Journal of Econometrics 178 (2014) 383–397

Contents lists available at ScienceDirect

Journal of Econometrics

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

Estimation and inference for distribution functions and quantile functions in treatment effect models☆

Stephen G. Donald [a], Yu-Chin Hsu [b,*]

[a] Department of Economics, University of Texas at Austin, United States
[b] Institute of Economics, Academia Sinica, Taiwan

ARTICLE INFO

Article history:
Received 16 June 2011
Received in revised form
4 October 2012
Accepted 21 March 2013
Available online 2 September 2013

JEL classification:
C01
C12
C21

Keywords:
Hypothesis testing
Stochastic dominance
Treatment effects
Propensity score

ABSTRACT

We propose inverse probability weighted estimators for the distribution functions of the potential outcomes under the unconfoundedness assumption and apply the inverse mapping to obtain the quantile functions. We show that these estimators converge weakly to zero mean Gaussian processes. A simulation method is proposed to approximate these limiting processes. Based on these results, we construct tests for stochastic dominance relations between the potential outcomes. Monte-Carlo simulations are conducted to examine the finite sample properties of our tests. We apply our test in an empirical example and find that a job training program had a positive effect on incomes.

1. Introduction

This paper is related to numerous papers in the treatment effects literature. Much of the literature to date has focused on examining the impact of a treatment by examining the effect of this treatment on certain features of the distributions of the treatment and control groups. Most papers focus on the mean treatment effects which are defined as the average treatment effect (ATE) and average treatment effect on the treated (ATT) (e.g., Rosenbaum and Rubin (1983, 1985), Heckman et al. (1997, 1998a,b), Hahn (1988), Rosenbaum (1987), Rubin and Thomas (1996) and Hirano et al. (2003, HIR hereafter)). More recently Firpo (2007a) has examined quantile treatment effects (QTE) and quantile treatment effects on the treated (QTT) which measure the difference between certain quantiles of the potential outcomes. Also, Firpo (2010) proposes estimators of the treatment effect on various measures of inequality. These include measures such as the coefficient of variation, interquartile range, Theil's index and the Gini coefficient. For recent reviews, please see Imbens (2004) and Imbens and Wooldridge (2009) among others.

In this paper we propose methods for estimating the entire distributions (CDF) of potential outcomes in a binary treatment effect model where the unconfoundedness assumption is satisfied. The estimation method is the inverse probability weighted (IPW) estimator of the CDF for potential outcomes and allows for the use of a nonparametric estimator of the propensity score. We also use these estimates and the inverse mapping to estimate the quantile processes associated with the distributions. All our estimators are treated as functions and our asymptotic analysis of the estimators shows that, under regularity assumptions, the estimators converge weakly to Gaussian stochastic processes at the usual parametric rate. Since it is well known that the covariance kernel of the associated Gaussian processes depend on the estimation error in the estimated propensity score, we propose simulation methods, based on the multiplier central limit theorem, for conducting inference that take this into account. We also propose estimators for the CDFs and quantile functions of the potential outcomes in the *treated* subpopulation and discuss simulation methods for conducting inference. To the best of our knowledge, this is the first paper to estimate the CDFs and the whole quantile functions of the potential outcomes under the unconfoundedness assumptions and to propose a method to approximate the limiting processes of the estimators. To demonstrate the usefulness of our results we propose Kolmogorov–Smirnov (KS) type tests for stochastic dominance relations between the distributions of potential outcomes. The importance of such tests has been

☆ A previous version was circulated under the title "Testing for Stochastic Dominance in Treatment Effects".

* Correspondence to: 128 Academia Road, Section 2, Nankang, Taipei, 115, Taiwan. Tel.: +886 2 27822791x322; fax: +886 2 2785 3946.
E-mail address: ychsu@econ.sinica.edu.tw (Y.-C. Hsu).

0304-4076/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jeconom.2013.03.010

discussed by Imbens and Wooldridge (2009). Our paper also examines the properties of the proposed methods in finite samples by conducting a small scale simulation as well as applying the methods for the KS tests to examine the effect of a job training program in an empirical example using data from the National Supported Work Demonstration (NSW) job training program. In the empirical example we find that real earnings under job training stochastically dominates real earnings without job training and we find evidence against the opposite relationship. Thus our methods suggest that positive effects of job training are felt over the entire distribution which is a stronger conclusion than one would reach based simply on the mean.

This paper is related to some other recent papers that examine distributional effects in treatment effect models. Abadie (2002) constructs tests for stochastic dominance relationship between potential outcomes when the treatment assignment is endogenous and one has a binary instrument. Maier (2011) focuses on KS tests for the stochastic dominance relation in a treatment effect situation similar to ours but uses a different method for inference based on the bootstrap. Lee and Whang (2009), and Hsu (2011) consider the null hypothesis that the stochastic dominance relation between the potential outcomes holds conditional on the all possible values of the covariances. Shen (2011) proposes nonparametric and semiparametric tests for conditional stochastic dominance relation between the potential outcomes at a given covariate value. Their focus is different from ours, because we focus on the stochastic dominance relation between the unconditional CDFs. As noted above Firpo (2007a) examines quantile treatment effects for specific quantiles while Firpo (2010) examines specific measures of inequality. Our paper provides results for the entire CDF and quantile process and can be used in principle, along with functional delta methods, to construct any functional of either and conduct inference with KS tests for stochastic dominance being one empirically important example. Additionally our paper provides results for the treated subpopulation.

This paper is also related to numerous papers in the stochastic dominance literature. Papers that consider inference for stochastic dominance between two distributions include McFadden (1989), Anderson (1996), Davidson and Duclos (2000), Barrett and Donald (2003), Linton et al. (2005), Bennett (2009), Linton et al. (2010), Donald and Hsu (2010) and Donald et al. (2012) among others. Our paper considers the empirically important case where one comparing two distributions in the treatment effect context where selection to treatment depends on observables and does not rely on completely random assignment as would be required in order to use results of Barrett and Donald's (2003).[1]

The remainder of this paper is organized as follows. In Section 2, we discuss the treatment effect model. We also examine the identification of the CDFs of potential outcomes and propose estimators of the CDFs as well as the associated quantile functions. Section 3 presents regularity conditions and asymptotic results for the estimators of the CDFs and quantile functions. In Section 4, we introduce a simulation method to approximate the limiting stochastic processes of the estimated CDFs and quantile functions. We propose KS tests for the stochastic dominance relation between the distributions of potential outcomes and show the size and power properties of our tests in Section 5. We extend our results to the treated group in Section 6. Monte-Carlo simulation results are summarized in Section 7. Section 8 presents the empirical application, and Section 9 concludes. All mathematical proofs are contained in the Appendix.

[1] Linton et al. (2005, 2010) and Donald et al. (2012) allow for covariates in their model in a different way. In their framework, the variable of interest, say Y, is allowed to be a function of covariates. That is, Y can be defined as $Y = g(X, \theta)$, where X denotes a vector of covariates and $g(\cdot, \theta)$ is a real-valued function known up to the finite dimensional parameter θ. However, the dependence of the selection to treatment on covariates is not allowed in their model.

2. Estimating CDFs and quantile functions of the potential outcomes

2.1. Model framework

We consider the standard treatment effects framework. Let T be a dummy variable such that $T = 1$ if the individual receives treatment; otherwise, $T = 0$. Define $Y(1)$ as the potential outcome for the individual under treatment and $Y(0)$ as that without treatment. Let $F_0(\cdot) = P(Y(0) \leq \cdot)$ and $F_1(\cdot) = P(Y(1) \leq \cdot)$ denote the unconditional CDFs of $Y(0)$ and $Y(1)$ respectively. We observe T, $Y = T \cdot Y(1) + (1-T) \cdot Y(0)$, and X, a vector of covariates. We have a random sample of size N.

2.2. Identification of F_0 and F_1

In the treatment effect model, we never observe $Y(0)$ and $Y(1)$ simultaneously, so there is an identification problem due to the missing variable. We impose the unconfoundedness assumption (Rosenbaum and Rubin, 1983) to obtain identification.

Assumption 2.1 (*Unconfoundedness Assumption*). $(Y(0), Y(1)) \perp T|X$.

The unconfoundedness assumption requires that conditional on the observed individual characteristics, the treatment assignment is independent of the potential outcomes. Let $p(x) = P(T = 1|X = x)$ denote the propensity score function (Rosenbaum and Rubin, 1983, 1985). Under Assumption 2.1, $F_1(z)$ and $F_0(z)$ are identified by

$$\begin{aligned} F_0(z) &= E\left[\frac{(1-T)\cdot 1(Y \leq z)}{1-p(X)}\right], \\ F_1(z) &= E\left[\frac{T \cdot 1(Y \leq z)}{p(X)}\right]. \end{aligned} \tag{1}$$

The identification results in (1) are similar to those of $E[Y(1)]$ and $E[Y(0)]$ in HIR after we replace Y with indicator function $1(Y \leq z)$.

2.3. Estimation of F_0 and F_1

Based on (1), the IPW estimators for $F_0(z)$ and $F_1(z)$ are:

$$\widehat{F}_0(z) = \sum_{i=1}^{N} \frac{(1-T_i)\cdot 1(Y_i \leq z)}{1-\hat{p}(X_i)} \Big/ \sum_{i=1}^{N} \frac{1-T_i}{1-\hat{p}(X_i)}, \tag{2}$$

$$\widehat{F}_1(z) = \sum_{i=1}^{N} \frac{T_i \cdot 1(Y_i \leq z)}{\hat{p}(X_i)} \Big/ \sum_{i=1}^{N} \frac{T_i}{\hat{p}(X_i)}, \tag{3}$$

where $\hat{p}(X_i)$ is a nonparametric estimator for $p(x)$. As in HIR, we use the Series Logit Estimator (SLE) to estimate $p(x)$ based on power series. Let $\lambda = (\lambda_1, \ldots, \lambda_r)' \in \mathbb{Z}_+^r$ be a r-dimensional vector of non-negative integers where $\mathbb{Z}_+$ denotes the set of non-negative integers, and define the norm for λ as $|\lambda| = \sum_{j=1}^{r} \lambda_j$. Let $\{\lambda(k)\}_{k=1}^{\infty}$ be a sequence including all distinct $\lambda \in \mathbb{Z}_+^r$ such that $|\lambda(k)|$ is non-decreasing in k and let $x^{\lambda} = \prod_{j=1}^{r} x_j^{\lambda_j}$. For any integer K, define $R^K(x) = (x^{\lambda(1)}, \ldots, x^{\lambda(K)})'$ as a vector of power functions. Let $\mathcal{L}(a) = \exp(a)/(1+\exp(a))$ be the logistic CDF. The SLE for $p(X_i)$ is defined as $\hat{p}(x) = \mathcal{L}\left(R^K(x)'\hat{\pi}_K\right)$ where

$$\begin{aligned} \hat{\pi}_K = \arg\max_{\pi_k} \frac{1}{N} \sum_{i=1}^{N} \Big(&T_i \cdot \ln \mathcal{L}\left(R^K(X_i)'\pi_K\right) \\ &+ (1-T_i)\cdot \ln\left(1-\mathcal{L}\left(R^K(X_i)'\pi_K\right)\right)\Big). \end{aligned}$$