Quantifying Sensitivity to Selection on Unobservables: Refining Oster's Coefficient of Proportionality

Kenneth A. Frank
Qinyun Lin
Spiro Maroulis
With
Shimeng Dai
Nicole Jess
Hung-Chang Lin
Yuqing Liu
Sarah Maestrales
Ellen Searle
Jordan Tait

6/20/2023

This work was supported by U.S. Department of Education Institute for Education Sciences through R305D220022 to Michigan State University. The opinions expressed are those of the authors and do not represent views of the Institute for education Sciences or the U.S. Department. of Education.

Abstract

Beginning with debates about the effects of smoking on lung cancer, sensitivity analyses characterizing the hypothetical unobserved conditions that can alter statistical inferences have had profound impacts on public policy. One of the most ascendant techniques for sensitivity analysis is Oster's (2019) coefficient of proportionality, which approximates how strong selection into a treatment on unobserved variables must be compared to selection on observed variables to change an inference. We refine Oster's asymptotic approximation by deriving expressions for the correlations associated with a latent omitted variable that reduce an estimated effect to a specified threshold, given a corresponding coefficient of determination (R^2). We verify our expressions through empirical examples and simulated data. We show that, because our calculations are exact, they apply regardless of sample size. In contrast, Oster's approximation is likely to overstate robustness when sample size is small and observed covariates account for a large portion of an estimated effect relative to a baseline model. Moreover, even in cases that produce similar values, our correlation-based expressions have the advantage of not depending on the analyst's choice of a baseline model. Our correlation-based expressions can be directly calculated from conventionally reported quantities through commands in R or Stata and an on-line app, and therefore can be applied to most published studies. We present best practices including making maximal use of observed covariates, caution (and an alternative correlation metric) when selection on observables is small and considering a minimum value of the maximal variance to be explained.

Keywords: sensitivity analysis; causal inference; coefficient of proportionality

INTRODUCTION

Cornfield et al. (1959) initiated sensitivity analysis in public policy to interpret inferences regarding the effect of smoking on lung cancer. In the context of lack of randomized experiments many questioned the *effect* of smoking on lung cancer. For example, the famous statistician R. A Fisher (1958) argued "both characteristics [smoking and lung cancer] might be largely influenced by a common cause [genotype]" (page 108). Cornfield et al., countered by calculating that to overturn the inference, an unobserved characteristic "would need to be a near perfect predictor of lung cancer and about nine times more common among smokers than among nonsmokers" (Rosenbaum, 2005, page 1809).

Cornfield et al. (1959) had a profound effect on tobacco policy. Cornfield was an acknowledged contributor to the U.S. Department of Health (1964) report on smoking and lung cancer and Cornfield et al (1959) is cited repeatedly (e.g., pages 141, 183) as a basis of causal inference in the report. The report, in turn, affected public policy concerning regulations of the use, sale, and advertising of tobacco products (Alberg et al., 2014, page 407) in the United States (U.S. Department of Health, 1989) and Britain (Berridge, 2006).

Since Cornfield et al., (1959), sensitivity analyses have proliferated to inform policy-relevant causal inferences in the social, health, and statistical sciences (e.g., Altonji, Elder & Taber, 2005; Cinelli & Haslett, 2020; Dorie et al., 2016; Frank, 2000; Frank et al, 2013; Imbens 2003; Kallus, Mao, & Zhou, 2018; Robins, Rotnitzky & Scharfstein, 2000; Rosenbaum & Rubin, 1983; Vanderweele & Arah, 2011). Perhaps the most ascendant technique for sensitivity analysis is Oster's (2019) coefficient of proportionality (cited 2788 times as of March 17, 2023) which represents how strong selection on unobservable covariates would have to be relative to selection on observed covariates to nullify an estimated effect. In the *Journal of Policy Analysis and*

Management, Oster's (2019) technique has been used to inform inferences regarding effects of emission standards and public transit on infant health (Ngo, 2017), state subsidies on low-income homeownership (Hembre, Moulton & Record, 2021), and temporary mortgage assistance for unemployed homeowners on mortgage default (Moulton et al., 2022). In each case the sensitivity analysis helps those interpreting the inference weigh the strength of the evidence relative to concerns about potentially omitted variables.

In this paper we refine Oster's approximation to the coefficient of proportionality by deriving expressions for correlations associated with a latent omitted variable required to obtain a specific effect size, given a threshold for inference and a targeted final R^2. Our approach has four advantages. First, we improve on Oster's asymptotic derivation by producing the exact specified conditions (e.g., estimated effect and R^2) had the omitted variable been included in the model. Second, we verify our expressions of sensitivity are exact through \empirical examples and simulations, showing that Oster (2019) overstates the robustness of inferences for small samples in which the observed covariates account for a large portion of an initial baseline estimated effect. Third, our approach does not rely on the analyst's choice of the baseline model required by Oster's approach to assess coefficient stability when observed covariates are added to a model. Fourth, our approach supports evaluation of robustness to omitted variables in the metric of a correlation as well as selection on unobservables relative to selection on observables. This is especially valuable when selection on observables is small. Our correlation-based expressions can also be directly calculated from conventionally reported quantities and are available in R and Stata as well as an on-line app, and therefore can be applied to most published studies.

Background

Altonji, Elder and Tabor (2002, 2005) introduced the coefficient of proportionality in the context of a two stage instrumental variables type estimation in which first selection into a treatment is predicted based on covariates (using a probit model) and then the predicted treatment is used to model an outcome. Recognizing that there are likely unobserved factors that predict the treatment that could be related to the outcome, Altonji Elder and Tabor examine a range of scenarios to evaluate how strong selection into the treatment on omitted variables would have to be in the first stage, such that if included in a final model they would nullify the estimated effect in the second stage. For example, Altonji, Elder, and Tabor (2005, page 176) report the estimated ratio of selection on unobservables relative to selection on observables to nullify the effect of attending a Catholic school on college graduation is 1.43. That is, selection on unobservables would have the 40% stronger than selection on observables to nullify the estimated effect. Altonji, Eleder and Tabor (2005) then draw on this result to inform a causal inference: "Since the ratio of selection on unobservables relative to selection on observables is likely less than one, part of the CH [Catholic High School] effect on college graduation is probably real" (pp., 176-177).

Oster (2019) extends Altonji Elder and Tabor in three fundamental ways. First, Oster (2019) considers more fully how covariates that might impact estimated effects should be interpreted in the context of the potential explanatory power (R^2) of any model, including a hypothetical model that included the unobserved covariates. Altonji, Elder and Tabor implicitly assume that all variance on an outcome can potentially be explained (Oster, 2019, page 188). But there may simply be aspects of an outcome that are unexplainable due to essential heterogeneity (Heckman, Urzua, & Vytlacil, 2006). Oster provides empirical validation for choosing the maximum R^2 of less than 1 based on an evaluation of inferences from randomized experiments that would hold for $R^2<1$.

Second, Oster (2019) verifies the derivation of the estimation of an effect based on coefficient stability with simulated data. Altonji Elder and Tabor (2005) had not provided such validation. This is critical to establish under what conditions and what information is necessary to support the derivation; we will replicate this validation below as part of a broader evaluation of Oster's derivation.

Third, Oster (2019) leverages the change in estimated treatment effects and R^2 when observed covariates are added to a model to anticipate the change in the estimated treatment effect if unobserved covariates were added. Oster (2019) then uses this conceptualization to formulate a sensitivity analysis. Oster's work consequently has enabled researchers to characterize the robustness of an estimated effect taking into account the explanatory power available given the context. For example, in the example we present below, Oster calculates that selection on unobservables would have to be greater than selection on observables to reduce the estimated effect of low birth weight and preterm delivery on *IQ* to zero, given a maximum R^2 of .61 in (1c).

While Oster makes important contributions to the conceptualization and application of the coefficient of proportionality it still leaves key limitations. First, while Oster verifies through simulation the estimator for the treatment effect used to calculate the coefficient of proportionality, the expression for the coefficient of proportionality itself is not directly verified. through broad simulation or specific numerical examples Second, the derivation is only approximate and it is not known how the technique performs for small sample sizes. Third, the conceptualization and derivation depend on the analyst's choice of a baseline model used to determine coefficient stability when observed covariates are added.

In this paper we will provide exact expressions to generate a specified estimated effect and corresponding R^2, the two conditions Oster uses to define the coefficient of proportionality. We will verify our exact expressions through a numerical example as well as simulation. We will also show that the exact expressions do not depend on sample size nor on a baseline model (before key observed covariates are added to a model). Furthermore, the expressions can be interpreted in terms of correlations associated with observed and unobserved covariates as well as proportional selection of unobservables to observables. Finally, the expressions derived here can be calculated from conventionally reported quantities and are available through the konfound commands in R and Stata as well as an on-line app (http://konfound-it.com). In the next section we provide the basic set-up required for the derivation.

OSTER'S CONCEPTUALIZATION: SENSITIVITY IN TERMS OF COEFFICIENT STABILITY

Setup and Notation

Following Oster (2019, page 192), consider the models:

$Y = \dot{\beta}_0 + \dot{\beta}_1 X$ (1a, unconditional baseline),

$Y = \tilde{\beta}_0 + \tilde{\beta}_1 X + \tilde{\boldsymbol{\beta}}_2 \mathbf{Z}$ (1b, intermediate with observed covariates, **Z**), and

$Y = \hat{\beta}_0 + \hat{\beta}_1 X + \hat{\boldsymbol{\beta}}_2 \mathbf{Z} + \hat{\beta}_3 CV$ (1c, hypothetical final, with unobserved covariate *CV*),

where *Y* is the outcome (or dependent variable), *X* is the focal predictor (or independent variable), **Z** is an observed vector of covariates added to the unconditional model, typically to control for factors related to both the focal predictor and the outcome, and *CV* is an unobserved covariate added to a hypothetical final model (1c) that was not included in (1b). We will later also consider extensions of this setup in which *CV* is a latent variable representing a set of covariates. For exposition, we focus on the specific concern that a positive $\tilde{\beta}_1$ in (1b) is larger than $\hat{\beta}_1$ in (1c) because the *CV* is omitted from (1b). That is, there is positive bias in $\tilde{\beta}_1$, and therefore the inference about β_1 from (1b) is invalid. Cases in which there is negative bias can be addressed by symmetry.

Leveraging Coefficient Stability to Approximate Conditions when Unobserved Covariates are Added to a Model

To inform inferences about β_1 based on $\dot{\beta}_1$ and $\tilde{\beta}_1$ Oster (2019) describes the common practice of interpreting an inference as robust if the estimated effect changes little from (1a) to (1b) when observed covariates are entered into the model. This is referred to as "coefficient stability" and reflects the practice of examining a model for changes in the estimate of interest as covariates are added, with high stability presumably implying a lower likelihood that unobserved variables would alter that estimate.

Formally, Oster (2019) uses the change in estimated effect from $\dot{\beta}_0$ to $\tilde{\beta}_1$ and the change in explained variance from $\dot{R}^2$ to $\tilde{R}^2$ from (1a) to (1b) to anticipate the conditions necessary to produce $\hat{\beta}_1$ and maximum R^2 in model (1c) when unobserved covariates are added to the model. Oster draws on three quantities to represent coefficient stability: the change in the estimated effect from the baseline model in (1a) to (1b) -- $\left(\dot{\beta}_1 - \tilde{\beta}_1\right)$; the change in R^2 from (1a) to (1b) -- $\left(\tilde{R} - \dot{R}\right)$; and the specified change in R^2 from model (1b) to model (1c) which includes the unobserved confounding variable -- $\left(R_{Max} - \tilde{R}\right)$. Consider (Oster, 2019, page 193):

$$\hat{\beta}_1 = \tilde{\beta}_1 - \delta(\dot{\beta}_1 - \tilde{\beta})\frac{R_{Max} - \tilde{R}}{\tilde{R} - \dot{R}}$$
$$\Rightarrow \frac{\tilde{\beta}_1 - \hat{\beta}_1}{(\dot{\beta}_1 - \tilde{\beta})} = \delta \frac{R_{Max} - \tilde{R}}{\tilde{R} - \dot{R}}. \quad (2)$$

Expression (2) shows how Oster conceptualizes the relative rate of change of the estimated effect as proportional to the relative rate of change in R^2, governed by δ.

Oster (2019, page 193) then makes the assumption that δ =1 indicating that selection on unobservables equals that of selection on observables. This yields the Restricted Estimator:

$$\beta_1^* \approx \tilde{\beta}_1 - \left[\dot{\beta}_1 - \tilde{\beta}_1\right]\left(\frac{R_{Max} - \tilde{R}}{\tilde{R} - \dot{R}}\right). \quad (3)$$

In Appendix A we replicate Oster's validation of the estimator in (3), showing that the estimator is less accurate for small sample sizes. For example, the interquartile range for n=1000 is roughly triple that for n=10000.

Sensitivity: The Approximate Coefficient of Proportionality, δ^*

In the spirit of Cornfield et al. (1959), Oster uses equation (3) to derive a measure of the sensitivity of the estimate and inference for β_1. Specifically, setting $\hat{\beta}_1 = \beta^\#$ (a threshold for evaluating $\hat{\beta}_1$) and solving (3) for δ^* yields

$$\delta^* = \left(\frac{\tilde{\beta}_1 - \beta^\#}{\dot{\beta}_1 - \tilde{\beta}_1}\right)\left(\frac{\tilde{R} - \dot{R}}{R_{Max} - \tilde{R}}\right). \quad (4)$$

Expression (4) allows one to calculate how strong selection on unobservables must be relative to selection on observables (δ^*) to make the estimated effect $\hat{\beta}_1$ in model (1c) equal to a threshold $\beta^\#$ for a specified R^2 (R_{Max}) in model (1c). In this way expression (4) quantifies the robustness of an inference by characterizing the conditions necessary to change that inference. But note that the expression in (4) is based on the derivation and validation for $\hat{\beta}_1$; Oster offers no direct validation or justification for δ^* to which we will turn in subsequent sections.

Consider Oster's empirical example of the effect of *low birth weight and preterm* on *IQ* (Oster's Table 3: Column 3) for illustration throughout this paper. It is the most robust inference (defined by of δ^*) among Oster's examples. The inference is of scientific interest relating to how birth conditions extrapolate throughout the life course (e.g., Breslau et al., 1994). The inference is of public policy interest because if *low birth weight and preterm* have an effect on *IQ* then public policy might attend more fully to the corresponding prenatal medical and social supports (Gross, Spiker & Haynes 1997; National Research Council, 2000).

To obtain δ^*, Oster (2019) used a baseline regression which included controls for *child sex* and *age* dummies (different from model 1a which has no covariates), but without seven key covariates (e.g., *race*, *education*, *income*). For this baseline model $\dot{\beta}_1$ =−.188, and $\dot{R}^2$ =.004. For the intermediate model (1b) with the seven key covariates $\tilde{\beta}_1 = -.125$, $\tilde{R}^2$ = .251. For the final model (1c), Oster specified R_{Max}=.61 and $\beta^{\#}$=0. Therefore, from (3) using $|\dot{\beta}_1|$ and $|\tilde{\beta}_1|$:

$$\delta^* = \left(\frac{\tilde{\beta}_1 - \beta^{\#}}{\dot{\beta}_1 - \tilde{\beta}_1}\right)\left(\frac{\tilde{R} - \dot{R}}{R_{Max} - \tilde{R}}\right) = \left(\frac{.125 - 0}{.188 - .125}\right)\left(\frac{.251 - .004}{.61 - .251}\right) = 1.365,$$

which rounds to 1.37 as reported by Oster (2019). The interpretation is that selection on unobserved variables would have to be 1.37 times greater than selection on observed variables to reduce the estimated effect of *low birth weight and preterm* on *IQ* to 0 under the condition that adding unobserved covariates to the model produces a final R^2 of .61. While the Restricted

Estimator in (4) is valuable for intuition, Oster also presented an Unrestricted Estimator that makes fewer assumptions that we use in our calculations below. [1]

While Oster's coefficient of proportionality has intuitive appeal and is derived from the estimator in (4), it has not been directly verified empirically or through simulation. In the next section we derive exact expressions for the correlations associated with the omitted variable necessary to produce a specified R_{max} and $\beta^{\#}$. We will then verify our expressions are exact through a numerical example and through simulation before using them to evaluate Oster's approximation.

EXACT EXPRESSIONS FOR SENSITIVITY IN TERMS OF CORRELATIONS

We follow other correlation-based approaches to sensitivity analysis (e.g., Cinelli & Hazlett, 2020; Frank, 2000) to derive the exact quantities that generate a specified estimated effect in (1c) and corresponding R^2 in terms of correlations associated with the omitted variable CV. We then use simulated data to verify our exact expressions and evaluate the accuracy of Oster's approximation, showing that Oster's approximation can be considerably inaccurate when $n < 10000$, as is often the case in practice. We also verify our exact expressions through an empirical example.

To derive the exact expressions that generate the specified results, note that Oster's framework implies two conditions. First, that an inference of an effect based on the intermediate regression in (1b) is invalid if $\hat{\beta}_1$ in (1c) falls below some threshold for making an inference, defined as $\beta^{\#}$. Second, that the coefficient of determination (R^2) in (1c) has a maximum value

[1] Adapting code from the R procedure o_delta_rsq_viz generated a value of 1.343 for the Unrestricted Estimator , less than 1% difference from the Restricted Estimator of δ^* of 1.365 based on (4).

$R_{max} = R^2_{Y \cdot X\mathbf{Z}CV}$, the total variance in Y explained by X, **Z**, and CV (this includes Altonji, Elder and Tabor's special case of R_{max} =1). For example, on page 198, Oster finds "…the value of δ that would produce β =0 under the assumed R_{max}…." Oster's two conditions can then be formalized as

$$\hat{\beta}_1 = \beta^{\#}, \qquad (6a)$$

$$R^2_{Y \cdot X\mathbf{Z}CV} = R_{max} < 1. \qquad (6b)$$

To derive a measure of sensitivity, we use the added variable concept (Weisberg, 2005) to rewrite (1c) conditioning each variable on **Z**:

$$Y \mid \mathbf{Z} = \hat{\beta}_0 + \hat{\beta}_1 X \mid \mathbf{Z} + \hat{\beta}_3 CV \mid \mathbf{Z}. \quad (7)$$

When using Ordinary Least Squares (OLS) estimation, as we do throughout the remainder of this paper, the added variable concept ensures that the estimate of β_1 in (7) equals that in model (1c).[2] Specifically, given $\hat{\sigma}_X$ and $\hat{\sigma}_Y$, the estimate for $\hat{\beta}_1$ can be obtained from the partial correlation matrix $\Sigma_{|\mathbf{Z}}$

$$\Sigma_{|\mathbf{Z}} = \begin{bmatrix} & Y \mid \mathbf{Z} & X \mid \mathbf{Z} & CV \mid \mathbf{Z} \\ Y \mid \mathbf{Z} & 1 & & \\ X \mid \mathbf{Z} & r_{X \cdot Y|\mathbf{Z}} & 1 & \\ CV \mid \mathbf{Z} & r_{CV \cdot Y|\mathbf{Z}} & r_{CV \cdot X|\mathbf{Z}} & 1 \end{bmatrix}$$

where r represents a sample correlation between two scalers and "|" represents conditional on. For example, $r_{X \cdot Y|\mathbf{Z}}$ represents the observed correlation between X and Y conditional on, or partialling for, **Z**. Thus, the expression for $\Sigma_{|\mathbf{Z}}$ shows that there are two

---

[2] Expressions in terms of partial correlations allow us to identify the conditions that satisfy (6a) and (6b) in terms of $r_{X \cdot Y|\mathbf{Z}}$ (which can be directly obtained from the reported $\tilde{\beta}_1$, its standard error, and the degrees of freedom) instead of $r_{X \cdot Y}$ which would require further assumptions (see Appendix B for details).

unknowns needed to estimate $\hat{\beta}_1$: $r_{CV \cdot Y|\mathbf{Z}}$ and $r_{CV \cdot X|\mathbf{Z}}$. Below we leverage the two equations in (6a) and (6b) to develop expressions for $r_{CV \cdot Y|\mathbf{Z}}$ and $r_{CV \cdot X|\mathbf{Z}}$ for a given value of $r_{X \cdot Y|\mathbf{Z}}$. These results are then used to directly obtain the coefficient of proportionality. The quantities $R_{X \cdot|\mathbf{Z}}$ and $R_{Y \cdot|\mathbf{Z}}$ will be necessary at several points in the derivation. They can either be directly obtained from the data, or see technical Appendix C for how they can be obtained from conventionally reported quantities: $\text{se}(\tilde{\beta}_1)$, $R^2_{Y \cdot X\mathbf{Z}}$, $\hat{\sigma}_X$, $\hat{\sigma}_Y$ and *n*, and *number of covariates*.

Assumptions

1. Omitted variable is orthogonal to observed covariates

Oster assumes (top of page 192) the observed covariates (Oster's W_1) and the unobserved covariates (Oster's W_2) are orthogonal. This implies that because W_1 is a function of the elements in *Z*, $R_{\mathbf{Z} \cdot CV}$=0 for each element in **Z** (Frank, 2000, page 165, makes the same assumption). Note that setting $R_{\mathbf{Z} \cdot CV}$=0 allows the strongest challenge to the inference of an effect of *X* on *Y* because if $R_{\mathbf{Z} \cdot CV} \neq 0$ then some of the impact of the unobserved covariate *CV* on $\hat{\beta}_1$ would be accounted for by the observed covariates in **Z**, weakening the challenge to the inference based on *CV*. For example, the relationship between *preterm and low birthweight* (*X*) on *IQ* (*Y*) might be challenged in terms of *quality of pre-natal care* (*CV*). But the challenge is weaker if *quality of pre-natal care* is accounted for in part by measured covariates including *income* and *education* (**Z**). Therefore it is conservative to assume the impacts of unobserved covariates are absorbed by the observed covariates (Frank, 2000).

2. Single omitted confounding variable

Oster (2019, pp. 191-192) represents all unobserved covariates with a single index, W_2. Similarly, for derivation we assume a single unobserved covariate *CV*, allowing us to express the coefficient of proportionality by comparing $r_{X \cdot CV}$ directly with $R_{X \cdot \mathbf{Z}}$. In Appendix B we show how a single *CV* can be generalized to a vector **CV** under the assumption that each omitted variable is equally correlated with *X* (i.e.., $r_{X \cdot CV1}=r_{X \cdot CV2} = \ldots = r_{X \cdot CV}$ for all *CV*) and each omitted variable is equally correlated with *Y* (i.e.., $r_{Y \cdot CV1}=r_{Y \cdot CV2} = \ldots = r_{Y \cdot CV}$ for all *CV*). That is, that each omitted variable is equally important regarding the estimation of the effect of *X* on *Y*. This is consistent with the thought experiment in which the elements in *CV* are unknown and therefore cannot easily be differentiated from one another. In this sense the *CV* can be considered a single latent variable that represents equal contributions of an arbitrary number of elements. Even without this assumption our derivation can generalize if the *CV* is considered a single latent variable that represents multiple unobserved variables.

Leveraging Correlations to Generate Exact Conditions when Unobserved Covariates are Added to a Model

Using the model in (7), condition (6a) implies:

$$\hat{\beta}_1 = \frac{\hat{\sigma}_{Y|\mathbf{Z}}}{\hat{\sigma}_{X|\mathbf{Z}}} \frac{r_{X \cdot Y|\mathbf{Z}} - r_{X \cdot CV|\mathbf{Z}} r_{Y \cdot CV|\mathbf{Z}}}{1 - r^2_{X \cdot CV|\mathbf{Z}}} = \beta^{\#}, \quad (8)$$

where $r_{X \cdot Y|\mathbf{Z}}$ is a function of observed quantities: $r_{X \cdot Y|\mathbf{Z}} = \frac{t(\tilde{\beta}_1)}{\sqrt{df + t(\tilde{\beta}_1)^2}}$ with $t(\tilde{\beta}_1) = \frac{\tilde{\beta}_1}{se(\tilde{\beta}_1)}$

and *df*=degrees of freedom=*number of observations*−*number of covariates* in **Z**−3 (acounting for the intercept, *X*, and *CV*). Note also that $\hat{\sigma}_{Y|\mathbf{Z}} = \hat{\sigma}_Y \sqrt{1 - R^2_{Y \cdot \mathbf{Z}}}$, and $\hat{\sigma}_{X|\mathbf{Z}} = \hat{\sigma}_X \sqrt{1 - R^2_{X \cdot \mathbf{Z}}}$. We also note that in (8), for $\tilde{\beta}_1 > \hat{\beta}_1 \geq 0$ (consistent with positive bias in $\tilde{\beta}_1$), the product $r_{X\cdot CV|\mathbf{Z}} r_{Y\cdot CV|\mathbf{Z}}$ must be positive. For situations in which $\tilde{\beta}_1 < \hat{\beta}_1 \leq 0$ (as in the empirical examples) the derivation applies except $r_{X\cdot CV|\mathbf{Z}}$ and $r_{Y\cdot CV|\mathbf{Z}}$ take opposite signs.

To satisfy (6b), the expression for the variance in Y explained by X and CV when all variables are conditioned on **Z** ($R^2_{Y\cdot XCV|\mathbf{Z}}$) associated with the model in (7) is

$$R^2_{Y\cdot XCV|\mathbf{Z}} = \frac{r^2_{X\cdot Y|\mathbf{Z}} + r^2_{Y\cdot CV|\mathbf{Z}} - 2 r_{X\cdot Y|\mathbf{Z}} r_{Y\cdot CV|\mathbf{Z}} r_{X\cdot CV|\mathbf{Z}}}{1 - r^2_{X\cdot CV|\mathbf{Z}}}, \quad (9)$$

where the term $R^2_{Y\cdot XCV|\mathbf{Z}}$ can be obtained from the specified value of R_{MAX} and $R^2_{Y\cdot\mathbf{Z}}$. Specifically, the total variance explained ($R^2_{Y\cdot XCV\mathbf{Z}}$) is the variance explained in Y by **Z** ($R^2_{Y\cdot\mathbf{Z}}$) plus the proportion of variance not explained by **Z** $(1 - R^2_{Y\cdot\mathbf{Z}})$ that is explained by X and CV ($R^2_{Y\cdot XCV|\mathbf{Z}}$):

$$R^2_{Y\cdot XCV\mathbf{Z}} = R_{MAX} = R^2_{Y\cdot\mathbf{Z}} + (1 - R^2_{Y\cdot\mathbf{Z}}) R^2_{Y\cdot XCV|\mathbf{Z}} \Rightarrow R^2_{Y\cdot XCV|\mathbf{Z}} = \frac{R_{MAX} - R^2_{Y\cdot\mathbf{Z}}}{(1 - R^2_{Y\cdot\mathbf{Z}})}. \quad (10)$$

Therefore, for specified values of $\hat{\beta}_1 = \beta^{\#}$ and $R^2_{Y\cdot XCV\mathbf{Z}} = R_{MAX}$ with $r_{X\cdot Y|\mathbf{Z}}$ and $R^2_{Y\cdot\mathbf{Z}}$ obtained from observed data there are two expressions in (6) and two unknowns: $r_{X\cdot CV|\mathbf{Z}}$ and $r_{Y\cdot CV|\mathbf{Z}}$.

To derive expressions for $r_{X\cdot CV|\mathbf{Z}}$ and $r_{Y\cdot CV|\mathbf{Z}}$, we begin by solving (9) for $r_{Y\cdot CV|\mathbf{Z}}$

$$r_{Y\cdot CV|\mathbf{Z}} = r_{X\cdot Y|\mathbf{Z}} r_{X\cdot CV|\mathbf{Z}} + \sqrt{(R^2_{Y\cdot XCV|\mathbf{Z}} - r^2_{X\cdot Y|\mathbf{Z}})(1 - r^2_{X\cdot CV|\mathbf{Z}})}, \quad (11)$$

where the positive of the root is taken because $r_{Y\cdot CV|\mathbf{Z}}$ must be greater than $r_{X\cdot Y|\mathbf{Z}} r_{X\cdot CV|\mathbf{Z}}$ to ensure that $r_{Y\cdot CV|\mathbf{Z}}$ is positive. The expression in (11) reveals that for an observed value of the relationship between X and Y conditional on **Z** ($r_{X\cdot Y|\mathbf{Z}}$), and specified total variance explained $\left(R^2_{Y\cdot XCV\mathbf{Z}}\right)$ which determines $R^2_{Y\cdot XCV|\mathbf{Z}}$, the relationship between the omitted confound and Y conditional on **Z** ($r_{Y\cdot CV|\mathbf{Z}}$) is completely determined by the selection on unobservables ($r_{X\cdot CV|\mathbf{Z}}$).

That is, selection on unobservables implies the specific relationship of *CV* to *Y* to satisfy (6a) and (6b).

Then, to solve for $r_{X\cdot CV|\mathbf{Z}}$, substitute the expression for $r_{Y\cdot CV|\mathbf{Z}}$ from (11) into (8):

$$\hat{\beta}_1 = \frac{\hat{\sigma}_{Y|\mathbf{Z}}}{\hat{\sigma}_{X|\mathbf{Z}}}\left(\frac{r_{X\cdot Y|\mathbf{Z}} - r_{X\cdot CV|\mathbf{Z}}\left(r_{X\cdot Y|\mathbf{Z}} r_{X\cdot CV|\mathbf{Z}} \pm \sqrt{(R^2_{Y\cdot X\cdot CV|\mathbf{Z}} - r^2_{X\cdot Y|\mathbf{Z}})(1-r^2_{X\cdot CV|\mathbf{Z}})}\right)}{1-r^2_{X\cdot CV|\mathbf{Z}}}\right) = \beta^{\#},$$

and solve for $r_{X\cdot CV|\mathbf{Z}}$:

$$r_{X\cdot CV|\mathbf{Z}} = \frac{r_{X\cdot Y|\mathbf{Z}} - \frac{\hat{\sigma}_{X|\mathbf{Z}}}{\hat{\sigma}_{Y|\mathbf{Z}}}\beta^{\#}}{\sqrt{\frac{\hat{\sigma}^2_{X|\mathbf{Z}}}{\hat{\sigma}^2_{Y|\mathbf{Z}}}\beta^{\#2} - 2r_{X\cdot Y|\mathbf{Z}}\frac{\hat{\sigma}_{X|\mathbf{Z}}}{\hat{\sigma}_{Y|\mathbf{Z}}}\beta^{\#} + R^2_{Y\cdot X\cdot CV|\mathbf{Z}}}}, \quad (12)$$

where $r_{X\cdot Y|\mathbf{Z}} - \frac{\hat{\sigma}_{X|\mathbf{Z}}}{\hat{\sigma}_{Y|\mathbf{Z}}}\beta^{\#}$ is taken in the numerator to ensure a positive value of $r_{X\cdot CV|\mathbf{Z}}$ when $r_{X\cdot Y|\mathbf{Z}}$ is greater than $\frac{\hat{\sigma}_{X|\mathbf{Z}}}{\hat{\sigma}_{Y|\mathbf{Z}}}\beta^{\#}$; there is positive bias in $\tilde{\beta}_1$ due to the omission of *CV*. Once $r_{X\cdot CV|\mathbf{Z}}$ is obtained in (12), $r_{Y\cdot CV|\mathbf{Z}}$ can be obtained from (11) (or by solving [8] for $r_{Y\cdot CV|\mathbf{Z}}$).

Thus (11) and (12) allow direct calculation of $r_{X\cdot CV|\mathbf{Z}}$ and $r_{Y\cdot CV|\mathbf{Z}}$ from the quantities $\tilde{\beta}_1$, $\text{se}(\tilde{\beta}_1)$, $R^2_{Y\cdot X\mathbf{Z}}, \hat{\sigma}_X$, $\hat{\sigma}_Y$, *number of covariates*, and *n* as well as the specified quantities $\beta^{\#}$ and R_{MAX} as in (6a) and (6b). Because the observed quantities are conventionally reported, one can apply the expressions to most published studies including any predictor *X* (continuous or binary) used in a linear model such as (1b). Expressions (11) and (12) then generate the exact values of $r_{X\cdot CV|\mathbf{Z}}$ and $r_{Y\cdot CV|\mathbf{Z}}$ that produce the desired $\hat{\beta}_1 = \beta^{\#}$ and $R^2 = R_{Max}$ associated with model (1c).

Sensitivity: The Exact Coefficient of Proportionality, δ

To obtain the coefficient of proportionality from our derivations for $r_{X\cdot CV|\mathbf{Z}}$ we first obtain a general expression for $r_{X\cdot CV}$ for any $\beta^{\#}$ to be compared directly with $R_{X\cdot \mathbf{Z}}$. We show in Appendix D that under the assumption that each element in **Z** is orthogonal to CV and that the elements of **Z** are orthogonal to one another (see Oster, 2019, page 192)

$$r_{X\cdot CV} = \sqrt{1-R^2_{X\cdot\mathbf{Z}}}\,r_{X\cdot CV|\mathbf{Z}} = \frac{\sqrt{1-R^2_{X\cdot\mathbf{Z}}}\left(r_{X\cdot Y|\mathbf{Z}} - \frac{\hat{\sigma}_{X|\mathbf{Z}}}{\hat{\sigma}_{Y|\mathbf{Z}}}\beta^{\#}\right)}{\sqrt{\frac{\hat{\sigma}^2_{X|\mathbf{Z}}}{\hat{\sigma}^2_{Y|\mathbf{Z}}}\beta^{\#2} - 2r_{X\cdot Y|\mathbf{Z}}\frac{\hat{\sigma}_{X|\mathbf{Z}}}{\hat{\sigma}_{Y|\mathbf{Z}}}\beta^{\#} + R^2_{Y\cdot X\cdot CV|\mathbf{Z}}}}, \text{ and (13a)}$$

$$r_{Y\cdot CV} = \sqrt{1-R^2_{Y\cdot\mathbf{Z}}}\,r_{Y\cdot CV|\mathbf{Z}} = \sqrt{1-R^2_{Y\cdot\mathbf{Z}}}\left(r_{X\cdot Y|\mathbf{Z}}r_{X\cdot CV|\mathbf{Z}} \pm \sqrt{(R^2_{Y\cdot XCV|\mathbf{Z}} - r^2_{X\cdot Y|\mathbf{Z}})(1-r^2_{X\cdot CV|\mathbf{Z}})}\right). \quad (13b)$$

Then, using (13a), the coefficient of proportionality, δ, can be defined directly as: [3]

$$\delta = \frac{r_{X\cdot CV}}{R_{X\cdot\mathbf{Z}}}. \quad (14)$$

We now turn to a general verification of our approach and then present a specific example based on Oster (2019).

VERIFICATION OF THE DERIVATIONS

Simulation Verification of the Derivations

Verifying the expressions for $r_{X\cdot CV|Z}$ and $r_{Y\cdot CV|Z}$

[3] Following Cinelli and Hazlett (2020) we define selection by the zero-order correlation associated with the confounding variable: $r_{X\cdot CV}$. The logic is that because $R_{X\cdot\mathbf{Z}}$ is not conditioned on other covariates neither should be $r_{X\cdot CV}$ when compared with $R_{X\cdot\mathbf{Z}}$. This generates a smaller value of δ (conservative) than using $r_{X\cdot CV|\mathbf{Z}}$ which is less than $r_{X\cdot CV}$ as in (13b).

We use simulation to verify the expressions for $r_{Y\cdot CV|\mathbf{Z}}$ and $r_{X\cdot CV\,|\mathbf{Z}}$ in (11) and (12) across a range of scenarios.[4] The following elements were fixed for all scenarios:

- σ_X =.5 (representing scenarios in which half the cases were assigned to a treatment);
- σ_Y=1.5 (i.e., specifically, we chose $\sigma_Y \neq 1$ to demonstrate the results do not depend on standardization);
- n=1000 (representing a moderately large sample but for which δ^* might not yet have achieved asymptotic convergence);
- seven covariates consistent with Oster's Table 3 (*age*, *child female*, *mother Black*, *mother age*, *mother education*, *mother income*, *mother married*).
- Se($\tilde{\beta}_1$) was set such that $R_{X\cdot\mathbf{Z}}$ took a minimum value of .1 occurring when $R^2_{Y\cdot X\mathbf{Z}} = .1$. Specifically, Se($\tilde{\beta}_1$)=0.179 which results in $R_{X\cdot\mathbf{Z}}$ = .1 given σ_X =.5, σ_Y=1.5, and n=1000, see expressions (C2) and (C3);

The key components to vary were the specified threshold value, $\beta^{\#}$; $R^2_{Y\cdot X\mathbf{Z}}$ (Oster's $\tilde{R}$); and $\tilde{\beta}_1$, the estimated effect of X in the intermediate regression in (1b) that does not include the omitted confounding variable (CV);. Specifically, we generated

- threshold values: $\beta^{\#}$=.1 and $\beta^{\#}$=0 (2 conditions);
- $R^2_{Y\cdot X\mathbf{Z}}$ ($\tilde{R}$) ranged from .1 to .7 by increments of .2 (4 conditions);
    - $R^2_{Y\cdot XCV\mathbf{Z}}$ (Oster's R_{Max}) was set equal to $1.3R^2_{Y\cdot X\mathbf{Z}}$ as per Oster's guideline.

---

[4] All analyses conducted with this R code: https://www.dropbox.com/s/kva2t6bxsekz2wt/generating%20regression%20coefficient%20implied%20by%20Oster%20delta%20from%20estimated%20effect%20simplify%20oster%20delta%20confirmed%20reduced%20with%20oster%20regression%20adding%20plots%20redo%2011%2013%2022%20fiddle.R?dl=0

- $\tilde{\beta}_1$ ranged from 0.179 (Se[$\tilde{\beta}_1$]) to 1.379 by increments of .3 (σ_Y /5=.3) (5 conditions).

The combinations of conditions produced 2x4x5=40 scenarios to evaluate the derivation. Four scenarios were removed because $R^2_{Y\cdot X\mathbf{Z}} < r^2_{Y\cdot X|\mathbf{Z}}$, resulting in 36 scenarios.

For each scenario $r_{Y\cdot CV|\mathbf{Z}}$ and $r_{X\cdot CV|\mathbf{Z}}$ were calculated as in (11) and (12) with $r_{X\cdot Y|\mathbf{Z}} = \frac{t(\tilde{\beta}_1)}{\sqrt{df + t(\tilde{\beta}_1)^2}}$, where *df*=*n*-7-3. The term $R_{CV\cdot \mathbf{Z}}$ was fixed at zero as in the derivation.

Then β_1 was estimated using (8) and R^2 using (9) as a function of $r_{Y\cdot CV|\mathbf{Z}}$, $r_{X\cdot Y|\mathbf{Z}}$ and $r_{X\cdot CV|\mathbf{Z}}$. Results were confirmed with the Lavaan procedure in the statistics package R (see appendix E for a specific example).

As shown in Figure 1, the expressions for $r_{X\cdot CV|\mathbf{Z}}$ and $r_{Y\cdot CV|\mathbf{Z}}$ in (11) and (12) generated the values of $\hat{\beta}_1$ and $R^2_{Y\cdot X\mathbf{Z}CV}$ specified as in (6a) and (6b) associated with model (1c). Specifically, in Figure 1 the realized values of $\hat{\beta}_1$ were all within .000001 of the specified values of zero or .1. Similarly, in Figure 2 the realized values of $R^2_{Y\cdot X\mathbf{Z}CV}$ were all within .000001 of the specified values of R_{Max} (represented by the darkness of the data points).

Figure 1. Exact Value of $\hat{\beta}_1$ Generated through Correlations Associated with the Omitted Variable

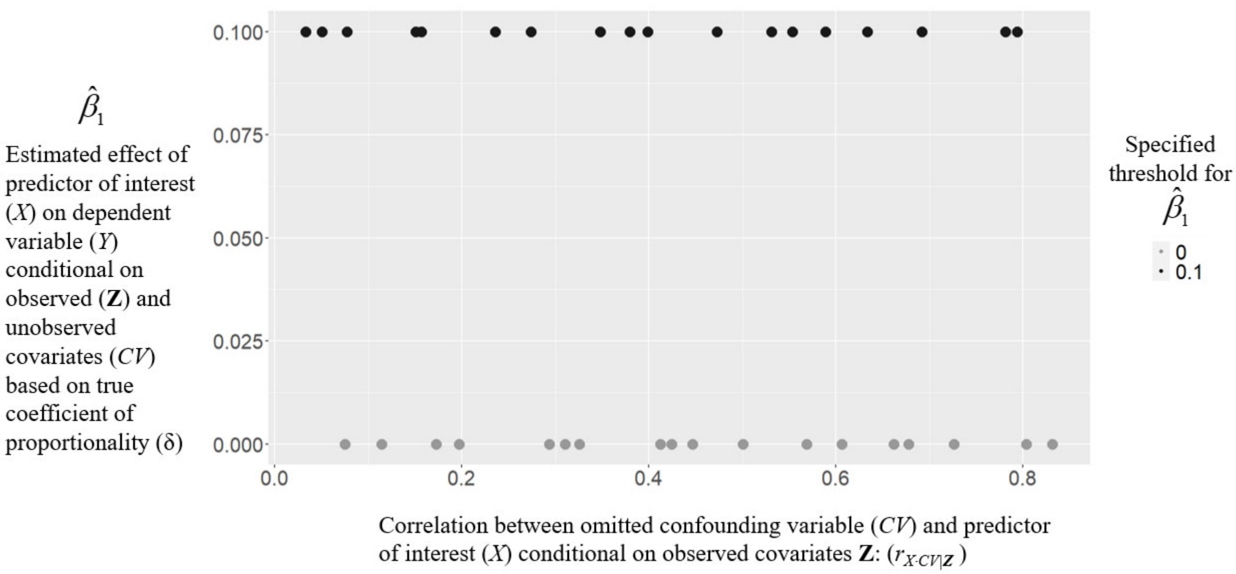
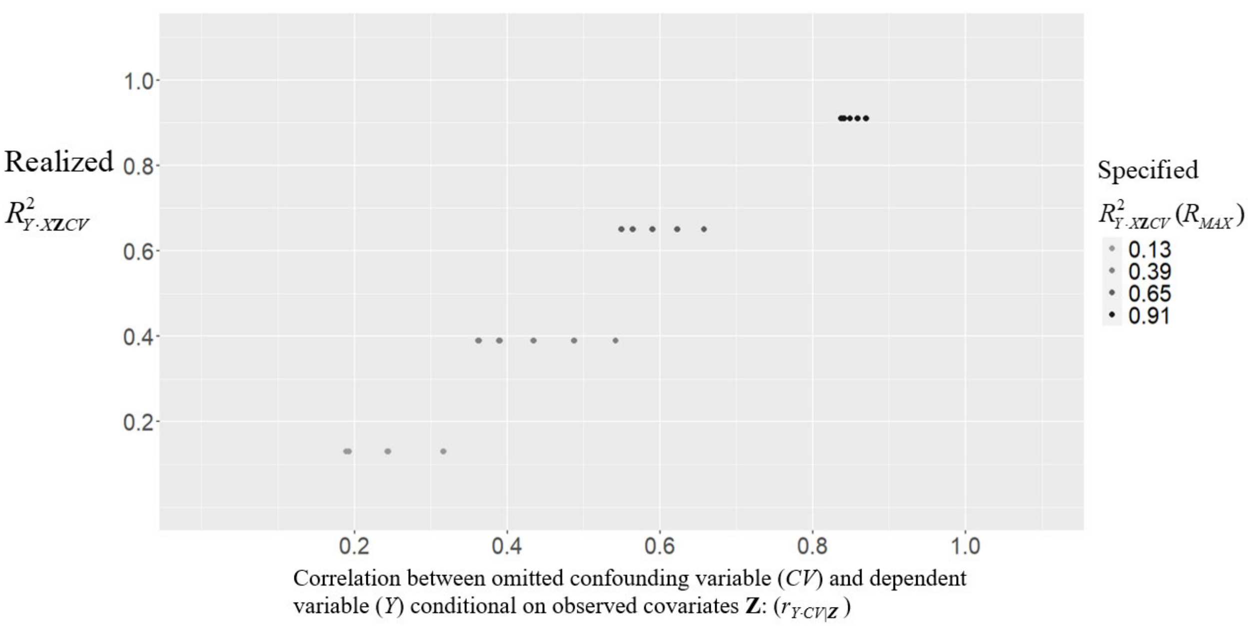


Figure 2. Exact Value of R^2 Generated through Correlations Associated with the Omitted Variable



Empirical Verification of Low Birthweight and Premature Status on IQ[5]

For the illustrative example, Oster (2019) reports the estimated effect of *low birthweight and preterm* on *IQ* is $\tilde{\beta}_1 = -.125$, standard error of .05, and $\tilde{R}^2$ of .251. Supplemental materials in Oster (2019) indicate $\hat{\sigma}_Y = .991$ and $\hat{\sigma}_X = .217$. The supplemental materials for Oster (2019) also indicate the sample size for *IQ* is 6962 while the sample size for *low birth weight and preterm* is 6174, with the difference presumably due to missing data. We use the sample size of 6174 assuming listwise deletion. Larger sample sizes increase $R_{X\cdot\mathbf{Z}}$ and therefore decrease δ, creating a greater difference between δ and δ^*.[6] The corresponding degrees of freedom are 6174-7-3=6164, accounting for the seven covariates used in Oster's model 1b and the intercept, the estimated effect of *X*, and the *CV*.

Our correlational framework reveals an important aspect of the example of *low birth weight and preterm* on *IQ*. Based on expression C2, the reported quantities ($\tilde{\beta}_1 = -.125$, standard error =.050, $\tilde{R}^2 = .251$, $\hat{\sigma}_Y = .991$, $\hat{\sigma}_X = .217$, *df*=6164) imply that $R^2_{X\cdot\mathbf{Z}} < 0$ and $R_{X\cdot\mathbf{Z}}$ is undefined. This is a result of rounding. Specifically, the smallest value of the standard error (to five digits after the decimal) that could be rounded to.050 and produce $R^2_{X\cdot\mathbf{Z}}>0$ is .05034 which would produce $R^2_{X\cdot\mathbf{Z}}$=.000115. Note that this corresponds to a substantively insignificant level

---

[5] Results can be obtained in R using the following commands: install.packages("devtools")
devtools::install_github("jrosen48/konfound")
library(konfound)
pkonfound(est_eff = .125, std_err = .05049, n_obs = 6174, n_covariates = 7, sdx = .217, sdy = .991, R2 = .251, eff_thr = 0, FR2max = .61, index = "COP", to_return = "raw_output")
or in Stata: test_cop 0.125 0.05049 6174 7 0.217 0.991 0.251, fr2max(0.61)
or using the konfound-it app or on the konfound-it spreadsheet.

[6] If *n*=6300, δ^* =1.37, an order of magnitude larger than the true value (δ=.12) instead of twice the value of δ=.583as reported below.

of selection on observables on which to assess the robustness of the estimated effect. Therefore, we use $se(\tilde{\beta}_1)$ =.05049 , the maximum value (to five decimals) that could be rounded to $se(\tilde{\beta}_1) = .050$. Even using $se(\tilde{\beta}_1)$ =.05049 generates $R^2_{X\cdot\mathbf{Z}}$ =.006 with the covariates still explaining less than one percent of the variance in *X*. This preliminary calculation of the value of $R^2_{X\cdot\mathbf{Z}}$ implied by reported quantities demonstrates the importance of conceptualizing the coefficient of proportionality in terms of correlations.

Obtaining $r_{X\cdot CV|Z}$ and $r_{Y\cdot CV|Z}$ to produce for specified values of $\hat{\beta}_1$ and $R^2_{Y\cdot XCV\mathbf{Z}}$

Oster (2019) specified $\hat{\beta}_1 = \beta^{\#}$=0 and $R^2_{Y\cdot XCV\mathbf{Z}}$=$R_{\text{Max}}$=.61 for the inference of an effect of *low birthweight and preterm* on *IQ*. To satisfy these conditions, we start by obtaining

$$r_{X\cdot Y|\mathbf{Z}} = \frac{t(\tilde{\beta}_1)}{\sqrt{df + t(\tilde{\beta}_1)^2}} = \frac{2.476}{\sqrt{6164 + 2.476^2}} = .032,$$

where $t(\tilde{\beta}_1) = \frac{.125}{.05049} = 2.476$, and

$$R^2_{Y\cdot XCV|\mathbf{Z}} = \frac{R^2_{Y\cdot XCVZ} - R^2_{Y\cdot\mathbf{Z}}}{(1 - R^2_{Y\cdot\mathbf{Z}})} = \frac{.61 - .500^2}{(1 - .500^2)} = .480,$$

where $R_{Y\cdot\mathbf{Z}} = .500$ is obtained in Appendix C.

Then using (12) to obtain $r_{X\cdot CV|\mathbf{Z}}$ for a threshold value of $\beta^{\#}$= 0 (using $|\tilde{\beta}_1|$):

$$r_{X\cdot CV|\mathbf{Z}} = \frac{r_{X\cdot Y|\mathbf{Z}} - \frac{\hat{\sigma}_{X|\mathbf{Z}}}{\hat{\sigma}_{Y|\mathbf{Z}}}\beta^{\#}}{\sqrt{\frac{\hat{\sigma}^2_{X|\mathbf{Z}}}{\hat{\sigma}^2_{Y|\mathbf{Z}}}\beta^{\#2} - 2r_{X\cdot Y|\mathbf{Z}}\frac{\hat{\sigma}_{X|\mathbf{Z}}}{\hat{\sigma}_{Y|\mathbf{Z}}}\beta^{\#} + R^2_{Y\cdot X\cdot CV|\mathbf{Z}}}} = \frac{r_{X\cdot Y|\mathbf{Z}}}{\sqrt{R^2_{Y\cdot X\cdot CV|\mathbf{Z}}}} = \frac{.032}{\sqrt{.480}} = .045 .$$

And from (11)

$$r_{Y\cdot CV|\mathbf{Z}} = r_{X\cdot Y|\mathbf{Z}} r_{X\cdot CV|\mathbf{Z}} + \sqrt{(R^2_{Y\cdot XCV|\mathbf{Z}} - r^2_{X\cdot Y|\mathbf{Z}})(1 - r^2_{X\cdot CV|\mathbf{Z}})}$$
$$= (.032)(.045) + \sqrt{(.480 - .032^2)(1 - .045^2)} = .693.$$

Note that

$$\hat{\sigma}_{Y|\mathbf{Z}} = \sqrt{1-R^2_{Y\cdot\mathbf{Z}}}\hat{\sigma}_Y = \sqrt{1-.500^2}.991 = .858, \text{ and}$$
$$\hat{\sigma}_{X|\mathbf{Z}} = \sqrt{1-R^2_{X\cdot\mathbf{Z}}}\hat{\sigma}_X = \sqrt{1-.078^2}.217 = .216.$$

Therefore, from (8)

$$\hat{\beta}_1 = \frac{\hat{\sigma}_{Y|\mathbf{Z}}}{\hat{\sigma}_{X|\mathbf{Z}}}\frac{r_{X\cdot Y|\mathbf{Z}} - r_{X\cdot CV|\mathbf{Z}}r_{Y\cdot CV|\mathbf{Z}}}{1-r^2_{X\cdot CV|\mathbf{Z}}} = \frac{.858}{.216}\frac{.032-(.045)(.693)}{1-.045^2} = 0 = \beta^{\#}.$$

The result holds (regardless of $\hat{\sigma}_{Y|\mathbf{Z}}$ and $\hat{\sigma}_{X|\mathbf{Z}}$) because $r_{X\cdot Y|\mathbf{Z}}$ - $r_{X\cdot CV|\mathbf{Z}}\, r_{Y\cdot CV|\mathbf{Z}}$= 0.03151536−(.0454969)(.6926925)=0, where we provide extra significant digits to verify the result. Note that for $\beta^{\#}$=0, $r_{Y\cdot CV|\mathbf{Z}}$ could be directly obtained as: $r_{Y\cdot CV|\mathbf{Z}} = r_{X\cdot Y|\mathbf{Z}} / r_{X\cdot CV|\mathbf{Z}}$. Note this is presented for $r_{X\cdot Y|\mathbf{Z}}$>0. Because $\tilde{\beta}_1$ is negative (the estimated effect of *low birthweight and preterm* is negative on *IQ*), $r_{X\cdot CV|\mathbf{Z}}$ and $r_{Y\cdot CV|\mathbf{Z}}$ would have to take opposite signs.

The combined R^2 is

$$R^2_{Y\cdot XCV\mathbf{Z}} = R^2_{Y\cdot\mathbf{Z}} + (1-R^2_{Y\cdot\mathbf{Z}})R^2_{Y\cdot XCV|\mathbf{Z}} = .500^2 + (1-.500^2).480 = .61.$$

This verifies in one empirical example that the expressions in (11) and (12) generate the specified values of $\hat{\beta}_1$=0 and $R^2_{Y\cdot XCVZ} = .61$ for model (1c).

For completeness,

$$se(\hat{\beta}_1) = \sqrt{\frac{1-R^2_{Y\cdot XCV|\mathbf{Z}}}{df}\times\frac{1}{1-r^2_{X\cdot CV|\mathbf{Z}}}} = \sqrt{\frac{1-.480}{6264}\times\frac{1}{1-.045^2}} = .009.$$

Note that the standard error would have to be corrected for *the number of covariates* in **Z** that are included in model (1c). The values generated here are confirmed with the Lavann procedure in Appendix E.

Obtaining the coefficient of proportionality

To obtain the coefficient of proportionality, start with

$$r_{X\cdot CV} = \sqrt{1-R^2_{X\cdot\mathbf{Z}}}r_{X\cdot CV|\mathbf{Z}} = \sqrt{1-.078^2}.0455 = .0454,$$

with extra digits provided to differentiate the unconditional from the conditional correlation. This expression already allows a direct evaluation of the robustness of the inference – a fairly small correlation of $r_{X\cdot CV}$=.0454 between an omitted variable and *low birthweight and preterm* would reduce the estimated effect to zero, indicating the inference is not highly robust.

To generate R^2=.61 for $r_{X\cdot CV}$=.0454, $r_{Y\cdot CV} = \sqrt{1-R^2_{Y\cdot \mathbf{Z}}}\, r_{Y\cdot CV|\mathbf{Z}} = \sqrt{1-.500^2}\,.691 = .599$ (again noting that $r_{X\cdot CV}$ and $r_{Y\cdot CV}$ would take opposite signs because $\tilde{\beta}_1 < 0$).

The value of δ corresponding to $r_{X\cdot CV}$=.0454 is (from Equation 14)

$$\delta = \frac{r_{X\cdot CV}}{R_{X\cdot \mathbf{Z}}} = \frac{.0454}{.078} = .583\,.$$

That is, selection on unobservables would have to be about 58% that of the very modest selection on observables ($R_{X\cdot \mathbf{Z}}$=.078) to reduce the estimated effect of *low birth weight and preterm* on *IQ* to zero.

To further evaluate the implications of using δ^*, we solve (14) for $r_{X\cdot CV}$ and replace δ with δ^* yielding (see Appendix F):

$$r_{X\cdot CV} = \delta^* R_{X\cdot \mathbf{Z}}\,. \quad (15)$$

In the empirical example, using the unrestricted estimator, $r_{X\cdot CV} = \delta^* R_{X\cdot \mathbf{Z}} = 1.343(.0778) = .104$ with corresponding $r_{Y\cdot CV}$=.599 (with opposite signs) from (13b) ensuring R_{Max}=.61. Using these values and the Lavaan procedure in the R software, $\hat{\beta}_1$ =-.164 for model (1c). That is, the value of δ^* larger than δ reduces $\hat{\beta}_1$ below the specified threshold of $\beta^{\#}$=0. This is consistent with δ^* overstating the robustness of the inference; δ^* need not be as large as 1.343 to reduce $\hat{\beta}_1$ to zero.

In the empirical example, both the Restricted (1.365) and Unrestricted (1.343) estimators for δ^* are greater than 1 (the rule of thumb value for robustness) and yield $\hat{\beta}_1$ <0 while the exact value of δ=.583 is less than one and yields $\hat{\beta}_1$ =0. However, again we note that our correlation-based approach reveals that the value of the coefficient of proportionality is sensitive to difference in $R_{X\cdot\mathbf{Z}}$ especially for low values of $R_{X\cdot\mathbf{Z}}$. Consequently, these qualitative differences may depend on choice of rounding. For example, using a value of se($\tilde{\beta}_1$)=.05034 would a very small value of $R_{X\cdot\mathbf{Z}}$==.0107 and therefore a large value of δ, 4.259). Nonetheless, the simulations demonstrate that δ^* can depart considerably from δ and produce $\hat{\beta}_1$ <0 independent of rounding in reported values.[7]

EVALUATION OF δ^*

Evaluation of δ^* through Simulation

From (15), for a specified value of δ^* and observed $R_{X\cdot\mathbf{Z}}$, one can calculate $r_{X\cdot CV}$. In turn, one can obtain $r_{Y\cdot CV}$ from (13b) generating the sufficient statistics necessary to estimate $\hat{\beta}_1$ as in (1c).[8,9] We apply this approach to evaluate δ^* based on the Unrestricted Estimator as in (5) which requires fewer assumptions and is more accurate than the Restricted Estimator, drawing on the same scenarios used to verify $r_{X\cdot CV|\mathbf{Z}}$ and $r_{Y\cdot CV|\mathbf{Z}}$ in Figures 1 and 2.

[7] Oster (2019) refers to an external study (Salt & Redshaw, 2006) consistent with the interpretation of a robust inference of an effect of *low birthweight and premature status* on *IQ*, but this validation is based on a synthesis of only observational studies since one cannot randomly assign infants to be *premature* or *low birthweight*. Therefore, a bias free comparison is not feasible.

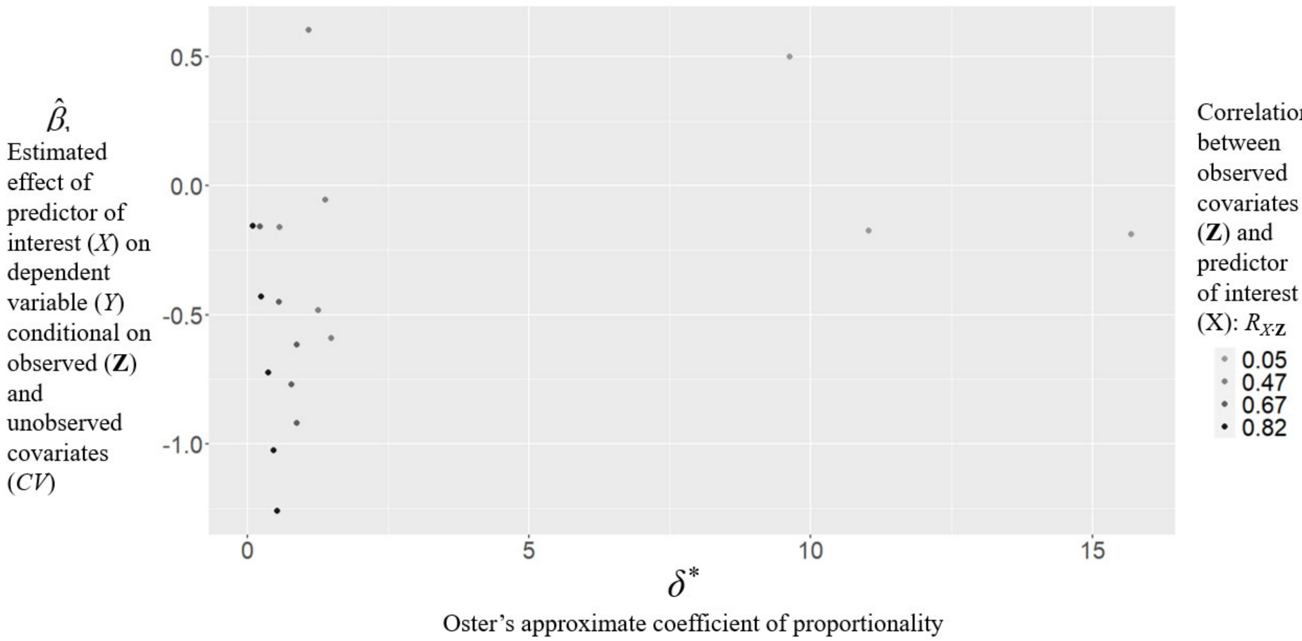
[8] We use the unconditional $r_{X\cdot CV}$ and $r_{Y\cdot CV}$ instead of the conditional $r_{X\cdot CV|\mathbf{Z}}$ and $r_{Y\cdot CV|\mathbf{Z}}$ because Oster's (2019) derivation is a function of a baseline regression which is not conditioned on **Z**.

[9] Oster provides no direct verification of (15) or (16) for δ^* in terms of estimates of β_1.

Corresponding to Oster's (2019) derivation, we use an unconditional baseline model in which $\dot{\beta}_1$ was set to be $r_{X \cdot Y}$ (σ_Y/σ_X) and $\dot{R} = r^2_{X \cdot Y}$.[10] As shown in Figure 3, when calculated over the simulated scenarios, Oster's δ^* produces many $\hat{\beta}_1$ considerably less than zero. In these cases, δ^* is larger than necessary to reduce $\hat{\beta}_1$ to the threshold value (of zero), ultimately inflating the expression of robustness. Consider the value of δ^* of 1.04 (for which $\hat{\beta}_1$=-.217) associated with δ =.65 (for which $\hat{\beta}_1$=0), resulting in an overstatement of the robustness of inference by a factor of 60%: (1.04-.65)/.65=60%. That is, using δ^* one would believe that selection on unobservables must be 1.04 (104%) that of selection on observables to reduce $\hat{\beta}_1$ to 0, when in fact selection on unobservables must be only .65 (65%) that of observables to reduce $\hat{\beta}_1$ to 0. This example is not unusual. More than 75% (28/36) of the $\hat{\beta}_1$ in Figure 3 are negative using δ. Moreover, none of the $\hat{\beta}_1$ are within .01 of zero when $\hat{\beta}_1$ is obtained based on δ^*. Note that because equation (13b) determines the specific $r_{Y \cdot CV}$ to satisfy condition 6b given *any* $r_{X \cdot CV}$, the realized $R^2_{Y \cdot X\mathbf{Z}CV}$ based on δ^* is exactly the specified R_{max} (as was the case for Figure 2 for the exact value of δ).

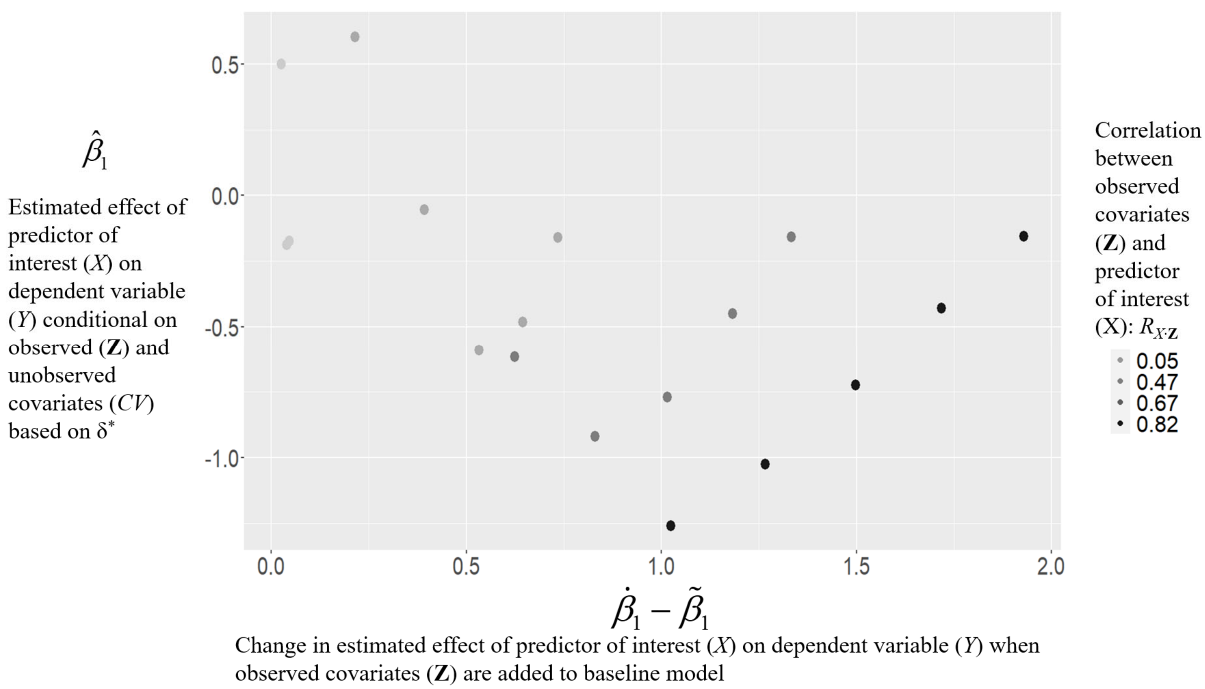
[10] The term $r_{X \cdot Y}$ is obtained as in Appendix A. Results would be very similar if we set $\dot{\beta}_1$ =($r_{X \cdot Y}$ +.05) (σ_Y/σ_X) and $\dot{R} = (r_{X \cdot Y} + .05)^2$ reflecting some practice in which the baseline model includes some covariates.

Figure 3. Value of $\hat{\beta}_1$ implied versus Oster's Approximate Coefficient of Proportionality δ^*



To further explore the properties of δ^*, we note that the inaccuracy of δ^* is related to $\dot{\beta}_1 - \tilde{\beta}_1$ as shown in Figure 4. For small changes in $\dot{\beta}_1 - \tilde{\beta}_1$, $\hat{\beta}_1$ is positive, indicating that δ^* was too small, understating the robustness of the inference. As $\dot{\beta}_1 - \tilde{\beta}_1$ increases, $\hat{\beta}_1$ decreases, with a strong negative trend (correlation of -.45). This is of concern because when $\hat{\beta}_1$ is less than zero δ^* is too large, overstating the robustness of the inference. In general, δ^* is overly responsive to changes in β_1 from (1a) to (1b) -- ($\dot{\beta}_1 - \tilde{\beta}_1$). This is inherent in the conceptualization of coefficient stability (Oster, 2019); unstable coefficients are indicative of an inference that is not robust. [11]

Figure 4. Value of $\hat{\beta}_1$ versus Change in Estimated Effect with Observed Covariates ($\dot{\beta}_1 - \tilde{\beta}_1$) Implied by Oster's Approximate Coefficient of Proportionality δ^*



To more fully evaluate the implications of Oster's (2019) approximation δ^*, we compare δ^* with δ in the simulated scenarios in Figure 5. About 89% (32/36) of the points are above the 45-degree line, indicating δ^* characterizes the inference as more robust than it is according to the

[11] Within large values of $R_{X\cdot\mathbf{Z}}$, the trend is opposite. For $R_{X\cdot\mathbf{Z}}$=.82, as δ^* increases $\hat{\beta}_1$ increases and approaches zero. This is part of the mechanism by which δ^* is asymptotically correct because $R_{X\cdot\mathbf{Z}}$ increases with sample size, all else being equal.

true δ. Generally, for larger original observed $\tilde{R}$ from the intermediate model in (1b), δ^* is more likely to overstate the robustness of the inference. In particular, in the light shaded green box $\delta^* >$ 1 but the true δ <1. In these cases, one would infer from δ^* that the estimate of β_1 is robust (Oster, 2019, page 195, uses a threshold of δ^*=1 to define a robust inference) but not so from δ (as in the example of the inference of an effect of *low birthweight and preterm* on *IQ*).

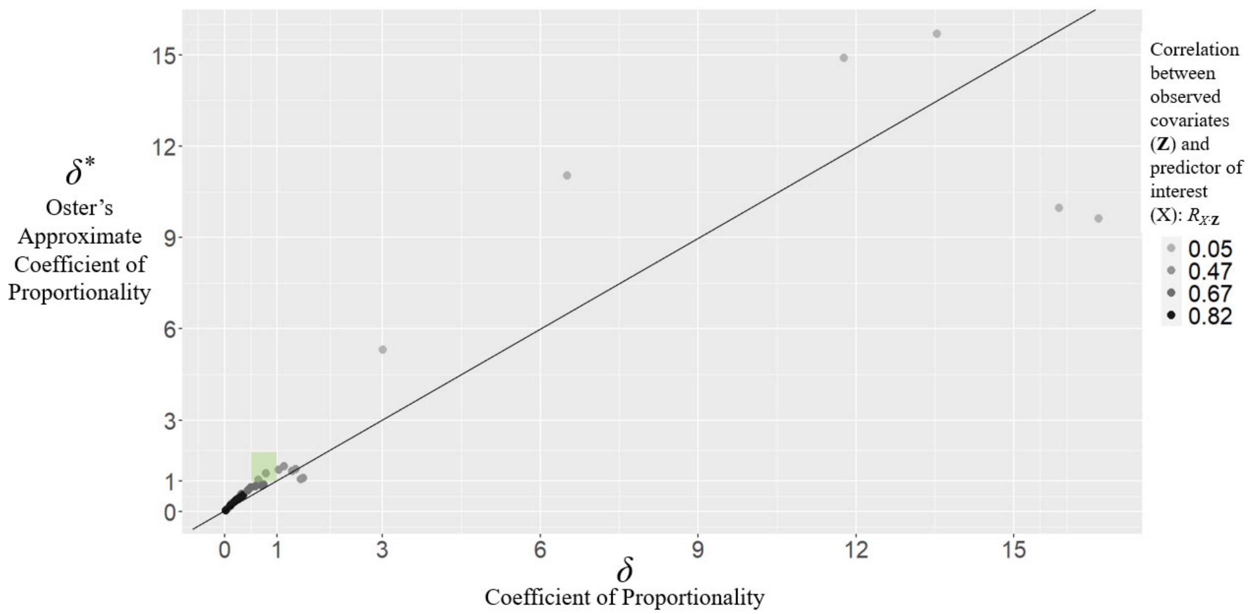


Figure 5. Oster's Approximate Coefficient of Proportionality δ^* by Exact Coefficient of Proportionality δ

Analytic Evaluation of δ^* Relative to δ

To analytically compare δ^* with δ we begin with the intuitive expression for the Restricted Estimator for δ^* and simplify by considering the threshold $\beta^{\#}$=0:

$$\delta^* = \tilde{\beta}_1 \left(\frac{\tilde{R} - \dot{R}}{(\dot{\beta}_1 - \tilde{\beta}_1)(R_{Max} - \tilde{R})} \right) = \frac{\tilde{R} - \dot{R}}{(\dot{\beta}_1 - \tilde{\beta}_1)} \left(\frac{\tilde{\beta}_1}{(R_{Max} - \tilde{R})} \right) \quad . \quad (16)$$

To compare with the expression for δ when $\beta^{\#}$=0, we note from (13a)

$$r_{X\cdot CV} = \sqrt{1-R^2_{X\cdot\mathbf{Z}}}\, r_{X\cdot CV|\mathbf{Z}} = \sqrt{1-R^2_{X\cdot\mathbf{Z}}}\, \frac{r_{X\cdot Y|\mathbf{Z}}}{\sqrt{R^2_{Y\cdot X\cdot CV|\mathbf{Z}}}} = (1-R^2_{X\cdot\mathbf{Z}})\left(\frac{\tilde{\beta}_1 \dfrac{\hat{\sigma}_X}{\hat{\sigma}_Y}}{\sqrt{R^2_{Max}-R^2_{\mathbf{Z}\cdot Y}}}\right). \quad (17)$$

Substituting the expression for $r_{X\cdot CV}$ in (17) into (14):

$$\delta = \frac{(1-R^2_{X\cdot\mathbf{Z}})}{R_{X\cdot\mathbf{Z}}}\left(\frac{\tilde{\beta}_1 \dfrac{\hat{\sigma}_X}{\hat{\sigma}_Y}}{\sqrt{R^2_{Y\cdot X\mathbf{Z}CV}-R^2_{\mathbf{Z}\cdot Y}}}\right) = \frac{(1-R^2_{X\cdot\mathbf{Z}})}{R_{X\cdot\mathbf{Z}}}\left(\frac{\tilde{\beta}_1 \dfrac{\hat{\sigma}_X}{\hat{\sigma}_Y}}{\sqrt{R^2_{Max}-R^2_{\mathbf{Z}\cdot Y}}}\right). \quad (18)$$

Both the expression for δ^* in (16) and the expression for δ in (18) are proportional to $\tilde{\beta}_1$ -- intuitively, the larger the estimated effect in model (1b) the more robust the inference. Both are also inversely proportional to the extra variation in Y accounted for by the unobserved covariates ($R_{\max}$). The smaller the increase in variance explained in the hypothetical final model (1c) including the unobserved covariate, the more robust the inference because it suggests that the observed covariates are not impactful. One distinction is that (16) evaluates $R_{\max}$ relative to $\tilde{R}$, the variance in Y explained by $\mathbf{Z}$ and X while (18) evaluates $R_{\max}$ relative to $R^2_{\mathbf{Z}\cdot Y}$ (the variance in Y explained only by $\mathbf{Z}$).

The greatest distinction between (16) and (18) is how they incorporate information about the strength of the covariates. Specifically, the expression for δ^* in (16) is inversely proportional to $\dot{\beta}_1 - \tilde{\beta}_1$, the change in estimated effect when observed covariates are added to model (1a) to generate model (1b). The stronger the observed covariates in terms of reducing the baseline estimated effect in (1a), the less robust the inference from (1b). The intuition is that if observed covariates account for much of the estimated effect then the same will be expected for unobserved covariates because the observed covariates represent the unobserved covariates (see also Altonji, Elder and Tabor, 2005, pages 169-170). Therefore, the inference is more robust when the observed covariates are relatively weak.

In comparison to (16), the expression in (18) is a direct function of the association between **Z** and *X* and not a function of $\dot{\beta}_1$ - $\tilde{\beta}_1$ relative to $\tilde{R} - \dot{R}$. Simply put, the stronger the selection on observables ($R_{X \cdot \mathbf{Z}}$) the weaker must be the proportional selection on unobservables ($r_{X \cdot CV}$) to reduce $\hat{\beta}_1$ to zero. In this way interpretation of the inference is based only on selection on unobservables relative to the observables included in model (1b); the expression in (18) does not depend on coefficient stability relative to the baseline model (1a). This is the advantage of the static conceptualization of estimation of models (1b) and (1c) in terms of a correlation or covariance matrix (see Appendix A) instead of the dynamic conceptualization that depends on coefficient stability relative to a baseline model.

The reliance of δ^* on a baseline model opens the general analytic question of how one should choose the covariates to be in baseline model. Oster (2019) does not include baseline covariates in the derivation of δ^* but includes some covariates in baseline models in empirical examples. Furthermore, when baseline covariates are included, it is not clear if they should include fixed effects. Oster (2019) did not include sibling fixed effects in baseline models, while Redding & Grissom (2021) included fixed effects for students in their baseline models. Thus, δ^* requires interpreters of an inference to adjudicate the choice of baseline and final covariates in models (1a) and (1b) while interpreting δ^*. In contrast, interpretation of δ does not require consideration of how model (1b) emerged from (1a).

BOUNDING OF $\hat{\beta}_1$

An alternative way of expressing the sensitivity of an inference is to generate a bound for the estimated effect. Indeed, many of the applications of Oster (2019) employ such bounding. We develop bounds for $\hat{\beta}_1$ under the condition that selection on unobservables is the same as that for observables: $r_{X\cdot CV}= R_{X\cdot \mathbf{Z}}$ (maintaining the assumption $R_{CV\cdot \mathbf{Z}}$=0). Under these assumptions, (11) and (12) become:

$$r_{Y\cdot CV|\mathbf{Z}} = \frac{r_{X\cdot Y|\mathbf{Z}} R_{X\cdot \mathbf{Z}}}{\sqrt{1-R^2_{X\cdot \mathbf{Z}}}} + \sqrt{(R^2_{Y\cdot XCV|\mathbf{Z}} - r^2_{X\cdot Y|\mathbf{Z}})(1-\frac{R^2_{X\cdot \mathbf{Z}}}{1-R^2_{X\cdot \mathbf{Z}}})} \text{, and} \quad (19a)$$

$$r_{X\cdot CV|\mathbf{Z}} = \frac{R_{X\cdot \mathbf{Z}}}{\sqrt{1-R^2_{X\cdot \mathbf{Z}}}} . \quad (19b)$$

Then substituting from (19a) and (19b) into (8) and simplifying:

$$\text{Bound}(\hat{\beta}_1) = \frac{\hat{\sigma}_{Y|\mathbf{Z}}}{\hat{\sigma}_{X|\mathbf{Z}}} \frac{r_{X\cdot Y|\mathbf{Z}} - \frac{R_{X\cdot \mathbf{Z}}}{\sqrt{1-R^2_{X\cdot \mathbf{Z}}}}\left(\frac{r_{X\cdot Y|\mathbf{Z}} R_{X\cdot \mathbf{Z}}}{\sqrt{1-R^2_{X\cdot \mathbf{Z}}}} + \sqrt{(R^2_{Y\cdot XCV|\mathbf{Z}} - r^2_{X\cdot Y|\mathbf{Z}})(1-\frac{R^2_{X\cdot \mathbf{Z}}}{1-R^2_{X\cdot \mathbf{Z}}})}\right)}{1-\frac{R^2_{X\cdot \mathbf{Z}}}{1-R^2_{X\cdot \mathbf{Z}}}}$$

$$= \frac{\hat{\sigma}_{Y|\mathbf{Z}}}{\hat{\sigma}_{X|\mathbf{Z}}}\left(r_{X\cdot Y|\mathbf{Z}} - R_{X\cdot \mathbf{Z}} \sqrt{\frac{(R^2_{Y\cdot XCV|\mathbf{Z}} - r^2_{Y\cdot X|\mathbf{Z}})}{1-2R^2_{X\cdot \mathbf{Z}}}} \right). \quad (20)$$

The applies for $\tilde{\beta}_1$ >0; the sign for $R_{X\cdot \mathbf{Z}}$ is reversed for $\tilde{\beta}_1$ <0.

For the negative estimated effect of *low birth weight and preterm* on *IQ*,

$$\text{Bound}(\hat{\beta}_1) = \frac{.858}{.216}\left(-.032 + .078\sqrt{\frac{(.479-.032^2)}{1-2(.078^2)}} \right) = .090 .$$

That is, $\hat{\beta}_1$ would be slightly positive if selection on unobservables were equal to selection on unobservables. This is consistent with the value of δ =.583, that selection on unobservables need be only 58% that of observables to make $\hat{\beta}_1$ =0.

Using the same approach, the bound for the estimated effect of *smoking* on *low birthweight* is −69.0396. That is, even if selection on unobservables were as strong as selection on observables the estimated effect would be negative. This is consistent with δ needing to be greater than one (δ =1.627) to reduce the estimated effect to 0.

DISCUSSION

Beginning with inferences about the effects of smoking on lung cancer (Cornfield et al., 1959), sensitivity analyses have been key to informing debates about causal inferences. In that context, Oster (2019) highlights the explanatory power of a model (R^2) in contextualizing the robustness of an estimated effect to unobserved variables. It is reasonable to consider that there are limits on the amount of variance that can be explained in a given study, which should be accounted for in assessing how strong selection on unobservables must be to change an inference. In this study, we have drawn on Oster's intuition to generate expressions for correlations associated with an unobserved covariate that reduce an estimated effect to a specific threshold (e.g., $\beta^{\#}$=0) in an estimated model with a specified maximum R^2. In contrast to Oster's original approximation, our derivation is exact, does not depend on the choice of the baseline model defining coefficient stability, and generates expressions of the omitted variable that can be evaluated in the absolute terms of a correlation coefficient. We verify our derivation in an empirical example and through simulation. Furthermore, our expressions can be directly calculated from conventionally reported quantities (e.g., $\tilde{\beta}_1$, se[$\tilde{\beta}_1$], $\hat{\sigma}_X$, $\hat{\sigma}_Y$, *n, number of covariates*, and $R^2_{\mathrm{Y \cdot X\mathbf{Z}}}$) and can also be used to bound the final estimated effect if an omitted variable were included in the model.

We emphasize that the derivation presented here yields the exact conditions that produce a specified estimated effect and R^2, whereas Oster's (2019) approach gives only an approximation. Specifically, compared to the exact solution for the coefficient of proportionality (δ), we show Oster's δ^* can overstate or understate robustness in empirical examples and through simulation. The approximation δ^* is most likely to overstate the robustness of an inference for strong designs in which the observed covariates account for a large change in an estimated effect from the baseline model which is a by-product of the conceptualization of robustness in terms of coefficient stability (Oster, 2019).

Best Practices

Make Maximal use of Observed Covariates and Analytic Techniques (adapted from Frank et al., 2022)

All sensitivity analyses assume that models have already been developed appropriately and are inclusive of alternative explanations represented by observed variables –**Stated plainly, the first best practice is that any sensitivity analysis should only be conducted after a researcher has fully specified a model including all relevant and available measures.** This especially applies to leveraging longitudinal data whenever possible (e.g., Shadish, Clark, and Steiner, 2008; see the review in Wong, Valentine, and Miller-Bains, 2017). See also recent work by Belloni et al., (2016), Young and Holsteen (2017) and Young (2018) regarding model selection. **It is misleading to apply any sensitivity analyses to models that have not already**

been rigorously vetted given the available data. But once the best models have been estimated and adjudicated, there may still be concerns about potential bias due to selection on unobservables. Debate about such concerns can be informed by calculating the coefficient of proportionality.

Carefully Examine Extent of Selection on Observables

A by-product of our conceptualization of selection in terms of correlations is a preliminary calculation of the correlation between observed covariates and the predictor of interest. We caution against over interpretation of proportional selection when selection on observed covariates is less than .05. Indeed, such a small correlation undermines the premise of using selection on observables to represent the likely selection on unobservables (e.g., Oster, 2019, pages 195-196). In such cases, one might simply report the correlation associated with the omitted variable necessary to produce the specified R^2 and estimated effect without expressing as a ratio to selection on observed covariates.

Consider a Minimum Value of the Maximum R^2

In deriving our expression for δ we adopted Oster's (2019) emphasis on realistic expectations for variance explained – $R^2_{Y\cdot X\mathbf{Z}CV}$ or R_{Max}. Oster establishes a guideline for $R_{Max} = 1.3R^2_{Y\cdot X\mathbf{Z}}$ based on the finding that most (97%) inferences from the randomized studies analyzed would have withstood a value of $R_{Max} = 1.3R^2_{Y\cdot X\mathbf{Z}}$. While we accept Oster's logic and empirical validation, we note that small values of $R^2_{Y\cdot X\mathbf{Z}}$ could generate unrealistically small values of R_{Max}, leading the inference to be overstated (the larger the value of R_{Max}, the smaller the δ indicating a less robust inference). In particular, one might consider an absolute minimum R_{Max} of .1.

CONCLUSION

The coefficient of proportionality informs debate about causal inferences by quantifying how strong selection on unobservables must be relative to observables to change an inference. By recognizing the importance of variance explained by a model, Oster (2019) accounts for the data collected and used in social science. We refine Oster's (2019) contribution to the coefficient of proportionality by deriving the exact conditions associated with unobservables that would generate an estimated effect at a specified threshold and associated with a specified R^2, and provide an expression for the coefficient of proportionality that can be used for any sample size. Ultimately, our intent is to inform causal inferences in public policy drawn from nonexperimental studies. But we emphasize sensitivity analyses do not, in and of themselves, establish the quality of a model or change an inference. What sensitivity analyses can do is formalize and quantify the hypothetical conditions necessary to change an inference to inform debate about that inference.

REFERENCES

Alberg, A. J., Shopland, D. R., & Cummings, K. M. (2014). The 2014 Surgeon General's report: commemorating the 50th Anniversary of the 1964 Report of the Advisory Committee to the US Surgeon General and updating the evidence on the health consequences of cigarette smoking. American journal of epidemiology, 179(4), 403-412.

Algina, J., & Seaman, S. (1984). Calculation of semipartial correlations. Educational and psychological measurement, 44(3), 547-549.

Altonji, J. G., Elder, T. E., & Taber. (2002). "Selection on Observed and Unobserved Variables: Assessing theEffectiveness of Catholic Schools." Manuscript (April), Dept. Econ., North-western Univ.

Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. Journal of political economy, 113(1), 151-184.

Baer, B. R., Gaudino, M., Charlson, M., Fremes, S. E., & Wells, M. T. Fragility indices for only sufficiently likely modifications. Proceedings of the National Academy of Sciences 2021; 118(49).

Belloni, A., Chernozhukov, V., & Wei, Y. (2016). Post-selection inference for generalized linear models with many controls. *Journal of Business & Economic Statistics*, *34*(4), 606-619.

Berridge, V. (2006). The policy response to the smoking and lung cancer connection in the 1950s and 1960s. The Historical Journal, 49(4), 1185-1209.

Breslau, N., DelDotto, J. E., Brown, G. G., Kumar, S., Ezhuthachan, S., Hufnagle, K. G., & Peterson, E. L. (1994). A gradient relationship between low birth weight and IQ at age 6 years. Archives of pediatrics & adolescent medicine, 148(4), 377-383.

Cinelli, C., & Hazlett, C. (2020). Making sense of sensitivity: Extending omitted variable bias. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 82(1), 39-67.

Cohen, J., Cohen, P. (1983). Applied multiple regression. Correlation Analysis for the Behavioral Sciences.London. Lawrence Erlbaum.

Cornfield, J., Haenszel, W., Hammond, E. C., Lilienfeld, A. M., Shimkin, M. B., & Wynder, E. L. (1959). Smoking and lung cancer: recent evidence and a discussion of some questions. Journal of the National Cancer institute, 22(1), 173-203.

Campos-Mercade, P., Meier, A. N., Schneider, F. H., & Wengström, E. (2021). Prosociality predicts health behaviors during the COVID-19 pandemic. Journal of public economics, 195, 104367.

Dorie, V., Harada, M., Carnegie, N. B., & Hill, J. (2016). A flexible, interpretable framework for assessing sensitivity to unmeasured confounding. Statistics in medicine, 35(20), 3453-3470.

Fisher, R. A. (1958). Lung cancer and cigarettes?. Nature, 182(4628), 108-108.

Frank, K. A. (2000). Impact of a confounding variable on the inference of a regression coefficient. Sociological Methods and Research, 29(2), 147-194.

Frank, K.A., Lin, Q., Xu, R., Maroulis, S.J., Mueller, A. (on-line first). Quantifying the Robustness of Causal Inferences: Sensitivity Analysis for Pragmatic Social Science. *Social Science Research. 110, 102815.*

* Frank, K. A., *Lin, Q., *Maroulis, S., *Mueller, A. S., Xu, R., Rosenberg, J. M., ... & Zhang, L. Hypothetical case replacement can be used to quantify the robustness of trial results. Journal of Clinical Epidemiology 2021; 134:150-159. *authors listed alphabetically.

Frank, K.A., Maroulis, S., Duong, M., and Kelcey, B. 2013. What would it take to Change an Inference?: Using Rubin's Causal Model to Interpret the Robustness of Causal Inferences. Education, Evaluation and Policy Analysis. Vol 35: 437-460.

*Frank, K. A. and Min, K. Indices of Robustness for Sample Representation. Sociological Methodology 2007;37:349-392. * co first authors.

Frank, K. A., Sykes, G., Anagnostopoulos, D., Cannata, M., Chard, L., Krause, A., & McCrory, R. (2008). Does NBPTS certification affect the number of colleagues a teacher helps with instructional matters?. Educational Evaluation and Policy Analysis, 30(1), 3-30.

Gastwirth, J. L., Krieger, A. M., & Rosenbaum, P. R. (1998). Dual and simultaneous sensitivity analysis for matched pairs. Biometrika, 85(4), 907-920.

Greenwood, B. N., Hardeman, R. R., Huang, L., & Sojourner, A. (2020). Physician–patient racial concordance and disparities in birthing mortality for newborns. Proceedings of the National Academy of Sciences, 117(35), 21194-21200.

Gross, R. T., Spiker, D., & Haynes, C. W. (1997). Helping low birth weight, premature babies: The Infant Health and Development Program. Stanford University Press.

Hack, M., Klein, N. K., & Taylor, H. G. (1995). Long-term developmental outcomes of low birth weight infants. The future of children, 176-196.

Heckman, J. J., Urzua, S., & Vytlacil, E. (2006). Understanding instrumental variables in models with essential heterogeneity. The Review of Economics and Statistics, 88(3), 389-432.

Hembre, E., Moulton, S., & Record, M. (2021). Low-Income Homeownership and the Role of State Subsidies: A Comparative Analysis of Mortgage Outcomes. Journal of Policy Analysis and Management, 40(1), 78-106.

Imbens, G. (2003). Sensitivity to exogeneity assumptions in program evaluation. American Economic Review, 93(2), 126-132.

Kallus, N. Xiaojie Mao, and Angela Zhou. Interval estimation of individual-level causal effects under unobserved confounding. In The 22nd international conference on artificial intelligence and statistics, pages 2281–2290. PMLR, 2019.

Lien, D. S., & Evans, W. N. (2005). Estimating the impact of large cigarette tax hikes the case of maternal smoking and infant birth weight. Journal of Human resources, 40(2), 373-392.

Mauro, R. (1990). Understanding LOVE (left out variables error): A method for estimating the effects of omitted variables. Psychological Bulletin, 108(2), 314.

Moulton, S., Chun, Y., Casey Pierce, S., Quercia, R., Riley, S., & Holtzen, H. (2022). Does Temporary Mortgage Assistance for Unemployed Homeowners Reduce Longer-Term Mortgage Default? An Analysis of the Hardest Hit Fund Program. Journal of Policy Analysis and Management, 41(2), 515-551.

National Research Council. (2000). From neurons to neighborhoods: The science of early childhood development.

Ngo, N. S. (2017). Emission standards, public transit, and infant health. Journal of Policy Analysis and Management, 36(4), 773-789.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. Journal of Business & Economic Statistics, 37(2), 187-204.

Redding, C., & Grissom, J. A. (2021). Do students in gifted programs perform better? Linking gifted program participation to achievement and nonachievement outcomes. Educational Evaluation and Policy Analysis, 01623737211008919.

Robins, J. M., Rotnitzky, A., & Scharfstein, D. O. (2000). Sensitivity analysis for selection bias and unmeasured confounding in missing data and causal inference models. In Statistical models in epidemiology, the environment, and clinical trials (pp. 1-94). Springer, New York, NY.

Rosenbaum, P. R. (2005). Sensitivity analysis in observational studies. Encyclopedia of statistics in behavioral science, 4, 1809-1814.

Rosenbaum, P. R., & Rubin, D. B. (1983). Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. Journal of the Royal Statistical Society: Series B (Methodological), 45(2), 212-218.

Salt, A., & Redshaw, M. (2006). Neurodevelopmental follow-up after preterm birth: follow up after two years. Early human development, 82(3), 185-197.

Shadish, W. R., Clark, M. H., & Steiner, P. M. (2008). Can nonrandomized experiments yield accurate answers? A randomized experiment comparing random to nonrandom assignment. *Journal of the American Statistical Association*, *103*(484), 1334-1344.

US Department of Health, Education, and Welfare. Smoking and Health: Report of the Advisory Committee to the Surgeon General of the Public Health Service. Washington, DC: US Department of Health, Education, and Welfare, Public Health Service; 1964. (Public Health Service Publication No. 1103).

United States. Office on Smoking, Center for Chronic Disease Prevention, & Health Promotion (US). Office on Smoking. (1989). Reducing the Health Consequences of Smoking: 25 Years of Progress: a Report of the Surgeon General.

VanderWeele, T. J., & Arah, O. A. (2011). Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments, and confounders. Epidemiology (Cambridge, Mass.), 22(1), 42-52.

Weisberg, S. (2005). Applied linear regression (Vol. 528). John Wiley & Sons.

Wong, V. C., Valentine, J. C., & Miller-Bains, K. (2017). Empirical performance of covariates in education observational studies. *Journal of Research on Educational Effectiveness*, *10*(1), 207-236.

Wooldridge, J. M. (2010). Econometric analysis of cross section and panel data. MIT press.

Xu, R., Frank, K. A., Maroulis, S. J., & Rosenberg, J. M. (2019). konfound: Command to quantify robustness of causal inferences. The Stata Journal, 19(3), 523-550.

Young, C., & Holsteen, K. (2017). Model uncertainty and robustness: A computational framework for multimodel analysis. *Sociological Methods & Research*, *46*(1), 3-40.

Young, C. (2018). Model uncertainty and the crisis in science. *Socius*, *4*.

Appendix A: Regenerating Oster's Validation of the Coefficient of Proportionality

We replicated Oster's (2019) validation of the estimator in (3) by simulating cases based on the information provided in the caption below Figure 2 in Oster (2019, p. 195) and β_1^* as in (3). To begin, Oster defines W_1 as a linear combination of two orthogonal observed covariates: $W_1=100Z_1+200Z_2$.

In Oster's specification $\hat{\sigma}_{Z_1}^2=\hat{\sigma}_{Z_2}^2=1$ (where $\hat{\sigma}^2$ represents a sample variance). Therefore $\hat{\sigma}_{W_1}^2=100^2(1)+200^2(1)=50000$. Also in Oster's specification, $\hat{\sigma}_{X\cdot Z_1}=\hat{\sigma}_{X\cdot Z_2}=.1$, where $\hat{\sigma}_{X\cdot}$ represents a sample covariance with X. Therefore $\hat{\sigma}_{X\cdot W_1}=100(.1)+200(.1)=30$, noting Z_1 and Z_2 are orthogonal. Note also that the specifications translate to a modest correlation ($r_{X\cdot W_1}$):

$$r_{X\cdot W_1}=\frac{\hat{\sigma}_{X\cdot W_1}}{\hat{\sigma}_X\hat{\sigma}_{W_1}}=\frac{30}{(1)(223.6)}=.134.$$

Oster then specifies $\hat{\sigma}_{CV}^2=250,000$, where CV represents an unobserved confounding variable (W_2 in Oster's notation). And because $\delta=1$

$$\frac{\hat{\sigma}_{X\cdot W_1}}{\hat{\sigma}_{W_1}^2}=\frac{\hat{\sigma}_{X\cdot CV}}{\hat{\sigma}_{CV}^2}\Rightarrow\hat{\sigma}_{X\cdot CV}=\hat{\sigma}_{CV}^2\frac{\hat{\sigma}_{X\cdot W_1}}{\hat{\sigma}_{W_1}^2}=250000\frac{30}{50000}=150.$$

The above specifications culminate in the covariance matrix

$$\begin{bmatrix} & X & W_1 & CV & e \\ X & 1 & 30 & 150 & 0 \\ W_1 & 30 & 50000 & 0 & 0 \\ CV & 150 & 0 & 250000 & 0 \\ e & 0 & 0 & 0 & 1 \end{bmatrix},$$

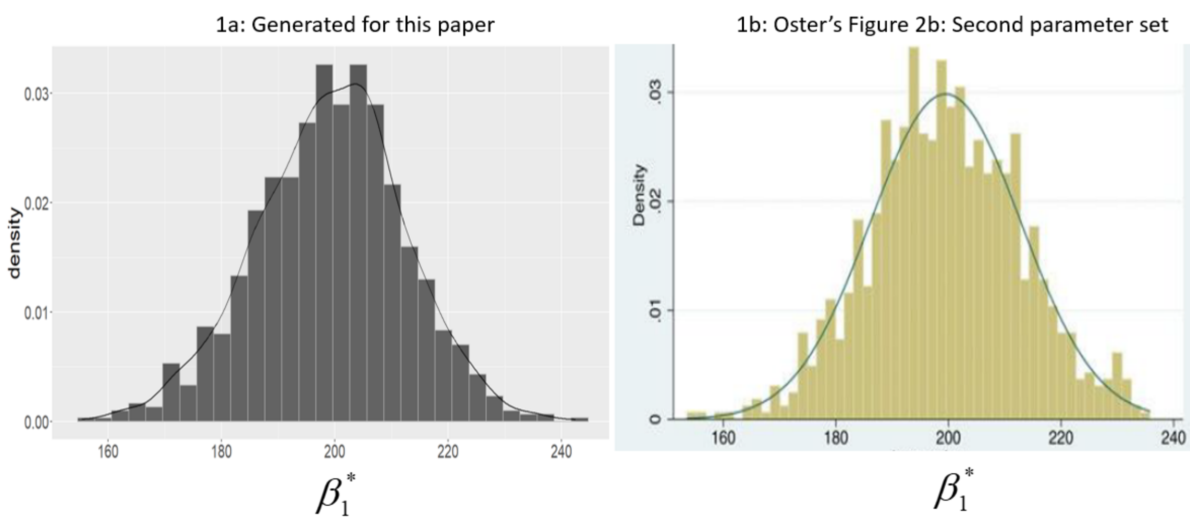
from which we generated 1,000 cases. For each case we generated

$$Y=\beta_1X+W_1+CV+e \quad \text{(A1)}$$

where the $e \sim N(200,1)$ and are orthogonal to W_1 and CV.[12] Then β_1 in (A1) was estimated for each case.

Using the above procedure, Figure 1a (on the left) closely reproduces Oster's Figure 2b (on the right). Specifically, each distribution is centered at the data generating value of $\beta_1=200$. The standard deviation of the estimates (the standard error) is about 13 with 95% of the estimates falling between 172 and 224, less than .05 of a standard deviation of *Y*. We also note that the Restricted Estimator in (3) generates estimates that are approximately normally distributed (which Oster notes can be leveraged as a basis for inference). Most importantly, the replication demonstrates that Oster's derivation can be expressed in terms of a static covariance matrix representing the relationships among *X*, *Y*, and **Z**.

Figure 1A. Regenerating Oster's Validation of the Estimator for β_1

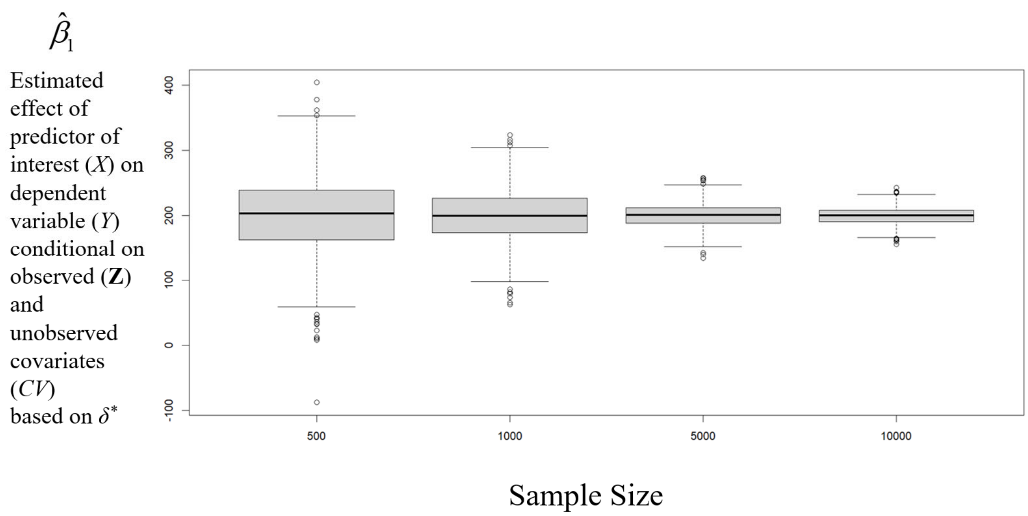


Note. Figure 1b is Figure 2b in Oster (2019).

[12] $Var(Y) = Var(\beta_1 X + W_1 + CV + e) = \beta_1^2 Var(X) + Var(W_1) + 2\beta_1 \text{Cov}(X, W_1) + \text{Var(CV)} + 2\beta_1 \text{Cov}(X, CV) + \text{Var}(e)$
$= 200^2 + 50000 + 2(200)(30) + 250000 + 2(200)(12) + 1 = 412001.$

To evaluate the asymptotics of the estimator in (3) we simulate 1,000 cases each for sample sizes of 500, 1,000, 5,000, and 10,000. The results in Figure 2 show the estimator in (3) is not precise for small *n*. For example, 25% of the β_1^* for sample size of 1000 are below 173 and 25% are above 226, roughly tripling the interquartile range for n=10000, for which 25% of the β_1^* are below 190 and 25% are above 207.

Figure 2A. Asymptotic Performance of Oster's Estimator for β_1



Appendix B: Application to multiple unobserved covariates

To evaluate whether the correlations associated with a single unobserved variable can represent multiple unobserved variables, we begin by considering a model in which there are no covariates **Z** and we seek the conditions under which

$$\hat{\beta}_1 = \frac{\hat{\sigma}_Y}{\hat{\sigma}_X}\frac{r_{X\cdot Y} - r_{X\cdot CV}r_{Y\cdot CV}}{1-r^2_{X\cdot CV}} = \frac{\hat{\sigma}_Y}{\hat{\sigma}_X}\frac{r_{X\cdot Y} - R_{X\cdot \mathbf{CV}}R_{Y\cdot \mathbf{CV}}}{1-R^2_{X\cdot \mathbf{CV}}}. \quad \text{(B1)}$$

To evaluate (B1) we develop expressions for $R^2_{X\cdot \mathbf{CV}}$, and $R^2_{Y\cdot \mathbf{CV}}$ and compare the corresponding expression for $\tilde{\beta}_1$ with a known expression for $\hat{\beta}_1$ from the right hand side of (B1) as a function of X and two covariates: **CV**=[*CV1*, *CV2*]).

The expressions for $R^2_{X\cdot \mathbf{CV}}$, and $R^2_{Y\cdot \mathbf{CV}}$ are

$$R^2_{X\cdot \mathbf{CV}} = \frac{r^2_{X\cdot CV1} + r^2_{X\cdot CV2} - 2r_{X\cdot CV1}r_{X\cdot CV2}r_{CV1\cdot CV2}}{1-r^2_{CV_1\cdot CV_2}}, \text{ and } R^2_{Y\cdot \mathbf{CV}} = \frac{r^2_{Y\cdot CV1} + r^2_{Y\cdot CV2} - 2r_{Y\cdot CV1}r_{Y\cdot CV2}r_{CV1\cdot CV2}}{1-r^2_{CV_1\cdot CV_2}}.$$

The expression for $\hat{\beta}_1$ is therefore

$$\hat{\beta}_1 = \frac{r_{X\cdot Y} - \sqrt{\left(\frac{r^2_{X\cdot CV1} + r^2_{X\cdot CV2} - 2r_{X\cdot CV1}r_{X\cdot CV2}r_{CV1\cdot CV2}}{1-r^2_{CV_1\cdot CV_2}}\right)\left(\frac{r^2_{Y\cdot CV1} + r^2_{Y\cdot CV2} - 2r_{Y\cdot CV1}r_{Y\cdot CV2}r_{CV1\cdot CV2}}{1-r^2_{CV_1\cdot CV_2}}\right)}}{1-\left(\frac{r^2_{X\cdot CV1} + r^2_{X\cdot CV2} - 2r_{X\cdot CV1}r_{X\cdot CV2}r_{CV1\cdot CV2}}{1-r^2_{CV_1\cdot CV_2}}\right)}$$

$$= \frac{r_{X\cdot Y}(1-r^2_{CV1\cdot CV2}) - \sqrt{\left(r^2_{X\cdot CV1} + r^2_{X\cdot CV2} - 2r_{X\cdot CV1}r_{X\cdot CV2}r_{CV1\cdot CV2}\right)\left(r^2_{Y\cdot CV1} + r^2_{Y\cdot CV2} - 2r_{Y\cdot CV1}r_{Y\cdot CV2}r_{CV1\cdot CV2}\right)}}{1-r^2_{X\cdot CV1} - r^2_{X\cdot CV2} + 2r_{X\cdot CV1}r_{X\cdot CV2}r_{CV1\cdot CV2}}. \quad \text{(B2)}$$

A direct expression for $\hat{\beta}_1$ with two covariates is (Mauro, 1990, page 316):

$$\hat{\beta}_1 = \frac{r_{X\cdot Y}(1-r^2_{CV1\cdot CV2}) + r_{Y\cdot CV1}(r_{X\cdot CV2}r_{CV1\cdot CV2} - r_{X\cdot CV1}) + r_{Y\cdot CV2}(r_{X\cdot CV1}r_{CV1\cdot CV2} - r_{X\cdot CV2})}{1 - r^2_{CV1\cdot CV2} - r^2_{X\cdot CV1} - r^2_{X\cdot CV2} + 2r_{X\cdot CV1}r_{CV1\cdot CV2}r_{X\cdot CV2}}$$
$$= \frac{r_{X\cdot Y}(1-r^2_{CV1\cdot CV2}) + r_{Y\cdot CV1}(r_{X\cdot CV2}r_{CV1\cdot CV2} - r_{X\cdot CV1}) + r_{Y\cdot CV2}(r_{X\cdot CV1}r_{CV1\cdot CV2} - r_{X\cdot CV2})}{1 - r^2_{X\cdot CV1} - r^2_{X\cdot CV2} + 2r_{X\cdot CV1}r_{CV1\cdot CV2}r_{X\cdot CV2}}, \text{(B3)}$$

assuming $r^2_{CV1\cdot CV2}$=0 because the unobserved covariates are orthogonal to each other.

Expressions (B2) and (B3) are equivalent if

$$-\sqrt{\left(r^2_{X\cdot CV1}+r^2_{X\cdot CV2}-2r_{X\cdot CV1}r_{X\cdot CV2}r_{CV1\cdot CV2}\right)\left(r^2_{Y\cdot CV1}+r^2_{Y\cdot CV2}-2r_{Y\cdot CV1}r_{Y\cdot CV2}r_{CV1\cdot CV2}\right)} = r_{Y\cdot CV1}(r_{X\cdot CV2}r_{CV1\cdot CV2} - r_{X\cdot CV1}) + r_{Y\cdot CV2}(r_{X\cdot CV1}r_{CV1\cdot CV2} - r_{X\cdot CV2}). \text{ (B4)}$$

Assume each covariate *CV* contributes equally to the correlations associated with *X*. Specifically, let $r_{X\cdot CV1} = r_{X\cdot CV2} = r_{X\cdot CV}$ (i.e., each covariate is equally correlated with *X*) and $r_{Y\cdot CV1} = r_{Y\cdot CV2} = r_{Y\cdot CV}$ (i.e., each covariate is equally correlated with *Y*). This is consistent with the assumption that the specific weights relating each element in **CV** to *X* and *Y* are not known, because the specific elements in *CV* are themselves not observed.

The left-hand side of (B4) is then:

$$-\sqrt{\left(2r^2_{Y\cdot CV} - 2r^2_{Y\cdot CV}r_{CV1\cdot CV2}\right)\left(2r^2_{X\cdot CV} - 2r^2_{X\cdot CV}r_{CV1\cdot CV2}\right)} = -2r_{Y\cdot CV}r_{X\cdot CV}(1 - r_{CV1\cdot CV2}).$$

And the right-hand side of (B4) is:

$$r_{Y\cdot CV}(r_{X\cdot CV}r_{CV1\cdot CV2} - r_{X\cdot CV}) + r_{Y\cdot CV}(r_{X\cdot CV}r_{CV1\cdot CV2} - r_{X\cdot CV}) = -2r_{Y\cdot CV}r_{X\cdot CV}(1 - r_{CV1\cdot CV2}),$$

and the equivalence in (B4) holds.

More generally, if the elements in **CV** are orthogonal or have been orthogonalized to represent unique contributions to *X* and *Y* (cf. Oster, 2019, page 192) such that $r_{CVi\cdot CVj}$ =0 for all *i≠j* then for *q* elements in **CV** the left-hand side of (B4) reduces to

$$-\sqrt{\left(r^2_{Y\cdot CV1} + r^2_{Y\cdot CV2} + ... + r^2_{Y\cdot CVq}\right)\left(r^2_{X\cdot CV1} + r^2_{X\cdot CV2} + ... + r^2_{X\cdot CVq}\right)} = -qr_{Y\cdot CV}r_{X\cdot CV}, \text{ (B5)}$$

and the ride hand side of (B4) reduces to the same:

$$-r_{Y\cdot CV1}r_{X\cdot CV1}-r_{Y\cdot CV2}r_{X\cdot CV2}-...-r_{Y\cdot CVq}r_{X\cdot CVq}=-qr_{Y\cdot CV}r_{X\cdot CV}.$$

Therefore, $\hat{\beta}_1=\frac{\hat{\sigma}_Y}{\hat{\sigma}_X}\frac{r_{X\cdot Y}-R_{X\cdot \mathbf{CV}}R_{Y\cdot \mathbf{CV}}}{1-R^2_{X\cdot \mathbf{CV}}}$ under the assumptions that

$r_{X\cdot CV1}=r_{X\cdot CV2}=...=r_{X\cdot CVq}=r_{X\cdot CV}$ and $r_{Y\cdot CV1}=r_{Y\cdot CV2}=...=r_{Y\cdot CVq}=r_{Y\cdot CV}$ for all q orthogonal elements in **CV**.[13] In summary, multiple unobserved covariates can be represented by a single covariate for which $r_{X\cdot CV}=R_{X\cdot \mathbf{CV}}$ and $r_{Y\cdot CV}=R_{Y\cdot \mathbf{CV}}$ under the assumption that the covariates are orthogonal to one another and each unobserved covariate is equally predictive of X and each is equally predictive of Y.

[13] Note that the equivalence will also hold if each covariate is equally correlated with X and Y. That is, that $r_{Y\cdot CV1}=r_{X\cdot CV1}=r_{.CV1}$, and $r_{Y\cdot CV2}=r_{X\cdot CV2}=r_{.CV2}$. Cinelli and Hazlett (2020, page 51) make a similar assumption and Oster (2019, page 192) employs a similar but slightly weaker assumption in deriving the Restricted Estimator. But this assumption would force $R_{\mathbf{CV}\cdot Y}=R_{\mathbf{CV}\cdot X}$, a highly restrictive condition.

Appendix C: Obtaining $R_{Y\cdot\mathbf{Z}}$, $R_{X\cdot\mathbf{Z}}$ and $r_{X\cdot\mathbf{Y}}$

We develop our derivation for $\tilde{\beta}$>0, with the case for $\tilde{\beta}$<0 obtained by symmetry.

Obtaining $R_{Y\cdot Z}$

From Cohen & Cohen (1983, page 143):

$$r^2_{X\cdot Y|\mathbf{Z}} = \frac{R^2_{Y\cdot X\mathbf{Z}} - R^2_{Y\cdot\mathbf{Z}}}{1 - R^2_{Y\cdot\mathbf{Z}}}$$

$$\Rightarrow R_{Y\cdot\mathbf{Z}} = \sqrt{\frac{R^2_{Y\cdot X\mathbf{Z}} - r^2_{X\cdot Y|\mathbf{Z}}}{(1 - r^2_{X\cdot Y|\mathbf{Z}})}} \text{ for } \frac{R^2_{Y\cdot X\mathbf{Z}} - r^2_{X\cdot Y|\mathbf{Z}}}{(1 - r^2_{X\cdot Y|\mathbf{Z}})} > 0. \text{ (C1)}$$

where $r_{X\cdot Y|\mathbf{Z}} = \frac{t(\tilde{\beta}_1)}{\sqrt{df + t(\tilde{\beta}_1)^2}}$, with *df*=*n*-*number of covariates* -2 (for *X* and the intercept) and

$t(\tilde{\beta}_1) = \frac{\tilde{\beta}_1}{se(\tilde{\beta}_1)}$. Thus $R_{Y\cdot\mathbf{Z}}$ can be calculated from reported quantities $\tilde{\beta}_1$, $\text{se}(\tilde{\beta}_1)$, *n*, *number of covariates* and $R^2_{Y\cdot X\mathbf{Z}}$ (the unadjusted R^2 when regressing *Y* on *X* and **Z**).

Obtaining $R_{X\cdot Z}$

The expression for $\text{Se}(\tilde{\beta}_1)$ from Cohen and Cohen (1983, page 109):

$$\text{Se}(\tilde{\beta}_1) = \frac{\hat{\sigma}_Y}{\hat{\sigma}_X}\sqrt{\frac{1 - R^2_{Y\cdot X,\mathbf{Z}}}{df}}\sqrt{\frac{1}{1 - R^2_{X\cdot\mathbf{Z}}}}. \qquad \text{(C2)}$$

Solving for $R_{X\cdot Y|\mathbf{Z}}$ yields

$$R_{X\cdot\mathbf{Z}} = \sqrt{1 - \frac{\hat{\sigma}^2_Y(1 - R^2_{Y\cdot X\mathbf{Z}})}{\hat{\sigma}^2_X df\left(\text{Se}(\tilde{\beta}_1)\right)^2}}, \text{ for } 1 - \frac{\hat{\sigma}^2_Y(1 - R^2_{Y\cdot X\mathbf{Z}})}{\hat{\sigma}^2_X df\left(\text{Se}(\tilde{\beta}_1)\right)^2} > 0 \quad . \quad \text{(C3)}$$

Note that all terms are positive, so the equalities hold. Therefore, $R_{X \cdot \mathbf{Z}}$ can be calculated from reported quantities $\text{se}(\tilde{\beta}_1)$, $R^2_{Y \cdot X\mathbf{Z}}, \hat{\sigma}_X$, $\hat{\sigma}_Y$ and n. Note that given $R_{X \cdot \mathbf{Z}}$ and $\hat{\sigma}_X$, one can also obtain Oster's $\tau_X = (1 - R^2_{X \cdot \mathbf{Z}})\hat{\sigma}^2_X$.

In summary, $R_{X \cdot \mathbf{Z}}$, and $R_{Y \cdot \mathbf{Z}}$ depend only on conventionally reported quantities $\hat{\sigma}_X$, $\hat{\sigma}_Y$ and n and *number of covariates* as well as $\tilde{\beta}_1, \text{se}(\tilde{\beta}_1)$ and $R^2_{Y \cdot X\mathbf{Z}}$. That is, one does not need direct access to **Z** to generate the sufficient statistics associated with **Z** to estimate model (1b) in the main text. This allows one to apply the analysis in the derivation in the main text to any analysis that reports $\hat{\sigma}_X$, $\hat{\sigma}_Y$, n and *number of covariates* as well as $\hat{\beta}_1, \text{se}(\tilde{\beta}_1)$ and $R^2_{Y \cdot X\mathbf{Z}}$ as we do below.

Obtaining $r_{X \cdot Y}$

Having obtained $R^2_{X \cdot \mathbf{Z}}$, and $R^2_{Y \cdot \mathbf{Z}}$ we seek to calculate $r_{X \cdot Y}$. This is necessary only to calculate a baseline regression for Oster's δ^* in the simulated scenarios. The term $r_{X \cdot Y}$ is not necessary to calculate the coefficient of proportionality, δ. In empirical examples $r_{X \cdot Y}$ can often be directly obtained from a reported zero-order correlation or from a baseline model in which no covariates are included. But $r_{X \cdot Y}$ can also be calculated from $\tilde{\beta}_1$, $\hat{\sigma}_Y$, $\hat{\sigma}_X, R_{X \cdot \mathbf{Z}}$, and $R_{Y \cdot \mathbf{Z}}$ under the assumption $r_{X \cdot Z1} = r_{X \cdot Z2} = ... = r_{X \cdot Z\cdot}$ and $r_{Y \cdot Z1} = r_{Y \cdot Z2} = ... = r_{Y \cdot Z\cdot}$ for elements in all **Z** (see Appendix B). That is, if we do not have information about the specific covariates we assume they are equally weighted in predicting *X* and in predicting *Y*. Under this assumption

$$\tilde{\beta}_1 = \frac{\hat{\sigma}_Y}{\hat{\sigma}_X} \frac{r_{X \cdot Y} - R_{Y \cdot \mathbf{Z}} R_{X \cdot \mathbf{Z}}}{1 - R^2_{X \cdot \mathbf{Z}}} \Rightarrow r_{X \cdot Y} = \frac{\hat{\sigma}_X}{\hat{\sigma}_Y} \tilde{\beta}_1 \left(1 - R^2_{X \cdot \mathbf{Z}}\right) + R_{Y \cdot \mathbf{Z}} R_{X \cdot \mathbf{Z}}. \quad \text{(C4)}$$

Numerical Example

This example is taken from Oster's (2019) analysis of *low birth weight and preterm* (*LBW+Preterm*) on *IQ*.

Oster's Table 3 shows:

	Panel A: Child IQ, standardized (N	
Treatment variable	(1) Baseline effect (Std. error), [R^2]	(2) Controlled effect (Std. error), [R^2]
Breastfeed (Months)	0.045*** (0.003) [0.045]	0.017***(0.002) [0.256]
Drink in Preg. (Any)	0.176***(0.026) [0.008]	0.050**(0.023) [0.249]
LBW + Preterm	– 0.188***(0.057) [0.004]	– 0.125***(0.050) [0.251]

For the model in the second column with controls, the estimated effect of *LBW+Preterm* is -.125, with standard error of .050 and R^2 of .251. The covariates include *age*, *child female*, *mother Black*, *mother age*, *mother education*, *mother income*, *mother married*. But the coefficients for the covariates are not reported.

Descriptive statistics are reported in a supplemental file with $\hat{\sigma}_X = .217$ and $\hat{\sigma}_Y = .991$.

Table 2: **Summary Statistics: Early Life and Child IQ**

Panel A: IQ Analysis			
	Mean	*Standard Deviation*	*Sample Size*
IQ (PIAT Score, Standardized)	0.025	0.991	6962
Breastfeeding Months	2.40	4.63	6514
LBW + Preterm	0.049	0.217	6174

We note the discrepancy in sample sizes for *IQ* (6962) versus *LBW* + *Preterm* (6174), presumably because of missing data. We choose a sample size of 6174 assuming listwise deletion. Note that larger sample sizes will produce smaller values of δ and a more extreme comparison between δ and δ^* in the main text. For sample size of 6174, $R^2_{X\cdot\mathbf{Z}}$ is very small and potentially less than zero depending on rounding of other reported values. In fact, for $se(\tilde{\beta}_1)$ =.050, $R^2_{X\cdot Z}$ is less than zero and $R_{X\cdot\mathbf{Z}}$ is not defined, and model (1b) cannot be estimated using conventional techniques. We note that when $R^2_{X\cdot Z}$ is small its calculation is very sensitive to the standard error (relative to $\frac{\hat{\sigma}_Y}{\hat{\sigma}_X}$ and for given sample size and $\tilde{R}^2$). Specifically, if $se(\tilde{\beta}_1)$ had been rounded from .05034 to .050 then $R^2_{X\cdot Z}$ would be .00011 with $R_{X\cdot\mathbf{Z}}$ =.0107. But such a small relationship between observed covariates **Z** and the predictor of interest *X* challenges the premise of using observed covariates as a baseline to evaluate unobserved covariates. Therefore, we set $se(\tilde{\beta}_1)$ =.05049, the largest possible value (to five decimals) that could be rounded to .050.

Thus, we have:

$\tilde{\beta}_1 = .125,\ \text{se}(\tilde{\beta}_1) = .050049,\ R^2 = .251, \hat{\sigma}_X = .217,\ \hat{\sigma}_Y = .991$ *n*= 6265 and *number of covariates*=7.

To generate $R^2_{Y\cdot\mathbf{Z}}$ and $R^2_{X\cdot\mathbf{Z}}$, begin with

$$t(\tilde{\beta}_1) = \frac{.125}{.05049} = 2.476\text{, and } r_{X\cdot Y|Z} = \frac{t(\tilde{\beta}_1)}{\sqrt{n-7-2+t(\tilde{\beta}_1)^2}} = \frac{2.476}{\sqrt{6174-7-2+2.476^2}} = .032.$$

Therefore, from (C1)

$$R_{Y\cdot\mathbf{Z}} = \sqrt{\frac{R^2_{Y\cdot X\mathbf{Z}} - r^2_{X\cdot Y|\mathbf{Z}}}{(1-r^2_{X\cdot Y|\mathbf{Z}})}} = \sqrt{\frac{.251-.032^2}{(1-.032^2)}} = .500\ .$$

And from (C3)

$$R_{X\cdot\mathbf{Z}}=\sqrt{1-\frac{\hat{\sigma}_Y^2\left(1-R_{Y\cdot X\mathbf{Z}}^2\right)}{\hat{\sigma}_X^2 df\left(\text{Se}(\tilde{\beta}_1)\right)^2}}=\sqrt{1-\frac{.991^2\left(1-.251\right)}{.217^2(6165)\left(.05049\right)^2}}=.078.$$

(Correspondingly for Oster's δ^*, $\tau_X=(1-R_{X\cdot\mathbf{Z}}^2)\hat{\sigma}_X^2=(1-.078^2).217^2=.047$).

A regression following model (1b) in the main text can be recovered from the values for $R_{X\cdot\mathbf{Z}}$ and $R_{Y\cdot\mathbf{Z}}$ as well as $r_{Y\cdot X|\mathbf{Z}}$ obtained from $t(\tilde{\beta}_1)$ above and $\hat{\sigma}_X$, $\hat{\sigma}_Y$, $R_{Y\cdot X\mathbf{Z}}^2$ n, and *number of covariates* (q) as reported:

$$\tilde{\beta}_1=\frac{\hat{\sigma}_{Y|\mathbf{Z}}}{\hat{\sigma}_{X|\mathbf{Z}}}r_{X\cdot Y|\mathbf{Z}}=\frac{\hat{\sigma}_Y\sqrt{1-R_{Y\cdot\mathbf{Z}}^2}}{\hat{\sigma}_X\sqrt{1-R_{X\cdot\mathbf{Z}}^2}}r_{X\cdot Y|\mathbf{Z}}=\frac{.991\sqrt{1-.500^2}}{.217\sqrt{1-.078^2}}.032=.125,$$

$$se(\tilde{\beta}_1)=\frac{\hat{\sigma}_Y}{\hat{\sigma}_X}\times\sqrt{\frac{1-R_{Y\cdot X\mathbf{Z}}^2}{n-q-2}\times\frac{1}{1-R_{X\cdot\mathbf{Z}}^2}}=\frac{.991}{.217}\times\sqrt{\frac{1-.251}{6174-7-2}\times\frac{1}{1-.078^2}}=.05049,\text{ and}$$

$$\tilde{R}=R_{Y\cdot X\mathbf{Z}}^2=R_{Y\cdot\mathbf{Z}}^2+\frac{r_{X\cdot Y|\mathbf{Z}}^2}{1-R_{X\cdot\mathbf{Z}}^2}=.500^2+\frac{.032^2}{1-.078^2}=.251.$$

These correspond to Oster's originally reported values.

Last, from (C4)

$$r_{X\cdot Y}=\frac{\hat{\sigma}_X}{\hat{\sigma}_Y}\tilde{\beta}_1\left(1-R_{X\cdot\mathbf{Z}}^2\right)+R_{X\cdot\mathbf{Z}}R_{Y\cdot\mathbf{Z}}=\frac{.217}{.991}.125\left(1-.078^2\right)+(.078)(.500)=.066$$

Correspondingly, the estimate of β_1 from an unconditional model is $\mathrm{r}_{X\cdot Y}\frac{\hat{\sigma}_Y}{\hat{\sigma}_X}=.066\frac{.991}{.217}=.302$ with R^2 of .004.

A regression for model (1b) including observed covariates based on zero-order correlations yields:

$$\tilde{\beta}_1=\frac{\hat{\sigma}_Y}{\hat{\sigma}_X}\frac{r_{X\cdot Y}-R_{X\cdot\mathbf{Z}}R_{Y\cdot\mathbf{Z}}}{1-R_{X\cdot\mathbf{Z}}^2}=\frac{.991}{.217}\frac{.066-(.078)(.500)}{1-.078^2}=.125,$$

$$\tilde{R}^2 = \frac{r_{X\cdot Y}^2 + R_{Y\cdot\mathbf{Z}}^2 - 2r_{X\cdot Y}R_{Y\cdot\mathbf{Z}}R_{X\cdot\mathbf{Z}}}{1-R_{X\cdot\mathbf{Z}}^2} = \frac{.066^2 + .500^2 - 2(.066)(.500)(.078)}{1-.078^2} = .251$$

$$se(\tilde{\beta}_1) = \frac{\hat{\sigma}_Y}{\hat{\sigma}_X} \times \sqrt{\frac{1-R_{Y\cdot X}^2}{n-q-1} \times \frac{1}{1-R_{X\cdot\mathbf{Z}}^2}} = \frac{.991}{.217} \times \sqrt{\frac{1-.251}{6174-7-2} \times \frac{1}{1-.078^2}} = .05049$$

We verify the results by constructing the correlation matrix:

$$\begin{bmatrix} & \text{IQ} & \text{LBW} & \text{Z} \\ \text{IQ} & 1 & .0530 & .5003 \\ \text{LBW} & .0530 & 1 & .0515 \\ \text{Z} & .5003 & .0515 & 1 \end{bmatrix}.$$

And then implemeting the Lavaan package in R:

```
Regressions:
                   Estimate  Std.Err  z-value  P(>|z|)
  Y ~
    X       (bet1)    0.125    0.050    2.478    0.013
    Z       (bet2)    0.494    0.011   45.088    0.000

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .Y                 0.735    0.013   55.561    0.000
```

with

$$R^2 = \frac{.991^2 - .735}{.991^2} = .251 \ .$$

These correspond to Oster's originally reported values $\tilde{\beta}_1 = .125$, $\text{se}(\tilde{\beta}_1) = .050$, and $\tilde{R}^2 = .251$.

Appendix D: Obtaining $r_{X\cdot CV}$ and $r_{Y\cdot CV}$ from $r_{X\cdot CV|Z}$ and $r_{Y\cdot CV|Z}$ assuming CV is Orthogonal to each Element in Z

We show how to obtain $r_{X\cdot CV}$ and $r_{Y\cdot CV}$ from $r_{X\cdot CV|\mathbf{Z}}$ and $r_{Y\cdot CV|\mathbf{Z}}$ assuming the confounding variable *CV* is orthogonal to each element in **Z**.

To begin, $r_{X\cdot CV|Z1Z2} = \dfrac{r_{X\cdot CV|Z2} - r_{X\cdot Z1|Z2}r_{CV\cdot Z1|Z2}}{\sqrt{(1-r^2_{X\cdot Z1|Z2})(1-r^2_{CV\cdot Z1|Z2})}} = \dfrac{r_{X\cdot CV|Z2}}{\sqrt{(1-r^2_{X\cdot Z1|Z2})}}$,

because $r_{CV\cdot Z1|Z2}=0$ by assumption of orthogonality of *CV* with *Z1* and *Z2.* And

$$r_{X\cdot CV|Z2} = \frac{r_{X\cdot CV} - r_{X\cdot Z2}r_{CV\cdot Z2}}{\sqrt{(1-r^2_{X\cdot Z2})(1-r^2_{CV\cdot Z2})}} = \frac{r_{X\cdot CV}}{\sqrt{(1-r^2_{X\cdot Z2})}}$$

because $r_{CV\cdot Z2}=0$ by assumption of orthogonality of *CV* with *Z2*.

Note

$$r_{X\cdot Z1|Z2} = \frac{r_{X\cdot Z1} - r_{X\cdot Z2}r_{Z1\cdot Z2}}{\sqrt{(1-r^2_{X\cdot Z2})(1-r^2_{Z1\cdot Z2})}}, \text{ and } r^2_{X\cdot Z1|Z2} = \frac{(r_{X\cdot Z1} - r_{X\cdot Z2}r_{Z1\cdot Z2})^2}{(1-r^2_{X\cdot Z2})(1-r^2_{Z1\cdot Z2})}.$$

Then

$$r_{X\cdot CV|Z1Z2} = \frac{r_{X\cdot CV|Z2}}{\sqrt{(1-r^2_{X\cdot Z1|Z2})}} = \frac{\dfrac{r_{X\cdot CV}}{\sqrt{(1-r^2_{X\cdot Z2})}}}{\sqrt{(1-\dfrac{(r_{X\cdot Z1} - r_{X\cdot Z2}r_{Z1\cdot Z2})^2}{(1-r^2_{X\cdot Z2})(1-r^2_{Z1\cdot Z2})})}}$$

$$= \frac{r_{X\cdot CV}}{\sqrt{\dfrac{1-r^2_{Z1\cdot Z2}}{1-r^2_{Z1\cdot Z2}} - \dfrac{r^2_{X\cdot Z2} + r^2_{X\cdot Z1} - 2r_{X\cdot Z1}r_{X\cdot Z2}r_{Z1\cdot Z2}}{1-r^2_{Z1\cdot Z2}}}} = \frac{r_{X\cdot CV}}{\sqrt{1-R^2_{X\cdot Z1Z2}}}.$$

Applying the approach recursively:

$$r_{X\cdot CV|\mathbf{Z}} = \frac{r_{X\cdot CV}}{\sqrt{1-R_{X\cdot\mathbf{Z}}^2}}.$$

This implies

$$r_{X\cdot CV} = \sqrt{1-R_{X\cdot\mathbf{Z}}^2}\, r_{X\cdot CV|\mathbf{Z}}.$$

Analogously, $r_{Y\cdot CV} = \sqrt{1-R_{Y\cdot\mathbf{Z}}^2}\, r_{Y\cdot CV|\mathbf{Z}}$.

Appendix E: Confirmation of the recovery of $\hat{\beta}_1$ and $R^2_{Y\cdot XCV|\mathbf{Z}}$ from $r_{X\cdot Y|\mathbf{Z}}$, $r_{X\cdot CV|\mathbf{Z}}$, and $r_{Y\cdot CV|\mathbf{Z}}$

We present the case for $\tilde{\beta}_1$>0 with the case for $\tilde{\beta}_1$<0 obtained by symmetry.

Calculations in the main text generate the following correlation matrix for $\beta^{\#}$=0 and R_{Max}=.61.

$$\begin{bmatrix} & IQ|\mathbf{Z} & LBW|\mathbf{Z} & CV|\mathbf{Z} \\ IQ|\mathbf{Z} & 1 & .032 & .693 \\ LBW|\mathbf{Z} & .032 & 1 & .045 \\ CV|\mathbf{Z} & .693 & .045 & 1 \end{bmatrix}.$$

Using this correlation matrix a linear model $Y|\mathbf{Z}=\beta_0+\beta_1X|\mathbf{Z}+\beta_2CV|\mathbf{Z}$ was estimated using the Lavaan procure in R.

```
Regressions:
                   Estimate  Std.Err  z-value  P(>|z|)
  YGZ ~
    XGZ     (bet1)    0.000    0.009    0.000    1.000
    CVGZ    (bet2)    0.693    0.009   75.387    0.000

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .YGZ               0.520    0.009   55.561    0.000
```

Note $\hat{\beta}_1$=0. Also, $R^2_{Y\cdot XCV|\mathbf{Z}}=1-\hat{\sigma}^2_{Y|XCV\mathbf{Z}}=1-.520=.480$. Therefore, the combined R^2 is

$R^2_{Y\cdot XCV\mathbf{Z}}=R^2_{Y\cdot\mathbf{Z}}+(1-R^2_{Y\cdot\mathbf{Z}})R^2_{Y\cdot XCV|\mathbf{Z}}=.500^2+(1-.500^2).480=.61$. Thus, the obtained values $r_{X\cdot CV|\mathbf{Z}}$ =.045 and $r_{Y\cdot CV|\mathbf{Z}}$ =.693 (for $R_{X\cdot\mathbf{Z}}$=.078 and $R_{Y\cdot\mathbf{Z}}$=.500) with $r_{X\cdot Y|\mathbf{Z}}$=.032 generate the specified values of $\hat{\beta}_1=\beta^{\#}$=0 and $R^2_{Y\cdot XCVZ}=.61$ for model (1c).

Apendix F: Oster's definition of δ

Oster (page 192) defines δ in terms of

$$\delta \frac{\sigma_{1X}}{\sigma_1^2} = \frac{\sigma_{2X}}{\sigma_2^2} \Rightarrow \delta = \frac{\frac{\sigma_{2X}}{\sigma_2^2}}{\frac{\sigma_{1X}}{\sigma_1^2}}, \text{(F1)}$$

where σ_{1X} is the covariance between X and the observed covariates represented by W_1; and σ_{2X} is the covariance between X and the unobserved covariates represented by W_2, our confounding variable (*CV*). The terms σ_1^2 and σ_1^2 are the variances of W_1 and W_2 respectively. More specifically, W_1 is defined as $W_1 = \mathbf{\Psi Z}$ and $W_2 = CV$ from the following model (Oster, pages 191-192):

$$Y = \beta_1 X + W_1 + W_2 = \beta_1 X + \mathbf{\Psi Z} + W_2,$$

where the absence of a coefficient associated with W_2 implies the coefficient is 1. Most importantly, the implication of setting $W_1 = \mathbf{\Psi Z}$ is that elements in **Z** representing selection on observables are weighted by their relationship to Y and that $\sigma_1^2 = R^2_{Y\cdot\mathbf{Z}|X} \sigma^2_{Y|X}$ because W_1 is a predicted value (see also Altonji Elder & Tabor, 2005, page 175). In these ways the relationship between **Z** and Y is taken into account in expressing the selection into X based on **Z** in (F1).

Note that for a single Z, Oster's definition of δ can be expressed as $\delta = \gamma_2/\gamma_1$ where γ_2 and γ_1 can be defined from:

$$X = \gamma_0 + \gamma_1 Z + \gamma_2 CV.$$

In this circumstance, δ can be expressed as

$$\delta = \frac{\dfrac{\hat{\sigma}_X}{\hat{\sigma}_{CV}} \dfrac{r_{X \cdot CV} - r_{CV \cdot Z} r_{X \cdot Z}}{1 - r_{CV \cdot Z}^2}}{\dfrac{\hat{\sigma}_X}{\hat{\sigma}_Z} \dfrac{r_{X \cdot Z} - r_{CV \cdot Z} r_{X \cdot CV}}{1 - r_{CV \cdot Z}^2}} = \frac{\dfrac{\hat{\sigma}_X}{\hat{\sigma}_{CV}} r_{X \cdot CV}}{\dfrac{\hat{\sigma}_X}{\hat{\sigma}_Z} r_{X \cdot Z}} = \frac{r_{X \cdot CV}}{r_{X \cdot Z}}, \quad \text{(F2)}$$

assuming $r_{CV \cdot Z}=0$, and that the scale for the unobserved variable *CV* is chosen such that $\hat{\sigma}_{CV} = \hat{\sigma}_Z$. Therefore, the expressions for δ in (F1) and (F2) differ only in how they weight the elements in **Z**, with Oster's definition in (F2) weighting in proportion to contributions to the prediction of *Y* through $W_1 = \mathbf{\Psi Z}$ and the definition in (F2) weighting in proportion to the contributions to *X* through $R_{X \cdot \mathbf{Z}}$. Note that Oster (page 192) assumes the contributions of **Z** to *X* and to *Y* are proportional to one another in deriving the Restricted Estimator, limiting the differences between the definitions in (F1) and (F2).