Supplementary Materials

Section 1

We extend these methods to the continuous exposure case. We define the potential outcome $Y(X)$ as the outcome under exposure level X, and assume the causal exposure effect of interest is

$$T(X) = g(E(Y(X)|X), E(Y(0)|X)).$$

This parameter relates the outcome for subjects with exposure level X to what their outcomes would be if $X = 0$. Note that this effect depends on the amount of exposure (X), as it should. Confounding is then quantified by

$$\Delta(X) = T(X) - T_m(X),$$

where $T_m(X) = g(E(Y^{obs}|X), E(Y^{obs}|X = 0))$ is the marginal effect of X units of exposure. Similarly, the non-linearity effect is quantified by

$$\Delta_{nl}(X) = T_c(X) - T(X),$$

where $T_c(X) = g(E(Y^{obs}|X, Z), E(Y^{obs}|X = 0, Z))$ is the effect of X units of exposure conditional on covariate value Z. Note that the conditional exposure effect is assumed to be constant over Z.

We make the conditional independence assumption that $Y(0)$ is independent of exposure given Z. That is, within a population with fixed Z, the expected outcome absent exposure is not dependent on the actual exposure level. Under this assumption, the causal exposure effect reduces to

$$\begin{aligned}
T(X) &= g(E(Y(X)|X), \int E(Y(0)|X, Z)dF_{Z|X}(z)) \\
&= g(E(Y^{obs}|X), \int E(Y(0)|X=0, Z)dF_{Z|X}(z)) \\
&= g(E(Y^{obs}|X), \int E(Y^{obs}|X=0, Z)dF_{Z|X}(z)),
\end{aligned}$$

which involves only directly observable quantities.

We propose a standardization estimate of confounding, following the methods of Hernan and Robins (2006) and Sato and Matsuyama (2003). Here $E(Y^{obs}|X)$ is estimated from a marginal binary regression model and $E(Y^{obs}|X, Z)$ from a binary regression model that includes Z as a covariate. The distribution of Z conditional on exposure can be estimated using a parametric model. In the exercises below we use logistic regression models and assume that Z is normally distributed conditional on X, where the parameters of the distribution are estimated using a normal linear model.

As in Section 5, we compare the performances of the simple and standardization estimates of confounding using simulations. We simulate a fixed $X \sim N(0, 1)$, and draw 3000 datasets of size 2000 from the following model:

$$\begin{aligned}
Z|X &\sim N(\alpha_0 + \alpha_1 X, 1) \\
\log\text{odds}\, E(Y|X, Z) &= \beta_0 + \beta_1 X + \beta_2 Z. \qquad (1\text{ a})
\end{aligned}$$

We fix the parameters at $\alpha_0 = 0$, $\beta_0 = -3$, and $e^{\beta_1} = 1.5$, a modest conditional exposure effect. The parameters α_1 and β_2 are varied to explore different scenarios. We quantify the causal exposure effect of 1 unit of exposure. The simulation results are shown in Table 1 of the Supplementary Materials. The conclusions are similar to the binary exposure case. The only qualitative difference is that here the corrected estimate of confounding incorporates very little extra variability. This is to be expected, as we have estimated the distribution of Z parametrically.

Section 2

We address the problem of quantifying incremental confounding, the amount by which a factor of interest (Z) confounds the exposure-outcome association over and above known confounders (C). We explore two approaches:

1. Quantify the amount of confounding due to (C, Z), and compare this to the amount of confounding due to C alone.

2. Quantify the amount of confounding due to Z conditional on C, and then marginalize this over C.

We show that these two approaches answer different questions. Consider method (1). The amount of confounding due to (C, Z) is

$$\Delta^{CZ} = g\left(E(Y^{obs}| = 1), \int E(Y^{obs}|X = 0, Z, C)\ dF_{C,Z|X=1}(c, z)\right) \\ - g\left(E(Y^{obs}|X = 1), E(Y^{obs}|X = 0)\right)$$

and the amount of confounding due to C alone is

$$\Delta^{C} = g\left(E(Y^{obs}|X = 1), \int E(Y^{obs}|X = 0, C)\ dF_{C|X=1}(c)\right) \\ - g\left(E(Y^{obs}|X = 1), E(Y^{obs}|X = 0)\right).$$

The measure of incremental confounding according to method (1) is then $\Delta^{CZ} - \Delta^{C}$. According to method (2), the measure of incremental confounding is

$$\Delta^{Z|C} = \int \Big[g\Big(E(Y^{obs}|X=1,C), \int E(Y^{obs}|X=0,C,Z)\, dF_{Z|X=1,C}(z) \Big) - g\big(E(Y^{obs}|X=1,C), E(Y^{obs}|X=0,C) \big) \Big]\, dF_{C|X=1}(c)).$$

If $g(a,b) = a - b$ is the risk difference, $\Delta^{Z|C} = \Delta^{CZ} - \Delta^{Z}$. More generally, however, the two measures differ.

Approaches (1) and (2) address different questions. The first approach quantifies the change in the amount of confounding of the marginal association due to considering Z in addition to C. The second approach quantifies the amount of confounding of the C-conditional association due to Z. The first approach focuses on the marginal association, the second on the association conditional on C.

Both measures (1) and (2) of incremental confounding can be estimated using the standardization or IPW approaches. The distribution of C can be estimated empirically, and bootstrapping used for inference.

Section 3

Suppose the causal exposure effect of interest is the average causal effect (ACE),

$$T = g(E(Y(1)), E(Y(0))),$$

the effect of exposure in the whole population. The magnitude of confounding can then be quantified by

$$\Delta = T - T_m = g(E(Y(1), E(Y(0))) - g(E(Y^{obs}|X=1), E(Y^{obs}|X=0)).$$

Under the assumption that $(Y(1), Y(0))$ is independent of exposure given a covariate Z (a stronger assumption than was used to define confounding for the causal effect among the exposed (Hernan and Robins 2006)), we rewrite Δ as

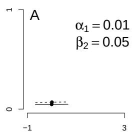
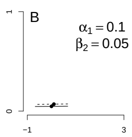
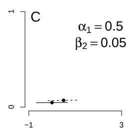
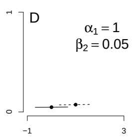
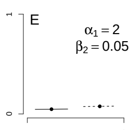
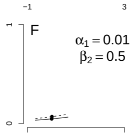
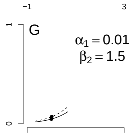
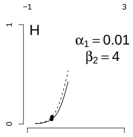
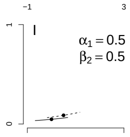
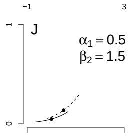
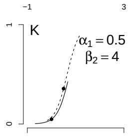
$$
\begin{aligned}
\Delta &= g\left(\int E(Y(1)|Z)dF_Z(z), \int E(Y(0)|Z)dF_Z(z)\right) - g(E(Y^{obs}|X=1), E(Y^{obs}|X=0)) \\
&= g\left(\int E(Y(1)|X=1,Z)dF_Z(z), \int E(Y(0)|X=0,Z)dF_Z(z)\right) - \\
&\qquad g(E(Y^{obs}|X=1), E(Y^{obs}|X=0)) \\
&= g\left(\int E(Y^{obs}|X=1,Z)dF_Z(z), \int E(Y^{obs}|X=0,Z)dF_Z(z)\right) - \\
&\qquad g(E(Y^{obs}|X=1), E(Y^{obs}|X=0)),
\end{aligned}
$$

where the second line follows from the conditional independence assumption. The parameter Δ then can be estimated using the standardization approach, where F_Z is estimated empirically and $E(Y^{obs}|X, Z)$ using binary regression, or using the IPW approach, the fact that

$$
E(Y(0)) = E\left(\frac{I(X=0)Y(0)}{P(X=0|Z)}\right) \text{ and } E(Y(1)) = E\left(\frac{I(X=1)Y(1)}{P(X=1|Z)}\right),
$$

and estimating $P(X|Z)$ using binary regression.

Table 1: 3000 simulations under model (1 a) to evaluate the performance of the simple and standardized estimates of confounding bias for one unit of a continuous exposure, $X \sim N(0,1)$. In each scenario, $E(Y|X,Z)$ is plotted against Z for $X=0$ and $X=1$. In all scenarios, $e^{\beta_1} = 1.5$. The mean Neuhaus et al. (1991) estimate of the non-linearity effect is also shown.

Scenario	$T(1)$	$\Delta(1) \times 10$	$\Delta_{nl}(1) \times 10$	Mean ×10 $\widehat{\Delta}_{si}(1)$	Mean ×10 $\widehat{\Delta}_{st}(1)$	Var × 1000 $\widehat{\Delta}_{si}(1)$	Var × 1000 $\widehat{\Delta}_{st}(1)$	Mean ×10 $\widehat{\Delta}_{nl}(1)$
	0.41	< 0.01	< 0.01	< 0.01	< 0.01	0.01	0.01	< 0.01
	0.41	-0.05	< 0.01	-0.04	-0.05	0.11	0.12	< 0.01
	0.41	-0.25	< 0.01	-0.22	-0.23	2.67	2.69	< 0.01
	0.41	-0.50	< 0.01	-0.45	-0.46	10.60	10.61	< 0.01
	0.41	-1.00	< 0.01	-0.90	-0.90	41.67	41.62	< 0.01
	0.40	-0.05	0.06	0.02	-0.06	0.17	0.16	0.10
	0.33	-0.12	0.81	0.67	-0.16	1.48	1.09	0.80
	0.16	-0.16	2.50	2.51	-0.19	4.54	1.68	2.50
	0.40	-2.46	0.08	-2.38	-2.46	2.12	2.33	0.10
	0.31	-5.94	0.97	-5.04	-6.01	2.12	1.89	0.90
	0.15	-7.52	2.58	-7.86	-7.86	5.84	2.21	2.60