

Submitted to Econometrica

On the Use of Instrumental Variables  
to Identify the Effect of a  
Mis-measured, Binary Regressor

Francis J. DiTraglia and Camilo Garcia-Jimeno

September 10, 2015

ON THE USE OF INSTRUMENTAL VARIABLES TO IDENTIFY THE  
EFFECT OF A MIS-MEASURED, BINARY REGRESSOR<sup>1</sup>

FRANCIS J. DiTRAGLIA<sup>a</sup> AND CAMILO GARCIA-JIMENO<sup>a</sup>

Abstract goes here.

KEYWORDS: Instrumental Variables, Measurement Error, Binary Regressor,  
Endogeneity.

1. INTRODUCTION

Introduction goes here.

2. A REVIEW OF THE LITERATURE

Should probably start off, possible in previous section, by writing out a  
general encompassing framework that will allow us to talk about all the  
papers in this section. Present our model before this section.

Many examples want to estimate effect of binary treatment. Often treat-  
ment is mis-measured and possibly endogenous. Measurement error in bi-  
nary regressor cannot be classical. This has been know for a while, see  
[Aigner \(1973\)](#) and [Bollinger \(1996\)](#). Intuition: can only mis-code true zero  
and one and true one as zero. Describe non-differential measurement error  
idea. Under this kind of measurement error, IV estimator is inconsistent  
for the causal effect: can remove the effect of endogeneity but not of mea-  
surement error. See for example [Black et al. \(2000\)](#); [Frazis and Loewenstein  
\(2003\)](#); [Kane et al. \(1999\)](#).

Now talk about [Kane et al. \(1999\)](#) and [Black et al. \(2000\)](#). Two measures  
of exogenous binary treatment with non-differential measurement error al-  
low one to identify treatment effect. Method of moments estimator *not* IV.

---

<sup>1</sup>We thank...

<sup>a</sup>Dept. of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA  
19104, U.S.A.; [fditra@econ.upenn.edu](mailto:fditra@econ.upenn.edu); [gcamilo@econ.upenn.edu](mailto:gcamilo@econ.upenn.edu)

Relies on discreteness of the problem: construct “cells” for  $E[y|z, T]$ . Talk about how the two papers differ in their contribution. [Black et al. \(2000\)](#) consider not only the binary case but a continuous version that isn’t identified. Need to figure out how [Card \(1996\)](#) relates to these as well.

[Frazis and Loewenstein \(2003\)](#) point out that an instrument can be used in place of a second measure of  $T^*$  provided that  $T^*$  is still exogenous. Essentially the same estimator as in [Black et al. \(2000\)](#) and [Kane et al. \(1999\)](#) but more general since the instrument need not be binary: can in fact be continuous. But they make a mistake. They assume  $E[zu] = 0$ ,  $E[T^*u] = 0$  and non-differential measurement error and claim that this is sufficient to consistently estimate  $\beta$ . However, this is incorrect: we need the additional assumption that  $E[zT^*u] = 0$  which is stronger. (They seem to think that this term only appears when you have an endogenous  $T^*$ .) While [Frazis and Loewenstein \(2003\)](#) are aware that there are some differences between two measures of  $T^*$ , as in [Black et al. \(2000\)](#), and an arbitrary instrument  $z$ , they seem to have missed one subtle point. The assumptions in [Black et al. \(2000\)](#) in fact imply that  $E[u|T^*, z] = 0$ .<sup>1</sup> From this it follows that  $E[zT^*u] = E[zu] = E[T^*u] = 0$ . However, if one takes the non-differential measurement assumption literally it is in fact sufficient in the case of two measures to assume only that  $E[zu] = E[T^*u] = 0$ :

$$\begin{aligned}
 (1) \quad E[zT^*u] &= E[(T^* + w)T^*u] = E[(T^*)^2u] + E[wuT^*] \\
 (2) \quad &= E[E(u|T^*)(T^*)^2] + E[E(wu|T^*)T^*] \\
 (3) \quad &= 0 + E[E(w|T^*)E(u|T^*)T^*] = 0
 \end{aligned}$$

using the fact that  $E[u|T^*] = 0$  and  $w$  is independent of  $u$  conditional on  $T^*$ . This argument does *not* necessarily apply to an arbitrary instrument  $z$ :  $E[zu] = E[T^*u] = 0$  does not imply that  $E[zT^*u] = 0$ .

---

<sup>1</sup>This follows from Assumptions A1 and A2 combined with Equation 3.

Put in our simple binary example.

While it might seem strange to assume in practice that  $E[zu] = E[T^*u] = 0$  are exogenous but not that  $E[zT^*u] = 0$  the point is merely that this is an additional assumption beyond the usual assumptions of lack of correlation.

Frazis and Loewenstein (2003) also briefly discuss case in which  $T^*$  is endogenous, basically conclude that all you can get in this case is bounds for the treatment effect.

### 3. NOTES ON MAHAJAN (2006)

Mahajan (2006) considers regression models of the form

$$(4) \quad E[y - g(x^*, z)] = 0$$

where  $x^*$  is an unobserved binary regressor and  $z$  is a  $d_z \times 1$  vector of control regressors. Rather than  $x^*$  we observe a noisy measure  $x$  called the “surrogate” and an additional variable  $v$  that acts, in essence, as an instrumental variable. Since  $v$  does not, strictly speaking, meet the traditional requirements for an instrument, Mahajan refers to it as an “instrument-like variable” or ILV for short. Throughout the paper, Mahajan assumes that  $v$  is binary although he claims that the same idea applies to arbitrary discrete variables. The paper considers two main cases: one in which  $x^*$  is assumed to be exogenous, and another in which it is not.

#### 3.1. The Case of Exogenous $x^*$

The first is based on the restriction

$$(5) \quad E[y - g(x^*, z) \mid x^*, x, z, v] = 0$$

### 3.2. The Case of Endogenous $x^*$

While the preceding case required  $x^*$  to be exogenous, Mahajan claims (page 640) that his identification results can be extended to account for endogeneity provided that one is willing to restrict attending to additively separable models of the form

$$(6) \quad y = g^*(x^*, z) + \varepsilon$$

In this case, the ILV is assumed to satisfy the usual instrumental variables mean independence assumption

$$(7) \quad E[\varepsilon|z, v] = 0$$

and Equation 5 is replaced by

$$(8) \quad E[y|x^*, x, z, v] = E[y|x^*, z]$$

Unfortunately, Mahajan's proof is incorrect and the model in Equation 6 is unidentified. The mistake stems from a false analogy with the identification proof in the case of exogenous  $x^*$ . In A.2 Mahajan argues, correctly, that under 6–8 knowledge of the mis-classification rates is sufficient to identify the model even when  $x^*$  is endogenous. He then appeals to Theorem 1 to argue that the mis-classification rates are indeed identified. The proof of Theorem 1, however, depends crucially on the assumption that  $x^*$  is exogenous. Without this assumption, the mis-classification rates are unidentified, as we now show. For ease of exposition we consider the case without covariates. Equivalently, one can interpret all of the expressions that follow as implicitly conditioned on  $z = z_a$  where  $z_a$  is a value in the support of  $z$ .<sup>2</sup>

---

<sup>2</sup>Because the covariates are held fixed throughout the proof of Mahajan's Theorem 1, there is no loss of generality.

Without covariates we can write

$$(9) \quad y = \alpha + \beta x^* + \varepsilon$$

where  $\alpha = g^*(0)$  and  $\beta = g^*(1) - g^*(0)$  and the mis-classification rates become  $\eta_0 = P(x = 1|x^* = 0)$  and  $\eta_1 = P(x = 0|x^* = 1)$ . Now define

$$(10) \quad m_{jk} = E[\varepsilon|x^* = j, v = k]$$

#### 4. LEWBEL (2007)

Lewbel shows that under an exogenous but missclassified treatment, and an instrument that takes on at least three values, the treatment effect is identified. The model is

$$\mathbb{E}[Y|T^*, T] = \alpha + \beta T^*$$

Using iterated expectations over the distribution of  $T^*$  given  $T$ ,

$$\begin{aligned} \mathbb{E}[Y|T] &= \mathbb{E}_{T^*|T}[\mathbb{E}[Y|T^*, T]] \\ &= \mathbb{P}(T^* = 1|T)\mathbb{E}[Y|T^* = 1] + \mathbb{P}(T^* = 0|T)\mathbb{E}[Y|T^* = 0] \\ &= \mathbb{P}(T^* = 1|T)(\alpha + \beta) + \mathbb{P}(T^* = 0|T)\alpha \\ &= \alpha + \mathbb{P}(T^* = 1|T)\beta \end{aligned}$$

which implies that

$$\beta_{OLS} = \mathbb{E}[Y|T = 1] - \mathbb{E}[Y|T = 0] = [\mathbb{P}(T^* = 1|T = 1) - \mathbb{P}(T^* = 1|T = 0)] \beta$$

Lewbel defines

$$M(\alpha_0, \alpha_1, p) = \mathbb{P}(T^* = 1|T = 1) - \mathbb{P}(T^* = 1|T = 0)$$

implying that  $\beta_{OLS} = \beta M(\alpha_0, \alpha_1, p)$ . It turns out that we can re-express  $M(\alpha_0, \alpha_1, p)$  as

$$M(\alpha_0, \alpha_1, p) = \frac{1}{1 - \alpha_0 - \alpha_1} \left[ 1 - \frac{(1 - \alpha_1)\alpha_0}{p} - \frac{(1 - \alpha_0)\alpha_1}{1 - p} \right]$$

To see why this is the case first note that, by Bayes' Rule,

$$M(\alpha_0, \alpha_1, p) = \frac{(1 - \alpha_1)p^*}{p} - \frac{\alpha_1 p^*}{1 - p} = p^* \left[ \frac{(1 - \alpha_1)(1 - p) - \alpha_1 p}{p(1 - p)} \right]$$

Now, by the Law of Total Probability,

$$\begin{aligned} p &= P(T = 1|T^* = 1)p^* + P(T = 1|T^* = 0)(1 - p^*) \\ &= (1 - \alpha_0)p^* + \alpha_0(1 - p^*) \\ &= (1 - \alpha_0 - \alpha_1)p^* + \alpha_0 \end{aligned}$$

Rearranging, we see that  $p^* = (p - \alpha_0)/(1 - \alpha_0 - \alpha_1)$ . Substituting this into the expression for  $M(\alpha_0, \alpha_1, p)$  and simplifying,

$$\begin{aligned} M(\alpha_0, \alpha_1, p) &= \frac{p - \alpha_0}{1 - \alpha_0 - \alpha_1} \left[ \frac{(1 - \alpha_1)(1 - p) - \alpha_1 p}{p(1 - p)} \right] \\ &= \frac{1}{1 - \alpha_0 - \alpha_1} \left[ \frac{(p - \alpha_0)(1 - \alpha_1)(1 - p) - (p - \alpha_0)\alpha_1 p}{p(1 - p)} \right] \\ &= \frac{1}{1 - \alpha_0 - \alpha_1} \left[ 1 - \frac{(1 - p)(1 - \alpha_1)\alpha_0 + p\alpha_1 - \alpha_0\alpha_1 p}{p(1 - p)} \right] \\ &= \frac{1}{1 - \alpha_0 - \alpha_1} \left[ 1 - \frac{(1 - p)(1 - \alpha_1)\alpha_0}{p(1 - p)} - \frac{p\alpha_1(1 - \alpha_0)}{p(1 - p)} \right] \\ &= \frac{1}{1 - \alpha_0 - \alpha_1} \left[ 1 - \frac{(1 - \alpha_1)\alpha_0}{p} - \frac{(1 - \alpha_0)\alpha_1}{1 - p} \right] \end{aligned}$$

Now, the instrument  $z$  is assumed to be discrete and to take on at least three distinct values. Let  $\beta_{OLS}^k$  denote the OLS estimator based only on observations for which  $z = z_k$ , where  $z_k$  is a particular value in the support

of  $z_k$ , that is

$$\beta_{OLS}^k = \frac{Cov(T, Y|z = z_k)}{Var(T|z = z_k)}$$

and let  $p_k = E(T|z = z_k)$ . The denominator of the expression for  $\beta_{OLS}^k$  is simply  $Var(T|z = z_k) = p_k(1 - p_k)$ . For the numerator, note that

$$\begin{aligned} Cov(T, y|z) &= E(Ty|z) - E(T|z)E(y|z) \\ &= E_{T|z} [E(y|T, z)T] - E(T|z)E_{T|z} [E(y|T, z)] \\ &= E(y|T = 1, z)E(T|z) \\ &\quad - E(T|z) \{E(T|z)E(y|T = 1, z) + [1 - E(T|z)]E(y|T = 0, z)\} \\ &= E(T|z) [1 - E(T|z)] \{E(y|T = 1, z) - E(y|T = 0, z)\} \end{aligned}$$

by iterated expectations over the distribution of  $T^*$  given  $T$  and  $z$ . Thus,

$$\beta_{OLS}^k = E(y|T = 1, z = z_k) - E(y|T = 0, z = z_k)$$

and finally, since  $E(y|T, z) = \alpha + \beta P(T^* = 1|T, z)$ , we see that

$$\beta_{OLS}^k = \beta \{P(T^* = 1|T = 1, z = z_k) - P(T^* = 1|T = 0, z = z_k)\}$$

Notice that this expression looks almost identical to  $\beta_{OLS} = \beta M(\alpha_0, \alpha_1, p)$  from above. The only difference is that we condition on  $z = z_k$ . Because we assume that the mis-classification probabilities are independent of  $z$ ,

$$\begin{aligned} P(T = 1|T^* = 1, z = z_k) &= P(T = 1|T^* = 1) = 1 - \alpha_1 \\ P(T = 1|T^* = 0, z = z_k) &= P(T = 1|T^* = 0) = \alpha_0 \end{aligned}$$



from which it follows by the Law of Total Probability that

$$p_k = P(T = 1|z = z_k) = (1 - \alpha_1)p_k^* + \alpha_0 \{1 - p_k^*\}$$

where  $p_k^* = P(T^* = 1|z = z_k)$  and thus,

$$p_k^* = \frac{p_k - \alpha_0}{1 - \alpha_0 - \alpha_1}$$

Now, by Bayes' Rule and again using the fact that the mis-classification probabilities do not depend on  $z$ , so that conditioning on  $z = z_k$  is superfluous given that we have already conditioned on  $T^*$ , we have

$$\begin{aligned} \frac{\beta_{OLS}^k}{\beta} &= P(T^* = 1|T = 1, z = z_k) - P(T^* = 1|T = 0, z = z_k) \\ &= \frac{P(T = 1|T^* = 1, z = z_k)p_k^*}{p_k} - \frac{P(T = 1|T^* = 1, z = z_k)p_k^*}{(1 - p_k)} \\ &= \left( \frac{p_k - \alpha_0}{1 - \alpha_0 - \alpha_1} \right) \left[ \frac{(1 - p_k)(1 - \alpha_1) - p_k\alpha_0}{p_k(1 - p_k)} \right] \end{aligned}$$

This is *exactly* the same expression as  $M(\alpha_0, \alpha_1, \cdot)$ , only evaluated at  $p_k$  rather than  $p$ . This means that we can re-use the algebra from above:

(11)

$$\beta_{OLS}^k = \beta M(\alpha_0, \alpha_1, p_k) = \beta \left\{ \frac{1}{1 - \alpha_0 - \alpha_1} \left[ 1 - \frac{\alpha_0(1 - \alpha_1)}{p_k} - \frac{(1 - \alpha_0)\alpha_1}{1 - p_k} \right] \right\}$$

which gives us an equation for *each* value  $z_k$  in the support of  $z$ . Solving the expressions for  $\beta_{OLS}$  and  $\beta_{OLS}^k$  for  $\beta$  and equating them yields an equation of the form

$$\beta_{OLS}^k M(\alpha_0, \alpha_1, p) = \beta_{OLS} M(\alpha_0, \alpha_1, p_k)$$

for *each* value of  $k$ . Using our expression for the function  $M$  from above and

multiplying through by  $1 - \alpha_0 - \alpha_1$  gives

$$\beta_{OLS} \left[ 1 - \frac{(1 - \alpha_1)\alpha_0}{p_k} - \frac{(1 - \alpha_0)\alpha_1}{1 - p_k} \right] = \beta_{OLS}^k \left[ 1 - \frac{(1 - \alpha_1)\alpha_0}{p} - \frac{(1 - \alpha_0)\alpha_1}{1 - p} \right] = 0$$

Rearranging this expression gives

$$(1 - \alpha_1)\alpha_0 \left[ \frac{\beta_{OLS}^k}{p} - \frac{\beta_{OLS}}{p_k} \right] + (1 - \alpha_0)\alpha_1 \left[ \frac{\beta_{OLS}^k}{1 - p} - \frac{\beta_{OLS}}{1 - p_k} \right] = \beta_{OLS}^k - \beta_{OLS}$$

which is an expression of the form

$$B_0 w_0^k + B_1 w_1^k = w_2^k$$

where the unknowns  $B_0, B_1$  are defined as

$$B_0 = \alpha_0(1 - \alpha_1)$$

$$B_1 = \alpha_1(1 - \alpha_0)$$

and the observable constants  $w_0^k, w_1^k, w_2^k$  are

$$w_0^k = \frac{\beta_{OLS}^k}{p} - \frac{\beta_{OLS}}{p_k}$$

$$w_1^k = \frac{\beta_{OLS}^k}{1 - p} - \frac{\beta_{OLS}}{1 - p_k}$$

$$w_2^k = \beta_{OLS}^k - \beta_{OLS}$$

Since we have an equation for each value of  $k$ , we have a linear system of  $k$  equations in two unknowns. One of these equations, however is redundant.

Need to prove this.

Thus,  $z$  must take at least *three* values for the system to have a solution.

In matrix form, we have

$$\begin{bmatrix} w_0^1 & w_1^1 \\ w_0^2 & w_1^2 \end{bmatrix} \begin{bmatrix} B_0 \\ B_1 \end{bmatrix} = \begin{bmatrix} w_2^1 \\ w_2^2 \end{bmatrix}$$

and as long as  $w_0^1 w_1^2 - w_0^2 w_1^1 \neq 0$  (Lewbel's Assumption A5) the solution is

$$\begin{bmatrix} B_0 \\ B_1 \end{bmatrix} = \frac{1}{w_0^1 w_1^2 - w_0^2 w_1^1} \begin{bmatrix} w_1^2 w_2^1 - w_1^1 w_2^2 \\ w_0^1 w_2^2 - w_0^2 w_2^1 \end{bmatrix}$$

Finally, given that  $B_0 = (1 - \alpha_1)\alpha_0$  and  $B_1 = (1 - \alpha_0)\alpha_1$ , we can solve for the mis-classification rates as follows. First, rearranging the definition of  $B_1$  gives  $\alpha_0 = 1 - B_1/\alpha_1$ . Substituting this into the definition of  $B_0$ , we see that  $B_0 = (1 - B_1/\alpha_1)(1 - \alpha_1)$ , yielding the following quadratic equation

$$\alpha_1^2 - (1 - B_0 + B_1)\alpha_1 + B_1 = 0$$

Solving, we find that

$$\alpha_1 = \frac{1}{2} \left[ 1 - B_0 + B_1 \pm \sqrt{(1 - B_0 + B_1)^2 - 4B_1} \right]$$

and since  $\alpha_0 = B_0/(1 - \alpha_1)$

$$\alpha_0 = \frac{B_0}{1 - \frac{1}{2} \left[ 1 - B_0 + B_1 \pm \sqrt{(1 - B_0 + B_1)^2 - 4B_1} \right]}$$

To determine the nature of these solutions, we re-express the discriminant:

$$\begin{aligned}
 (1 - B_0 + B_1)^2 - 4B_1 &= [1 - \alpha_0(1 - \alpha_1) + (1 - \alpha_0)\alpha_1]^2 - 4\alpha_1(1 - \alpha_0) \\
 &= (1 - \alpha_0 + \alpha_1)^2 - 4\alpha_1(1 - \alpha_0) \\
 &= [(1 - \alpha_0 - \alpha_1) + 2\alpha_1]^2 - 4\alpha_1(1 - \alpha_0) \\
 &= [(1 - \alpha_0 - \alpha_1)^2 - 4\alpha_1(1 - \alpha_0)] - 4(\alpha_1 - \alpha_0\alpha_1) \\
 &= (1 - \alpha_0 - \alpha_1)^2
 \end{aligned}$$

Since the discriminant is necessarily positive, the solutions are always real. But more importantly, we have established that

$$1 - \alpha_0 - \alpha_1 = \pm \sqrt{(1 - B_0 + B_1)^2 - 4B_1}$$

Since  $\beta = \beta_{IV}(1 - \alpha_0 - \alpha_1)$ , this expression *alone* is sufficient to identify  $\beta$ : there is no need to substitute the solutions for  $(\alpha_0, \alpha_1)$  into  $M(\alpha_0, \alpha_1, p)$ . Moreover, it clarifies the role of the assumption  $1 - \alpha_0 - \alpha_1 > 0$ : without this restriction we learn the *magnitude* of  $\beta$  but not the sign.

In page 544 Lewbel observes that if his instrument were binary (if we only had one instrument in our case), identification could be achieved with one additional restriction on the missclassification rates. One such restriction is implied by homoskedasticity on the instrument, which he does not mention.

I like the way that Lewbel sets things up. I wonder whether a similar way of doing things would simplify some of our other calculations.

## 5. IDENTIFICATION BY HOMOSKEDASTICITY

This section uses our notation rather than Mahajan's. We'll have to decide what notation we want to use in the paper itself but for the moment I'm trying to avoid confusion by talking about Mahajan's proofs using his own notation while keeping our derivations in the same notation we used on the whiteboard. I think that by assuming the instrument takes on three values (as in Lewbell) and imposing our homoskedasticity assumption we'll get identification in the case where  $T^*$  is endogenous so I've written out this derivation for arbitrary discrete  $z$ .

Now suppose that one is prepared to assume that

$$(12) \quad E[u^2|z] = E[u^2].$$

When combined with the usual IV assumption,  $E[u|z] = 0$ , this implies  $Var(u|z) = Var(u)$ . Whether this assumption is reasonable, naturally, depends on the application. When  $z$  is the offer of treatment in a randomized controlled trial, for example, Equation 12 holds automatically as a consequence of the randomization. Similarly, in studies based on a “natural” rather than controlled experiment one typically argues that the instrument is not merely uncorrelated with  $u$  but *independent* of it, so that Equation 12 follows.

To see why homoskedasticity with respect to the instrument provides additional identifying information, first express the conditional variance of  $y$  as follows

$$(13) \quad Var(y|z) = \beta^2 Var(T^*|z) + Var(u|z) + 2\beta Cov(T^*, u|z)$$

Under 12,  $Var(u|z)$  does not depend on  $z$ . Hence the *difference* of conditional variances evaluated at two values  $z_a$  and  $z_b$  in the support of  $z$  is

simply

$$(14) \quad \Delta Var(y|z_a, z_b) = \beta^2 \Delta Var(T^*|z_a, z_b) + 2\beta \Delta Cov(T^*, u|z_a, z_b)]$$

Where  $\Delta Var(y|z_a, z_b) = Var(y|z = z_a) - Var(y|z = z_b)$ , and we define  $\Delta Var(T^*|z_a, z_b)$  and  $\Delta Cov(T^*, u|z_a, z_b)$  analogously.

First we simplify the  $\Delta Var(T^*|z_a, z_b)$  term. Since  $T$  is conditionally independent of  $z$  given  $T^*$ ,

$$\begin{aligned} P(T = 1|z) &= E_{T^*|z} [E(T|z, T^*)] = E_{T^*|z} [E(T|T^*)] \\ &= P(T^* = 1|z) (1 - \alpha_1) + [1 - P(T^* = 1|z)] \alpha_0 \\ &= \alpha_0 + (1 - \alpha_0 - \alpha_1) P(T^* = 1|z) \end{aligned}$$

Rearranging,

$$(15) \quad P(T^* = 1|z) = \frac{P(T = 1|z) - \alpha_0}{1 - \alpha_0 - \alpha_1}$$

and accordingly,

$$(16) \quad Var(T^*|z) = \frac{[P(T = 1|z) - \alpha_0] [1 - P(T = 1|z) - \alpha_1]}{(1 - \alpha_0 - \alpha_1)^2}$$

Thus, evaluating Equation 16 at  $z_a$  and  $z_b$  and simplifying,

$$(17) \quad \Delta Var(T^*|z_a, z_b) = \frac{\Delta Var(T|z_a, z_b) + (\alpha_0 - \alpha_1) \Delta E(T|z_a, z_b)}{(1 - \alpha_0 - \alpha_1)^2}$$

Turning our attention to  $\Delta Cov(T^*, u|z_a, z_b)$  first note that

$$(18) \quad Cov(T^*, u|z) = E_{T^*|z} [E(T^* u|z, T^*)] = P(T^* = 1|z) E(u|T^* = 1, z)$$

since  $E[z|u] = 0$ . Combining this with Equation 15 and evaluating at  $z_a$

and  $z_b$  gives

$$(19) \quad \Delta Cov(T^*, u|z_a, z_b) = \frac{[E(T|z_a) - \alpha_0] m_{1a} - [E(T|z_b) - \alpha_0] m_{1b}}{1 - \alpha_0 - \alpha_1}$$

where  $m_{1a} = E[u|T^* = 1, z_a]$  and  $m_{1b} = E[u|T^* = 1, z_b]$ .

Both Equations 17 and 19 involve only observable quantities and the mis-classification rates  $\alpha_0$  and  $\alpha_1$ . Equation 14, however, also involves  $\beta$ . Fortunately we can eliminate this quantity as follows. First, let  $\mathcal{W}(z_a, z_b)$  denote the Wald Estimator of  $\beta$  given by

$$(20) \quad \mathcal{W}(z_a, z_b) = \frac{E(y|z_a) - E(y|z_b)}{E(T|z_a) - E(T|z_b)}$$

Since  $E(u|z) = 0$ ,

$$E(y|z_a) - E(y|z_b) = \beta [E(T^*|z_a) - E(T^*|z_b)]$$

and by Equation 15,

$$E(T|z_a) - E(T|z_b) = (1 - \alpha_0 - \alpha_1) [E(T^*|z_a) - E(T^*|z_b)]$$

thus we find that

$$(21) \quad \beta = (1 - \alpha_0 - \alpha_1) \mathcal{W}(z_a, z_b).$$

Finally, combining Equations 14, 17, 19 and 21 we have

$$(22) \quad \begin{aligned} \Delta Var(y|z_a, z_b) &= \mathcal{W}(z_a, z_b)^2 \{ \Delta Var(T|z_a, z_b) + (\alpha_0 - \alpha_1) \Delta E(T|z_a, z_b) \} \\ &\quad + 2\mathcal{W}(z_a, z_b) \{ [E(T|z_a) - \alpha_0] m_{1a} - [E(T|z_b) - \alpha_0] m_{1b} \} \end{aligned}$$

an equation relating  $\alpha_0, \alpha_1, m_{1a}$  and  $m_{1b}$  to various observable quantities.

Equation 22 provides an additional identifying restriction for each unique pair of values  $(z_a, z_b)$  in the support of  $z$ . If  $z$  takes on two values it provides one restriction, whereas if  $z$  takes on three values it provides two restrictions, and so on. To take a particularly simple example, suppose that  $z$  is binary and Mahajan's (2006) assumption that  $E[u|z, T^*] = 0$  holds. Then Equation 22 reduces to

$$\Delta Var(y|1, 0) = \left[ \frac{Cov(z, y)}{Cov(z, T)} \right]^2 \left\{ \Delta Var(T|1, 0) + (\alpha_0 - \alpha_1) \left[ \frac{Cov(z, T)}{Var(z)} \right] \right\}$$

Rearranging, we see that

$$\alpha_0 - \alpha_1 = \Delta Var(y|1, 0) \left[ \frac{Cov(z, T) Var(z)}{Cov(z, y)^2} \right] - \Delta Var(T|1, 0) \left[ \frac{Var(z)}{Cov(z, T)} \right]$$

In other words, the homoskedasticity restriction identifies the *difference* between the mis-classification rates. This makes intuitive sense. Provided that the variance of  $u$  is unrelated to  $z$  the only way that the variance of  $y$  can differ across values of  $z$  is if some values of  $z$  provide *more* information about the distribution of  $T^*$  than others. This is only possible if the mis-classification rates differ.

Of course, one need not impose the restriction that  $E[u|z, T^*] = 0$  to use the identifying information provided by Equation 22. Indeed, by exploiting homoskedasticity with respect to the instrument we can identify  $\beta$  using weaker conditions than Mahajan (2006) without requiring that  $z$  take on three or more values, as in Lewbel (2007). Moreover, when  $z$  does take on three or more values we can identify  $\beta$  even when  $T^*$  is endogenous.

I'm pretty sure this is true, but we do still need to prove it!

In the general case where we do not impose Mahajan's assumption that  $E[u|z, T^*] = 0$  the purpose of the homoskedasticity restrictions is to eliminate a quantity that appears in the moment condition that arises from the



“modified IV estimator” in which  $\tilde{z} \equiv T(z - E[z])$  is used as an instrument for  $T$ . We showed previously that

$$\tilde{\beta}_{IV} = \beta \left[ \frac{(1-p-\alpha_1) + \alpha_0}{(1-p)(1-\alpha_0-\alpha_1)} \right] + \left[ \frac{(1-\alpha_0-\alpha_1) \{E[zT^*u] - E[z]E[T^*u]\}}{(1-p)Cov(z, T)} \right]$$

First consider the case in which  $T^*$  is exogenous, so that  $E(T^*u) = 0$ , and  $z$  is binary. Then the preceding reduces to

$$\tilde{\beta}_{IV} = \beta \left[ \frac{(1-p-\alpha_1) + \alpha_0}{(1-p)(1-\alpha_0-\alpha_1)} \right] + \left[ \frac{(1-\alpha_0-\alpha_1) E[zT^*u]}{(1-p)Cov(z, T)} \right]$$

where

$$E[zT^*u] = E_{T^*,z} [E(zT^*u|z, T^*)] = p_{11}m_{11}$$

where  $p_{jk} = P(T^* = j, z = k)$  and  $m_{jk} = E[u|T^* = j, z = k]$ . Note that by the definition of conditional probability we can equivalently express this as

$$E[zT^*u] = E(T^*|z = 1)P(z = 1)m_{11}$$

Thus we can rewrite the numerator of the second term in the expression for  $\tilde{\beta}$  from above as

$$\begin{aligned} C &= (1-\alpha_0-\alpha_1)E(zT^*u) \\ &= (1-\alpha_0-\alpha_1)E(T^*|z = 1)P(z = 1)m_{11} \\ &= [E(T|z = 1) - \alpha_0] P(z = 1)m_{11} \end{aligned}$$

using Equation 15. Thus, when  $T^*$  is exogenous and  $z$  is binary, the expression for  $\tilde{\beta}_{IV}$  can be written as

$$(23) \quad \tilde{\beta}_{IV} = \beta \left[ \frac{(1-p-\alpha_1) + \alpha_0}{(1-p)(1-\alpha_0-\alpha_1)} \right] + \left[ \frac{P(z = 1) [E(T|z = 1) - \alpha_0] m_{11}}{(1-p)Cov(z, T)} \right]$$

Now we will show that the second term from Equation 22 can be expressed

in a similar fashion. The term in question is:

$$D = [E(T|z = 1) - \alpha_0] m_{11} - [E(T|z = 0) - \alpha_0] m_{10}$$

Imposing  $Cov(T^*, u) = 0$  gives  $p_{10}m_{10} + p_{11}m_{11} = 0$ . Thus  $m_{10} = -p_{11}m_{11}/p_{10}$ .

Now, by Equation 15,

$$-\frac{p_{11}}{p_{10}} = -\frac{P(T^* = 1|z = 1)P(z = 1)}{P(T^* = 1|z = 0)P(z = 0)} = -\frac{[E(T|z = 1) - \alpha_0] P(z = 1)}{[E(T|z = 0) - \alpha_0] P(z = 0)}$$

Substituting this into the expression for  $D$ , we have

$$D = \left[ \frac{E(T|z = 1) - \alpha_0}{P(z = 0)} \right] m_{11}$$

and therefore, in the case where  $z$  is binary and  $T^*$  is exogenous Equation 22 simplifies to

$$(24) \quad \Delta Var(y|z) = \mathcal{W}^2 \{ \Delta Var(T|z) + (\alpha_0 - \alpha_1) \Delta E(T|z) \} + 2\mathcal{W} \left\{ \frac{E(T|z = 1) - \alpha_0}{P(z = 0)} \right\} m_{11}$$

I think this will make things easier to solve because we could treat the quantity  $[E(T|z = 1) - \alpha_0] m_{11}$  as a unit and eliminate it from the system. But I could be wrong...

## 6. THE CASE OF A CONTINUOUS INSTRUMENT

I think we'll be able to say something helpful about the case in which the instrument is continuous rather than discrete. For one, we can explain how one could discretize a continuous instrument provided that one is willing to assume that it is independent of the regression error term.

But it also looks like some of the equations we wrote down go through

without modification when  $z$  is continuous. For example define the first-stage functions  $\pi^*(z) = E(T^*|z)$  and  $\pi(z) = E(T|z)$ . Then by precisely the same argument as given above in the section on identification by homoskedasticity,

$$(25) \quad E[T^*|z] = \frac{E[T|z] - \alpha_0}{1 - \alpha_0 - \alpha_1}$$

so the observable first-stage is just a shifted and scaled version of the “true” first-stage. Proceeding similarly for the reduced form, we have

$$(26) \quad E[y|z] = \alpha + \beta E[T^*|z]$$

Substituting the expression for the first-stage gives

$$(27) \quad E[y|z] = \alpha + \beta \left( \frac{E[T|z] - \alpha_0}{1 - \alpha_0 - \alpha_1} \right)$$

and since

$$(28) \quad E[y] = \alpha + \beta E[T^*] = \alpha + \beta \left( \frac{E[T] - \alpha_0}{1 - \alpha_0 - \alpha_1} \right)$$

we see that

$$(29) \quad \frac{E[y|z] - E[y]}{E[T|z] - E[T]} = \frac{\beta}{1 - \alpha_0 - \alpha_1}$$

which is very similar to the expressions we wrote down before for IV but in this case involves an *arbitrary* reduced form function  $E[y|z]$  and an *arbitrary* first-stage  $E[T|z]$ . What is particularly interesting about the preceding expression is that it is *just as informative* when  $z$  is binary as when it takes on three values or is continuous, assuming the model is correct. However it provides over-identifying information since we can evaluate the function at

any value of  $z$ : we should get the same result in each case.

I wonder if we can proceed like this for all of the other conditions we use in our analysis. I also wonder whether a continuous instrument makes identification any easier. Can we get by without the homoskedasticity assumption? My guess is that we can since we could try to extend Lewbel to an instrument with more than three values, the answer is yes.

I think there's no avoiding the fact that we get identification from a non-linearity in the first stage, which is something that Angrist and Pischke, for example, criticize in *Mostly Harmless*. For example, I think the Lewbel determinant condition must imply this.

It would be nice to link what we do with the Hausman et al work and also situations in which people use a first-stage probit, etc.

## 7. CONCLUSION

Conclusion goes here.

## REFERENCES

- Aigner, D. J., 1973. Regression with a binary independent variable subject to errors of observation. *Journal of Econometrics* 1, 49–60.
- Black, D. A., Berger, M. C., Scott, F. A., September 2000. Bounding parameter estimates with nonclassical measurement error. *Journal of the American Statistical Association* 95 (451), 739–748.
- Bollinger, C. R., 1996. Bounding mean regressions when a binary regressor is mismeasured. *Journal of Econometrics* 73, 387–399.
- Card, D., July 1996. The effects of unions on the structure of wages: a longitudinal analysis. *Econometrica* 64 (4), 957–979.
- Frazis, H., Loewenstein, M. A., 2003. Estimating linear regressions with mismeasured, possibly endogenous, binary explanatory variables. *Journal of Econometrics* 117, 151–178.
- Kane, T. J., Rouse, C. E., Staiger, D., July 1999. Estimating returns to schooling when schooling is misreported. Tech. rep., National Bureau of Economic Research, nBER Working Paper 7235.