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

July 20, 2015

Submitted to Econometrica

ON THE USE OF INSTRUMENTAL VARIABLES TO IDENTIFY THE EFFECT OF A MIS-MEASURED, BINARY REGRESSOR[1]

FRANCIS J. DITRAGLIA[a] AND CAMILO GARCIA-JIMENO[a]

Abstract goes here.

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

1. INTRODUCTION

Introduction goes here.

2. NOTES ON MAHAJAN (2006)

Mahajan (2006) considers regression models of the form

$$E\left[y - g(x^*, z)\right] = 0 \tag{1}$$

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.

[1]We thank...

[a]Dept. of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104, U.S.A.; fditra@econ.upenn.edu; gcamilo@econ.upenn.edu

2.1. *The Case of Exogenous x^**

The first is based on the restriction

$$E\left[y - g(x^*, z) \mid x^*, x, z, v\right] = 0 \tag{2}$$

2.2. *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

$$y = g^*(x^*, z) + \varepsilon \tag{3}$$

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

$$E\left[\varepsilon | z, v\right] = 0 \tag{4}$$

and Equation 2 is replaced by

$$E\left[y | x^*, x, z, v\right] = E\left[y | x^*, z\right] \tag{5}$$

Unfortunately, Mahajan's proof is incorrect and the model in Equation 3 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 3–5 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.[1]

Without covariates we can write

$$y = \alpha + \beta x^* + \varepsilon \tag{6}$$

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

$$m_{jk} = E\left[\varepsilon | x^* = j, v = k\right] \tag{7}$$

3. LEWBEL (2007)

Lewbel shows that under an exogenous but missclasified 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,

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

which implies that

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

[1] Because the covariates are held fixed throughout the proof of Mahajan's Theorem 1, there is no loss of generality.

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 β^k_{OLS} 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

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

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

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

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

$$\beta^k_{OLS} = E\left(y|T = 1, z = z_k\right) - E(y|T = 0, z = z_k)$$

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

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

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

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

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

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

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:

$$\beta^k_{OLS} = \beta M(\alpha_0, \alpha_1, p_k) = \beta\left\{\frac{1}{1-\alpha_0-\alpha_1}\left[1-\frac{\alpha_0(1-\alpha_1)}{p_k} - \frac{(1-\alpha_0)\alpha_1}{1-p_k}\right]\right\} \tag{8}$$

which gives us an equation for *each* value z_k in the support of z. Solving the expressions for β_{OLS} and β^k_{OLS} for β 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

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

Still need to type in the part about the terms under the square root. Show that there are always two real solutions: one positive and one negative, both equal in magnitude. To tell which is the solution we need to impose the condition that the measurement error isn't too severe.

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:

$$B_1 = \alpha_1 - \alpha_1 \frac{B_0}{1 - \alpha_1}$$

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

and

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

Given knowledge of (α_0, α_1) we can solve for β using $\beta = \beta_{OLS}/M(\alpha_0, \alpha_1, p)$.

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

4. 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

$$E[u^2|z] = E[u^2]. \tag{9}$$

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 9 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 9 follows.

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

y as follows

$$(10) \qquad Var(y|z) = \beta^2 Var\left(T^*|z\right) + Var(u|z) + 2\beta Cov(T^*, u|z)$$

Under 9, $Var(u|z)$ 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

$$(11) \qquad \Delta Var(y|z_a, z_b) = \beta^2 \Delta Var(T^*|z_a, z_b) + 2\beta \Delta Cov(T^*, u|z_a, z_b)]$$

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^*,

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

Rearranging,

$$(12) \qquad P(T^* = 1|z) = \frac{P(T=1|z) - \alpha_0}{1 - \alpha_0 - \alpha_1}$$

and accordingly,

$$(13) \qquad Var(T^*|z) = \frac{\left[P\left(T=1|z\right) - \alpha_0\right]\left[1 - P\left(T=1|z\right) - \alpha_1\right]}{\left(1-\alpha_0-\alpha_1\right)^2}$$

Thus, evaluating Equation 13 at z_a and z_b and simplifying,

$$(14) \qquad \Delta Var(T^*|z_a, z_b) = \frac{\Delta Var(T|z_a, z_b) + (\alpha_0 - \alpha_1)\, \Delta E(T|z_a, z_b)}{\left(1-\alpha_0-\alpha_1\right)^2}$$

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

$$Cov(T^*, u|z) = E_{T^*|z}\left[E\left(T^* u|z, T^*\right)\right] = P(T^* = 1|z)E\left(u|T^* = 1, z\right) \tag{15}$$

since $E[z|u] = 0$. Combining this with Equation 12 and evaluating at z_a and z_b gives

$$\Delta Cov(T^*, u|z_a, z_b) = \frac{\left[E\left(T|z_a\right) - \alpha_0\right] m_{1a} - \left[E(T|z_b) - \alpha_0\right] m_{1b}}{1 - \alpha_0 - \alpha_1} \tag{16}$$

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

Both Equations 14 and 16 involve only observable quantities and the mis-classification rates α_0 and α_1. Equation 11, however, also involves β. Fortunately we can eliminiate this quantity as follows. First, let $\mathcal{W}(z_a, z_b)$ denote the Wald Estimator of β given by

$$\mathcal{W}(z_a, z_b) = \frac{E(y|z_a) - E(y|z_b)}{E(T|z_a) - E(T|z_b)} \tag{17}$$

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

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

and by Equation 12,

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

thus we find that

$$\beta = (1 - \alpha_0 - \alpha_1)\mathcal{W}(z_a, z_b). \tag{18}$$

Finally, combining Equations 11, 14, 16 and 18 we have

(19)
$$\Delta Var(y|z_a, z_b) = \mathcal{W}\left(z_a, z_b\right)^2 \left\{\Delta Var(T|z_a, z_b) + (\alpha_0 - \alpha_1)\Delta E(T|z_a, z_b)\right\}$$
$$+ 2\mathcal{W}(z_a, z_b)\left\{\left[E\left(T|z_a\right) - \alpha_0\right] m_{1a} - \left[E(T|z_b) - \alpha_0\right] m_{1b}\right\}$$

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

Equation 19 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 if 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 19 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 19. Indeed, by exploiting homoskedasticity with respect to the instrument we can identify β 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 β 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 $\widetilde{z} \equiv T\,(z - E[z])$ is used as an instrument for T. We showed previously that

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

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

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

where

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

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 $\widetilde{\beta}$ from above as

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

using Equation 12. Thus, when T^* is exogenous and z is binary, the expression for $\widetilde{\beta}_{IV}$ can be written as

$$(20) \quad \widetilde{\beta}_{IV} = \beta\left[\frac{(1-p-\alpha_1)+\alpha_0}{(1-p)(1-\alpha_0-\alpha_1)}\right] + \left[\frac{P(z=1)\left[E(T|z=1)-\alpha_0\right]m_{11}}{(1-p)Cov(z,T)}\right]$$

Now we will show that the second term from Equation 19 can be expressed in a similar fashion. The term in question is:

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

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 12,

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

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 19 simplifies to

(21)

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

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

5. 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,

$$E[T^*|z] = \frac{E[T|z] - \alpha_0}{1 - \alpha_0 - \alpha_1} \tag{22}$$

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

$$E[y|z] = \alpha + \beta E[T^*|z] \tag{23}$$

Substituting the expression for the first-stage gives

$$E[y|z] = \alpha + \beta \left(\frac{E[T|z] - \alpha_0}{1 - \alpha_0 - \alpha_1} \right) \tag{24}$$

and since

$$E[y] = \alpha + \beta E[T^*] = \alpha + \beta \left(\frac{E[T] - \alpha_0}{1 - \alpha_0 - \alpha_1} \right) \tag{25}$$

we see that

$$\frac{E[y|z] - E[y]}{E[T|z] - E[T]} = \frac{\beta}{1 - \alpha_0 - \alpha_1} \tag{26}$$

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 nonlinearity in the first stage, which is something that Angrist and Pishcke, 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.

6. CONCLUSION

Conclusion goes here.