Biometrika, *pp.* 1–8

Printed in Great Britain

Online supplement to bias attenuation results for nondifferentially mismeasured ordinal and coarsened confounders

BY ELIZABETH L. OGBURN AND TYLER J. VANDERWEELE

1. PROOFS OF LEMMAS

LEMMA 1. *If* $E(Y \mid A=1) \geq E_{C'}(Y \mid A=1) \geq E(Y_1)$ *and* $E(Y \mid A=0) \leq E_{C'}(Y \mid A=0) \leq E(Y_0)$ *or if* $E(Y \mid A=1) \leq E_{C'}(Y \mid A=1) \leq E(Y_1)$ *and* $E(Y \mid A=0) \geq E_{C'}(Y \mid A=0) \geq E(Y_0)$*, then the observed adjusted effect is between the crude and true effects for any effect measure that can be written as* $h[g\{E(Y_1)\} - g\{E(Y_0)\}]$ *where* g *and* h *are monotonic functions.*

The proof of this lemma is immediate.

LEMMA 2. *If* $E(Y \mid A, C)$ *and* $E(A \mid C)$ *are either both nonincreasing or both nondecreasing in* C*, then* $E(Y \mid A=1) \geq E(Y_1)$ *and* $E(Y \mid A=0) \leq E(Y_0)$*. If one of* $E(Y \mid A, C)$ *and* $E(A \mid C)$ *is nonincreasing and the other nondecreasing in* C*, then* $E(Y \mid A=1) \leq E(Y_1)$ *and* $E(Y \mid A=0) \geq E(Y_0)$*.*

The proof of this lemma is due to VanderWeele (2008). An analogous result holds comparing the observed adjusted and crude expectations by replacing C with C'.

LEMMA 3. *Suppose* C *is nondifferentially misclassified with respect to* A *and* Y*. If* $E(Y \mid A, C)$ *and* $E(A \mid C)$ *are both nondecreasing or both nonincreasing in* C*, then* $E_{C'}(Y \mid A=1) \geq E(Y_1)$ *and* $E_{C'}(Y \mid A=0) \leq E(Y_0)$*. If one of* $E(Y \mid A, C)$ *and* $E(A \mid C)$ *is nondecreasing and the other nonincreasing in* C*, then* $E_{C'}(Y \mid A=1) \leq E(Y_1)$ *and* $E_{C'}(Y \mid A=0) \geq E(Y_0)$*.*

Proof. We will prove that $E_{C'}(Y|A=1) \geq E(Y_1)$ when $E(Y|A,C)$ and $E(A|C)$ are both nondecreasing; the other cases follow by analogous arguments.

Assume that $E(Y|A,C)$ and $E(A|C)$ are nondecreasing in C. We want to show that

$$\sum_{j=1}^{K} E(Y \mid A=1, C=j) \left\{ \sum_{i=1}^{K} \mathrm{pr}(C=j \mid A=1, C'=i)\, \mathrm{pr}(C'=i) \right\} \geq \sum_{j=1}^{K} E(Y \mid A=1, C=j)\, \mathrm{pr}(C=j), \quad (1)$$

where it can easily be shown that the left hand side is equal to $E_{C'}(Y \mid A=1)$ and the right hand side is equal to $E(Y_1)$. It suffices to show that (1) holds for the special case of $E(Y \mid A=1, C=j) = 0$ for $j < k$ and $E(Y \mid A=1, C=j) = 1$ for $k \leq j \leq K$, because any sequence of nondecreasing conditional expectations $E(Y \mid A=1, C=j)$, $j = 1, \ldots, K$ can be written as a linear combination of $K-$dimensional vectors comprised of a sequence of 0's followed by a sequence of 1's. We now turn our attention to this special case, for which (1) reduces to

$$\sum_{j=k}^{K}\left\{\sum_{i=1}^{K}\mathrm{pr}\left(C=j\mid A=1,C'=i\right)\mathrm{pr}\left(C'=i\right)\right\}\geq\sum_{j=k}^{K}\mathrm{pr}\left(C=j\right)$$

or, equivalently,

$$\sum_{i=1}^{K}\mathrm{pr}\left(C'=i\right)\sum_{j=k}^{K}\mathrm{pr}\left(C=j\mid A=1,C'=i\right)\geq\sum_{i=1}^{K}\mathrm{pr}\left(C'=i\right)\sum_{j=k}^{K}\mathrm{pr}\left(C=j\mid C'=i\right).$$

Because the first sum is the same on the left hand side and on the right hand side of this inequality, it suffices to show that for each i the second sum on the left hand side, $\sum_{j=k}^{K}\mathrm{pr}\left(C=j\mid A=1,C'=i\right)=\left\{\sum_{j=k}^{K}\mathrm{pr}(A=1\mid C=j)\mathrm{pr}(C=j)\mathrm{pr}(C'=i\mid C=j)\right\}/\left\{\sum_{j}\mathrm{pr}(A=1\mid C=j)\mathrm{pr}(C=j)\mathrm{pr}(C'=i\mid C=j)\right\}$, is greater than or equal to the second sum on the right hand side, $\sum_{j=k}^{K}\mathrm{pr}\left(C=j\mid C'=i\right)=\left\{\sum_{j=k}^{K}\mathrm{pr}(C=j)\mathrm{pr}(C'=i\mid C=j)\right\}/\left\{\sum_{j}\mathrm{pr}(C=j)\mathrm{pr}(C'=i\mid C=j)\right\}$. This holds because, by the assumption that $\mathrm{pr}\left(A=1\mid C\right)$ is nondecreasing, $\left\{\sum_{j=k}^{K}\mathrm{pr}(A=1\mid C=j)\mathrm{pr}(C=j)\mathrm{pr}(C'=i\mid C=j)\right\}/\left\{\sum_{j}\mathrm{pr}(A=1\mid C=j)\mathrm{pr}(C=j)\mathrm{pr}(C'=i\mid C=j)\right\}$ $\geq\left\{\sum_{j=k}^{K}\mathrm{pr}(C=j)\mathrm{pr}(C'=i\mid C=j)\right\}/\left\{\sum_{j}\mathrm{pr}(C=j)\mathrm{pr}(C'=i\mid C=j)\right\}$ for all i. □

LEMMA 4. *Suppose C is nondifferentially misclassified with respect to A and Y with tapered misclassification probabilities. If $E\left(Y\mid A,C\right)$ and $E\left(A\mid C\right)$ are both nondecreasing or both nonincreasing in C, then $E_{C'}\left(Y\mid A=1\right)\leq E\left(Y\mid A=1\right)$ and $E_{C'}\left(Y\mid A=0\right)\geq E\left(Y\mid A=0\right)$. If one of $E\left(Y\mid A,C\right)$ and $E\left(A\mid C\right)$ is nondecreasing and the other nonincreasing in C, then $E_{C'}\left(Y\mid A=1\right)\geq E\left(Y\mid A=1\right)$ and $E_{C'}\left(Y\mid A=0\right)\leq E\left(Y\mid A=0\right)$.*

Proof. We will prove that $E_{C'}\left(Y\mid A=1\right)\leq E\left(Y\mid A=1\right)$ when $E\left(Y\mid A,C\right)$ and $E\left(A\mid C\right)$ are both nondecreasing; the other cases follow by analogous arguments.

Assume $E\left(Y\mid A,C\right)$ and $E\left(A\mid C\right)$ are nondecreasing in C. We will show that $E\left(Y\mid A=1\right)-E_{C'}\left(Y\mid A=1\right)$ is minimized at 0. We minimize this difference with respect to $E\left(Y\mid A=1,C=k\right)$, $k=1,...,K$; at its minimum, $E\left(Y\mid A=1\right)-E_{C'}\left(Y\mid A=1\right)=0$ for all possible values of $E\left(A\mid C=k\right)$, $\mathrm{pr}\left(C=k\right)$, and $\mathrm{pr}\left(C'=i\mid C=j\right)$, $i,j=1,...,K$. The derivative of $E\left(Y\mid A=1\right)-E_{C'}\left(Y\mid A=1\right)$ with respect to $E\left(Y\mid A=1,C=k\right)$ is

$$\mathrm{pr}(C=k)\mathrm{pr}(A=1\mid C=k)\left\{\frac{1}{E(A)}-\sum_{i}\frac{\mathrm{pr}\left(C'=i\mid C=k\right)}{\mathrm{pr}(A=1\mid C'=i)}\right\}.$$

We will prove that $\sum_{i}\left\{\mathrm{pr}\left(C'=i\mid C=k\right)/\mathrm{pr}(A=1\mid C'=i)\right\}$ is nonincreasing in k. Then, because $1/E(A)$ is constant in k, one of the following cases must hold: the derivative of $E\left(Y\mid A=1\right)-E_{C'}\left(Y\mid A=1\right)$ with respect to $E\left(Y\mid A=1,C=k\right)$ is positive for all k; the derivative is negative for all k; or the derivative is negative for k below some cutoff and positive for k above the cutoff. In all three cases, $E\left(Y\mid A=1\right)-E_{C'}\left(Y\mid A=1\right)$ is minimized by setting $E\left(Y\mid A=1,C=k\right)$ to be constant in k, which results in $E\left(Y\mid A=1\right)=E_{C'}\left(Y\mid A=1\right)$.

It suffices to show that $\mathrm{pr}(A=1\mid C'=i)$ is nondecreasing in i, because then, by the assumption of tapering misclassification probabilities, $\sum_{i}\mathrm{pr}\left(C'=i\mid C=k\right)/\mathrm{pr}(A=1\mid C'=i)$ places less weight on smaller values of $1/\mathrm{pr}(A=1\mid C'=i)$ and more weight on larger values than does $\sum_{i}\mathrm{pr}\left(C'=i\mid C=m\right)/\mathrm{pr}(A=1\mid C'=i)$, for $m>k$. We now argue that $\mathrm{pr}(A=1\mid C'=i)\leq\mathrm{pr}(A=1\mid C'=l)$ for $l>i$ or, equivalently, that

$$\frac{\sum_j \text{pr}(A=1 \mid C=j)\text{pr}(C=j)\text{pr}\,(C'=i \mid C=j)}{\sum_j \text{pr}(C=j)\text{pr}\,(C'=i \mid C=j)} \leq \frac{\sum_j \text{pr}(A=1 \mid C=j)\text{pr}(C=j)\text{pr}\,(C'=l \mid C=j)}{\sum_j \text{pr}(C=j)\text{pr}\,(C'=l \mid C=j)}.$$

This inequality holds because the numerator of the left side of the inequality pairs small values of $\text{pr}(A=1 \mid C=j)$ with larger values of $\text{pr}\,(C'=i \mid C=j)$ than does the numerator of the right side, and it pairs large values of $\text{pr}(A=1 \mid C=j)$ with smaller values of $\text{pr}\,(C'=i \mid C=j)$ than does the numerator of the right side. □

LEMMA 5. *Suppose that A is binary and C ordinal and coarsened. If $E\,(Y \mid A, C)$ and $E\,(A \mid C)$ are both nondecreasing or both nonincreasing in C, then $E\,(Y \mid A=1) \geq E_{C'}\,(Y \mid A=1) \geq E\,(Y_1)$ and $E\,(Y \mid A=0) \leq E_{C'}\,(Y \mid A=0) \leq E\,(Y_0)$. If one of $E\,(Y \mid A, C)$ and $E\,(A \mid C)$ is nonincreasing and the other nondecreasing, then $E\,(Y \mid A=1) \leq E_{C'}\,(Y \mid A=1) \leq E\,(Y_1)$ and $E\,(Y \mid A=0) \geq E_{C'}\,(Y \mid A=0) \geq E\,(Y_0)$.*

Proof. That $E(A \mid C')$ and $E(Y \mid A, C')$ are monotonic and in the same direction as $E(A \mid C)$ and $E(Y \mid A, C)$, respectively, is immediate. Therefore, by Lemma 2, we know that the desired relations hold between $E(Y \mid A=a)$ and $E_{C'}(Y \mid A=a)$ and between $E(Y \mid A=a)$ and $E(Y_a)$, $a=0,1$. It remains to be shown that $E_{C'}(Y \mid A=1) \geq E(Y_1)$ and $E_{C'}(Y \mid A=0) \leq E(Y_0)$ when $E(A \mid C)$ and $E(Y \mid A, C)$ are both nondecreasing or both nonincreasing and that $E_{C'}(Y \mid A=1) \leq E(Y_1)$ and $E_{C'}(Y \mid A=0) \geq E(Y_0)$ when one of $E(A \mid C)$ and $E(Y \mid A, C)$ is nondecreasing and the other nonincreasing. We will prove that $E_{C'}\,(Y \mid A=1) \geq E\,(Y_1)$ under nondecreasing monotonicity of both $E\,(Y \mid A, C)$ and $E\,(A \mid C)$. The other cases follow by analogous arguments.

We will show that, at its minimum, $E_{C'}\,(Y \mid A=1) - E\,(Y_1) = 0$. The derivative of $E_{C'}\,(Y \mid A=1) - E\,(Y_1)$ with respect to $E\,(Y \mid A=1, C=k)$ is

$$\frac{\text{pr}\,(A=1 \mid C=k)\,\text{pr}\,(C=k)\sum_{i=l}^{m}\text{pr}(C=i)}{\sum_{i=l}^{m}\text{pr}\,(A=1 \mid C=i)\,\text{pr}(C=i)} - \frac{\text{pr}\,(A=1 \mid C=k)\,\text{pr}(C=k)\sum_{i=l}^{m}\text{pr}(C=i)}{\sum_{i=l}^{m}\text{pr}\,(A=1 \mid C=k)\,\text{pr}(C=i)}$$

where $l \leq k \leq m$ and levels l through m of C comprise a single level of C'. This derivative will be negative for $k=l$, positive for $k=m$, and nondecreasing in k for intervening values. Therefore, in order to minimize $E_{C'}\,(Y \mid A=1) - E\,(Y_1)$ for each stratum of C', we maximize $E\,(Y \mid A=1, C=k)$ for k at the lower end of the levels of C comprising the stratum and minimize it for k at the higher end. However, since we are constrained by nondecreasing monotonicity, this is tantamount to setting $E\,(Y \mid A=1, C=k)$ to be constant within each stratum of C'. Under this constraint, $E_{C'}\,(Y \mid A=1) = E\,(Y_1)$. □

2. COUNTEREXAMPLES

Table 1 gives a counterexample to the partial control result when $E\,(Y \mid A, C)$ is monotonic but $E\,(A \mid C)$ is not. The observed $2 \times 2 \times 3$ table was generated from the true by the same tapered misclassification probabilities as Table 1 in the main text. Monotonicity holds for $E\,(A \mid C)$: $E\,(A \mid C=1) = 0{\cdot}55$, $E\,(A \mid C=2) = 0{\cdot}71$, and $E\,(A \mid C=3) = 0{\cdot}81$. Monotonicity fails to hold for $E\,(Y \mid A, C)$: $E\,(Y \mid A=1, C=1) = 0{\cdot}83$, $\text{E}(Y \mid A=1, C=2) = 0{\cdot}81$ and $E\,(Y \mid A=1, C=3) = 0{\cdot}85$, while $E\,(Y \mid A=0, C=1) = 0{\cdot}33$, $\text{E}(Y \mid A=0, C=2) = 0{\cdot}20$ and $E\,(Y \mid A=0, C=3) = 0{\cdot}83$. Instead of the partial control ordering, we observe that the observed adjusted measure is greater than the crude which is greater than the true measure on all three scales, for both the average treatment effect and the effect of treatment on the treated.

Table 1. *The partial control result is violated when $E\,(Y \mid A, C)$ is not monotonic in C.*

Crude

	$A=0$	$A=1$
$Y=0$	221	76
$Y=1$	110	372

$\mathrm{RD}_{\mathrm{crude}}$	0·498
$\mathrm{RR}_{\mathrm{crude}}$	2·499
$\mathrm{OR}_{\mathrm{crude}}$	9·834

$\mathrm{RD}^{T}_{\mathrm{crude}}$	0·498
$\mathrm{RR}^{T}_{\mathrm{crude}}$	2·499
$\mathrm{OR}^{T}_{\mathrm{crude}}$	9·834

True

$C=1$	$A=0$	$A=1$
$Y=0$	200	60
$Y=1$	100	300

$C=2$	$A=0$	$A=1$
$Y=0$	20	12
$Y=1$	5	50

$C=3$	$A=0$	$A=1$
$Y=0$	1	4
$Y=1$	5	22

$\mathrm{RD}_{\mathrm{true}}$	0·492
$\mathrm{RR}_{\mathrm{true}}$	2·451
$\mathrm{OR}_{\mathrm{true}}$	9·579

$\mathrm{RD}^{T}_{\mathrm{true}}$	0·486
$\mathrm{RR}^{T}_{\mathrm{true}}$	2·415
$\mathrm{OR}^{T}_{\mathrm{true}}$	9·338

Obs

$C'=1$	$A=0$	$A=1$
$Y=0$	200	60
$Y=1$	100	300

$C'=2$	$A=0$	$A=1$
$Y=0$	12·4	8·8
$Y=1$	5	38·8

$C'=3$	$A=0$	$A=1$
$Y=0$	8·6	7·2
$Y=1$	5	33·2

$\mathrm{RD}_{\mathrm{obs}}$	0·499
$\mathrm{RR}_{\mathrm{obs}}$	2·504
$\mathrm{OR}_{\mathrm{obs}}$	9·900

$\mathrm{RD}^{T}_{\mathrm{obs}}$	0·499
$\mathrm{RR}^{T}_{\mathrm{obs}}$	2·505
$\mathrm{OR}^{T}_{\mathrm{obs}}$	9·869

Table 2 demonstrates that the partial control result may not hold if the misclassification probabilities are not tapered, even if $E\,(Y \mid A, C)$ and $E\,(A \mid C)$ are monotonic in C and if the probability of misclassification is less than $0{\cdot}5$ for any given subject. The observed $2 \times 2 \times 3$ table was generated from the true by the following misclassification probabilities: $\text{pr}\,(C' = 1 \mid C = 1) =$0·97, $\text{pr}\,(C' = 2 \mid C = 1) =$0·02, $\text{pr}\,(C' = 3 \mid C = 1) =$ pr$(C' = 1 \mid C = 2) =$0·01, $\text{pr}\,(C' = 2 \mid C = 2) =$0·51, $\text{pr}\,(C' = 3 \mid C = 2) =$ pr$(C' = 1 \mid C = 3) =$0·48, $\text{pr}\,(C' = 2 \mid C = 3) =$0·01, and $\text{pr}\,(C' = 3 \mid C = 3) =$0·51. Monotonicity holds for $E\,(Y \mid A, C)$ because $E\,(Y \mid A = 1, C = c)$ and $E\,(Y \mid A = 0, C = c)$ are both nondecreasing in c: $E\,(Y \mid A = 1, C = 1) = E\,(Y \mid A = 1, C = 2) = 0{\cdot}6$ and $E\,(Y \mid A = 1, C = 3) = 0{\cdot}67$, while $E\,(Y \mid A = 0, C = 1) = E\,(Y \mid A = 0, C = 2) = 0{\cdot}4$ and $E\,(Y \mid A = 0, C = 3) = 0{\cdot}67$. Monotonicity also holds for $E\,(A \mid C)$: $E\,(A \mid C = 1) = 0{\cdot}99$, $E\,(A \mid C = 2) = 0{\cdot}56$, and $E\,(A \mid C = 3) = 0{\cdot}4$. Instead of the partial control ordering, we observe that the observed adjusted measure is less than the crude which is less than the true measure on all three scales, for both the average treatment effect and the effect of treatment on the treated.

Table 2. *The partial control result is violated when the misclassification probabilities are not tapered.*

Crude

	$A=0$	$A=1$				
$Y=0$	126	298	$\mathrm{RD}_{\mathrm{crude}}$	0·107	$\mathrm{RD}^{T}_{\mathrm{crude}}$	0·107
$Y=1$	124	452	$\mathrm{RR}_{\mathrm{crude}}$	1·215	$\mathrm{RR}^{T}_{\mathrm{crude}}$	1·215
			$\mathrm{OR}_{\mathrm{crude}}$	1·541	$\mathrm{OR}^{T}_{\mathrm{crude}}$	1·541

True

$C=1$	$A=0$	$A=1$	$C=2$	$A=0$	$A=1$	$C=3$	$A=0$	$A=1$	$\mathrm{RD}_{\mathrm{true}}$	0·168	$\mathrm{RD}^{T}_{\mathrm{true}}$	0·181
$Y=0$	3	200	$Y=0$	93	78	$Y=0$	30	20	$\mathrm{RR}_{\mathrm{true}}$	1·382	$\mathrm{RR}^{T}_{\mathrm{true}}$	1·430
$Y=1$	2	295	$Y=1$	62	117	$Y=1$	60	40	$\mathrm{OR}_{\mathrm{true}}$	1·974	$\mathrm{OR}^{T}_{\mathrm{true}}$	2·083

Obs

$C'=1$	$A=0$	$A=1$	$C'=2$	$A=0$	$A=1$	$C'=3$	$A=0$	$A=1$	$\mathrm{RD}_{\mathrm{obs}}$	0·048	$\mathrm{RD}^{T}_{\mathrm{obs}}$	0·026
$Y=0$	18·24	204·38	$Y=0$	47·79	43·98	$Y=0$	59·97	49·64	$\mathrm{RR}_{\mathrm{obs}}$	1·086	$\mathrm{RR}^{T}_{\mathrm{obs}}$	1·046
$Y=1$	31·36	306·52	$Y=1$	32·26	65·97	$Y=1$	60·38	79·51	$\mathrm{OR}_{\mathrm{obs}}$	1·217	$\mathrm{OR}^{T}_{\mathrm{obs}}$	1·116

REFERENCES

VANDERWEELE, T. (2008). The sign of the bias of unmeasured confounding. *Biometrics* **64**, 702–706.