

# Notes for Paper on Mis-measured, Binary, Endogenous Regressors

Francis J. DiTraglia & Camilo García-Jimeno

May 12, 2017

## 1 Model and Notation

### Probabilities

$$\begin{aligned} p_{tk}^* &= P(T^* = t, Z = k) \\ p_{tk} &= P(T = t, Z = k) \\ p_k^* &= P(T^* = 1|Z = k) \\ p_k &= P(T = 1|Z = k) \\ q &= P(Z = 1) \end{aligned}$$

$$\begin{aligned} p_{00}^* &= P(T^* = 0|Z = 0)P(Z = 0) = (1 - p_0^*)(1 - q) = \left( \frac{1 - p_0 - \alpha_1}{1 - \alpha_0 - \alpha_1} \right) (1 - q) \\ p_{10}^* &= P(T^* = 1|Z = 0)P(Z = 0) = p_0^*(1 - q) = \left( \frac{p_0 - \alpha_0}{1 - \alpha_0 - \alpha_1} \right) (1 - q) \\ p_{01}^* &= P(T^* = 0|Z = 1)P(Z = 1) = (1 - p_1^*)q = \left( \frac{1 - p_1 - \alpha_1}{1 - \alpha_0 - \alpha_1} \right) q \\ p_{11}^* &= P(T^* = 1|Z = 1)P(Z = 1) = p_1^*(1 - q) = \left( \frac{p_1 - \alpha_0}{1 - \alpha_0 - \alpha_1} \right) q \end{aligned}$$

**CDFs** For  $t, Z \in \{0, 1\}$  define

$$\begin{aligned} F_{tk}^*(\tau) &= P(Y \leq \tau | T^* = t, Z = k) \\ F_{tk}(\tau) &= P(Y \leq \tau | T = t, Z = k) \\ F_k(\tau) &= P(Y \leq \tau | Z = k) \end{aligned}$$

Note that the second two are observed for all  $t, k$  while the first is never observed since it depends on the unobserved RV  $T^*$ .

## 2 Weakest Bounds on $\alpha_0, \alpha_1$

Assume that  $\alpha_0 + \alpha_1 < 1$  that  $T$  is independent of  $Z$  conditional on  $T^*$ . These standard assumptions turn out to yield informative bounds on  $\alpha_0$  and  $\alpha_1$  without *any further restrictions of any kind*. In particular, we assume nothing about the validity of the instrument  $Z$  and nothing about the relationship between the mis-classification error and the outcome  $Y$ : we impose only that the mis-classification error rates do not depend on  $z$  and that the mis-classification is not so bad that  $1 - T$  is a better measure of  $T^*$  than  $T$ .

By the Law of Total Probability and the assumption that  $T$  is conditionally independent of  $Z$  given  $T^*$ ,

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

and similarly

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

and hence

$$\begin{aligned} p_k - \alpha_0 &= (1 - \alpha_0 - \alpha_1)p_k^* \\ (1 - p_k) - \alpha_1 &= (1 - \alpha_0 - \alpha_1)(1 - p_k^*) \end{aligned}$$

Now, since  $p_k^*$  and  $(1 - p_k^*)$  are probabilities they are between zero and one which means that the sign of  $p_k - \alpha_0$  as well as that of  $(1 - p_k) - \alpha_1$  are both determined by that of  $1 - \alpha_0 - \alpha_1$ .

Accordingly, provided that  $1 - \alpha_0 - \alpha_1 < 1$ , we have

$$\begin{aligned}\alpha_0 &< p_k \\ \alpha_1 &< (1 - p_k)\end{aligned}$$

so long as  $p_k^*$  does not equal zero or one, which is not a realistic case for any example that we consider. Since these bounds hold for all  $k$ , we can take the tightest bound over all values of  $Z$ .

Important: using these to bound  $\beta$  gives  $\beta \in [\text{ITT}, \text{Wald}]$ .

### 3 Stronger Bounds for $\alpha_0, \alpha_1$

Now suppose we add the assumption that  $T$  is conditionally independent of  $Y$  given  $T^*$ . This is essentially the non-differential measurement error assumption although it is slightly stronger than the version used by Mahajan (2006) who assumes only conditional mean independence. This assumption allows us to considerably strengthen the bounds from the preceding section by exploiting information contained in the conditional distribution of  $Y$  given  $T$  and  $Z$ . The key ingredient is a relationship that we can derive between the unobservable distributions  $F_{tk}^*$  and the observable distributions  $F_{tk}$  using this new conditional independence assumption. To begin, note that by Bayes' rule we have

$$\begin{aligned}P(T^* = 1|T = 1, Z = k) &= P(T = 1|T^* = 1) \left( \frac{p_k^*}{p_k} \right) = (1 - \alpha_1) \left( \frac{p_k^*}{p_k} \right) \\ P(T^* = 1|T = 0, Z = k) &= P(T = 0|T^* = 1) \left( \frac{p_k^*}{1 - p_k} \right) = \alpha_1 \left( \frac{p_k^*}{1 - p_k} \right) \\ P(T^* = 0|T = 1, Z = k) &= P(T = 1|T^* = 0) \left( \frac{1 - p_k^*}{p_k} \right) = \alpha_0 \left( \frac{1 - p_k^*}{p_k} \right) \\ P(T^* = 0|T = 0, Z = k) &= P(T = 0|T^* = 0) \left( \frac{1 - p_k^*}{1 - p_k} \right) = (1 - \alpha_0) \left( \frac{1 - p_k^*}{1 - p_k} \right)\end{aligned}$$

Now, by the conditional independence assumption

$$\begin{aligned}P(Y \leq \tau|T^* = 0, T = t, Z = k) &= P(Y \leq \tau|T^* = 0, Z = k) = F_{0k}^*(\tau) \\ P(Y \leq \tau|T^* = 1, T = t, Z = k) &= P(Y \leq \tau|T^* = 1, Z = k) = F_{1k}^*(\tau)\end{aligned}$$

Finally, putting everything together using the Law of Total Probability, we find that

$$\begin{aligned}(1 - p_k)F_{0k}(\tau) &= (1 - \alpha_0)(1 - p_k^*)F_{0k}^*(\tau) + \alpha_1 p_k^* F_{1k}^*(\tau) \\ p_k F_{1k}(\tau) &= \alpha_0(1 - p_k^*)F_{0k}^*(\tau) + (1 - \alpha_1)p_k^* F_{1k}^*(\tau)\end{aligned}$$

for all  $k$ . Defining the shorthand

$$\begin{aligned}\tilde{F}_{0k}(\tau) &\equiv (1 - p_k)F_{0k}(\tau) \\ \tilde{F}_{1k}(\tau) &\equiv p_k F_{1k}(\tau)\end{aligned}$$

this becomes

$$\tilde{F}_{0k}(\tau) = (1 - \alpha_0)(1 - p_k^*)F_{0k}^*(\tau) + \alpha_1 p_k^* F_{1k}^*(\tau) \quad (1)$$

$$\tilde{F}_{1k}(\tau) = \alpha_0(1 - p_k^*)F_{0k}^*(\tau) + (1 - \alpha_1)p_k^* F_{1k}^*(\tau) \quad (2)$$

Now, solving Equation 1 for  $p_k^* F_{1k}^*(\tau)$  we have

$$p_k^* F_{1k}^*(\tau) = \frac{1}{\alpha_1} \left[ \tilde{F}_{0k}(\tau) - (1 - \alpha_0)(1 - p_k^*)F_{0k}^*(\tau) \right]$$

Substituting this into Equation 2,

$$\begin{aligned}\tilde{F}_{1k}(\tau) &= \alpha_0(1 - p_k^*)F_{0k}^*(\tau) + \frac{1 - \alpha_1}{\alpha_1} \left[ \tilde{F}_{0k}(\tau) - (1 - \alpha_0)(1 - p_k^*)F_{0k}^*(\tau) \right] \\ &= \frac{1 - \alpha_1}{\alpha_1} \tilde{F}_{0k}(\tau) + \left[ \alpha_0 - \frac{(1 - \alpha_1)(1 - \alpha_0)}{\alpha_1} \right] (1 - p_k^*)F_{0k}^*(\tau) \\ &= \frac{1 - \alpha_1}{\alpha_1} \tilde{F}_{0k}(\tau) + \left[ \frac{\alpha_0 \alpha_1 - (1 - \alpha_1)(1 - \alpha_0)}{\alpha_1} \right] (1 - p_k^*)F_{0k}^*(\tau) \\ &= \frac{1 - \alpha_1}{\alpha_1} \tilde{F}_{0k}(\tau) - \left[ \frac{(1 - \alpha_1)(1 - \alpha_0) - \alpha_0 \alpha_1}{\alpha_1} \right] (1 - p_k^*)F_{0k}^*(\tau) \\ &= \frac{1 - \alpha_1}{\alpha_1} \tilde{F}_{0k}(\tau) - \left[ \frac{1 - \alpha_1 - \alpha_0}{\alpha_1} \right] \left( \frac{1 - p_k - \alpha_1}{1 - \alpha_0 - \alpha_1} \right) F_{0k}^*(\tau)\end{aligned}$$

and therefore

$$\tilde{F}_{1k}(\tau) = \frac{1 - \alpha_1}{\alpha_1} \tilde{F}_{0k}(\tau) - \frac{1 - p_k - \alpha_1}{\alpha_1} F_{0k}^*(\tau) \quad (3)$$

Equation 3 relates the observable  $\tilde{F}_{1k}(\tau)$  to the mis-classification error rate  $\alpha_1$  and the unobservable CDF  $F_{0k}^*(\tau)$ . Since  $F_{0k}^*(\tau)$  is a CDF, however, it lies in the interval  $[0, 1]$ .

Accordingly, substituting 0 in place of  $F_{0k}^*(\tau)$  gives

$$\tilde{F}_{1k}(\tau) \leq \frac{1 - \alpha_1}{\alpha_1} \tilde{F}_{0k}(\tau) \quad (4)$$

while substituting 1 gives

$$\tilde{F}_{1k}(\tau) \geq \frac{1 - \alpha_1}{\alpha_1} \tilde{F}_{0k}(\tau) - \frac{1 - p_k - \alpha_1}{\alpha_1} \quad (5)$$

Rearranging Equation 4

$$\begin{aligned} \alpha_1 \tilde{F}_{1k}(\tau) &\leq (1 - \alpha_1) \tilde{F}_{0k}(\tau) \\ \alpha_1 \tilde{F}_{1k}(\tau) &\leq \tilde{F}_{0k}(\tau) - \alpha_1 \tilde{F}_{0k}(\tau) \\ \alpha_1 [\tilde{F}_{0k}(\tau) + \tilde{F}_{1k}(\tau)] &\leq \tilde{F}_{0k}(\tau) \end{aligned}$$

since  $\alpha_1 \in [0, 1]$  and therefore

$$\alpha_1 \leq \frac{\tilde{F}_{0k}(\tau)}{\tilde{F}_{0k}(\tau) + \tilde{F}_{1k}(\tau)} = (1 - p_k) \left[ \frac{F_{0k}(\tau)}{F_k(\tau)} \right] \quad (6)$$

since  $\tilde{F}_{1k}(\tau) + \tilde{F}_{0k}(\tau) \geq 0$ . Proceeding similarly for Equation 5,

$$\begin{aligned} \alpha_1 \tilde{F}_{1k}(\tau) &\geq (1 - \alpha_1) \tilde{F}_{0k}(\tau) - (1 - p_k - \alpha_1) \\ \alpha_1 [\tilde{F}_{1k}(\tau) + \tilde{F}_{0k}(\tau) - 1] &\geq \tilde{F}_{0k}(\tau) - (1 - p_k) \\ -\alpha_1 [1 - \tilde{F}_{1k}(\tau) - \tilde{F}_{0k}(\tau)] &\geq -[1 - \tilde{F}_{0k}(\tau) - p_k] \\ \alpha_1 [1 - \tilde{F}_{1k}(\tau) - \tilde{F}_{0k}(\tau)] &\leq 1 - \tilde{F}_{0k}(\tau) - p_k \end{aligned}$$

Now since  $\tilde{F}_{1k}(\tau) = p_k F_{1k}(\tau) \leq p_k$  and  $\tilde{F}_{0k}(\tau) = (1 - p_k) F_{0k}(\tau) \leq (1 - p_k)$  it follows that  $1 - \tilde{F}_{1k}(\tau) - \tilde{F}_{0k}(\tau) \geq 0$  and hence

$$\alpha_1 \leq \frac{1 - \tilde{F}_{0k}(\tau) - p_k}{1 - \tilde{F}_{1k}(\tau) - \tilde{F}_{0k}(\tau)} = (1 - p_k) \left[ \frac{1 - F_{0k}(\tau)}{1 - F_k(\tau)} \right] \quad (7)$$

The bounds given in Equations 6 and 7 relate  $\alpha_1$  to observable quantities *only* and hold for all values of  $\tau$  for which their respective denominators are non-zero. Moreover, these bounds hold for any value  $k$  that the instrument takes on.

We can proceed similarly for  $\alpha_0$ . First solve Equation 1 for  $(1 - p_k^*)F_{0k}^*(\tau)$ :

$$(1 - p_k^*)F_{0k}^*(\tau) = \frac{1}{1 - \alpha_0} \left[ \tilde{F}_{0k}(\tau) - \alpha_1 p_k^* F_{1k}^*(\tau) \right]$$

and then substitute into Equation 2:

$$\begin{aligned} \tilde{F}_{1k}(\tau) &= \frac{\alpha_0}{1 - \alpha_0} \left[ \tilde{F}_{0k}(\tau) - \alpha_1 p_k^* F_{1k}^*(\tau) \right] + (1 - \alpha_1) p_k^* F_{1k}^*(\tau) \\ &= \frac{\alpha_0}{1 - \alpha_0} \tilde{F}_{0k}(\tau) + \left[ (1 - \alpha_1) - \frac{\alpha_0 \alpha_1}{1 - \alpha_0} \right] p_k^* F_{1k}^*(\tau) \\ &= \frac{\alpha_0}{1 - \alpha_0} \tilde{F}_{0k}(\tau) + \left[ \frac{(1 - \alpha_1)(1 - \alpha_0) - \alpha_0 \alpha_1}{1 - \alpha_0} \right] p_k^* F_{1k}^*(\tau) \\ &= \frac{\alpha_0}{1 - \alpha_0} \tilde{F}_{0k}(\tau) + \left[ \frac{1 - \alpha_0 - \alpha_1}{1 - \alpha_0} \right] \frac{p_k - \alpha_0}{1 - \alpha_0 - \alpha_1} F_{1k}^*(\tau) \end{aligned}$$

and therefore

$$\tilde{F}_{1k}(\tau) = \frac{\alpha_0}{1 - \alpha_0} \tilde{F}_{0k}(\tau) + \frac{p_k - \alpha_0}{1 - \alpha_0} F_{1k}^*(\tau) \quad (8)$$

Now we can again obtain two bounds by substituting the smallest and largest possible values of  $F_{1k}^*(\tau)$ . Substituting zero gives

$$\tilde{F}_{1k}(\tau) \geq \frac{\alpha_0}{1 - \alpha_0} \tilde{F}_{0k}(\tau) \quad (9)$$

while substituting one gives

$$\tilde{F}_{1k}(\tau) \leq \frac{\alpha_0}{1 - \alpha_0} \tilde{F}_{0k}(\tau) + \frac{p_k - \alpha_0}{1 - \alpha_0} \quad (10)$$

Now, rearranging Equation 9,

$$\begin{aligned} (1 - \alpha_0) \tilde{F}_{1k}(\tau) &\geq \alpha_0 \tilde{F}_{0k}(\tau) \\ \tilde{F}_{1k}(\tau) &\geq \alpha_0 \left[ \tilde{F}_{0k}(\tau) + \tilde{F}_{1k}(\tau) \right] \end{aligned}$$

since  $1 - \alpha_0 \geq 0$ . Therefore,

$$\alpha_0 \leq \frac{\tilde{F}_{1k}(\tau)}{\tilde{F}_{0k}(\tau) + \tilde{F}_{1k}(\tau)} = p_k \left[ \frac{F_{1k}(\tau)}{F_k(\tau)} \right] \quad (11)$$

since  $\left[\tilde{F}_{0k}(\tau) + \tilde{F}_{1k}(\tau)\right] \geq 0$ . Similarly, rearranging Equation 10

$$\begin{aligned} (1 - \alpha_0)\tilde{F}_{1k}(\tau) &\leq \alpha_0\tilde{F}_{0k}(\tau) + p_k - \alpha_0 \\ \tilde{F}_{1k}(\tau) - p_k &\leq \alpha_0 \left[\tilde{F}_{0k}(\tau) + \tilde{F}_{1k}(\tau) - 1\right] \\ - \left[1 - \tilde{F}_{1k}(\tau) - (1 - p_k)\right] &\leq -\alpha_0 \left[1 - \tilde{F}_{0k}(\tau) - \tilde{F}_{1k}(\tau)\right] \\ \left[1 - \tilde{F}_{1k}(\tau) - (1 - p_k)\right] &\geq \alpha_0 \left[1 - \tilde{F}_{0k}(\tau) - \tilde{F}_{1k}(\tau)\right] \end{aligned}$$

Therefore

$$\alpha_0 \leq \frac{1 - \tilde{F}_{1k}(\tau) - (1 - p_k)}{1 - \tilde{F}_{0k}(\tau) - \tilde{F}_{1k}(\tau)} = p_k \left[ \frac{1 - F_{1k}(\tau)}{1 - F_k(\tau)} \right] \quad (12)$$

**Putting Everything Together** For all  $k$  we have

$$\alpha_0 \leq p_k \min_{\tau} \left\{ \left[ \frac{F_{1k}(\tau)}{F_k(\tau)} \right] \wedge \left[ \frac{1 - F_{1k}(\tau)}{1 - F_k(\tau)} \right] \right\} \leq p_k \quad (13)$$

$$\alpha_1 \leq (1 - p_k) \min_{\tau} \left\{ \left[ \frac{F_{0k}(\tau)}{F_k(\tau)} \right] \wedge \left[ \frac{1 - F_{0k}(\tau)}{1 - F_k(\tau)} \right] \right\} \leq (1 - p_k) \quad (14)$$

Note that these bounds can only improve upon those derived in the previous section since the ratio of CDFs tends to one as  $\tau \rightarrow \infty$ . To derive these tighter bounds we have made no assumption regarding the relationship between  $Z$  and the error term  $\varepsilon$ . These bounds use only the assumption that  $\alpha_0 + \alpha_1 < 1$ , and the assumption that  $T$  is conditionally independent of  $Z, Y$  given  $T^*$ . Notice that that the bounds are related. In particular,

$$p_k \left[ \frac{F_{1k}(\tau)}{F_k(\tau)} \right] = 1 - (1 - p_k) \left[ \frac{F_{0k}(\tau)}{F_k(\tau)} \right]$$

and

$$p_k \left[ \frac{1 - F_{1k}(\tau)}{1 - F_k(\tau)} \right] = 1 - (1 - p_k) \left[ \frac{1 - F_{0k}(\tau)}{1 - F_k(\tau)} \right]$$

## 4 Even Stronger Bounds on $\alpha_0, \alpha_1$

Try applying the stochastic dominance conditions from our simulation study.

## 5 Independent Instrument

Assume that  $Z \perp U$ . The model is  $Y = \beta T^* + U$  and

$$F_U(\tau) = P(U \leq \tau) = P(Y - \beta T^* \leq \tau)$$

but if  $Z$  is independent of  $U$  then it follows that

$$\begin{aligned} F_U(\tau) &= F_{U|Z=k}(\tau) = P(U \leq \tau | Z = k) = P(Y - \beta T^* \leq \tau | Z = k) \\ &= P(Y \leq \tau | T^* = 0, Z = k)(1 - p_k^*) + P(Y \leq \tau + \beta | T^* = 1, Z = k)p_k^* \\ &= (1 - p_k^*)F_{0k}^*(\tau) + p_k^*F_{1k}^*(\tau + \beta) \end{aligned}$$

for all  $k$  by the Law of Total Probability. Similarly,

$$F_k(\tau) = (1 - p_k^*)F_{0k}^*(\tau) + p_k^*F_{1k}^*(\tau)$$

and rearranging

$$(1 - p_k^*)F_{0k}^*(\tau) = F_k(\tau) - p_k^*F_{1k}^*(\tau)$$

Substituting this expression into the equation for  $F_U(\tau)$  from above, we have

$$F_U(\tau) = F_k(\tau) + p_k^*[F_{1k}^*(\tau + \beta) - F_{1k}^*(\tau)]$$

for all  $k$  and all  $\tau$ . Evaluating at two values  $k$  and  $\ell$  in the support of  $Z$  and equating

$$F_k(\tau) + p_k^*[F_{1k}^*(\tau + \beta) - F_{1k}^*(\tau)] = F_\ell(\tau) + p_\ell^*[F_{1\ell}^*(\tau + \beta) - F_{1\ell}^*(\tau)]$$

or equivalently

$$F_k(\tau) - F_\ell(\tau) = p_\ell^*[F_{1\ell}^*(\tau + \beta) - F_{1\ell}^*(\tau)] - p_k^*[F_{1k}^*(\tau + \beta) - F_{1k}^*(\tau)] \quad (15)$$

for all  $\tau$ . Now we simply need to re-express all of the “star” quantities, namely  $p_k^*, p_\ell^*$  and  $F_{1k}^*, F_{1\ell}^*$  in terms of  $\alpha_0, \alpha_1$  and the *observable* probability distributions  $F_{1k}$  and  $F_{1\ell}$  and observable probabilities  $p_k, p_\ell$ . To do this, we use the fact that

$$\begin{aligned} F_{0k}(\tau) &= \frac{1 - \alpha_0}{1 - p_k}(1 - p_k^*)F_{0k}^*(\tau) + \frac{\alpha_1}{1 - p_k}p_k^*F_{1k}^*(\tau) \\ F_{1k}(\tau) &= \frac{\alpha_0}{p_k}(1 - p_k^*)F_{0k}^*(\tau) + \frac{1 - \alpha_1}{p_k}p_k^*F_{1k}^*(\tau) \end{aligned}$$



for all  $k$  by Bayes' rule. Solving these equations,

$$p_k^* F_{1k}^*(\tau) = \frac{1 - \alpha_0}{1 - \alpha_0 - \alpha_1} p_k F_{1k}(\tau) - \frac{\alpha_0}{1 - \alpha_0 - \alpha_1} (1 - p_k) F_{0k}(\tau)$$

for all  $k$ . Combining this with Equation 15, we find that

$$(1 - \alpha_0 - \alpha_1) [F_k(\tau) - F_\ell(\tau)] = \alpha_0 \{ (1 - p_k) [F_{0k}(\tau + \beta) - F_{0k}(\tau)] - (1 - p_\ell) [F_{0\ell}(\tau + \beta) - F_{0\ell}(\tau)] \} \\ - (1 - \alpha_0) \{ p_k [F_{1k}(\tau + \beta) - F_{1k}(\tau)] - p_\ell [F_{1\ell}(\tau + \beta) - F_{1\ell}(\tau)] \}$$

Now, define

$$\Delta_{tk}^\tau(\beta) = F_{tk}(\tau + \beta) - F_{tk}(\tau) = E \left[ \frac{\mathbf{1}\{T = t, Z = k\}}{p_{tk}} (\mathbf{1}\{Y \leq \tau + \beta\} - \mathbf{1}\{Y \leq \tau\}) \right]$$

and note that we can express  $F_k(\tau) - F_\ell(\tau)$  similarly as

$$F_k(\tau) - F_\ell(\tau) = E \left[ \mathbf{1}\{Y \leq \tau\} \left( \frac{\mathbf{1}\{Z = k\}}{q_k} - \frac{\mathbf{1}\{Z = \ell\}}{q_\ell} \right) \right]$$

Using this notation, we can write the preceding as

$$(1 - \alpha_0 - \alpha_1) [F_k(\tau) - F_\ell(\tau)] = \alpha_0 [(1 - p_k) \Delta_{0k}^\tau(\beta) - (1 - p_\ell) \Delta_{0\ell}^\tau(\beta)] - (1 - \alpha_0) [p_k \Delta_{1k}^\tau(\beta) - p_\ell \Delta_{1\ell}^\tau(\beta)]$$

or in moment-condition form

$$E \left[ (1 - \alpha_0 - \alpha_1) \mathbf{1}\{Y \leq \tau\} \left( \frac{\mathbf{1}\{Z = k\}}{q_k} - \frac{\mathbf{1}\{Z = \ell\}}{q_\ell} \right) - (\mathbf{1}\{Y \leq \tau + \beta\} - \mathbf{1}\{Y \leq \tau\}) \left\{ \right. \\ \alpha_0 \left( (1 - p_k) \frac{\mathbf{1}\{T = 0, Z = k\}}{p_{0k}} - (1 - p_\ell) \frac{\mathbf{1}\{T = 0, Z = \ell\}}{p_{0\ell}} \right) \\ \left. \left. - (1 - \alpha_0) \left( p_k \frac{\mathbf{1}\{T = 1, Z = k\}}{p_{1k}} - p_\ell \frac{\mathbf{1}\{T = 1, Z = \ell\}}{p_{1\ell}} \right) \right\} \right] = 0$$

Each value of  $\tau$  yields a moment condition.

## 6 Special Case: $\alpha_0 = 0$

In this case the expressions from above simplify to

$$(1 - \alpha_1) [F_k(\tau) - F_\ell(\tau)] = [p_\ell F_{1\ell}(\tau + \beta) - p_k F_{1k}(\tau + \beta) - p_\ell F_{1\ell}(\tau) + p_k F_{1k}(\tau)] \quad (16)$$

for all  $\tau$ . Now, provided that all of the CDFs are differentiable we have<sup>1</sup>

$$e^{i\omega\tau}(1 - \alpha_1) [f_k(\tau) - f_\ell(\tau)] = e^{i\omega\tau} [p_\ell f_{1\ell}(\tau + \beta) - p_k f_{1k}(\tau + \beta) - p_\ell f_{1\ell}(\tau) + p_k f_{1k}(\tau)]$$

where we have pre-multiplied both sides by  $e^{i\omega\tau}$ . Finally, integrating both sides with respect to  $\tau$  over  $(-\infty, \infty)$ , we have

$$(1 - \alpha_1) [\varphi_k(\omega) - \varphi_\ell(\omega)] = \left\{ \int_{-\infty}^{\infty} e^{i\omega\tau} [p_\ell f_{1\ell}(\tau + \beta) - p_k f_{1k}(\tau + \beta)] d\tau - p_\ell \varphi_{1\ell}(\omega) + p_k \varphi_{1k}(\omega) \right\}$$

where  $\varphi_k$  is the conditional characteristic function of  $Y$  given  $Z = k$  and  $\varphi_{1k}$  is the conditional characteristic function of  $Y$  given  $T = 1, Z = k$ . Finally,

$$\begin{aligned} \int_{-\infty}^{\infty} e^{i\omega\tau} p_\ell f_{1\ell}(\tau + \beta) d\tau &= e^{i\omega\beta} p_\ell \int_{u=-\infty+\beta}^{u=\infty+\beta} e^{i\omega u} f_{1\ell}(u) du \\ &= e^{-i\omega\beta} p_\ell \varphi_{1\ell}(\omega) \end{aligned}$$

using the substitution  $u = \tau + \beta$ . Changing subscripts, the same holds for  $k$  and thus

$$(1 - \alpha_1) [\varphi_k(\omega) - \varphi_\ell(\omega)] = e^{-i\omega\beta} [p_\ell \varphi_{1\ell}(\omega) - p_k \varphi_{1k}(\omega)] + [p_k \varphi_{1k}(\omega) - p_\ell \varphi_{1\ell}(\omega)]$$

which, after collecting terms, simplifies to

$$(1 - \alpha_1) [\varphi_k(\omega) - \varphi_\ell(\omega)] = (e^{-i\omega\beta} - 1) [p_\ell \varphi_{1\ell}(\omega) - p_k \varphi_{1k}(\omega)] \quad (17)$$

for all  $\omega$ . Equation 17 contains exactly the same information as Equation 16 but gives us a more convenient way to prove identification since  $\beta$  enters in a simpler way. Leibniz's formula for the  $r$ th derivative of a product of two functions  $f$  and  $g$  is:

$$(fg)^{(r)} = \sum_{s=0}^r \binom{r}{s} f^{(s)} g^{(r-s)}$$

---

<sup>1</sup>There must be a way to generalize this using Lebesgue.

where  $f^{(r)}$  denotes the  $r$ th derivative of the function  $f$  and  $g^{(r-s)}$  denotes the  $(r-s)$ th derivative of the function  $g$ . Applying this to the RHS,  $R(\omega)$  of Equation 17 gives

$$\begin{aligned} \frac{d}{d\omega^r} R(\omega) &= \sum_{s=0}^r \binom{r}{s} \frac{d}{d\omega^s} (e^{-i\omega\beta} - 1) \frac{d}{d\omega^{r-s}} [p_\ell \varphi_{1\ell}(\omega) - p_k \varphi_{1k}(\omega)] \\ &= (e^{-i\omega\beta} - 1) [p_\ell \varphi_{1\ell}^{(r)}(\omega) - p_k \varphi_{1k}^{(r)}(\omega)] + e^{-i\omega\beta} \sum_{s=1}^r \binom{r}{s} (-i\beta)^s [p_\ell \varphi_{1\ell}^{(r-s)}(\omega) - p_k \varphi_{1k}^{(r-s)}(\omega)] \end{aligned}$$

where we split off the  $s = 0$  term because our generic expression for the  $s$ th derivative of  $(e^{-i\omega\beta} - 1)$  only applies for  $s \geq 1$ . Evaluating at zero:

$$\frac{d}{d\omega^r} R(0) = \sum_{s=1}^r \binom{r}{s} (-i\beta)^s [p_\ell \varphi_{1\ell}^{(r-s)}(0) - p_k \varphi_{1k}^{(r-s)}(0)]$$

Combining this with the LHS of Equation 17, also differentiated  $r$  times and evaluated at zero, we have

$$(1 - \alpha_1) [\varphi_k^{(r)}(0) - \varphi_\ell^{(r)}(0)] = \sum_{s=1}^r \binom{r}{s} (-i\beta)^s [p_\ell \varphi_{1\ell}^{(r-s)}(0) - p_k \varphi_{1k}^{(r-s)}(0)]$$

Now, recall that if  $\varphi(\omega)$  is the characteristic function of  $Y$  then  $\varphi^{(r)}(0) = i^r E[Y^r]$  provided that the expectation exists where  $\varphi^{(r)}$  denotes the  $r$ th derivative of  $\varphi$ . The same applies for the conditional characteristic functions we consider here. Hence, provided that the  $r$ th moments exist,

$$i^r (1 - \alpha_1) \{E[Y^r|Z = k] - E[Y^r|Z = \ell]\} = \sum_{s=1}^r \binom{r}{s} (-i\beta)^s i^{r-s} (p_\ell E[Y^{r-s}|T = 1, Z = \ell] - p_k E[Y^{r-s}|T = 1, Z = k])$$

After simplifying the terms involving  $i$  and cancelling them from both sides,

$$(1 - \alpha_1) (E[Y^r|Z = k] - E[Y^r|Z = \ell]) = \sum_{s=1}^r \binom{r}{s} (-\beta)^s (p_\ell E[Y^{r-s}|T = 1, Z = \ell] - p_k E[Y^{r-s}|T = 1, Z = k])$$

again provided that the moments exist. Abbreviating the conditional expectations according to  $E[Y^r|Z = k] = E_k[Y^r]$  and  $E[Y^r|T = t, Z = k] = E_{tk}[Y^r]$ , this becomes

$$(1 - \alpha_1) (E_k[Y^r] - E_\ell[Y^r]) = \sum_{s=1}^r \binom{r}{s} (-\beta)^s (p_\ell E_{1\ell}[Y^{r-s}] - p_k E_{1k}[Y^{r-s}]) \quad (18)$$

Equation 18 can be used to generate moment equations that are implied by the Equation 17 and the equivalent representation in terms of CDFs: Equation 16. Assuming that the

conditional first moments exist, we can evaluate Equation 18 at  $r = 1$ , yielding

$$\begin{aligned} (1 - \alpha_1) (E_k[Y] - E_\ell[Y]) &= \sum_{s=1}^1 \binom{1}{s} (-\beta)^s (p_\ell E_{1\ell} [Y^{1-s}] - p_k E_{1k} [Y^{1-s}]) \\ &= -\beta (p_\ell - p_k) \end{aligned}$$

Rearranging, this gives us the expression for the probability limit of the Wald estimator

$$\mathcal{W} \equiv \frac{E_k[Y] - E_\ell[Y]}{p_k - p_\ell} = \frac{\beta}{1 - \alpha_1} \quad (19)$$

Evaluating Equation 18 at  $r = 2$ , we have

$$\begin{aligned} (1 - \alpha_1) (E_k[Y^2] - E_\ell[Y^2]) &= \sum_{s=1}^2 \binom{2}{s} (-\beta)^s (p_\ell E_{1\ell} [Y^{2-s}] - p_k E_{1k} [Y^{2-s}]) \\ &= 2\beta (p_k E_{1k}[Y] - p_\ell E_{1\ell}[Y]) - \beta^2 (p_k - p_\ell) \end{aligned}$$

Rearranging, we have

$$E_k[Y^2] - E_\ell[Y^2] = \frac{\beta}{1 - \alpha_1} [2 (p_k E_{1k}[Y] - p_\ell E_{1\ell}[Y]) - \beta (p_k - p_\ell)] \quad (20)$$

Substituting Equation 19, we can replace  $\beta/(1 - \alpha_1)$  with a function of observables only, namely  $\mathcal{W}$ . Solving, we find that

$$\beta = \frac{2 (p_k E_{1k}[Y] - p_\ell E_{1\ell}[Y])}{p_k - p_\ell} - \frac{E_k[Y^2] - E_\ell[Y^2]}{E_k[Y] - E_\ell[Y]} \quad (21)$$

This allows us to state low-level sufficient conditions for identification:

- (a)  $\alpha_1 < 1$
- (b)  $p_k \neq p_\ell$
- (c)  $E_k[Y] \neq E_\ell[Y]$
- (d)  $E_{1k}[|Y|], E_{1\ell}[|Y|], E_k[|Y^2|], E_\ell[|Y^2|] < \infty$ .

Note that, although  $\beta = 0$  is always a solution of Equation 16 this solution is ruled out by the assumption that  $E_k[Y] \neq E_\ell[Y]$  via Equation 19. The mis-classification error rate  $\alpha_1$  is likewise uniquely identified under these assumptions. Substituting  $\beta/\mathcal{W} = 1 - \alpha_1$  into

Equation 21

$$\begin{aligned}(1 - \alpha_1) &= \left\{ \frac{p_k - p_\ell}{E_k[Y] - E_\ell[Y]} \right\} \left\{ \frac{2(p_k E_{1k}[Y] - p_\ell E_{1\ell}[Y])}{p_k - p_\ell} - \frac{E_k[Y^2] - E_\ell[Y^2]}{E_k[Y] - E_\ell[Y]} \right\} \\ &= \frac{2(p_k E_{1k}[Y] - p_\ell E_{1\ell}[Y])}{E_k[Y] - E_\ell[Y]} - (p_k - p_\ell) \left\{ \frac{E_k[Y^2] - E_\ell[Y^2]}{(E_k[Y] - E_\ell[Y])^2} \right\}\end{aligned}$$

and thus

$$\alpha_1 = 1 + (p_k - p_\ell) \left\{ \frac{E_k[Y^2] - E_\ell[Y^2]}{(E_k[Y] - E_\ell[Y])^2} \right\} - \frac{2(p_k E_{1k}[Y] - p_\ell E_{1\ell}[Y])}{E_k[Y] - E_\ell[Y]}$$

## 7 Identification in the General Case

## 8 Characteristic Functions

Recall from above that in the general case an independent instrument combined with non-differential measurement error implies that

$$\begin{aligned}(1 - \alpha_0 - \alpha_1) [F_k(\tau) - F_\ell(\tau)] &= \alpha_0 \{ (1 - p_k) [F_{0k}(\tau + \beta) - F_{0k}(\tau)] - (1 - p_\ell) [F_{0\ell}(\tau + \beta) - F_{0\ell}(\tau)] \} \\ &\quad - (1 - \alpha_0) \{ p_k [F_{1k}(\tau + \beta) - F_{1k}(\tau)] - p_\ell [F_{1\ell}(\tau + \beta) - F_{1\ell}(\tau)] \}\end{aligned}$$

Using the same steps as in the preceding section, we can convert this expression into characteristic function form by differentiating each side, multiplying by  $e^{i\omega\tau}$  and then integrating with respect to  $\tau$ , yielding

$$\begin{aligned}(1 - \alpha_0 - \alpha_1) [\varphi_k(\omega) - \varphi_\ell(\omega)] &= \alpha_0 \{ (1 - p_k) (e^{-i\omega\beta} - 1) \varphi_{0k}(\omega) - (1 - p_\ell) (e^{-i\omega\beta} - 1) \varphi_{0\ell}(\omega) \} \\ &\quad - (1 - \alpha_0) \{ p_k (e^{-i\omega\beta} - 1) \varphi_{1k}(\omega) - p_\ell (e^{-i\omega\beta} - 1) \varphi_{1\ell}(\omega) \}\end{aligned}$$

which simplifies to

$$\varphi_k(\omega) - \varphi_\ell(\omega) = (e^{-i\omega\beta} - 1) \left( \frac{\alpha_0 [(1 - p_k) \varphi_{0k}(\omega) - (1 - p_\ell) \varphi_{0\ell}(\omega)] - (1 - \alpha_0) [p_k \varphi_{1k}(\omega) - p_\ell \varphi_{1\ell}(\omega)]}{1 - \alpha_0 - \alpha_1} \right).$$

As above, we will differentiate both sides of this expression  $r$  times and evaluate at  $\omega = 0$ . Steps nearly identical to those given above yield

$$(1 - \alpha_0 - \alpha_1) (E_k[Y^r] - E_\ell[Y^r]) = \alpha_0 \sum_{s=1}^r \binom{r}{s} (-\beta)^s \{ (1 - p_k) E_{0k}[Y^{r-s}] - (1 - p_\ell) E_{0\ell}[Y^{r-s}] \} \\ - (1 - \alpha_0) \sum_{s=1}^r \binom{r}{s} (-\beta)^s \{ p_k E_{1k}[Y^{r-s}] - p_\ell E_{1\ell}[Y^{r-s}] \}$$

**First Moments** Taking  $r = 1$  gives

$$(1 - \alpha_0 - \alpha_1) (E_k[Y] - E_\ell[Y]) = \beta(p_k - p_\ell)$$

Simplifying,

$$\mathcal{W} \equiv \frac{E_k[Y] - E_\ell[Y]}{p_k - p_\ell} = \frac{\beta}{1 - \alpha_0 - \alpha_1} \quad (22)$$

**Second Moments** Now, taking  $r = 2$  gives

$$(1 - \alpha_0 - \alpha_1) (E_k[Y^2] - E_\ell[Y^2]) = \alpha_0 \{ [(1 - p_k) E_{0k}[Y] - (1 - p_\ell) E_{0\ell}] - \beta^2 (p_k - p_\ell) \} \\ - (1 - \alpha_0) \{ -2\beta (p_k E_{1k}[Y] - p_\ell E_{1\ell}[Y]) + \beta^2 (p_k - p_\ell) \} \\ = -2\beta\alpha_0 \{ (1 - p_k) E_{0k}[Y] - (1 - p_\ell) E_{0\ell}[Y] p_k E_{1k}[Y] + p_\ell E_{1\ell}[Y] \} \\ + 2\beta (p_k E_{1k}[Y] - p_\ell E_{1\ell}[Y]) - (p_k - p_\ell) \beta^2 (\alpha_0 + 1 - \alpha_0) \\ = -2\beta \{ \alpha_0 (E_k[Y] - E_\ell[Y]) - (p_k E_{1k}[Y] - p_\ell E_{1\ell}[Y]) \} - \beta^2 (p_k - p_\ell)$$

Now, simplifying

$$(1 - \alpha_0 - \alpha_1) \left( \frac{E_k[Y^2] - E_\ell[Y^2]}{p_k - p_\ell} \right) = -2\beta\alpha_0 \left( \frac{E_k[Y] - E_\ell[Y]}{p_k - p_\ell} \right) + 2\beta \left( \frac{p_{1k} E_{1k}[Y] - p_{\ell} E_{1\ell}[Y]}{p_k - p_\ell} \right) - \beta^2$$

and substituting Equation 22 to eliminate  $\beta$ , this becomes

$$(1 - \alpha_0 - \alpha_1) \left( \frac{E_k[Y^2] - E_\ell[Y^2]}{p_k - p_\ell} \right) = -2\alpha_0(1 - \alpha_0 - \alpha_1) \mathcal{W}^2 + 2\mathcal{W}(1 - \alpha_0 - \alpha_1) \left( \frac{p_{1k} E_{1k}[Y] - p_{\ell} E_{1\ell}[Y]}{p_k - p_\ell} \right) \\ - (1 - \alpha_0 - \alpha_1)^2 \mathcal{W}^2 \\ \left( \frac{E_k[Y^2] - E_\ell[Y^2]}{p_k - p_\ell} \right) = -2\alpha_0 \mathcal{W}^2 + 2\mathcal{W} \left( \frac{p_{1k} E_{1k}[Y] - p_{\ell} E_{1\ell}[Y]}{p_k - p_\ell} \right) - (1 - \alpha_0 - \alpha_1) \mathcal{W}^2$$

And thus, simplifying

$$\begin{aligned} -2\alpha_0\mathcal{W}^2 - (1 - \alpha_0 - \alpha_1)\mathcal{W}^2 &= \left( \frac{E_k[Y^2] - E_\ell[Y^2]}{p_k - p_\ell} \right) - 2\mathcal{W} \left( \frac{p_{1k}E_{1k}[Y] - p_\ell E_{1\ell}[Y]}{p_k - p_\ell} \right) \\ \alpha_1 - \alpha_0 &= 1 + \left[ \frac{E_k[Y^2] - E_\ell[Y^2]}{\mathcal{W}^2(p_k - p_\ell)} \right] - 2 \left[ \frac{p_{1k}E_{1k}[Y] - p_\ell E_{1\ell}[Y]}{\mathcal{W}(p_k - p_\ell)} \right] \end{aligned}$$

and therefore

$$\alpha_1 - \alpha_0 = 1 + (p_k - p_\ell) \left[ \frac{E_k[Y^2] - E_\ell[Y^2]}{(E_k[Y] - E_\ell[Y])^2} \right] - 2 \left[ \frac{p_{1k}E_{1k}[Y] - p_\ell E_{1\ell}[Y]}{E_k[Y] - E_\ell[Y]} \right] \quad (23)$$

**“Product” Moments** Recall that in our initial draft of the paper we worked with moments such as  $E[TY|Z = k]$ ,  $E[TY|Z = \ell]$  and  $E[TY^2|Z = k]$ ,  $E[TY^2|Z = \ell]$ . In the notation of this document, we can express these quantities as follows:

$$\begin{aligned} E[TY^r|z = k] &= E[TY^r|T = 1, z = k]p_k + E[TY^r|T = 0, z = k](1 - p_k) \\ &= p_k E[Y^r|T = 1, z = k] + 0 \\ &= p_k E_{1k}[Y^r] \end{aligned}$$

for any  $r$ . We will use this relationship to motivate some shorthand notation below.

**Some Shorthand** The notation above is becoming very cumbersome and we haven’t even looked at the third moments yet! To make life easier, define the following:

$$\begin{aligned} \widetilde{y_{1k}^r} &= p_k E_{1k}[Y^r] \\ \widetilde{y_{0k}^r} &= (1 - p_k) E_{1k}[Y^r] \\ \Delta \overline{y^r} &= E_k[Y^r] - E_\ell[Y^r] \\ \Delta \overline{T y^r} &= p_k E_{1k}[Y^r] - p_\ell E_{1\ell}[Y^r] = \widetilde{y_{1k}^r} - \widetilde{y_{1\ell}^r} \\ \mathcal{W} &= (E_k[Y] - E_\ell[Y]) / (p_k - p_\ell) \end{aligned}$$

for all  $r$ . When no  $r$  superscript is given this means  $r = 1$ . Note, moreover, that when  $r = 0$  we have  $\widetilde{y_{1k}^0} = p_k$  and  $\widetilde{y_{0k}^0} = (1 - p_k)$ . Thus  $\Delta \overline{T y^0} = p_k - p_\ell$ . In contrast,  $\Delta y^0 = 0$ .

Among other things, this notation will make it easier for us to link the derivations here to our earlier derivations from the first draft of the paper that used slightly different notation and did not work explicitly with the independence of the instrument.

**Simplifying the Moment Equalities** Using the final two pieces of notation defined in the preceding section, we can re-rewrite the collection of moment equalities arising from the characteristic function equations as

$$(1 - \alpha_0 - \alpha_1)\Delta\bar{y}^r = \sum_{s=1}^r \binom{r}{s} (-\beta)^s \left[ \alpha_0 \left( \widetilde{y_{0k}^{r-s}} - \widetilde{y_{0\ell}^{r-s}} \right) - (1 - \alpha_0) \left( \widetilde{y_{1k}^{r-s}} - \widetilde{y_{1\ell}^{r-s}} \right) \right]$$

Now, simplifying the terms in the square brackets,

$$\begin{aligned} \alpha_0 \left( \widetilde{y_{0k}^{r-s}} - \widetilde{y_{0\ell}^{r-s}} \right) - (1 - \alpha_0) \left( \widetilde{y_{1k}^{r-s}} - \widetilde{y_{1\ell}^{r-s}} \right) &= \alpha_0 \left[ \left( \widetilde{y_{0k}^{r-s}} + \widetilde{y_{1k}^{r-s}} \right) - \left( \widetilde{y_{0\ell}^{r-s}} + \widetilde{y_{1\ell}^{r-s}} \right) \right] - \left( \widetilde{y_{1k}^{r-s}} - \widetilde{y_{1\ell}^{r-s}} \right) \\ &= \alpha_0 \left( E_k[Y^{r-s}] - E_\ell[Y^{r-s}] \right) - \Delta\bar{T}y^{r-s} \\ &= \alpha_0\Delta\bar{y}^{r-s} - \Delta\bar{T}y^{r-s} \end{aligned}$$

and hence

$$(1 - \alpha_0 - \alpha_1)\Delta\bar{y}^r = \sum_{s=1}^r \binom{r}{s} (-\beta)^s \left( \alpha_0\Delta\bar{y}^{r-s} - \Delta\bar{T}y^{r-s} \right) \quad (24)$$

**Third Moments** Evaluating Equation 24 at  $r = 3$

$$\begin{aligned} (1 - \alpha_0 - \alpha_1)\Delta\bar{y}^3 &= \sum_{s=1}^3 \binom{3}{s} (-\beta)^s \left( \alpha_0\Delta\bar{y}^{3-s} - \Delta\bar{T}y^{3-s} \right) \\ &= -3\beta \left( \alpha_0\Delta\bar{y}^2 - \Delta\bar{T}y^2 \right) + 3\beta^2 \left( \alpha_0\Delta\bar{y} - \Delta\bar{T}y \right) + \beta^3(p_k - p_\ell) \end{aligned}$$

**Solving the System** Using  $\mathcal{W} = \beta/(1 - \alpha_0 - \alpha_1)$  we can re-write the third moment expression as follows

$$\begin{aligned} \Delta\bar{y}^3 &= -3\mathcal{W} \left( \alpha_0\Delta\bar{y}^2 - \Delta\bar{T}y^2 \right) + 3\beta\mathcal{W} \left( \alpha_0\Delta\bar{y} - \Delta\bar{T}y \right) + \beta^2\mathcal{W}(p_k - p_\ell) \\ \frac{\Delta\bar{y}^3}{\mathcal{W}(p_k - p_\ell)} &= \beta^2 + 3\beta \left( \frac{\alpha_0\Delta\bar{y} - \Delta\bar{T}y}{p_k - p_\ell} \right) - 3 \left( \frac{\alpha_0\Delta\bar{y}^2 - \Delta\bar{T}y^2}{p_k - p_\ell} \right) \\ \frac{\Delta\bar{y}^3 - 3\mathcal{W}\Delta\bar{y}^2\bar{T}}{\mathcal{W}(p_k - p_\ell)} &= \beta^2 + 3\beta \left( \frac{\alpha_0\Delta\bar{y} - \Delta\bar{T}y}{p_k - p_\ell} \right) - 3 \left( \frac{\alpha_0\Delta\bar{y}^2}{p_k - p_\ell} \right) \end{aligned}$$

Now, translating the second moment equation into the shorthand notation defined above, we have



**Simplifying the Characteristic Function Equation** From above, we have

$$\varphi_k(\omega) - \varphi_\ell(\omega) = (e^{-i\omega\beta} - 1) \left( \frac{\alpha_0 [(1 - p_k)\varphi_{0k}(\omega) - (1 - p_\ell)\varphi_{0\ell}(\omega)] - (1 - \alpha_0) [p_k\varphi_{1k}(\omega) - p_\ell\varphi_{1\ell}(\omega)]}{1 - \alpha_0 - \alpha_1} \right).$$

Using the fact that  $\varphi_k = p_k\varphi_{1k} + (1 - p_k)\varphi_{0k}$ , we can simplify this further, yielding

$$(1 - \alpha_0 - \alpha_1) = (e^{-i\omega\beta} - 1) [\alpha_0 - \xi(\omega)]$$

where we define

$$\xi(\omega) \equiv \frac{\varphi_k(\omega) - \varphi_\ell(\omega)}{p_k\varphi_{1k}(\omega) - p_\ell\varphi_{1\ell}(\omega)}$$

Now, re-arranging

$$(1 - \alpha_1) - \xi(\omega) = e^{-i\omega\beta} [\alpha_0 - \xi(\omega)]$$

or equivalently

$$e^{i\omega\beta} [(1 - \alpha_1) - \xi(\omega)] = \alpha_0 - \xi(\omega)$$

or

$$e^{i\omega\beta} = \frac{\alpha_0 - \xi(\omega)}{(1 - \alpha_1) - \xi(\omega)}$$

provided the denominator does not vanish. By taking differences or ratios evaluated at  $\omega_1$  and  $\omega_2$  we can eliminate  $\beta$ ,  $\alpha_0$  or  $\alpha_1$  but it's not clear how or if we can prove identification in terms of a restriction on the characteristic functions.

Suppose we consider three values  $\omega_1, \omega_2$  and  $\omega_3$  for which that yield to distinct, non-zero values  $\xi_1, \xi_2$  and  $\xi_3$  of  $\xi(\omega)$ .

$$e^{i\omega_1\beta} [(1 - \alpha_1) - \xi_1] - e^{i\omega_2\beta} [(1 - \alpha_1) - \xi_2] = \xi_2 - \xi_1$$

## 8.1 Simplifying the Characteristic CDF Equation

Recall from above that

$$(1 - \alpha_0 - \alpha_1) [F_k(\tau) - F_\ell(\tau)] = \alpha_0 \{ (1 - p_k) [F_{0k}(\tau + \beta) - F_{0k}(\tau)] - (1 - p_\ell) [F_{0\ell}(\tau + \beta) - F_{0\ell}(\tau)] \} \\ - (1 - \alpha_0) \{ p_k [F_{1k}(\tau + \beta) - F_{1k}(\tau)] - p_\ell [F_{1\ell}(\tau + \beta) - F_{1\ell}(\tau)] \}$$

We can simplify the RHS as follows

$$\text{RHS} = \alpha_0 \{ [F_k(\tau + \beta) - F_\ell(\tau + \beta)] - [F_k(\tau) - F_\ell(\tau)] \} \\ - \{ [p_k F_{1k}(\tau + \beta) - p_\ell F_{1\ell}(\tau + \beta)] - [p_k F_{1k}(\tau) - p_\ell F_{1\ell}(\tau)] \}$$

Now, define

$$\begin{aligned}\Delta(\tau) &= F_k(\tau) - F_\ell(\tau) \\ \tilde{\Delta}_1(\tau) &= p_k F_{1k}(\tau) - p_\ell F_{1\ell}(\tau)\end{aligned}$$

Using this notation, our equation becomes

$$(1 - \alpha_0 - \alpha_1)\Delta(\tau) = \alpha_0 [\Delta(\tau + \beta) - \Delta(\tau)] - [\tilde{\Delta}_1(\tau + \beta) - \tilde{\Delta}_1(\tau)]$$

which simplifies to

$$\tilde{\Delta}_1(\tau + \beta) - \tilde{\Delta}_1(\tau) = \alpha_0 \Delta(\tau + \beta) - (1 - \alpha_1)\Delta(\tau)$$

**Suppose  $\alpha_0 = 0$ :** In this case we obtain

$$(1 - \alpha_1) = \frac{\tilde{\Delta}_1(\tau) - \tilde{\Delta}_1(\tau + \beta)}{\Delta(\tau)}$$

Now, evaluating at two values of  $\tau$  and taking differences, we find

$$\frac{\tilde{\Delta}_1(\tau) - \tilde{\Delta}_1(\tau + \beta)}{\Delta(\tau)} - \frac{\tilde{\Delta}_1(\tau') - \tilde{\Delta}_1(\tau' + \beta)}{\Delta(\tau')} = 0$$

**Suppose  $\alpha_1 = 0$ :** In this case we obtain

$$\alpha_0 = \frac{\tilde{\Delta}_1(\tau + \beta) - \tilde{\Delta}_1(\tau) + \Delta(\tau)}{\Delta(\tau + \beta)}$$

Again, taking differences evaluated at two values of  $\tau$ ,

$$\frac{\tilde{\Delta}_1(\tau + \beta) - \tilde{\Delta}_1(\tau) + \Delta(\tau)}{\Delta(\tau + \beta)} - \frac{\tilde{\Delta}_1(\tau' + \beta) - \tilde{\Delta}_1(\tau') + \Delta(\tau')}{\Delta(\tau' + \beta)} = 0$$

**Some Equations to Check Numerically** We can use the same basic idea when either  $\alpha_0$  or  $\alpha_1$  is known but nonzero. This isn't realistic in practice, but can be used to check our

equations:

$$\alpha_0 = \frac{\tilde{\Delta}_1(\tau + \beta) - \tilde{\Delta}_1(\tau) + (1 - \alpha_1)\Delta(\tau)}{\Delta(\tau + \beta)}$$

$$(1 - \alpha_1) = \frac{\alpha_0\Delta(\tau + \beta) + \tilde{\Delta}_1(\tau) - \tilde{\Delta}_1(\tau + \beta)}{\Delta(\tau)}$$

As above, after substituting the true value of either  $\alpha_1$  or  $\alpha_0$ , we can eliminate the remaining mis-classification probability by evaluating at two quantiles  $\tau$ ,  $\tau'$  and taking differences.

These appear to work just fine!

**What if  $\alpha_0$  and  $\alpha_1$  are both unknown?** Suppose we take differences at two quantiles  $\tau$  and  $\nu$  to eliminate  $\alpha_1$ :

$$\begin{aligned} & \left[ \frac{\alpha_0\Delta(\tau + \beta) + \tilde{\Delta}_1(\tau) - \tilde{\Delta}_1(\tau + \beta)}{\Delta(\tau)} \right] - \left[ \frac{\alpha_0\Delta(\nu + \beta) + \tilde{\Delta}_1(\nu) - \tilde{\Delta}_1(\nu + \beta)}{\Delta(\nu)} \right] = 0 \\ \alpha_0 & \left[ \frac{\Delta(\tau + \beta)}{\Delta(\tau)} - \frac{\Delta(\nu + \beta)}{\Delta(\nu)} \right] - \left[ \frac{\tilde{\Delta}_1(\tau) - \tilde{\Delta}_1(\tau + \beta)}{\Delta(\tau)} - \frac{\tilde{\Delta}_1(\nu) - \tilde{\Delta}_1(\nu + \beta)}{\Delta(\nu)} \right] = 0 \end{aligned}$$

**The Equation that Didn't Work...**

$$\frac{[\tilde{\Delta}_1(\tau + \beta) - \tilde{\Delta}_1(\tau)] - [\tilde{\Delta}_1(\tau' + \beta) - \tilde{\Delta}_1(\tau')]}{\Delta(\tau + \beta) - \Delta(\tau' + \beta)} - \frac{[\tilde{\Delta}_1(\nu + \beta) - \tilde{\Delta}_1(\nu)] - [\tilde{\Delta}_1(\nu' + \beta) - \tilde{\Delta}_1(\nu')]}{\Delta(\nu + \beta) - \Delta(\nu' + \beta)} = 0$$

where  $\Delta(\nu) = \Delta(\nu')$  and  $\Delta(\tau) = \Delta(\tau')$ .

## 9 New Results from September 2016

### 9.0.1 Relationship between observed and unobserved CDFs

Let

$$F_{tk}^*(\tau) = P(Y \leq \tau | T^* = t, z_k)$$

$$F_{tk}(\tau) = P(Y \leq \tau | T = t, z_k)$$

Now, by the assumption of non-differential measurement error,

$$\begin{aligned} p_k F_{1k}(\tau) &= (1 - \alpha_1) p_k^* F_{1k}^*(\tau) + \alpha_0 (1 - p_k^*) F_{0k}^*(\tau) \\ (1 - p_k) F_{0k}(\tau) &= \alpha_1 p_k^* F_{1k}^*(\tau) + (1 - \alpha_0) (1 - p_k^*) F_{0k}^*(\tau) \end{aligned}$$

Solving the linear system as above, we find that

$$\begin{aligned} F_{0k}^*(\tau) &= F_{0k}(\tau) + \left( \frac{\alpha_1 p_k}{1 - p_k - \alpha_1} \right) [F_{0k}(\tau) - F_{1k}(\tau)] \\ F_{1k}^*(\tau) &= F_{1k}(\tau) + \left( \frac{\alpha_0 (1 - p_k)}{p_k - \alpha_0} \right) [F_{1k}(\tau) - F_{0k}(\tau)] \end{aligned}$$

## 9.1 Can we relax the measurement error assumptions?

Suppose that we continue to assume that  $P(Y|T^*, T, z) = P(Y|T^*, z)$  but relax the assumption that  $P(T|T^*, z) = P(T|T^*)$ . Define:

$$\begin{aligned} \alpha_{0k} &= P(T = 1|T^* = 1, z_k) \\ \alpha_{1k} &= P(T = 1|T^* = 0, z_k) \end{aligned}$$

As before, the Wald estimator converges in probability to

$$\mathcal{W} = \frac{E[Y|z_k] - E[Y|z_\ell]}{p_k - p_\ell}$$

but the relationship between  $p_1 - p_0$  and the unobserved  $p_1^* - p_0^*$  changes. By the law of total probability

$$\begin{aligned} p_k &= P(T = 1|z_k) = P(T = 1|T^* = 1, z_k)P(T^* = 1|z_k) + P(T = 1|T^* = 0, z_k)P(T^* = 0|z_k) \\ &= (1 - \alpha_{1k})p_k^* + \alpha_{0k}(1 - p_k^*) = p_k^*(1 - \alpha_{0k} - \alpha_{1k}) + \alpha_{0k} \end{aligned}$$

and thus

$$p_k^* = \frac{p_k - \alpha_{0k}}{1 - \alpha_{0k} - \alpha_{1k}}, \quad 1 - p_k^* = \frac{1 - p_k - \alpha_{1k}}{1 - \alpha_{0k} - \alpha_{1k}}.$$

Thus, we have

$$\begin{aligned} p_k^* - p_\ell^* &= \left( \frac{p_k - \alpha_{0k}}{1 - \alpha_{0k} - \alpha_{1k}} \right) - \left( \frac{p_0 - \alpha_{0\ell}}{1 - \alpha_{0\ell} - \alpha_{1\ell}} \right) \\ &= \frac{(p_k - \alpha_{0k})(1 - \alpha_{0\ell} - \alpha_{1\ell}) - (p_0 - \alpha_{0\ell})(1 - \alpha_{0k} - \alpha_{1k})}{(1 - \alpha_{0k} - \alpha_{1k})(1 - \alpha_{0\ell} - \alpha_{1\ell})} \end{aligned}$$

## 9.2 Is there a LATE interpretation of our results?

Let  $J \in \{a, c, d, n\}$  index an individual's *type*: always-taker, complier, defier, or never-taker. Let  $\pi_a, \pi_c, \pi_d, \pi_n$  denote the population proportions of always-takers, compliers, defiers, and never-takers. The unconfounded type assumption is  $P(J = j|z = 1) = P(J = j|z = 0)$ . Combined with the law of total probability, this gives

$$\begin{aligned} p_1^* &= P(T^* = 1|z = 1) = \pi_a + \pi_c \\ 1 - p_1^* &= P(T^* = 0|z = 1) = \pi_d + \pi_n \\ p_0^* &= P(T^* = 1|z = 0) = \pi_d + \pi_a \\ 1 - p_0^* &= P(T^* = 0|z = 0) = \pi_n + \pi_c \end{aligned}$$

Imposing no-defiers,  $\pi_d = 0$ , these expressions simplify to

$$\begin{aligned} p_1^* &= \pi_a + \pi_c \\ 1 - p_1^* &= \pi_n \\ p_0^* &= \pi_a \\ 1 - p_0^* &= \pi_n + \pi_c \end{aligned}$$

Solving for  $\pi_c$ , we see that

$$\begin{aligned} \pi_c &= p_1^* - p_0^* \\ \pi_a &= p_0^* \\ \pi_n &= 1 - p_1^* \end{aligned}$$

Now, let  $Y(1)$  indicate the potential outcome when  $T^* = 1$  and  $Y(0)$  indicate the potential outcome when  $T^* = 0$ . The standard LATE assumptions (no defiers, mean exclusion,

unconfounded type) imply

$$\begin{aligned}
\mathbb{E}(Y|T^* = 1, z = 1) &= \left(\frac{p_0^*}{p_1^*}\right) \mathbb{E}[Y(1)|J = a] + \left(\frac{p_1^* - p_0^*}{p_1^*}\right) \mathbb{E}[Y(1)|J = c] \\
\mathbb{E}(Y|T^* = 0, z = 0) &= \left(\frac{p_1^* - p_0^*}{1 - p_0^*}\right) \mathbb{E}[Y(0)|J = c] + \left(\frac{1 - p_1^*}{1 - p_0^*}\right) \mathbb{E}[Y(0)|J = n] \\
\mathbb{E}(Y|T^* = 1, z = 0) &= \mathbb{E}[Y(1)|J = a] \\
\mathbb{E}(Y|T^* = 0, z = 1) &= \mathbb{E}[Y(0)|J = n]
\end{aligned}$$

### 9.2.1 LATE Version of Theorem 2 from the Draft

$$\begin{aligned}
\Delta \overline{yT} &= \mathbb{E}(yT|z = 1) - \mathbb{E}(yT|z = 0) \\
&= (1 - \alpha_1) [p_1^* \mathbb{E}(y|T^* = 1, z = 1) - p_0^* \mathbb{E}(y|T^* = 1, z = 0)] \\
&\quad + \alpha_0 [(1 - p_1^*) \mathbb{E}(y|T^* = 0, z = 1) - (1 - p_0^*) \mathbb{E}(y|T^*, z = 0)]
\end{aligned}$$

So we find that

$$\begin{aligned}
\Delta \overline{yT} &= (p_1^* - p_0^*) \{ (1 - \alpha_1) \mathbb{E}[Y(1)|J = c] - \alpha_0 \mathbb{E}[Y(0)|J = c] \} \\
&= (1 - \alpha_1) \left\{ \frac{\mathbb{E}[Y(1) - Y(0)|J = c]}{1 - \alpha_0 - \alpha_1} (p_1 - p_0) \right\} + (p_1 - p_0) \mathbb{E}[Y(0)|J = c]
\end{aligned}$$

Recall that the analogous expression in the homogeneous treatment effect case is

$$\begin{aligned}
\Delta \overline{yT} &= (1 - \alpha_1) \mathcal{W}(p_1 - p_0) + \mu_{10}^* \\
&= (1 - \alpha_1) \left( \frac{\beta}{1 - \alpha_0 - \alpha_1} \right) (p_1 - p_0) + (p_1 - \alpha_0) m_{11}^* - (p_0 - \alpha_0) m_{10}^*
\end{aligned}$$

while the expression for the difference of variances is

$$\Delta \overline{y^2} = \beta \mathcal{W}(p_1 - p_0) + 2\mathcal{W}\mu_{10}^*$$

From above we see that the analogue of  $\mu_{10}^*$  in the heterogeneous treatment effects setting is  $(p_1 - p_0) \mathbb{E}[Y(0)|J = c]$  and since the LATE is  $\mathbb{E}[Y(1) - Y(0)|J = c]$ , the analogue of  $\mathcal{W}$  is

$$\frac{\mathbb{E}[Y(1) - Y(0)|J = c]}{1 - \alpha_0 - \alpha_1}$$

so *if* we could establish that

$$\Delta \overline{y^2} = \left( \frac{p_1 - p_0}{1 - \alpha_0 - \alpha_1} \right) \mathbb{E} [Y(1) - Y(0)|J = c] \cdot \mathbb{E} [Y(1) + Y(0)|J = c]$$

in the heterogeneous treatment effects case, the proof of Theorem 2 would go through immediately. Now, if we assume an exclusion restriction on the *second* moment of  $y$  an argument almost identical to the standard LATE derivation gives

$$\Delta \overline{y^2} = \frac{\mathbb{E} [Y^2(1) - Y^2(0)|J = c]}{p_1^* - p_0^*} = \left( \frac{p_1 - p_0}{1 - \alpha_0 - \alpha_1} \right) \mathbb{E} [Y^2(1) - Y^2(0)|J = c]$$

so we see that the necessary and sufficient condition for our proof to go through is

$$\mathbb{E} [Y^2(1) - Y^2(0)|J = c] = \mathbb{E} [Y(1) - Y(0)|J = c] \cdot \mathbb{E} [Y(1) + Y(0)|J = c]$$

Rearranging, this in turn is equivalent to

$$\text{Var} [Y(1)|J = c] = \text{Var} [Y(0)|J = c]$$

### 9.3 Partial Identification Under Independence Assumption

Suppose we only make the LATE independence assumption  $Y(T^*, z) = Y(T^*)$  rather than the conditional independence assumption  $P(Y < \tau|T^*, z_k) = P(Y < \tau|T^*, z_\ell)$ . Then we still obtain

$$\begin{aligned} \mathbb{P}(Y|T^* = 1, z = 1) &= \left( \frac{p_0^*}{p_1^*} \right) \mathbb{P}[Y(1)|J = a] + \left( \frac{p_1^* - p_0^*}{p_1^*} \right) \mathbb{P}[Y(1)|J = c] \\ \mathbb{P}(Y|T^* = 0, z = 0) &= \left( \frac{p_1^* - p_0^*}{1 - p_0^*} \right) \mathbb{P}[Y(0)|J = c] + \left( \frac{1 - p_1^*}{1 - p_0^*} \right) \mathbb{P}[Y(0)|J = n] \\ \mathbb{P}(Y|T^* = 1, z = 0) &= \mathbb{P}[Y(1)|J = a] \\ \mathbb{P}(Y|T^* = 0, z = 1) &= \mathbb{P}[Y(0)|J = n] \end{aligned}$$

From above, we also know that

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

The notation is getting a bit unwieldy so let  $\pi_{tk}^*(y) = P(Y = y|T^* = t, z_k)$  and similarly define  $\pi_{tk}(y) = P(Y = y|T = t, z_k)$ . Using this new notation, we have

$$\begin{aligned}(1 - p_k - \alpha_1)\pi_{0k}^*(y) &= (1 - p_k - \alpha_1)\pi_{0k}(y) + \alpha_1 p_k [\pi_{0k}(y) - \pi_{1k}(y)] \\ (p_k - \alpha_0)\pi_{1k}^*(y) &= (p_k - \alpha_0)\pi_{1k}(y) + \alpha_0(1 - p_k) [\pi_{1k}(y) - \pi_{0k}(y)]\end{aligned}$$

Writing these out for all values of  $k$ ,

$$\begin{aligned}(p_1 - \alpha_0)\pi_{11}^*(y) &= (p_1 - \alpha_0)\pi_{11}(y) + \alpha_0(1 - p_1) [\pi_{11}(y) - \pi_{01}(y)] \\ (1 - p_0 - \alpha_1)\pi_{00}^*(y) &= (1 - p_0 - \alpha_1)\pi_{00}(y) + \alpha_1 p_0 [\pi_{00}(y) - \pi_{10}(y)] \\ (p_0 - \alpha_0)\pi_{10}^*(y) &= (p_0 - \alpha_0)\pi_{10}(y) + \alpha_0(1 - p_0) [\pi_{10}(y) - \pi_{00}(y)] \\ (1 - p_1 - \alpha_1)\pi_{01}^*(y) &= (1 - p_1 - \alpha_1)\pi_{01}(y) + \alpha_1 p_1 [\pi_{01}(y) - \pi_{11}(y)]\end{aligned}$$

Similarly, using the fact that  $p_k^* = (p_k - \alpha_0)/(1 - \alpha_0 - \alpha_1)$ ,

$$\begin{aligned}\pi_{11}^*(y) &= \left(\frac{p_0 - \alpha_0}{p_1 - \alpha_0}\right) P[Y(1)|J = a] + \left(\frac{p_1 - p_0}{p_1 - \alpha_0}\right) P[Y(1)|J = c] \\ \pi_{00}^*(y) &= \left(\frac{p_1 - p_0}{1 - p_0 - \alpha_1}\right) P[Y(0)|J = c] + \left(\frac{1 - p_1 - \alpha_1}{1 - p_0 - \alpha_1}\right) P[Y(0)|J = n] \\ \pi_{10}^*(y) &= P[Y(1)|J = a] \\ \pi_{01}^*(y) &= P[Y(0)|J = n]\end{aligned}$$

or equivalently,

$$\begin{aligned}(p_1 - \alpha_0)\pi_{11}^*(y) &= (p_0 - \alpha_0) P[Y(1)|J = a] + (p_1 - p_0) P[Y(1)|J = c] \\ (1 - p_0 - \alpha_1)\pi_{00}^*(y) &= (p_1 - p_0) P[Y(0)|J = c] + (1 - p_1 - \alpha_1) P[Y(0)|J = n] \\ (p_0 - \alpha_0)\pi_{10}^*(y) &= (p_0 - \alpha_0) P[Y(1)|J = a] \\ (1 - p_1 - \alpha_1)\pi_{01}^*(y) &= (1 - p_1 - \alpha_1) P[Y(0)|J = n]\end{aligned}$$

Equating,

$$\begin{aligned}(p_0 - \alpha_0) P[Y(1)|J = a] + (p_1 - p_0) P[Y(1)|J = c] &= (p_1 - \alpha_0)\pi_{11}(y) + \alpha_0(1 - p_1) [\pi_{11}(y) - \pi_{01}(y)] \\ (p_1 - p_0) P[Y(0)|J = c] + (1 - p_1 - \alpha_1) P[Y(0)|J = n] &= (1 - p_0 - \alpha_1)\pi_{00}(y) + \alpha_1 p_0 [\pi_{00}(y) - \pi_{10}(y)] \\ (p_0 - \alpha_0) P[Y(1)|J = a] &= (p_0 - \alpha_0)\pi_{10}(y) + \alpha_0(1 - p_0) [\pi_{10}(y) - \pi_{00}(y)] \\ (1 - p_1 - \alpha_1) P[Y(0)|J = n] &= (1 - p_1 - \alpha_1)\pi_{01}(y) + \alpha_1 p_1 [\pi_{01}(y) - \pi_{11}(y)]\end{aligned}$$



and substituting the third and fourth equalities into the first and second we obtain

$$\begin{aligned}
(p_0 - \alpha_0)\pi_{10}(y) + \alpha_0(1 - p_0)[\pi_{10}(y) - \pi_{00}(y)] + (p_1 - p_0)P[Y(1)|J = c] &= (p_1 - \alpha_0)\pi_{11}(y) + \alpha_0(1 - p_1)[\pi_{11}(y) - \pi_{01}(y)] \\
(p_1 - p_0)P[Y(0)|J = c] + (1 - p_1 - \alpha_1)\pi_{01}(y) + \alpha_1 p_1[\pi_{01}(y) - \pi_{11}(y)] &= (1 - p_0 - \alpha_1)\pi_{00}(y) + \alpha_1 p_0[\pi_{00}(y) - \pi_{10}(y)] \\
(p_0 - \alpha_0)P[Y(1)|J = a] &= (p_0 - \alpha_0)\pi_{10}(y) + \alpha_0(1 - p_0)[\pi_{10}(y) - \pi_{00}(y)] \\
(1 - p_1 - \alpha_1)P[Y(0)|J = n] &= (1 - p_1 - \alpha_1)\pi_{01}(y) + \alpha_1 p_1[\pi_{01}(y) - \pi_{11}(y)]
\end{aligned}$$

Simplifying and re-arranging,

$$\begin{aligned}
P[Y(1) = y|J = c] &= \left[ \frac{p_1\pi_{11}(y) - p_0\pi_{10}(y)}{p_1 - p_0} \right] - \alpha_0 \left[ \frac{p_1\pi_{11}(y) - p_0\pi_{10}(y) + (1 - p_1)\pi_{01}(y) - (1 - p_0)\pi_{00}(y)}{p_1 - p_0} \right] \\
P[Y(0) = y|J = c] &= \left[ \frac{(1 - p_0)\pi_{00}(y) - (1 - p_1)\pi_{01}(y)}{p_1 - p_0} \right] - \alpha_1 \left[ \frac{(1 - p_0)\pi_{00}(y) - (1 - p_1)\pi_{01}(y) + p_0\pi_{10}(y) - p_1\pi_{11}(y)}{p_1 - p_0} \right] \\
P[Y(1) = y|J = a] &= \pi_{10}(y) + \left[ \frac{\alpha_0(1 - p_0)}{p_0 - \alpha_0} \right] [\pi_{10}(y) - \pi_{00}(y)] \\
P[Y(0) = y|J = n] &= \pi_{01}(y) + \left[ \frac{\alpha_1 p_1}{1 - p_1 - \alpha_1} \right] [\pi_{01}(y) - \pi_{11}(y)]
\end{aligned}$$

Notice that the first two equations can be simplified as follows

$$\begin{aligned}
P[Y(1) = y|J = c] &= \left[ \frac{P(Y = y, T = 1|z = 1) - P(Y, T = 1|z = 0)}{p_1 - p_0} \right] - \alpha_0 \left[ \frac{P(Y = y|z = 1) - P(Y = y|z = 0)}{p_1 - p_0} \right] \\
P[Y(0) = y|J = c] &= \left[ \frac{P(Y = y, T = 0|z = 1) - P(Y = y, T = 0|z = 0)}{p_1 - p_0} \right] - \alpha_1 \left[ \frac{P(Y = y|z = 0) - P(Y = y|z = 1)}{p_1 - p_0} \right]
\end{aligned}$$

Now, since probabilities must be between zero and one, we obtain the bounds

$$\begin{aligned}
0 &\leq \left[ \frac{P(Y = y, T = 1|z = 1) - P(Y = y, T = 1|z = 0)}{p_1 - p_0} \right] - \alpha_0 \left[ \frac{P(Y = y|z = 1) - P(Y = y|z = 0)}{p_1 - p_0} \right] \leq 1 \\
0 &\leq \left[ \frac{P(Y = y, T = 0|z = 1) - P(Y = y, T = 0|z = 0)}{p_1 - p_0} \right] - \alpha_1 \left[ \frac{P(Y = y|z = 0) - P(Y = y|z = 1)}{p_1 - p_0} \right] \leq 1
\end{aligned}$$

which we abbreviate

$$\begin{aligned}
0 &\leq \left[ \frac{\Delta P(Y = y, T = 1)}{p_1 - p_0} \right] - \alpha_0 \left[ \frac{\Delta P(Y = y)}{p_1 - p_0} \right] \leq 1 \\
0 &\leq \alpha_1 \left[ \frac{\Delta P(Y = y)}{p_1 - p_0} \right] - \left[ \frac{\Delta P(Y = y, T = 0)}{p_1 - p_0} \right] \leq 1
\end{aligned}$$

where

$$\begin{aligned}
\Delta P(Y = y) &= P(Y = y|z = 1) - P(Y = y|z = 0) \\
\Delta P(Y = y, T = t) &= P(Y = y, T = t|z = 1) - P(Y = y, T = t|z = 0).
\end{aligned}$$

To manipulate these bounds, we need to know the sign of  $R = \Delta P(Y = y)/(p_1 - p_0)$ . Presumably this will be positive for most values of  $y$ , but it could be negative.

**Case I:  $R$  is positive.**

$$\begin{aligned} \frac{\Delta P(Y = y, T = 1) - (p_1 - p_0)}{\Delta P(Y = y)} &\leq \alpha_0 \leq \frac{\Delta P(Y = y, T = 1)}{\Delta P(Y = y)} \\ \frac{\Delta P(Y = y, T = 0)}{\Delta P(Y = y)} &\leq \alpha_1 \leq \frac{\Delta P(Y = y, T = 0) + (p_1 - p_0)}{\Delta P(Y = y)} \end{aligned}$$

**Case II:  $R$  is negative.**

$$\begin{aligned} \frac{\Delta P(Y = y, T = 1)}{\Delta P(Y = y)} &\leq \alpha_0 \leq \frac{\Delta P(Y = y, T = 1) - (p_1 - p_0)}{\Delta P(Y = y)} \\ \frac{\Delta P(Y = y, T = 0) + (p_1 - p_0)}{\Delta P(Y = y)} &\leq \alpha_1 \leq \frac{\Delta P(Y = y, T = 0)}{\Delta P(Y = y)} \end{aligned}$$

Note that we *two-sided* bounds for the misclassification probabilities. These may be trivial in some cases, but I don't think it's obvious that they always will be.

Do these bounds have anything to do with the testability of the LATE assumptions? That is, do we get a lower bound for measurement error *precisely when* we would otherwise violate a testable LATE assumption?

Note that we also obtain bounds from the potential outcome distributions of always-takers and never-takers, namely

$$\begin{aligned} 0 &\leq \pi_{10}(y) + \left[ \frac{\alpha_0(1 - p_0)}{p_0 - \alpha_0} \right] [\pi_{10}(y) - \pi_{00}(y)] \leq 1 \\ 0 &\leq \pi_{01}(y) + \left[ \frac{\alpha_1 p_1}{1 - p_1 - \alpha_1} \right] [\pi_{01}(y) - \pi_{11}(y)] \leq 1 \end{aligned}$$

but these are redundant. From the assumption of non-differential measurement error, we already have

$$\begin{aligned} \pi_{0k}^* &= \pi_{0k} + \left( \frac{\alpha_1 p_k}{1 - p_k - \alpha_1} \right) (\pi_{0k} - \pi_{1k}) \\ \pi_{1k}^* &= \pi_{1k} + \left( \frac{\alpha_0(1 - p_k)}{p_k - \alpha_0} \right) (\pi_{1k} - \pi_{0k}) \end{aligned}$$

for all  $k$  as given at the beginning of this section. These expressions imply

$$\begin{aligned} 0 &\leq \pi_{0k} + \left( \frac{\alpha_1 p_k}{1 - p_k - \alpha_1} \right) (\pi_{0k} - \pi_{1k}) \leq 1 \\ 0 &\leq \pi_{1k} + \left( \frac{\alpha_0 (1 - p_k)}{p_k - \alpha_0} \right) (\pi_{1k} - \pi_{0k}) \leq 1 \end{aligned}$$

Re-arranging, we have

$$\begin{aligned} 0 &\leq (1 - p_k)\pi_{0k} - \alpha_1\pi_{0k} + \alpha_1 p_k (\pi_{0k} - \pi_{1k}) \leq 1 - p_k - \alpha_1 \\ 0 &\leq p_k\pi_{1k} - \alpha_0\pi_{1k} + \alpha_0(1 - p_k) (\pi_{1k} - \pi_{0k}) \leq p_k - \alpha_0 \end{aligned}$$

and thus

$$\begin{aligned} 0 &\leq (1 - p_k)\pi_{0k} - \alpha_1 [(1 - p_k)\pi_{0k} + p_k\pi_{1k}] \leq 1 - p_k - \alpha_1 \\ 0 &\leq p_k\pi_{1k} - \alpha_0 [p_k\pi_{1k} + (1 - p_k)\pi_{0k}] \leq p_k - \alpha_0 \end{aligned}$$

Now consider the first inequality. Re-arranging the right-hand side we obtain

$$\alpha_1 \leq \frac{(1 - p_k)(1 - \pi_{0k})}{1 - [(1 - p_k)\pi_{0k} + p_k\pi_{1k}]} = (1 - p_k) \left[ \frac{P(Y = 0|T = 0, z = k)}{P(Y = 0|z = k)} \right]$$

and re-arranging the left-hand side we find

$$\alpha_1 \leq \frac{(1 - p_k)\pi_{0k}}{(1 - p_k)\pi_{0k} + p_k\pi_{1k}} = (1 - p_k) \left[ \frac{P(Y = 1|T = 0, z = k)}{P(Y = 1|z = k)} \right]$$

For the second inequality, the left-hand side gives

$$\alpha_0 \leq \frac{p_k\pi_{1k}}{p_k\pi_{1k} + (1 - p_k)\pi_{0k}} = p_k \left[ \frac{P(Y = 1|T = 1, z = k)}{P(Y = 1|z_k)} \right]$$

while the right-hand side gives

$$\alpha_0 \leq \frac{p_k(1 - \pi_{1k})}{1 - [p_k\pi_{1k} + (1 - p_k)\pi_{0k}]} = p_k \left[ \frac{P(Y = 0|T = 1, z = k)}{P(Y = 0|z = k)} \right]$$

These are analogous to our CDF bounds from above although they may not be tighter than the bounds

$$\alpha_0 \leq p_k, \quad \alpha_1 \leq (1 - p_k)$$

because we cannot argue, as we did above, about a limit in which the ratio of CDFs ap-

proaches one. As before, however, we can take the tightest bound over  $k = 0, 1$ .

## 9.4 Bounding the LATE

Even if we didn't know anything about  $\alpha_0$  and  $\alpha_1$  beyond the fact that they are probabilities, it looks like we could still bound the LATE. I think we can do this without using the independence of the instrument, that is only using the mean exclusion restriction. Write out the LATE expressions with the  $\alpha_0$  and  $\alpha_1$  in them and then just plug in zero and one. Could then tighten the bounds by imposing additional assumptions to get bounds for  $\alpha_0$  and  $\alpha_1$ , from weakest to strongest. If you have an independent instrument, you also get bounds for the outcome distributions. Need to think some more about this...

## 9.5 Stochastic Dominance Conditions

What if we imposed a stochastic ordering, e.g.  $Y(1) > Y(0)$  for compliers? Presumably this would give joint bounds for  $\alpha_0$  and  $\alpha_1$  from the LATE expressions from above. Alternatively, perhaps one would choose to impose an ordering on the  $Y(0)$  distributions for compliers versus never-takers or the  $Y(1)$  distributions for the compliers versus always-takers. This might be interesting in situations where one is concerned that the assumption we need for identification does not in fact hold and should give additional bounds.

# 10 Outline For New Draft

### 1. Introduction / Literature Review

#### (a) Why is this an important question?

- Treatments of interest in economics usually endogenous and often binary.
- Randomized encouragement designs are common in applied work.
- Treatment status is often self-reported.
- This problem is much more challenging than people realize.

#### (b) Why are we different from Ura?

- Main difference is that we, in line with the existing literature, study the case of non-differential measurement error. This allows us to obtain point identification under certain assumptions.
- In contrast, Ura considers arbitrary forms of mis-classification but as a consequence presents only partial identification results.

- Second, while we do provide results for LATE in Section blah, we mainly focus on additively separable model in which heterogeneity is captured by observed covariates while Ura considers only a LATE model. (And also doesn't allow for covariates.)

## 2. Mahajan/Lewbell-style Assumptions

### (a) Setup and Assumptions:

- Homogenous treatment effect model (additively separable)
- Conditional mean version of non-differential measurement error assumption.
- Conditional mean independence for IV.

(b) Show that the model is not identified, regardless of (discrete) support of IV.

(c) Derive sharp bounds for  $\alpha_0, \alpha_1$  and treatment effect.

(d) Show that second and third-moment independence for IV identifies this model?  
Maybe this isn't interesting in and of itself?

## 3. Independence Assumption

### (a) Motivation

- Showed above that stronger assumptions are needed for identification, but the additional moment restrictions seem a bit artificial.
- When instruments are derived from economic theory that yields conditional mean independence only, we wouldn't want to use them.
- They would make sense, however, in an RCT or natural experiment.
- The whole point in these settings is *not* to rely on functional form assumptions. It would be strange to say that  $z$  is an instrument for  $y$  but not  $\log y$ .
- This points towards an *independence* assumption for the instrument.
- Can make a similar argument for measurement error: seems strange to assume that  $T$  is non-differential for  $y$  but not  $\log y$ .

(b) Sharp Bounds for  $\alpha_0$  and  $\alpha_1$  without valid instrument

- Assume "independence" version of non-differential measurement error.
- Derive CDF bounds.

(c) Conditional Independence for Instrument

- Exactly what assumptions do we need here?
- Characteristic functions.

- Identification conditions?
- Overidentifying restrictions? Test model?

#### 4. LATE Model

##### (a) Introduction

- Most of the existing mis-classification literature focuses on a homogeneous treatment effects model.
- What if we don't have an additively separable model?
- These results complement Ura because we work under the assumption of non-differential measurement error while he asks what can be learned when one is unwilling to make any assumptions about the form of the mis-classification.

##### (b) Mahajan/Lewbel Setup

- Presumably the partial identification results go through for a LATE.
- The second and third moment conditions would require restrictions on form of heterogeneity. These would seem to be satisfied by a generalized Roy model.

##### (c) Independence Assumptions

- Presumably the CDF bounds go through as before but need to state exact form of independence assumption in terms of potential outcomes.
- Kitagawa-style independence assumption for IV:  $Y(T^*, z) = Y(T^*)$ . This gives bounds for all quantile treatment effects.
- Stochastic dominance conditions?

#### 5. Estimation / Inference

#### 6. Simulation Study

#### 7. Empirical Examples

- Try to look at a number of examples under different assumptions to illustrate both point and partial identification results. Don't forget about Oreopoulos: the sample size is so huge that inference isn't a major concern.

## 11 Weak Identification

### 11.1 Moment Equations

First we write the moment equations in a more familiar GMM-style form.

**First Moment Condition:** This is simply the IV moment condition:

$$\text{Cov}(y, z)/\text{Cov}(T, z) = \beta/(1 - \alpha_0 - \alpha_1)$$

Rearranging gives a more “canonical” GMM form:

$$\text{Cov}(y, z) - \left( \frac{\beta}{1 - \alpha_0 - \alpha_1} \right) \text{Cov}(T, z) = 0$$

**Second Moment Condition:** The equations used to identify  $(\alpha_0 - \alpha_1)$  in the paper are

$$\begin{aligned}\mu_{k\ell}^* &= (p_k - \alpha_0)m_{1k}^* - (p_\ell - \alpha_0)m_{1\ell}^* \\ \Delta \overline{y^2} &= \beta \mathcal{W}(p_k - p_\ell) + 2\mathcal{W}\mu_{k\ell}^* \\ \Delta \overline{yT} &= (1 - \alpha_1)\mathcal{W}(p_k - p_\ell) + \mu_{k\ell}^*\end{aligned}$$

Re-arranging the third equation,  $\mu_{k\ell}^* = \Delta \overline{yT} - (1 - \alpha_1)\mathcal{W}(p_k - p_\ell)$ . Substituting into the second equation,

$$\begin{aligned}\Delta \overline{y^2} &= \mathcal{W} [\beta(p_k - p_\ell) + 2\mu_{k\ell}^*] \\ &= \mathcal{W} \{ \beta(p_k - p_\ell) + 2 [\Delta \overline{yT} - (1 - \alpha_1)\mathcal{W}(p_k - p_\ell)] \} \\ &= \mathcal{W} \{ (p_k - p_\ell) [\beta - 2(1 - \alpha_1)\mathcal{W}] + 2\Delta \overline{yT} \}\end{aligned}$$

Now, substituting  $\mathcal{W} = \beta/(1 - \alpha_0 - \alpha_1)$

$$\begin{aligned}\Delta \overline{y^2} &= \frac{\beta}{1 - \alpha_0 - \alpha_1} \left\{ \beta(p_k - p_\ell) \left[ 1 - \frac{2(1 - \alpha_1)}{1 - \alpha_0 - \alpha_1} \right] + 2\Delta \overline{yT} \right\} \\ &= \frac{\beta}{1 - \alpha_0 - \alpha_1} \left\{ 2\Delta \overline{yT} - \beta(p_k - p_\ell) \left( \frac{1 + \alpha_0 - \alpha_1}{1 - \alpha_0 - \alpha_1} \right) \right\}\end{aligned}$$

We now write this in a more standard form. Let  $w$  be any random variable. Then,

$$\begin{aligned}\text{Cov}(w, z) &= E(wz) - E(w)E(z) = [1 \times E(w|z = 1)q + 0 \times E(w|z = 0)(1 - q)] - E(w)q \\ &= qE(w|z = 1) - qE(w) = qE(w|z = 1) - q[E(w|z = 1)q + E(w|z = 0)(1 - q)] \\ &= q[E(w|z = 1)(1 - q) + E(w|z = 0)(1 - q)] \\ &= q(1 - q)[E(w|z = 1) - E(w|z = 0)]\end{aligned}$$

Using this fact, we can express the quantities that appear in the second moment equality in terms of covariances as follows

$$\Delta \overline{y^2} = \frac{\text{Cov}(y^2, z)}{q(1-q)}, \quad \Delta \overline{yT} = \frac{\text{Cov}(yT, z)}{q(1-q)}, \quad (p_k - p_\ell) = \frac{\text{Cov}(T, z)}{q(1-q)}$$

leading to

$$\frac{\text{Cov}(y^2, z)}{q(1-q)} = \frac{\beta}{1 - \alpha_0 - \alpha_1} \left\{ \frac{2\text{Cov}(yT, z)}{q(1-q)} - \beta \frac{\text{Cov}(T, z)}{q(1-q)} \left( \frac{1 + \alpha_0 - \alpha_1}{1 - \alpha_0 - \alpha_1} \right) \right\}$$

Or, multiplying through by  $q(1-q)$  and re-arranging,

$$\text{Cov}(y^2, z) - \frac{\beta}{1 - \alpha_0 - \alpha_1} \left\{ 2\text{Cov}(yT, z) - \beta \text{Cov}(T, z) \left( \frac{1 + \alpha_0 - \alpha_1}{1 - \alpha_0 - \alpha_1} \right) \right\} = 0$$

**Third Moment Condition:** The third and final set of moment conditions is

$$\begin{aligned} \Delta \overline{y^3} &= \beta^2 [\mathcal{W}(p_k - p_\ell)] + 3\beta [\mathcal{W}\mu_{k\ell}^*] + 3\mathcal{W}\lambda_{k\ell}^* \\ \Delta \overline{y^2T} &= \beta(1 - \alpha_1) \mathcal{W}(p_k - p_\ell) + 2(1 - \alpha_1) \mathcal{W}\mu_{k\ell}^* + \lambda_{k\ell}^* \end{aligned}$$

To put this into a more familiar format, we first eliminate  $\mu_{k\ell}^*$  using

$$\mu_{k\ell}^* = \Delta \overline{yT} - (1 - \alpha_1) \mathcal{W}(p_k - p_\ell)$$

from the derivation of the second moment equation from above, yielding

$$\begin{aligned} \Delta \overline{y^3} &= \beta^2 [\mathcal{W}(p_k - p_\ell)] + 3\beta \mathcal{W} [\Delta \overline{yT} - (1 - \alpha_1) \mathcal{W}(p_k - p_\ell)] + 3\mathcal{W}\lambda_{k\ell}^* \\ \Delta \overline{y^2T} &= \beta(1 - \alpha_1) \mathcal{W}(p_k - p_\ell) + 2(1 - \alpha_1) \mathcal{W} [\Delta \overline{yT} - (1 - \alpha_1) \mathcal{W}(p_k - p_\ell)] + \lambda_{k\ell}^* \end{aligned}$$

Re-arranging and factoring the first equation gives

$$\Delta \overline{y^3} = \mathcal{W}(p_k - p_\ell) \left\{ \beta^2 + \frac{3\beta \Delta \overline{yT}}{p_k - p_\ell} - 3\beta \mathcal{W}(1 - \alpha_1) + \frac{3\lambda_{k\ell}^*}{p_k - p_\ell} \right\}$$

Now, by re-arranging the second equation we find that

$$\begin{aligned} \lambda_{k\ell}^* &= \Delta \overline{y^2T} - \beta(1 - \alpha_1) \mathcal{W}(p_k - p_\ell) - 2(1 - \alpha_1) \mathcal{W} [\Delta \overline{yT} - (1 - \alpha_1) \mathcal{W}(p_k - p_\ell)] \\ &= \Delta \overline{y^2T} - \beta(1 - \alpha_1) \mathcal{W}(p_k - p_\ell) - 2(1 - \alpha_1) \mathcal{W} \Delta \overline{yT} + 2(1 - \alpha_1)^2 \mathcal{W}^2(p_k - p_\ell) \end{aligned}$$



and thus

$$\frac{3\lambda_{k\ell}^*}{p_k - p_\ell} = 3 \left( \frac{\Delta \overline{y^2 T}}{p_k - p_\ell} \right) - 3\beta(1 - \alpha_1)\mathcal{W} - 6(1 - \alpha_1)\mathcal{W} \left( \frac{\Delta \overline{y T}}{p_k - p_\ell} \right) + 6(1 - \alpha_1)^2 \mathcal{W}^2$$

so that

$$\begin{aligned} \frac{\Delta \overline{y^3}}{\mathcal{W}(p_k - p_\ell)} &= \left\{ \beta^2 - 6\beta\mathcal{W}(1 - \alpha_1) + 6\mathcal{W}^2(1 - \alpha_1)^2 + \left( \frac{\Delta \overline{y T}}{p_k - p_\ell} \right) [3\beta - 6\mathcal{W}(1 - \alpha_1)] + 3 \left( \frac{\Delta \overline{y^2 T}}{p_k - p_\ell} \right) \right\} \\ &= \left\{ \beta^2 \left[ 1 - \frac{6(1 - \alpha_1)}{1 - \alpha_0 - \alpha_1} + \frac{6(1 - \alpha_1)^2}{(1 - \alpha_0 - \alpha_1)^2} \right] + 3\beta \left[ 1 - \frac{2(1 - \alpha_1)}{1 - \alpha_0 - \alpha_1} \right] \left( \frac{\Delta \overline{y T}}{p_k - p_\ell} \right) + 3 \left( \frac{\Delta \overline{y^2 T}}{p_k - p_\ell} \right) \right\} \end{aligned}$$

Simplifying, we find that

$$\left[ 1 - \frac{2(1 - \alpha_1)}{1 - \alpha_0 - \alpha_1} \right] = \frac{(1 - \alpha_0 - \alpha_1) - 2(1 - \alpha_1)}{1 - \alpha_0 - \alpha_1} = \frac{1 - \alpha_0 - \alpha_1 - 2 + 2\alpha_1}{1 - \alpha_0 - \alpha_1} = \frac{\alpha_1 - \alpha_0 - 1}{1 - \alpha_0 - \alpha_1}$$

and

$$\begin{aligned} \left[ 1 - \frac{6(1 - \alpha_1)}{1 - \alpha_0 - \alpha_1} + \frac{6(1 - \alpha_1)^2}{(1 - \alpha_0 - \alpha_1)^2} \right] &= 1 - \left[ \frac{6(1 - \alpha_1)}{1 - \alpha_0 - \alpha_1} \right] \left[ 1 - \frac{1 - \alpha_1}{1 - \alpha_0 - \alpha_1} \right] \\ &= 1 - \left[ \frac{6(1 - \alpha_1)}{1 - \alpha_0 - \alpha_1} \right] \left[ \frac{(1 - \alpha_0 - \alpha_1) - (1 - \alpha_1)}{1 - \alpha_0 - \alpha_1} \right] \\ &= 1 + \frac{6\alpha_0(1 - \alpha_1)}{(1 - \alpha_0 - \alpha_1)^2} \end{aligned}$$

so that

$$\frac{\Delta \overline{y^3}}{\mathcal{W}(p_k - p_\ell)} = \left\{ \beta^2 \left[ 1 + \frac{6\alpha_0(1 - \alpha_1)}{(1 - \alpha_0 - \alpha_1)^2} \right] - 3\beta \left[ \frac{1 - (\alpha_1 - \alpha_0)}{1 - \alpha_0 - \alpha_1} \right] \left( \frac{\Delta \overline{y T}}{p_k - p_\ell} \right) + 3 \left( \frac{\Delta \overline{y^2 T}}{p_k - p_\ell} \right) \right\}$$

Therefore, re-arranging and multiplying through by  $q(1 - q)$ ,

$$\text{Cov}(z, y^3) = \left( \frac{\beta}{1 - \alpha_0 - \alpha_1} \right) \left\{ \beta^2 \left[ 1 + \frac{6\alpha_0(1 - \alpha_1)}{(1 - \alpha_0 - \alpha_1)^2} \right] \text{Cov}(z, T) - 3\beta \left[ \frac{1 - (\alpha_1 - \alpha_0)}{1 - \alpha_0 - \alpha_1} \right] \text{Cov}(z, yT) + 3\text{Cov}(z, y^2 T) \right\}$$

### Full Set of Moment Conditions

$$\begin{aligned} \text{Cov}(y, z) - \left( \frac{\beta}{1 - \alpha_0 - \alpha_1} \right) \text{Cov}(T, z) &= 0 \\ \text{Cov}(y^2, z) - \frac{\beta}{1 - \alpha_0 - \alpha_1} \left\{ 2\text{Cov}(yT, z) - \beta\text{Cov}(T, z) \left( \frac{1 + \alpha_0 - \alpha_1}{1 - \alpha_0 - \alpha_1} \right) \right\} &= 0 \\ \text{Cov}(y^3, z) - \left( \frac{\beta}{1 - \alpha_0 - \alpha_1} \right) \left\{ \beta^2 \left[ 1 + \frac{6\alpha_0(1 - \alpha_1)}{(1 - \alpha_0 - \alpha_1)^2} \right] \text{Cov}(T, z) - 3\beta \left[ \frac{1 - (\alpha_1 - \alpha_0)}{1 - \alpha_0 - \alpha_1} \right] \text{Cov}(yT, z) + 3\text{Cov}(y^2 T, z) \right\} &= 0 \end{aligned}$$

## 11.2 Simple Special Case: $\alpha_0 = 0$

Suppose that  $\alpha_0$ . Then the model is identified using the first and second moment equalities, which simplify to

$$\begin{aligned} \text{Cov}(y, z) - \left( \frac{\beta}{1 - \alpha_1} \right) \text{Cov}(T, z) &= 0 \\ \text{Cov}(y^2, z) - \left( \frac{\beta}{1 - \alpha_1} \right) [2\text{Cov}(yT, z) - \beta\text{Cov}(T, z)] &= 0 \end{aligned}$$

In this simple special case, it is easy to solve for  $\beta$  by substituting the first moment condition into the second:

$$\beta = \frac{2\text{Cov}(yT, z)}{\text{Cov}(T, z)} - \frac{\text{Cov}(y^2, z)}{\text{Cov}(y, z)}$$

I checked this equation in our simulation experiment and it is indeed correct

Notice that if  $\beta \approx 0$  then both  $\text{Cov}(y^2, z)$  and  $\text{Cov}(y, z)$  are close to zero so their ratio becomes extremely noisy.

**Standard GMM form:** To express this system in the standard GMM form, we need to agument these moment equalities with expressions for the means of  $z, y, y^2, T$ , and  $yT$  as follows. Let  $\mathbf{w}_i = (y_i, z_i, T_i)'$ ,  $\theta = (\beta, \alpha_1)'$  and  $\gamma = (q, p, \mu, s, r)'$  where

$$\begin{aligned} q &= \mathbb{E}[z] \\ p &= \mathbb{E}[T] \\ \mu &= \mathbb{E}[y] \\ s &= \mathbb{E}[y^2] \\ r &= \mathbb{E}[yT]. \end{aligned}$$

We can express our problem in terms of two blocks of moment conditions, namely

$$f(\mathbf{w}; \theta, \gamma) = \begin{bmatrix} g(\mathbf{w}; \theta, \gamma) \\ h(\mathbf{w}; \gamma) \end{bmatrix}$$

where

$$g(\mathbf{w}; \theta, \gamma) = \begin{bmatrix} (zy - q\mu) - \left( \frac{\beta}{1 - \alpha_1} \right) (zT - qp) \\ (zy^2 - qs) - 2 \left( \frac{\beta}{1 - \alpha_1} \right) (zyT - qr) + \left( \frac{\beta^2}{1 - \alpha_1} \right) (zT - qp) \end{bmatrix}$$

and

$$h(\mathbf{w}; \gamma) = \begin{bmatrix} z - q \\ T - p \\ y - \mu \\ y^2 - s \\ yT - r \end{bmatrix}$$

We can view this as a two-step or “plug-in” GMM estimation problem where  $\hat{\gamma}$  solves the sample moment condition

$$\frac{1}{n} \sum_{i=1}^n h(\mathbf{w}_i; \gamma) = 0$$

and  $\hat{\theta}$  solves

$$\frac{1}{n} \sum_{i=1}^n g(\mathbf{w}_i; \theta, \hat{\gamma}) = 0.$$

Unfortunately, in our example the first-step estimation affects the asymptotic variance of the second since an inconsistent estimator of  $\gamma$  yields an inconsistent estimator of  $\theta$ .<sup>2</sup> This means that we will have to proceed “the hard way.”

Under standard regularity conditions, a GMM estimator based on the sample analogue  $f_n(\theta, \gamma)$  of  $\mathbb{E}[f(\mathbf{w}; \theta, \gamma)] = 0$  using a weighting matrix  $\widehat{W} \rightarrow_p W$  converges in distribution to

$$-(F'WF)^{-1}F'WM, \quad M \sim N(0, \Omega)$$

where  $\sqrt{n}f_n(\theta_0, \gamma_0) \rightarrow_d M$  and  $F = \mathbb{E}[\nabla'_\theta f(\mathbf{w}; \theta_0, \gamma_0), \nabla'_\gamma f(\mathbf{w}; \theta_0, \gamma_0)]$ . The present example, however, is just-identified which means that  $F$  is square and hence

$$-(F'WF)^{-1}F'W = F^{-1}W^{-1}(F')^{-1}F'W = -F^{-1}$$

Now, given the special structure of our example,

$$F = \begin{bmatrix} \mathbb{E}\{\nabla'_\theta g(\mathbf{w}; \theta_0, \gamma_0)\} & \mathbb{E}\{\nabla'_\gamma g(\mathbf{w}; \theta_0, \gamma_0)\} \\ \mathbf{0} & -\mathbf{I} \end{bmatrix} \equiv \begin{bmatrix} G_\theta & G_\gamma \\ \mathbf{0} & -\mathbf{I} \end{bmatrix}$$

because  $h$  does not involve  $\theta$  and  $\nabla'_\gamma h(\mathbf{w}, \gamma) = -\mathbf{I}$ . Inverting, we have

$$-F^{-1} = \begin{bmatrix} -G_\theta & -G_\gamma \\ \mathbf{0} & \mathbf{I} \end{bmatrix}^{-1} = \begin{bmatrix} -G_\theta^{-1} & -G_\theta^{-1}G_\gamma \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$$

---

<sup>2</sup>See Newey & McFadden (1994), Section 6.

We see from this expression that if  $G_\gamma$  were zero, the first step-estimation would not affect the limit distribution of  $\hat{\theta}$ . Differentiating,

$$\begin{bmatrix} \nabla g_\beta & \nabla g_{\alpha_1} \end{bmatrix} = \begin{bmatrix} -\left(\frac{zT - qp}{1 - \alpha_1}\right) & -\left\{\frac{\beta(zT - qp)}{(1 - \alpha_1)^2}\right\} \\ 2\left\{\frac{\beta(zT - qp) - (zyT - qr)}{1 - \alpha_1}\right\} & \frac{\beta^2(zT - qp) - 2\beta(zyT - qr)}{(1 - \alpha_1)^2} \end{bmatrix}$$

and thus, taking expectations,

$$G_\theta = \begin{bmatrix} \frac{-\text{Cov}(z, T)}{1 - \alpha_1} & \frac{-\beta \text{Cov}(z, T)}{(1 - \alpha_1)^2} \\ 2\left\{\frac{\beta \text{Cov}(z, T) - \text{Cov}(yT, z)}{1 - \alpha_1}\right\} & \frac{\beta^2 \text{Cov}(z, T) - 2\beta \text{Cov}(yT, z)}{(1 - \alpha_1)^2} \end{bmatrix}$$

Now, for  $G_\gamma$  we have

$$\begin{aligned} G_\gamma &= \mathbb{E} \begin{bmatrix} \nabla_q g & \nabla_p \mu & \nabla_\mu g & \nabla_s g & \nabla_r g \end{bmatrix} \\ &= \begin{bmatrix} \left(\frac{p\beta}{1 - \alpha_1} - \mu\right) & \left(\frac{q\beta}{1 - \alpha_1}\right) & -q & 0 & 0 \\ \left(\frac{\beta}{1 - \alpha_1}\right)(2r - \beta p) - s & \frac{-q\beta^2}{1 - \alpha_1} & 0 & -q & \frac{2\beta q}{1 - \alpha_1} \end{bmatrix} \end{aligned}$$

The next step is to invert  $G_\theta$ . First we calculate the determinant. For the purposes of this calculation, use the shorthand  $C = \text{Cov}(z, T)$  and  $D = \text{Cov}(yT, z)$ . We have:

$$\begin{aligned} |G_\theta| &= \left[\frac{-C}{1 - \alpha_1}\right] \left[\frac{\beta^2 C - 2\beta D}{(1 - \alpha_1)^2}\right] - \left[\frac{-\beta C}{(1 - \alpha_1)^2}\right] \left[\frac{2\beta C - 2D}{1 - \alpha_1}\right] \\ &= \left(\frac{1}{1 - \alpha_1}\right)^3 [2\beta CD - \beta^2 C^2 + 2\beta^2 C^2 - 2\beta CD] \\ &= \frac{\beta^2 \text{Cov}(z, T)^2}{(1 - \alpha_1)^3} \end{aligned}$$

Thus,

$$\begin{aligned}
G_{\theta}^{-1} &= \frac{(1 - \alpha_1)^3}{\beta^2 \text{Cov}(z, T)^2} \begin{bmatrix} \frac{\beta^2 \text{Cov}(z, T) - 2\beta \text{Cov}(yT, z)}{(1 - \alpha_1)^2} & \frac{\beta \text{Cov}(z, T)}{(1 - \alpha_1)^2} \\ -2 \left\{ \frac{\beta \text{Cov}(z, T) - \text{Cov}(yT, z)}{1 - \alpha_1} \right\} & \frac{-\text{Cov}(z, T)}{1 - \alpha_1} \end{bmatrix} \\
&= \begin{bmatrix} \left\{ \frac{1 - \alpha_1}{\text{Cov}(z, T)} \right\} \left\{ 1 - \frac{2\text{Cov}(yT, z)}{\beta \text{Cov}(z, T)} \right\} & \frac{1 - \alpha_1}{\beta \text{Cov}(z, T)} \\ \frac{2(1 - \alpha_1)^2}{\beta \text{Cov}(z, T)} \left\{ \frac{\text{Cov}(yT, z)}{\beta \text{Cov}(z, T)} - 1 \right\} & \frac{-(1 - \alpha_1)^2}{\beta^2 \text{Cov}(z, T)} \end{bmatrix}
\end{aligned}$$

The next step is to calculate  $\Omega$ :

$$\Omega = \lim_{n \rightarrow \infty} \text{Var} [\sqrt{n} f_n(\theta_0, \gamma_0)] = \lim_{n \rightarrow \infty} \text{Var} \left[ \frac{1}{\sqrt{n}} \sum_{i=1}^n f(\mathbf{w}_i; \theta_0, \gamma_0) \right]$$

If  $\mathbf{w}_i$  is an iid sequence of RVs, then

$$\Omega = \lim_{n \rightarrow \infty} \frac{1}{n} \text{Var} \left[ \sum_{i=1}^n f(\mathbf{w}_i; \theta_0, \gamma_0) \right] = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \text{Var} [f(\mathbf{w}_i; \theta_0, \gamma_0)] = \text{Var} [f(\mathbf{w}_i; \theta_0, \gamma_0)]$$

And assuming that our model is correctly specified, so that  $\mathbb{E} [f(\mathbf{w}_i; \theta_0, \gamma_0)] = 0$ ,

$$\begin{aligned}
\text{Var} [f(\mathbf{w}_i; \theta_0, \gamma_0)] &= \mathbb{E} \begin{bmatrix} g(\mathbf{w}_i; \theta_0, \gamma_0) g(\mathbf{w}_i; \theta_0, \gamma_0)' & g(\mathbf{w}_i; \theta_0, \gamma_0) h(\mathbf{w}_i; \theta_0, \gamma_0)' \\ h(\mathbf{w}_i; \theta_0, \gamma_0) g(\mathbf{w}_i; \theta_0, \gamma_0)' & h(\mathbf{w}_i; \theta_0, \gamma_0) h(\mathbf{w}_i; \theta_0, \gamma_0)' \end{bmatrix} \\
&\equiv \begin{bmatrix} \Omega_{gg} & \Omega_{gh} \\ \Omega_{gh} & \Omega_{hh} \end{bmatrix}
\end{aligned}$$

We now calculate each block.

I don't think this is actually going to give us anything interpretable. The expressions are quite involved and it seems unlikely that they'll cancel in a useful way. This doesn't matter for implementation, of course, since it's easy to calculate the estimate of  $\hat{\Omega}$  by plugging the GMM estimates into the sample moment conditions, taking the outer product, and averaging.

We are only interested in the asymptotic variance matrix  $V_{\theta}$  of our parameters of interest

$\theta = (\beta, \alpha_1)$ . We calculate this as follows:

$$\begin{aligned}
V_\theta &= \begin{bmatrix} G_\theta^{-1} & G_\theta^{-1}G_\gamma \end{bmatrix} \begin{bmatrix} \Omega_{gg} & \Omega_{gh} \\ \Omega_{hg} & \Omega_{hh} \end{bmatrix} \begin{bmatrix} (G_\theta^{-1})' \\ (G_\theta^{-1}G_\gamma)' \end{bmatrix} \\
&= \begin{bmatrix} G_\theta^{-1}(\Omega_{gg} + G_\gamma\Omega_{gh}) & G_\theta^{-1}(\Omega_{gh} + G_\gamma\Omega_{hh}) \end{bmatrix} \begin{bmatrix} (G_\theta^{-1})' \\ G_\gamma'(G_\theta^{-1})' \end{bmatrix} \\
&= G_\theta^{-1}(\Omega_{gg} + G_\gamma\Omega_{gh})(G_\theta^{-1})' + G_\theta^{-1}(\Omega_{gh} + G_\gamma\Omega_{hh})G_\gamma'(G_\theta^{-1})' \\
&= G_\theta^{-1}(\Omega_{gg} + G_\gamma\Omega_{hg} + \Omega_{gh}G_\gamma' + G_\gamma\Omega_{hh}G_\gamma')(G_\theta^{-1})'
\end{aligned}$$

### 11.3 Easier(?) Derivation of Simple Special Case: $\alpha_0 = 0$

Recall that we could eliminate  $\alpha_1$  from the moment conditions, yielding,

$$\beta = \frac{2\text{Cov}(yT, z)}{\text{Cov}(T, z)} - \frac{\text{Cov}(y^2, z)}{\text{Cov}(y, z)}$$

We can treat this as our  $g$  block of moment conditions with a parameter vector  $\theta$  that is simply  $\beta$ . This gives

$$g(\mathbf{w}; \beta, \gamma) = \left[ 2 \left( \frac{zTy - qr}{zT - qp} \right) - \frac{zy^2 - qs}{zy - q\mu} - \beta \right]$$

The  $h$  block of moment conditions is unchanged. Now, we have

$$F = \begin{bmatrix} \mathbb{E} \{ \nabla'_\beta g(\mathbf{w}; \beta_0, \gamma_0) \} & \mathbb{E} \{ \nabla'_\gamma g(\mathbf{w}; \beta_0, \gamma_0) \} \\ \mathbf{0} & -\mathbf{I} \end{bmatrix} \equiv \begin{bmatrix} G_\beta & G_\gamma \\ \mathbf{0} & -\mathbf{I} \end{bmatrix} = \begin{bmatrix} -1 & G_\gamma \\ \mathbf{0} & -\mathbf{I} \end{bmatrix}$$

since the derivative of  $g$  with respect to  $\beta$  is -1 and that of  $h$  with respect to  $\gamma$  is  $-\mathbf{I}$ . Inverting,

$$-F^{-1} = \begin{bmatrix} -G_\beta & -G_\gamma \\ \mathbf{0} & \mathbf{I} \end{bmatrix}^{-1} = \begin{bmatrix} -G_\beta^{-1} & -G_\beta^{-1}G_\gamma \\ \mathbf{0} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} 1 & G_\gamma \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$$

We calculate  $G_\gamma$  as follows:

$$G_\gamma = \mathbb{E} \begin{bmatrix} \nabla_q g & \nabla_p g & \nabla_\mu g & \nabla_s g & \nabla_r g \end{bmatrix}$$

$$\begin{aligned} \nabla_q g(\mathbf{w}; \beta_0, \gamma_0) &= \nabla_q \left[ 2 \left( \frac{zTy - qr}{zT - qp} \right) - \frac{zy^2 - qs}{zy - q\mu} - \beta \right] \\ &= 2 \left[ \frac{-r(zT - qp) + p(zTy - qr)}{(zT - qp)^2} \right] - \frac{-s(zy - q\mu) + \mu(zy^2 - qs)}{(zy - q\mu)^2} \\ &= \frac{2zT(py - r)}{(zT - qp)^2} - \frac{zy(\mu y - s)}{(zy - q\mu)^2} \end{aligned}$$

$$\nabla_p g(\mathbf{w}; \beta_0, \gamma_0) = \frac{2q(zTy - qr)}{(zT - qp)^2}$$

$$\nabla_\mu g(\mathbf{w}; \beta_0, \gamma_0) = \frac{-q(zy^2 - qs)}{(zy - q\mu)^2}$$

$$\nabla_s g(\mathbf{w}; \beta_0, \gamma_0) = \frac{q}{zy - q\mu}$$

$$\nabla_r g(\mathbf{w}; \beta_0, \gamma_0) = \frac{-2q}{zT - qp}$$

The next step is to calculate  $\Omega$ :

$$\begin{aligned} \Omega &= \text{Var} [f(\mathbf{w}_i; \theta_0, \gamma_0)] = \mathbb{E} \begin{bmatrix} g(\mathbf{w}_i; \theta_0, \gamma_0)g(\mathbf{w}_i; \theta_0, \gamma_0)' & g(\mathbf{w}_i; \theta_0, \gamma_0)h(\mathbf{w}_i; \theta_0, \gamma_0)' \\ h(\mathbf{w}_i; \theta_0, \gamma_0)g(\mathbf{w}_i; \theta_0, \gamma_0)' & h(\mathbf{w}_i; \theta_0, \gamma_0)h(\mathbf{w}_i; \theta_0, \gamma_0)' \end{bmatrix} \\ &\equiv \begin{bmatrix} \Omega_{gg} & \Omega_{gh} \\ \Omega_{gh} & \Omega_{hh} \end{bmatrix} \end{aligned}$$

We now calculate each block.

Still need to do this!

## 11.4 Two-Step Inference Idea

If  $\beta$  is small, then confidence intervals based on the GMM limit distribution from above will perform badly. But even in this case, inference for the *identified set*  $[\beta_{RF}, \beta_{IV}]$  should still be well-behaved, so long as the instrument is strong. The idea of this section is to

explore a two-step procedure that chooses between reporting the GMM confidence interval or inference for the identified set based on a pre-test of  $\beta_{RF}$ . Presumably conducting valid inference based on such a procedure will require a Bonferroni correction. The first step, however, is to determine the joint limiting behavior of the reduced form, IV, and GMM estimators.

**Reduced Form Estimator** The reduced form is given by

$$\begin{aligned} y &= \gamma_0 + \gamma_1 z + \eta \\ \gamma_0 &= \mathbb{E}[y|z = 0] \\ \gamma_1 &= \mathbb{E}[y|z = 0] - \mathbb{E}[y|z = 1] \end{aligned}$$

Now, we need to write  $\eta$  in terms of the “primitives” of our model. The first stage and main equation are

$$\begin{aligned} y &= c + \beta T^* + \varepsilon \\ T^* &= \pi_0 + \pi_1 z + v \\ \pi_1 &= p_1^* - p_0^* \end{aligned}$$

which implies

$$y = (c + \beta\pi_0) + (\beta\pi_1)z + (\varepsilon + \beta v)$$

so that

$$\begin{aligned} \gamma_1 &= \beta(p_1^* - p_0^*) \\ \eta &= \varepsilon + \beta v \end{aligned}$$

Now, define  $W = (\mathbf{1}, \mathbf{z})$  and  $\boldsymbol{\gamma} = (\gamma_0, \gamma_1)'$ . Then the reduced form estimator is

$$\hat{\boldsymbol{\gamma}} = (W'W)^{-1} W'\mathbf{y} = (W'W)^{-1} W'(W\boldsymbol{\gamma} + \boldsymbol{\eta}) = \boldsymbol{\gamma} + (W'W)^{-1} W'\boldsymbol{\eta}$$

and hence

$$\sqrt{n}(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}) = \left( \frac{W'W}{n} \right)^{-1} \frac{W'\boldsymbol{\eta}}{\sqrt{n}}$$

Now,

$$\left( \frac{W'W}{n} \right)^{-1} = \begin{bmatrix} 1 & \bar{\mathbf{z}} \\ \bar{\mathbf{z}} & \mathbf{z}'\mathbf{z}/n \end{bmatrix}^{-1} \xrightarrow{p} \begin{bmatrix} 1 & q \\ q & q \end{bmatrix}^{-1} = \frac{1}{q(1-q)} \begin{bmatrix} q & -q \\ -q & 1 \end{bmatrix}$$



and by the Central Limit Theorem,

$$\frac{W'\boldsymbol{\eta}}{\sqrt{n}} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \eta_i \begin{bmatrix} 1 \\ z_i \end{bmatrix} = \frac{1}{\sqrt{n}} \sum_{i=1}^n (\varepsilon_i + \beta v_i) \begin{bmatrix} 1 \\ z_i \end{bmatrix} \rightarrow_d N(\mathbf{0}, \Sigma)$$

where

$$\begin{aligned} \Sigma &= \mathbb{E} \begin{bmatrix} \eta_i^2 & z_i \eta_i^2 \\ z_i \eta_i^2 & z_i^2 \eta_i^2 \end{bmatrix} = \mathbb{E} \begin{bmatrix} (\varepsilon_i + \beta v_i)^2 & z_i (\varepsilon_i + \beta v_i)^2 \\ z_i (\varepsilon_i + \beta v_i)^2 & z_i^2 (\varepsilon_i + \beta v_i)^2 \end{bmatrix} \\ &= \mathbb{E} \begin{bmatrix} \varepsilon_i^2 + 2\beta \varepsilon_i v_i + \beta^2 v_i^2 & z_i^2 \varepsilon_i + 2\beta z_i \varepsilon_i v_i + \beta^2 z_i v_i^2 \\ z_i^2 \varepsilon_i + 2\beta z_i \varepsilon_i v_i + \beta^2 z_i v_i^2 & z_i^2 \varepsilon_i^2 + 2\beta z_i^2 \varepsilon_i v_i + \beta^2 z_i^2 v_i^2 \end{bmatrix} \end{aligned}$$

The next step is to work out the joint distribution of  $v$  and the other primitives of our model:

$$\begin{aligned} T^* = 0, z = 0 &\implies 0 = \pi_0 + v \implies v = -\pi_0 = -p_0^* \\ T^* = 1, z = 0 &\implies 1 = \pi_0 + v \implies v = 1 - \pi_0 = 1 - p_0^* \\ T^* = 0, z = 1 &\implies 0 = \pi_0 + \pi_1 + v \implies v = -(\pi_0 + \pi_1) = -p_1^* \\ T^* = 1, z = 1 &\implies 1 = \pi_0 + \pi_1 + v \implies v = 1 - (\pi_0 + \pi_1) = 1 - p_1^* \end{aligned}$$

Thus,

$$\begin{aligned} \mathbb{P}(T^* = 0, z = 0) &= \mathbb{P}(v = -p_0^*) = (1 - p_0^*)(1 - q) \\ \mathbb{P}(T^* = 1, z = 0) &= \mathbb{P}(v = 1 - p_0^*) = p_0^*(1 - q) \\ \mathbb{P}(T^* = 0, z = 1) &= \mathbb{P}(v = -p_1^*) = (1 - p_1^*)q \\ \mathbb{P}(T^* = 1, z = 1) &= \mathbb{P}(v = 1 - p_1^*) = p_1^*q \end{aligned}$$

Notice that, as must be true *by construction*,  $v$  is mean zero:

$$\begin{aligned} \mathbb{E}[v] &= -p_0^*(1 - p_0^*)(1 - q) + (1 - p_0^*)p_0^*(1 - q) - p_1^*(1 - p_1^*)q + (1 - p_1^*)p_1^*q \\ &= (1 - q)(1 - p_0^*)(p_0^* - p_0^*) + q(1 - p_1^*)(p_1^* - p_1^*) = 0 \end{aligned}$$

and uncorrelated with  $v$ :

$$\begin{aligned} \mathbb{E}[zv] &= q\mathbb{E}[v|z = 1] = q\{p_1^*\mathbb{E}[v|T^* = 1, z = 1] + (1 - p_1^*)\mathbb{E}[v|T^* = 0, z = 1]\} \\ &= q[p_1^*(1 - p_1^*) - p_1^*(1 - p_1^*)] = 0 \end{aligned}$$

Now, to calculate  $\Sigma$ , we need  $\mathbb{E}[v^2]$ ,  $\mathbb{E}[\varepsilon v]$ ,  $\mathbb{E}[z\varepsilon v]$ ,  $\mathbb{E}[z^2\varepsilon v]$ ,  $\mathbb{E}[zv^2]$ , and  $\mathbb{E}[z^2v^2]$ :

$$\begin{aligned}\mathbb{E}[v^2] &= p_0^{*2}(1-p_0^*)(1-q) + (1-p_0^*)^2 p_0^*(1-q) + p_1^{*2}(1-p_1^*)q + (1-p_1^*)^2 p_1^*q \\ &= (1-q)p_0^*(1-p_0^*)[p_0^* + (1-p_0^*)] + qp_1^*(1-p_1^*)[p_1^* + (1-p_1^*)] \\ &= (1-q)p_0^*(1-p_0^*) + qp_1^*(1-p_1^*)\end{aligned}$$

$$\begin{aligned}\mathbb{E}[\varepsilon v] &= \mathbb{E}[v\mathbb{E}(\varepsilon|v)] \\ &= -p_0^*\mathbb{P}(v = -p_0^*)\mathbb{E}(\varepsilon|v = -p_0^*) - p_1^*\mathbb{P}(v = -p_1^*)\mathbb{E}(\varepsilon|v = -p_1^*) \\ &\quad + (1-p_0^*)\mathbb{P}(v = 1-p_0^*)\mathbb{E}(\varepsilon|v = 1-p_0^*) + (1-p_1^*)\mathbb{P}(v = 1-p_1^*)\mathbb{E}(\varepsilon|v = 1-p_1^*) \\ &= -p_0^*(1-p_0^*)(1-q)\mathbb{E}[\varepsilon|T^* = 0, z = 0] - p_1^*(1-p_1^*)q\mathbb{E}[\varepsilon|T^* = 0, z = 1] \\ &\quad + (1-p_0^*)p_0^*(1-q)\mathbb{E}[\varepsilon|T^* = 1, z = 0] + (1-p_1^*)p_1^*q\mathbb{E}[\varepsilon|T^* = 1, z = 1] \\ &= -p_0^*(1-p_0^*)(1-q)(m_{00}^* - c) - p_1^*(1-p_1^*)q(m_{01}^* - c) \\ &\quad + (1-p_0^*)p_0^*(1-q)(m_{10}^* - c) + (1-p_1^*)p_1^*q(m_{11}^* - c) \\ &= (1-q)p_0^*(1-p_0^*)(m_{10}^* - m_{00}^*) + qp_1^*(1-p_1^*)(m_{11}^* - m_{01}^*)\end{aligned}$$

$$\begin{aligned}\mathbb{E}[z\varepsilon v] &= \mathbb{E}[z^2\varepsilon v] = \mathbb{E}[z^2\mathbb{E}[\varepsilon v|z]] = q\mathbb{E}[\varepsilon v|z = 1] = q\mathbb{E}_{T^*|z=1}[\mathbb{E}[\varepsilon v|T^* = 1, z = 1]] \\ &= q\{p_1^*\mathbb{E}[\varepsilon v|T^* = 1, z = 1] + (1-p_1^*)\mathbb{E}[\varepsilon v|T^* = 0, z = 1]\} \\ &= q\{p_1^*(1-p_1^*)\mathbb{E}[\varepsilon|T^* = 1, z = 1] - p_1^*(1-p_1^*)\mathbb{E}[\varepsilon|T^* = 0, z = 1]\} \\ &= qp_1^*(1-p_1^*)[m_{11}^* - m_{01}^*]\end{aligned}$$

$$\begin{aligned}\mathbb{E}[zv^2] &= \mathbb{E}[z^2v^2] = \mathbb{E}[z^2\mathbb{E}[v^2|z]] = q\mathbb{E}[v^2|z = 1] = q\mathbb{E}_{T^*|z=1}[\mathbb{E}[v^2|T^*, z = 1]] \\ &= q\{p_1^*\mathbb{E}[v^2|T^* = 1, z = 1] + (1-p_1^*)\mathbb{E}[v^2|T^* = 0, z = 1]\} \\ &= q\{p_1^*(1-p_1^*)^2 + p_1^{*2}(1-p_1^*)\} \\ &= qp_1^*(1-p_1^*)[(1-p_1^*) + p_1^*] \\ &= qp_1^*(1-p_1^*)\end{aligned}$$

Using these calculations, we find the elements of  $\Sigma$  as follows:

$$\begin{aligned}
\mathbb{E}[\eta^2] &= \mathbb{E}[\varepsilon^2 + 2\beta\varepsilon v + \beta^2 v^2] \\
&= \sigma_\varepsilon^2 + 2\beta [(1-q)p_0^*(1-p_0^*)(m_{10}^* - m_{00}^*) + qp_1^*(1-p_1^*)(m_{11}^* - m_{01}^*)] \\
&\quad + \beta^2 [(1-q)p_0^*(1-p_0^*) + qp_1^*(1-p_1^*)] \\
&= \sigma_\varepsilon^2 + (1-q)p_0^*(1-p_0^*)\beta [\beta + 2(m_{10}^* - m_{00}^*)] + qp_1^*(1-p_1^*)\beta [\beta + 2(m_{11}^* - m_{01}^*)] \\
\mathbb{E}[z^2\eta^2] &= \mathbb{E}[z\eta^2] = \mathbb{E}[z\varepsilon^2 + 2\beta z\varepsilon v + \beta^2 zv^2] = \sigma_\varepsilon^2 + 2\beta\mathbb{E}[z\varepsilon v] + \beta^2\mathbb{E}[zv^2] \\
&= q\sigma_\varepsilon^2 + 2\beta qp_1^*(1-p_1^*)(m_{11}^* - m_{01}^*) + \beta^2 qp_1^*(1-p_1^*) \\
&= q \{ \sigma_\varepsilon^2 + p_1^*(1-p_1^*)\beta [\beta + 2(m_{11}^* - m_{01}^*)] \}
\end{aligned}$$

Now, using the fact that  $\mathbb{E}[z^2\eta] = \mathbb{E}[z\eta]$ , the asymptotic variance of the reduced form estimator can be written as:

$$\begin{aligned}
\text{AVAR} [\sqrt{n}(\hat{\gamma} - \gamma)] &= \frac{1}{q^2(1-q)^2} \begin{bmatrix} q & -q \\ -q & 1 \end{bmatrix} \begin{bmatrix} \mathbb{E}(\eta^2) & \mathbb{E}(z\eta^2) \\ \mathbb{E}(z\eta^2) & \mathbb{E}(z\eta^2) \end{bmatrix} \begin{bmatrix} q & -q \\ -q & 1 \end{bmatrix} \\
&= \frac{1}{q^2(1-q)^2} \begin{bmatrix} q & -q \\ -q & 1 \end{bmatrix} \begin{bmatrix} q \{ \mathbb{E}(\eta^2) - \mathbb{E}(z\eta^2) \} & \mathbb{E}(z\eta^2) - q\mathbb{E}(\eta^2) \\ 0 & (1-q)\mathbb{E}(z\eta^2) \end{bmatrix} \\
&= \frac{1}{q^2(1-q)^2} \begin{bmatrix} q^2 \{ \mathbb{E}(\eta^2) - \mathbb{E}(z\eta^2) \} & -q^2 \{ \mathbb{E}(\eta^2) - \mathbb{E}(z\eta^2) \} \\ -q^2 \{ \mathbb{E}(\eta^2) - \mathbb{E}(z\eta^2) \} & q^2\mathbb{E}(\eta^2) + (1-2q)\mathbb{E}(z\eta^2) \end{bmatrix}
\end{aligned}$$

Now, we are only interested in the reduced form slope coefficient  $\gamma_1$ . The asymptotic variance of its OLS estimator is:

$$\begin{aligned}
\text{AVAR} [\sqrt{n}(\hat{\gamma}_1 - \gamma_1)] &= \frac{1}{q^2(1-q)^2} [q^2\mathbb{E}(\eta^2) + (1-2q)\mathbb{E}(z\eta^2)] \\
&= \frac{\mathbb{E}(\eta^2)}{(1-q)^2} + \left[ \frac{1-2q}{q^2(1-q)^2} \right] \mathbb{E}(z\eta^2)
\end{aligned}$$

In general, this will lead to quite a complicated expression. In our simulation design, however,

$$\begin{aligned}
q &= 1/2 \\
p_0^*(1-p_0^*) &= \delta(1-\delta) \\
p_1^*(1-p_1^*) &= (1-\delta)\delta \\
(m_{10}^* - m_{00}^*) &= (m_{11}^* - m_{01}^*) > 0
\end{aligned}$$

leading to the following simplifications:

$$\begin{aligned}\text{AVAR} [\sqrt{n}(\hat{\gamma}_1 - \gamma_1)] &= 4\mathbb{E}(\eta^2) \\ &= 4 \left\{ \sigma_\varepsilon^2 + 1/2 \times \delta(1 - \delta)\beta [\beta + 2(m_{10}^* - m_{00}^*)] + 1/2 \times \delta(1 - \delta)\beta [\beta + 2(m_{11}^* - m_{01}^*)] \right\} \\ &= 4 \left\{ \sigma_\varepsilon^2 + \delta(1 - \delta)\beta [\beta + 2(m_{11}^* - m_{01}^*)] \right\}\end{aligned}$$

Notice that, since  $(m_{11}^* - m_{01}^*)$  is positive, as in the simulation design, the asymptotic variance is *smallest* when  $\beta = 0$  so that there is no treatment effect.

**Robust Standard Errors** When implementing the reduced form estimator in practice we base our inference on sample residuals:

$$\hat{\eta}_i = y_i - \hat{\gamma}_0 - \hat{\gamma}_1 z_i$$

As we saw above, however, the errors  $\eta$  are heteroskedastic:  $\mathbb{E}(z\eta^2) \neq \mathbb{E}(z)\mathbb{E}(\eta^2)$  etc. For this reason, we must use robust standard errors:

$$\hat{\Sigma} = \frac{1}{n} \sum_{i=1}^n \begin{bmatrix} \hat{\eta}_i^2 & z_i \hat{\eta}_i^2 \\ z_i \hat{\eta}_i^2 & z_i^2 \hat{\eta}_i^2 \end{bmatrix}$$

leading to

$$\widehat{\text{AVAR}} [\sqrt{n}(\hat{\gamma}_1 - \gamma_1)] = \frac{\widehat{\mathbb{E}}(\eta^2)}{(1 - \hat{q})^2} + \left[ \frac{1 - 2\hat{q}}{\hat{q}^2(1 - \hat{q})^2} \right] \widehat{\mathbb{E}}(z\eta^2)$$

Note that  $\hat{q}$  is fixed and equal to 1/2 in our simulation design, so this becomes  $4\hat{\sigma}_\eta^2$ .

**Values of  $m_{tk}^*$  in the Simulation** To calculate  $\Sigma$  in our simulation design, we'll need to know the values of  $m_{tk}^*$  in the threshold-crossing model with bivariate normal errors:

$$\begin{bmatrix} \varepsilon \\ \xi \end{bmatrix} \sim N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right)$$

$$T^* = \mathbf{1} \{ \kappa_0 + \kappa_1 z + \xi > 0 \}$$

$$\kappa_0 = \Phi^{-1}(\delta)$$

$$\kappa_1 = \Phi^{-1}(1 - \delta) - \Phi^{-1}(\delta)$$

$$\begin{aligned}
m_{00}^* - c &= \mathbb{E}[\varepsilon | T^* = 0, z = 0] = \mathbb{E}[\varepsilon | \xi \leq -\kappa_0] \\
m_{01}^* - c &= \mathbb{E}[\varepsilon | T^* = 0, z = 1] = \mathbb{E}[\varepsilon | \xi \leq -(\kappa_0 + \kappa_1)] \\
m_{10}^* - c &= \mathbb{E}[\varepsilon | T^* = 1, z = 0] = \mathbb{E}[\varepsilon | \xi > -\kappa_0] \\
m_{11}^* - c &= \mathbb{E}[\varepsilon | T^* = 1, z = 1] = \mathbb{E}[\varepsilon | \xi > -(\kappa_0 + \kappa_1)]
\end{aligned}$$

$$\xi | \varepsilon \sim N(\rho\varepsilon, 1 - \rho^2)$$

$$\begin{aligned}
\mathbb{P}(\varepsilon \leq x | \xi \leq a) &= \frac{\mathbb{P}(\varepsilon \leq x, \xi \leq a)}{\mathbb{P}(\xi \leq a)} = \frac{\int_{-\infty}^x \int_{-\infty}^a f(\xi | \varepsilon) f(\varepsilon) d\xi d\varepsilon}{\Phi(a)} \\
&= \frac{\int_{-\infty}^x F_{\xi | \varepsilon}(a) f(\varepsilon) d\varepsilon}{\Phi(a)} = \frac{\int_{-\infty}^x \varphi(\varepsilon) \Phi\left(\frac{a - \rho\varepsilon}{\sqrt{1 - \rho^2}}\right) d\varepsilon}{\Phi(a)}
\end{aligned}$$

$$f(\varepsilon | \xi \leq a) = \frac{\varphi(\varepsilon)}{\Phi(a)} \Phi\left(\frac{a - \rho\varepsilon}{\sqrt{1 - \rho^2}}\right)$$

$$\mathbb{E}[\varepsilon | \xi \leq a] = \frac{1}{\Phi(a)} \int_{-\infty}^{\infty} \varepsilon \varphi(\varepsilon) \Phi\left(\frac{a - \rho\varepsilon}{\sqrt{1 - \rho^2}}\right) d\varepsilon$$

$$\begin{aligned}
\mathbb{P}(\varepsilon \leq x | \xi > a) &= \frac{\mathbb{P}(\varepsilon \leq x, \xi > a)}{\mathbb{P}(\xi > a)} = \frac{\int_{-\infty}^x \int_a^{\infty} f(\xi | \varepsilon) f(\varepsilon) d\xi d\varepsilon}{1 - \Phi(a)} \\
&= \frac{\int_{-\infty}^x [1 - F_{\xi | \varepsilon}(a)] f(\varepsilon) d\varepsilon}{1 - \Phi(a)} = \frac{\int_{-\infty}^x \varphi(\varepsilon) \left[1 - \Phi\left(\frac{a - \rho\varepsilon}{\sqrt{1 - \rho^2}}\right)\right] d\varepsilon}{1 - \Phi(a)}
\end{aligned}$$

$$f(\varepsilon | \xi > a) = \frac{\varphi(\varepsilon)}{1 - \Phi(a)} \left[1 - \Phi\left(\frac{a - \rho\varepsilon}{\sqrt{1 - \rho^2}}\right)\right]$$

$$\mathbb{E}[\varepsilon | \xi > a] = \frac{1}{1 - \Phi(a)} \int_{-\infty}^{\infty} \varepsilon \varphi(\varepsilon) \left[1 - \Phi\left(\frac{a - \rho\varepsilon}{\sqrt{1 - \rho^2}}\right)\right] d\varepsilon$$

I've checked all of these integrals in R and they're definitely correct.

**IV Estimator** First we calculate the probability limits of the IV estimators of  $\beta$  and  $c$ . Notice that *both* are inconsistent:

$$\begin{aligned}\hat{\beta}_{IV} &= \left( \tilde{Z}'\tilde{T}/n \right)^{-1} \left( \tilde{Z}'\mathbf{y}/n \right) = \left( \tilde{Z}'\tilde{T}/n \right)^{-1} \tilde{Z}' \left( \tilde{T}^*\tilde{\beta} + \boldsymbol{\varepsilon} \right) /n \\ &= \left( \tilde{Z}'\tilde{T}/n \right)^{-1} \left( \left[ \tilde{Z}'\tilde{T}^*/n \right] \tilde{\beta} + \tilde{Z}'\boldsymbol{\varepsilon}/n \right)\end{aligned}$$

$$\tilde{Z}'\boldsymbol{\varepsilon}/n \rightarrow_p \mathbb{E} \begin{bmatrix} \varepsilon \\ z\varepsilon \end{bmatrix} = \mathbf{0}$$

$$\left( \tilde{Z}'\tilde{T}/n \right)^{-1} = \begin{bmatrix} 1 & \bar{T} \\ \bar{z} & \mathbf{z}'\mathbf{T}/n \end{bmatrix} \rightarrow_p \begin{bmatrix} 1 & p \\ q & p_1q \end{bmatrix} = \frac{1}{q(p_1 - p)} \begin{bmatrix} p_1q & -p \\ -q & 1 \end{bmatrix}$$

$$\tilde{Z}'\tilde{T}^*/n = \begin{bmatrix} 1 & \bar{T}^* \\ \bar{z} & \mathbf{z}'\mathbf{T}^*/n \end{bmatrix} \rightarrow_p \begin{bmatrix} 1 & p^* \\ q & p_1^*q \end{bmatrix}$$

$$\left( \tilde{Z}'\tilde{T}/n \right)^{-1} \left( \tilde{Z}'\tilde{T}^*/n \right) \rightarrow_p \frac{1}{q(p_1 - p)} \begin{bmatrix} qp_1 & -p \\ -q & 1 \end{bmatrix} \begin{bmatrix} 1 & p^* \\ q & qp_1^* \end{bmatrix} = \begin{bmatrix} 1 & (p_1p^* - p_1^*p)/(p_1 - p) \\ 0 & (p_1^* - p^*)/(p_1 - p) \end{bmatrix}$$

$$\frac{p_1^* - p^*}{p_1 - p} = \frac{1}{p_1 - p} \left( \frac{p_1 - \alpha_0}{1 - \alpha_0 - \alpha_1} - \frac{p - \alpha_0}{1 - \alpha_0 - \alpha_1} \right) = \frac{1}{1 - \alpha_0 - \alpha_1}$$

$$\frac{p_1p^* - p_1^*p}{p_1 - p} = \frac{1}{p_1 - p} \left[ \frac{p_1(p - \alpha_0)}{1 - \alpha_0 - \alpha_1} - \frac{p(p_1 - \alpha_0)}{1 - \alpha_0 - \alpha_1} \right] = \frac{-\alpha_0}{1 - \alpha_0 - \alpha_1}$$

$$\boldsymbol{\beta}_{IV} \rightarrow_p \begin{bmatrix} 1 & (p_1p^* - p_1^*p)/(p_1 - p) \\ 0 & (p_1^* - p^*)/(p_1 - p) \end{bmatrix} \begin{bmatrix} c \\ \beta \end{bmatrix} = \frac{1}{1 - \alpha_0 - \alpha_1} \begin{bmatrix} c - \alpha_0\beta \\ \beta \end{bmatrix} \equiv \begin{bmatrix} c_{IV} \\ \beta_{IV} \end{bmatrix}$$

Now that we have the probability limit of the IV estimator, we can work out the error term  $\omega$  that corresponds to it as a function of the “primitives” of our model. We have:

$$\begin{aligned}
\zeta &= y - (c_{IV} + \beta_{IV}T) = (c + \beta T^* + \varepsilon) - \frac{1}{1 - \alpha_0 - \alpha_1} [(c - \alpha_0\beta) + \beta T] \\
&= \varepsilon + c - \left( \frac{c - \alpha_0\beta}{1 - \alpha_0 - \alpha_1} \right) + \beta \left( T^* - \frac{T}{1 - \alpha_0 - \alpha_1} \right) \\
&= \varepsilon + \left[ \frac{\alpha_0(\beta - c) - \alpha_1 c}{1 - \alpha_0 - \alpha_1} \right] + \beta \left[ \frac{(T^* - T) - (\alpha_0 + \alpha_1)T^*}{1 - \alpha_0 - \alpha_1} \right] \\
&= \varepsilon + \left[ \frac{\alpha_0(\beta - c) - \alpha_1 c}{1 - \alpha_0 - \alpha_1} \right] + \beta \left[ \frac{w - (\alpha_0 + \alpha_1)(\pi_0 + \pi_1 z + v)}{1 - \alpha_0 - \alpha_1} \right]
\end{aligned}$$

This looks pretty complicated. I’m also not sure we need to work this out analytically. All we really need is to write out the joint limit distribution of the IV and RF...

**Joint Distribution of IV and RF** To carry out inference for the identified set, we need to work out the joint distribution of the reduced form and IV estimators:

$$\begin{aligned}
\hat{\beta}_{IV} &= \left( \tilde{Z}'\tilde{T}/n \right)^{-1} \left( \tilde{Z}'\mathbf{y}/n \right) \\
\hat{\gamma} &= \left( \tilde{Z}'\tilde{Z} \right)^{-1} \left( \tilde{Z}'\mathbf{y}/n \right)
\end{aligned}$$

The probability limits of each estimator define an associated error term, each of which depends in a complicated way on our “primitive” model parameters:

$$\begin{aligned}
y &= c_{IV} + \beta_{IV}T + \zeta \\
y &= \gamma_0 + \gamma_1 z + \eta
\end{aligned}$$

But because these two estimators depend only on observable quantities, we can use their residuals to work out the covariance matrix of  $(\zeta, \eta)$ . Accordingly, we proceed as follows:

$$\begin{aligned}
\hat{\beta}_{IV} &= \left( \tilde{Z}'\tilde{T}/n \right)^{-1} \left( \tilde{Z}' \left[ \tilde{T}\beta_{IV} + \boldsymbol{\zeta} \right] /n \right) = \beta_{IV} + \left( \tilde{Z}'\tilde{T}/n \right)^{-1} \left( \tilde{Z}'\boldsymbol{\zeta}/n \right) \\
\hat{\gamma} &= \left( \tilde{Z}'\tilde{Z} \right)^{-1} \left( \tilde{Z}' \left[ \tilde{Z}\gamma + \boldsymbol{\eta} \right] /n \right) = \gamma + \left( \tilde{Z}'\tilde{Z}/n \right)^{-1} \left( \tilde{Z}'\boldsymbol{\eta}/n \right)
\end{aligned}$$

yielding

$$\begin{aligned}\sqrt{n}(\hat{\beta}_{IV} - \beta_{IV}) &= (\tilde{Z}'\tilde{T}/n)^{-1} (\tilde{Z}'\zeta/\sqrt{n}) \\ \sqrt{n}(\hat{\gamma} - \gamma) &= (\tilde{Z}'\tilde{Z}/n)^{-1} (\tilde{Z}'\eta/\sqrt{n})\end{aligned}$$

## 11.5 Other Stuff...

- Manski was interested in a heterogenous treatment effect model and whether we could bound *ATE* rather than *LATE*.
- Would it help in the performance of the GMM estimator if we enforced the bounds on  $\alpha_0$  and  $\alpha_1$ ?

**Exogenous Covariates in a Linear Model:** These should be very easy to handle because we can just stack the GMM moment conditions to include an IV estimator for the parameter on the exogenous covariates in the main equation. Recall that the usual IV estimator for these parameters is unaffected by measurement error. We should write this out since it's the case that many people will use in practice given the extreme sample size demands of fully non-parametric estimation!

**More About Weak Identification:** Sophocles pointed out in an email exchange that I had been assuming (incorrectly) that  $\text{Cov}(y, z)$  is always well-behaved. This is not the case if  $z$  is a weak instrument. I don't think we can simply assume we have a strong instrument and consider the weak identification that arises from  $\beta \approx 0$  in isolation. I think the two problems of  $\beta \approx 0$  and weak  $z$  interact in an important way since, as we saw from above, the determinant  $|G_\theta|$  that measures the strength of identification depends on the *product* of  $\beta$  and  $\text{Cov}(z, T)$ . I think the correct interpretation of this is that the magnitude of  $\beta$  that gives strong identification should be always relative to the strength of  $z$ . If  $z$  is very strong, then  $\beta$  can be smaller without causing problems. But if  $z$  is weak then I think  $\beta$  needs to be really large to get strong identification. If I recall correctly, we uncovered something in our simulations that appears to agree with this intuition but I need to go back and check.

To see why  $\text{Cov}(y, z)$  is badly behaved when  $z$  is weak, write out an explicit first-stage equation for our model as follows:

$$T^* = \pi_0 + \pi_1 z + v$$



where

$$\begin{aligned}\pi_0 &= \mathbb{E}[T^*|z=0] = p_0^* \\ \pi_1 &= \mathbb{E}[T^* = 1|z=1] - \mathbb{E}[T^*|z=0] = p_1^* - p_0^*\end{aligned}$$

and  $\mathbb{E}[zv] = 0$  by construction. Now,

$$\begin{aligned}\text{Cov}(z, y) &= \mathbb{E}(zy) - \mathbb{E}(z)\mathbb{E}(y) \\ &= \mathbb{E}[z(c + \beta T^* + \varepsilon)] - q\mathbb{E}(c + \beta T^* + \varepsilon) \\ &= \mathbb{E}[z\{c + \beta(\pi_0 + \pi_1 z + v) + \varepsilon\}] - q\mathbb{E}[c + \beta(\pi_0 + \pi_1 z + v) + \varepsilon] \\ &= q(c + \beta\pi_0) + \beta\pi_1\mathbb{E}(z^2) - q(c + \beta\pi_0 + \beta\pi_1\mathbb{E}[z]) \\ &= q(c + \beta\pi_0 + \beta\pi_1) - q(c + \beta\pi_0 + \beta\pi_1 q) \\ &= q(\beta\pi_1 - \beta\pi_1 q) = \beta\pi_1 q(1 - q) \\ &= \beta(p_1^* - p_0^*)q(1 - q)\end{aligned}$$

**Auxiliary Moment Inequalities** Notice that if  $\beta = 0$ , then the preceding moment equalities do *not* identify  $\alpha_1$ . However, we do have auxiliary moment *inequalities* that partially identify  $\alpha_1$  regardless of the value of  $\beta$ . The simplest of these comes from the relationship

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

where  $p_k = P(T = 1|z_k)$  and  $p_k^* = P(T^* = 1|z_k)$ . (This follows from the Law of Total Probability and our assumption that the mis-classification probabilities rates depend only on  $T^*$ , not  $z$ .) Under our assumption that  $\alpha_0 + \alpha_1 < 1$ , we obtain  $\alpha_0 < \min_k p_k$  and  $\alpha_1 < \min_k (1 - p_k)$ . If  $\alpha_0 = 0$ , as we assume in the present special case, then without any assumption on the true value of  $\alpha_1$  we have

$$0 \leq \alpha_1 < \min_k (1 - p_k) = 1 - \max_k p_k.$$

Is there some way to use these moment inequalities in estimation?

**Under Normality** In our simulation for the CDF bounds on  $\alpha_0$  and  $\alpha_1$ , we found that the upper bounds were in fact equal to the true parameter values. This is very surprising and is very likely comes from the specific parametric model from which we simulated. This happens to have been a model with normally distributed errors. Can we say anything about such a model theoretically? Perhaps try to write down the likelihood function? This could

also be a useful way to look at the weak identification problem.

## 12 April 2017 – New GMM Formulation

Consider the special case in which  $\alpha_0 = 0$  so the model is identified from

$$\begin{aligned}\text{Cov}(y, z) - \left( \frac{\beta}{1 - \alpha_1} \right) \text{Cov}(T, z) &= 0 \\ \text{Cov}(y^2, z) - \left( \frac{\beta}{1 - \alpha_1} \right) [2\text{Cov}(yT, z) - \beta\text{Cov}(T, z)] &= 0\end{aligned}$$

Above we wrote this in a standard GMM form by adding auxiliary moment equations to identify  $\mathbb{E}[y^2]$ ,  $\mathbb{E}[yT]$ , etc. But there's a simpler and more transparent way to do this. Under our assumptions and  $\alpha_0 = 0$ , some algebra shows that

Add the algebra from the whiteboard notes later

$$\mathbb{E} \left[ y - \frac{\beta}{1 - \alpha_1} T \right] = c$$

and that

$$\mathbb{E} \left[ y^2 - \frac{\beta}{1 - \alpha_1} 2yT + \frac{\beta^2}{1 - \alpha_1} T^2 \right] = \sigma_{\varepsilon\varepsilon} + c^2$$

where  $c$  is the intercept from the regression model and  $\sigma_{\varepsilon\varepsilon} = \text{Var}(\varepsilon)$ . Now, re-writing the first covariance equation using the linearity of expectation,

$$\begin{aligned}\mathbb{E} \left[ yz - \frac{\beta}{1 - \alpha_1} Tz \right] - \mathbb{E}[z] \mathbb{E} \left[ y - \frac{\beta}{1 - \alpha_1} T \right] &= 0 \\ \mathbb{E} \left[ yz - \frac{\beta}{1 - \alpha_1} Tz \right] - \mathbb{E}[z]c &= 0 \\ \mathbb{E} \left[ \left\{ y - c - \frac{\beta}{1 - \alpha_1} T \right\} z \right] &= 0\end{aligned}$$

and proceeding similarly for the second covariance equation

$$\begin{aligned}\mathbb{E} \left[ y^2 z - \frac{\beta}{1 - \alpha_1} 2yTz + \frac{\beta^2}{1 - \alpha_1} T^2 z \right] - \mathbb{E}[z] \mathbb{E} \left[ y^2 - \frac{\beta}{1 - \alpha_1} 2yT + \frac{\beta^2}{1 - \alpha_1} T^2 \right] &= 0 \\ \mathbb{E} \left[ y^2 z - \frac{\beta}{1 - \alpha_1} 2yTz + \frac{\beta^2}{1 - \alpha_1} T^2 z \right] - \mathbb{E}[z] (\sigma_{\varepsilon\varepsilon} + c^2) &= 0 \\ \mathbb{E} \left[ \left\{ y^2 - \sigma_{\varepsilon\varepsilon} - c^2 - \frac{\beta}{1 - \alpha_1} 2yT + \frac{\beta^2}{1 - \alpha_1} T^2 \right\} z \right] &= 0\end{aligned}$$

Thus, we can express our estimator in terms of the following four moment equations:

$$\begin{aligned}\mathbb{E} \left[ y - c - \frac{\beta}{1 - \alpha_1} T \right] &= 0 \\ \mathbb{E} \left[ y^2 - \sigma_{\varepsilon\varepsilon} - c^2 - \frac{\beta}{1 - \alpha_1} 2yT + \frac{\beta^2}{1 - \alpha_1} T \right] &= 0 \\ \mathbb{E} \left[ \left\{ y - c - \frac{\beta}{1 - \alpha_1} T \right\} z \right] &= 0 \\ \mathbb{E} \left[ \left\{ y^2 - \sigma_{\varepsilon\varepsilon} - c^2 - \frac{\beta}{1 - \alpha_1} 2yT + \frac{\beta^2}{1 - \alpha_1} T \right\} z \right] &= 0\end{aligned}$$

To simplify the notation, let  $\boldsymbol{\theta} = (\alpha_1, \beta, c, \sigma_{\varepsilon\varepsilon})'$  and define

$$\begin{aligned}u(\boldsymbol{\theta}) &= y - c - \frac{\beta}{1 - \alpha_1} T \\ v(\boldsymbol{\theta}) &= y^2 - \sigma_{\varepsilon\varepsilon} - c^2 - \frac{\beta}{1 - \alpha_1} 2yT + \frac{\beta^2}{1 - \alpha_1} T\end{aligned}$$

Then we can express the four moment equalities as

$$\mathbb{E} [g_1(\mathbf{x}, \boldsymbol{\theta})] = \mathbb{E} \begin{bmatrix} u(\boldsymbol{\theta}) \\ v(\boldsymbol{\theta}) \end{bmatrix} = \mathbf{0}, \quad \mathbb{E} [g_2(\mathbf{x}, \boldsymbol{\theta})] = \mathbb{E} \begin{bmatrix} u(\boldsymbol{\theta})z \\ v(\boldsymbol{\theta})z \end{bmatrix} = \mathbf{0}$$

We also have two moment inequalities, namely  $\alpha_1 \leq 1 - p_1$  and  $\alpha_1 \leq 1 - p_0$ . After some algebra (see the whiteboard), we can show that these are equivalent to

$$\mathbb{E} [h(\mathbf{x}, \theta)] = \mathbb{E} \begin{bmatrix} (1 - \alpha_1) - T(1 - z)/(1 - q) \\ (1 - \alpha_1) - Tz/q \end{bmatrix} \geq \mathbf{0}$$

where  $q = \mathbb{E}[z]$ . We will *condition* on  $z$ , i.e. hold it fixed in repeated samples, so we will not add an extra moment condition for  $q$ . Instead we will simply substitute the sample analogue.

To formulate the GMM estimator with moment inequalities as in Moon and Schorfheide (2009), we introduce some further notation. Let  $\boldsymbol{\lambda}$  denote the slack in  $h(\mathbf{x}, \theta)$ , so that

$$\mathbb{E}[h(\mathbf{x}, \theta)] = \boldsymbol{\lambda} \geq \mathbf{0} \iff \mathbb{E}[h(\mathbf{x}, \theta) - \boldsymbol{\lambda}] = \mathbf{0}$$

Further define

$$f(\mathbf{x}) = \begin{bmatrix} g(\mathbf{x}, \theta) \\ h(\mathbf{x}, \theta) \end{bmatrix}, \quad \psi(\mathbf{x}, \boldsymbol{\theta}, \boldsymbol{\lambda}) = \begin{bmatrix} g(\mathbf{x}, \boldsymbol{\theta}) \\ h(\mathbf{x}, \boldsymbol{\theta}) - \boldsymbol{\lambda} \end{bmatrix}$$

and let  $\Theta = \{\boldsymbol{\theta}: \alpha_1 \geq 0, \sigma_{\varepsilon\varepsilon} \geq 0\}$ . Then, the GMM estimator based on our moment equalities

and inequalities is

$$(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\lambda}}) = \arg \min_{\boldsymbol{\theta} \in \Theta, \boldsymbol{\lambda} \geq \mathbf{0}} Q_n(\boldsymbol{\theta}, \boldsymbol{\lambda})$$

where

$$Q_n(\boldsymbol{\theta}, \boldsymbol{\lambda}) = \frac{1}{2} \bar{\boldsymbol{\psi}}_n(\mathbf{x}, \boldsymbol{\theta}, \boldsymbol{\lambda})' \mathbf{W}_n \bar{\boldsymbol{\psi}}_n(\mathbf{x}, \boldsymbol{\theta}, \boldsymbol{\lambda})$$

$$\bar{\boldsymbol{\psi}}_n(\boldsymbol{\theta}, \boldsymbol{\lambda}) = \frac{1}{n} \sum_{i=1}^n \boldsymbol{\psi}(\mathbf{x}_i, \boldsymbol{\theta}, \boldsymbol{\lambda})$$

and  $\mathbf{W}_n$  is a weighting matrix that should be irrelevant in our case because we're just identified. We scale the criterion by  $1/2$  so that the derivative looks nice.

Is this still the case with the moment inequalities? Mechanically we have introduced two new parameters to match the two extra conditions.

At various points we will need the derivatives of the moment equations. Let

$$\mathbf{G}(\boldsymbol{\theta}) = \mathbb{E}[\nabla_{\boldsymbol{\theta}'} g(\mathbf{x}, \boldsymbol{\theta})], \quad \mathbf{H}(\boldsymbol{\theta}) = \mathbb{E}[\nabla_{\boldsymbol{\theta}'} h(\mathbf{x}, \boldsymbol{\theta})]$$

and define

$$\begin{aligned} \boldsymbol{\Psi}(\boldsymbol{\theta}, \boldsymbol{\lambda}) &= \mathbb{E} \begin{bmatrix} \nabla_{\boldsymbol{\theta}'} \boldsymbol{\psi}(\mathbf{x}, \boldsymbol{\theta}, \boldsymbol{\lambda}) & \nabla_{\boldsymbol{\lambda}'} \boldsymbol{\psi}(\mathbf{x}, \boldsymbol{\theta}, \boldsymbol{\lambda}) \end{bmatrix} \\ &= \mathbb{E} \begin{bmatrix} \textit{fill} & \textit{this} \\ \textit{in} & \textit{later} \end{bmatrix} \end{aligned}$$

## 13 April 1–15, 2017 – Andrews & Soares (2010)

Our moment equalities from above do not identify  $\alpha$  when  $\beta = 0$ . More generally, the estimator based on them performs poorly when  $\beta$  is relatively small compared to the error variance. Continue to assume that  $\alpha_0 = 0$  so the moment conditions simplify.

We now consider an inference procedure following Andrews & Soares. The basic idea is to “isolate” the problematic parameters, in our case  $\alpha$  and  $\beta$ , and carry out joint inference for these using the Anderson-Rubin test statistic. This is constructed by substituting a null hypothesis  $H_0: \boldsymbol{\theta} = \boldsymbol{\theta}_0$  into the sample analogue of the GMM moment conditions and relying on the fact that this sample analogue remains “well-behaved” even in situations where inference for the GMM *parameter estimator* breaks down. Examples include parameters on the boundary, and parameters that may not be identified, e.g.  $\alpha_1$  if  $\beta = 0$

### 13.1 Simple Example: $\alpha_0 = 0$ , $c = 0$ , and $\sigma_{\varepsilon\varepsilon} = 1$

In our problem  $c$  and  $\sigma_{\varepsilon\varepsilon}$  are essentially nuisance parameters. Fortunately, they are *always identified* from our moment conditions regardless of the values of  $\beta$  and  $\alpha$ , as we will discuss further below. For the moment, we will suppose that  $c$  is known to equal zero and  $\varepsilon_{\varepsilon\varepsilon}$  is known to equal one as is the case in our baseline simulation. Later we will estimate them which requires only a small modification of the procedure we know outline. With the simplifications  $\alpha_0, c = 0, \sigma_{\varepsilon\varepsilon} = 1$  the equality moment conditions become

$$\begin{aligned}\mathbb{E} \left[ \left\{ y - \frac{\beta}{1 - \alpha_1} T \right\} z \right] &= 0 \\ \mathbb{E} \left[ \left\{ y^2 - 1 - \frac{\beta}{1 - \alpha_1} 2yT + \frac{\beta^2}{1 - \alpha_1} T \right\} z \right] &= 0\end{aligned}$$

since we no longer need the  $g_1$  block of moment conditions to identify the “intercepts”  $c$  and  $\sigma_{\varepsilon\varepsilon}$ . For this simplified set of moment conditions, our parameter vector is  $\boldsymbol{\theta} = (\alpha_1, \beta)'$  and the residuals are given by

$$\begin{aligned}u(\boldsymbol{\theta}) &= y - \frac{\beta}{1 - \alpha_1} T \\ v(\boldsymbol{\theta}) &= y^2 - 1 - \frac{\beta}{1 - \alpha_1} 2yT + \frac{\beta^2}{1 - \alpha_1}\end{aligned}$$

and we can write the equality moment conditions as

$$\mathbb{E} [g(\mathbf{x}, \boldsymbol{\theta})] = \mathbb{E} \begin{bmatrix} u(\boldsymbol{\theta})z \\ v(\boldsymbol{\theta})z \end{bmatrix} = \mathbf{0}$$

The inequality moment conditions are unchanged from above, namely

$$\mathbb{E} [h(\mathbf{x}, \theta)] = \mathbb{E} \begin{bmatrix} (1 - \alpha_1) - T(1 - z)/(1 - q) \\ (1 - \alpha_1) - Tz/q \end{bmatrix} \geq \mathbf{0}$$

where  $q = \mathbb{E}[z]$ . As above we will *condition* on  $z$ , i.e. hold it fixed in repeated samples, so we will not add an extra moment condition for  $q$ . Instead we will simply substitute the sample analogue. Note that this means we should hold  $q$  *fixed* when bootstrapping below. We now introduce some notation from Andrews and Soares (2010).

## Population Moment Conditions

$$\mathbb{E}[m_j(\mathbf{w}_i, \theta_0)] \begin{cases} \geq 0 & \text{for } j = 1, \dots, p \\ = 0 & \text{for } j = p+1, \dots, k \text{ where } k = p+v \end{cases}$$

where  $p$  is the number of inequality moment conditions (in our case  $p = 2$ ),  $v$  is the number of equality moment conditions (in our case  $v = 2$ ),  $\theta_0$  is the true parameter vector, and  $\mathbf{w}_i$  is the vector of observations for individual  $i$  (in our case  $\mathbf{w}_i = (T_i, z_i, y_i)$ ).

## Sample Moment Functions, etc.

$$\bar{m}_n(\theta) = \begin{bmatrix} \bar{m}_{n,1}(\theta) \\ \vdots \\ \bar{m}_{n,k}(\theta) \end{bmatrix}, \quad \bar{m}_{n,j} = \frac{1}{n} \sum_{i=1}^n m_j(\mathbf{w}_i, \theta) \text{ for } j = 1, \dots, k$$

Now, let  $\Sigma(\theta_0)$  denote the asymptotic variance of  $\sqrt{n} \bar{m}_n(\theta)$ . We estimate this quantity using  $\hat{\Sigma}_n(\theta)$ . For iid observations, as in our example, the estimator is

$$\hat{\Sigma}(\theta) = \frac{1}{n} \sum_{i=1}^n [m(\mathbf{w}_i, \theta) - \bar{m}_n(\theta)] [m(\mathbf{w}_i, \theta) - \bar{m}_n(\theta)]', \quad m(\mathbf{w}_i, \theta) = \begin{bmatrix} m_1(\mathbf{w}_i, \theta) \\ \vdots \\ m_k(\mathbf{w}_i, \theta) \end{bmatrix}$$

**Test Statistic** The test statistic takes the form  $T_n(\theta) = S\left(\sqrt{n} \bar{m}_n(\theta), \hat{\Sigma}(\theta)\right)$  for some real-valued function  $S$ . The example we will use is  $S_1$ , defined by

$$S_1(m, \Sigma) = \sum_{j=1}^p [m_j/\sigma_j]_-^2 + \sum_{j=p+1}^{p+v} (m_j/\sigma_j)^2$$

where  $m = (m_1, \dots, m_k)'$ ,

$$[x]_- = \begin{cases} x, & \text{if } x < 0 \\ 0, & \text{if } x \geq 0 \end{cases}$$

and  $\sigma_j^2$  is the  $j$ th diagonal element of  $\Sigma$ . Notice that  $S_1$  only gives weight to inequality moment conditions that are *violated*.

**Basic Idea of the Test** Let's return to our specific example for a moment. The idea is essentially to plug a null hypothesis  $\theta^* = (\alpha_1^*, \beta^*)'$  into the sample analogue:

$$\sqrt{n} \bar{m}_n(\alpha_1^*, \beta^*) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \begin{bmatrix} (1 - \alpha_1^*) - T_i(1 - z_i)/(1 - q) \\ (1 - \alpha_1^*) - T_i z_i/q \\ \left(y_i - \frac{\beta^*}{1 - \alpha_1^*} T_i\right) z_i \\ \left(y_i^2 - 1 - \frac{\beta^*}{1 - \alpha_1^*} 2y_i T_i + \frac{\beta^{*2}}{1 - \alpha_1^*} T_i\right) z_i \end{bmatrix}$$

and see if the result is “large” after standardizing and squaring the individual elements. We only give weight to an inequality if it is violated. The variance matrix of the sample analogue is calculated *under the null*, i.e. assuming that  $\theta = \theta^*$ . Note that we also use the *centered variance matrix estimator*.

We reject the null if the test statistic is too large. This gives us *joint inference* for  $\alpha$  and  $\beta$  *simultaneously*. To construct a joint confidence region, we need to test pairs  $(\alpha_1, \beta)$ . Of course, we restrict  $\alpha_1$  to lie in  $[0, 1)$ . The resulting confidence region *need not be convex*. In fact it could even be disconnected! However, in our particular example, it might be possible to prove that one gets a connected or even convex region. This is something we should think about since it would reduce the computational burden substantially. To get *marginal* inference, say for  $\beta$  only, one projects the joint confidence set. This is necessarily conservative, but may not be too bad in practice. We'll have to see...

A particularly salient null hypothesis is  $\beta = 0$ . Imposing this yields

$$\sqrt{n} \bar{m}_n(\alpha_1^*, 0) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \begin{bmatrix} (1 - \alpha_1^*) - T_i(1 - z_i)/(1 - q) \\ (1 - \alpha_1^*) - T_i z_i/q \\ y_i z_i \\ (y_i^2 - 1) z_i \end{bmatrix}$$

We see that this function depends on  $\alpha_1^*$  *only via the moment inequalities*. What is more, the test statistic based on  $S_1$  does *not* depend on  $\alpha_1^*$  unless the inequality constraints are violated.

**Calculating the Critical Value** The test statistic we will use to test  $\theta = \theta^*$  is fairly simple to compute: we simply substitute into the GMM sample analogue. The critical value

for the test, however, is much more complicated. Following Andrews & Soares (2010), we use the following bootstrap procedure. First we define some additional notation. All of the test statistics considered in Andrews & Soares (2010) satisfy

$$T_n = S \left( \sqrt{n} \bar{m}_n(\theta), \hat{\Sigma}(\theta) \right) = S \left( \hat{D}^{-1/2}(\theta) \sqrt{n} \bar{m}_n(\theta), \hat{\Omega}_n(\theta) \right)$$

where

$$\hat{D}_n(\theta) = \text{diag} \left( \hat{\Sigma}(\theta) \right), \quad \hat{\Omega}_n(\theta) = \hat{D}_n^{-1/2}(\theta) \hat{\Sigma}(\theta) \hat{D}_n^{-1/2}(\theta)$$

Now, let  $\{\mathbf{w}_i^*\}_{i=1}^n$  denote a bootstrap sample and define the associated bootstrap quantities

$$\begin{aligned} M_n^*(\theta) &= \sqrt{n} \left( \hat{D}^*(\theta) \right)^{-1/2} (\bar{m}_n^*(\theta) - \bar{m}_n(\theta)) \\ \hat{\Omega}^*(\theta) &= \left( \hat{D}^*(\theta) \right)^{-1/2} \hat{\Sigma}_n^*(\theta) \left( \hat{D}^*(\theta) \right)^{-1/2} \\ \hat{D}^*(\theta) &= \text{diag} \left( \hat{\Sigma}_n^*(\theta) \right) \\ m_n^*(\theta) &= \frac{1}{n} \sum_{i=1}^n m(\mathbf{w}_i^*, \theta) \\ \hat{\Sigma}(\theta)^* &= \frac{1}{n} \sum_{i=1}^n [m(\mathbf{w}_i^*, \theta) - \bar{m}_n^*(\theta)] [m(\mathbf{w}_i^*, \theta) - \bar{m}_n^*(\theta)]' \end{aligned}$$

Note that  $M_n^*(\theta)$  is centered around the *non-bootstrap* sample analogue  $\bar{m}_n(\theta)$ : this is *very important!* Now we describe the procedure for calculating the bootstrap critical value:

1. Calculate  $\sqrt{n} \bar{m}_n(\theta_0)$  and  $\hat{\Sigma}(\theta_0)$  under the null hypothesis  $H_0: \theta = \theta_0$ .
2. Determine which inequality moment conditions are “far from binding” as follows:
  - Let  $j \in J = \{1, \dots, p\}$  index the inequality moment conditions.
  - Let  $\hat{\sigma}_{n,j}(\theta_0)^2$  denote the  $(j, j)$  element of  $\hat{\Sigma}(\theta_0)$
  - For each  $j \in J$  calculate the “t-statistic”  $t_{n,j} = \sqrt{n} \bar{m}_j(\theta_0) / \hat{\sigma}_{n,j}(\theta_0)$
  - Let  $\mathcal{FB}$  denote the subset of  $J$  for which  $t_{n,j} > \sqrt{\log n}$ . These are the inequality moment conditions that are “far from binding” under  $H_0: \theta = \theta_0$ .
3. Calculate the test statistic  $T_n = S_1 \left( \sqrt{n} \bar{m}_n(\theta_0), \hat{\Sigma}(\theta_0) \right)$ .
4. Calculate the bootstrap critical value for the test as follows:
  - Draw  $R$  bootstrap samples – each with sample size  $n$ .



- For each bootstrap sample,  $r$ , calculate  $M_{n,r}^{**}(\theta_0)$  and  $\widehat{\Omega}_{n,r}^{**}(\theta_0)$  – the bootstrap versions of  $M_n(\theta_0)$  and  $\widehat{\Omega}(\theta_0)$ , defined above but with a slight change: *drop* any moment inequality  $j \in \mathcal{FB}$ . That is, drop any inequality that we determined was far from binding *on the basis of the real data* (i.e. *not* this bootstrap sample!)
- For each bootstrap sample  $r$  calculate  $T_{n,r}^{**} = S_1 \left( M_{n,r}^{**}(\theta_0), \widehat{\Omega}_{n,r}^{**}(\theta_0) \right)$ .
- Set  $\widehat{c}_n(\theta_0, 1 - \delta)$  equal to the  $1 - \delta$  sample quantile of the  $\{T_{n,r}^{**}\}_{r=1}^R$

5. Reject  $H_0: \theta = \theta_0$  if  $T_n > \widehat{c}_n(\theta_0, 1 - \delta)$

6. To construct a  $(1 - \delta) \times 100\%$  confidence set, invert the test of  $H_0: \theta = \theta_0$  for  $\theta_0 \in \Theta$ .

## 14 April 15–18, 2017

### Second Moment Inequalities

**Step 1: Second Moment Bounds** Our model has  $\varepsilon = y - c - bT^*$  where  $\mathbb{E}[\varepsilon] = 0$ . Conditional on  $T^* = 0$ ,  $\varepsilon = y - c$  and conditional on  $T^* = 1$ ,  $\varepsilon = y - c - \beta$ . Hence,

$$\begin{aligned}\mathbb{E}[\varepsilon^2 | T^* = 0, z = k] &= \mathbb{E}[y^2 - 2cy | T^* = 0, z = k] + c^2 \\ \mathbb{E}[\varepsilon^2 | T^* = 1, z = k] &= \mathbb{E}[y^2 - 2(\beta + c)y | T^* = 1, z = k] + (\beta + c)^2\end{aligned}$$

The bounds will impose that  $\text{Var}(\varepsilon | T^*, z) > 0$  which is equivalent to  $\mathbb{E}[\varepsilon^2 | T^*, z] > 0$  since  $\varepsilon$  is mean zero:

$$\begin{aligned}\mathbb{E}[y^2 | T^* = 0, z = k] &> 2c\mathbb{E}[y | T^* = 0, z = k] - c^2 \\ \mathbb{E}[y^2 | T^* = 1, z = k] &> 2(\beta + c)\mathbb{E}[y | T^* = 1, z = k] - (\beta + c)^2\end{aligned}$$

These are not in fact the bounds we want! I messed up since  $\varepsilon$  is *not* mean zero *conditional* on  $T^*$  and  $z$ . These bounds are still correct, but they're not the best we can do and may not in fact be better than the simple first-moment bounds...

**Step 2: Relate  $\mathbb{E}[y^r | T^*, z]$  to  $\mathbb{E}[y^r | T, z]$**  By iterated expectations and the assumption that  $\mathbb{E}[y^r | T, T^*, z] = \mathbb{E}[y^r | T^*, z]$ ,

$$\begin{aligned}\mathbb{E}[y^r | T = 0, z_k] &= \mathbb{E}[y^r | T^* = 0, z_k] \mathbb{P}(T^* = 0 | T = 0, z_k) + \mathbb{E}[y^r | T^* = 1, z_k] \mathbb{P}(T^* = 1 | T = 0, z_k) \\ \mathbb{E}[y^r | T = 1, z_k] &= \mathbb{E}[y^r | T^* = 0, z_k] \mathbb{P}(T^* = 0 | T = 1, z_k) + \mathbb{E}[y^r | T^* = 1, z_k] \mathbb{P}(T^* = 1 | T = 1, z_k)\end{aligned}$$

The preceding is a linear system of the form

$$\begin{aligned} a &= \pi x + (1 - \pi)y \\ b &= (1 - \delta)x + \delta y \end{aligned}$$

and hence its solution is

$$\begin{aligned} x &= \left[ \frac{\delta}{\pi\delta - (1 - \pi)(1 - \delta)} \right] a + \left[ \frac{-(1 - \pi)}{\pi\delta - (1 - \pi)(1 - \delta)} \right] b \\ y &= \left[ \frac{-(1 - \delta)}{\pi\delta - (1 - \pi)(1 - \delta)} \right] a + \left[ \frac{\pi}{\pi\delta - (1 - \pi)(1 - \delta)} \right] b \end{aligned}$$

by Bayes' rule, as we show in the appendix to sick-instruments,

$$\begin{aligned} \pi &= \mathbb{P}(T^* = 0 | T = 0, z_k) = (1 - \alpha_0)(1 - p_k^*) / (1 - p_k) \\ 1 - \pi &= \mathbb{P}(T^* = 1 | T = 0, z_k) = \alpha_1 p_k^* / (1 - p_k) \\ \delta &= \mathbb{P}(T^* = 1 | T = 1, z_k) = (1 - \alpha_1) p_k^* / p_k \\ 1 - \delta &= \mathbb{P}(T^* = 0 | T = 1, z_k) = \alpha_0 (1 - p_k^*) / p_k \end{aligned}$$

Some algebra shows that

$$\pi\delta - (1 - \pi)(1 - \delta) = \frac{(p_k - \alpha_0)(1 - p_k - \alpha_1)}{1 - \alpha_0 - \alpha_1}$$

Hence, after simplifying and rearranging, it follows that

$$\begin{aligned} p_k(1 - p_k)(1 - p_k - \alpha_1)\mathbb{E}[y^r | T^* = 0, z_k] &= (1 - \alpha_1)(1 - p_k)\mathbb{E}[y^r | T = 0, z_k] - \alpha_1 p_k \mathbb{E}[y^r | T = 1, z_k] \\ p_k(1 - p_k)(p_k - \alpha_0)\mathbb{E}[y^r | T^* = 1, z_k] &= (1 - \alpha_0)p_k \mathbb{E}[y^r | T = 1, z_k] - \alpha_0(1 - p_k)\mathbb{E}[y^r | T = 0, z_k] \end{aligned}$$

**Step 3: Convert Conditional to Unconditional Moments** By iterated expectations and the assumption that  $\mathbb{E}[y^r | T, T^*, z] = \mathbb{E}[y^r | T^*, z]$

$$\begin{aligned} p_k &= \mathbb{E}[T | z = k] = \mathbb{E}[T \mathbf{1}(z = k)] / \mathbb{P}(z = k) \\ \mathbb{E}[y^r | T = 0, z = k] &= \mathbb{E}[y^r (1 - T) \mathbf{1}(z = k)] / [(1 - p_k)\mathbb{P}(z = k)] \\ \mathbb{E}[y^r | T = 1, z = k] &= \mathbb{E}[y^r T \mathbf{1}(z = k)] / [p_k \mathbb{P}(z = k)] \end{aligned}$$

**Step 4: Substitute Steps 2–3 into Step 1** After some algebra, we find that

$$\begin{aligned}\mathbb{E} [\mathbf{1}(z = k) \{T - (1 - \alpha_1)\} (y^2 - 2cy)] &< c^2 \mathbb{P}(z = k) p_k (1 - p_k) (1 - p_k - \alpha_1) \\ \mathbb{E} [\mathbf{1}(z = k) (\alpha_0 - T) \{y^2 - 2(\beta + c)y\}] &< (\beta + c)^2 \mathbb{P}(z = k) p_k (1 - p_k) (p_k - \alpha_0)\end{aligned}$$

**Step 5: Complete the Square**

$$\begin{aligned}\mathbb{E} [\mathbf{1}(z = k) \{1 - \alpha_1 - T\} (y - c)^2] &> c^2 \mathbb{P}(z = k) (1 - p_k - \alpha_1) [1 - p_k (1 - p_k)] \\ \mathbb{E} [\mathbf{1}(z = k) (T - \alpha_0) \{y - (\beta + c)\}^2] &> (\beta + c)^2 \mathbb{P}(z = k) (p_k - \alpha_0) [1 - p_k (1 - p_k)]\end{aligned}$$

## 15 April 21–23

### The Full Set of Moment Inequalities

Above we considered the special case in which  $\alpha_0 = 0$  so that first and second moments were sufficient to identify the model. We now derive the GMM-style moment conditions for the general case in which  $\alpha_0$  may not equal zero and we need to use third moments for identification.

Note that we have checked all of these equalities numerically in our simulation study and they are indeed correct!

**Notation and Identification** Define the following re-parameterization

$$\begin{aligned}\theta_1 &= \beta / (1 - \alpha_0 - \alpha_1) \\ \theta_2 &= \theta_1^2 [1 + (\alpha_0 - \alpha_1)] \\ \theta_3 &= \theta_1^3 [(1 - \alpha_0 - \alpha_1)^2 + 6\alpha_0 (1 - \alpha_1)]\end{aligned}$$

Using this notation, the covariance equations from above become

$$\begin{aligned}\text{Cov}(y, z) - \theta_1 \text{Cov}(T, z) &= 0 \\ \text{Cov}(y^2, z) - \theta_1 2 \text{Cov}(yT, z) + \theta_2 \text{Cov}(T, z) &= 0 \\ \text{Cov}(y^3, z) - \theta_1 3 \text{Cov}(y^2 T, z) + \theta_2 3 \text{Cov}(yT, z) - \theta_3 \text{Cov}(T, z) &= 0\end{aligned}$$

Note that it is trivial to prove that  $\theta_1, \theta_2$  and  $\theta_3$  are identified. To show that  $\beta, \alpha_0$  and  $\alpha_1$  are identified, write

$$\begin{aligned}\theta_2/\theta_1^2 &= 1 + (\alpha_0 - \alpha_1) \\ \theta_3/\theta_1^3 &= (1 - \alpha_0 - \alpha_1)^2 + 6\alpha_0(1 - \alpha_1)\end{aligned}$$

which we can always do provided that  $\theta_1 \neq 0$  which is equivalent to  $\beta \neq 0$ . Re-arranging the first equation allows us to solve for  $\alpha_0$  as a function of  $\alpha_1$ . Substituting this into the second gives us a quadratic in  $\alpha_1$  only. This should be identical to the quadratic we obtained in our original identification proof.

**First Moment Equalities** Expanding,

$$\begin{aligned}\text{Cov}(y, z) - \theta_1 \text{Cov}(T, z) &= 0 \\ \mathbb{E}[yz - \theta_1 Tz] - \mathbb{E}[z]\mathbb{E}[y - \theta_1 T] &= 0\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}[y - \theta_1 T] &= c + \beta \mathbb{E}[T^* - T/(1 - \alpha_0 - \alpha_1)] \\ &= c + \beta \left\{ \frac{\mathbb{E}[T] - \alpha_0}{1 - \alpha_0 - \alpha_1} - \frac{\mathbb{E}[T]}{1 - \alpha_0 - \alpha_1} \right\} \\ &= c - \alpha_0 \theta_1\end{aligned}$$

Thus,

$$\mathbb{E}[\{y - (c - \alpha_0 \theta_1) - \theta_1 T\} z] = 0$$

Now define

$$\begin{aligned}\kappa_1 &= c - \alpha_0 \theta_1 \\ u_1 &= y - \kappa_1 - \theta_1 T\end{aligned}$$

Using this notation, we obtain the unconditional moment equalities

$$\mathbb{E} \begin{bmatrix} u_1(\kappa_1, \theta_1) \\ u_1(\kappa_1, \theta_1)z \end{bmatrix} = \mathbf{0}$$

## Second Moment Equalities

$$\begin{aligned} \text{Cov}(y^2, z) - \theta_1 2\text{Cov}(yT, z) + \theta_2 \text{Cov}(T, z) &= 0 \\ \mathbb{E} [y^2 z - \theta_1 2yTz + \theta_2 Tz] - \mathbb{E} [z] \mathbb{E} [y^2 - \theta_1 2yT + \theta_2 T] &= 0 \end{aligned}$$

Now, using some lemmas from the notes in our sick-instruments paper,

$$\begin{aligned} \mathbb{E}[T\varepsilon] &= \text{Cov}(T, \varepsilon) \\ &= \text{Cov}(T^*, \varepsilon) + \text{Cov}(w, \varepsilon) \\ &= \text{Cov}(T^*, \varepsilon) - \text{Cov}(T^*, \varepsilon)(\alpha_0 + \alpha_1) \\ &= \mathbb{E}(T^* \varepsilon)(1 - \alpha_0 - \alpha_1) \end{aligned}$$

hence,

$$\begin{aligned} \mathbb{E}[yT] &= c\mathbb{E}[T] + \beta\mathbb{E}[TT^*] + \mathbb{E}[T, \varepsilon] \\ &= cp + \beta\mathbb{P}(T = 1, T^* = 1) + \mathbb{E}[T^* \varepsilon](1 - \alpha_0 - \alpha_1) \\ &= cp + \beta(1 - \alpha_1)p^* + \mathbb{E}[T^* \varepsilon](1 - \alpha_0 - \alpha_1) \end{aligned}$$

Combining this with

$$\mathbb{E}[y^2] = c^2 + \beta^2 p^* + \sigma_{\varepsilon\varepsilon} + 2c\beta p^* + 2\beta\mathbb{E}[T^* \varepsilon]$$

we find that

$$\mathbb{E} [y^2 - \theta_1 2yT + \theta_2 T] = \dots = c^2 + \sigma_{\varepsilon\varepsilon} + \alpha_0(\theta_2 - 2c\theta_1)$$

For the steps, see our whiteboard notes from 2017-04-21 17.06.32

Now, define

$$\begin{aligned} \kappa_2 &= c^2 + \sigma_{\varepsilon\varepsilon} + \alpha_0(\theta_2 - 2c\theta_1) \\ u_2 &= y^2 - \kappa_2 - \theta_1 2yT + \theta_2 T \end{aligned}$$

Using this notation, we obtain the unconditional moment equalities

$$\mathbb{E} \begin{bmatrix} u_2(\kappa_2, \theta_1, \theta_2) \\ u_2(\kappa_2, \theta_1, \theta_2)z \end{bmatrix} = \mathbf{0}$$

### Third Moment Equalities

$$\begin{aligned} \text{Cov}(y^3, z) - \theta_1 3\text{Cov}(y^2 T, z) + \theta_2 3\text{Cov}(y T, z) - \theta_3 \text{Cov}(T, z) &= 0 \\ \mathbb{E}[y^3 z - \theta_1 3y^2 T z + \theta_2 3y T z - \theta_3 T z] - \mathbb{E}[z] \mathbb{E}[y^3 - \theta_1 3y^2 T + \theta_2 3y T - \theta_3 T] &= 0 \end{aligned}$$

$$\mathbb{E}[y^3] = \dots = c^3 + \beta p^* (3c^2 + 3c\beta + \beta^2) + 3\beta \mathbb{E}[\varepsilon T^*] (2c + \beta) + 3c\sigma_{\varepsilon\varepsilon} + 3\beta \mathbb{E}[\varepsilon^2 T^*] + \mathbb{E}[\varepsilon^3]$$

For derivations of the following, see our whiteboard notes from April 21st and 22nd, 2017

$$\mathbb{E}[y^2 T] = \dots = c^2 p + \beta(1 - \alpha_1)p^* + \mathbb{E}[T\varepsilon^2] + 2c\beta(1 - \alpha_1)p^* + 2c\mathbb{E}[T\varepsilon] + 2\beta \mathbb{E}[TT^*\varepsilon]$$

and

$$\begin{aligned} \mathbb{E}[T\varepsilon^2] &= \dots = \alpha_0 \sigma_{\varepsilon\varepsilon} + (1 - \alpha_0 - \alpha_1) \mathbb{E}[\varepsilon^2 T^*] \\ \mathbb{E}[TT^*\varepsilon] &= \dots = (1 - \alpha_1) \mathbb{E}[T^*\varepsilon] \end{aligned}$$

We then need to calculate  $\mathbb{E}[y^3 - \theta_1 3y^2 T + \theta_2 3y T - \theta_3 T]$ . Using the preceding expressions, we can show after some algebra that the  $\mathbb{E}[T^*\varepsilon]$  and  $\mathbb{E}[T^*\varepsilon^2]$  terms drop out, hence

$$\begin{aligned} \mathbb{E}[y^3 - \theta_1 3y^2 T + \theta_2 3y T - \theta_3 T] &= \{c^3 + \beta p^* (3c^2 + 3c\beta + \beta^2)\} + 3c\sigma_{\varepsilon\varepsilon} + \mathbb{E}[\varepsilon^3] - \theta_1 3\alpha_0 \sigma_{\varepsilon\varepsilon} - \theta_3 p \\ &\quad - \theta_1 3 \{c^2 p + \beta^2(1 - \alpha_1)p^* + 2c\beta(1 - \alpha_1)p^*\} + \theta_2 3 \{cp + \beta(1 - \alpha_1)p^*\} \end{aligned}$$

After *even more algebra* we can show that this expression depends neither on  $p$  nor on  $p^*$ :

$$\begin{aligned} \mathbb{E}[y^3 - \theta_1 3y^2 T + \theta_2 3y T - \theta_3 T] &= c^3 + 3(c - \theta_1 \alpha_0) \sigma_{\varepsilon\varepsilon} + \mathbb{E}[\varepsilon^3] - \alpha_0 \theta_3 - 3c\alpha_0 \left[ \theta_1 (c + \beta) - 2\theta_2 \frac{1 - \alpha_1}{1 + \alpha_0 - \alpha_1} \right] \\ &= c^3 + 3(c - \theta_1 \alpha_0) \sigma_{\varepsilon\varepsilon} + \mathbb{E}[\varepsilon^3] - \alpha_0 \theta_3 - 3c\alpha_0 [\theta_1 (c + \beta) - 2\theta_1^2 (1 - \alpha_1)] \end{aligned}$$

Now let

$$\begin{aligned} \kappa_3 &= c^3 + 3(c - \theta_1 \alpha_0) \sigma_{\varepsilon\varepsilon} + \mathbb{E}[\varepsilon^3] - \alpha_0 \theta_3 - 3c\alpha_0 [\theta_1 (c + \beta) - 2\theta_1^2 (1 - \alpha_1)] \\ u_3 &= y^3 - \kappa_3 - \theta_1 3y^2 T + \theta_2 3y T - \theta_3 T \end{aligned}$$

we obtain the following moment equalities

$$\mathbb{E} \begin{bmatrix} u_3(\kappa_3, \theta_1, \theta_2, \theta_3) \\ u_3(\kappa_3, \theta_1, \theta_2, \theta_3)z \end{bmatrix} = \mathbf{0}$$

**Putting Everything Together** Define

$$\begin{aligned} \boldsymbol{\kappa} &= (\kappa_1, \kappa_2, \kappa_3)' \\ \boldsymbol{\theta} &= (\theta_1, \theta_2, \theta_3)' \\ \mathbf{u}(\boldsymbol{\kappa}, \boldsymbol{\theta}) &= \begin{bmatrix} u_1(\kappa_1, \theta_1) & u_2(\kappa_2, \theta_2) & u_3(\kappa_3, \theta_3) \end{bmatrix}' \end{aligned}$$

where

$$\begin{aligned} \kappa_1 &= c - \alpha_0 \theta_1 \\ \kappa_2 &= c^2 + \sigma_{\varepsilon\varepsilon} + \alpha_0(\theta_2 - 2c\theta_1) \\ \kappa_3 &= c^3 + 3(c - \theta_1\alpha_0)\sigma_{\varepsilon\varepsilon} + \mathbb{E}[\varepsilon^3] - \alpha_0\theta_3 - 3c\alpha_0[\theta_1(c + \beta) - 2\theta_1^2(1 - \alpha_1)] \end{aligned}$$

and

$$\begin{aligned} u_1(\kappa_1, \theta_1) &= y - \kappa_1 - \theta_1 T \\ u_2(\kappa_2, \theta_1, \theta_2) &= y^2 - \kappa_2 - \theta_1 2yT + \theta_2 T \\ u_3(\kappa_3, \theta_1, \theta_2, \theta_3) &= y^3 - \kappa_3 - \theta_1 3y^2 T + \theta_2 3yT - \theta_3 T \end{aligned}$$

Then the full system of moment equalities is given by

$$\mathbb{E} \begin{bmatrix} \mathbf{u}(\boldsymbol{\kappa}, \boldsymbol{\theta}) \\ \mathbf{u}(\boldsymbol{\kappa}, \boldsymbol{\theta})z \end{bmatrix} = \mathbf{0}$$

where

$$\begin{aligned} \theta_1 &= \beta / (1 - \alpha_0 - \alpha_1) \\ \theta_2 &= \theta_1^2 [1 + (\alpha_0 - \alpha_1)] \\ \theta_3 &= \theta_1^3 [(1 - \alpha_0 - \alpha_1)^2 + 6\alpha_0(1 - \alpha_1)] \end{aligned}$$

**What about exogenous covariates?** At one extreme, we could simply carry out the above *conditional* on  $\mathbf{x}$ . This would entail treating  $\boldsymbol{\kappa}$  and  $\boldsymbol{\theta}$  as functions of  $\mathbf{x}$  such that  $\alpha_0, \alpha_1$  and  $\beta$  would be allowed to depend on  $\mathbf{x}$ . To implement this, one would need to

estimate a number of conditional mean functions:  $\mathbb{E}[y|\mathbf{x}]$ ,  $\mathbb{E}[T|\mathbf{x}]$ ,  $\mathbb{E}[yT|\mathbf{x}]$ ,  $\mathbb{E}[y^3|\mathbf{x}]$ , and  $\mathbb{E}[y^2T|\mathbf{x}]$ . It would appear that this leads to a standard two-step estimation problem although clearly you'd need a lot of data to have any chance! Another idea would be to impose some restrictions on the way in which  $\mathbf{x}$  affects  $\alpha_0$  etc, leading to a semi-parametric estimator. At the other extreme, we could impose the assumption that  $\mathbf{x}$  affects  $y$  linearly and  $\alpha_0, \alpha_1$  do not vary with  $\mathbf{x}$ :

$$y = c + \beta T^* + \mathbf{x}'\boldsymbol{\gamma} + \varepsilon$$

Under this assumption the result of Frazis & Loewenstein (2003) holds – the IV estimator would recover  $\kappa_1, \theta_1, \boldsymbol{\gamma}$ . Since  $(y - \mathbf{x}'\boldsymbol{\gamma}) = c + \beta T^* + \varepsilon$ , all of the preceding moment conditions should still hold only with  $(y - \mathbf{x}'\boldsymbol{\gamma})$  in place of  $y$ . To identify  $\boldsymbol{\gamma}$  we add an extra moment condition in which  $\mathbf{x}$  multiplies the  $u_1$  error term, re-defined to subtract  $\mathbf{x}'\boldsymbol{\gamma}$ .

## Correcting the Second Moment Inequalities

In the derivation from above I mistakenly claimed that since  $\varepsilon$  is mean zero, it is sufficient to work with the conditional means of  $\varepsilon^2$  given  $T^*$  and  $z$ . This is not correct since the *conditional* means of  $\varepsilon$  are not zero. The bounds from above are not wrong, but they can be improved by looking at the conditional variance of  $\varepsilon$  rather than simply  $\mathbb{E}[\varepsilon^2]$ . Steps 2–3 from above are still usable, but we need to change step 1.

Our inequalities come from the fact that

$$\text{Var}(\varepsilon|T^*, z) > 0 \iff \mathbb{E}[\varepsilon^2|T^*, z] > \{\mathbb{E}[\varepsilon|T^*, z]\}^2$$

We've already done the calculations for  $\varepsilon^2|T^*, z$ . Now we simply need to do the same for  $\varepsilon|T^*, z$ , namely

$$\begin{aligned}\mathbb{E}[\varepsilon|T^* = 0, z = k] &= \mathbb{E}[y|T^* = 0, z = k] - c \\ \mathbb{E}[\varepsilon|T^* = 1, z = k] &= \mathbb{E}[y|T^* = 1, z = k] - (c + \beta)\end{aligned}$$

Now we relate  $\mathbb{E}[y|T^*, z]$  to  $\mathbb{E}[y|T, z]$  using the calculations from above:

$$\begin{aligned}\mathbb{E}[y|T^* = 0, z_k] &= \frac{(1 - \alpha_1)}{p_k(1 - p_k - \alpha_1)}\mathbb{E}[y|T = 0, z_k] - \frac{\alpha_1}{(1 - p_k)(1 - p_k - \alpha_1)}\mathbb{E}[y|T = 1, z_k] \\ \mathbb{E}[y|T^* = 1, z_k] &= \frac{(1 - \alpha_0)}{(1 - p_k)(p_k - \alpha_0)}\mathbb{E}[y|T = 1, z_k] - \frac{\alpha_0}{p_k(p_k - \alpha_0)}\mathbb{E}[y|T = 0, z_k]\end{aligned}$$



After lots of algebra (see our handwritten notes), we obtain:

$$\begin{aligned} p_k(1 - p_k)(1 - p_k - \alpha_1) \{ (1 - \alpha_1)(1 - p_k)\mathbb{E}[y^2|T = 0, z_k] - \alpha_1 p_k \mathbb{E}[y^2|T = 1, z_k] \} \\ > \{ (1 - \alpha_1)(1 - p_k)\mathbb{E}[y|T = 0, z_k] - \alpha_1 p_k \mathbb{E}[y|T = 1, z_k] \}^2 \end{aligned}$$

and

$$\begin{aligned} p_k(1 - p_k)(p_k - \alpha_0) \{ (1 - \alpha_0)p_k \mathbb{E}[y^2|T = 1, z_k] - \alpha_0(1 - p_k)\mathbb{E}[y^2|T = 0, z_k] \} \\ > \{ (1 - \alpha_0)p_k \mathbb{E}[y|T = 1, z_k] - \alpha_0(1 - p_k)\mathbb{E}[y|T = 0, z_k] \}^2 \end{aligned}$$

Finally, we need to convert these into expressions involving unconditional moments via

$$\begin{aligned} p_k &= \mathbb{E}[T|z = k] = \mathbb{E}[T\mathbf{1}(z = k)] / \mathbb{P}(z = k) \\ \mathbb{E}[y^r|T = 0, z = k] &= \mathbb{E}[y^r(1 - T)\mathbf{1}(z = k)] / [(1 - p_k)\mathbb{P}(z = k)] \\ \mathbb{E}[y^r|T = 1, z = k] &= \mathbb{E}[y^r T \mathbf{1}(z = k)] / [p_k \mathbb{P}(z = k)] \end{aligned}$$

These appear to match our expressions from the sick-instruments paper.

## Estimated Parameters in Andrews & Soares

In our simple example from above applying the Andrews and Soares (2010) generalized moment selection procedure to the case where  $\alpha_0 = 0$ , we assumed for simplicity that  $c$  and  $\sigma_{\varepsilon\varepsilon}$  were known. Without  $\alpha_0 = 0$ , the analogous assumption would be that  $\kappa$  is known. In practice, however, this parameter is estimated from the data. Accounting for this fact requires a change in the procedure for calculating the variance matrix  $\Sigma$  of the moment conditions under the null hypothesis. We know explain how this works.

## 16 May 5th, 2017 – Todo for Revised Paper

1. New notation to accomodate covariates etc. as in Mahajan and Lewbel
2. Move Mahajan stuff into appendix; possibly convert back to his notation later – I have handwritten notes on this.
3. Identification results: identification from higher moments, lack of identification from conditional means alone
4. Explain briefly how to do fully non-parametric estimation, specialize to some simple cases, e.g. linear model
5. Explain about inference versus estimation.

6. Possibly show that Mahajan, Lewbel, etc. also suffer from a weak identification problem. This might be helpful for selling the paper. Would need to think about what inequalities to use in their case. Presumably we could still use the “weak” bounds, but there might be others we could exploit.
7. Testing  $\beta = 0$  should be very easy: don’t need Andrews & Soares at all. I think we can either just use the Stock & Wright GMM-AR or even a plain old-fashioned GMM inference for testing a linear restriction.
8. Canay et al. projection inference for  $\beta$ ? Need to figure out if this actually gives a speed-up in our case. If so, could be very useful for simulations. Would be nice to show that we have power to reject the probability limit of the IV and reduced form. Is it easier to handle the strongly-identified parameters using their method?
9. Could also talk briefly about the continuum of moment conditions idea and how one obtains identification from quantiles. There’s also a question of using more inequalities by imposing the independence assumption for the measurement error. Could also proceed without our higher moment assumptions and just get inference for the identified set.
10. Is there any way to handle treatment effect heterogeneity? Should be ok if it’s modeled heterogeneity but maybe there are some other special cases we can handle? What about that quantile treatment effect idea?
11. Simulations: want to show that we can actually learn something useful. Things we want to show are that we have more power than just the RF for testing  $\beta = 0$ , that we have power to reject the RF and IV plims in certain cases.
12. Need to show at least in simulations that our confidence regions aren’t insane. Would be good if we could formally argue that they must be convex and connected, for example. Maybe the key is to show that our inequalities are always well-behaved and that adding the equalities can’t make the problem suddenly become badly behaved.
13. Andrews & Soares with estimates of the strongly identified parameters to get a joint region for  $\alpha_0, \alpha_1$ , and  $\beta$ . But this is slow, so we probably can’t do simulations although it would be interesting in the empirical examples.
14. Try to get the projection inference for  $\beta$  working
15. Empirical example or examples: Oreopolous? Maybe the Heckman JPTA dataset? Petra suggested emailing someone about this but I forget who it was...

## 17 Lewbel (2007)

**Notation** Observe  $Y, Z$  and  $T$  where  $T$  is a proxy for  $T^*$  and

$$\begin{aligned} h^*(X, T^*) &= E(Y|X, T^*) \\ T^* &= \text{unobserved binary regressor} \\ X &= \text{vector of covariates} \\ Y &= \text{outcome} \end{aligned}$$

Since  $T^*$  is binary, without loss of generality,

$$\begin{aligned} h^*(X, T^*) &= h_0^*(X) + \tau^*(x)T^* \\ h_0^*(X) &= h^*(X, 0) \\ \tau^*(x) &= h^*(x, 1) - h^*(x, 0) \end{aligned}$$

Goal is to estimate  $\tau^*(x)$ . Below we will partition  $X$  according to  $X = (V, Z)$  where  $V$  is an “instrument-like variable” and  $Z$  is the set of remaining covariates.

**Assumption A1** There exists  $E(Y|X, T^*, T) = E(Y|X, T^*)$ . Equivalently,  $Y$  is mean-independent of  $T - T^*$ , conditional on  $X, T^*$  – mis-classification does not affect true expected outcome. This rules out placebo effects, etc.

### More Notation

$$\begin{aligned} r^*(x) &= E(T^*|X = x) = P(T^* = 1|X = x) \\ b_0(x) &= P(T = 1|T^* = 0, X = x) \\ b_1(x) &= P(T = 0|T^* = 1, X = x) \\ r(x) &= E(T|X = x) \\ \tau(x) &= h(x, 1) - h(x, 0) \\ h(x, t) &= E(Y|X = x, T = t) \end{aligned}$$

**Assumption A2** There exist  $b_0(x) + b_1(x) < 1$  and  $0 < r^*(x) < 1$  for all  $x$  in the support of  $X$ .

**Theorem 1** Under Assumption A1 there exists a function  $m(x)$  such that  $|m(x)| \leq 1$  and  $\tau(x) = \tau^*(x)m(x)$ . If we add Assumption A2 then  $m(x) > 0$  as well.

- Analogous to attenuation bias under classical measurement error: A1 implies that  $\tau(x)$  is a lower bound for the magnitude of  $\tau^*(x)$ . Adding A2, sign of  $\tau(x)$  agrees with that of  $\tau^*(x)$ .

- Assumptions A1 and A2 imply that  $\tau^*(x) = 0$  iff  $\tau(x) = 0$ . Thus, if we only want to test  $\tau^*(x) = 0$  we can simply ignore mis-classification.
- As shown below:

$$\begin{aligned}
m(x) &= P(T^* = 1|T = 1, X = x) - P(T^* = 1|T = 0, X = x) \implies |m(x)| < 1 \\
m(x) &= M[b_0(x), b_1(x), r(x)] = \frac{1}{1 - b_1(x) - b_0(x)} \left\{ 1 - \frac{[1 - b_1(x)] b_0(x)}{r(x)} - \frac{[1 - b_0(x)] b_1(x)}{1 - r(x)} \right\} \\
m(x) &= \frac{[1 - r^*(x)] r^*(x)}{[1 - r(x)] r(x)} [1 - b_0(x) - b_1(x)] \implies m(x) > 0 \text{ if } 1 - b_0(x) + b_1(x) > 0
\end{aligned}$$

**Proof of Theorem 1.** Let  $p_t(X) = P(T^* = 1|X, T = t)$ . By A1 we have

$$E(Y|X, T^*, T) = E(Y|X, T^*) = h_0^*(X) + \tau^*(X)T^*$$

combining this with iterated expectations,

$$\begin{aligned}
E(Y|X, T = t) &= E_{T^*|X, T=t} [E(Y|X, T = t, T^*)] = E_{T^*|X, T=t} [h_0^*(X) + \tau^*(X)T^*] \\
&= p_t(X) [h_0^*(X) + \tau^*(X)] + [1 - p_t(X)] h_0^*(X) \\
&= h_0^*(X) + p_t(X)\tau^*(X)
\end{aligned}$$

Taking the difference of the preceding expression evaluated at  $T = 1$  and  $T = 0$

$$E(Y|X, T = 1) - E(Y|X, T = 0) = [p_1(X) - p_0(X)] \tau^*(X)$$

whereas we defined

$$\tau(X) = h(X, 1) - h(X, 0) = E(Y|X, T = 1) - E(Y|X, T = 0)$$

Combining these two equations, we see that

$$\tau(X) = [p_1(X) - p_0(X)] \tau^*(X)$$

so that the function  $m(x)$  defined in Theorem 1 equals  $p_1(x) - p_0(x)$ . Since  $m$  is a difference of probabilities, it follows immediately that  $-1 \leq m(x) \leq 1$ . Now, by Bayes' Rule,

$$\begin{aligned}
p_0(x) &= P(T^* = 1|X = x, T = 0) = \frac{P(T = 0|X = x, T^* = 1)P(T^* = 1|X = x)}{P(T = 0|X = x)} = \frac{b_1(x)r^*(x)}{1 - r(x)} \\
p_1(x) &= P(T^* = 1|X = x, T = 1) = \frac{P(T = 1|X = x, T^* = 1)P(T^* = 1|X = x)}{P(T = 1|X = x)} = \frac{[1 - b_1(x)]r^*(x)}{r(x)}
\end{aligned}$$

and by iterated expectations,

$$\begin{aligned} r(x) &= E(T|X = x) = E(T|X = x, T^* = 1)P(T^* = 1|X = x) + E(T|X = x, T^* = 0)P(T^* = 0|X = x) \\ &= [1 - b_1(x)]r^*(x) + b_0(x)[1 - r^*(x)] = b_0(x) + r^*(x)[1 - b_0(x) - b_1(x)] \end{aligned}$$

Using this expression, if  $b_0(x) + b_1(x) = 1$  then  $r(x) = b_0$ . If instead  $b_0(x) + b_1(x) \neq 1$ , then we can divide through by  $1 - b_0(x) - b_1(x)$  to solve for  $r^*(x)$  and  $1 - r^*(x)$  as follows:

$$\begin{aligned} r^*(x) &= \frac{r(x) - b_0(x)}{1 - b_0(x) - b_1(x)} \\ 1 - r^*(x) &= \frac{[1 - b_0(x) - b_1(x)] - [r(x) - b_0(x)]}{1 - b_0(x) - b_1(x)} = \frac{1 - b_1(x) - r(x)}{1 - b_0(x) - b_1(x)} \end{aligned}$$

Using the expressions we have just derived for  $p_0, p_1, r^*$  and  $1 - r^*$  and suppressing the dependence on  $x$  for simplicity, it follows that

$$\begin{aligned} m = p_1 - p_0 &= \frac{(1 - b_1)r^*}{r} - \frac{b_1r^*}{1 - r} = \frac{(1 - r)(1 - b_1)r^* - rb_1r^*}{r(1 - r)} \\ &= \frac{r^*(1 - r - b_1 + rb_1 - rb_1)}{r(1 - r)} = \frac{r^*(1 - r - b_1)}{r(1 - r)} \end{aligned}$$

Rearranging,

$$(1 - r)rm = r^*(1 - r - b_1)$$

and combining this with  $1 - r^* = (1 - b_1 - r)/(1 - b_0 - b_1)$ , we have

$$(1 - r)rm = (1 - r^*)r^*(1 - b_0 - b_1)$$

Both  $r$  and  $r^*$  are strictly between zero and one. Thus, if  $b_0 + b_1 = 1$  we have  $m = 0$ . If instead  $b_0 + b_1 < 1$  we have  $m > 0$  and if  $b_0 + b_1 > 1$  then  $m < 0$ . Finally,

$$m = \frac{r^*(1 - r - b_1)}{r(1 - r)} = \frac{r - b_0}{1 - b_0 - b_1} \left[ \frac{1 - r - b_1}{r(1 - r)} \right]$$

and since

$$\begin{aligned} (r - b_0)(1 - r - b_1) &= r(1 - r) - b_0(1 - r) - b_1r + b_1b_0 \\ &= r(1 - r) - b_0(1 - r) - b_1r + b_1b_0 + (b_1b_0r - b_1b_0r) \\ &= r(1 - r) - b_0(1 - r) - b_1r + b_1b_0(1 - r) + b_1b_0r \\ &= r(1 - r) - b_0(1 - r) - b_1r(1 - b_0) + b_1b_0(1 - r) \\ &= r(1 - r) + b_0(1 - r)(b_1 - 1) - b_1r(1 - b_0) \\ &= r(1 - r) - b_0(1 - r)(1 - b_1) - b_1r(1 - b_0) \end{aligned}$$

we find that

$$\begin{aligned} m = M(b_0, b_1, r) &= \frac{1}{1 - b_0 - b_1} \left[ \frac{r(1 - r) - (1 - b_1)b_0(1 - r) - (1 - b_0)b_1r}{r(1 - r)} \right] \\ &= \frac{1}{1 - b_0 - b_1} \left[ 1 - \frac{(1 - b_1)b_0}{r} - \frac{(1 - b_0)b_1}{1 - r} \right]. \end{aligned}$$

□

**Assumption A3** Assume that  $r(x)$  and  $\tau(x)$  are identified. Note that this only requires that we can consistently estimate conditional expectations of observable quantities.

**Assumption A4** Suppose that we have partitioned  $X$  into two subvectors:  $X = (V, Z)$ . We assume that for each  $z$  in the support of  $Z$  there exists a subset  $\Omega_z$  of the support of  $V$  such that:

- (i) for all  $v, v' \in \Omega_z$ ,  $b_0(v, z) = b_0(v', z)$ ,  $b_1(v, z) = b_1(v', z)$ , and  $\tau^*(v, z) = \tau^*(v', z)$
- (ii) for all  $v, v' \in \Omega_z$  such that  $v \neq v'$ ,  $r^*(v, z) \neq r^*(v', z)$ .

The basic idea here is that  $V$  affects the probability of being treated  $r^*$  but not the treatment effect  $\tau^*$  after we have conditioned on  $Z$ . As sufficient condition for the restriction on  $\tau^*$  to hold is  $E(Y|Z = z, V = v, T^* = t) = s_1(z, t) + s_2(z, v)$  for some functions  $s_1$  and  $s_2$ .

**Some More Notation** Let  $b_0(z), b_1(z), \tau^*(z)$  denote  $b_0(v, z), b_1(v, z), \tau^*(v, z)$  for  $v \in \Omega_z$  since, under A4, these do not vary with  $v$ .

**Assumption A5** Each set  $\Omega_z$  from A4 contains three elements  $v_k, k = 0, 1, 2$  such that

$$\left[ \frac{\tau(v_0, z)}{r(v_1, z)} - \frac{\tau(v_1, z)}{r(v_0, z)} \right] \left[ \frac{\tau(v_0, z)}{1 - r(v_2, z)} - \frac{\tau(v_2, z)}{1 - r(v_0, z)} \right] \neq \left[ \frac{\tau(v_0, z)}{r(v_2, z)} - \frac{\tau(v_2, z)}{r(v_0, z)} \right] \left[ \frac{\tau(v_0, z)}{1 - r(v_1, z)} - \frac{\tau(v_1, z)}{1 - r(v_0, z)} \right]$$

- The key requirement is that  $V$  takes on at least three values.
- Assumption A5 depends only on observables, so we can test it.
- An equivalent way to state A5 is:  $\tau^*(z) \neq 0$ ,  $b_0(z) + b_1(z) \neq 1$  and an inequality involving  $r$  and  $r^*$  which is explained below.
- The triplets  $(v_0, v_1, v_2)$  are allowed to depend on  $z$ .

**Theorem 2** Under A1–A5, the mis-classification probabilities  $b_0(x), b_1(x)$  are identified as is the probability of treatment  $r^*(x)$  and treatment effect  $\tau^*(x)$ . If we replace  $b_0(x) + b_1(x) < 1$  from A2 with  $b_0(x) + b_1(x) \neq 1$ , then  $\tau^*(x)$  is identified up to sign.

- For simplicity, suppress dependence on  $z$ . By Theorem 1 and A4,

$$\tau(v_k)M[b_0, b_1, r(v_k)] = \tau(v_0)M[b_0, b_1, r(v_0)]$$

- Again suppressing dependence on  $z$ , recall from above that

$$M[b_0, b_1, r(v_k)] = \frac{1}{1 - b_1 - b_0} \left\{ 1 - \frac{(1 - b_1)b_0}{r(v_k)} - \frac{(1 - b_0)b_1}{1 - r(v_k)} \right\}$$

- Evaluating the two preceding expressions at  $k = 1$  and  $k = 2$  gives two equations relating the identified functions  $r$  and  $\tau$  to the unknown mis-classification probabilities  $b_0$  and  $b_1$ .
- Since each equation involves  $v_0$ , we need  $V$  to take on at least three values to get as many equations as unknowns (two).
- The proof of Theorem 2 shows that the two equations admit a unique solution, so that  $b_0$  and  $b_1$  are identified.
- Given knowledge of  $b_0$  and  $b_1$ , Theorem 1 allows us to solve for  $r^*$  and  $\tau^*$ .
- If  $V$  only took on two values, we could not identify  $b_0$  and  $b_1$  without further restrictions. However if either  $b_0$  or  $b_1$  were known, e.g. known to be zero as in a one-sided mis-classification setting, then we could identify the model using only a binary  $V$ .

**Proof of Theorem 2.** For a fixed value of the covariates  $z$ , Assumption A4 ensures that there is a subset  $\Omega_z$  of the support of  $V$  – a subset that may depend on  $z$  – such that for all  $v, v' \in \Omega_z$  we have  $b_0(v, z) = b_0(v', z)$ ,  $b_1(v, z) = b_1(v', z)$  and  $\tau^*(v, z) = \tau^*(v', z)$ . Assumption A5 ensures that  $\Omega_z$  contains at least three values:  $v_0, v_1$  and  $v_2$ . Suppress dependence on the covariates  $z$ : let  $r_k = r(v_k)$  and  $\tau_k = \tau(v_k)$  for  $k = 0, 1, 2$ . Since  $b_0$  and  $b_1$  do not depend on  $v$  for  $v \in \Omega_z$ , there is no subscript on these quantities. By Theorem 1,  $\tau(v, z) = \tau^*(v, z)m(v, z)$ . Again suppressing dependence on  $z$  and using the expression for  $m$  derived in the proof of Theorem 1,  $\tau_k = M(b_0, b_1, r_k)\tau_k^*$  for all  $k = 0, 1, 2$ . But by Assumption A4,  $\tau_0^* = \tau_1^* = \tau_2^*$ , which yields

$$\frac{\tau_k}{M(b_0, b_1, r_k)} = \frac{\tau_\ell}{M(b_0, b_1, r_\ell)}$$

for any  $k, \ell$ . Rearranging,

$$M(b_0, b_1, r_k)\tau_\ell = M(b_0, b_1, r_\ell)\tau_k$$

For  $k \neq \ell$  this yields a nontrivial equation relating  $b_0$  and  $b_1$  to observables:  $\tau_k, r_k$  and  $\tau_\ell, r_\ell$ . In

particular, take  $\ell = 0$  and  $k = 1, 2$ . We obtain,

$$\begin{aligned}
0 &= M(b_0, b_1, r_k)\tau_0 - M(b_0, b_1, r_0)\tau_k \\
0 &= \frac{1}{1 - b_0 - b_1} \left[ 1 - \frac{(1 - b_1)b_0}{r_k} - \frac{(1 - b_0)b_1}{1 - r_k} \right] \tau_0 - \frac{1}{1 - b_0 - b_1} \left[ 1 - \frac{(1 - b_1)b_0}{r_0} - \frac{(1 - b_0)b_1}{1 - r_0} \right] \tau_k \\
0 &= \left[ 1 - \frac{(1 - b_1)b_0}{r_k} - \frac{(1 - b_0)b_1}{1 - r_k} \right] \tau_0 - \left[ 1 - \frac{(1 - b_1)b_0}{r_0} - \frac{(1 - b_0)b_1}{1 - r_0} \right] \tau_k \\
0 &= (1 - b_1)b_0 \left( \frac{\tau_k}{r_0} - \frac{\tau_0}{r_k} \right) + (1 - b_0)b_1 \left( \frac{\tau_k}{1 - r_0} - \frac{\tau_0}{1 - r_k} \right) + (\tau_0 - \tau_k) \\
0 &= (1 - b_1)b_0 \left( \frac{\tau_0}{r_k} - \frac{\tau_k}{r_0} \right) + (1 - b_0)b_1 \left( \frac{\tau_0}{1 - r_k} - \frac{\tau_k}{1 - r_0} \right) + (\tau_k - \tau_0)
\end{aligned}$$

This is an equation of the form

$$0 = B_0 w_{0k} + B_1 w_{1k} + w_{2k}$$

where the unknowns  $B_0, B_1$  are defined as

$$B_0 = b_0(1 - b_1)$$

$$B_1 = b_1(1 - b_0)$$

and the observable constants  $w_{0k}, w_{1k}, w_{2k}$  are

$$w_{0k} =$$

$$w_{1k} = \frac{\tau_0}{1 - r_k} - \frac{\tau_k}{1 - r_0}$$

$$w_{2k} = \tau_k - \tau_0$$

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.  $\square$

## 18 Notes on Mahajan (2006)

It should be easy to adapt the proof for Lewbel (2007) above to get the result from Mahajan (2006). We just need to replace Lewbel's assumption that  $\tau^*(Z, V)$  does not depend on  $V$  with the stronger assumption that  $\tau^*(Z, V)$  and  $h_0^*(Z, V)$  do not depend on  $V$ . Then we only need a binary  $V$ . Does Mahajan's assumption imply that  $\tau_k = \tau_\ell$ ? (Rather than simply that  $\tau_k^* = \tau_\ell^*$ ?)

Mahajan (2006) considers regression models of the form

$$E[y - g(x^*, z)] = 0 \quad (25)$$

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.

## 18.1 The Case of Exogenous $x^*$

The first is based on the restriction

$$E[y - g(x^*, z) \mid x^*, x, z, v] = 0 \quad (26)$$

## 18.2 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

$$y = g^*(x^*, z) + \varepsilon \quad (27)$$

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

$$E[\varepsilon \mid z, v] = 0 \quad (28)$$

and Equation 26 is replaced by

$$E[y \mid x^*, x, z, v] = E[y \mid x^*, z] \quad (29)$$

Unfortunately, Mahajan’s proof is incorrect and the model in Equation 27 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 27–29 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>3</sup>

Without covariates we can write

$$y = \alpha + \beta x^* + \varepsilon \quad (30)$$

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

$$m_{jk} = E[\varepsilon|x^* = j, v = k] \quad (31)$$

## 19 Notation and Results for New Draft – May 8, 2017

Additively separable model

$$y = m(T^*, \mathbf{x}) + \varepsilon$$

where  $\varepsilon$  is a mean-zero error term,  $T^*$  is an endogenous binary regressor of interest and  $\mathbf{x}$  is a vector of exogenous controls. Since  $T^*$  is binary, we can re-write this as linear in  $T^*$  conditional on  $\mathbf{x}$

$$\begin{aligned} y &= c(\mathbf{x}) + \beta(\mathbf{x})T^* + \varepsilon \\ \beta(\mathbf{x}) &= m(1, \mathbf{x}) - m(0, \mathbf{x}) \\ c(\mathbf{x}) &= m(0, \mathbf{x}) \end{aligned}$$

Goal is to use an instrumental variable  $z$  to identify  $\beta(\mathbf{x})$  when we observe not  $T^*$  but a mis-measured binary surrogate  $T$ . Define

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

Identification will only rely on two values for  $z$  so throughout the remainder of the paper we assume that  $z$  is binary and takes on values 0 and 1.

### Assumption 19.1.

- (i)  $y = c(\mathbf{x}) + \beta(\mathbf{x})T^* + \varepsilon$  where  $\mathbb{E}[\varepsilon] = 0$ .
- (ii)  $\alpha_0(\mathbf{x}, z) = \alpha_1(\mathbf{x}), \alpha_1(\mathbf{x}, z) = \alpha_1(\mathbf{x})$
- (iii)  $\mathbb{E}[\varepsilon|\mathbf{x}, z, T^*, T] = \mathbb{E}[\varepsilon|\mathbf{x}, z, T^*]$
- (iv)  $\mathbb{E}[T^*|\mathbf{x}, z = 0] \neq \mathbb{E}[T^*|\mathbf{x}, z = 1]$

---

<sup>3</sup>Because the covariates are held fixed throughout the proof of Mahajan's Theorem 1, there is no loss of generality.

(v)  $\alpha_0(\mathbf{x}) + \alpha_1(\mathbf{x}) \neq 1$  ( $T$  is relevant for  $T^*$ )

(vi)  $\mathbb{E}[\varepsilon|\mathbf{x}, z] = 0$

**Assumption 19.2.**

(i)  $\mathbb{E}[\varepsilon^2|\mathbf{x}, z, T^*, T] = \mathbb{E}[\varepsilon^2|\mathbf{x}, z, T^*]$

(ii)  $\mathbb{E}[\varepsilon^2|\mathbf{x}, z] = \mathbb{E}[\varepsilon^2|\mathbf{x}]$

**Assumption 19.3.**

(i)  $\mathbb{E}[\varepsilon^3|\mathbf{x}, z, T^*, T] = \mathbb{E}[\varepsilon^3|\mathbf{x}, z, T^*]$

(ii)  $\mathbb{E}[\varepsilon^3|\mathbf{x}, z] = \mathbb{E}[\varepsilon^3|\mathbf{x}]$

**Assumption 19.4.**  $\alpha_0(\mathbf{x}) + \alpha_1(\mathbf{x}) < 1$  ( $T$  is positively correlated with  $T^*$ )

**Notation**

$$\theta_1(\mathbf{x}) = \beta(\mathbf{x}) [1 - \{\alpha_0(\mathbf{x}) + \alpha_1(\mathbf{x})\}]^{-1} \quad (32)$$

$$\theta_2(\mathbf{x}) = [\theta_1(\mathbf{x})]^2 [1 + \{\alpha_0(\mathbf{x}) - \alpha_1(\mathbf{x})\}] \quad (33)$$

$$\theta_3(\mathbf{x}) = [\theta_1(\mathbf{x})]^3 \left[ (1 - \{\alpha_0(\mathbf{x}) + \alpha_1(\mathbf{x})\})^2 + 6\alpha_0(\mathbf{x}) \{1 - \alpha_1(\mathbf{x})\} \right] \quad (34)$$

$$\text{Cov}(y, z|\mathbf{x}) - \text{Cov}(T, z|\mathbf{x})\theta_1(\mathbf{x}) = 0$$

$$\text{Cov}(y^2, z|\mathbf{x}) - 2\text{Cov}(yT, z|\mathbf{x})\theta_1(\mathbf{x}) + \text{Cov}(T, z|\mathbf{x})\theta_2(\mathbf{x}) = 0$$

$$\text{Cov}(y^3, z|\mathbf{x}) - 3\text{Cov}(y^2T, z|\mathbf{x})\theta_1(\mathbf{x}) + 3\text{Cov}(yT, z|\mathbf{x})\theta_2(\mathbf{x}) - \text{Cov}(T, z|\mathbf{x})\theta_3(\mathbf{x}) = 0$$

$$q(\mathbf{x}) = \mathbb{P}(z = 1|\mathbf{x})$$

$$\pi(\mathbf{x}) = \text{Cov}(T, z|\mathbf{x})$$

$$\eta_j(\mathbf{x}) = \text{Cov}(y^j, z|\mathbf{x})$$

$$\tau_j(\mathbf{x}) = \text{Cov}(Ty^j, z|\mathbf{x})$$

$$\eta_1(\mathbf{x}) = \pi(\mathbf{x})\theta_1(\mathbf{x})$$

$$\eta_2(\mathbf{x}) = 2\tau_1(\mathbf{x})\theta_1(\mathbf{x}) - \pi(\mathbf{x})\theta_2(\mathbf{x})$$

$$\eta_3(\mathbf{x}) = 3\tau_2(\mathbf{x})\theta_1(\mathbf{x}) - 3\tau_1(\mathbf{x})\theta_2(\mathbf{x}) + \pi(\mathbf{x})\theta_3(\mathbf{x})$$

**Lemma 19.1** (Various bits and pieces, e.g.  $\eta_1$  equation).

**Lemma 19.2** (Lemma for Appendix only with Bayes' Rule). *For mis-classification probabilities*

$$\begin{aligned} P(T^* = 1|T = 1, Z = k) &= P(T = 1|T^* = 1) \left( \frac{p_k^*}{p_k} \right) = (1 - \alpha_1) \left( \frac{p_k^*}{p_k} \right) \\ P(T^* = 1|T = 0, Z = k) &= P(T = 0|T^* = 1) \left( \frac{p_k^*}{1 - p_k} \right) = \alpha_1 \left( \frac{p_k^*}{1 - p_k} \right) \\ P(T^* = 0|T = 1, Z = k) &= P(T = 1|T^* = 0) \left( \frac{1 - p_k^*}{p_k} \right) = \alpha_0 \left( \frac{1 - p_k^*}{p_k} \right) \\ P(T^* = 0|T = 0, Z = k) &= P(T = 0|T^* = 0) \left( \frac{1 - p_k^*}{1 - p_k} \right) = (1 - \alpha_0) \left( \frac{1 - p_k^*}{1 - p_k} \right) \end{aligned}$$

**Theorem 19.1** (Non-identification Result).

**Proof of Theorem 19.1.** □

**Proposition 19.1.** *Under Assumptions ???,  $\eta_2(\mathbf{x}) = 2\tau_1(\mathbf{x})\theta_1(\mathbf{x}) - \pi(\mathbf{x})\theta_2(\mathbf{x})$ .*

**Proof of Proposition 19.1.** Since  $z$  is binary,  $\text{Cov}(w, z|\mathbf{x}) = q(\mathbf{x})[1 - q(\mathbf{x})] \Delta_z \mathbb{E}(w|\mathbf{x}, z)$  for any  $w$ , where we define  $\Delta_z \mathbb{E}(w|\mathbf{x}, z) \equiv \mathbb{E}(w|\mathbf{x}, z = 1) - \mathbb{E}(w|\mathbf{x}, z = 0)$ . Hence it suffices to show that

$$\Delta_z \mathbb{E}(y^2|\mathbf{x}, z) = 2\Delta_z \mathbb{E}(Ty|\mathbf{x}, z)\theta_1(\mathbf{x}) - \Delta_z \mathbb{E}(T|\mathbf{x}, z)\theta_2(\mathbf{x}).$$

By iterated expectations

$$\begin{aligned} \mathbb{E}(y^2|\mathbf{x}, z) &= \mathbb{E}(y^2|\mathbf{x}, T^* = 0, z)\mathbb{P}(T^* = 0|\mathbf{x}, z) + \mathbb{E}(y^2|\mathbf{x}, T^* = 1, z)\mathbb{P}(T^* = 1|\mathbf{x}, z) \\ \mathbb{E}(Ty|\mathbf{x}, z) &= \mathbb{E}(Ty|\mathbf{x}, T^* = 0, z)\mathbb{P}(T^* = 0|\mathbf{x}, z) + \mathbb{E}(Ty|\mathbf{x}, T^* = 1, z)\mathbb{P}(T^* = 1|\mathbf{x}, z) \end{aligned}$$

and since  $y = c(\mathbf{x}) + \beta(\mathbf{x})T^* + \varepsilon$ ,

$$\begin{aligned} \mathbb{E}(y^2|\mathbf{x}, T^* = 1, z) &= [c(\mathbf{x}) + \beta(\mathbf{x})]^2 + 2[c(\mathbf{x}) + \beta(\mathbf{x})] \mathbb{E}(\varepsilon|\mathbf{x}, T^* = 1, z) + \mathbb{E}(\varepsilon^2|\mathbf{x}, T^* = 1, z) \\ \mathbb{E}(Ty|\mathbf{x}, T^* = 1, z) &= [c(\mathbf{x}) + \beta(\mathbf{x})] \mathbb{E}(T|\mathbf{x}, T^* = 1, z) + \mathbb{E}(T\varepsilon|\mathbf{x}, T^* = 1, z) \\ \mathbb{E}(y^2|\mathbf{x}, T^* = 0, z) &= c(\mathbf{x})^2 + 2c(\mathbf{x}) \mathbb{E}(\varepsilon|\mathbf{x}, T^* = 0, z) + \mathbb{E}(\varepsilon^2|\mathbf{x}, T^* = 0, z) \\ \mathbb{E}(Ty|\mathbf{x}, T^* = 0, z) &= c(\mathbf{x}) \mathbb{E}(T|\mathbf{x}, T^* = 0, z) + \mathbb{E}(T\varepsilon|\mathbf{x}, T^* = 0, z). \end{aligned}$$

using Assumptions ?? and 19.2. □

**Proposition 19.2.** *Under Assumptions ???,  $\eta_3(\mathbf{x}) = 3\tau_2(\mathbf{x})\theta_1(\mathbf{x}) - 3\tau_1(\mathbf{x})\theta_2(\mathbf{x}) + \pi(\mathbf{x})\theta_3(\mathbf{x})$*

**Proof of Proposition 19.2.** □

**Corollary 19.1.** *Suppose that  $\eta_1(\mathbf{x}), \eta_2(\mathbf{x}), \eta_3(\mathbf{x}), \tau_1(\mathbf{x}), \tau_2(\mathbf{x})$  and  $\pi(\mathbf{x})$  are identified and that  $\pi(\mathbf{x}) \neq 0$ . Then, under the conditions of Lemmas ???,  $\theta_1(\mathbf{x}), \theta_2(\mathbf{x})$ , and  $\theta_3(\mathbf{x})$  are identified.*

**Proof of Corollary 19.1.** For ease of notation we suppress dependence on  $\mathbf{x}$  throughout this argument. Propositions ??? yield a triangular linear system of equations in  $\theta_1, \theta_2, \theta_3$ , namely

$$\begin{aligned}\pi\theta_1 &= \eta_1 \\ 2\tau_1\theta_1 - \pi\theta_2 &= \eta_2 \\ 3\tau_2\theta_1 - 3\tau_1\theta_2 + \pi\theta_3 &= \eta_3\end{aligned}$$

By inspection, the determinant of the system is  $-\pi^3$  so a unique solution exists if  $z$  is correlated with  $T^*$ . In particular,

$$\begin{aligned}\theta_1 &= \eta_1/\pi \\ \theta_2 &= 2\tau_1\eta_1/\pi^2 - \eta_2/\pi \\ \theta_3 &= \eta_3/\pi - 3(\tau_2\eta_1 + \tau_1\eta_2)/\pi^2 + 6\tau_1\eta_1/\pi^3.\end{aligned}$$

□

**Theorem 19.2** (Identification of  $\beta, \alpha_0, \alpha_1$ ).

**Proof of Theorem 19.2.** For ease of notation we suppress dependence on  $\mathbf{x}$  throughout this argument. So long as  $\beta \neq 0$ , we can rearrange Equations 33 and 34 to obtain

$$A = \theta_2/\theta_1^2 = 1 + (\alpha_0 - \alpha_1) \tag{35}$$

$$B = \theta_3/\theta_1^3 = (1 - \alpha_0 - \alpha_1)^2 + 6\alpha_0(1 - \alpha_1) \tag{36}$$

Equation 35 gives  $(1 - \alpha_1) = A - \alpha_0$ . Hence  $(1 - \alpha_0 - \alpha_1) = A - 2\alpha_0$  and  $\alpha_0(1 - \alpha_1) = \alpha_0(A - \alpha_0)$ . Substituting into Equation 36 and simplifying,  $(A^2 - B) + 2A\alpha_0 - 2\alpha_0^2 = 0$ . Substituting for  $\alpha_0$  analogously yields a quadratic in  $(1 - \alpha_1)$  with *identical* coefficients. It follows that one root of  $(A^2 - B) + 2Ar - 2r^2 = 0$  is  $\alpha_0$  and the other is  $1 - \alpha_1$ . Solving,

$$r = \frac{A}{2} \pm \sqrt{3A^2 - 2B} = \frac{1}{\theta_1^2} \left( \frac{\theta_2}{2} \pm \sqrt{3\theta_2^2 - 2\theta_1\theta_3} \right). \tag{37}$$

By Equations 33 and 34,

$$\begin{aligned}3\theta_2^2 - 2\theta_1\theta_3 &= 3[\theta_1^2(1 + \alpha_0 - \alpha_1)]^2 - 2\theta_1\{\theta_1^3[(1 - \alpha_0 - \alpha_1)^2 + 6\alpha_0(1 - \alpha_1)]\} \\ &= \theta_1^4\{3(1 + \alpha_0 - \alpha_1)^2 - 2[(1 - \alpha_0 - \alpha_1)^2 + 6\alpha_0(1 - \alpha_1)]\}.\end{aligned}$$

Expanding the first term we find that

$$\begin{aligned}3(1 + \alpha_0 - \alpha_1)^2 &= 3[1 + 2(\alpha_0 - \alpha_1) + (\alpha_0 - \alpha_1)^2] \\ &= 3 + 6\alpha_0 - 6\alpha_1 + 3\alpha_0^2 + 3\alpha_1^2 - 6\alpha_0\alpha_1\end{aligned}$$

and expanding the second

$$\begin{aligned} 2 \left[ (1 - \alpha_0 - \alpha_1)^2 + 6\alpha_0(1 - \alpha_1) \right] &= 2 \left[ 1 - 2(\alpha_0 + \alpha_1) + (\alpha_0 + \alpha_1)^2 + 6\alpha_0 - 6\alpha_0\alpha_1 \right] \\ &= 2 + 8\alpha_0 - 4\alpha_1 + 2\alpha_0^2 + 2\alpha_1^2 - 8\alpha_0\alpha_1. \end{aligned}$$

Therefore

$$\begin{aligned} 3\theta_2^2 - 2\theta_1\theta_3 &= \theta_1^4 \{ 1 - 2\alpha_0 - 2\alpha_1 + \alpha_0^2 - \alpha_1^2 + 2\alpha_0\alpha_1 \} \\ &= \theta_1^4 \left[ (1 - \alpha_0 - \alpha_1)^2 \right] \end{aligned}$$

which is strictly greater than zero since  $\theta_1 \neq 0$  and  $\alpha_0 + \alpha_1 \neq 0$ . It follows that both roots of the quadratic are real. Moreover,  $3\theta_2^2/\theta_1^4 - 2\theta_3/\theta_1^3$  identifies  $(1 - \alpha_0 - \alpha_1)^2$ . Substituting into Equation [32](#), it follows that  $\beta$  is identified up to sign. If  $\alpha_0 + \alpha_1 < 1$  then  $\text{sign}(\beta) = \text{sign}(\theta_1)$  so that both the sign and magnitude of  $\beta$  are identified. If  $\alpha_0 + \alpha_1 > 1$  then  $1 - \alpha_1 > \alpha_0$  so  $(1 - \alpha_1)$  is the larger root of  $(A^2 - B) + 2Ar - 2r^2 = 0$  and  $\alpha_0$  is the smaller root.  $\square$