Journal of Econometrics 178 (2014) 45–56

Contents lists available at ScienceDirect

Journal of Econometrics

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

Nonparametric inference for counterfactual means: Bias-correction, confidence sets, and weak IV

Yanqin Fan [a,*], Sang Soo Park [b]

[a] Department of Economics, University of Washington, Box 353330, Seattle, WA 98195, United States

[b] Department of Economics, Korea University, Anam-dong, Sungbuk-gu, Seoul 136-701, Republic of Korea

ARTICLE INFO

Article history:
Available online 12 August 2013

JEL classification:
C14
C15
C19

Keywords:
Average treatment effect
Counterfactual mean outcome
Kernel estimation
Partial identification

ABSTRACT

This paper supplements Manski (1990) and Manski and Pepper (2000) and contributes to the literature by introducing the concept of weak IV for the partially identified mean counterfactual outcomes when an instrumental variable (IV) or a monotone instrumental variable (MIV) is available (IV or MIV assumption respectively); developing asymptotically uniformly valid confidence sets for the counterfactual mean outcomes and average treatment effects under the assumptions; correcting biases of estimates of bounds on the counterfactual mean outcomes under the assumptions. We apply the confidence sets to further examining the effect of family intactness on a child's high school graduation originally studied in Manski et al. (1992).

Published by Elsevier B.V.

1. Introduction

For bounded outcomes, Manski (1989) established sharp bounds on the counterfactual mean outcomes and average treatment effects (ATEs) known as the worst-case bounds. These bounds can be tightened when an instrumental variable (IV) is available, see Manski (1990), or when a monotone instrumental variable (MIV) is available, see Manski and Pepper (2000). As noted in Manski (1990) and Manski and Pepper (2000), in contrast to inference for the counterfactual means and ATEs based on the worst-case bounds, inference on the counterfactual mean outcomes and ATEs based on IV and MIV bounds poses technical challenges because of their sup / inf structure; see also Manski (2003). This paper supplements Manski (1990) and Manski and Pepper (2000) by introducing bias-correction and inference tools when an IV or an MIV is available.[1] In addition, we introduce the concept of weak IV and explore implications of weak IV on the estimation and inference for partially identified mean counterfactual outcomes.

The issue of constructing confidence sets (CSs) for the counterfactual mean outcomes and ATEs belongs to the recently fast growing area of inference for partially identified parameters. Much attention has been devoted to CSs for parameters defined by a finite number of moment inequalities; see e.g., Imbens and Manski (2004), Bugni (2010), Canay (2010), Chernozhukov et al. (2007), Galichon and Henry (2009), Romano and Shaikh (2008), Stoye (2009), Rosen (2008), Beresteanu and Molinari (2009), Andrews and Guggenberger (2009), Andrews and Soares (2010), and Fan and Park (2010). CSs in these works can be modified to obtain CSs for the counterfactual mean outcomes and ATEs when the IV or MIV is discrete taking a finite number of values. The modifications are entailed by the presence of nonparametric regression functions in the corresponding moment inequalities.

To construct CSs for the case of a continuous IV/MIV, we first characterize the identified interval for the counterfactual mean outcomes in terms of a continuum of conditional moment inequalities and then construct CSs for the counterfactual mean outcomes by using an integral-type criterion function. For continuous IV/MIV, our paper is related to Chernozhukov et al. (2013), Kim (2009), and Andrews and Shi (2010).[2] The set-up in Chernozhukov et al. (2013) is the same as that of this paper. Their CSs differ from ours and are applicable to the case without additional covariates besides IV/MIV. The difference between our approach and that of Chernozhukov et al. (2013) is similar to the difference between integrated conditional moment (ICM) tests (see e.g., Bierens and Ploberger, 1997, Stute, 1997) and smoothing-based model specification tests (see e.g., Hardle and Mammen, 1993, Fan, 1994, Fan

* Corresponding author.
E-mail addresses: fany88@u.washington.edu (Y. Fan), starpac@korea.ac.kr (S. Park).

[1] The first draft of this paper was completed independently of and concurrently with related work by Chernozhukov et al. (2013).

[2] All these papers are independently and concurrently done.

0304-4076/$ – see front matter. Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.jeconom.2013.08.005

and Li, 1996).[3] Kim (2009) and Andrews and Shi (2010) developed asymptotically uniformly valid CSs for parameters defined by conditional moment inequalities. When applied to our set-up, their CSs are applicable to the case without covariates besides IV/MIV.

As noted in Manski and Pepper (2000), the simple plug-in estimators of the IV and MIV bounds tend to have upward/downward bias in finite samples; see also Haile and Tamer (2003) and Krieder and Pepper (2007). For a binary IV/MIV, we introduce analytically bias-corrected estimators of the bounds and show that their first order asymptotic biases approach zero. For a general IV/MIV, analytical bias-correction is more tedious. Instead, we introduce novel simulation bias-corrected estimators of the bounds.[4]

An interesting finding of this paper is the role of the strength of the IV in the identification and inference procedures for the counterfactual mean outcomes. Motivated by the observation that even when the IV bounds are tighter than the worst-case bounds of Manski (1989), the two sets of bounds may be very close to each other, we introduce the concepts of very weak, weak, and strong IVs for partially identified mean counterfactual outcomes in Definition 2.1 in Section 2. The strength of the IV is measured by the difference between the worst-case bounds and the IV bounds. We explore implications of the presence of a weak or very weak IV on various CSs for the mean counterfactual outcomes and the asymptotic bias of the plug-in estimators of the IV bounds. For a binary IV, we show analytically that when the IV is very weak or weak, the plug-in estimators of the IV bounds have non-negligible first order asymptotic bias.

On the empirical side, we follow up the study of Manski et al. (1992) on the effect of family structure (intact, non-intact) on a child's high school graduation (yes, no). Manski et al. (1992) reported the worst-case bounds in their Table 4, categorizing the parent years of schooling covariate into three groups ($<$12 years, 12 years, $>$12 years). Using parents' schooling as IV, we estimate IV bounds and apply the CSs introduced in this paper to further examining the effect of family structure on a child's high school graduation. Compared with the worst-case bounds, the IV bounds are tighter. Based on our CSs, we are unable to reject the null hypothesis that parents' schooling is an IV.

The rest of this paper is organized as follows. In Section 2, we review IV and MIV bounds and introduce our concept of weak IV. In Section 3, we consider the finite IV/MIV case. We construct asymptotically uniformly valid CSs for the counterfactual mean outcomes, present a detailed analysis of the asymptotic bias of the plug-in estimators of the bounds, and develop bias-corrected estimators. Section 4 extends results in Section 3 to the continuous IV/MIV case. Section 5 presents our empirical results and the last section concludes. Some technical proofs are relegated to the Appendix.

Throughout this paper, we use $\Longrightarrow$ to denote weak convergence. All limits are taken as the sample size n goes to ∞ unless stated otherwise.

2. IV/MIV bounds and weak IV

This paper uses the same formal setup as Manski (1997), Manski and Pepper (2000), and Manski (2003). There is a probability space $(\mathcal{J}, \mathcal{B}, P)$ of individuals. Each member j of population $\mathcal{J}$ has observable covariate $x_j \in \mathcal{X}$ and a response function $y_j(\cdot) : \mathcal{T} \to \mathcal{Y}$ mapping the mutually exclusive and exhaustive treatments $t \in \mathcal{T}$ into outcomes $y_j(t) \in \mathcal{Y}$. Person j has a realized treatment $z_j \in \mathcal{T}$ and a realized outcome $y_j \equiv y_j(z_j)$, both of which are observable. The latent outcomes $y_j(t)$, $t \neq z_j$, are not observable. An empirical researcher learns the distribution $P(x, z, y)$ of covariates, realizes treatments, and realizes outcomes by observing a random sample of the population. The researcher's problem is to combine this empirical evidence with assumptions in order to learn about the distribution $P[y(\cdot)]$ of response functions, or perhaps the conditional distributions $P[y(\cdot)|x]$. In this paper, we focus on the mean counterfactual outcomes conditional on x, $E[y(t)|x]$,[5] $t \in \mathcal{T}$. Let $[K_L, K_U]$ denote the range of y, where $-\infty < K_L < K_U < +\infty$.

In this section, we first review the IV and MIV bounds of Manski (1990) and Manski and Pepper (2000) and then introduce the concept of a weak IV.

2.1. A review of IV/MIV bounds

Let $x = (v, w)$ and $\mathcal{X} = \mathcal{V} \times \mathcal{W}$, where $\mathcal{V} \subset R^p$ and $\mathcal{W} \subset \mathbb{R}^d$. Each value of (v, w) defines an observable subpopulation of persons. We consider two types of IV assumptions: the familiar mean-independence form of IV assumption and the monotone instrumental variable form of IV assumption introduced in Manski and Pepper (2000).

IV *Assumption*: Covariate v is an IV in the sense of mean-independence if, for each $t \in \mathcal{T}$, each value of w, and all $(u, u') \in (\mathcal{V} \times \mathcal{V})$,

$$E\left[y(t)|v = u', w\right] = E\left[y(t)|v = u, w\right].$$

MIV *Assumption*: Covariate v is an MIV in the sense of mean-monotonicity if, for each $t \in \mathcal{T}$, each value of w, and all $(u, u') \in (\mathcal{V} \times \mathcal{V})$ such that $u' \geq u$,

$$E\left[y(t)|v = u', w\right] \geq E\left[y(t)|v = u, w\right].$$

Manski (1990, 1994) provide sharp bounds on $E[y(t)|w]$ under the IV assumption and Manski and Pepper (2000) provide sharp bounds on $E[y(t)|v, w]$ under the MIV assumption. The following lemma is adapted from Proposition 2.4 in Manski (2003).

Lemma 2.1. *Let the* IV *assumption hold. Then for each $t \in \mathcal{T}$, $w \in \mathcal{W}$, we have*

$$\sup_{u\in\mathcal{V}} E\left[y(t)I\{z = t\} + K_L I\{z \neq t\}|v = u, w\right] \leq E[y(t)|w] \leq \inf_{u\in\mathcal{V}} E\left[y(t)I\{z = t\} + K_U I\{z \neq t\}|v = u, w\right].$$

In the absence of other information, this bound is sharp.

It is interesting to note that when $E[y(t)|w]$ is point-identified, the availability of an IV v according to the above IV assumption does not help identification and traditionally has been regarded as an 'irrelevant' or 'excluded' variable. Comparing the above IV bounds with the worst-case bounds on $E[y(t)|w]$ in Manski (1989)

$$E\left[y(t)I\{z = t\} + K_L I\{z \neq t\}|w\right] \leq E[y(t)|w] \leq E\left[y(t)I\{z = t\} + K_U I\{z \neq t\}|w\right]; \quad (1)$$

however, we observe that when $E[y(t)|w]$ is not point-identified, the IV has identifying power in the sense that the bounds in Lemma 2.1 are tighter than the worst-case bounds in (1), unless for each $l = L, U$, $E(\{y(t)I\{z = t\} + K_l I\{z \neq t\}\}|v = u, w)$ does not depend on u.

The following lemma is from Manski and Pepper (2000); see also Manski (2003).

[3] Fan and Li (2000) provides a detailed comparison of these two classes of tests.

[4] Chernozhukov et al. (2009) constructed a (downward/upward) median unbiased estimator of the (upper/lower) bounds.

[5] Since we do not consider unconditional mean counterfactual outcomes, we will simply refer to the conditional mean counterfactual outcomes as mean counterfactual outcomes in this paper.