.doi.org/10.1016/j.jeconom.2016.09.014

examples later in the paper, but prioritize to the NPQIV one, which is the focus of our Monte Carlo simulations and empirical study.

Model (1) encompasses semi-nonparametric models of the form:

$$E\left[m(Y,\ \beta_0,\ h_0(X))|Z\right] = 0 \tag{2}$$

where β is a finite-dimensional parameter and $h(\cdot)$ is an unknown function. Indeed, any parametric component β can be regarded as a subvector of $\theta(\cdot)$ containing only constant functions. (So $\theta(\cdot) = (\beta,\ h(\cdot)) \in \mathcal{B} \times \mathcal{H}$.) In most parts of our analysis, we do not intend to develop any results specialized for the parametric component β. For this reason, as well as for the ease of notation, we treat (2) as equivalent to Model (1).

Although this paper is primarily concerned with potential partial identification, it is also closely related to a huge literature on point-identified moment models, popularized after Hansen (1982). In fact, Model (1) and its special forms have been extensively studied under point-identification assumptions. This line of work is reported in Newey and Powell (2003), Ai and Chen (2003), Horowitz and Lee (2007), Chen and Pouzo (2009, 2012), and Darolles et al. (2011), among others. In particular, Chen and Pouzo (2009, 2012) considered very similar models as ours. As it will be shown in the following sections, we borrow insights from Chen and Pouzo's analysis. However, our work differs from theirs in two major aspects: First, we allow for partial identification; Second, we focus on testing instead of estimation. Therefore, our work complements theirs.

For a conjectured restriction on θ, we develop a general procedure for testing the hypothesis that the restriction is satisfied by at least one element of the identified set of θ. This procedure is applicable to a broad family of restrictions. As a direct application, we show how to test the model specification. We also show how to construct confidence set (CS) for some feature or any given component of θ, such as CS for $\theta(\cdot)$ at a given point, and CS for the finite-dimensional component β. Throughout our analysis, we maintain the random sample assumption, i.e., we observe an independent and identically distributed (i.i.d.) sample of $(X,\ Y,\ Z)$, whose joint probability distribution is unknown.

Our CSs are designed to cover the true value with correct asymptotic probabilities. This notion of CSs was first introduced by Imbens and Manski (2004) for partial identification analysis, as opposed to an earlier notion of CSs which aims at covering the whole identified set. For instance, our CSs for β, denoted by $\mathrm{CS}_n\,(1-\alpha)$, have the following property:

$$\liminf_{n\to\infty} \inf_{\beta\in\mathcal{B}_I} Pr\left(\beta \in \mathrm{CS}_n\,(1-\alpha)\right) \geq 1-\alpha \tag{3}$$

for a targeted level $1-\alpha$, where $\mathcal{B}_I$ denotes the identified set for β. The infimum in (3) is carried out before the limit is taken, which implies uniformity of correct asymptotic coverage probability over Θ_I.

The main technical challenge of our analysis is to derive the asymptotic behavior of a Kolmogorov–Smirnov (KS) type test statistic under partial identification. In Santos (2012), analyzing the local parameter space around the identified set is an essential approach for deriving the asymptotics of their test statistic corresponding to the NPIV. This approach is still essential for our setting. Nevertheless, nontrivial changes are needed to accommodate the added generality. A crucial step in our derivation is the linearization of certain population moments with respect to (w.r.t.) the local deviation of θ from its identified set. The crux of this step is to develop regularity conditions under which the corresponding second order derivative term is $o\left(n^{-1/2}\right)$ — a result holds automatically in Santos (2012) because $m(\cdot)$ is linear in θ. To this end, we borrow the idea from Chen and Pouzo (2009, 2012) to establish sufficient convergence rate in a pseudo norm (weaker than the L^2 norm) induced by pathwise derivatives, which in turn guarantees the desired rate for the second order derivative term. Here, the convergence is in the sense that the distance between any minimizer of certain sample criterion function and the identified set approaches zero as sample size increases. Our regularity conditions allow for ill-posedness to a certain degree, but rule out severely ill-posed problems.

The remainder of this paper is organized as follows. In Section 2 we discuss motivating examples. In Section 3 we formally define the parameter space and describe the notion of test to be considered. In Section 4 we construct the test statistic, derive its asymptotic distribution, and provide bootstrap critical values. We conclude this section by discussing sufficient conditions for the required assumption to hold in three specific models. In Section 5 we study the finite sample performance by Monte Carlo simulations. In Section 6, we provide an empirical illustration to study the Brazilian gasoline quantile Engel curves. And we briefly conclude in Section 7. All mathematical proofs are gathered in Appendices A and B.

Notation. For any positive number a, $[a]$ represents the largest integer that is not larger than a; $\mathcal{X} \subseteq \mathbb{R}^{d_x}$, $\mathcal{Y} \subseteq \mathbb{R}^{d_y}$, and $\mathcal{Z} \subseteq \mathbb{R}^{d_z}$ are supports of X, Y, and Z, respectively; $\|\cdot\|_E$ represents the Euclidean norm; For any measurable function $f : X \to R$, $\|\cdot\|_{L^2}$ represents the L^2 norm $\|f\|_{L^2}^2 \equiv E\left[(f(X))^2\right]$, and $\|f\|_\infty \equiv \sup_{x\in\mathcal{X}} |f(x)|$; When $\theta(\cdot) = (\beta,\ h(\cdot))$, for any (pseudo) norm $\|\cdot\|$ defined on a functional space of h, it should be understood that $\|\theta\| \equiv \|\beta\|_E + \|h\|$[2]; For any two positive sequences $\{a_n\}$ and $\{b_n\}$, $a_n \preceq b_n$ if $a_n \leq cb_n$ for some constant $c > 0$, and $a_n \asymp b_n$ if both $a_n \preceq b_n$ and $b_n \preceq a_n$ hold.

2. Motivating examples

We now provide three motivating examples of partially identified econometric models. In all three examples, the parameters are characterized by moment restrictions with nonlinear residual functions. These moment restrictions are special cases of Model (1).

Example 1 (*NPQIV*)**.** A NPQIV model specifies that

$$E\left[1\left\{Y \leq \theta_0(X)\right\}|Z\right] = \tau \text{ for a constant } \tau \in (0,\ 1) \tag{4}$$

which is a special case of Model (1).

In Model (4), identification of $\theta(\cdot)$ is not guaranteed without making ad hoc assumptions on the data generating process (DGP), such as some completeness assumptions. For example, under the following DGP, $\theta(\cdot)$ is only partially identified:

$$Y = \sin(\pi X) + U \tag{5}$$

$$U = \frac{1}{10}\left[E\left(X^2|Z\right) - X^2\right] - \tau + \varepsilon \tag{6}$$

$$E\left(X \mid Z\right) = 0 \tag{7}$$

$$\varepsilon \mid X, Z \sim \text{Uniform}\,[0,\ 1]\,. \tag{8}$$

Lemma 2.1. *Let X have a bounded support such that $\tau - \frac{1}{10}\left[E\left(X^2\,|\,Z\right) - X^2\right] \in (0,\ 1)$. Under DGP specified by (5)–(8), for any sufficiently small λ such that $\tau - \frac{1}{10}\left[E\left(X^2\,|\,Z\right) - X^2 + \lambda X\right] \in [0,\ 1]$, $\theta(x) = \sin(x) + \lambda x$ satisfies the conditional moment restriction in Model (4). Therefore, $\theta(\cdot)$ is only partially identified.*

Example 2 (*A Binary Choice Dynamic Panel Data Model*)**.** Consider the following binary choice panel data model:

$$Y_{is} = 1\left\{Y_{is-1}\gamma_0 + X_{is}'\beta_0 + u_i \geq \varepsilon_{is}\right\} \quad (s = 1, \ldots, S) \tag{9}$$

[2] For example: $\|\theta\|_{L^2} \equiv \|\beta\|_E + \|h\|_{L^2}$