TECHNICAL WORKING PAPER SERIES

INSTRUMENTAL VARIABLES, SELECTION MODELS, AND TIGHT BOUNDS ON THE AVERAGE TREATMENT EFFECT

James J. Heckman
Edward J. Vytlacil

Technical Working Paper 259
http://www.nber.org/papers/T0259

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2000

This research was supported by NIH:R01-HD3498-01, NIH:R01-HD3208-03, NSF97-09-873, the Donner Foundation, and research support from the American Bar Foundation. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

Instrumental Variables, Selection Models, and Tight Bounds
on the Average Treatment Effect
James J. Heckman and Edward J. Vytlacil
NBER Technical Working Paper No. 259
August 2000
JEL No. C50, H43

ABSTRACT

This paper exposits and relates two distinct approaches to bounding the average treatment effect. One approach, based on instrumental variables, is due to Manski (1990, 1994), who derives tight bounds on the average treatment effect under a mean independence form of the instrumental variables (IV) condition. The second approach, based on latent index models, is due to Heckman and Vytlacil (1999, 2000a), who derive bounds on the average treatment effect that exploit the assumption of a nonparametric selection model with an exclusion restriction. Their conditions imply the instrumental variable condition studied by Manski, so that their conditions are stronger than the Manski conditions. In this paper, we study the relationship between the two sets of bounds implied by these alternative conditions. We show that: (1) the Heckman and Vytlacil bounds are tight given their assumption of a nonparametric selection model; (2) the Manski bounds simplify to the Heckman and Vytlacil bounds under the nonparametric selection model assumption.

James J. Heckman
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
and NBER
j-heckman@uchicago.edu

Edward J. Vytlacil
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
e-vytlacil@uchicago.edu

1 Introduction

A basic problem in evaluating social programs is that we do not observe the same individual in both the treated and untreated state at the same time. A variety of econometric assumptions are invoked to undo the consequences of this missing data. The traditional approach to this problem is to invoke sufficient assumptions about outcome equations, treatment selection equations, and their interrelationship to point identify the treatment parameters. A more recent approach to identification of treatment effects is to conduct sensitivity or bounding analyses to present ranges of estimates for estimated treatment parameters.

This paper exposits and relates two distinct approaches to bounding the average treatment effect. One approach, based on instrumental variables, is due to Manski (1990, 1994), who derives tight bounds on the average treatment effect under a mean independence form of the instrumental variables (IV) condition.[1] The second approach, based on latent index models, is due to Heckman and Vytlacil (1999,2000a), who derive bounds on the average treatment effect that exploit the assumption of a nonparametric selection model with an exclusion restriction. Their conditions imply the instrumental variable condition studied by Manski, so that their conditions are stronger than the Manski conditions. In this paper, we study the relationship between the two sets of bounds implied by these alternative conditions. We show that: (1) the Heckman and Vytlacil

[1]Manski also refers to this condition as a level-set restriction. See Robins (1989) and Balke and Pearl (1997) for bounds that exploit a statistical independence version of the instrumental variables assumption. See Manski and Pepper (2000) for bounds that exploit a weakened version of the instrumental variables assumption. Heckman, Smith and Clements (1997) consider bounds on the distribution of treatment effects in a randomized experiment. See Heckman and Vytlacil (2000b) for a discussion of alternative approaches to the evaluation of treatment effects, including a survey of the bounding literature.

bounds are tight given their assumption of a nonparametric selection model; (2) the Manski bounds simplify to the Heckman and Vytlacil bounds under the nonparametric selection model assumption.

This paper is organized in the following way. In Section 2, we introduce notation and the basic framework. We review the Manski IV bounds in Section 3, and review the Heckman and Vytlacil nonparametric selection model bounds in Section 4. In Section 5, we show that the Heckman and Vytlacil bounds are tight under the nonparametric selection model assumption. We compare the Manski bounds to the Heckman and Vytlacil bounds in Section 6, and show that the Manski bounds simplify to the Heckman and Vytlacil bounds under the nonparametric selection model assumption. The paper concludes in Section 7 by relating the analysis of this paper to the analysis of Balke and Pearl (1997).

2 Switching Regression Framework

For each person i, we observe (Y_i, D_i, W_i), where Y_i is the outcome variable, D_i is an indicator variable for receipt of treatment, and W_i is a vector of covariates. We assume that the outcome variable is generated by a switching regression,

$$Y_i = D_i Y_{1i} + (1 - D_i) Y_{0i},$$

where Y_{0i} is the potential outcome if the individual does not receive treatment and Y_{1i} is the potential outcome if the individual does receive treatment. Y_{1i} is observed if $D_i = 1$ but not otherwise; Y_{0i} is observed if $D_i = 0$ but not otherwise. We assume access to an i.i.d. sample, and henceforth supress the i subscript. For any random variable A, we will use $\mathcal{A}$ to denote the support of A, a to denote a potential realization of A, and F_A to denote the distribution function for A. In this paper, we will maintain the assumption that the outcome variables are bounded with probability one:

Assumption B. For $j = 0, 1$, and for a.e. $w \in \mathcal{W}$, there exists $y^l_{w,j}, y^u_{w,j} \in \Re$ such that:

$$\Pr(y^l_{w,j} \leq Y_j \leq y^u_{w,j} | W = w) = 1.$$

In this paper, we examine bounds on the average treatment effect (ATE), defined for $w \in \mathcal{W}$ as[2]

$$E(Y_1 - Y_0 | W = w).$$

By the law of iterated expectations:

$$\begin{aligned}
&E(Y_1 - Y_0|W = w) \\
&\quad = \Big[\Pr[D = 1|W = w]E(Y_1|W = w, D = 1) + \Pr[D = 0|W = w]E(Y_1|W = w, D = 0)\Big] \\
&\qquad - \Big[\Pr[D = 1|W = w]E(Y_0|W = w, D = 1) + \Pr[D = 0|W = w]E(Y_0|W = w, D = 0)\Big].
\end{aligned}$$

The central identification problem in recovering this parameter from observational samples is that we do not observe Y_0 for individuals with $D = 1$, and we do not observe Y_1 for individuals with

[2]Another potential parameter of interest is the effect of treatment on the treated, $E(Y_1 - Y_0|W = w, D = 1)$. Heckman and Vytlacil (1999,2000a) construct bounds for the treatment on the treated parameter given the nonparametric selection model assumption. Manski's analysis can be easily extended to this parameter as well. One can extend the results of this paper to show that the Heckman and Vytlacil bounds on treatment on the treated are tight given the assumption of a nonparametric selection model, and to show that the Manski bounds adapted to the treatment on the treated parameter simplify to the Heckman and Vytlacil bounds on the treatment on the treated parameter under the assumption of a nonparametric selection model.

$D = 0$. Thus, we can identify $\Pr[D = 1|W = w]$, $E(Y_1|W = w, D = 1)$, and $E(Y_0|W = w, D = 0)$, but cannot identify the counterfactual means $E(Y_1|W = w, D = 0)$ or $E(Y_0|W = w, D = 1)$.

Assumption (B) immediately implies that $E(Y_1|W = w, D = 0)$ and $E(Y_0|W = w, D = 1)$ are bounded, and thus we can follow Manski (1989) and Robins (1989) in bounding the ATE parameter as follows,[3]

$$B_w^L \leq E(Y_1 - Y_0|W = w) \leq B_w^U,$$

where

$$\begin{aligned}
B_w^L &= \Big[\Pr[D = 1|W = w]E(Y_1|D = 1, W = w) + \Pr[D = 0|W = w]y_{w,1}^l\Big] \\
&\quad - \Big[\Pr[D = 0|W = w]E(Y_0|D = 0, W = w) + \Pr[D = 1|W = w]y_{w,0}^u\Big] \\
B_w^U &= \Big[\Pr[D = 1|W = w]E(Y_1|D = 1, W = w) + \Pr[D = 0|W = w]y_{w,1}^u\Big] \\
&\quad - \Big[\Pr[D = 0|W = w]E(Y_0|D = 0, W = w) + \Pr[D = 1|W = w]y_{w,0}^l\Big].
\end{aligned}$$

For every value in the interval $[B_w^L, B_w^u]$, one can trivially construct a distribution of (Y_1, Y_0, D, W) which is consistent with the observed distribution of (Y, D, W) and such that the average treatment effect equals the specified value. Thus, every point in the interval $[B_w^L, B_w^u]$ must be contained in any bounds on the average treatment effect, and thus these bounds are tight under the given

[3] Smith and Welch (1986) construct analogous bounds on $E(Y_1)$ using the law of iterated expectations and the restriction that $\frac{1}{2}E(Y_1|W = w, D = 1) \leq E(Y_1|W = w, D = 0) \leq E(Y_1|W = w, D = 1)$, where the lower bound is assumed to be known a priori.

information structure. Note that the width of the bounds is

$$\Pr[D=1|W=w](y^u_{w,0}-y^l_{w,0})+(1-\Pr[D=1|W=w])(y^u_{w,1}-y^l_{w,1}).$$

Note that the width of the bounds depends only on $\Pr[D=1|W=w]$ and $y^u_{w,j}, y^l_{w,j}$, $j=0,1$.

3 Bounds Under an IV Condition

We first review the analysis of Manski (1990).[4] Partition W as $W=[X,Z]$, where Z denotes the instrument(s). He considers identification or bounding of the average treatment effect under a mean-independence form of the IV assumption:

Assumption IV. $E(Y_j|X,Z)=E(Y_j|X)$ for $j=0,1$.

Note that Assumption IV immediately implies that the average treatment effect depends only on X, $E(Y_1-Y_0|X=x,Z=z)=E(Y_1-Y_0|X=x)$. Let $\mathcal{Z}_x$ denote the support of Z conditional on $X=x$. Let

$$P(z,x)=\Pr[D=1|Z=z,X=x].$$

[4] See also Manski (1994) for a further development of these bounds.

Using the law of iterated expectations and the assumption that $E(Y_1|X,Z) = E(Y_1|X)$, for any x in the support of X and $z \in \mathcal{Z}_x$,

$$P(z,x)E(Y_1|D=1,X=x,Z=z) + (1-P(z,x))y^l_{(x,z),1}$$
$$\leq E(Y_1|X=x) \leq P(z,x)E(Y_1|D=1,X=x,Z=z) + (1-P(z,x))y^u_{(x,z),1}.$$

Since these bounds hold for all $z \in \mathcal{Z}_x$, we have

$$\sup_{z\in\mathcal{Z}_x}\{P(z,x)E(Y_1|D=1,X=x,Z=z) + (1-P(z,x))y^l_{(x,z),1}\}$$
$$\leq E(Y_1|X=x) \leq \inf_{z\in\mathcal{Z}_x}\{P(z,x)E(Y_1|D=1,X=x,Z=z) + (1-P(z,x))y^u_{(x,z),1}\}.$$

Following the parallel argument for $E(Y_0|X=x)$, Manski derives the the following sharp bounds on the average treatment effect under the mean independence assumption:

$$I^L_x \leq E(Y_1 - Y_0|X=x) \leq I^U_x,$$

with

$$I^L_x = \sup_{z\in\mathcal{Z}_x}\{P(z,x)E(Y_1|D=1,X=x,Z=z) + (1-P(z,x))y^l_{(x,z),1}\}$$
$$-\inf_{z\in\mathcal{Z}_x}\{(1-P(z,x))(E(Y_0|D=0,X=x,Z=z) + P(z,x)y^u_{(x,z),0}\}.$$

$$I_x^U = \inf_{z \in \mathcal{Z}_x} \{P(z,x)E(Y_1|D=1, X=x, Z=z) + (1-P(z,x))y^u_{(x,z),1}\}$$
$$- \sup_{z \in \mathcal{Z}_x} \{(1-P(z,x))(E(Y_0|D=0, X=x, Z=z) + P(z,x)y^l_{(x,z),0}\},$$

Let $\mathcal{P}_x$ denote the support of $P(Z,X)$ conditional on $X=x$. Let $p_x^u = \sup \mathcal{P}_x$ and $p_x^l = \inf \mathcal{P}_x$. The width of the bounds is $I_x^U - I_x^L$, a complicated expression to evaluate. Note that the above bounds exactly identify the average treatment effect if $I_x^U = I_x^L$. A trivial modification of Corollary 1 and Corollary 2 of Proposition 6 of Manski (1994) shows that, under assumptions (B) and (IV),

(i) $p_x^u \geq \frac{1}{2}$ and $p_x^l \leq \frac{1}{2}$ is a necessary condition for $I_x^L = I_x^U$.

(ii) If $Y \perp\!\!\!\perp D|X$, then $p_x^u = 1$, $p_x^l = 0$ is a necessary and sufficient condition for $I_x^L = I_x^U$.

Note that it is neither necessary nor sufficient for $P(z,x)$ to be a nontrivial function of z for these bounds to improve upon the $[B_w^L, B_w^U]$ bounds of Section 2. Evaluating the bounds and the width of the bounds for a given x requires knowledge of $P(z,x)$, $E(Y_1|D=1, X=x, Z=z)$, $E(Y_0|D=1, X=x, Z=z)$, and $y^l_{(x,z),j}, y^u_{(x,z),j}$, $j=0,1$, for each $z \in \mathcal{Z}_x$.

4 Bounds Under the Nonparametric Selection Model

We now review the analysis of Heckman and Vytlacil (1999,2000a). They use a nonparametric selection model to identify or bound the average treatment effect, where the nonparametric selection model is defined through the following assumption:

Assumption S $D = 1[\mu(Z,X) \geq U]$, with $Z \perp\!\!\!\perp (U, Y_0, Y_1)|X$.

This is clearly a stronger assumption than Assumption IV because of the treatment assignment rule, because of the independence (rather than mean independence) between Z and (Y_0, Y_1) given X, and because of the assumed independence between U and Z given X. Without loss of generality, they impose the normalization that $\mu(z,x) = P(z,x)$ so that $\Pr[U \leq P(Z,X)|Z = z, X = x] = P(z,x)$. Note that $Z \perp\!\!\!\perp (Y_0, Y_1)|X$ immediately implies that the average treatment effect depends only on X, $E(Y_1 - Y_0|X = x, Z = z) = E(Y_1 - Y_0|X = x)$, and that $y^k_{(x,z),j} = y^k_{x,j}$ for $j = 0, 1$, $k = u, l$.

Note that $DY = DY_1$ is an observed random variable, and thus for any $x \in \mathrm{Supp}(X)$, $p \in \mathcal{P}_x$, we identify the expectation of DY_1 given $X = x, P(Z,X) = p$,

$$\begin{aligned} E\left(DY_1 \middle| X = x, P(Z,X) = p\right) &= E(Y_1|X = x, P(Z,X) = p, D = 1)p \\ &= E(Y_1|X = x, P(Z,X) = p, P(Z,X) \geq U)p \\ &= E(Y_1|X = x, p \geq U)p \\ &= \int_0^p E(Y_0|X = x, U = u)dF_{U|X}(u|x), \end{aligned} \tag{1}$$

where the third equality follows from $Z \perp\!\!\!\perp (U, Y_0, Y_1) \mid X$, and the fourth equality follows from the law of iterated expectations. By similar reasoning,

$$E\left((1-D)Y_0 \middle| X = x, P(Z,X) = p\right) = \int_p^1 E(Y_0|X = x, U = u)dF_{U|X}(u|x). \tag{2}$$

We can evaluate (1) at $p = p^u_x$ and evaluate (2) at $p = p^l_x$. The distribution of (D, Y, X, Z) contains no information on $\int_{p^u_x}^1 E(Y_1|X = x, U = u)dF_{U|X}(u|x)$ and $\int_0^{p^l_x} E(Y_0|X = x, U = u)dF_{U|X}(u|x)$, but we can bound these quantities:

$$
\begin{aligned}
(1-p_x^u)y_{x,1}^l \quad &\leq \quad \int_{p_x^u}^1 E(Y_1|X=x,U=u)dF_{U|X}(u|x) \quad \leq \quad (1-p_x^u)y_{x,1}^u \\
p_x^l y_{x,0}^l \quad &\leq \quad \int_0^{p_x^l} E(Y_0|X=x,U=u)dF_{U|X}(u|x) \quad \leq \quad p_x^l y_{x,0}^u,
\end{aligned}
\tag{3}
$$

where we use the fact that $\Pr[U > p_x^u|X = x] = 1 - p_x^u$, and $\Pr[U \leq p_x^l|X = x] = p_x^l$. Since $Z \perp\!\!\!\perp (Y_0, Y_1) \mid X$, it follows that $E(Y_1 - Y_0|X = x, Z = z) = E(Y_1 - Y_0|X = x)$. These inequalities allow Heckman and Vytlacil to bound $E(Y_1 - Y_0|X = x)$ as in the following way:

$$
S_x^L \leq E(Y_1 - Y_0|X = x) \leq S_x^u,
$$

where

$$
\begin{aligned}
S_x^L \quad &= \quad p_x^u \Big[E(Y_1|X=x, P(Z,X)=p_x^u, D=1) \Big] + (1-p_x^u)y_{x,1}^l \\
&\qquad -(1-p_x^l)\Big[E(Y_0|X=x, P(Z,X)=p_x^l, D=0) \Big] - p_x^l y_{x,0}^u, \\
S_x^U \quad &= \quad p_x^u \Big[E(Y_1|X=x, P(Z,X)=p_x^u, D=1) \Big] + (1-p_x^u)y_{x,1}^u \\
&\qquad -(1-p_x^l)\Big[E(Y_0|X=x, P(Z,X)=p_x^l, D=0) \Big] - p_x^l y_{x,0}^l.
\end{aligned}
$$

The width of the bounds is

$$
S_x^U - S_x^L = (1-p_x^u)(y_{x,1}^u - y_{x,1}^l) + p_x^l(y_{x,0}^u - y_{x,0}^l).
$$

Trivially, $p_x^u = 1$, $p_x^l = 0$ is necessary and sufficient for $S_x^L = S_x^U$.[5] Note that it is both necessary and sufficient for $P(z, x)$ to be a nontrivial function of z for these bounds to improve upon the $[B_w^L, B_w^U]$ bounds of Section 2. Evaluating the width of the bounds for a given x requires knowledge only of p_x^l, p_x^u, and $y_{x,j}^l, y_{x,j}^u$, $j = 0, 1$. The only additional information required to evaluate the bounds for a given x is $E(Y_0|X = x, P(Z, X) = p_x^l, D = 0)$ and $E(Y_1|X = x, P(Z, X) = p_x^u, D = 1)$. The simpler structure for the Heckman-Vytlacil bounds compared to the Manski bounds is a consequence of the selection model structure imposed by Heckman and Vytlacil.

5 Tight Bounds

We now show that the Heckman and Vytlacil bounds are tight given the assumption that the outcomes are bounded (Assumption B) and the nonparametric selection model (Assumption S).

Theorem 1 *Impose the nonparametric selection model, Assumption S, and impose that the outcome variables are bounded, Assumption B. Then the Heckman-Vytlacil bounds on ATE are tight.*

Proof.

The logic of the proof is as follows. We show that the Heckman-Vytlacil bounds are tight by showing that for any point $s \in [S_x^L, S_x^U]$, there exists a distribution with the following properties: (i) the distribution is consistent with the observed data; (ii) the distribution is consistent with all of the Heckman-Vytlacil assumptions; and (iii) $E(Y_1 - Y_0|X)$ evaluated

[5] That $p_x^u = 1$, $p_x^l = 0$ is sufficient for point identification of the average treatment effects is shown by Heckman (1990).

under the distribution equals s. Thus, the point s must be contained in any bounds on the average treatment effect. Since this holds for every $s \in [S_x^L, S_x^U]$, we have that the interval $[S_x^L, S_x^U]$ must be contained in any bounds on the the average treatment effect, and thus $[S_x^L, S_x^U]$ are tight bounds on the average treatment effect. We prove the existence of such a distribution by constructing one that conforms to conditions (i)-(iii) for any given $s \in [S_x^L, S_x^U]$.

For any random variable A, let F_A^0 denote the "true" CDF of A, and let $F_{A|B}^0(\cdot|b)$ denote the true CDF of A conditional on $B = b$. Let s denote any given element of $[S_x^L, S_x^U]$. Note that any element $s \in [S_x^L, S_x^U]$ can be written as

$$\begin{aligned} s = p_x^u \Big[E(Y_1|X = x, P(Z,X) = p_x^u, D = 1) \Big] + (1 - p_x^u) q_x^1 \\ - (1 - p_x^l) \Big[E(Y_0|X = x, P(Z,X) = p_x^l, D = 0) \Big] - p_x^l q_x^0 \end{aligned}$$

for some q_x^0, q_x^1 s.t. $y_x^l \leq q_x^j \leq y_x^u$, $j = 0, 1$.

For $(u, x) \in \text{Supp}(U, X)$, define

$$F_{Y_1|U,X}(y_1|u,x) = \begin{cases} F^0_{Y_1|U,X}(y_1|u,x) & \text{if } u \leq p_x^u \\ 1[y_1 \geq q_x^1] & \text{if } u > p_x^u \end{cases}$$

$$F_{Y_0|U,X}(y_0|u,x) = \begin{cases} F^0_{Y_0|U,X}(y_0|u,x) & \text{if } u \geq p_x^l \\ 1[y_0 \geq q_x^0] & \text{if } u < p_x^l. \end{cases}$$

Define

$$F_{Y_0,Y_1,U,X,Z}(y_0, y_1, u, x, z)$$
$$= \int \left[\int_0^u F_{Y_0|U,X}(y_0|t_u, t_x) F_{Y_1|U,X}(y_1|t_u, t_x) dF^0_{U|X}(t_u|t_x) \right]$$
$$\times 1[t_x \leq x, t_z \leq z] dF^0_{X,Z}(t_x, t_z).$$

Where $F^0_{X,Z}$ and $F^0_{U|X}$ are the "true" distributions of (X, Z) and of U conditional on X. Note that F is a proper CDF and that F is a distribution satisfying the conditions that Y_1, Y_0 are bounded conditional on X, and satisfying the property that Z is independent of (Y_0, Y_1, U) conditional on X.

By construction, $F_{X,Z,U}(x, z, u) = F^0_{X,Z,U}(x, z, u)$ so that $F_{X,Z,D}(x, z, d) = F^0_{X,Z,D}(x, z, d)$. In addition, using the fact that $F_{Y_1|U,X}(y_1|u, x) = F^0_{Y_1|U,X}(y_1|u, x)$ for $u \leq p^u_x$, we have

$$F_{Y_1|X,Z,D}(y_1|x, z, 1) = \frac{1}{P(z, x)} \int_0^{P(z,x)} F^0_{Y_1|U,X}(y_1|u, x) dF^0_{U|X}(u|x) = F^0_{Y_1|X,Z,D}(y_1|x, z, 1)$$

for $(x, z) \in \text{Supp}(X, Z|D = 1)$. By a parallel argument,

$$F_{Y_0|X,Z,D}(y_0|x, z, 0) = F^0_{Y_0|X,Z,D}(y_0|x, z, 0)$$

for $(x,z) \in \text{Supp}(X,Z|D=0)$. Combining these results, we have

$$F_{Y,X,Z,D}(y,x,z,d) = F^0_{Y,X,Z,D}(y,x,z,d),$$

where $Y = DY_1 + (1-D)Y_0$. Thus, F is observationally equivalent to the true F^0.

The expected value of $Y_1 - Y_0$ under F equals the given point $s \in [S^L_x, S^U_x]$:

$$\begin{aligned}
E(Y_1 - Y_0|X) &= \int \left[\int y_1 dF_{Y_1|U,X}(y_1|u,x) \right] dF^0_{U|X}(u|x) \\
&\quad - \int \left[\int y_0 dF_{Y_0|U,X}(y_0|u,x) \right] dF^0_{U|X}(u|x) \\
&= \Pr[U \le p^u_x] \int \left[\int_0^{p^u_x} y_1 dF^0_{Y_1|U,X}(y_1|u,x) \right] dF^0_{U|X}(u|x) + \Pr[U > p^u_x] q^1_x \\
&\quad - \Pr[U > p^l_x] \int \left[\int_{p^l_x}^1 y_0 dF^0_{Y_0|U,X}(y_0|u,x) \right] dF^0_U(u) - \Pr[U \le p^l_x] q^0_x \\
&= p^u_x E(Y_1|X, P(Z) = p^u_x, D=1) + (1-p^u_x) q^1_x \\
&\quad - p^l_x E(Y_0|X, P(Z) = p^l_x, D=0) - p^l_x q^0_x \\
&= s.
\end{aligned}$$

Since the expected value of $Y_1 - Y_0$ under F equals s, and since F satisfies all of the required properties of the nonparametric selection model and is observationally equivalent to the true F^0, we have that the point s must be contained in any bounds on the average treatment effect. Since this holds for any point $s \in [S^L_x, S^U_x]$, we have that every point in $[S^L_x, S^U_x]$ must be contained in any bounds on the average treatment effect, and thus the bounds $[S^L_x, S^U_x]$ are tight. ■

6 Comparing the Bounds

We now compare the Heckman and Vytlacil bounds that exploit the nonparametric selection model to the Manski bounds that exploit an instrumental variables assumption. The nonparametric selection model of Heckman and Vytlacil implies the mean independence conditions of Manski, so that Manksi's bounds hold under the Heckman and Vytlacil conditions. We now show that, under the nonparametric selection model, the Manski bounds simplify to the simpler form of the Heckman and Vytlacil bounds.

Theorem 2 *Impose the nonparametric selection model, Assumption S, and impose that the outcome variables are bounded, Assumption B. The Manski mean-independence bounds coincide with the Heckman-Vytlacil bounds.*

Proof.

We first show that the first term of the Heckman-Vytlacil upper bound on Y_1 coincides with the first term of the Manski upper bound on Y_1:

$$\inf_{z\in\mathcal{Z}_x}\{P(z,x)E(Y_1|D=1,X=x,Z=z)+(1-P(z,x))y^u_{(x,z),1}\}$$
$$= p^u_x E(Y_1|D=1,X=x,P(Z)=p^u_x)+(1-p^u_x)y^u_{x,1}.$$

Note that $Z \perp\!\!\!\perp (Y_0,Y_1) \mid X$ implies that $y^u_{(x,z),1}=y^u_{x,1}$. Fix any $x\in \mathrm{Supp}(X)$ and fix any

$z \in \mathcal{Z}_x$.

$$
\begin{aligned}
& \left[p_x^u E(Y_1|D=1, X=x, P(Z,X)=p_x^u) + (1-p_x^u) y_{x,1}^u\right] \\
& \quad - \left[P(z,x) E(Y_1|D=1, X=x, Z=z) + (1-P(z,x)) y_{x,1}^u\right] \\
= \; & \left[\int_0^{p_x^u} E(Y_1|X=x, U=u) dF_{U|X}(u|x) + (1-p_x^u) y_{x,1}^u\right] \\
& \quad - \left[\int_0^{P(z,x)} E(Y_1|X=x, U=u) dF_{U|X}(u|x) + (1-P(z,x)) y_{x,1}^u\right] \\
= \; & \int_{P(z,x)}^{p_x^u} E(Y_1|X=x, U=u) dF_{U|X}(u|x) - (p_x^u - P(z,x)) y_{x,1}^u \\
= \; & \int_{P(z,x)}^{p_x^u} \left[E(Y_1|X=x, U=u) - y_{x,1}^u\right] dF_{U|X}(u|x) \\
\leq \; & 0.
\end{aligned}
$$

Since this inequality holds for any $z \in \mathcal{Z}_x$, we have

$$
\begin{aligned}
& p_x^u E(Y_1|D=1, X=x, P(Z,X)=p_x^u) + (1-p_x^u) y_{x,1}^u \\
& \qquad \leq \inf_{z \in \mathcal{Z}_x} \{P(z,x) E(Y_1|D=1, X=x, Z=z) + (1-P(z,x)) y_{x,1}^u\}
\end{aligned}
$$

Using the fact that $E(Y_1|X=x, U=u) - y_{x,1}^u$ is bounded and the definition of p_x^u, we have that

$$
\begin{aligned}
& p_x^u E(Y_1|D=1, X=x, P(Z,X)=p_x^u) + (1-p_x^u) y_{x,1}^u \\
& \qquad \geq \inf_{z \in \mathcal{Z}_x} \{P(z,x) E(Y_1|D=1, X=x, Z=z) + (1-P(z,x)) y_{x,1}^u\}
\end{aligned}
$$

and thus

$$p_x^u E(Y_1|D=1, X=x, P(Z,X)=p_x^u) + (1-p_x^u) y_{x,1}^u$$
$$= \inf_{z \in \mathcal{Z}_x} \{P(z,x) E(Y_1|D=1, X=x, Z=z) + (1-P(z,x)) y_{x,1}^u\}.$$

By the parallel argument, all other terms of the two sets of bounds coincide. ■

Thus, under the assumption of a nonparametric selection model, the Manski bounds simplify to the same form as the Heckman and Vytlacil bounds. This result is related to Corollary 2 of Proposition 6 of Manski (1994), which shows the same simplification of the bounds under the strong assumption that the treatment choice is *exogenous* so that $Y \perp\!\!\!\perp D|X$. Note that the Manski bounds do not simplify if one does not impose additional restrictions. One can easily construct examples where the Manski bounds do not simplify when the mean independence condition holds but not the nonparametric selection model does not hold.

Somewhat suprisingly, the assumption of a nonparametric selection model does not narrow the bounds compared to what is produced from the weaker mean-independence assumption. However, imposing the nonparametric selection model substantially simplifies the tight mean-independence bounds. Note that this simplification implies the following results for the tight mean-independence bounds under the nonparametric selection model:

1. $p_x^u = 1$, $p_x^l = 0$ is necessary and sufficient for point identification.

2. It is both necessary and sufficient for $P(z,x)$ to be a nontrivial function of z for the bounds

to improve upon the bounds that only impose that the outcome is bounded, $[B_w^L, B_w^U]$.

3. Evaluating the width of the bounds for a given x requires knowledge only of p_x^l, p_x^u, and $y_{x,j}^l, y_{x,j}^u$, $j = 0, 1$.

4. Evaluating the bounds for a given x requires knowledge only of p_x^l, p_x^u, $y_{x,j}^l, y_{x,j}^u$, $j = 0, 1$, $E(Y_0|X = x, P(Z, X) = p_x^l, D = 0)$ and $E(Y_1|X = x, P(Z, X) = p_x^u, D = 1)$.

In each case, the result does not hold in general if the nonparametric selection model is not imposed.

7 Applications to Other Bounds

Our results can be related to the analysis of Balke and Pearl (1997). For the case where Y and Z are binary, Balke and Pearl consider bounds that impose the same statistical independence condition as used by Imbens and Angrist (1994):

$$(Y_1, Y_0, D_0, D_1) \perp\!\!\!\perp Z|X$$

where D_z denotes the counterfactual choice that would have been observed if Z had been externally set to z. Note that this independence condition strengthens the Manski assumptions not only by imposing statistical independence of potential outcomes from Z, instead of mean-independence from Z, but also by imposing independence of the counterfactual choices from Z. When Z and Y are binary, Balke and Pearl show that the sharp bounds under their statistical independence condition are narrower in general than the Manski bounds, although their bounds and the Manski bounds coincide for some distributions of the observed data. In the context of binary Z and Y, Balke and Pearl discuss the Imbens and Angrist monotonicity condition: either $D_1 \geq D_0$ everywhere or $D_1 \leq D_0$ everywhere. They show that this assumption imposes constraints on the observed data which imply that their bounds and the Manski mean-independence bounds coincide.[6]

As demonstrated by Vytlacil (2000), imposing nonparametric selection model (Assumption S) is equivalent to imposing the independence and monotonicity conditions of Imbens and Angrist. The Heckman and Vytlacil analysis imposes the nonparametric selection model. Thus, for the nonparametric selection model, we have from the analysis of Balke and Pearl that the tight bounds when Y and Z are binary are the Manski mean-independence bounds. Thus, the analysis of this paper can be seen as an extension of the Balke and Pearl analysis of the special case of binary Y and Z under the independence and monotonicity conditions. They show that the tight bounds for binary Y and Z under the independence and monotonicity conditions coincide with the Manski mean-independence bounds. Our analysis shows that under the independence and monotonicity conditions, the tight bounds for Y and Z with any support coincide with the Manski mean-independence bounds while having a much simpler and more readily implemented form than the Manski mean-independence bounds.

[6]Robins (1989) also constructs the same bounds under the same conditions for the case of Z and Y binary, but he does not prove that the bounds are tight.

References

[1] Balke, A., and Pearl, J., 1997, "Bounds on Treatment Effects From Studies with Imperfect Compliance," *Journal of the American Statistical Association,* **92**, 1171-1176.

[2] Heckman, J., 1990, "Varieties of Selection Bias," *American Economic Review*, **80**, 313-318.

[3] Heckman, J., J. Smith, and N. Clements, 1997, "Making the Most Out of Programme Evaluations and Social Experiments: Accounting for Hetergeneity in Programme Impacts," *Review of Economic Studies*, 64(4):487-535.

[4] Heckman, J., and E. Vytlacil, 1999, "Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects," *Proceedings of the National Academy of Sciences,* **96**, 4730-4734.

[5] _____, 2000a, "Local Instrumental Variables" with J. Heckman, forthcoming in C. Hsiao, K. Morimune, and J. Powell, eds., *Nonlinear Statistical Inference: Essays in Honor of Takeshi Amemiya*, (Cambridge University Press: Cambridge).

[6] _____, 2000b, "Econometric Evaluation of Social Programs," in J. Heckman and E. Leamer, eds., *Handbook of Econometrics, Volume 5*, (North-Holland: Amsterdam), forthcoming.

[7] Imbens, G., and J. Angrist, 1994, "Identification and Estimation of Local Average Treatment Effects", *Econometrica*, **62**, 467-476.

[8] Manski, C., 1989, "Anatomy of the Selection Problem," *Journal of Human Resources*, **24**, 343-360.

[9] ——, 1990, "Nonparametric Bounds on Treatment Effects," *American Economic Review, Papers and Proceedings,* **80**, 319-323.

[10] ——, 1994, "The Selection Problem," in C. Sims, ed., *Advances in Econometrics: Sixth World Congress*, (Cambridge: Cambridge University Press), 143-170.

[11] Manski, C. and J. Pepper, 2000, "Monotone Instrumental Variables: With an Application to the Returns to Schooling," *Econometrica,* forthcoming.

[12] Robins, J., 1989, "The Analysis of Randomized and Non-randomized AIDS Treatment Trials Using a New Approach to Causal Inference in Longitudinal Studies", in L. Sechrest, H. Freeman and A. Mulley, eds., *Health service Research Methodology: A Focus on AIDS* (U.S. Public Health Service, Washington, DC), 113-159.

[13] Smith, J. and F. Welch, 1986, *Closing The Gap: Forty Years of Economic Progress for Blacks*, (Rand Corporation, Santa Monica, CA).

[14] Vytlacil, E., 2000, "Independence, Monotonicity, and Latent Variable Models: An Equivalence Result," working paper, University of Chicago.