Available online at www.sciencedirect.com

ScienceDirect

Economics Letters 98 (2008) 23–28

economics letters

www.elsevier.com/locate/econbase

Endogenous binary choice models with median restrictions: A comment ☆

Azeem M. Shaikh [a], Edward Vytlacil [b,*]

[a] Department of Economics, University of Chicago, United States
[b] Department of Economics, Columbia University, MC:3308, 420 West 118th Street, New York, NY 10027, United States

Received 25 February 2005; received in revised form 1 March 2007; accepted 15 March 2007
Available online 29 June 2007

Abstract

Hong and Tamer [Hong, H. and Tamer, E. (2003). Endogenous binary choice model with median restrictions. Economics Letters, 80 219–225] provide a sufficient condition for identification of a binary choice model with endogenous regressors. For a special case of their model, we show that this condition essentially requires that the endogenous regressor is degenerate conditional on the instrument with positive probability. Moreover, under weak assumptions, we show that this condition fails to rule out any possible value for the coefficient on the endogenous regressor.

Keywords: Binary response; Endogeneity; Instrumental variables

JEL classification: C14; C25; C35

1. Introduction

Hong and Tamer (2003) consider the following binary choice model:

$$\begin{aligned} y &= \mathbf{1}\{y^* \geq 0\} \\ y^* &= x'\beta + \epsilon, \end{aligned} \tag{1}$$

☆ The authors gratefully acknowledge financial support from the National Science Foundation, SES-05-51089. The authors thank two anonymous referees and the editor Eric Maskin for useful comments. Aprajit Mahajan also provided helpful discussions.
* Corresponding author. Tel.: +1 212 854 2512; fax: +1 212 854 8059.
E-mail address: ev2156@columbia.edu (E. Vytlacil).

0165-1765/$ - see front matter
doi:10.1016/j.econlet.2007.03.006

where $\mathbf{1}\{A\}$ denotes the indicator function of the event A, $x \in \mathbf{R}^k$ is a vector of observed covariates and ϵ is an unobserved random variable. Let x_j denote the jth element of x and let β_j denote the corresponding coefficient. The first element of x, x_0, will be assumed to be identically equal to 1. While the maximum score approach of Manski (1985) imposes $\mathrm{Med}(\epsilon|x)=0$, Hong and Tamer (2003) instead propose an approach based on the weaker assumption that $\mathrm{Med}(\epsilon|z)=0$ for some random vector $z \in \mathbf{R}^d$. Thus, they allow for x to be endogenous in the sense that $\mathrm{Med}(\epsilon|x) \neq 0$ and impose median independence in terms of a vector of instruments z instead.

For ease of exposition, we will consider a simple example with one exogenous regressor and one endogenous regressor; that is,

$$\begin{aligned} y &= \mathbf{1}\{y^* \geq 0\} \\ y^* &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon, \end{aligned} \tag{2}$$

with $z=(x_1, w)$ and $\mathrm{Med}(\epsilon|z)=0$. In addition, we will suppose further that β_1 is known to be strictly positive, so that without loss of generality we can impose the normalization $\beta_1=1$. Our results, however, do not depend on these assumptions and will therefore be valid more generally.

Denote by $\mathbf{B}$ the parameter space for β. In this case, $\mathbf{B}=\{(b_0, 1, b_2):(b_0, b_2)\in\mathbf{R}^2\}$. Under the assumptions that the model is given by Eq. (1), $\mathrm{Med}(\epsilon|z)=0$, and the conditional density of ϵ given (x, z) is continuous and bounded away from zero in a neighborhood of zero uniformly in x and z, Hong and Tamer (2003) show that β is point identified if for all $b \in \mathbf{B}$ such that $b \neq \beta$,

$$\Pr\{z : \Pr\{x'b < 0 \leq x'\beta | z\} = 1 \cup \Pr\{x'\beta < 0 \leq x'b | z\} = 1\} > 0. \tag{3}$$

Hong and Tamer (2003) note that sufficient conditions for Eq. (3) can be characterized using support conditions on the distribution of (x, z). In particular, when $x=z$ (so the regressors are instruments for themselves), Eq. (3) reduces to the condition for identification found in Manski (1985). Yet the analysis of Hong and Tamer (2003) leaves open the question of the extent to which Eq. (3) can be satisfied if $x \neq z$, that is, if the regressors are not instruments for themselves. We show that their analysis for point identification only allows a mild relaxation of the assumption that the regressors are instruments for themselves. In fact, in the special case in which x_2 is a discrete endogenous regressor, a necessary condition for Eq. (3) to hold for all $b \in \mathbf{B}$ such that $b \neq \beta$ is that x_2 is degenerate conditional on z for a set of z values with positive probability. More generally, point identification requires that for all $\delta > 0$ the support of x_2 conditional on z is contained in an interval of length δ for a set of z values with positive probability. When x_2 is discrete, this condition reduces to the one above; when x_2 is not discrete, however, the condition requires that the distribution of x_2 conditional on z is arbitrarily close to degenerate for a set of z values with positive probability.

Hong and Tamer (2003) point out that even when Eq. (3) fails to point identify β, it may still be possible to partially identify β using Eq. (3). In particular, it follows from their analysis that β does not equal any value of $b \in \mathbf{R}$ that satisfies Eq. (3). Thus, $\beta \in \mathbf{B}_I$, where

$$\mathbf{B}_I = \{b \in \mathbf{B} : \Pr\{z : \Pr\{x'b < 0 \leq x'\beta | z\} = 1 \cup \Pr\{x'\beta < 0 \leq x'b \,|\, z\} = 1\} = 0\}.$$

It follows that β_2, the coefficient on the endogenous regressors, lies in the set $\mathbf{B}_{I,2}$, where

$$\mathbf{B}_{I,2} = \{b_2 \in \mathbf{R} : \exists\ b_0 \text{ with } (b_0, 1, b_2) \in \mathbf{B}\}.$$

Under mild restrictions on the joint distribution of the endogenous variable and the instrument, we show that $\mathbf{B}_{I,2}=\mathbf{R}$, so the bounds on β_2 derived from the restriction (3) are typically trivial.

The rest of the paper is organized as follows. In Section 2, we consider a special case of the model (2) in which the endogenous regressor, x_2, is a dummy variable. This special case helps us illustrate ideas and also allows us to draw comparisons of this approach to dummy endogenous regressors in binary choice models with approaches based on "identification-at-infinity" arguments. We relax the assumption that x_2 is a dummy endogenous regressor in Section 3. Section 4 concludes.

2. Special case: dummy endogenous variable

In this section, we consider the special case of the model (2) in which the endogenous regressor, x_2, is a dummy variable. The next proposition shows that unless x_2 is non-degenerate conditional on z for a set of z values with positive probability, Eq. (3) will not hold for some $b \in \mathbf{B}$ such that $b \neq \beta$. In fact, it shows further that unless this condition is satisfied, $\mathbf{B}_{I,2} = \mathbf{R}$, so the bounds on β_2 are vacuous.

Proposition 2.1. *Suppose that y is determined according to Eq.*(2), x_2 *is a Bernoulli random variable,* $Med(\epsilon|z) = 0$, *and* $z = (x_1, w)$. *If*

$$\Pr\{z : \Pr\{x_2 = 0|z\} > 0\} = 1 \text{ or } \Pr\{z : \Pr\{x_2 = 1|z\} > 0\} = 1, \tag{4}$$

then Eq.(3) *does not hold for some* $b \in \mathbf{B}$ *such that* $b \neq \beta$. *Furthermore,* $\mathbf{B}_{I,2} = \mathbf{R}$.

Proof. First suppose $\Pr\{z : \Pr\{x_2 = 0|z\} > 0\} = 1$. Let $b_2^* \in \mathbf{R}$. For $b = (\beta_0, 1, b_2^*)$, we have that

$$\begin{aligned}
\Pr\{x'b < 0 \leq x'\beta|z\} &= \Pr\{x_2 = 1|z\}\Pr\{\beta_0 + x_1 + b_2^* < 0 \leq \beta_0 + x_1 + \beta_2|z, x_2 = 1\} \\
&\quad + \Pr\{x_2 = 0|z\}\Pr\{\beta_0 + x_1 < 0 \leq \beta_0 + x_1|z, x_2 = 0\} \\
&= \Pr\{x_2 = 1|z\}\mathbf{1}\{\beta_0 + x_1 + b_2^* < 0 \leq \beta_0 + x_1 + \beta_2\} \\
&\quad + \Pr\{x_2 = 0|z\}\mathbf{1}\{\beta_0 + x_1 < 0 \leq \beta_0 + x_1\}.
\end{aligned}$$

It follows that for any z such that $\Pr\{x_2 = 0|z\} > 0$ we have $\Pr\{x'b < 0 \leq x'\beta|z\} = 1$ only if $\beta_0 + x_1 < 0 \leq \beta_0 + x_1$. The same argument *mutatis mutandis* shows that $\Pr\{x'\beta < 0 \leq x'b|z\} = 1$ only if $\beta_0 + x_1 < 0 \leq \beta_0 + x_1 < 0$. Thus,

$$\Pr\{z : (\Pr\{x'b < 0 \leq x'\beta|z\} = 1 \cup \Pr\{x'\beta < 0 \leq x'b|z\} = 1) \cap \Pr\{x_2 = 0|z\} > 0\} = 0.$$

Hence, Eq. (3) cannot hold for such a value of b. Moreover, since the choice of b_2^* was arbitrary, we have that $\mathbf{B}_{I,2} = \mathbf{R}$.

Now suppose $\Pr\{z : \Pr\{x_2 = 1|z\} > 0\} = 1$. As before, let $b_2^* \in \mathbf{R}$. For $b = (\beta_0 + \beta_2 - b_2^*, 1, b_2^*)$, we have that

$$\begin{aligned}
\Pr\{x'b < 0 \leq x'\beta|z\} &= \Pr\{x_2 = 1|z\}\Pr\{\beta_0 + x_1 + \beta_2 < 0 \leq \beta_0 + x_1 + \beta_2|z, x_2 = 1\} \\
&\quad + \Pr\{x_2 = 0|z\}\Pr\{\beta_0 + \beta_2 - b_2^* + x_1 < 0 \leq \beta_0 + x_1|z, x_2 = 0\} \\
&= \Pr\{x_2 = 1|z\}\mathbf{1}\{\beta_0 + x_1 + \beta_2 < 0 \leq \beta_0 + x_1 + \beta_2\} \\
&\quad + \Pr\{x_2 = 0|z\}\mathbf{1}\{\beta_0 + \beta_2 - b_2^* + x_1 < 0 \leq \beta_0 + x_1\}.
\end{aligned}$$

It follows that for any z such that $\Pr\{x_2=1|z\}>0$ we have $\Pr\{x'b<0\le x'\beta|z\}=1$ only if $\beta_0+x_1+\beta_2<0\le\beta_0+x_1+\beta_2$. The same argument *mutatis mutandis* shows that $\Pr\{x'\beta<0\le x'b|z\}=1$ only if $\beta_0+x_1+\beta_2<0\le\beta_0+x_1+\beta_2$. Thus,

$$\Pr\{z:(\Pr\{x'b<0\le x'\beta|z\}=1\cup\Pr\{x'\beta<0\le x'b|z\}=1)\cap\Pr\{x_2=1|z\}>0\}=0.$$

Hence, Eq. (3) cannot hold for such a value of b. Moreover, since the choice of b_2^* was arbitrary, we have that $\mathbf{B}_{I,2}=\mathbf{R}$. □

Remark 2.1. It is interesting to compare the approach proposed by Hong and Tamer (2003) with the "identification-at-infinity" approach for selection models with a dummy endogenous regressor (see, for example, Heckman (1990)).[1] Proposition 2.1 shows that in order to identify the coefficient on the dummy endogenous regressor, Hong and Tamer (2003) require that the probability that the endogenous variable equals one conditional on covariates is exactly equal to one and zero with positive probability. In contrast, the identification-at-infinity approach only requires the weaker condition that the probability that the endogenous variable equals one conditional on covariates is arbitrarily close to one and zero with positive probability. On the other hand, this approach also imposes several restrictions not imposed by Hong and Tamer (2003); in particular, it requires that the error term is fully independent (instead of just median independent) of the instruments and that the dummy endogenous variable is determined by a threshold crossing model.[2] Of course, both approaches allow for the possibility of partial identification when the sufficient conditions for point identification are relaxed, but Proposition 2.1 shows that the restriction (3) does not rule out any possible value for the coefficient on the endogenous regressor unless the endogenous regressor is degenerate conditional on the instrument with positive probability. In contrast, it is possible to derive nontrivial bounds for certain parameters without such strong requirements on the joint distribution of the endogenous regressor and the instrument when the conditions for point identification under the identification-at-infinity approach are relaxed (see, for example, Shaikh and Vytlacil (2007)). □

3. General case

For the case of a dummy endogenous variable, we have shown in the preceding section that a necessary condition for Eq. (3) to be satisfied for all $b\in\mathbf{B}$ such that $b\ne\beta$ is that the endogenous variable is degenerate conditional on the instrument with positive probability. In this section, we first show that for Eq. (3) to be satisfied for all $b\in\mathbf{B}$ such that $b\ne\beta$ the endogenous variable must be arbitrarily close to degenerate conditional on the instrument with positive probability. We also show under a mild restriction on the joint distribution of the endogenous regressor and the instrument that Eq. (3) fails to rule out any possible value for β_2; that is, $\mathbf{B}_{I,2}=\mathbf{R}$.

[1] Heckman (1990) formally assumes that the outcome equation is additively separable in the regressors and the error term, but his analysis extends immediately to the case of a binary choice model. See also Cameron and Heckman (1998) and Aakvik et al. (1998) for identification-at-infinity arguments in the context of a system of discrete choice equations.

[2] It is worthwhile to note that since Eq. (3) is only a sufficient condition for identification, it does not preclude the possibility of an identification-at-infinity strategy for the model considered by Hong and Tamer (2003).

Proposition 3.1. *Suppose that y is determined according to Eq.*(2), *Med*($\epsilon|z$) = 0, *and* $z = (x_1, w)$.

(i) *If* $\beta_2 \neq 0$ *and there exists* $\delta > 0$ *such that for a.e. value of z the support of* x_2 *conditional on z is not contained in any interval of* δ, *then Eq.*(3) *does not hold for some* $b \in \mathbf{B}$ *such that* $b \neq \beta$.

(ii) *If there exists a value* x_2^* *such that* x_2^* *is contained in the support of* x_2 *conditional on z for a.e. value of z, then Eq.*(3) *does not hold for some* $b \in \mathbf{B}$ *such that* $b \neq \beta$. *Furthermore,* $\mathbf{B}_{I,2} = \mathbf{R}$.

Proof. (i) Suppose $\beta_2 > 0$. The same argument *mutatis mutandis* will establish the result for $\beta_2 < 0$. Let b_0 satisfy

$$0 < |b_0 - \beta_0| < \delta\beta_2 \tag{5}$$

and define $b = (b_0, 1, \beta_2)$. First consider z such that $\Pr\{x'\beta \geq 0|z\} = 1$. It follows that

$$\Pr\{x'\beta \geq \delta\beta_2|z\} > 0. \tag{6}$$

To see this, suppose by way of contradiction that $\Pr\{x'\beta \geq \delta\beta_2|z\} = 0$, which in turn implies that $\Pr\{x'\beta < \delta\beta_2|z\} = 1$. Since $\Pr\{x'\beta \geq 0|z\} = 1$ by assumption, we have as a result that $\Pr\{0 \leq x'\beta < \delta\beta_2|z\} = 1$. Thus,

$$\Pr\left\{\frac{-(\beta_0 + x_1)}{\beta_2} \leq x_2 \leq \frac{-(\beta_0 + x_1)}{\beta_2} + \delta \,|\, z\right\} = 1,$$

which implies that the support of x_2 conditional on z is contained in an interval of length δ. This contradiction establishes that $\Pr\{x'\beta \geq \delta\beta_2|z\} > 0$. Next note that Eq. (5) implies that

$$\beta_0 \leq b_0 + \delta\beta_2.$$

It therefore follows from Eq. (6) that $\Pr\{x'b > 0|z\} > 0$, which in turn implies that $\Pr\{x'b < 0 \leq x'\beta|z\} < 1$. Hence, $\Pr\{z : \Pr\{x'b < 0 \leq x'\beta|z\} = 1\} = 0$.

Now consider z such that $\Pr\{x'\beta < 0|z\} = 1$. An argument parallel to the one given above shows that $\Pr\{z : \Pr\{x'\beta < 0 \leq x'b|z\} = 1\} = 0$. Thus, Eq. (3) does not hold for such a value of b.

(ii) Let $b_2^* \in \mathbf{R}$. Suppose z is such that x_2^* is an element of the support of x_2 conditional on z. Consider $b = (b_0^*, 1, b_2^*)$ with $b_0^* = \beta_0 + (\beta_2 - b_2^*)x_2^*$. For such b, we have that

$$\Pr\{x'b < 0 \leq x'\beta|z\} = \int \mathbf{1}\{b_0^* + x_1 + b_2^* x_2 < 0 \leq \beta_0 + x_1 + \beta_2 x_2\} \mathrm{d}F(x_2|z).$$

Thus, in order to have $\Pr\{x'b < 0 \leq x'\beta|z\} = 1$, we must have that

$$b_0^* + x_1 + b_2^* x_2 < 0 \leq \beta_0 + x_1 + \beta_2 x_2 \tag{7}$$

for a.e. x_2 conditional on z. In particular, Eq. (7) must hold for $x_2 = x_2^*$, which implies that $b_0^* + x_1 + b_2^* x_2^* = \beta_0 + x_1 + \beta_2 x_2^* < \beta_0 + x_1 + \beta_2 x_2^*$. The same argument *mutatis mutandis* shows that in order to have $\Pr\{x'b < 0 \leq x'\beta|z\} = 1$ we must have $\beta_0 + x_1 + \beta_2 x_2^* < \beta_0 + x_1 + \beta_2 x_2^*$. Since x_2^* is an element of the support of x_2 conditional on z for a.e. value of z, it follows that Eq. (3) cannot hold for such a value of b. Moreover, since the choice of b_2^* was arbitrary, we have that $\mathbf{B}_{I,2} = \mathbf{R}$. □

Remark 3.1. Note that in the special case in which x_2 is a dummy endogenous regressor the hypotheses of both part (i) and part (ii) are satisfied when Eq. (4) holds. Thus, Proposition 3.1 generalizes the results in the preceding section. □

Remark 3.2. As noted by Hong and Tamer (2003) and demonstrated by the above proposition, Eq. (3) is indeed a very strong condition. More surprising, however, is the fact that under the very mild restriction that there exists a value x_2^* such that x_2^* is contained in the support of x_2 conditional on z for a.e. value of z the restriction (3) fails to rule out any value of β_2. For example, this condition is satisfied trivially when $x_2=\gamma z+u$ and u is distributed with support equal to the real line. Thus, the suggestion of Hong and Tamer (2003) to use Eq. (3) to partially identify β_2 may not be useful. Of course, when there is reason to believe that the hypotheses of part (ii) of Proposition 3.1 do not hold, then it may still be possible to follow this suggestion. □

Remark 3.3. Denote by $\mathbf{Z}$ a set of values for z such that x_2 is degenerate conditional on z when $z\in\mathbf{Z}$. Obviously, $\mathrm{Med}(\epsilon|z)=0$ implies $\mathrm{Med}(\epsilon|x)=0$ for any $z\in\mathbf{Z}$. Thus, if such a set $\mathbf{Z}$ with positive probability is known to the researcher, then one may proceed by following the approach of Manski (1985) after restricting attention to observations satisfying $z\in\mathbf{Z}$. □

4. Conclusion

In a recent paper, Hong and Tamer (2003) suggest an approach for endogenous regressors in a binary choice model that exploits median independence of the latent error term from a vector of instruments. In this paper, we consider a special case of their model in which there is a single endogenous regressor and show that a necessary condition for point identification in this model is essentially that the endogenous regressor is degenerate conditional on the instrument with positive probability. Furthermore, under weak restrictions on the joint distribution of the endogenous regressor and the instrument, we show that this approach fails to rule out any possible value for the coefficient on the endogenous regressor.

References

Aakvik, A., Heckman, J., Vytlacil, E., 1998. Semiparametric program evaluation lessons from an evaluation of a Norwegian training program. Mimeo, University of Chicago.

Cameron, S., Heckman, J., 1998. Life cycle schooling and dynamic selection bias: models and evidence for five cohorts of American males. Journal of Political Economy 106, 262–333.

Heckman, J., 1990. Varieties of selection bias. American Economic Review 80, 313–318.

Hong, H., Tamer, E., 2003. Endogenous binary choice model with median restrictions. Economics Letters 80, 219–225.

Manski, C.F., 1985. Semiparametric analysis of discrete response: asymptotic properties of the maximum score estimator. Journal of Econometrics 27, 313–333.

Shaikh, A., Vytlacil, E., 2007. Threshold crossing models and bounds on treatment effects: a nonparametric analysis. Mimeo, University of Chicago and Columbia University.