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

Francis J. DiTraglia & Camilo García-Jimeno

May 11, 2016

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

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

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 α_0, α_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 α_0 and α_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.

3   Stronger Bounds for α_0, α_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

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

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

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

for all k. Defining the shorthand

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

this becomes

$$\widetilde{F}_{0k}(\tau) = (1 - \alpha_0)(1 - p_k^*)F_{0k}^*(\tau) + \alpha_1 p_k^* F_{1k}^*(\tau) \tag{3.1}$$
$$\widetilde{F}_{1k}(\tau) = \alpha_0(1 - p_k^*)F_{0k}^*(\tau) + (1 - \alpha_1)p_k^* F_{1k}^*(\tau) \tag{3.2}$$

Now, solving Equation 3.1 for $p_k^* F_{1k}^*(\tau)$ we have

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

Substituting this into Equation 3.2,

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

4

and therefore

$$\widetilde{F}_{1k}(\tau) = \frac{1 - \alpha_1}{\alpha_1}\widetilde{F}_{0k}(\tau) - \frac{1 - p_k - \alpha_1}{\alpha_1}F_{0k}^*(\tau) \tag{3.3}$$

Equation 3.3 relates the observable $\widetilde{F}_{1k}(\tau)$ to the mis-classification error rate α_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

$$\widetilde{F}_{1k}(\tau) \leq \frac{1 - \alpha_1}{\alpha_1}\widetilde{F}_{0k}(\tau) \tag{3.4}$$

while substituting 1 gives

$$\widetilde{F}_{1k}(\tau) \geq \frac{1 - \alpha_1}{\alpha_1}\widetilde{F}_{0k}(\tau) - \frac{1 - p_k - \alpha_1}{\alpha_1} \tag{3.5}$$

Rearranging Equation 3.4

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

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

$$\alpha_1 \leq \frac{\widetilde{F}_{0k}(\tau)}{\widetilde{F}_{0k}(\tau) + \widetilde{F}_{1k}(\tau)} = (1 - p_k)\left[\frac{F_{0k}(\tau)}{F_k(\tau)}\right] \tag{3.6}$$

since $\widetilde{F}_{1k}(\tau) + \widetilde{F}_{1k}(\tau) \geq 0$. Proceeding similarly for Equation 3.5,

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

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

$$\alpha_1 \leq \frac{1 - \widetilde{F}_{0k}(\tau) - p_k}{1 - \widetilde{F}_{1k}(\tau) - \widetilde{F}_{0k}(\tau)} = (1 - p_k)\left[\frac{1 - F_{0k}(\tau)}{1 - F_k(\tau)}\right] \tag{3.7}$$

The bounds given in Equations 3.6 and 3.7 relate α_1 to observable quantities *only* and hold for all values of τ 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 α_0. First solve Equation 3.1 for $(1 - p_k^*)F_{0k}^*(\tau)$:

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

and then substitute into Equation 3.2:

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

and therefore

$$\widetilde{F}_{1k}(\tau) = \frac{\alpha_0}{1 - \alpha_0}\widetilde{F}_{0k}(\tau) + \frac{p_k - \alpha_0}{1 - \alpha_0}F_{1k}^*(\tau) \tag{3.8}$$

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

$$\widetilde{F}_{1k}(\tau) \geq \frac{\alpha_0}{1 - \alpha_0}\widetilde{F}_{0k}(\tau) \tag{3.9}$$

while substituting one gives

$$\widetilde{F}_{1k}(\tau) \leq \frac{\alpha_0}{1 - \alpha_0}\widetilde{F}_{0k}(\tau) + \frac{p_k - \alpha_0}{1 - \alpha_0} \tag{3.10}$$

Now, rearranging Equation 3.9,

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

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

$$\alpha_0 \leq \frac{\widetilde{F}_{1k}(\tau)}{\widetilde{F}_{0k}(\tau) + \widetilde{F}_{1k}(\tau)} = p_k \left[\frac{F_{1k}(\tau)}{F_k(\tau)} \right] \tag{3.11}$$

since $\left[\widetilde{F}_{0k}(\tau) + \widetilde{F}_{1k}(\tau) \right] \geq 0$. Similarly, rearranging Equation 3.10

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

Therefore

$$\alpha_0 \leq \frac{1 - \widetilde{F}_{1k}(\tau) - (1 - p_k)}{1 - \widetilde{F}_{0k}(\tau) - \widetilde{F}_{1k}(\tau)} = p_k \left[\frac{1 - F_{1k}(\tau)}{1 - F_k(\tau)} \right] \tag{3.12}$$

Putting Everything Together For all k we have

$$\alpha_0 \leq p_k \min_\tau \left\{ \left[\frac{F_{1k}(\tau)}{F_k(\tau)} \right] \wedge \left[\frac{1 - F_{1k}(\tau)}{1 - F_k(\tau)} \right] \right\} \leq p_k \tag{3.13}$$

$$\alpha_1 \leq (1 - p_k) \min_\tau \left\{ \left[\frac{F_{0k}(\tau)}{F_k(\tau)} \right] \wedge \left[\frac{1 - F_{0k}(\tau)}{1 - F_k(\tau)} \right] \right\} \leq (1 - p_k) \tag{3.14}$$

Note that these bounds can only improve upon those derived in the previous section since the ratio of CDFs tends to one as $\tau \to \infty$. To derive these tighter bounds we have made no assumption regarding the relationship between Z and the error term ε. 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 α_0, α_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

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

for all k and all τ. Evaluating at two values k and ℓ in the support of Z and equating

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

or equivalently

$$F_k(\tau) - F_\ell(\tau) = p_\ell^* \left[F_{1\ell}^*(\tau + \beta) - F_{1\ell}^*(\tau) \right] - p_k^* \left[F_{1k}^*(\tau + \beta) - F_{1k}^*(\tau) \right] \qquad (5.1)$$

for all τ. Now we simply need to re-express all of the "star" quantities, namely p_k^*, p_ℓ^* and $F_{1k}^*, F_{1\ell}^*$ in terms of α_0, α_1 and the *observable* probability distributions F_{1k} and $F_{1\ell}$

and observable probabilities p_k, p_ℓ. To do this, we use the fact that

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

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

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

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

for all k. Combining this with Equation 5.1, we find that

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

Now, define

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

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

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

or in moment-condition form

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

Each value of τ yields a moment condition.

6 Special Case: $\alpha_0 = 0$

In this case the expressions from above simplify to

$$(1 - \alpha_1)\left[F_k(\tau) - F_\ell(\tau)\right] = \left[p_\ell F_{1\ell}(\tau + \beta) - p_k F_{1k}(\tau + \beta) - p_\ell F_{1\ell}(\tau) + p_k F_{1k}(\tau)\right] \quad (6.1)$$

for all τ. Now, provided that all of the CDFs are differentiable we have[1]

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

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

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

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

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

$$= e^{-i\omega\beta} p_\ell \varphi_{1\ell}(\omega)$$

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

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

which, after collecting terms, simplifies to

$$(1 - \alpha_1)\left[\varphi_k(\omega) - \varphi_\ell(\omega)\right] = \left(e^{-i\omega\beta} - 1\right)\left[p_\ell\varphi_{1\ell}(\omega) - p_k\varphi_{1k}(\omega)\right] \quad (6.2)$$

for all ω. Equation 6.2 contains exactly the same information as Equation 6.1 but gives us a more convenient way to prove identification since β enters in a simpler way.

---

[1]There must be a way to generalize this using Lebesgue.

Leibniz's formula for the rth derivative of a product of two functions f and g is:

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

where $f^{(r)}$ denotes the rth 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 6.2 gives

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

$$= \left(e^{-i\omega\beta} - 1 \right) \left[p_\ell \varphi_{1\ell}^{(r)}(\omega) - p_k \varphi_{1k}^{r}(\varphi) \right] + e^{-i\omega\beta} \sum_{s=1}^{r} \binom{r}{s} (-i\beta)^s \left[p_\ell \varphi_{1\ell}^{(r-s)}(\omega) - p_k \varphi_{1k}^{(r-s)}(\omega) \right]$$

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

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

Combining this with the LHS of Equation 6.2, also differentiated r times and evaluated at zero, we have

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

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 rth derivative of φ. The same applies for the conditional characteristic functions we consider here. Hence, provided that the rth moments exist,

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

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

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

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

$$(1-\alpha_1)\left(E_k[Y^r] - E_\ell[Y^r]\right) = \sum_{s=1}^{r} \binom{r}{s}(-\beta)^s \left(p_\ell E_{1\ell}\left[Y^{r-s}\right] - p_k E_{1k}\left[Y^{r-s}\right]\right) \quad (6.3)$$

Equation 6.3 can be used to generate moment equations that are implied by the Equation 6.2 and the equivalent representation in terms of CDFs: Equation 6.1. Assuming that the conditional first moments exist, we can evaluate Equation 6.3 at $r=1$, yielding

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

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

$$\mathcal{W} \equiv \frac{E_k[Y] - E_\ell[Y]}{p_k - p_\ell} = \frac{\beta}{1-\alpha_1} \quad (6.4)$$

Evaluating Equation 6.3 at $r=2$, we have

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

Rearranging, we have

$$E_k[Y^2] - E_\ell[Y^2] = \frac{\beta}{1-\alpha_1}\left[2\left(p_k E_{1k}[Y] - p_\ell E_{1\ell}[Y]\right) - \beta(p_k - p_\ell)\right] \quad (6.5)$$

Substituting Equation 6.4, we can replace $\beta/(1-\alpha_1)$ with a function of observables only, namely $\mathcal{W}$. Solving, we find that

$$\beta = \frac{2\left(p_k E_{1k}[Y] - p_\ell E_{1\ell}[Y]\right)}{p_k - p_\ell} - \frac{E_k[Y^2] - E_\ell[Y^2]}{E_k[Y] - E_\ell[Y]} \quad (6.6)$$

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 6.1 this solution is ruled out by the assumption that $E_k[Y] \neq E_\ell[Y]$ via Equation 6.4. The mis-classification error rate α_1 is likewise uniquely identified under these assumptions. Substituting $\beta/\mathcal{W} = 1 - \alpha_1$ into Equation 6.6

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

and thus

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

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

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

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 τ, yielding

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

which simplifies to

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

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

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

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

First Moments   Taking $r = 1$ gives

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

Simplifying,

$$\mathcal{W} \equiv \frac{E_k[Y] - E_\ell[Y]}{p_k - p_\ell} = \frac{\beta}{1 - \alpha_0 - \alpha_1} \tag{8.1}$$

Second Moments   Now, taking $r = 2$ gives

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

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

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

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

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

Now, simplifying

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

14

and substituting Equation 8.1 to eliminate β, 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

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

and therefore

$$\alpha_1 - \alpha_0 = 1 + (p_k - p_\ell)\left[\frac{E_k[Y^2] - E_\ell[Y^2]}{(E_k[Y] - E_\ell[Y])^2}\right] - 2\left[\frac{p_{1k}E_{1k}[Y] - p_\ell E_{1\ell}[Y]}{E_k[Y] - E_\ell[Y]}\right] \qquad (8.2)$$

"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:

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

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:

$$\widetilde{y_{1k}^r} = p_k E_{1k}[Y^r]$$

$$\widetilde{y_{0k}^r} = (1 - p_k)E_{1k}[Y^r]$$

$$\Delta\overline{y^r} = E_k[Y^r] - E_\ell[Y^r]$$

$$\Delta\overline{Ty^r} = p_k E_{1k}[Y^r] - p_\ell E_{1\ell}[Y^r] = \widetilde{y_{1k}^r} - \widetilde{y_{1k}^r}$$

$$\mathcal{W} = (E_k[Y] - E_\ell[Y])/(p_k - p_\ell)$$

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 $\Delta\overline{Ty^0} = p_k - p_\ell$. 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

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

Now, simplifying the terms in the square brackets,

$$\alpha_0\left(\widetilde{y_{0k}^{r-s}} - \widetilde{y_{0\ell}^{r-s}}\right) - (1 - \alpha_0)\left(\widetilde{y_{1k}^{r-s}} - \widetilde{y_{1\ell}^{r-s}}\right) = \alpha_0\left[\left(\widetilde{y_{0k}^{r-s}} + \widetilde{y_{1k}^{r-s}}\right) - \left(\widetilde{y_{0\ell}^{r-s}} + \widetilde{y_{1\ell}^{r-s}}\right)\right] - \left(\widetilde{y_{1k}^{r-s}} - \widetilde{y_{1\ell}^{r-s}}\right)$$

$$= \alpha_0\left(E_k[Y^{r-s}] - E_\ell[Y^{r-s}]\right) - \Delta\overline{Ty^{r-s}}$$

$$= \alpha_0\Delta\overline{y^{r-s}} - \Delta\overline{Ty^{r-s}}$$

and hence

$$(1 - \alpha_0 - \alpha_1)\Delta\overline{y^r} = \sum_{s=1}^{r}\binom{r}{s}(-\beta)^s\left(\alpha_0\Delta\overline{y^{r-s}} - \Delta\overline{Ty^{r-s}}\right) \tag{8.3}$$

16

Third Moments   Evaluating Equation 8.3 at $r = 3$

$$(1 - \alpha_0 - \alpha_1)\Delta\overline{y^3} = \sum_{s=1}^{3} \binom{3}{s}(-\beta)^s \left(\alpha_0\Delta\overline{y^{3-s}} - \Delta\overline{Ty^{3-s}}\right)$$

$$= -3\beta\left(\alpha_0\Delta\overline{y^2} - \Delta\overline{Ty^2}\right) + 3\beta^2\left(\alpha_0\Delta\overline{y} - \Delta\overline{Ty}\right) + \beta^3(p_k - p_\ell)$$

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

$$\Delta\overline{y^3} = -3\mathcal{W}\left(\alpha_0\Delta\overline{y^2} - \Delta\overline{Ty^2}\right) + 3\beta\mathcal{W}\left(\alpha_0\Delta\overline{y} - \Delta\overline{Ty}\right) + \beta^2\mathcal{W}(p_k - p_\ell)$$

$$\frac{\Delta\overline{y^3}}{\mathcal{W}(p_k - p_\ell)} = \beta^2 + 3\beta\left(\frac{\alpha_0\Delta\overline{y} - \Delta\overline{Ty}}{p_k - p_\ell}\right) - 3\left(\frac{\alpha_0\Delta\overline{y^2} - \Delta\overline{Ty^2}}{p_k - p_\ell}\right)$$

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

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