eAppendix for "Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments and confounders"

1. Causal diagrams and the causal interpretation of sensitivity parameters

Each of the bias formulas given in Theorem 1 make use of expressions of the form $E(Y|a,x,u)-E(Y|a,x,u')$. Such expression can sometimes, but not always, be interpreted causally. Note that the expression $E(Y|a,x,u)-E(Y|a,x,u')$ cannot be interpreted as the total effect of U on Y since the expression $E(Y|a,x,u)-E(Y|a,x,u')$ is conditional on A, and here U is a cause of A; we are thus conditioning on a variable, namely A, that is on the causal pathway from U to Y. Let Y_{ua} denote the counterfactual value of Y if, possibly contrary to fact, U were set to u and A were set to a. To interpret $E(Y|a,x,u)-E(Y|a,x,u')$ as a controlled direct effect two conditions are needed. The first condition is that the effect of U on Y is unconfounded given X, i.e. that $Y_{ua} \coprod U|X$. The second condition is that $Y_{ua} \coprod A|X,U$; note that this second condition $Y_{ua} \coprod A|X,U$ is slightly different than the condition assumed throughout the paper that $Y_a \coprod A|X,U$. If both of these conditions hold, $Y_{ua} \coprod U|X$ and $Y_{ua} \coprod A|X,U$ then $E(Y|a,x,u)-E(Y|a,x,u')$ can be interpreted, within strata $X=x$, as the controlled direct effect of U on Y with A set to a.[33] In other words, if we have both $Y_{ua} \coprod U|X$ and $Y_{ua} \coprod A|X,U$ then $E(Y|a,x,u)-E(Y|a,x,u')$ is the effect of U on Y not mediated through A. Causal directed acyclic graphs can be helpful in clarifying whether these additional conditions hold.[33] Graphically, on a causal directed acyclic graph we will have that $Y_{ua} \coprod U|X$ if all backdoor paths from U to Y are blocked by X and we will have that $Y_{ua} \coprod A|X,U$ if all backdoor paths from A to Y are blocked by (X,U). See Pearl[33] for the definition of blocked paths and further discussion of the rules concerning causal directed acyclic graphs and the identification of causal effects.

The condition that $Y_{ua} \coprod U|X$ (i.e. that the effect of U on Y is unconfounded given X) will not always hold; the condition is not implied by the effect of A on Y being unconfounded given (X,U). Figure 1 below gives an example in which the effect of U on Y is unconfounded given X.

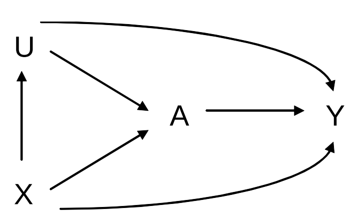

Fig. 1. Causal DAG in which the effect of U on Y is unconfounded given X.

Figure 2 below gives an example in which the effect of U on Y is not unconfounded given X.

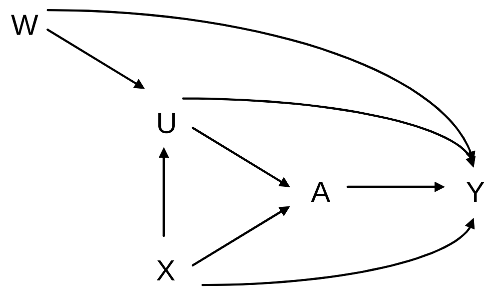

Fig. 2. Causal DAG in which the effect of U on Y is not unconfounded given X but in which the effect of A on Y in unconfounded given (X,U).

In Figure 2, suppose that data is available for neither W nor U. On the causal directed acyclic graph given in Figure 2 it is the case that the effect of A on Y is unconfounded given (X, U); control need not be made for W; the set (X, U) blocks all backdoor paths from A to Y. There is, however, an unblocked backdoor path from U to Y (through W) and thus the effect of the effect of U on Y is not unconfounded given X (although it would be unconfounded given (X, W)). Although the sensitivity analysis and external adjustment procedures described in section 3 could still be used by considering only U and ignoring W (since $Y_a \coprod A|X, U$), the expression $E(Y|a, x, u) - E(Y|a, x, u')$ cannot be interpreted as a causal direct effect. However, as noted in the paper, the expression $E(Y|a, x, u) - E(Y|a, x, u')$ need not have a causal interpretation to be used in sensitivity analysis.

If U is a cause of some of the variables in X as in the causal directed acyclic graph given in Figure 3 then the effect of U on Y will not unconfounded given X because some variables in X are on the pathway from U to Y; we will not have $Y_u \coprod U|X$; and $E(Y|a, x, u) - E(Y|a, x, u')$ cannot be interpreted as the controlled direct effect of U on Y with A set to a.

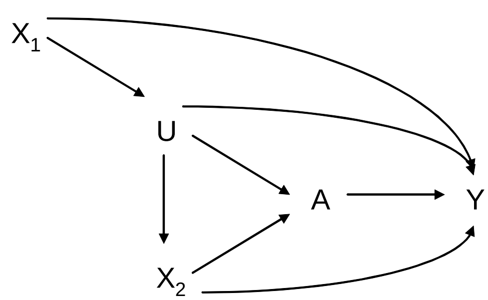

Fig. 3. Causal DAG in which U is a cause of some subset of X.

If, however, X can be partitioned into two sets, X_1 and X_2, so that $Y_{ux_2a} \coprod U|X_1$ and $Y_{ux_2a} \coprod X_2|X_1, U$ and $Y_{ux_2a} \coprod A|X, U$ where Y_{ux_2a} denotes the counterfactual value of Y if, possibly contrary to fact, U were set to u, X_2 were set to x_2 and A were set to a, then $E(Y|a, x, u) - E(Y|a, x, u')$ can be interpreted, within strata $X_1 = x_1$, as the controlled direct effect of U on Y with A set to a and X_2 set to x_2.[33] On a causal directed acyclic graph the condition $Y_{ux_2a} \coprod U|X_1$ will hold if X_1 blocks all backdoor paths from U to Y, the condition $Y_{ux_2a} \coprod X_2|X_1, U$ will hold if (X_1, U) block all backdoor paths from X_2 to Y; the condition $Y_{ux_2a} \coprod A|X, U$

will hold if (X, U) block all backdoor paths from A to Y. Thus for the causal directed acyclic graph given in Figure 3, $E(Y|a, x, u) - E(Y|a, x, u')$ can be interpreted as the controlled direct effect of U on Y with A set to a and X_2 set to x_2.

2. Bias Formulas for the Marginal and Conditional Odds Ratio

The conditional causal odds ratios in the total population or amongst those receiving treatment a_1 or a_0 are defined respectively by

$$\frac{E(Y_{a_1}|x)/\{1 - E(Y_{a_1}|x)\}}{E(Y_{a_0}|x)/\{1 - E(Y_{a_0}|x)\}}$$
$$\frac{E(Y_{a_1}|a_1, x)/\{1 - E(Y_{a_1}|a_1, x)\}}{E(Y_{a_0}|a_1, x)/\{1 - E(Y_{a_0}|a_1, x)\}}$$
$$\frac{E(Y_{a_1}|a_0, x)/\{1 - E(Y_{a_1}|a_0, x)\}}{E(Y_{a_0}|a_0, x)/\{1 - E(Y_{a_0}|a_0, x)\}}.$$

Define the bias expressions $d_{a_+}^{OR}(x)$, $d_{a_1}^{OR}(x)$ and $d_{a_0}^{OR}(x)$ as follows:

$$\begin{aligned}
d_{a_+}^{OR}(x) &= \frac{E(Y|a_1, x)/\{1 - E(Y|a_1, x)\}}{E(Y|a_0, x)/\{1 - E(Y|a_0, x)\}} \Big/ \frac{E(Y_{a_1}|x)/\{1 - E(Y_{a_1}|x)\}}{E(Y_{a_0}|x)/\{1 - E(Y_{a_0}|x)\}} \\
d_{a_1}^{OR}(x) &= \frac{E(Y|a_1, x)/\{1 - E(Y|a_1, x)\}}{E(Y|a_0, x)/\{1 - E(Y|a_0, x)\}} \Big/ \frac{E(Y_{a_1}|a_1, x)/\{1 - E(Y_{a_1}|a_1, x)\}}{E(Y_{a_0}|a_1, x)/\{1 - E(Y_{a_0}|a_1, x)\}} \\
d_{a_0}^{OR}(x) &= \frac{E(Y|a_1, x)/\{1 - E(Y|a_1, x)\}}{E(Y|a_0, x)/\{1 - E(Y|a_0, x)\}} \Big/ \frac{E(Y_{a_1}|a_0, x)/\{1 - E(Y_{a_1}|a_0, x)\}}{E(Y_{a_0}|a_0, x)/\{1 - E(Y_{a_0}|a_0, x)\}}.
\end{aligned}$$

If the outcome is rare in all strata of a, x and u then the bias formulas for the risk ratio will hold approximately true for the odds ratio. Under the rare outcome assumption we may also replace risk ratios with odds ratios and thus if $Y_a \coprod A|X, U$ and if u' is any chosen reference value for U then the following formulas will hold approximately:

$$\begin{aligned}
d_{a_+}^{OR}(x) &\cong \frac{\sum_u \{\frac{E(Y|a_1,x,u)}{1-E(Y|a_1,x,u)} / \frac{E(Y|a_1,x,u')}{1-E(Y|a_1,x,u')}\} P(u|a_1, x)}{\sum_u \{\frac{E(Y|a_1,x,u)}{1-E(Y|a_1,x,u)} / \frac{E(Y|a_1,x,u')}{1-E(Y|a_1,x,u')}\} P(u|x)} \Big/ \frac{\sum_u \{\frac{E(Y|a_0,x,u)}{1-E(Y|a_0,x,u)} / \frac{E(Y|a_0,x,u')}{1-E(Y|a_0,x,u')}\} P(u|a_0, x)}{\sum_u \{\frac{E(Y|a_0,x,u)}{1-E(Y|a_0,x,u)} / \frac{E(Y|a_0,x,u')}{1-E(Y|a_0,x,u')}\} P(u|x)} \\
d_{a_1}^{OR}(x) &\cong \frac{\sum_u \{\frac{E(Y|a_0,x,u)}{1-E(Y|a_0,x,u)} / \frac{E(Y|a_0,x,u')}{1-E(Y|a_0,x,u')}\} P(u|a_1, x)}{\sum_u \{\frac{E(Y|a_0,x,u)}{1-E(Y|a_0,x,u)} / \frac{E(Y|a_0,x,u')}{1-E(Y|a_0,x,u')}\} P(u|a_0, x)} \\
d_{a_0}^{OR}(x) &\cong \frac{\sum_u \{\frac{E(Y|a_1,x,u)}{1-E(Y|a_1,x,u)} / \frac{E(Y|a_1,x,u')}{1-E(Y|a_1,x,u')}\} P(u|a_1, x)}{\sum_u \{\frac{E(Y|a_1,x,u)}{1-E(Y|a_1,x,u)} / \frac{E(Y|a_1,x,u')}{1-E(Y|a_1,x,u')}\} P(u|a_0, x)}.
\end{aligned}$$

Note that to obtain the approximate bias formulas for the marginal causal odds ratio above, the rare outcome assumption must be invoked; thus if the rare outcome assumption does not hold, the approximations may deviate considerably from the true bias formulas. If the outcome is not rare then we can instead use the following exact result.

Theorem 4: If $Y_a \coprod A|X, U$ then

$$
\begin{aligned}
d_{a_+}^{OR}(x) &= \frac{\frac{\sum_u E(Y|a_1,x,u)P(u|a_1,x)}{\sum_u \{1-E(Y|a_1,x,u)\}P(u|a_1,x)}}{\frac{\sum_u E(Y|a_0,x,u)P(u|a_0,x)}{\sum_u \{1-E(Y|a_0,x,u)\}P(u|a_0,x)}} \Big/ \frac{\frac{\sum_u E(Y|a_1,x,u)P(u|x)}{\sum_u \{1-E(Y|a_1,x,u)\}P(u|x)}}{\frac{\sum_u E(Y|a_0,x,u)P(u|x)}{\sum_u \{1-E(Y|a_0,x,u)\}P(u|x)}} \\
d_{a_1}^{OR}(x) &= \frac{\sum_u E(Y|a_0,x,u)P(u|a_1,x)}{\sum_u \{1-E(Y|a_0,x,u)\}P(u|a_1,x)} \Big/ \frac{\sum_u E(Y|a_0,x,u)P(u|a_0,x)}{\sum_u \{1-E(Y|a_0,x,u)\}P(u|a_0,x)} \\
d_{a_0}^{OR}(x) &= \frac{\sum_u E(Y|a_1,x,u)P(u|a_1,x)}{\sum_u \{1-E(Y|a_1,x,u)\}P(u|a_1,x)} \Big/ \frac{\sum_u E(Y|a_1,x,u)P(u|a_0,x)}{\sum_u \{1-E(Y|a_1,x,u)\}P(u|a_0,x)}.
\end{aligned}
$$

Using these exact expressions in Theorem 4 requires a more involved approach to external adjustment than using the approximate formulas. The exact approach is more involved because we must specify $E(Y|a,x,u)$ for each u. We could do so by obtaining $E(Y|a,x,u')$ for some u' as well as

$$
\frac{E(Y|a,x,u)/\{1-E(Y|a,x,u)\}}{E(Y|a,x,u')/\{1-E(Y|a,x,u')\}}
$$

for each u from external information. From these two external quantities, it is then straightforward to obtain $E(Y|a,x,u)$ for each u. Note that for a particular value x, this exact approach will require one additional quantity, namely $E(Y|a,x,u')$ for some u', as compared with the approximate approach.

The marginal causal odds ratio or "standardized" causal odds ratio in the total population or amongst those receiving treatment a_1 or a_0 are defined respectively by

$$
\begin{gathered}
\frac{E(Y_{a_1})/\{1-E(Y_{a_1})\}}{E(Y_{a_0})/\{1-E(Y_{a_0})\}} \\
\frac{E(Y_{a_1}|a_1)/\{1-E(Y_{a_1}|a_1)\}}{E(Y_{a_0}|a_1)/\{1-E(Y_{a_0}|a_1)\}} \\
\frac{E(Y_{a_1}|a_0)/\{1-E(Y_{a_1}|a_0)\}}{E(Y_{a_0}|a_0)/\{1-E(Y_{a_0}|a_0)\}}.
\end{gathered}
$$

Define the bias expressions $d_{a_+}^{OR}$, $d_{a_1}^{OR}$ and $d_{a_0}^{OR}$ as follows:

$$
\begin{aligned}
d_{a_+}^{OR} &= \frac{\sum_x E(Y|a_1,x)P(x)/\{1-\sum_x E(Y|a_1,x)P(x)\}}{\sum_x E(Y|a_0,x)P(x)/\{1-\sum_x E(Y|a_0,x)P(x)\}} \Big/ \frac{E(Y_{a_1})/\{1-E(Y_{a_1})\}}{E(Y_{a_0})/\{1-E(Y_{a_0})\}} \\
d_{a_1}^{OR} &= \frac{\sum_x E(Y|a_1,x)P(x|a_1)/\{1-\sum_x E(Y|a_1,x)P(x|a_1)\}}{\sum_x E(Y|a_0,x)P(x|a_1)/\{1-\sum_x E(Y|a_0,x)P(x|a_1)\}} \Big/ \frac{E(Y_{a_1}|a_1)/\{1-E(Y_{a_1}|a_1)\}}{E(Y_{a_0}|a_1)/\{1-E(Y_{a_0}|a_1)\}} \\
d_{a_0}^{OR} &= \frac{\sum_x E(Y|a_1,x)P(x|a_0)/\{1-\sum_x E(Y|a_1,x)P(x|a_0)\}}{\sum_x E(Y|a_0,x)P(x|a_0)/\{1-\sum_x E(Y|a_0,x)P(x|a_0)\}} \Big/ \frac{E(Y_{a_1}|a_0)/\{1-E(Y_{a_1}|a_0)\}}{E(Y_{a_0}|a_0)/\{1-E(Y_{a_0}|a_0)\}}.
\end{aligned}
$$

If the outcome is rare in all strata of a, x and u then the bias formulas for the marginal causal risk ratios will hold approximately true for the marginal causal odds ratios. Under the rare outcome assumption, we may also replace risk ratios with odds ratios and thus if $Y_a \coprod A|X, U$, and if u' is any chosen reference value for U, and if x' is any chosen reference value for X, then the following formulas will hold approximately:

$$\begin{aligned}
d_{a_+}^{OR} &\cong \frac{\sum_x \frac{E(Y|a_1,x)/\{1-E(Y|a_1,x)\}}{E(Y|a_1,x')/\{1-E(Y|a_1,x')\}}P(x)}{\sum_x r_1^{OR}(x)^{-1}\frac{E(Y|a_1,x)/\{1-E(Y|a_1,x)\}}{E(Y|a_1,x')/\{1-E(Y|a_1,x')\}}P(x)} \Big/ \frac{\sum_x \frac{E(Y|a_0,x)/\{1-E(Y|a_0,x)\}}{E(Y|a_0,x')/\{1-E(Y|a_0,x')\}}P(x)}{\sum_x r_0^{OR}(x)^{-1}\frac{E(Y|a_0,x)/\{1-E(Y|a_0,x)\}}{E(Y|a_0,x')/\{1-E(Y|a_0,x')\}}P(x)} \\
d_{a_1}^{OR} &\cong \frac{\sum_x d_{a_1}^{OR}(x)\frac{E(Y|a_0,x)/\{1-E(Y|a_0,x)\}}{E(Y|a_0,x')/\{1-E(Y|a_0,x')\}}P(x|a_1)}{\sum_x \frac{E(Y|a_0,x)/\{1-E(Y|a_0,x)\}}{E(Y|a_0,x')/\{1-E(Y|a_0,x')\}}P(x|a_1)} \\
d_{a_0}^{OR} &\cong \frac{\sum_x \frac{E(Y|a_1,x)/\{1-E(Y|a_1,x)\}}{E(Y|a_1,x')/\{1-E(Y|a_1,x')\}}P(x|a_0)}{\sum_x d_{a_0}^{OR}(x)^{-1}\frac{E(Y|a_1,x)/\{1-E(Y|a_1,x)\}}{E(Y|a_1,x')/\{1-E(Y|a_1,x')\}}P(x|a_0)}.
\end{aligned}$$

where

$$\begin{aligned}
r_1^{OR}(x) &= \frac{\sum_u \frac{E(Y|a_1,x,u)/\{1-E(Y|a_1,x,u)\}}{E(Y|a_1,x,u')/\{1-E(Y|a_1,x,u')\}}P(u|a_1,x)}{\sum_u \frac{E(Y|a_1,x,u)/\{1-E(Y|a_1,x,u)\}}{E(Y|a_1,x,u')/\{1-E(Y|a_1,x,u')\}}P(u|x)} \\
r_0^{OR}(x) &= \frac{\sum_u \frac{E(Y|a_0,x,u)/\{1-E(Y|a_0,x,u)\}}{E(Y|a_0,x,u')/\{1-E(Y|a_0,x,u')\}}P(u|a_0,x)}{\sum_u \frac{E(Y|a_0,x,u)/\{1-E(Y|a_0,x,u)\}}{E(Y|a_0,x,u')/\{1-E(Y|a_0,x,u')\}}P(u|x)}
\end{aligned}$$

and $d_{a_1}^{OR}(x)$ and $d_{a_0}^{OR}(x)$ are the bias formulas for the conditional odds ratios given in section 3 of the online appendix.

The expressions of the form

$$\frac{E(Y|a,x,u)/\{1-E(Y|a,x,u)\}}{E(Y|a,x,u')/\{1-E(Y|a,x,u')\}}$$

and $P(u|a,x)$ could be obtained through external information; the expressions of the form

$$\frac{E(Y|a,x)/\{1-E(Y|a,x)\}}{E(Y|a,x')/\{1-E(Y|a,x')\}}$$

can be obtained from the data.

Note that to obtain the approximate bias formulas for the marginal causal odds ratio above, the rare outcome assumption must be invoked repeatedly; thus if the rare outcome assumption does not hold the approximations may deviate considerably from the true bias formulas. If the rare outcome assumption does not hold then we can instead use the following exact result.

Theorem 5: If $Y_a \coprod A|X, U$ then

$$
\begin{aligned}
d_{a_+}^{OR} &= \frac{\frac{\sum_x \sum_u E(Y|a_1,x,u)P(u|a_1,x)P(x)}{\sum_x \sum_u \{1-E(Y|a_1,x,u)\}P(u|a_1,x)P(x)}}{\frac{\sum_x \sum_u E(Y|a_0,x,u)P(u|a_0,x)P(x)}{\sum_x \sum_u \{1-E(Y|a_0,x,u)\}P(u|a_0,x)P(x)}} \Big/ \frac{\frac{\sum_x \sum_u E(Y|a_1,x,u)P(u|x)P(x)}{\sum_x \sum_u \{1-E(Y|a_1,x,u)\}P(u|x)P(x)}}{\frac{\sum_x \sum_u E(Y|a_0,x,u)P(u|x)P(x)}{\sum_x \sum_u \{1-E(Y|a_0,x,u)\}P(u|x)P(x)}} \\
d_{a_1}^{OR} &= \frac{\frac{\sum_x \sum_u E(Y|a_0,x,u)P(u|a_1,x)P(x|a_1)}{\sum_x \sum_u \{1-E(Y|a_0,x,u)\}P(u|a_1,x)P(x|a_1)}}{\frac{\sum_x \sum_u E(Y|a_0,x,u)P(u|a_0,x)P(x|a_1)}{\sum_x \sum_u \{1-E(Y|a_0,x,u)\}P(u|a_0,x)P(x|a_1)}} \\
d_{a_0}^{OR} &= \frac{\frac{\sum_x \sum_u E(Y|a_1,x,u)P(u|a_1,x)P(x|a_0)}{\sum_x \sum_u \{1-E(Y|a_1,x,u)\}P(u|a_1,x)P(x|a_0)}}{\frac{\sum_x \sum_u E(Y|a_1,x,u)P(u|a_0,x)P(x|a_0)}{\sum_x \sum_u \{1-E(Y|a_1,x,u)\}P(u|a_0,x)P(x|a_0)}}.
\end{aligned}
$$

Using the exact expression in Theorem 5 requires a more involved approach to external adjustment than using the approximate formulas. The exact approach is more involved because we must specify every $E(Y|a, x, u)$ for each x and u. We could do so by obtaining all the odds ratios

$$\frac{E(Y|a, x, u)/\{1 - E(Y|a, x, u)\}}{E(Y|a, x, u')/\{1 - E(Y|a, x, u')\}}$$

as well as $E(Y|a, x, u')$ from external information. From these external quantities, it is then straightforward to obtain $E(Y|a, x, u)$ for each x and u. Note that this exact approach will require not only each odds ratio

$$\frac{E(Y|a, x, u)/\{1 - E(Y|a, x, u)\}}{E(Y|a, x, u')/\{1 - E(Y|a, x, u')\}}$$

as in the approximate formulas but also $E(Y|a, x, u')$ for some u' and each x.

3. Complete Proofs of Theorems 1-5.

Proof of Theorem 1. We have that

$$
\begin{aligned}
d_{a_1} &= \textstyle\sum_x \{E(Y|a_1, x) - E(Y|a_0, x)\}P(x|a_1) - \{E(Y_{a_1}|a_1) - E(Y_{a_0}|a_1)\} \\
&= \textstyle\sum_x \sum_u E(Y|a_1, x, u)P(u|a_1, x)P(x|a_1) - \sum_x \sum_u E(Y|a_0, x, u)P(u|a_0, x)P(x|a_1) \\
&\quad \textstyle - \sum_x \sum_u E(Y_{a_1}|a_1, x, u)P(u|a_1, x)P(x|a_1) + \sum_x \sum_u E(Y_{a_0}|a_1, x, u)P(u|a_1, x)P(x|a_1) \\
&= \textstyle\sum_x \sum_u E(Y_{a_0}|a_1, x, u)P(u|a_1, x)P(x|a_1) - \sum_x \sum_u E(Y|a_0, x, u)P(u|a_0, x)P(x|a_1) \\
&= \textstyle\sum_x \sum_u E(Y|a_0, x, u)P(u|a_1, x)P(x|a_1) - \sum_x \sum_u E(Y|a_0, x, u)P(u|a_0, x)P(x|a_1) \\
&= \textstyle\sum_x \sum_u E(Y|a_0, x, u)\{P(u|a_1, x) - P(u|a_0, x)\}P(x|a_1) \\
&= \textstyle\sum_x \sum_u \{E(Y|a_0, x, u) - E(Y|a_0, x, u')\}\{P(u|a_1, x) - P(u|a_0, x)\}P(x|a_1).
\end{aligned}
$$

The proof for d_{a_0} is similar. As noted in the paper, for d_{a_+} we have that

$$\begin{aligned}
d_{a_+} &= \{\textstyle\sum_x E(Y|a_1,x)P(x) - \sum_x E(Y|a_0,x)P(x)\} - \{E(Y_{a_1}) - E(Y_{a_0})\} \\
&= \textstyle\sum_x \sum_u E(Y|a_1,x,u)P(u|a_1,x)P(x) - \sum_x \sum_u E(Y|a_0,x,u)P(u|a_0,x)P(x) \\
&\quad - \textstyle\sum_x \sum_u E(Y_{a_1}|x,u)P(u|x)P(x) + \sum_x \sum_u E(Y_{a_0}|x,u)P(u|x)P(x) \\
&= \textstyle\sum_x \sum_u E(Y|a_1,x,u)P(u|a_1,x)P(x) - \sum_x \sum_u E(Y|a_0,x,u)P(u|a_0,x)P(x) \\
&\quad - \textstyle\sum_x \sum_u E(Y_{a_1}|a_1,x,u)P(u|x)P(x) + \sum_x \sum_u E(Y_{a_0}|a_0,x,u)P(u|x)P(x) \\
&= \textstyle\sum_x \sum_u E(Y|a_1,x,u)\{P(u|a_1,x) - P(u|x)\}P(x) \\
&\quad - \textstyle\sum_x \sum_u E(Y|a_0,x,u)\{P(u|a_0,x) - P(u|x)\}P(x) \\
&= \textstyle\sum_x \sum_u \{E(Y|a_1,x,u) - E(Y|a_1,x,u')\}\{P(u|a_1,x) - P(u|x)\}P(x) \\
&\quad - \textstyle\sum_x \sum_u \{E(Y|a_0,x,u) - E(Y|a_0,x,u')\}\{P(u|a_0,x) - P(u|x)\}P(x). \blacksquare
\end{aligned}$$

Proof of Theorem 2. We have that

$$\begin{aligned}
d^{RR}_{a_1}(x) &= \frac{E(Y|a_1,x)/E(Y|a_0,x)}{E(Y_{a_1}|a_1,x)/E(Y_{a_0}|a_1,x)} \\
&= \frac{E(Y_{a_0}|a_1,x)}{E(Y|a_0,x)} \\
&= \frac{\sum_u E(Y_{a_0}|a_1,x,u)P(u|a_1,x)}{\sum_u E(Y|a_0,x,u)P(u|a_0,x)} \\
&= \frac{\sum_u E(Y|a_0,x,u)P(u|a_1,x)}{\sum_u E(Y|a_0,x,u)P(u|a_0,x)} \\
&= \frac{\sum_u \frac{E(Y|a_0,x,u)}{E(Y|a_0,x,u')}P(u|a_1,x)}{\sum_u \frac{E(Y|a_0,x,u)}{E(Y|a_0,x,u')}P(u|a_0,x)}.
\end{aligned}$$

The proof for $d^{RR}_{a_0}(x)$ is similar. For $d^{RR}_{a_+}(x)$ we have that

$$\begin{aligned}
d^{RR}_{a_+}(x) &= \frac{E(Y|a_1,x)/E(Y|a_0,x)}{E(Y_{a_1}|x)/E(Y_{a_0}|x)} \\
&= \frac{\sum_u E(Y|a_1,x,u)P(u|a_1,x)}{\sum_u E(Y_{a_1}|x,u)P(u|x)} \Big/ \frac{\sum_u E(Y|a_0,x,u)P(u|a_0,x)}{\sum_u E(Y_{a_0}|x,u)P(u|x)} \\
&= \frac{\sum_u E(Y|a_1,x,u)P(u|a_1,x)}{\sum_u E(Y|a_1,x,u)P(u|x)} \Big/ \frac{\sum_u E(Y|a_0,x,u)P(u|a_0,x)}{\sum_u E(Y|a_0,x,u)P(u|x)} \\
&= \frac{\sum_u \frac{E(Y|a_1,x,u)}{E(Y|a_1,x,u')}P(u|a_1,x)}{\sum_u \frac{E(Y|a_1,x,u)}{E(Y|a_1,x,u')}P(u|x)} \Big/ \frac{\sum_u \frac{E(Y|a_0,x,u)}{E(Y|a_0,x,u')}P(u|a_0,x)}{\sum_u \frac{E(Y|a_0,x,u)}{E(Y|a_0,x,u')}P(u|x)}. \blacksquare
\end{aligned}$$

Proof of Theorem 3. We have that

$$\begin{aligned}
d^{RR}_{a_1} &= \frac{\sum_x E(Y|a_1,x)P(x|a_1)/\sum_x E(Y|a_0,x)P(x|a_1)}{E(Y_{a_1}|a_1)/E(Y_{a_0}|a_1)} \\
&= \frac{E(Y_{a_0}|a_1)}{\sum_x E(Y|a_0,x)P(x|a_1)} \\
&= \frac{\sum_x E(Y_{a_0}|a_1,x)P(x|a_1)}{\sum_x E(Y|a_0,x)P(x|a_1)} \\
&= \frac{\sum_x \frac{E(Y|a_1,x)/E(Y|a_0,x)}{E(Y_{a_1}|a_1,x)/E(Y_{a_0}|a_1,x)} E(Y|a_0,x)P(x|a_1)}{\sum_x E(Y|a_0,x)P(x|a_1)} \\
&= \frac{\sum_x d^{RR}_{a_1}(x)E(Y|a_0,x)P(x|a_1)}{\sum_x E(Y|a_0,x)P(x|a_1)} \\
&= \frac{\sum_x d^{RR}_{a_1}(x)\frac{E(Y|a_0,x)}{E(Y|a_0,x')}P(x|a_1)}{\sum_x \frac{E(Y|a_0,x)}{E(Y|a_0,x')}P(x|a_1)}.
\end{aligned}$$

The proof for $d^{RR}_{a_0}$ is similar. For $d^{RR}_{a_+}$ we have that

$$\begin{aligned}
d^{RR}_{a_+} &= \frac{\sum_x E(Y|a_1,x)P(x)/\sum_x E(Y|a_0,x)P(x)}{E(Y_{a_1})/E(Y_{a_0})} \\
&= \frac{\sum_x E(Y|a_1,x)P(x)}{\sum_x\sum_u E(Y|a_1,x,u)P(u|x)P(x)} \Big/ \frac{\sum_x E(Y|a_0,x)P(x)}{\sum_x\sum_u E(Y|a_0,x,u)P(u|x)P(x)} \\
&= \frac{\sum_x E(Y|a_1,x)P(x)}{\sum_x \frac{\sum_u E(Y|a_1,x,u)P(u|x)}{E(Y|a_1,x)}E(Y|a_1,x)P(x)} \Big/ \frac{\sum_x E(Y|a_0,x)P(x)}{\sum_x \frac{\sum_u E(Y|a_0,x,u)P(u|x)}{E(Y|a_0,x)}E(Y|a_0,x)P(x)} \\
&= \frac{\sum_x E(Y|a_1,x)P(x)}{\sum_x \frac{\sum_u E(Y|a_1,x,u)P(u|x)}{\sum_u E(Y|a_1,x,u)P(u|a_1,x)}E(Y|a_1,x)P(x)} \Big/ \frac{\sum_x E(Y|a_0,x)P(x)}{\sum_x \frac{\sum_u E(Y|a_0,x,u)P(u|x)}{\sum_u E(Y|a_0,x,u)P(u|a_0,x)}E(Y|a_0,x)P(x)} \\
&= \frac{\sum_x E(Y|a_1,x)P(x)}{\sum_x \frac{\sum_u \frac{E(Y|a_1,x,u)}{E(Y|a_1,x,u')}P(u|x)}{\sum_u \frac{E(Y|a_1,x,u)}{E(Y|a_1,x,u')}P(u|a_1,x)}E(Y|a_1,x)P(x)} \Big/ \frac{\sum_x E(Y|a_0,x)P(x)}{\sum_x \frac{\sum_u \frac{E(Y|a_0,x,u)}{E(Y|a_0,x,u')}P(u|x)}{\sum_u \frac{E(Y|a_0,x,u)}{E(Y|a_0,x,u')}P(u|a_0,x)}E(Y|a_0,x)P(x)} \\
&= \frac{\sum_x E(Y|a_1,x)P(x)}{\sum_x r_1(x)^{-1}E(Y|a_1,x)P(x)} \Big/ \frac{\sum_x E(Y|a_0,x)P(x)}{\sum_x r_0(x)^{-1}E(Y|a_0,x)P(x)} \\
&= \frac{\sum_x \frac{E(Y|a_1,x)}{E(Y|a_1,x')}P(x)}{\sum_x r_1(x)^{-1}\frac{E(Y|a_1,x)}{E(Y|a_1,x')}P(x)} \Big/ \frac{\sum_x \frac{E(Y|a_0,x)}{E(Y|a_0,x')}P(x)}{\sum_x r_0(x)^{-1}\frac{E(Y|a_0,x)}{E(Y|a_0,x')}P(x)}. \blacksquare
\end{aligned}$$

Proof of Thoerem 4. We have that

$$\begin{aligned}
d^{OR}_{a_1}(x) &= \frac{E(Y|a_1,x)/\{1-E(Y|a_1,x)\}}{E(Y|a_0,x)/\{1-E(Y|a_0,x)\}} \Big/ \frac{E(Y_{a_1}|a_1,x)/\{1-E(Y_{a_1}|a_1,x)\}}{E(Y_{a_0}|a_1,x)/\{1-E(Y_{a_0}|a_1,x)\}} \\
&= \frac{E(Y_{a_0}|a_1,x)/\{1-E(Y_{a_0}|a_1,x)\}}{E(Y|a_0,x)/\{1-E(Y|a_0,x)\}} \\
&= \frac{\sum_u E(Y_{a_0}|a_1,x,u)P(u|a_1,x)}{1-\sum_u E(Y_{a_0}|a_1,x,u)P(u|a_1,x)} \Big/ \frac{\sum_u E(Y|a_0,x,u)P(u|a_0,x)}{1-\sum_u E(Y|a_0,x,u)P(u|a_0,x)} \\
&= \frac{\sum_u E(Y|a_0,x,u)P(u|a_1,x)}{\sum_u\{1-E(Y|a_0,x,u)\}P(u|a_1,x)} \Big/ \frac{\sum_u E(Y|a_0,x,u)P(u|a_0,x)}{\sum_u\{1-E(Y|a_0,x,u)\}P(u|a_0,x)}.
\end{aligned}$$

The proof for $d^{OR}_{a_0}(x)$ is similar. For $d^{OR}_{a_+}(x)$ we have that

$$\begin{aligned}
d^{OR}_{a_+}(x) &= \frac{E(Y|a_1,x)/\{1-E(Y|a_1,x)\}}{E(Y|a_0,x)/\{1-E(Y|a_0,x)\}} \Big/ \frac{E(Y_{a_1}|x)/\{1-E(Y_{a_1}|x)\}}{E(Y_{a_0}|x)/\{1-E(Y_{a_0}|x)\}} \\
&= \frac{\frac{\sum_u E(Y|a_1,x,u)P(u|a_1,x)}{1-\sum_u E(Y|a_1,x,u)P(u|a_1,x)}}{\frac{\sum_u E(Y|a_0,x,u)P(u|a_0,x)}{1-\sum_u E(Y|a_0,x,u)P(u|a_0,x)}} \Big/ \frac{\frac{\sum_u E(Y_{a_1}|x,u)P(u|x)}{1-\sum_u E(Y_{a_1}|x,u)P(u|x)}}{\frac{\sum_u E(Y_{a_0}|x,u)P(u|x)}{1-\sum_u E(Y_{a_0}|x,u)P(u|x)}} \\
&= \frac{\frac{\sum_u E(Y|a_1,x,u)P(u|a_1,x)}{1-\sum_u E(Y|a_1,x,u)P(u|a_1,x)}}{\frac{\sum_u E(Y|a_0,x,u)P(u|a_0,x)}{1-\sum_u E(Y|a_0,x,u)P(u|a_0,x)}} \Big/ \frac{\frac{\sum_u E(Y|a_1,x,u)P(u|x)}{1-\sum_u E(Y|a_1,x,u)P(u|x)}}{\frac{\sum_u E(Y|a_0,x,u)P(u|x)}{1-\sum_u E(Y|a_0,x,u)P(u|x)}} \\
&= \frac{\frac{\sum_u E(Y|a_1,x,u)P(u|a_1,x)}{\sum_u \{1-E(Y|a_1,x,u)\}P(u|a_1,x)}}{\frac{\sum_u E(Y|a_0,x,u)P(u|a_0,x)}{\sum_u \{1-E(Y|a_0,x,u)\}P(u|a_0,x)}} \Big/ \frac{\frac{\sum_u E(Y|a_1,x,u)P(u|x)}{\sum_u \{1-E(Y|a_1,x,u)\}P(u|x)}}{\frac{\sum_u E(Y|a_0,x,u)P(u|x)}{\sum_u \{1-E(Y|a_0,x,u)\}P(u|x)}}. \blacksquare
\end{aligned}$$

Proof of Theorem 5. We have that

$$\begin{aligned}
d^{OR}_{a_1} &= \frac{\sum_x E(Y|a_1,x)P(x|a_1)/\{1-\sum_x E(Y|a_1,x)P(x|a_1)\}}{\sum_x E(Y|a_0,x)P(x|a_1)/\{1-\sum_x E(Y|a_0,x)P(x|a_1)\}} \Big/ \frac{E(Y_{a_1}|a_1)/\{1-E(Y_{a_1}|a_1)\}}{E(Y_{a_0}|a_1)/\{1-E(Y_{a_0}|a_1)\}} \\
&= \frac{E(Y_{a_0}|a_1)/\{1-E(Y_{a_0}|a_1)\}}{\sum_x E(Y|a_0,x)P(x|a_1)/\{1-\sum_x E(Y|a_0,x)P(x|a_1)\}} \\
&= \frac{\sum_x\sum_u E(Y_{a_0}|a_1,x,u)P(u|a_1,x)P(x|a_1)}{1-\sum_x\sum_u E(Y_{a_0}|a_1,x,u)P(u|a_1,x)P(x|a_1)} \Big/ \frac{\sum_x\sum_u E(Y|a_0,x,u)P(u|a_0,x)P(x|a_1)}{1-\sum_x\sum_u E(Y|a_0,x,u)P(u|a_0,x)P(x|a_1)} \\
&= \frac{\sum_x\sum_u E(Y|a_0,x,u)P(u|a_1,x)P(x|a_1)}{\sum_x\sum_u \{1-E(Y|a_0,x,u)\}P(u|a_1,x)P(x|a_1)} \Big/ \frac{\sum_x\sum_u E(Y|a_0,x,u)P(u|a_0,x)P(x|a_1)}{\sum_x\sum_u \{1-E(Y|a_0,x,u)\}P(u|a_0,x)P(x|a_1)}.
\end{aligned}$$

The proof for $d^{OR}_{a_0}$ is similar. For $d^{OR}_{a_+}$ we have that

$$\begin{aligned}
d^{OR}_{a_+} &= \frac{\sum_x E(Y|a_1,x)P(x)/\{1-\sum_x E(Y|a_1,x)P(x)\}}{\sum_x E(Y|a_0,x)P(x)/\{1-\sum_x E(Y|a_0,x)P(x)\}} \Big/ \frac{E(Y_{a_1})/\{1-E(Y_{a_1})\}}{E(Y_{a_0})/\{1-E(Y_{a_0})\}} \\
&= \frac{\frac{\sum_x\sum_u E(Y|a_1,x,u)P(u|a_1,x)P(x)}{1-\sum_x\sum_u E(Y|a_1,x,u)P(u|a_1,x)P(x)}}{\frac{\sum_x\sum_u E(Y|a_0,x,u)P(u|a_0,x)P(x)}{1-\sum_x\sum_u E(Y|a_0,x,u)P(u|a_0,x)P(x)}} \Big/ \frac{\frac{\sum_x\sum_u E(Y_{a_1}|x,u)P(u|x)P(x)}{1-\sum_x\sum_u E(Y_{a_1}|x,u)P(u|x)P(x)}}{\frac{\sum_x\sum_u E(Y_{a_0}|x,u)P(u|x)P(x)}{1-\sum_x\sum_u E(Y_{a_0}|x,u)P(u|x)P(x)}} \\
&= \frac{\frac{\sum_x\sum_u E(Y|a_1,x,u)P(u|a_1,x)P(x)}{1-\sum_x\sum_u E(Y|a_1,x,u)P(u|a_1,x)P(x)}}{\frac{\sum_x\sum_u E(Y|a_0,x,u)P(u|a_0,x)P(x)}{1-\sum_x\sum_u E(Y|a_0,x,u)P(u|a_0,x)P(x)}} \Big/ \frac{\frac{\sum_x\sum_u E(Y|a_1,x,u)P(u|x)P(x)}{1-\sum_x\sum_u E(Y|a_1,x,u)P(u|x)P(x)}}{\frac{\sum_x\sum_u E(Y|a_0,x,u)P(u|x)P(x)}{1-\sum_x\sum_u E(Y|a_0,x,u)P(u|x)P(x)}} \\
&= \frac{\frac{\sum_x\sum_u E(Y|a_1,x,u)P(u|a_1,x)P(x)}{\sum_x\sum_u \{1-E(Y|a_1,x,u)\}P(u|a_1,x)P(x)}}{\frac{\sum_x\sum_u E(Y|a_0,x,u)P(u|a_0,x)P(x)}{\sum_x\sum_u \{1-E(Y|a_0,x,u)\}P(u|a_0,x)P(x)}} \Big/ \frac{\frac{\sum_x\sum_u E(Y|a_1,x,u)P(u|x)P(x)}{\sum_x\sum_u \{1-E(Y|a_1,x,u)\}P(u|x)P(x)}}{\frac{\sum_x\sum_u E(Y|a_0,x,u)P(u|x)P(x)}{\sum_x\sum_u \{1-E(Y|a_0,x,u)\}P(u|x)P(x)}}. \blacksquare
\end{aligned}$$

4. Application of Sensitivity Analysis Approach to Doubly Robust Estimation

To further demonstrate the flexibility of our approach, we apply it to a recent study[40] employing a doubly robust estimator for treatment effects. There are multiple ways to implement doubly robust estimators[41–44] for a binary treatment A and continuous outcome Y but essentially all methods specify models for $E(Y|a, x)$ and for $E(A|x)$ and will give consistent estimates for the adjusted mean outcome difference $\int_x \{E(Y|a_1, x) - E(Y|a_0, x)\}dP(x)$ provided at least of one $E(Y|a, x)$ or $E(A|x)$ are correctly specified. Lambert and Pregibon[40] consider the effect of treatment A, the introduction of a new online advertising feature by an online search system, on change in log (base 2) spending by advertisers eight weeks later, Y. The new feature is not offered to all advertisers; however, the authors control for covariates X including the length of time the advertiser has been a customer, the way by which the advertiser became a customer, country, and baseline measures of spending, number of reports requested by the advertiser, mean number of ad impressions shown per day, ratio of mean daily clicks to ads, variance in daily spending, and mean spending per 1000 impressions. The analysis is stratified by service level for the advertiser (tier 1 versus non-tier 1). They implement a doubly robust estimator for the effect of the new feature on change in log spending using estimators from Robins et al.[41] and, for non-tier 1 advertisers, obtain an estimate of 1.09 (95% CI: $0.33, 1.85$) i.e. a spending ratio of 2.13 ($= 2^{1.09}$) for those offered the new feature.

We can apply the sensitivity analysis technique of the paper to the estimate obtained through the doubly robust procedure. We might hypothesize an unmeasured confounding variable U denoting whether or not anyone working for the advertiser has a personal friend working for the online search system; such a variable was not controlled for in the analysis and could affect both spending and whether or not the advertiser has access to the new feature. Using the simple sensitivity analysis technique described in the paper we can see that to completely eliminate the effect estimate one could hypothesize, for example, that the likelihood of there being a personal friend is 75 percentage points higher for those advertisers given access to the new feature (e.g. 85% versus 10%) across strata of x, $P(U = 1|a_1, x) - P(U = 1|a_0, x) = 0.75$, and that the effect of having a personal friend on change in log spending over eight weeks was 1.45 across all strata of a, x, $E(Y|a, x, U = 1) - E(Y|a, x, U = 0) = 1.45$ (i.e. a $2^{1.45} = 2.73$ times greater change in spending having a personal friend working for the online search system); if this were the case, the effect estimate would be reduced to 0 (95% CI: $-0.76, 0.76$) since $(0.75)(1.45) = 1.09$. However, because having a personal friend working for the online search system would likely affect baseline spending and since this is controlled for in x, the additional effect on change in spending of having a personal friend at the online search system (i.e. 2.73 times greater change) might be viewed as implausibly high in which case it would be difficult to attribute the entire effect to unmeasured confounding. Lambert and Pregibon also consider retention of the advertiser after eight weeks as another outcome and, using a doubly robust estimator, found that eight week retention rates for non-tier 1 advertisers were 9% higher (95% CI: 1%, 18%) for advertisers given access to the new feature. A similar sensitivity analysis suggests that if the likelihood of there being a personal friend is 75% higher and having such a personal friend increases the likelihood of retention by 12% then this would suffice to explain away the estimated effect since $(0.75)(12\%) = 9\%$. Arguably, the sensitivity analysis parameters needed to explain away the retention result are more plausible than those required to explain away the spending result. In the application of Lambert and Pregibon, sensitivity analysis techniques that rely on a regression[11,14] or on propensity score strata[6] are inapplicable because of the doubly robust estimation approach but the simple sensitivity analysis described in the previous section can still be employed; other sensitivity analysis techniques for doubly robust estimator could alternatively be employed[12] but the implementation of these is less straightforward than the approach discussed above.