Economics Letters 105 (2009) 256–260

Contents lists available at ScienceDirect

Economics Letters

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

Bounding the effect of a dichotomous regressor with arbitrary measurement errors

Ping Deng [a,1], Yingyao Hu [b,*]

[a] Department of Economics, Nanchang Institute of Technology, 289 Tianxiang Avenue, Nanchang HI-Tech Development Zone, Nanchang, Jiangxi 330099, China
[b] Department of Economics, Johns Hopkins University, 440 Mergenthaler Hall, 3400 North Charles Street, Baltimore, MD 21218, United States

ARTICLE INFO

Article history:
Received 15 October 2007
Received in revised form 25 July 2009
Accepted 13 August 2009
Available online 25 August 2009

Keywords:
Arbitrary measurement error
Misclassification
Bounds
Dichotomous variable

JEL classification:
C2

ABSTRACT

This note considers a nonlinear regression model containing a 0–1 dichotomous regressor when it is subject to arbitrary measurement errors in the sample. The key identification assumption requires that the third conditional moment of the regression error is zero. This note suggests that the effect of the latent dichotomous variable may be bounded away from zero using the observed moments.

1. Introduction

Measurement errors in survey samples have been discussed in many theoretical and empirical studies (see, e.g., Bound et al. (2001) and Chen et al. (2007) for a survey). When a variable of interest is misreported, the misreported values are considered as a proxy of the latent true values. The identification of the latent model is usually achieved by imposing restrictions on measurement errors in the proxy (e.g., Chen et al. (2008) and Lewbel (2007)). For example, the measurement error is assumed to be independent of the true values (e.g., Schennach (2004, 2007)), or to have zero mean or zero mode (e.g., Hu (2008) and Hu and Schennach (2008)). This note considers the case where measurement errors in the proxy of the latent variable may be arbitrary. In the extreme case, the proxy may be a white noise or a constant. In other words, the latent variable of interest is totally missing in the sense that neither the latent variable itself nor its informative proxies are observed in the sample. The key identification question in such a situation then becomes whether certain restrictions on the latent model may lead to the identification of the latent model.

We focus on a nonlinear regression model containing a misreported 0–1 dichotomous regressor and other accurately-measured regressors. The parameter of interest is the effect of the latent dichotomous variable on the dependent variable. Since the measurement error in the misreported dichotomous regressor may be arbitrary, the useful sample information only contains the joint distribution of the dependent variable and other regressors. We also impose all the assumptions on the regression model in order to handle arbitrary measurement errors. The key identification assumption is that the third conditional moment of the regression error is zero. Such an assumption is reasonable when the regression error has a symmetric distribution. Our main results suggest that the effect of the latent dichotomous variable may be bounded away from zero using the observed moments of the dependent variable conditional on other observed regressors. Such bounds may be useful to test the significance of the effect of the latent variable on the dependent variable given an incomplete sample. Bounds on the conditional mean of the latent dichotomous regressor are also provided.

This note is organized as follows: Section 2 provides the description of the model, data structure, and the main results; Section 3 discusses some properties of the bounds; Section 4 concludes the note; the proof is in the Appendix.

2. Main results

Consider a nonlinear regression model

$$Y = m(X^*, W) + \eta, \tag{1}$$

where Y is the dependent variable, X^* is a latent 0–1 dichotomous regressor and W contains other regressors. In a random sample, we observe the joint distribution of Y, W, and X, where X is a proxy of X^*. Since we allow it to be arbitrary, the measurement error in the proxy X may be correlated with Y, W, and X^*. In the extreme case, X may just be a constant or a white noise. Therefore, our identification results have to

* Corresponding author. Tel.: +1 410 516 7610; fax: +1 410 516 7600.
E-mail addresses: dengping1010@msn.com (P. Deng), yhu@jhu.edu (Y. Hu).
[1] Tel.: +86 791 2126289; fax: +86 791 2126288.

0165-1765/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.econlet.2009.08.009

be based on the sample information from the joint distribution of Y and W. In other words, we only need a sample of (Y,W).

Our assumptions are imposed on the latent model in Eq. (1) as follows:

Assumption 1. $E(\eta|X^*, W) = 0$;

Assumption 2. $E(\eta^2|X^*, W) = E(\eta^2|W)$;

Assumption 3. $E(\eta^3|W) = 0$.

Assumption 1 is a standard normalization condition of a nonlinear regression model. Assumption 2 requires that the variance of the regression error is the same for different values of X^* conditional on other regressors W. This condition is a generalization of the homoscadasticity assumption in the classic regression model. For different values of W, the variance of the regression error may still be different. Assumption 3 is the key assumption in this paper. It is implied by a stronger condition $E(\eta^3|X^*, W) = 0$. Its popular sufficient condition is that the distribution of the regression error η conditional on the regressors X^* and W is symmetric. On the one hand, such a symmetric regression error is a generalization of the normal regression error in the classic regression model. On the other hand, empirical evidences may support such a symmetry assumption. For example, the wage distribution is shown to be close to a log normal distribution. Therefore, the regression error would have a symmetric distribution when the log wage is the dependent variable.

The unknown parameters of interest include the effect $m_\Delta(w)$ of the latent dichotomous variable X^* and its mean $p(w)$ defined as follows:

$$m_\Delta(w) = m(1, w) - m(0, w),$$
$$p(w) = Pr(X^* = 1 | W = w).$$

The observables used in the identification contain the first three conditional moments of the dependent variable:

$$\mu_{Y|W}(w) = E(Y|W = w),$$
$$\sigma_{Y|W}(w) = \left[E\left(\left[Y - \mu_{Y|W}(W)\right]^2 | W = w\right)\right]^{1/2},$$
$$\upsilon_{Y|W}(w) = \left[E\left(\left[Y - \mu_{Y|W}(W)\right]^3 | W = w\right)\right]^{1/3}.$$

In addition, we define

$$\tau(w) = \left(\frac{\upsilon_{Y|W}(w)}{\sigma_{Y|W}(w)}\right)^6.$$

We show in the Appendix that the observables and the unknowns are associated as follows:

$$\sigma^2_{Y|W}(w) \geq p(w)(1 - p(w))m^2_\Delta(w), \tag{2}$$

and

$$\upsilon^3_{Y|W}(w) = p(w)(1 - p(w))(1 - 2p(w))m^3_\Delta(w). \tag{3}$$

Notice that

$$E\left([X^* - p(w)]^3 | W = w\right) = p(w)(1 - p(w))(1 - 2p(w)).$$

Eq. (2) is derived from the fact that the variance of the regression error is nonnegative. The intuition of Eq. (3) is that the observed skewness of Y conditional on W is related to that of the latent dichotomous variable X^* and the effect of interest $m_\Delta(w)$. Since the regression error has a zero third moment and the distribution of a dichotomous variable only has one parameter, i.e., its mean, a nonzero observed skewness $\upsilon^3_{Y|W}(w)$ may provide informative bounds on the effect of interest $m_\Delta(w)$. We summarize the main results as follows:

Theorem 1. Suppose that Assumptions 1–3 hold. We have

1. if $\upsilon_{Y|W}(w) \neq 0$, then $p(w)$ and $m_\Delta(w)$ are bounded as follows:

$$p(w) \in \left(0, \frac{1}{2} - \frac{1}{2}\sqrt{\frac{\tau(w)}{4 + \tau(w)}}\right] \cup \left[\frac{1}{2} + \frac{1}{2}\sqrt{\frac{\tau(w)}{4 + \tau(w)}}, 1\right)$$

and

$$|m_\Delta(w)| \geq C(w)|\upsilon_{Y|W}(w)|,$$

where

$$C(w) = \begin{cases} \tau(w)^{-1/6}\sqrt{4 + \tau(w)} > C_0 & \text{if } \tau(w) \in (2, \infty) \\ C_0 & \text{if } \tau(w) \in (0, 2] \end{cases}$$

with $C_0 = \left(6\sqrt{3}\right)^{1/3} \approx 2.182$. These bounds are sharp. Moreover,

$$\text{sign}\{m_\Delta(w)\} = \text{sign}\left\{\upsilon_{Y|W}(w)\right\} \times \text{sign}\left\{\frac{1}{2} - p(w)\right\}.$$

2. if $\upsilon_{Y|W}(w) = 0$, then either $p(w) \in \left\{0, \frac{1}{2}, 1\right\}$ or $m_\Delta(w) = 0$ holds.

Proof. See the Appendix. □

First of all, the condition $\upsilon_{Y|W}(w) = 0$ is directly testable from the data. Second, When $\upsilon_{Y|W}(w) \neq 0$, this theorem suggests that the effect of the latent variable $m_\Delta(w)$ may be bounded away from zero. That means we may be able to test whether the variable X^* has a significant effect on the dependent variable Y even if X^* or its proxies are not observed. Third, the sign of the effect may be determined if we know whether the probability of $X^* = 1$ is larger than a half or not. For example, suppose $X^* = 1$ stands for the union participation. If we know from another source that less than half of the population of interest are union members for a given value w. Then the sign of the effect $m_\Delta(w)$ is the same as $\upsilon_{Y|W}(w)$. In fact, the probability of $X^* = 1$ is bounded away from a half, which may make it relatively easy to determine the sign of $\left\{\frac{1}{2} - p(w)\right\}$. For example, suppose that we know $\tau(w) = 2$ from the data. Then Theorem 1 implies that the probability of $X^* = 1$ is in $(0, 0.212) \cup (0.788, 1)$. That means we only need to determine whether the percentage of union membership is lower than 21.2% or higher than 78.8%. Finally, if we believe that the probability of $X^* = 1$ is not equal to zero, a half, or one, then $\upsilon_{Y|W}(w) = 0$ implies that the latent regressor X^* has no impact on the dependent variable for a given value w. We summarize the discussion above in the following corollary:

Corollary 1. Suppose that assumptions in Theorem 1 hold and that $p(w) \in \left(0, \frac{1}{2}\right)$. Then,

$$p(w) \in \begin{cases} \left(0, \frac{1}{2} - \frac{1}{2}\sqrt{\frac{\tau(w)}{4 + \tau(w)}}\right] & \text{if } \upsilon_{Y|W}(w) \neq 0 \\ \left(0, \frac{1}{2}\right) & \text{if } \upsilon_{Y|W}(w) = 0 \end{cases}$$

and

$$m_\Delta(w) \begin{cases} \geq C(w)\upsilon_{Y|W}(w) & \text{if } \upsilon_{Y|W}(w) > 0 \\ = 0 & \text{if } \upsilon_{Y|W}(w) = 0 \\ \leq C(w)\upsilon_{Y|W}(w) & \text{if } \upsilon_{Y|W}(w) < 0 \end{cases}.$$

Moreover, Theorem 1 also implies that the sign of the effect $m_\Delta(w)$ is also useful to determine the range of the probability of $X^* = 1$. For example, if X^* is the education level in a wage equation, it is reasonable to assume the effect of education is positive. Then, the sign of the observed $\upsilon_{Y|W}(w)$ may determine whether the probability of $X^*=1$ is larger than a half or not. The results in this case are as follows:

Corollary 2. Suppose that assumptions in Theorem 1 hold and that $m(1, w)>m(0, w)$. Then,

$$p(w)\in\begin{cases}\left(0,\frac{1}{2}-\frac{1}{2}\sqrt{\frac{\tau(w)}{4+\tau(w)}}\right] & \text{if } \upsilon_{Y|W}(w)>0\\ \left\{0,\frac{1}{2},1\right\} & \text{if } \upsilon_{Y|W}(w)=0\\ \left[\frac{1}{2}+\frac{1}{2}\sqrt{\frac{\tau(w)}{4+\tau(w)}},1\right) & \text{if } \upsilon_{Y|W}(w)<0\end{cases}$$

and

$$m_\Delta(w)\begin{cases}\geq C(w)|\upsilon_{Y|W}(w)| & \text{if } \upsilon_{Y|W}(w)\neq 0\\ >0 & \text{if } \upsilon_{Y|W}(w)=0\end{cases}.$$

3. Discussion

The results in Theorem 1 imply that the structure of the model is important in terms of finding the correct bounds. For example, we consider a wage regression model as follows:

$$\ln G = a + bX^* + \eta, \quad (4)$$

where G is the wage and X^* is a 0–1 indicator for the college level education. For simplicity, we assume that η is independent of X^*. Suppose that the regression error η satisfies the assumptions in Theorem 1. The bounds on a and b then follow.

Now suppose that we use the wage itself instead of the log wage in the regression model as follows:

$$G = a' + b'X^* + \eta', \quad (5)$$

where

$$a' = e^a Ee^\eta,$$
$$b' = e^a\left(e^b-1\right)Ee^\eta,$$
$$\eta' = e^{a+bX^*}\left(e^\eta - Ee^\eta\right).$$

Notice that bounds on a' and b' may still imply bounds on b if η' satisfies the assumptions in Theorem 1. We then consider Assumptions 1–3 with η replaced by η'. Assumption 1 is satisfied by the definition of η' and the independence between η and X^*. However, Assumption 2 holds only if $b=0$. Moreover, Assumption 3 is unlikely to hold with η' if it holds with η. Following the proof of Theorem 1, one can show that the only inequality still holds is

$$\mathrm{Var}(G)\geq p(1-p)b'^2,$$

which does not provide any informative bounds on p or b'. Therefore, the results in Theorem 1 don't hold with Eq. (5) if Eq. (4) satisfies the assumptions and provides informative bounds. This example implies that the structure of the model is important in order to obtain the correct and informative bounds.

Another issue that we want to discuss here is the inference for the partially identified parameters p and m_Δ. There are many studies on this topic in the last decade (for example, Chernozhukov et al. (2007) and Beresteanu and Molinari (2008)). Given the length restriction on this note, we will only refer to Chernozhukov et al. (2007) on this issue instead of providing a lengthy discussion. We show that the bounds developed in our note fall into the category discussed in their paper. Therefore, the inference for the identified set follows from their results.

For simplicity, we consider a regression model without constant and covariates as follows:

$$Y = m_\Delta(X^* - p) + \eta.$$

As shown in the proof of Theorem 1, the partially identified set is characterized by

$$E\left[p(1-p)m_\Delta^2 - Y^2\right]\leq 0,$$
$$E\left[p(1-p)(1-2p)m_\Delta^3 - Y^3\right] = 0.$$

Define $\theta=(p,m_\Delta)^T\in\Theta$. The corresponding moment-inequality restrictions take the form

$$E[m_i(\theta)]\leq 0,$$

where

$$m_i(\theta) = \begin{pmatrix} p(1-p)m_\Delta^2 - Y_i^2\\ p(1-p)(1-2p)m_\Delta^3 - Y_i^3\\ -p(1-p)(1-2p)m_\Delta^3 + Y_i^3\end{pmatrix}.$$

The identified set is then $\Theta_I=\{\theta\in\Theta\colon E[m_i(\theta)]\leq 0\}$. As shown in Chernozhukov et al. (2007), the set Θ_I can be characterized as the set of minimizers of the criterion

$$Q(\theta) = \|E[m_i(\theta)]'W^{1/2}(\theta)\|_+^2,$$

where $||x||_+=||\max(x,0)||$ and $W(\theta)$ is a continuous and diagonal matrix with strictly positive diagonal elements for each $\theta\in\Theta$. Therefore, inference on Θ_I may be achieved from the empirical analog of Q as follows:

$$Q_n(\theta) = \|E_n[m_i(\theta)]'W_n^{1/2}(\theta)\|_+^2,$$

where $E_n[m_i(\theta)] = \frac{1}{n}\sum_{t=1}^n m_t(\theta)$ and $W_n(\theta)$ is a consistent estimate of $W(\theta)$. The consistent estimate of Θ_I they proposed takes the form of a contour set of level c, i.e., $C_n(c)=\{\theta\in\Theta\colon a_n Q_n(\theta)\leq c\}$, where a_n is a normalizing sequence. The set $C_n(c)$ may also be a confidence region for Θ_I in the sense that $\lim_{n\to\infty} P(\Theta_I\subseteq C_n(c))=\alpha$ for a specified confidence level α.

Furthermore, the bounds may degenerate when $\upsilon_{Y|W}$ (and therefore τ) is arbitrarily close to 0 so that the bounds on p, i.e.,

$$p\in\left(0,\frac{1}{2}-\frac{1}{2}\sqrt{\frac{\tau}{4+\tau}}\right]\cup\left[\frac{1}{2}+\frac{1}{2}\sqrt{\frac{\tau}{4+\tau}},1\right),$$

becomes less informative. One can imagine that when $\upsilon_{Y|W}$ (and therefore τ) is very close to zero the confidence region will be very likely to contain $\frac{1}{2}$, which makes the confidence region uninformative. Theoretically, we may let $\upsilon_{Y|W}$ converge to zero as the sample size goes to infinity. In that case, the confidence sets on p converge to the uninformative $(0,1)$ and those on $m_\Delta(w)$ converge to $(0,+\infty)$. With a real data set, the estimate of $\upsilon_{Y|W}(w)$ is unlikely to be exactly equal to 0. We may obtain informative bounds if we can reject the null hypothesis that $\upsilon_{Y|W}=0$. But when we can't reject the null, the confidence region may not be very informative or useful but it does not degenerate as long as the estimate of $\upsilon_{Y|W}$ is not zero.

4. Conclusion

This note considers a nonlinear regression model containing a 0–1 dichotomous variable when it is misreported with arbitrary measurement errors. Informative bounds are provided on the effect of the latent true regressor on the dependent variable. We impose all the assumptions on the latent model instead of the measurement error distribution. In other words, our results provide useful bounds even if the misreported values of the latent regressor are a white noise or a constant. The results imply that certain information on the latent model may be useful for the identification of the parameter of interest even if the measurement error is extremely severe. It would be interesting to extend such results to a more general measurement error model.

Appendix

Proof. (Theorem 1) By the definitions of $\upsilon_{Y|W}(w)$, $m_\Delta(W)$, and $p(w)$, the regression model may be written as

$$\left(Y-\mu_{Y|W}(w)\right)=m_\Delta(W)(X^*-p(w))+\eta.$$

For simplicity, we omit the argument w when it doesn't cause any confusion.

Given the definitions of $\sigma_{Y|W}(w)$ and $\upsilon_{Y|W}(w)$, we then consider the second and the third moments of Y conditional on W as follows:

$$\sigma^2_{Y|W}=p(1-p)m^2_\Delta+E\left(\eta^2|W=w\right) \tag{6}$$

and

$$\upsilon^3_{Y|W}=p(1-p)(1-2p)m^3_\Delta+E\left(\eta^3|W=w\right).$$

Assumption 3 then leads to

$$\upsilon^3_{Y|W}=p(1-p)(1-2p)m^3_\Delta. \tag{7}$$

When $\upsilon_{Y|W}\neq 0$, Eq. (7) implies that the sign of m_Δ satisfies

$$\text{sign}\{m_\Delta\}=\text{sign}\left\{\upsilon_{Y|W}\right\}\times\text{sign}\left\{\frac{1}{2}-p\right\}.$$

The next step is to use Eqs. (6) and (7) to derive bounds on p and m_Δ. It is obvious that Eq. (7) implies that $p\in\{0,1,1/2\}$ or $m_\Delta=0$ when $\upsilon_{Y|W}=0$. We then focus on the case where $\upsilon_{Y|W}\neq 0$, which implies that p is not equal to 0,1, or 1/2 and

$$m_\Delta=\left[\frac{1}{p(1-p)(1-2p)}\right]^{1/3}\upsilon_{Y|W}. \tag{8}$$

The bounds on p are implied by the condition $E(\eta^2|W=w)\geq 0$ through Eq. (6) as follows:

$$\sigma^2_{Y|W}\geq p(1-p)m^2_\Delta.$$

Eliminating m_Δ using Eq. (8) leads to

$$\frac{(1-2p)^2}{p(1-p)}\geq\left(\frac{\upsilon_{Y|W}}{\sigma_{Y|W}}\right)^6\equiv\tau,$$

which implies $\frac{1}{4+\tau}-p+p^2\geq 0$. Therefore, we have $p\in S_p$ with

$$S_p=\left(0,\frac{1}{2}-\frac{1}{2}\sqrt{\frac{\tau}{4+\tau}}\right]\cup\left[\frac{1}{2}+\frac{1}{2}\sqrt{\frac{\tau}{4+\tau}},1\right).$$

The bounds on m_Δ are then determined by the behavior of the function $p(1-p)(1-2p)$ over the set S_p derived above. Define $g(p)=p(1-p)(1-2p)$. Note that $\frac{d}{dp}g(p)=1-6p+6p^2$. Therefore, $\frac{d}{dp}g(p_0)=0$ leads to $p_0=\frac{1}{2}\pm\frac{1}{2}\sqrt{\frac{1}{3}}$. That means for $p\in[0,1]$

$$g\left(\frac{1}{2}+\frac{1}{2}\sqrt{\frac{1}{3}}\right)\leq g(p)\leq g\left(\frac{1}{2}-\frac{1}{2}\sqrt{\frac{1}{3}}\right),$$

where $g\left(\frac{1}{2}+\frac{1}{2}\sqrt{\frac{1}{3}}\right)=-\frac{\sqrt{3}}{18}$ and $g\left(\frac{1}{2}-\frac{1}{2}\sqrt{\frac{1}{3}}\right)=\frac{\sqrt{3}}{18}$. In other words, we have $|\frac{1}{g(p)}|\geq 6\sqrt{3}$ for $p\in[0,1]$ and

$$|m_\Delta|\geq C_0|\upsilon_{Y|W}|$$

with $C_0=\left(6\sqrt{3}\right)^{1/3}$. However, we have to consider the case where the minimizer or the maximizer $\frac{1}{2}\pm\frac{1}{2}\sqrt{\frac{1}{3}}$ may not be in the set S_p. Notice that $\sqrt{\frac{\tau}{4+\tau}}>\sqrt{\frac{1}{3}}$ if and only if $\tau>2$. Therefore, we have

$$|m_\Delta|\geq C|\upsilon_{Y|W}|,$$

where

$$C=\begin{cases}C_1 & \text{if } \tau>2\\ C_0 & \text{if } \tau\leq 2\end{cases}$$

with $C_1=\left[g\left(\frac{1}{2}-\frac{1}{2}\sqrt{\frac{\tau}{4+\tau}}\right)\right]^{-1/3}=\tau^{-1/6}\sqrt{4+\tau}.$

Furthermore, the bounds generated in this procedure are sharp because there exists a possible value of unobservable, i.e., $E(\eta^2|W=w)$, to support any p and m_Δ in the feasible region, including the bounds themselves, given the observables $\sigma_{Y|W}$ and $\upsilon_{Y|W}$. In other words, the sharpness of the bounds on p and m_Δ can be shown by finding possible values of the unobservable which lead to given values of p and m_Δ in the feasible region.

Next, we show the bounds on p are sharp. Suppose that for a given $\sigma_{Y|W}$ and $\upsilon_{Y|W}$ we pick a $\tilde{p}$ such that Theorem 1(1) holds, i.e.,

$$\tilde{p}\in\left(0,\frac{1}{2}-\frac{1}{2}\sqrt{\frac{\tau}{4+\tau}}\right]\cup\left[\frac{1}{2}+\frac{1}{2}\sqrt{\frac{\tau}{4+\tau}},1\right).$$

The value of corresponding m_Δ is then

$$m_\Delta=\left(\frac{\upsilon^3_{Y|W}}{\tilde{p}(1-\tilde{p})(1-2\tilde{p})}\right)^{1/3},$$

which is guaranteed to satisfy the bounds on m_Δ. The unobserved $E(\eta^2|W=w)$ then equals

$$E\left(\eta^2|W=w\right)=\sigma^2_{Y|W}-\tilde{p}(1-\tilde{p})m^2_\Delta.$$

The derivation of the bounds guarantees that the right hand side is nonnegative.

The sharpness of the bounds on m_Δ can be shown in a similar way. Suppose that for a given $\sigma_{Y|W}$ and $\upsilon_{Y|W}$ we pick an $\tilde{m}_\Delta$ such that Theorem 1(1) holds, i.e.,

$$|\tilde{m}_\Delta|\geq C(w)|\upsilon_{Y|W}|,$$

where

$$C(w)=\begin{cases}\tau^{-1/6}\sqrt{4+\tau(w)}>C_0 & \text{if } \tau\in(2,\infty)\\ C_0 & \text{if } \tau\in(0,2]\end{cases}.$$

The value of p is then a root of

$$p(1-p)(1-2p)-\frac{v^3_{Y|W}}{\tilde{m}^3_\Delta}=0.$$

We may pick any one of the three roots, which must satisfy the bounds on p. The value of $E(\eta^2|W=w)$ is equal to

$$E\left(\eta^2|W=w\right)=\sigma^2_{Y|W}-p(1-p)\tilde{m}^2_\Delta.$$

Again, the derivation of the bounds guarantees that the right hand side is nonnegative. Therefore, the bounds on p and m_Δ are sharp. The detailed algebraic derivation in the discussion above is straightforward but tedious, and therefore, is omitted. A similar and detailed proof can be found in Hu (2006). □

References

Beresteanu, A., Molinari, F., 2008. Asymptotic properties for a class of partially identified models. Econometrica 76 (4).

Bound, J., Brown, C., Mathiowetz, N., 2001. In: Heckman, J., Leamer, E. (Eds.), Measurement Error in Survey Data. In: Handbook of Econometrics, vol. 5. North Holland.

Chen, X., Hong, H., Nekipelov, D., 2007. "Measurement error models," working paper of New York University and Stanford University. A Survey Prepared for the Journal of Economic Literature.

Chen, X., Hu, Y., Lewbel, A., 2008. Nonparametric identification of regression models containing a misclassified dichotomous regressor without instruments. Economics Letters 100 (3), 381–384.

Chernozhukov, V., Hong, H., Tamer, E., 2007. Estimation and confidence regions for parameter sets in econometric models. Econometrica 75, 1243–1284.

Hu, Y., 2006. Bounding parameters in a linear regression model with a mismeasured regressor using additional information. Journal of Econometrics 133 (1), 51–70.

Hu, Y., 2008. Identification and estimation of nonlinear models with misclassification error using instrumental variables: a general solution. Journal of Econometrics 144 (1), 27–61.

Hu, Y., Schennach, S., 2008. Instrumental variable treatment of nonclassical measurement error models. Econometrica 76 (1), 195–216.

Lewbel, A., 2007. Estimation of average treatment effects with misclassification. Econometrica 75, 537–551.

Schennach, S., 2004. Estimation of nonlinear models with measurement error. Econometrica 72, 33–75.

Schennach, S., 2007. Instrumental variable estimation of nonlinear errors-in-variables models. Econometrica 75, 201–239.