

The Effect of Omitted Variables on the Sign of Regression Coefficients*

Matthew A. Masten[†] Alexandre Poirier[‡]

August 1, 2022

Abstract

Omitted variables are a common concern in empirical research. We show that “Oster’s delta” (Oster 2019), a commonly reported measure of regression coefficient robustness to the presence of omitted variables, does not capture sign changes in the parameter of interest. Specifically, we show that any time this measure is large—suggesting that omitted variables may be unimportant—a much smaller value can actually reverse the sign of the parameter of interest. Relatedly, we show that selection bias adjusted estimands can be extremely sensitive to the choice of the sensitivity parameter. Specifically, researchers commonly compute a bias adjustment under the assumption that Oster’s delta equals one. Under the alternative assumption that delta is very close to one, but not exactly equal to one, we show that the bias can instead be arbitrarily large. To address these concerns, we propose a modified measure of robustness that accounts for such sign changes, and discuss best practices for assessing sensitivity to omitted variables. We demonstrate this sign flipping behavior in an empirical application to social capital and the rise of the Nazi party, where we show how it can overturn conclusions about robustness, and how our proposed modifications can be used to regain robustness. We implement our proposed methods in the companion Stata module `regsensitivity` for easy use in practice.

JEL classification: C14; C18; C21; C25; C51

Keywords: Identification, Treatment Effects, Partial Identification, Sensitivity Analysis, Unconfoundedness

*We thank Paul Diegert for many helpful conversations. We thank Hongchang Guo and Paul Diegert for excellent research assistance. Masten thanks the National Science Foundation for research support under Grant 1943138.

[†]Department of Economics, Duke University, matt.masten@duke.edu

[‡]Department of Economics, Georgetown University, alexandre.poirier@georgetown.edu

1 Introduction

Most analysis of causality relies on untestable assumptions, like the absence of unobserved omitted variables that could bias one’s findings. A large literature in statistics and econometrics now provides many tools that allow researchers to perform sensitivity analyses to assess the importance of these assumptions. Many of these methods use *breakdown points* as quantitative measures of the robustness of one’s conclusions to departures from the baseline identifying assumptions. These points tell us how much we can depart from our identifying assumptions before our empirical conclusions break down in some way.

In this paper, we distinguish between two specific kinds of breakdown points. The first is the *explain away* breakdown point. It answers the question

What is the smallest value of the sensitivity parameter required for the data to be consistent with a zero causal effect?

The second is the *sign change* breakdown point. It answers the question

What is the smallest value of the sensitivity parameter required for the data to be consistent with a causal effect that has a different sign from the causal effect we found in the baseline model?

In section 2 we show that these two breakdown points are in general *not* equivalent. This typically occurs when researchers consider *deviations* from a baseline model, which implies that the identified set—the set of parameter values consistent with the data and the assumptions—does not monotonically increase as we consider models farther and farther from the baseline model. In contrast, when researchers consider *relaxations* of a baseline model, the identified set for the parameter of interest increases monotonically. When the identified set for these relaxations is an interval, we show that the two breakdown points are equivalent. We formalize these concepts and statements in section 2.

In sections 3 and 4 we study the specific sensitivity analysis of Oster (2019), designed to assess the importance of omitted variables in regression analysis. This approach has been extremely influential, with about 2600 Google Scholar citations as of July 2022. Finkelstein, Gentzkow, and Williams (2021, *AER*, page 2706) describe it as providing “the now-standard methodology” and “the standard approach” to adjusting for selection on unobservables. Moreover, from 2019–2021, 23 papers published in the top five economics journals formally compute and report the results from this method, while another 8 papers informally reference this method in support of their analysis. Based on the data collected by Blandhol, Bonney, Mogstad, and Torgovitsky (2022), this is larger than the number of papers that use 2SLS per year, among papers published in top five economics journals. We discuss in detail how Oster’s (2019) results are used in empirical economics in section 7.

In section 3 we review Oster’s (2019) method. The main sensitivity parameter for this method, denoted by δ , is commonly interpreted as the ratio of the magnitude of selection on unobservables to the magnitude of selection on observables. We then study breakdown points for this sensitivity parameter in section 4. Oster’s Proposition 2 characterizes the explain away breakdown point. Following her recommendation, empirical researchers now routinely report estimates of this explain away breakdown point as a measure of robustness to omitted variables. We show that this explain away breakdown point is very different from the sign change breakdown point. In fact, we show that any time the explain away breakdown point is large—suggesting that omitted variables may be unimportant—a much smaller value of her sensitivity parameter can actually change the sign of the parameter of interest. In particular, we prove that the sign change breakdown point is bounded above by 1. This is a substantial concern since 1 is often viewed as the cutoff for a robust result.

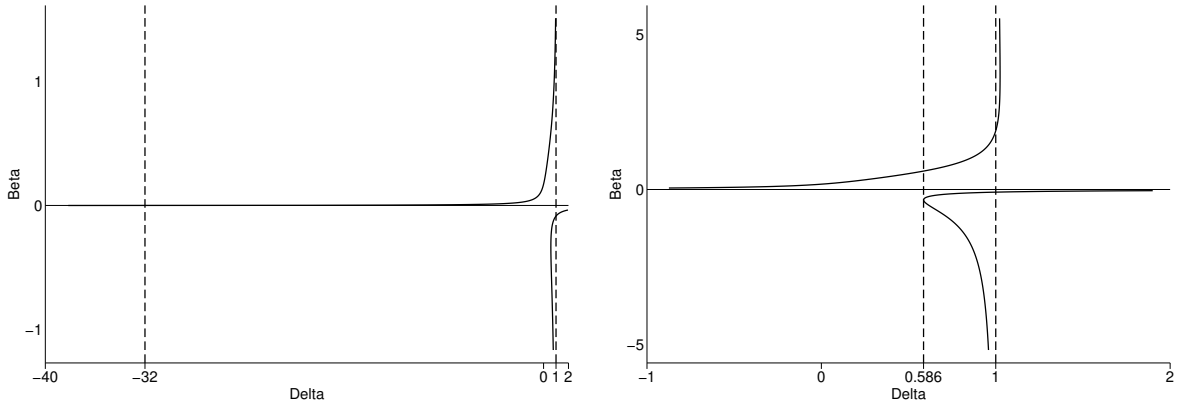


Figure 1: Example of Oster’s identified set $\mathcal{B}_I(\delta, R_{\text{long}}^2)$ as a function of the sensitivity parameter δ (see section 3 for a definition of this set and a description of the R_{long}^2 parameter). As we discuss in section 3.3, this identified set always has either one, two, or three elements. Both plots show the same set, but with different horizontal axis ranges. In this example, the explain away breakdown point is $|-32|$. This is shown on the left plot. In contrast, the sign change breakdown point is 0.586, as shown in the right plot. The plot also shows the vertical asymptote at $\delta = 1$ which we discuss in section 4.1.

Consequently, if we maintain 1 as the cutoff, our result implies that *no empirical results are robust to sign changes*, using Oster’s method. Put differently: The true causal effect can have the opposite sign as the baseline estimand any time the magnitude of selection on unobservables is at least as large as the magnitude of selection on observables.

In section 5 we study selection bias adjusted estimands. We show that the most commonly used adjustment is extremely sensitive to perturbations of Oster’s sensitivity parameter. We also show that one of the key assumptions this adjustment relies on is testable and develop a simple hypothesis test for it. We conclude section 5 by explaining a common but incorrect approach to computing explain away breakdown points.

To build intuition for our main results in section 4, consider figure 1. This is an example from our empirical application in section 6. Both plots show the set of regression coefficient values β consistent with the data and the assumptions—this is called the *identified set*—as we vary Oster’s main sensitivity parameter δ . Both plots show the same curves, but with different horizontal axis ranges. The baseline estimate is the vertical intercept at $\delta = 0$. This estimate is positive. How robust is this estimate to the presence of omitted variables? In the left plot we see that we must go all the way to $\delta = -32$ before the value $\beta = 0$ is consistent with the data and assumptions. This is the number that researchers commonly report as a measure of robustness. However, from the right plot, we see that we only need to go to $\delta = 0.586$ to find a value of β that is negative. This value $\delta = 0.586$ is the sign change breakdown point. Oster (2019) does not discuss this breakdown point, the companion Stata package `psacalc` does not compute it, and none of the top five papers we survey in section 7 report it. Yet, as we see here, the sign change breakdown point is substantially smaller than $|-32|$, the explain away breakdown point. Indeed, it is below $\delta = 1$, the commonly used cutoff for a robust result. Consequently, in this empirical example, focusing on the sign change breakdown point overturns the authors’ conclusion of robustness: Only a small amount of selection on unobservables relative to observables is necessary to flip the sign of the baseline estimate.

Figure 1 is also useful for building intuition for our results in section 5 on the sensitivity of bias adjust-

ments. In the figure, we see there is an asymptote at $\delta = 1$. This implies that bias adjustments computed at δ exactly equal to one—which is currently standard empirical practice—are *not* representative of bias adjustments computed at other values of δ . This explains our discussion in section 5.6 on why Oster’s (2019) “bounding set” does not correctly represent the values of the regression coefficient consistent with any δ between 0 and 1. Importantly, *all* of the coefficient values plotted in figure 1, even the seemingly extreme values, are equally consistent with the data and assumptions. Under the assumptions of Oster’s Proposition 2 (see our Theorem 2), which is used to obtain these plots, there is no justification for ignoring any of these values. The only way to eliminate values is by making additional assumptions. One reasonable assumption, which we consider in section 4.2, is to restrict the possible magnitudes of the coefficient of interest. With this additional assumption, we can justifiably eliminate some of the most extreme values. This assumption does not overturn our main conclusions, however. For example, in figure 1 the coefficient value consistent with $\delta = 0.586$ is not extreme at all—it is the same order of magnitude as the baseline estimate. So in this example, the sign change breakdown point is still very different from the explain away breakdown point, even with a substantial restriction on the magnitude of the regression coefficient. Moreover, magnitude restrictions require researchers to choose a tuning parameter—just how much should they restrict the magnitude? If the magnitude is restricted too much, then this assumption is essentially equivalent to assuming a priori that there is no omitted variable bias, in which case a sensitivity analysis would be unnecessary. But as soon as reasonably sized magnitudes are allowed, the concerns we discuss in sections 4 and 5 arise.

To address the issues described in sections 4 and 5, we propose several simple modifications to the methods in Oster (2019). First, we develop the sign change breakdown point for Oster’s main sensitivity parameter and recommend that researchers report this rather than, or in addition to, the explain away breakdown point. Second, to avoid the sensitivity inherent in the bias adjusted estimands, we recommend gathering all values of the coefficient consistent with a range of sensitivity parameters, and presenting this set. Furthermore, in section 4.2, we provide extensions to these two recommendations to formally incorporate the magnitude restrictions discussed above. Specifically, we consider restrictions on the absolute magnitude of the omitted variable bias. We show that the sign change breakdown point can be strictly larger than 1 when a magnitude restriction is imposed. Therefore it is possible to recover findings of robustness even without changing the cutoff for robustness. This is not always the case however, as we saw in figure 1.

In section 6 we illustrate our results using data from Satyanath, Voigtländer, and Voth (2017, *JPE*), who study the effect of social capital on the rise of the Nazi party. There we show just how different the sign change and explain away breakdown points can be in practice. We then show how magnitude restrictions can sometimes help retain findings of robustness. We also illustrate the extreme sensitivity of bias adjusted estimands to perturbations of the sensitivity parameter. In section 8 we conclude with a discussion of best practices for regression sensitivity analysis.

Additional Technical Contribution

In section 3 we show that the set defined in Oster’s Proposition 2 is in fact sharp, with one minor technical caveat, and therefore equals the identified set under the assumptions of that proposition. In contrast, Oster (2019) only showed that this set was an outer set, containing the true identified set, but did not discuss whether this set was sharp. This sharpness result is important to ensure that Oster’s breakdown analysis is not unnecessarily conservative. Similarly, in section 5 we provide a new proof that the result in Oster’s Proposition 1 is sharp.

Related Literature

A wide variety of papers propose methods for analyzing the sensitivity of regression based findings. This includes Mauro (1990), Murphy and Topel (1990), Frank (2000), Imbens (2003), Altonji, Elder, and Taber (2005), Clarke (2009), Bellows and Miguel (2009), Hosman, Hansen, and Holland (2010), Gonzalez and Miguel (2015), Krauth (2016), Oster (2019), Cinelli and Hazlett (2020), and Diegert, Masten, and Poirier (2022). Many other papers propose sensitivity analyses for nonlinear models, including Rosenbaum and Rubin (1983), Rosenbaum (1995, 2002), Robins, Rotnitzky, and Scharfstein (2000), Altonji, Elder, and Taber (2005, 2008), Masten and Poirier (2018), and Altonji, Conley, Elder, and Taber (2019).

Our analysis primarily focuses on Oster’s (2019) approach, since it is by far the most popular method for regression sensitivity analysis in economics. That said, our distinction between explain away and sign change breakdown points in section 2 is more general, and therefore can be applied to other methods too. For example, Imbens (2003) studies a model of treatment effects which parameterizes deviations from conditional random assignment. Figures 1–4 in Imbens (2003) plot exact value breakdown curves (curves rather than points because he uses two sensitivity parameters). Because a similar asymptotic phenomenon as shown in our figure 1 occurs in Imbens’ model, these exact value breakdown curves do not capture sign changes. Thus the sign change and explain away breakdown points will generally be quite different in his model as well. As we discuss in section 2, a similar remark generally applies for any sensitivity analysis based on deviations rather than interval relaxations. One exception is Krauth (2016). Like Oster (2019), he starts with a deviations based analysis. His deviations also induce asymptotes (see his figure 1), but he generally works with a convex hull of the identified set (his equations 15 and 16) which ensures that explain away and sign change breakdown points are the same.

The general idea of an explain away type sensitivity analysis goes back to Cornfield, Haenszel, Hammond, Lilienfeld, Shimkin, and Wynder (1959). Sign change type breakdown analysis has also been studied before; for example, see the detailed literature review in Masten and Poirier (2020, section 1 and Appendix D). However, we are unaware of any previous papers that explicitly draw the distinction between explain away and sign change breakdown points, and emphasize how different they can be when one is using a deviation based sensitivity analysis.

Notation Remark

For random vectors A and B , define $A^{\perp B} = A - [\text{var}(B)^{-1} \text{cov}(B, A)]'B$. This is the sum of the residual from a linear projection of A onto $(1, B)$ and the intercept in that projection.

2 Measuring Robustness: Sign Changes versus Explaining Away

In this section we formally define and analyze the two kinds of breakdown points discussed in section 1. We then apply this analysis in sections 3 and 4. Consider a model indexed by a scalar sensitivity parameter $r \geq 0$. We focus on the scalar case for simplicity; see Masten and Poirier (2020) for a study of breakdown analysis for vector relaxations. Let $\mathcal{B}_I(r^*) \subseteq \mathbb{R}$ denote the identified set for the parameter of interest $\beta \in \mathbb{R}$ under the assumption $r = r^*$. Suppose $r = 0$ represents the baseline model, while values $r > 0$ represent alternative assumptions about the model.

Definition 1. Say that r represents a *relaxation* of the baseline model if $\mathcal{B}_I(r_1) \subseteq \mathcal{B}_I(r_2)$ for any $r_1 \leq r_2$. Otherwise, say that r represents a *deviation* from the baseline model. Say r is an *interval* relaxation if it is

a relaxation and if $\mathcal{B}_I(r)$ is an interval for any $r \geq 0$.

The sensitivity parameter r is a relaxation when larger values lead to weakly less information about the parameter of interest, in the sense that the identified set simply becomes larger. In contrast, the sensitivity parameter r is a deviation when larger values lead to *different* information about the parameter of interest. These deviations are typically inconsistent with smaller values of r , in the sense that the identified sets $\mathcal{B}_I(r_1)$ and $\mathcal{B}_I(r_2)$ for $r_1 \leq r_2$ are not nested and may even be disjoint. As we will show, this distinction is important for understanding when the explain away and sign change breakdown points differ.

Define the *exact value breakdown point*

$$r^{\text{bp},=}(b) = \inf\{r \geq 0 : b \in \mathcal{B}_I(r)\}.$$

This is the smallest value of the sensitivity parameter such that the identified set contains b . Consequently, for all $r < r^{\text{bp},=}(b)$, we know that b is not in the identified set $\mathcal{B}_I(r)$. The *exact zero* or *explain away* breakdown point is the special case where $b = 0$. Define

$$r^{\text{bp},<}(\bar{b}) = \sup\{r \geq 0 : b < \bar{b} \text{ for all } b \in \mathcal{B}_I(r)\}.$$

This is the largest value of the sensitivity parameter such that we can conclude for sure that $\beta < \bar{b}$. That is, for all $r < r^{\text{bp},<}(\bar{b})$, we know $\beta < \bar{b}$. Similarly, define

$$r^{\text{bp},>}(\underline{b}) = \sup\{r \geq 0 : b > \underline{b} \text{ for all } b \in \mathcal{B}_I(r)\}.$$

This is the largest value of the sensitivity parameter such that we can conclude for sure that $\beta > \underline{b}$. That is, for all $r < r^{\text{bp},>}(\underline{b})$ we know $\beta > \underline{b}$. Suppose β is point identified in the baseline model, so that $\mathcal{B}_I(0) = \{\beta_{\text{baseline}}\}$ is a singleton. Define the *sign change breakdown point* as the value

$$r^{\text{bp},\text{sign}} = \begin{cases} r^{\text{bp},>}(0) & \text{if } \beta_{\text{baseline}} > 0 \\ r^{\text{bp},<}(0) & \text{if } \beta_{\text{baseline}} < 0. \end{cases}$$

This is the largest value of the sensitivity parameter such that we can conclude that β has the same sign as our baseline estimand. That is, for all $r < r^{\text{bp},\text{sign}}$, $\text{sign}(b) = \text{sign}(\beta_{\text{baseline}})$ for all b in the identified set $\mathcal{B}_I(r)$. For any larger values of r , the sign of the true causal effect can be different from the sign of the baseline estimand. Note that $r^{\text{bp},\text{sign}}$ is not defined if $\beta_{\text{baseline}} = 0$, a case in which a sign change sensitivity analysis is unnecessary.

Our first result relates the two kinds of breakdown points defined above.

Theorem 1. Suppose $\mathcal{B}_I(0) = \{\beta_{\text{baseline}}\}$. Suppose r is an interval relaxation. Then

$$\bar{r}^{\text{bf},=}(b) = \begin{cases} \bar{r}^{\text{bf},>}(b) & \text{for } b < \beta_{\text{baseline}} \\ \bar{r}^{\text{bf},<}(b) & \text{for } b > \beta_{\text{baseline}} \end{cases}$$

for $b \neq \beta_{\text{baseline}}$.

In particular, when r is an interval relaxation, the explain away breakdown point and the sign change breakdown point are the same. However, when r is a non-interval relaxation or a deviation, the result in

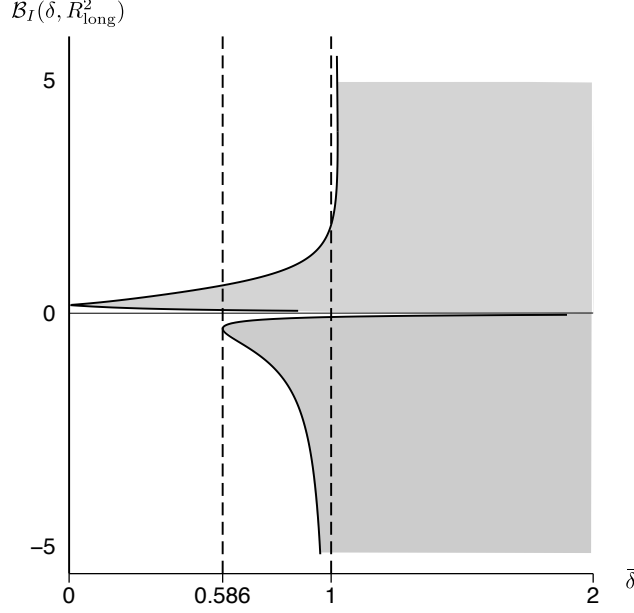


Figure 2: The cumulative identified set for the example of figure 1. Here we see that for large values of $\bar{\delta}$ the identified set includes every value in \mathbb{R} except a small interval around 0. The set includes 0 for $\bar{\delta} \geq 32$, the explain away breakdown point. But the sign change breakdown point occurs much earlier, at $\delta = 0.586$. Moreover, note that the identified set at $\bar{\delta} = 1$ is $(-\infty, -0.0855] \cup [0.0432, 1.8947]$. At $\bar{\delta} = 1.05$ it is $(-\infty, -0.0798] \cup [0.0415, \infty)$.

Theorem 1 generally does not hold. Consequently, in these cases, the explain away breakdown point and sign change breakdown points are generally different. We explore the implications of this difference in a specific model in the next few sections.

Figure 1 illustrates the difference between these two breakdown points, using data from our empirical application in section 6. Here we plot an example of the identified set we study in section 3. In this example, the explain away breakdown point is much larger than the sign change breakdown point. This arises because the identified set $\mathcal{B}_I(r)$ is discontinuous in the sensitivity parameter r , in the sense that there is a point r_0 such that the Hausdorff distance between $\mathcal{B}_I(r)$ and $\mathcal{B}_I(r_0)$ does not converge to zero as $r \rightarrow r_0$.

Another way to understand this distinction between the sign change and explain away breakdown points is to consider the *cumulative* identified set,

$$\mathcal{B}_I(\bar{r}) = \bigcup_{|r| \leq \bar{r}} \mathcal{B}_I(r).$$

Figure 2 plots this set for the example in figure 1. We discuss this specific application more in section 4.2. The sensitivity parameter \bar{r} is a relaxation since the identified sets $\mathcal{B}_I(\bar{r})$ weakly increase as \bar{r} increases. However, it is not an interval relaxation. Consequently, the sign change and explain away breakdown points are still different.

We conclude this section with one additional useful feature of interval relaxations.

Proposition 1. Suppose r is an interval relaxation. Then

$$\mathcal{B}_I(r) = \{b \in \mathbb{R} : r^{\text{bf},=}(b) \leq r\}.$$

Proposition 1 shows that knowledge of the exact value breakdown point can be used to characterize identified sets, when r is an interval relaxation. This is a kind of duality result, which shows that one can always go back and forth between the identified set and the exact value breakdown point.

3 Setup and Identification Analysis

In this section we review the main identification results in Oster (2019), including her breakdown analysis. Importantly, we provide one new result in this section: We show that the set of possible biases due to omitted variables given in her Proposition 2 is sharp, up to a minor technical caveat. Consequently, that set can be correctly interpreted as the identified set for the bias. Her proof of Proposition 2 only showed that this set contained the true bias; it did not prove sharpness. This sharpness result is important because it shows that breakdown analysis based on this set is not unnecessarily conservative.

3.1 Regressions of Interest

Let Y be an outcome variable, X a treatment variable, W_1 a vector of observed covariates, and W_2 a vector of unobserved covariates. The following assumption ensures that the OLS estimands we will consider are well defined.

Assumption A1. $\text{var}(Y, X, W_1, W_2)$ is finite. $\text{var}(X, W_1, W_2)$ and $\text{var}(Y, X, W_1)$ are positive definite.

A1 allows Y to be perfectly collinear with (X, W_1, W_2) , but not with (X, W_1) . Collinearity of Y with (X, W_1) would violate A3 below.

Consider three OLS regressions:

1. The **short regression** of

$$Y \text{ on } (1, X).$$

Let β_{short} denote the corresponding population regression coefficient on X . Let R_{short}^2 denote the corresponding R-squared.

2. The **medium regression** of

$$Y \text{ on } (1, X, W_1).$$

Let β_{med} denote the corresponding population regression coefficient on X . Let R_{med}^2 denote the corresponding R-squared. Let $\gamma_{1,\text{med}}$ denote the coefficients on W_1 .

3. The **long regression** of

$$Y \text{ on } (1, X, W_1, W_2).$$

Let β_{long} denote the corresponding population regression coefficient on X . Let R_{long}^2 denote the corresponding R-squared. Note $R_{\text{long}}^2 \in [R_{\text{med}}^2, 1]$. Let $(\gamma_{1,\text{long}}, \gamma_{2,\text{long}})$ denote the coefficients on (W_1, W_2) .

Suppose β_{long} is the parameter of interest. In particular, we can write the long regression as

$$Y = \beta_{\text{long}}X + \gamma'_{1,\text{long}}W_1 + \gamma'_{2,\text{long}}W_2 + Y^{\perp X, W} \quad (1)$$

where $Y^{\perp X, W}$ is defined to be the OLS residual plus the intercept term, and hence is uncorrelated with each component of (X, W_1, W_2) by construction. The identification problem is that W_2 is not observed. Instead,

we observe the joint distribution of (Y, X, W_1) . This allows us to point identify β_{short} and β_{med} . These coefficients usually do not equal β_{long} , however. The difference between β_{long} and either β_{short} or β_{med} is omitted variable bias. We next consider assumptions that will help us bound the magnitude of this omitted variable bias.

3.2 Comparing Selection on Observables and on Unobservables

Following the analysis of Altonji et al. (2005, pages 175–176), Oster (2019) recommends that we measure the magnitude of selection on unobservables via the parameter

$$\delta = \frac{\text{cov}(X, \gamma'_{2,\text{long}} W_2)}{\text{var}(\gamma'_{2,\text{long}} W_2)} \bigg/ \frac{\text{cov}(X, \gamma'_{1,\text{long}} W_1)}{\text{var}(\gamma'_{1,\text{long}} W_1)}.$$

To allow for cases where the denominator is zero, our formal analysis uses the following assumption.

Assumption A2. There is a known $\delta \in \mathbb{R}$ such that

$$\delta \cdot \frac{\text{cov}(X, \gamma'_{1,\text{long}} W_1)}{\text{var}(\gamma'_{1,\text{long}} W_1)} = \frac{\text{cov}(X, \gamma'_{2,\text{long}} W_2)}{\text{var}(\gamma'_{2,\text{long}} W_2)}. \quad (2)$$

The following assumption, combined with A1, ensures that we are not dividing by zero in equation (2).

Assumption A3. $\gamma_{1,\text{long}} \neq 0$ and $\gamma_{2,\text{long}} \neq 0$.

The parameter δ depends on two terms:

1. (“Selection on unobservables”) Regress X on $(1, \gamma'_{2,\text{long}} W_2)$ and get the coefficient on the index $\gamma'_{2,\text{long}} W_2$.
2. (“Selection on observables”) Regress X on $(1, \gamma'_{1,\text{long}} W_1)$ and get the coefficient on the index $\gamma'_{1,\text{long}} W_1$.

δ is the ratio of these two regression coefficients. δ is not known from the data since it depends on W_2 , which is not observed, and $(\gamma_{1,\text{long}}, \gamma_{2,\text{long}})$ which are also unknown. Instead, we will study what can be said about β_{long} under assumptions on δ . Several papers propose alternative sensitivity parameters, including Cinelli and Hazlett (2020) and Diegert et al. (2022). In this paper we instead use the same sensitivity parameter δ as Oster (2019), since our focus here is to study the implications of using different kinds of breakdown points. See Oster (2019), Cinelli and Hazlett (2020, section 6.3), and Diegert et al. (2022, appendix A.1) for further discussion of the interpretation of δ .

Baseline covariates

δ is defined to calibrate the magnitude of selection on unobservables by using the observed covariates W_1 . Typically we will have additional covariates W_0 which we want to include in the analysis, but *not* use for calibration. Denote these baseline controls by W_0 . To incorporate these covariates, replace (X, W_1, W_2) with $(X^{\perp W_0}, W_1^{\perp W_0}, W_2^{\perp W_0})$ throughout the entire analysis. See section 3.3 of Diegert et al. (2022) for a justification of this approach and further discussion. For simplicity, we omit W_0 throughout our theoretical analysis.

3.3 Identification of β_{long}

In this section we state a version of the main and most general identification result in Oster (2019). That paper also analyzes several special cases which we do not discuss for brevity. The main substantive assumption is the following.

Assumption A4 (Exogenous controls). $\text{cov}(W_1, W_2) = 0$.

A4 requires that all of the observed covariates are uncorrelated with the omitted variables. Appendix A.1 in Oster (2019) briefly discusses one approach to relaxing A4; see Appendix A of Diegert et al. (2022) for a detailed discussion of this approach. Diegert et al. (2022) also develop an alternative approach to sensitivity analysis that does not rely on A4. In this paper we maintain A4 throughout because our main focus is to study different notions of breakdown. For simplicity, from here on we assume W_2 is a scalar. All of our results can be generalized to vector W_2 using the arguments in appendix B of Diegert et al. (2022). We also maintain the following normalization.

Assumption A5 (Normalization). $\text{var}(W_2) = 1$.

To state the main result, it is also useful to consider the regression of X on $(1, W_1)$. Let π_1 denote the coefficient on W_1 . Let $X^{\perp W_1} = X - \pi_1' W_1$.

Theorem 2. Suppose the joint distribution of (Y, X, W_1) is known. Suppose A1–A5 hold. Suppose $R_{\text{long}}^2 \in (R_{\text{med}}^2, 1]$ is known. Then the identified set for β_{long} is

$$\mathcal{B}_I(\delta, R_{\text{long}}^2) = \{b \in \mathbb{R} : f(\beta_{\text{med}} - b, \delta, R_{\text{long}}^2) = 0\} \setminus \mathcal{B}_{\text{A3fail}}$$

where

$$\mathcal{B}_{\text{A3fail}} = \{b \in \mathbb{R} : \gamma_{1,\text{med}} + (\beta_{\text{med}} - b)\pi_1 = 0\}$$

and $f(B, \delta, R_{\text{long}}^2) = f_0(B) + \delta \cdot f_1(B, R_{\text{long}}^2)$ with

$$\begin{aligned} f_0(B) &= -B \text{var}(X^{\perp W_1}) (\text{var}(\gamma'_{1,\text{med}} W_1) + 2B \text{cov}(X, \gamma'_{1,\text{med}} W_1) + B^2 \text{var}(\pi'_1 W_1)) \\ f_1(B, R_{\text{long}}^2) &= (R_{\text{long}}^2 - R_{\text{med}}^2) \text{var}(Y) \text{cov}(X, \gamma'_{1,\text{med}} W_1) + B(R_{\text{long}}^2 - R_{\text{med}}^2) \text{var}(Y) \text{var}(\pi'_1 W_1) \\ &\quad + B^2 \text{var}(X^{\perp W_1}) \text{cov}(X, \gamma'_{1,\text{med}} W_1) + B^3 \text{var}(X^{\perp W_1}) \text{var}(\pi'_1 W_1). \end{aligned}$$

Theorem 2 is a version of Oster’s (2019) Proposition 2. Most importantly, Oster’s Proposition 2 did not prove sharpness. Instead, that result only showed that $\{b \in \mathbb{R} : f(\beta_{\text{med}} - b, \delta, R_{\text{long}}^2) = 0\}$ is an outer set that contains the true parameter. Our Theorem 2 shows that this set is in fact essentially sharp and therefore can correctly be interpreted as the identified set. As we discuss below, this sharpness is important to ensure that breakdown analysis is not unnecessarily conservative. The only caveat is that the set $\mathcal{B}_{\text{A3fail}}$ must be subtracted to obtain sharpness. This set will typically be empty, however. This is a minor technical point so we discuss it at the end of this subsection.

There are several other minor differences between our Theorem 2 and Oster’s (2019) Proposition 2. We do not impose the normalization that $\text{var}(W_1)$ is diagonal. Oster’s exogenous controls assumption is stated as $\text{cov}(W_1, \gamma'_{2,\text{long}} W_2) = 0$, which is equivalent to our A4 given that we are assuming for simplicity W_2 is scalar throughout the analysis, and since $\gamma_{2,\text{long}} \neq 0$ by A3.

As Oster (2019) discusses, the set

$$\{b \in \mathbb{R} : f(\beta_{\text{med}} - b, \delta, R_{\text{long}}^2) = 0\}$$

is the set of solutions to a cubic equation. Hence the identified set $\mathcal{B}_I(\delta, R_{\text{long}}^2)$ has at most three elements. Figure 1 shows an example of this set, as a function of δ .

Theorem 2 assumed $R_{\text{long}}^2 > R_{\text{med}}^2$. The following result explains why.

Proposition 2. Suppose assumption A1 holds. Suppose $R_{\text{long}}^2 = R_{\text{med}}^2$ and that δ is unknown. Then $\beta_{\text{long}} = \beta_{\text{med}}$.

Finally, the analysis simplifies to the baseline case when $\delta = 0$.

Corollary 1. Suppose A1–A5 hold. Suppose $\delta = 0$ and $\gamma_{1,\text{med}} \neq 0$. Then $\beta_{\text{long}} = \beta_{\text{med}}$.

The set $\mathcal{B}_{\text{A3fail}}$

Here we explain how to interpret $\mathcal{B}_{\text{A3fail}}$ and how to characterize when it is empty. By Lemma 1 in Appendix A,

$$\gamma_{1,\text{long}} = \gamma_{1,\text{med}} + (\beta_{\text{med}} - \beta_{\text{long}})\pi_1. \quad (3)$$

So $\mathcal{B}_{\text{A3fail}}$ is the set of β_{long} values which would imply $\gamma_{1,\text{long}} = 0$, which violates A3; recall that δ is not defined when A3 fails. This explains why such values must be removed. Next, notice that $\mathcal{B}_{\text{A3fail}}$ is nonempty if and only if there exists a $C \in \mathbb{R}$ such that $\gamma_{1,\text{med}} = C\pi_1$. That is, if and only if π_1 and $\gamma_{1,\text{med}}$ are proportional to each other. When W_1 is a vector, as is usually the case in empirical practice, this proportionality condition will typically not hold, and hence $\mathcal{B}_{\text{A3fail}}$ will be empty. When W_1 is a scalar, however, π_1 and $\gamma_{1,\text{med}}$ are scalars and hence are always proportional. In this case, the singleton $\mathcal{B}_{\text{A3fail}} = \{\beta_{\text{med}} + \gamma_{1,\text{med}}/\pi_1\}$ must be removed.

3.4 Oster’s Breakdown Analysis

Thus far we have characterized the identified set for β_{long} when δ and R_{long}^2 are known. Next we consider breakdown analysis. In this section we review the main breakdown point result in Oster (2019). In section 4, we then apply our general analysis in section 2 to better understand the behavior of Oster’s breakdown analysis.

Let $\beta_{\text{hypo}} \in \mathbb{R}$. Define the *exact value breakdown point*

$$\delta^{\text{bp},=}(\beta_{\text{hypo}}, R_{\text{long}}^2) = \inf\{|\delta| : \delta \in \mathbb{R}, \beta_{\text{hypo}} \in \mathcal{B}_I(\delta, R_{\text{long}}^2)\}.$$

This is the smallest magnitude of the sensitivity parameter δ that is compatible with $\beta_{\text{long}} = \beta_{\text{hypo}}$. That is, for all δ with $|\delta| < \delta^{\text{bp},=}(\beta_{\text{hypo}}, R_{\text{long}}^2)$, we know that β_{hypo} is not in the identified set $\mathcal{B}_I(\delta, R_{\text{long}}^2)$. The next result gives a closed form solution for this value.

Theorem 3. Suppose the joint distribution of (Y, X, W_1) is known. Suppose A1–A5 hold. Suppose $R_{\text{long}}^2 \in (R_{\text{med}}^2, 1]$ is known. Suppose $\beta_{\text{long}} = \beta_{\text{hypo}}$ for a known $\beta_{\text{hypo}} \in \mathbb{R} \setminus \mathcal{B}_{\text{A3fail}}$. Suppose $\beta_{\text{short}} \neq \beta_{\text{med}}$. Then δ is point identified and equals

$$\frac{-f_0(\beta_{\text{med}} - \beta_{\text{hypo}})}{f_1(\beta_{\text{med}} - \beta_{\text{hypo}}, R_{\text{long}}^2)}.$$

Consequently, $\delta^{\text{bp},=}(\beta_{\text{hypo}}, R_{\text{long}}^2)$ equals its absolute value.

Theorem 3 is very similar to Oster’s Proposition 3, with only minor differences. As Oster remarked, this result shows that, even though knowledge of δ does not point identify β_{long} , knowledge of β_{long} does point identify δ . Figure 1 explains why: Fixing δ is equivalent to placing a vertical line on the plot. This vertical line may intersect the displayed curve up to three times, since $\mathcal{B}_I(\delta, R_{\text{long}}^2)$ can have up to three elements. Fixing the value β_{long} to be equal to some known β_{hypo} , however, is equivalent to placing a horizontal line on the plot. This line can only intersect the displayed curve once. The absolute value of the horizontal coordinate at this intersection is precisely $\delta^{\text{bp},=}(\beta_{\text{hypo}}, R_{\text{long}}^2)$.

Setting $\beta_{\text{hypo}} = 0$, we get the explain away breakdown point

$$\delta^{\text{bp},\text{exact}}(R_{\text{long}}^2) \equiv \delta^{\text{bp},=}(0, R_{\text{long}}^2) = \frac{|f_0(\beta_{\text{med}})|}{|f_1(\beta_{\text{med}}, R_{\text{long}}^2)|}.$$

Its signed value, $-f_0(\beta_{\text{med}})/f_1(\beta_{\text{med}}, R_{\text{long}}^2)$, is often called *Oster’s delta*. This value is commonly reported as a measure of the robustness of the baseline model ($\delta = 0$) to the presence of omitted variables. However, Oster’s δ is a *deviation* from the baseline model, not a relaxation (following definition 1). Therefore, as discussed in section 2, the explain away and sign change breakdown points are generally not equal. We compare these two breakdown points in the next section.

4 Measuring the Robustness of $\text{sign}(\beta_{\text{long}})$ to Omitted Variables

In this section we analyze the robustness of conclusions about the sign of β_{long} to the presence of omitted variables. In section 4.1 we first show that the sign change breakdown point for δ can never be larger than 1. This is a concern since $\delta = 1$ is often considered the cutoff for robustness, and therefore this result implies that conclusions about the sign of β_{long} are always non-robust. In section 4.2 we propose a simple modification to Oster’s analysis that allows the sign change breakdown point to be larger than 1.

4.1 The Sign Change Breakdown Point for δ

Oster’s delta answers the following question:

Suppose we knew the exact value of R_{long}^2 . What is the smallest value of δ that is required for the data to be consistent with $\beta_{\text{long}} = 0$?

We call this the *explain away* breakdown point. Researchers may also be interested in answering the following question:

Suppose we knew the exact value of R_{long}^2 . What is the smallest value of δ such that β_{long} has a different sign from β_{med} ?

We call this the *sign change* breakdown point. Because Oster (2019) considers deviations from the baseline model rather than relaxations, Oster’s delta does not answer this question. In fact, as we next show, the sign change breakdown point is typically much smaller than the explain away breakdown point.

To formalize this claim, let

$$\delta^{\text{bp},\text{sign}}(R_{\text{long}}^2) = \begin{cases} \delta^{\text{bp},>}(R_{\text{long}}^2) & \text{if } \beta_{\text{med}} > 0 \\ \delta^{\text{bp},<}(R_{\text{long}}^2) & \text{if } \beta_{\text{med}} < 0 \end{cases}$$

where

$$\begin{aligned}\delta^{\text{bp},<}(R_{\text{long}}^2) &= \sup\{|\delta| : \delta \in \mathbb{R}, b < 0 \text{ for all } b \in \mathcal{B}_I(\delta, R_{\text{long}}^2)\} \\ \delta^{\text{bp},>}(R_{\text{long}}^2) &= \sup\{|\delta| : \delta \in \mathbb{R}, b > 0 \text{ for all } b \in \mathcal{B}_I(\delta, R_{\text{long}}^2)\}.\end{aligned}$$

We now state our main theoretical result.

Theorem 4. Suppose A1–A5 hold. Suppose $\beta_{\text{short}} \neq \beta_{\text{med}}$. Then for any $R_{\text{long}}^2 \in (R_{\text{med}}^2, 1]$,

$$\delta^{\text{bp,sign}}(R_{\text{long}}^2) \leq 1.$$

Here we assume $\beta_{\text{short}} \neq \beta_{\text{med}}$ for simplicity only. Theorem 4 shows that, in Oster’s sensitivity analysis, the sign change breakdown point *can never be larger than one*. In practice, researchers typically consider the value $\delta = 1$ to be the cutoff for determining the robustness of a result. That is, they interpret values of δ greater than 1 as robust and values less than 1 as sensitive. Based on this criterion, Theorem 4 therefore shows that the *sign* of the parameter of interest is *never robust* (assuming we require δ strictly larger than 1 to conclude that the result is robust). In contrast, the explain away breakdown point $\delta^{\text{bp,exact}}(R_{\text{long}}^2)$ is often substantially larger than one. We illustrate this difference empirically in section 6.

Figure 1 illustrates the intuition for this result: The graph $\{(\delta, b) \in \mathbb{R}^2 : f(\beta_{\text{med}} - b, \delta, R_{\text{long}}^2) = 0\}$ has an asymptote at $\delta = 1$. Consequently, values of δ close to one—but not equal to one—yield identified sets that contain arbitrarily large or arbitrarily small values. Hence for δ values near one we cannot point identify the sign of β_{long} . Since we are only considering a deviation, however, much larger values of δ may be needed before the graph crosses the horizontal axis $b = 0$ exactly. This explains why the explain away breakdown point can often be much larger than the sign change breakdown point for this model.

4.2 Incorporating Magnitude Restrictions on β_{long}

We have just shown that the sign change breakdown point for δ can never be larger than one. This negative result can be overturned by adding additional identifying assumptions. We consider the following particularly simple assumption.

Assumption A6. $|\beta_{\text{long}} - \beta_{\text{med}}| \leq M$ for some known $M \geq 0$.

A6 says that the magnitude of the omitted variable bias is no larger than M . We discuss the choice of M below. This assumption can be viewed as formalizing Oster’s (2019) statement that researchers may be “willing to assume that the bias is fairly small” (page 194).

Corollary 2. Suppose the joint distribution of (Y, X, W_1) is known. Suppose A1–A6 hold. Suppose $R_{\text{long}}^2 \in (R_{\text{med}}^2, 1]$ is known. Suppose $\delta \in \mathbb{R}$ is known. Then the identified set for β_{long} is

$$\mathcal{B}_I(\delta, R_{\text{long}}^2, M) = \mathcal{B}_I(\delta, R_{\text{long}}^2) \cap [\beta_{\text{med}} - M, \beta_{\text{med}} + M].$$

Let $\delta^{\text{bp,sign}}(R_{\text{long}}^2, M)$ denote the corresponding sign change breakdown point. For $0 \leq M < \infty$, it is now possible for $\delta^{\text{bp,sign}}(R_{\text{long}}^2, M) > 1$. We illustrate this empirically in section 6.

Thus far we have followed Oster (2019) by considering deviations from the baseline model—our identification results (Theorem 2 and Corollary 2) have assumed the exact value of δ is known. Consequently, when

we also make the magnitude restriction A6 the model can be falsified. To avoid this, we next convert the analysis to a relaxation of the baseline model. Specifically, we make the following assumption.

Assumption A7. There is a known constant $\bar{\delta} \geq 0$ such that $|\delta| \leq \bar{\delta}$.

This leads to the following identification result.

Corollary 3. Suppose the joint distribution of (Y, X, W_1) is known. Suppose A1–A5 and A7 hold. Suppose $R_{\text{long}}^2 \in (R_{\text{med}}^2, 1]$ is known. Then the identified set for β_{long} is

$$\mathcal{B}_I(\bar{\delta}, R_{\text{long}}^2) = \bigcup_{|\delta| \leq \bar{\delta}} \mathcal{B}_I(\delta, R_{\text{long}}^2).$$

Suppose A6 also holds. Then the identified set for β_{long} is

$$\mathcal{B}_I(\bar{\delta}, R_{\text{long}}^2, M) = \bigcup_{|\delta| \leq \bar{\delta}} \mathcal{B}_I(\delta, R_{\text{long}}^2, M).$$

This corollary simply notes that we take the union of the identified sets from Theorem 2 over all allowed values of the sensitivity parameter δ with magnitude less than $\bar{\delta}$ to obtain the identified set. Measuring sensitivity with $\bar{\delta}$ rather than δ ensures that the model which also imposes A6 is no longer falsifiable, regardless of the value of $M \geq 0$.

Choosing M

M is the largest difference between the known β_{med} and the unknown β_{long} . The choice of this parameter should therefore reflect the researcher’s beliefs about this difference. If $M = 0$ then $\beta_{\text{long}} = \beta_{\text{med}}$ is point identified regardless of the magnitude of selection on unobservables. Therefore M should not be chosen too small. If M is large, the magnitude restriction will not bind and therefore will not affect the sensitivity analysis. We recommend that researchers consider a range of magnitude restrictions, as we illustrate empirically in section 6. One simple approach is to choose M to be multiples of β_{med} . That is, let $M = p \cdot \beta_{\text{med}}$ for $p \geq 0$. For this choice, A6 is equivalent to

$$\beta_{\text{long}} \in [(1 - p)\beta_{\text{med}}, (1 + p)\beta_{\text{med}}].$$

Every element of this set has the same sign if and only if $p \leq 1$. For this reason, since we are interested in robustness to sign changes, p should be chosen strictly larger than 1. For example, in section 6 we consider $p = 2, 3$, and 10.

5 The Sensitivity of Bias Adjusted Estimands

Thus far we have focused on breakdown analysis. Researchers also commonly compute point estimands that adjust for omitted variable bias. See our survey in section 7 for examples from the empirical economics literature. When the sensitivity parameter is a deviation, as in Oster’s (2019) analysis, we show that this kind of bias adjustment can be extremely sensitive to perturbations in the choice of sensitivity parameter. We also discuss several additional concerns about common practice when using Oster’s (2019) results to do bias adjustment. We then make a recommendation which addresses these concerns. We conclude by

discussing a common mistake some researchers make when computing explain away breakdown points, along with a remark about the “bounding set” defined in Oster (2019).

5.1 Adding Assumptions to Point Identify β_{long}

In this subsection we describe the most common approach by which empirical researchers use Oster’s (2019) results to compute bias adjusted estimands. Theorem 2 generally does not point identify β_{long} , which means there is not a unique way to compute a bias adjusted estimand. However, as Oster shows, the identified set in Theorem 2 simplifies to a singleton under two additional assumptions. The first assumption is $\delta = 1$. The second assumption is the following, which is equivalent to Oster’s assumption 2 (page 192).¹

Assumption A8. There is a constant $C \in \mathbb{R}$ such that $\gamma_{1,\text{long}} = C \cdot \pi_1$.

Recall that π_1 denotes the coefficient on W_1 in the regression of X on $(1, W_1)$. A8 holds immediately when W_1 is a scalar, but otherwise is a restriction. In fact, the following result implies that A8 is refutable and confirmable (as defined in Breusch 1986) under exogenous controls. Hence the data alone tells us whether it holds.

Proposition 3. Suppose A1 and A4 hold. Then A8 holds if and only if there is a constant $C_{\text{med}} \in \mathbb{R}$ such that $\gamma_{1,\text{med}} = C_{\text{med}} \cdot \pi_1$.

For vector W_1 , Oster (2019) said A8 “is very unlikely to hold except in pathological cases” (page 192). Proposition A3 shows that researchers do not need to guess about this assumption—it can be directly tested in the data. We develop this test in section 5.3.

Under these two additional assumptions, we can now show that β_{long} is point identified.

Proposition 4. Suppose the joint distribution of (Y, X, W_1) is known. Suppose A1, A3, A4, A5, and A8 hold. Suppose A2 holds with $\delta = 1$. Suppose $R_{\text{long}}^2 \in (R_{\text{med}}^2, 1]$ is known. Suppose $\beta_{\text{short}} \neq \beta_{\text{med}}$. Then β_{long} is point identified and

$$\beta_{\text{long}} = \beta_{\text{med}} + (\beta_{\text{med}} - \beta_{\text{short}}) \frac{R_{\text{long}}^2 - R_{\text{med}}^2}{R_{\text{med}}^2 - R_{\text{short}}^2}. \quad (4)$$

This result is a version of Oster’s Proposition 1. Oster’s (2019) proof is slightly incomplete, however, since it only shows that the value in equation (4) is consistent with the data. It does not show that this is the only value that is consistent with the data. Our proof of Proposition 4 shows this last step. In this proof we show how the result follows directly from our Theorem 2: The cubic function $f(\beta_{\text{med}} - b, \delta, R_{\text{long}}^2)$ is quadratic with two solutions when $\delta = 1$ and one of the two solutions is eliminated by subtracting $\mathcal{B}_{\text{A3fail}}$. The remaining solution is equation (4). Finally, in this result we assume $\beta_{\text{short}} \neq \beta_{\text{med}}$ to ensure that $R_{\text{short}}^2 \neq R_{\text{med}}^2$ and hence that the right hand side of (4) is well defined.

5.2 The Sensitivity of the Equation (4) Adjustment

Many papers use equation (4) to produce bias adjusted estimates. There are three issues with this approach. First, it is only valid when A8 holds. Fortunately, as we showed in Proposition 3, we can use the data to

¹It is unclear to us whether Oster’s (2019) assumption 2 is meant to be an assumption relating $\gamma_{1,\text{med}}$ to π_1 or relating $\gamma_{1,\text{long}}$ to π_1 (using our notation), since that paper uses the same notation Ψ to denote both $\gamma_{1,\text{long}}$ and $\gamma_{1,\text{med}}$ in different places (for example, compare her equation (1) with the definition of ψ_i in the statement of her assumption 2). In general, $\gamma_{1,\text{med}} \neq \gamma_{1,\text{long}}$. Regardless, our Proposition 3 shows that the two forms of this assumption are equivalent under the exogenous controls assumption A4. Also note that Oster (2019) does not discuss testability of this assumption.

directly test this assumption. We develop such a test in section 5.3 and recommend that anyone who wants to use equation (4) also use this test to make sure that A8 holds. Second, as discussed in Diegert et al. (2022), equation (4) relies on the exogenous controls assumption A4.

Third, equation (4) is only valid when $\delta = 1$. Theorem 2 characterizes the feasible values of β_{long} when $\delta \neq 1$. In our earlier analysis, we saw that when δ is very close to one, but not equal to one, the identified set for β_{long} will contain values with arbitrarily large magnitudes; note that this is true even when A8 holds. Consequently, bias adjustments based on equation (4) are extremely sensitive to perturbations of δ . Therefore, for researchers to view bias adjusted estimands with $\delta = 1$ as useful robustness check, they must believe that δ *exactly* equals 1, and is not merely close to 1. We illustrate this sensitivity in table 3 in our empirical application.

Importantly, these large values are equally as valid as the smallest element of the identified set. That is, under the assumptions of Theorem 2, there is no justification for ruling out the possibility that the true bias is one of these large values. The only way to rule out these values is to make additional assumptions. One such assumption is the magnitude restriction A6 that we discuss in section 4.2. The identified set $\mathcal{B}_I(\bar{\delta}, R_{\text{long}}^2, M)$ that we describe in that section formally incorporates this magnitude restriction. However, even if we impose this magnitude restriction, the cumulative identified set $\mathcal{B}_I(\bar{\delta}, R_{\text{long}}^2, M)$ can still easily contain both positive and negative elements even with $\bar{\delta} = 1$. This can be seen by superimposing the bound implied by A6 onto figure 2 or 3. Consequently, even with a magnitude restriction, bias adjustments based on equation (4) will be sensitive to changes in δ .

5.3 Testing A8

In this section we sketch a simple hypothesis test for A8 using finite sample data. Let $\{(Y_i, X_i, W_i)\}_{i=1}^n$ be an iid sample from (Y, X, W) . Let $\hat{\pi}_1$ be the OLS estimator for the coefficient on W_1 in the regression of X on $(1, W_1)$. Let $\hat{\gamma}_{1,\text{med}}$ be the OLS estimator for the coefficient on W_1 in the regression of Y on $(1, X, W_1)$. Let $d_1 = \dim(W_1)$ denote the dimension of the observed covariates W_1 . Let $\hat{V}_0 \in \mathbb{R}^{2d_1 \times 2d_1}$ be the estimated asymptotic covariance matrix of $(\hat{\gamma}_{1,\text{med}}, \hat{\pi}_1)$.

By Proposition 3, testing A8 is equivalent to testing the existence of a scalar $C \in \mathbb{R}$ that solves $\gamma_{1,\text{med}} = C\pi_1$. Let

$$\tilde{C} = \underset{C \in \mathbb{R}}{\operatorname{argmin}} (\hat{\gamma}_{1,\text{med}} - C\hat{\pi}_1)'(\hat{\gamma}_{1,\text{med}} - C\hat{\pi}_1) \quad (5)$$

denote an initial estimator of this special C value. Our test is a kind of over-identification test since the idea is that the objective function in equation (5) should converge to zero as $n \rightarrow \infty$ when A8 holds. To construct the test statistic, let $\hat{V}_1(C) \in \mathbb{R}^{d_1 \times d_1}$ be the estimated asymptotic covariance matrix of $\hat{\gamma}_{1,\text{med}} - C\hat{\pi}_1$, as derived from \hat{V}_0 . Define

$$J = \min_{C \in \mathbb{R}} n(\hat{\gamma}_{1,\text{med}} - C\hat{\pi}_1)' \hat{V}_1(\tilde{C})^{-1} (\hat{\gamma}_{1,\text{med}} - C\hat{\pi}_1).$$

This is a J -test statistic for over-identifying restrictions. Using arguments from Newey and McFadden (1994), for example, we have $J \xrightarrow{d} \chi_{d_1-1}^2$ as $n \rightarrow \infty$ under standard regularity conditions. We thus have the following test of H_0 : A8 holds versus H_1 : A8 does not hold: Reject if $J > c(1 - \alpha)$, where $c(1 - \alpha)$ is the $1 - \alpha$ quantile of the $\chi_{d_1-1}^2$ distribution. Note that this test requires $d_1 > 1$ because otherwise, as noted earlier, A8 always holds when $d_1 = 1$. It also requires assuming that $\pi_1 \neq 0$, which always holds if $\beta_{\text{short}} \neq \beta_{\text{med}}$.

5.4 Recommendations

We have shown that bias corrections based on equation (4) are either not valid because A4 or A8 fails, or are highly dependent on δ being exactly 1, and therefore are extremely sensitive to perturbations of δ . A simple alternative is to present a plot of the identified set $\mathcal{B}_I(\bar{\delta}, R_{\text{long}}^2)$ for β_{long} as a function of $\bar{\delta} \geq 0$. We described this set in section 4.2. This set describes the values of β_{long} consistent with any amount of selection on unobservables δ with magnitude less than $\bar{\delta}$. Hence, unlike a single point estimand for $\delta = 1$, it shows how the feasible values of β_{long} vary for a range of δ values. Moreover, since it is based on Theorem 2, this set does not require assumption A8 to be valid. See figure 2 for an example, and figure 3 for the simpler convex hull of this set.

5.5 A Common Mistake when Computing Explain Away Breakdown Points

As pointed out in Diegert et al. (2022), several papers compute the explain away breakdown point $\delta^{\text{bp,exact}}(R_{\text{long}}^2)$ incorrectly. We conclude this section by describing the source of this mistake. It can be avoided by using software based on the correct results, such as Oster’s package `psacalc` or our package `regsensitivity` in Stata.

Proposition 4 shows how to point identify β_{long} if $\delta = 1$, R_{long}^2 is known, and A8 is assumed. The identification result in Theorem 2 is strictly more general since it allows for any value of δ and does not use A8. Because this general theorem holds for any δ , it should be used for breakdown analysis, as we did earlier. However, on page 193 Oster (2019) states the following equation:

$$\beta_{\text{long}} \approx \beta_{\text{med}} + \delta(\beta_{\text{med}} - \beta_{\text{short}}) \frac{R_{\text{long}}^2 - R_{\text{med}}^2}{R_{\text{med}}^2 - R_{\text{short}}^2} \quad (6)$$

where δ is no longer necessarily equal to one. We are not aware of any formal results which justify this equation—Oster (2019) does not provide any and does not explain the nature of the approximation \approx in this equation. To the contrary, as we explained earlier, Theorem 2 implies that β_{long} can be arbitrarily far from its value at $\delta = 1$ even for δ values arbitrarily close to 1. Moreover, even at $\delta = 1$, β_{long} is not necessarily point identified when A8 is not assumed. Consequently, equation (6) should not be used to perform sensitivity analysis (except when δ is exactly equal to 1 and under the other conditions of Proposition 4).

Several researchers, however, have apparently used equation (6) to compute breakdown points. This includes Satyanath et al. (2017, *JPE*) and Bazzi, Fiszbein, and Gebresilasse (2020, *ECMA*). Specifically, they fix R_{long}^2 and solve

$$\beta_{\text{med}} + \delta(\beta_{\text{med}} - \beta_{\text{short}}) \frac{R_{\text{long}}^2 - R_{\text{med}}^2}{R_{\text{med}}^2 - R_{\text{short}}^2} = 0$$

for δ . This solution does not equal the explain away breakdown point $\delta^{\text{bp,exact}}(R_{\text{long}}^2)$. Note that this is true even if A8 holds—Proposition 4 is only about the case $\delta = 1$ and does not describe the bias when $\delta \neq 1$. The correct value of $\delta^{\text{bp,exact}}(R_{\text{long}}^2)$ is given by Theorem 3.

Equation (6) also explains why some papers do not report the magnitude of their computed $\delta^{\text{bp,exact}}$ when it is negative. For example, Satyanath et al. (2017) report “[< 0]” in their Table A.27 results rather than providing the exact magnitude of negative estimates. In the table footnote, they say

“The entry [<0] indicates that the respective Altonji et al. ratio is negative... suggesting a downward bias for our OLS estimates due to unobservables”. (online appendix page 44)

If equation (6) were correct for all δ , then the sign of the omitted variable bias would be determined by the sign of δ and $\beta_{\text{med}} - \beta_{\text{short}}$. So it is incorrect to conclude that the true bias $\beta_{\text{med}} - \beta_{\text{long}}$ is negative whenever this breakdown point is negative. Regardless, this line of reasoning is all based on equation (6) which, as we explained above, generally does not hold except when δ is exactly equal to 1 and under the other conditions of Proposition 4.

5.6 Remark on Oster’s “Bounding Set”

Based on the discussion on pages 194–196, Oster (2019) defines

$$\beta^*(\delta, R_{\text{long}}^2) = \underset{b \in \mathcal{B}_I(\delta, R_{\text{long}}^2)}{\operatorname{argmin}} |\beta_{\text{med}} - b|.$$

Using this, she defines the set $[\beta_{\text{med}}, \beta^*(\delta, R_{\text{long}}^2)]$ (or the reverse if β_{med} is the larger value). She calls this a “bounding set” on page 196, but elsewhere calls it the “identified set” (for example, see her Table 3). The discussion on page 196 suggests that this interval should be interpreted as the identified set for β_{long} under the assumption that $0 \leq \delta \leq 1$. As we have shown, this is not the correct interpretation. This follows because Oster’s model is a deviation, not a relaxation, and hence the identified set $\mathcal{B}_I(\delta, R_{\text{long}}^2)$ for a known δ is not monotonic in δ . Note that this is generally true even if we impose the magnitude constraint A6. The correct identified set under the assumption that $0 \leq \delta \leq 1$ is given by $\mathcal{B}_I(\bar{\delta}, R_{\text{long}}^2)$ from Corollary 3 with $\bar{\delta} = 1$. Figure 2 plots an example of this set; figure 3 plots its convex hull. We recommend that researchers present this set, rather than the “bounding set” described above.

Finally, note that some authors (for example, Mian and Sufi 2014, *ECMA*, online appendix page 2) define $\beta^*(\delta, R_{\text{long}}^2)$ as the value from equation (4), from Oster’s Proposition 1. This is equivalent to the definition above when $\delta = 1$ and A8 holds, but otherwise they are different. Regardless, this approach also leads to an interval that does not have the correct interpretation as an identified set.

6 Empirical Application: Social Capital and the Rise of the Nazi Party

Social capital is typically viewed as a good which leads to good outcomes for countries. For example, de Tocqueville (1835) famously argued that social capital was a key reason why American democracy thrived. More recently, Putnam (2000) argued that the *decline* in social capital threatened American democracy. Satyanath et al. (2017, *JPE*) argue the converse: “Our paper is the first to show—on the basis of detailed cross-sectional data—that social capital can undermine and help to destroy a democratic system” (page 482). They specifically look at the rise of the Nazi Party (the National Socialist German Worker’s Party [NSDAP]) in 1933. Overall, they argue that “vibrant networks of clubs and associations facilitated the rise of the Nazi Party” (page 520). In contrast, earlier literature blamed the rise of the Nazis

“on a ‘civic non-age’ of low social capital (Stern 1972) and Nazi entry on rootless, isolated individuals in a modernized society (Shirer 1960; Arendt 1973)” (page 520)

They also qualify their results by arguing that

“where government was more stable, the link between association density and Nazi Party entry was weaker” and hence “the effects of social capital depend crucially on the political and

institutional context”

Their paper has a variety of analyses. We focus on their main results, which use a selection on observables identification strategy paired with linear models.

6.1 Data

We first describe the data. See section 3 of Satyanath et al. (2017) for further details. The units of analysis are towns/cities in Germany. The sample size is $n = 229$. The outcome variable Y is a measure of how many people in that city joined the Nazi party. The treatment variable X is a measure of social capital. Thus we are interested in whether there was a causal effect of having a large amount of social capital within a city on the rate at which new members joined the Nazi party within that city. Satyanath et al. (2017) use three different measures of social capital. All three are based on “association density,” the number of social clubs associations per 1,000 city inhabitants. There are 22,127 different associations in the sample. The authors classify these into three types:

1. *Military* associations: Stahlhelm (“steel helmet”), veterans associations.
2. *Civic* associations (have “a clearly nonmilitaristic/nationalist outlook”): Animal breeding, music, chess, hiking, women’s, citizens’, and homeland (heimat) clubs, and some others (predominantly civic clubs, many with an artistic or creative pursuit such as gardening, theatre, or photography).
3. *Other*: Sports, choirs, gymnastics, shooting, students/fraternities, lodges, youth, oldfellows, hunting, gentlemen, corps.

X is then defined to be the density of all associations, of civic associations only, or of military associations only.

Next consider the control variables. Recall from section 3.2 that we distinguish between two kinds of observed control variables: Those which we use for comparison in the sensitivity analysis (W_1) and those which we do not (W_0). The baseline controls W_0 are population, share Catholic, and share blue-collar, all defined for 1925. The comparison covariates W_1 are

- Socioeconomic controls: share of Jews (1925), share unemployed (1933), welfare recipients per 1000 (1933), war participants per 1000 (1933), social insurance pensioners per 1000 (1933), log average income tax payment (1933), log average property tax payment (1933).
- Political controls: Hitler speeches per 1000 (1932), average DNVP votes 1920–1928, average DVP votes 1920–1928, average SPD votes 1920–1928, average KPD votes 1920–1928.

In their appendix G, Satyanath et al. (2017) also consider using the variables in W_0 for calibration; we omit this analysis for brevity. In their baseline analysis, they also include state fixed effects as covariates. They do not include these in their sensitivity analysis, however.

6.2 Baseline Model Results

Table 1 shows the baseline regression results. Column (1) shows the estimated coefficient $\hat{\beta}_{\text{short}}$ on X from OLS of Y on $(1, X, W_0)$, along with its standard error. Panels A–C show this coefficient for the three different definitions of the treatment variable X . First consider Panel A. This is their most broad definition

Table 1: Baseline results for Satyanath et al. (2017).

Dependent Variable: Nazi Party Entry, 1925–January 1933			
	(1)	(2)	(3)
A. Treatment = All Associations			
	0.160*** (0.053)	0.170*** (0.050)	0.087* (0.047)
B. Treatment = Civic Associations			
	0.430*** (0.130)	0.439*** (0.120)	0.284*** (0.100)
C. Treatment = Military Associations			
	0.829*** (0.264)	0.855*** (0.263)	0.613* (0.330)
Controls:			
Baseline	X	X	X
Socioeconomic		X	X
Political		X	X
State Fixed Effects			X

Note: These estimates replicate part of table 3 in Satyanath et al. (2017): Column (1) replicates Panel A, columns (4)–(6). Column (3) replicates Panel B, columns (4)–(6). Panel C, column (2) replicates Panel B, column (3). Panels A and B, column (2) were not presented in table 3 of Satyanath et al. (2017). ***, **, *: statistically significant at 1%, 5%, 10%.

of treatment, accounting for all kinds of associations. We see that the estimated coefficient is positive and statistically significantly different from zero at the 1% level. Thus, adjusting for the baseline controls W_0 , cities with more social capital, as measured by density of all associations, had higher rates of new membership in the Nazi party.

Does this result reflect a broad causal effect of social capital on entry into the Nazi party? One counter-argument is that it may merely reflect the effect of a specific kind of social capital—membership in military specific associations. Indeed, in Panel C we see that, when restricting attention to military associations, the estimated coefficient is again positive and statistically significantly different from zero at the 1% level. However, suppose we restrict attention to civic associations, which are defined to be non-militaristic and non-nationalistic. Panel B shows these estimates. There we continue to see a positive point estimate that is statistically significantly different from zero at the 1% level. This finding suggests that the effect is not driven solely by military specific social capital.

Are these results driven by omitted variable bias? The authors consider a variety of alternative methods to answer this question. We focus on just two: (i) Using additional controls and (ii) Using formal econometric methods for sensitivity analysis. Column (2) of Table 1 includes the additional controls W_1 in the regressions. Specifically, it shows the estimated coefficient $\hat{\beta}_{\text{med}}$ on X from OLS of Y on $(1, X, W_0, W_1)$, along with its standard error. Panels A–C again show this coefficient for the three different versions of treatment. We again see that all of the point estimates are positive and statistically significantly different from zero at the 1% level. Column (3) adds state fixed effects. The point estimates are smaller but still positive and statistically significantly different from zero at the 10% level. On page 500 the authors explain that this drop in precision

Table 2: Robustness of Satyanath et al. (2017) Results to Omitted Variables.

	R_{long}^2	Oster (2019)		This Paper				DMP (2022)
		$\hat{\delta}^{\text{bp,exact}}(R_{\text{long}}^2)$		$\hat{\delta}^{\text{bp,sign}}(R_{\text{long}}^2, M)$ with M equal to				\hat{r}_X^{bp}
		Reported	Correct	$+\infty$	$10\hat{\beta}_{\text{med}}$	$3\hat{\beta}_{\text{med}}$	$2\hat{\beta}_{\text{med}}$	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
All	1	< 0	-32.0	0.586	0.586	0.586	0.676	0.538
	$1.3\hat{R}_{\text{med}}^2$	-	-331.5	0.953	0.954	1.476	2.829	
Civic		< 0	0.98	0.736	0.736	0.736	0.745	0.648
		-	9.78	0.994	1.036	1.806	2.986	
Military		< 0	1.42	0.899	0.899	0.909	0.977	0.761
		-	10.65	0.993	1.035	1.821	3.058	

Note: For all specifications, the W_0 includes the baseline controls while the calibration covariates W_1 includes the socioeconomic and political controls. See the text for a complete list of these variables. For comparison with column (1), in column (2) we also report the sign of the estimate obtained from Theorem 3. Column (7) reports the estimated breakdown point proposed in the alternative analysis of Diegert et al. (2022).

arises from the fact that some of the historical policies which drove variation in association density often varied at the state level.

6.3 Assessing Selection on Observables with Sensitivity Analysis

Interpreting the results of Table 1 as causal effects of social capital requires a selection on observables assumption: Association density is randomly assigned conditional on the included covariates. In their section 6, Satyanath et al. (2017) provide a wide variety of robustness checks to support this causal interpretation. This includes an application of Oster (2019) in subsection 4.D, with details provided in their appendix G. In this section we present the findings from this sensitivity analysis, as informed by the analysis in our paper.

Breakdown Analysis

Table 2 shows the main breakdown results. This table replicates and extends the bottom half of Panel B in Table A.27 of appendix G in Satyanath et al. (2017). That part of Panel B only uses W_1 for calibration. The top half uses both W_0 and W_1 for calibration; we omit this analysis for brevity. The three rows of Table 2 correspond to the three different versions of treatment. All of the cell entries are different kinds of breakdown points.

First consider columns (1) and (2). These show estimates of Oster’s explain away breakdown point. We present estimates for two different choices of R_{long}^2 , either 1 or Oster’s rule of thumb $1.3\hat{R}_{\text{med}}^2$. Column (1) shows the values reported by Satyanath et al. (2017). As described in section 5.5, these values are incorrect. Column (2) reports the correct values. If we judged robustness by comparing the magnitude of these values of the explain away breakdown point to 1, then all of the results would be deemed robust except one, civic associations with $R_{\text{long}}^2 = 1$, which gives an explain away breakdown point of 0.98.

However, as we have discussed in this paper, the explain away breakdown point can be very different from the sign change breakdown point. Columns (3)–(6) show estimated sign change breakdown points, for four increasingly restrictive choices of the constraint on the magnitude M of OVB: $+\infty$ (no constraint), $10\hat{\beta}_{\text{med}}$,

$3\hat{\beta}_{\text{med}}$, and $2\hat{\beta}_{\text{med}}$. First consider column (3), which does not impose any magnitude restriction. As required by Theorem 4, we see that all of the point estimates are bounded above by 1. Nonetheless, as explained in section 3.3 of Diegert et al. (2022), 1 should not necessarily be viewed as a universal cutoff for robustness. Indeed, for the rule of thumb choice of R_{long}^2 , all three sign change breakdown points are very close to 1, their largest possible value. This suggests that findings of positive coefficients in Table 1 are substantially robust to the presence of omitted variables. This finding is somewhat muted, however, if we consider different choices of R_{long}^2 . Specifically, for the largest possible choice, $R_{\text{long}}^2 = 1$, the sign change breakdown points shrink a fair amount. Still, the sign change breakdown point when examining civic associations is 0.736, which is reasonably far from zero.

As explained in section 4.2, we can obtain sign change breakdown points larger than 1 if we impose additional restrictions. Columns (4)–(6) show the estimated sign change breakdown points when we restrict the magnitude of OVB to be below M . For the relatively unrestrictive value of $M = 10\hat{\beta}_{\text{med}}$, this additional restriction does not change the sign change breakdown points much. However, it is enough such that two of the breakdown points are now slightly above 1. For the more restrictive values $M = 3\hat{\beta}_{\text{med}}$ and $M = 2\hat{\beta}_{\text{med}}$, the breakdown points are now all substantially about 1 for the rule of thumb choice $R_{\text{long}}^2 = 1.3\hat{R}_{\text{med}}^2$. For $R_{\text{long}}^2 = 1$, however, the magnitude restrictions do not affect the values of the sign change breakdown points very much.

The last column of Table 2 shows the estimated breakdown point \hat{r}_X^{bp} developed in Diegert et al. (2022). This point does not require the exogenous controls assumption A4. It is based on an interval relaxation, and hence it is a sign change breakdown point. It also does not require a choice of R_{long}^2 . Like the breakdown points based on Oster (2019), larger values represent larger magnitudes of selection on unobservables relative to selection on observables. Like the sign change breakdown point in column (3), it is also bounded above by 1. In column (7) we see that this alternative sensitivity analysis produces a similar pattern of robustness as the sign change breakdown points in Oster’s model: The results for military associations are most robust, followed by civic associations, and then followed by the overall measure (recall that the overall measure includes associations which are neither military nor civic).

Selection Bias Adjustments

Satyanath et al. (2017) did not present selection bias adjusted estimands; as we document in section 7, it is rare for empirical researchers to present both breakdown analysis and bias adjustments. To illustrate the results of section 5, here we also use bias adjustments to analyze the sensitivity of the results in Satyanath et al. (2017).

Table 3 summarizes our the bias adjustment results. For brevity, here we focus on just the “all associations” treatment variable and the specification in column (2) of table 1, and we only consider $R_{\text{long}}^2 = 1$. Our findings are similar for the other treatment variables, specifications, and R_{long}^2 values. For simplicity we also do not present confidence sets for these point estimates. Panel A represents current empirical practice. The first row shows the baseline estimate. The second row shows the selection bias adjusted estimate obtained from Oster’s Proposition 1 (our Proposition 4), which assumes $\delta = 1$ and requires A8. Based on this estimate, the researcher would conclude that selection bias only works in their favor—it has the same sign as their baseline estimate and indeed is even larger.

However, now consider the results in Panel B. If we maintain $\delta = 1$ but drop A8, then there are now two values of the coefficient that are consistent with the assumptions and the data, and one of them is negative. Hence even this minor change from current practice overturns the robustness finding based on

Table 3: The Sensitivity of Bias Adjustments.

Method	Assumptions	β_{long} Values
	(1)	(2)
A. Current Practice		
Baseline Estimate	$\delta = 0$	0.170
Oster’s Prop 1 (our Prop 4)	$\delta = 1$, A8	0.532
B. Sensitivity to δ Assumptions		
Oster’s Prop 2 (our Thm 2), $\widehat{\mathcal{B}}_I(\delta, R_{\text{long}}^2)$	$\delta = 1$	$\{-0.0855, 1.8947\}$
Oster’s Prop 2, $\widehat{\mathcal{B}}_I(\delta, R_{\text{long}}^2)$	$\delta = 0.99$	$\{-18.66, -0.0868, 1.736\}$
Oster’s Prop 2, $\widehat{\mathcal{B}}_I(\delta, R_{\text{long}}^2)$	$\delta = 1.01$	$\{-0.0843, 2.133, 15.64\}$
Corollary 3, $\widehat{\mathcal{B}}_I(\bar{\delta}, R_{\text{long}}^2)$	$ \delta \leq \bar{\delta} = 1$	$(-\infty, -0.0855] \cup [0.0432, 1.8947]$
Corollary 3, $\widehat{\mathcal{B}}_I(\bar{\delta}, R_{\text{long}}^2)$	$ \delta \leq \bar{\delta} = 1.05$	$(-\infty, -0.0798] \cup [0.0415, \infty)$

Note: All estimates are for the “all associations” specification from Satyanath et al. (2017). All sensitivity methods use $R_{\text{long}}^2 = 1$ and require exogenous controls, A4.

Oster’s Proposition 1 alone. Moreover, since the value 0.532 is no longer an element of the identified set, the test we develop in section 5.3 would likely reject A8. Now, a researcher may claim that the value -0.0855 is close to zero, and perhaps is not statistically significantly different from zero, in which case one may argue that Oster’s Proposition 2 with $\delta = 1$ still weakly supports a positive effect. So next we consider perturbing δ a slight amount: If we set $\delta = 0.99$ then the identified set now includes a large negative value, -18.66 . Likewise, if we set $\delta = 1.01$ then the identified set includes a large positive value, 15.64 . Thus we see that the magnitude of the bias adjustment is very sensitive to the specific choice of δ . Finally, the last two rows of Panel B show the identified sets when we replace the assumption that δ is known exactly with the assumption that we only have a bound $\bar{\delta}$ on its magnitude (A7). Here we see that for $\bar{\delta} = 1$, almost every negative value is consistent with the data and the assumptions, and many positive values are as well, including some bigger and smaller than the baseline estimate. If we increase this bound slightly to $\bar{\delta} = 1.05$ then arbitrarily large values of the coefficient are possible too. In Table 3 we have considered just a few values of δ or $\bar{\delta}$. Figure 2 shows the identified sets from Oster’s proposition 2 for a large range of δ ’s. Figure 3 shows the convex hull of the identified sets from our Corollary 3 for a large range of $\bar{\delta}$ ’s. Note that the magnitude constraint A6 can be imposed by simply superimposing horizontal lines at the largest and smallest feasible values on this plot.

To better understand the estimates in Tables 2 and 3, we next plot estimated identified sets $\widehat{\mathcal{B}}_I(\delta, R_{\text{long}}^2)$ as a function of δ for the 3 choices of treatment variables and 2 choices of R_{long}^2 . Figure 4 shows these plots. These plots show several of the features we discussed earlier in the paper. For example, at $\delta = 0$, the identified set is the singleton containing $\widehat{\beta}_{\text{med}}$. This is displayed as the horizontal intercept. The vertical intercept, in contrast, is the explain away breakdown point, as reported in column (2) of Table 2. Only two of these intercepts are visible given the range of the plots, however.

As δ varies, the identified set can contain either 1, 2, or 3 elements. For example, in the top left graph, there are two feasible values of β_{long} at $\delta = 1$. We can also see the vertical asymptote at $\delta = 1$, which implies that for δ ’s very close to 1, the identified set will contain either very large positive values, very large negative values, or both. This explains why the identified set under the restriction $|\delta| \leq \bar{\delta}$ is bounded and finite for $\bar{\delta} < 1$ but unbounded for $\bar{\delta} > 1$. It also explains why the bias adjustments in Table 3 are so sensitive to the

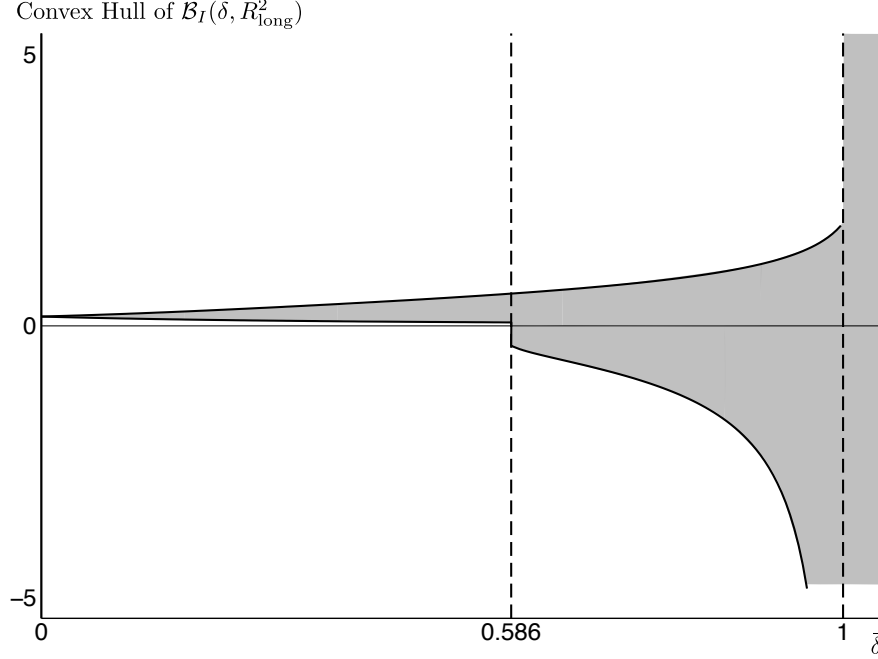


Figure 3: The convex hull of the estimated identified set for β_{long} under the assumption that R_{long}^2 is known and $|\delta| \leq \bar{\delta}$. Here we consider the “all associations” specification from Satyanath et al. (2017) and set $R_{\text{long}}^2 = 1$. This plot is the convex hull of the plot in figure 2.

exact choice of δ .

6.4 Empirical Conclusions

Overall, Table 2 shows just how different the explain away and sign change breakdown points can be in practice. By only examining the explain away breakdown point, Satyanath et al. (2017) conclude that

“selection on unobservables would have to be substantially stronger than selection on observables for our main result to be overturned.” (online appendix page 44)

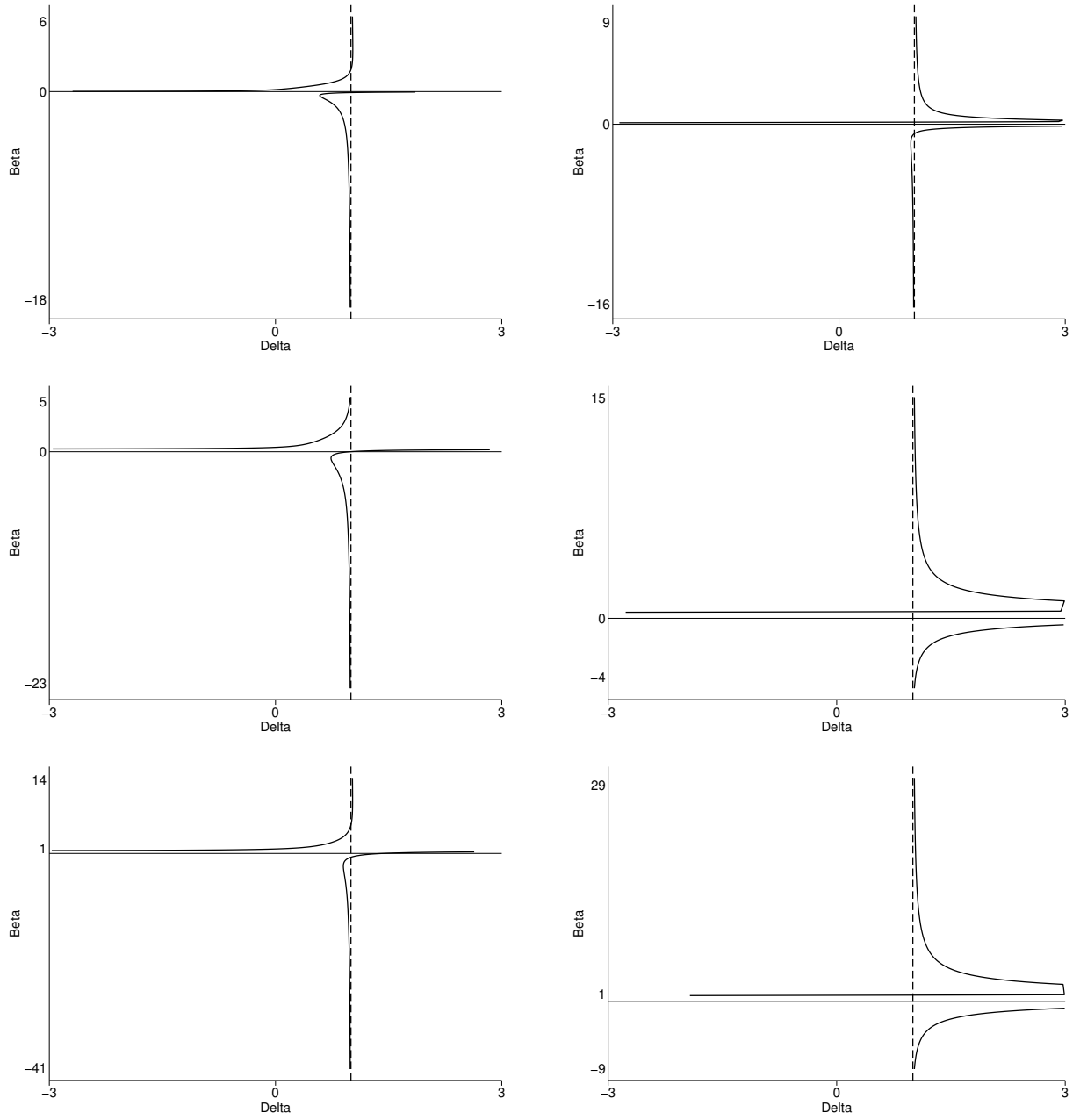
The language “for our main result to be overturned”, however, suggests that they are also interested in the robustness of their point estimates to sign changes. Indeed, if the true coefficient β_{long} was negative then higher social capital would lead to a *decrease* in Nazi party membership, which would support the opposite of their conclusion that social capital supported the Nazi party and therefore helped to undermine an existing democracy.

When examining the relative robustness of their results for the three versions of the treatment variable, a second concern arises. Based on the explain away breakdown point, they say:

“Note also that for the logic of our argument, the coefficient for military associations is not the most important—what matters is that the civic clubs and associations have an important effect on Nazi Party entry. This is never in doubt in the Altonji-Elder-Taber/Oster exercise.” (online appendix page 44)

When examining the sign change breakdown points, however, we see that the coefficient for military associations is the most robust, among the three versions of treatment. The coefficient for civic clubs and

Figure 4: Estimated identified sets $\widehat{\mathcal{B}}_I(\delta, R_{\text{long}}^2)$ for assessing the robustness of Satyanath et al. (2017) results to omitted variables. Left column: $R_{\text{long}}^2 = 1$. Right column: $R_{\text{long}}^2 = 1.3\widehat{R}_{\text{med}}^2$. Top row: Treatment variable includes all associations. Middle row: Treatment variable includes civic associations only. Bottom row: Treatment variable includes military associations only.



associations, in contrast, is not so obviously robust to omitted variables. Thus the sign change breakdown points suggest a more tentative conclusion about robustness than the definitive one drawn by the authors.

Moreover, our analysis of selection bias adjustments in Table 3 illustrates how sensitive these kinds of adjustments can be the choice of δ in practice. The bounds presented in Figure 3 provide a modified sensitivity analysis method that accurately represents the impact of omitted variables on the range of possible coefficient values.

7 A Survey of Regression Sensitivity Analysis in Practice

To understand how empirical researchers are using Oster (2019) in practice, we performed a survey of the empirical economics literature. We constructed the sample by looking at all papers published in the three year period from 2019 to 2021, which cite Oster (2019), and which were published in either *AER*, *JPE*, *QJE*, *REStud*, or *ECMA*. This gives 35 papers. 1 of these is an econometric theory which cited Oster (2019) as a related method only, and so we drop it from the sample. 3 of these papers used the methods in Oster (2019), but only in working paper drafts, not in the final publication version. We drop these 3 from the sample as well. This leaves 31 papers remaining. These papers are listed in Appendix B.

This sample size is similar to the survey of Blandhol et al. (2022), who found 112 papers that use 2SLS in the same five journals, over the 19 year span from January 2000 to October 2018. In fact, since we are only looking at 3 years of data, Oster’s (2019) method is more popular than 2SLS, on average, since 112 papers over 19 years is only about 18 papers every three years.² This is perhaps not surprising since Oster’s results can be applied to regressions based on many different identification strategies, including unconfoundedness, instrumental variables, and difference-in-differences (see section 4 of Diegert et al. 2022 for a related discussion).

There are three different ways in which researchers apply the results of Oster (2019):

1. Report the explain away breakdown point.
2. Report a bias adjusted estimand.
3. Informally discuss coefficient stability and cite Oster (2019) as justification.

Table 4 summarizes the distribution of these three methods across the 31 papers in our dataset. We see that roughly 75% of the papers use and report the results of one of the two formal sensitivity analysis methods. Among these, roughly half only report breakdown points, and about half report bias adjusted estimands. Only a few papers do both. The remaining roughly 25% of papers only informally discuss coefficient stability, and refer to Oster (2019) in their discussion. Some of those discussions suggest that they may have used one of the formal methods, but if so they did not report the findings. Although it is not the focus of this paper (cf. Diegert et al. 2022), none of the 31 papers discuss the validity of the exogenous controls assumption A4.

Consider the papers that report the explain away breakdown point. Very few also vary R_{long}^2 . Most use either $1.3\hat{R}_{\text{med}}^2$ or 1. One paper also used $2\hat{R}_{\text{med}}^2$. One paper presented the function $\hat{\delta}^{\text{bp,exact}}(R_{\text{long}}^2)$ as R_{long}^2 varied from \hat{R}_{med}^2 to 1. When discussing their results, researchers used a variety of phrases to describe their estimates of $\delta^{\text{bp,exact}}(R_{\text{long}}^2)$: It is the amount of selection on unobservables relative to observables needed...

²Blandhol et al. (2022) gathered their data from Web of Science, whereas we use Google Scholar. Using our search criteria, Web of Science finds 28 papers. So this difference in counts is not due to differences in the original data sources.

Table 4: How Empirical Researchers Use Oster (2019).

	Proportion	Count
Total	100%	31
1. Compute Explain Away Breakdown Points	42%	13
And vary R_{long}^2	15% (conditional)	2
2. Compute Bias Adjusted Estimates	39%	12
And vary δ	17% (conditional)	2
And vary R_{long}^2	42% (conditional)	5
3. Discuss Informally Only	26%	8
Both 1 and 2	6%	2

Note: Sample contains 31 papers published in *AER*, *QJE*, *JPE*, *ECMA*, or *REStud* between 2019–2021 and which use Oster (2019).

“to fully explain away the coefficient on the explanatory variable of interest” (Arbatli, Ashraf, Galor, and Klemp 2020, *ECMA*, page 747)

“to make our core estimate go away” (Dippel and Heblich 2021, *AER*, page 491)

“to drive the results observed” (Valencia Caicedo 2019, *QJE*, online appendix page 6)

“for our results to be spuriously generated” (Grosfeld, Sakalli, and Zhuravskaya 2020, *REStud*, page 296)

“to reverse the sign of the results” (Hoffman and Tadelis 2021, *JPE*, page 272)

to give a few examples from our survey. As we have shown in this paper, the last quote is incorrect, because the sign change and explain away breakdown points are not equivalent. The first quote correctly describes it as an explain away breakdown point. The other three quotes are more ambiguous, and the interpretation depends on what it means for an estimate to “go away,” or “to drive the results,” or for the “results to be spuriously generated.” If these statements refer to the sign of the main finding, rather than the conclusion that the effect is not exactly zero (but might be positive or negative), then these descriptions are incorrect. Our paper clarifies that empirical researchers must specify whether they are interested in robustness to sign changes, or merely to an exact zero.

Next consider the papers that report bias adjusted estimands. For any given specification, all of these papers report a scalar bias adjusted estimand. Most papers explicitly say that this is because they use Oster’s Proposition 1 (see our Prop. 4) and assume $\delta = 1$. Others do not clearly explain which result they are using to obtain their bias correction. None of the papers that use Oster’s Proposition 1 discuss validity of assumption A8. Nine of the twelve papers use just the single value $\delta = 1$. One paper considers $\delta = 0.5$ only (and no other values), but also uses the incorrect equation (6) as described in section 5. Two of the papers consider multiple δ values, but one of them also uses the incorrect equation (6). The other paper uses a variation of Oster’s results, which we have not formally analyzed, although we conjecture that the concerns we raise in our paper will apply to that extension as well. About half of the papers only consider one value of R_{long}^2 , always either $1.3\hat{R}_{\text{med}}^2$ or 1. In the other half, four papers also consider the second value $2\hat{R}_{\text{med}}^2$, while one paper considers up to 4 different values.

Overall, the results in our paper clarify and improve upon all three of the main ways in which empirical researchers use the methods of Oster (2019).

8 Conclusion: Best Practices for Regression Sensitivity Analysis

We conclude by providing several recommendations for best practices. Here we focus on Oster’s analysis, given its widespread adoption, but these recommendations often apply to other methods as well. Moreover, all of these recommendations are easy to implement using our companion Stata package `regsensitivity`.

1. Use sign change breakdown points to measure robustness

Throughout this paper we have emphasized the difference between the explain away and sign change breakdown points. Empirical researchers who are interested in the robustness of the sign of β_{long} to omitted variables should use sign change breakdown points. For Oster’s (2019) model, we developed this sign change breakdown point in section 4.

2. Present identified sets for a range of sensitivity value parameters

We showed that bias adjusted estimands can be extremely sensitive to small perturbations of the sensitivity parameter. We therefore recommend that, rather than presenting a single bias corrected estimand, researchers plot sharp bounds on the parameter of interest for a wide range of the sensitivity parameters.

3. Do not rely on coefficient stability alone to assess the role of omitted variables

Oster’s (2019) Proposition 1 (our Prop. 4) is a key result that links omitted variable bias to coefficient stability. She remarks, however, that her analysis “make[s] clear the limits of coefficient stability: it is possible that the coefficient will be completely unchanged with the addition of controls but there may still be a large bias on the treatment effect” (page 188). She later concludes that her “core insight is to recognize that coefficient stability on its own is at best uninformative and at worst very misleading” (page 203). Her caveats are based primarily on two features of her analysis: (i) A8 could be violated (see the first sentence of page 195), in which case her Proposition 1 (our Prop. 4) no longer holds and (ii) Even when A8 holds, the magnitude of the bias in that proposition also depends on R_{long}^2 .

Our analysis pushes her conclusion even further: We show that, even if the true value of R_{long}^2 was known, and even if A8 holds, then the bias adjustment in Oster’s Proposition 1 is extremely sensitive to the assumption that $\delta = 1$. In particular, regardless of how close the short and medium regression coefficients are, the true bias can be arbitrarily large when δ is close but not equal to 1. Moreover, as Diegert et al. (2022) emphasize, Oster’s Proposition 1 also requires the exogenous controls assumption A4.

Overall, the econometric theory evidence to date does not support the conclusion that coefficient stability alone is informative about the magnitude of omitted variable bias. This does not, however, imply that tests of coefficient stability are uninformative more generally. For example, Chen and Pearl (2015) show how coefficient stability tests can be viewed as falsification tests for causal graphs with enough exclusion restrictions to yield several different sets of observed covariates that lead to conditionally independent treatment assignment. Future work in econometrics would be helpful to further clarify the exact nature and value of coefficient stability exercises.

4. Use a range of values for auxiliary sensitivity parameters

Several sensitivity analysis methods focus on a primary sensitivity parameter, like Oster’s δ , but also require assumptions about an auxiliary sensitivity parameter, like R_{long}^2 . As we saw in section 7, researchers commonly fix this auxiliary parameter to a single value. In section 6, however, we saw that empirical results can often be quite sensitive to the choice of these auxiliary sensitivity parameters. We therefore recommend exploring a range of values for these auxiliary sensitivity parameters.

5. Consider using a variety of methods for sensitivity analysis

There is now a large menu of options for researchers interested in performing a regression sensitivity analysis. These methods are typically not nested, and make different assumptions about the role of omitted variables. Consequently, it can be useful to use several different methods together. We illustrated this in section 6 where we also applied the analysis of Diegert et al. (2022) to address the exogenous controls assumption. In that empirical application we saw that the modifications to Oster’s breakdown analysis developed in section 4 generally leads to similar conclusions about the robustness of the empirical results as Diegert et al. (2022). When a variety of sensitivity analyses all point in the same direction, researchers can be more confident in the robustness of their results.

References

- ALTONJI, J. G., T. CONLEY, T. E. ELDER, AND C. R. TABER (2019): “Methods for using selection on observed variables to address selection on unobserved variables,” *Working paper*.
- ALTONJI, J. G., T. E. ELDER, AND C. R. TABER (2005): “Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools,” *Journal of Political Economy*, 113, 151–184.
- (2008): “Using selection on observed variables to assess bias from unobservables when evaluating swan-ganz catheterization,” *American Economic Review P&P*, 98, 345–50.
- BAZZI, S., M. FISZBEIN, AND M. GEBRESILASSE (2020): “Frontier culture: The roots and persistence of “rugged individualism” in the United States,” *Econometrica*, 88, 2329–2368.
- BELLOWS, J. AND E. MIGUEL (2009): “War and local collective action in Sierra Leone,” *Journal of Public Economics*, 93, 1144–1157.
- BLANDHOL, C., J. BONNEY, M. MOGSTAD, AND A. TORGOVITSKY (2022): “When is TSLS actually late?” *Working paper*.
- BREUSCH, T. S. (1986): “Hypothesis testing in unidentified models,” *Review of Economic Studies*, 53, 635–651.
- CHEN, B. AND J. PEARL (2015): “Exogeneity and robustness,” *Working paper*.
- CINELLI, C. AND C. HAZLETT (2020): “Making sense of sensitivity: Extending omitted variable bias,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 82, 39–67.
- CLARKE, K. A. (2009): “Return of the phantom menace: Omitted variable bias in political research,” *Conflict Management and Peace Science*, 26, 46–66.
- CORNFIELD, J., W. HAENSZEL, E. C. HAMMOND, A. M. LILIENFELD, M. B. SHIMKIN, AND E. L. WYNDER (1959): “Smoking and lung cancer: Recent evidence and a discussion of some questions,” *Journal of the National Cancer Institute*, 22, 173–203.
- DE TOCQUEVILLE, A. (1835): *De La Démocratie en Amérique*.

- DIEGERT, P., M. A. MASTEN, AND A. POIRIER (2022): “Assessing omitted variable bias when the controls are endogenous,” *arXiv preprint arXiv:2206.02303*.
- FINKELSTEIN, A., M. GENTZKOW, AND H. WILLIAMS (2021): “Place-based drivers of mortality: Evidence from migration,” *American Economic Review*, 111, 2697–2735.
- FRANK, K. A. (2000): “Impact of a confounding variable on a regression coefficient,” *Sociological Methods & Research*, 29, 147–194.
- GONZALEZ, F. AND E. MIGUEL (2015): “War and local collective action in Sierra Leone: A comment on the use of coefficient stability approaches,” *Journal of Public Economics*, 128, 30–33.
- HOSMAN, C. A., B. B. HANSEN, AND P. W. HOLLAND (2010): “The sensitivity of linear regression coefficients’ confidence limits to the omission of a confounder,” *The Annals of Applied Statistics*, 4, 849–870.
- IMBENS, G. W. (2003): “Sensitivity to exogeneity assumptions in program evaluation,” *American Economic Review P&P*, 126–132.
- KRAUTH, B. (2016): “Bounding a linear causal effect using relative correlation restrictions,” *Journal of Econometric Methods*, 5, 117–141.
- MASTEN, M. A. AND A. POIRIER (2018): “Identification of treatment effects under conditional partial independence,” *Econometrica*, 86, 317–351.
- (2020): “Inference on breakdown frontiers,” *Quantitative Economics*, 11, 41–111.
- MAURO, R. (1990): “Understanding LOVE (left out variables error): A method for estimating the effects of omitted variables,” *Psychological Bulletin*, 108, 314.
- MIAN, A. AND A. SUFI (2014): “What explains the 2007–2009 drop in employment?” *Econometrica*, 82, 2197–2223.
- MURPHY, K. M. AND R. H. TOPEL (1990): “Efficiency wages reconsidered: Theory and evidence,” in *Advances in the Theory and Measurement of Unemployment*, Springer, 204–240.
- NEWKEY, K. AND D. MCFADDEN (1994): “Large sample estimation and hypothesis,” *Handbook of Econometrics*, 2112–2245.
- OSTER, E. (2019): “Unobservable selection and coefficient stability: Theory and evidence,” *Journal of Business & Economic Statistics*, 37, 187–204.
- PUTNAM, R. D. (2000): *Bowling Alone: The Collapse and Revival of American Community*, Simon and Schuster.
- ROBINS, J. M., A. ROTNITZKY, AND D. O. SCHARFSTEIN (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*, Springer, 1–94.
- ROSENBAUM, P. R. (1995): *Observational Studies*, Springer.
- (2002): *Observational Studies*, Springer, second ed.
- ROSENBAUM, P. R. AND D. B. RUBIN (1983): “Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome,” *Journal of the Royal Statistical Society, Series B*, 212–218.
- SATYANATH, S., N. VOIGTLÄNDER, AND H.-J. VOTH (2017): “Bowling for fascism: Social capital and the rise of the Nazi Party,” *Journal of Political Economy*, 125, 478–526.

A Proofs

Let $\mathbb{L}(A \mid B)$ denote the linear projection of the scalar random variable A on the random column vector B , defined as $\mathbb{L}(A \mid B) = \mathbb{E}(AB')\mathbb{E}(BB')^{-1}B$.

Section 2

Proof of Theorem 1. Let $b < \beta_{\text{baseline}}$. Define the following sets of relaxations:

$$\begin{aligned} R_1 &= \{r \geq 0 : (-\infty, b] \cap \mathcal{B}_I(r) = \emptyset\} \\ R_2 &= \{r \geq 0 : (-\infty, b] \cap \mathcal{B}_I(r) \neq \emptyset\}. \end{aligned}$$

These two sets partition $[0, +\infty)$ since $R_1 \cup R_2 = [0, +\infty)$ and $R_1 \cap R_2 = \emptyset$. Moreover, $R_1 \neq \emptyset$ since $b < \beta_{\text{baseline}}$ and $\mathcal{B}_I(0) = \{\beta_{\text{baseline}}\}$.

Since r is an interval relaxation, R_1 is a half interval starting at 0 (and is potentially infinite). This follows since for any $r' \leq r$, $\mathcal{B}_I(r') \subseteq \mathcal{B}_I(r)$ and hence

$$b < b' \text{ for all } b' \in \mathcal{B}_I(r) \quad \text{implies} \quad b < b' \text{ for all } b' \in \mathcal{B}_I(r').$$

Taking complements, $R_2 = [0, +\infty) \setminus R_1$ is an interval (which is potentially empty). Next,

$$\begin{aligned} r^{\text{bp},>}(b) &= \sup\{r \geq 0 : b' > b \text{ for all } b' \in \mathcal{B}_I(r)\} \\ &= \sup(R_1). \end{aligned}$$

The first line follows by definition of $r^{\text{bp},>}(b)$ and the second line follows by definition of R_1 . Similarly,

$$\begin{aligned} r^{\text{bp},=}(b) &= \inf\{r \geq 0 : b \in \mathcal{B}_I(r)\} \\ &= \inf(R_2). \end{aligned}$$

The first line follows by definition of the exact value breakdown point and the second line follows by definition of R_2 and since $\mathcal{B}_I(r)$ is an interval. Now consider two cases:

1. Suppose $R_1 = [0, \infty)$. Then $R_2 = \emptyset$. So $r^{\text{bp},>}(b) = \sup([0, \infty)) = +\infty = \inf(\emptyset) = r^{\text{bp},=}(b)$.
2. Suppose $R_1 = [0, \bar{r}]$ or $R_1 = [0, \bar{r})$ for finite \bar{r} . Then $R_2 = (\bar{r}, \infty)$ or $[\bar{r}, \infty)$. So $\sup(R_1) = \bar{r} = \inf(R_2)$ and thus $r^{\text{bp},>}(b) = r^{\text{bp},=}(b)$.

A similar proof holds when $b > \beta_{\text{baseline}}$. □

Proof of Proposition 1. We have

$$\begin{aligned} \{b \in \mathbb{R} : r^{\text{bp},=}(b) \leq r\} &= \{b \in \mathbb{R} : \inf\{\tilde{r} \geq 0 : b \in \mathcal{B}_I(\tilde{r})\} \leq r\} \\ &= \{b \in \mathbb{R} : b \in \mathcal{B}_I(r)\} \\ &= \mathcal{B}_I(r). \end{aligned}$$

The first equality follows by definition of $r^{\text{bp},=}(b)$. The second equality follows from the monotonicity of $\mathcal{B}_I(\tilde{r})$ in \tilde{r} . Specifically, the statement

$$\inf\{\tilde{r} \geq 0 : b \in \mathcal{B}_I(\tilde{r})\} \leq r$$

means there exists an $\tilde{r} \leq r$ such that $b \in \mathcal{B}_I(\tilde{r})$. So monotonicity implies that $b \in \mathcal{B}_I(r)$ since $r \geq \tilde{r}$. This is an equivalence because $b \in \mathcal{B}_I(r)$ implies there exists $\tilde{r} \leq r$ (namely, $\tilde{r} = r$) such that $b \in \mathcal{B}_I(\tilde{r})$. □

Section 3

It is useful to also consider the linear projection of X onto $(1, W_1, W_2)$:

$$X = \pi'_{1,\text{long}} W_1 + \pi_{2,\text{long}} W_2 + X^{\perp W}. \quad (7)$$

The exogenous controls assumption A4 implies that

$$\begin{aligned} \pi_1 &\equiv \text{var}(W_1)^{-1} \text{cov}(W_1, X) \\ &= \text{var}(W_1)^{-1} \text{cov}(W_1, W'_1 \pi_1 + W_2 \pi_2 + V) \\ &= \text{var}(W_1)^{-1} (\text{var}(W_1) \pi_1 + \text{cov}(W_1, W_2) \pi_2 + \text{cov}(W_1, V)) \\ &= \pi_{1,\text{long}}. \end{aligned}$$

Similarly, if we let π_2 denote the coefficient on W_2 from the linear projection of X onto $(1, W_2)$ then $\pi_2 = \pi_{2,\text{long}}$. Given these results, we use (π_1, π_2) to represent the common values of the regression coefficients throughout the proofs.

We denote the omitted variable bias by $B = \beta_{\text{med}} - \beta_{\text{long}}$. We begin by showing the following identities that we'll use throughout our proofs.

Lemma 1 (Regression Algebra). Suppose A1, A4, and A5 hold. Then the following hold:

$$\gamma_{1,\text{long}} = \gamma_{1,\text{med}} + B\pi_1 \quad (8)$$

$$\pi_2 \gamma_{2,\text{long}} = B \text{var}(X^{\perp W_1}) \quad (9)$$

$$\text{cov}(X, \gamma'_{1,\text{med}} W_1) = (\beta_{\text{short}} - \beta_{\text{med}}) \text{var}(X) \quad (10)$$

$$(R_{\text{med}}^2 - R_{\text{short}}^2) \text{var}(Y) = \text{var}(\gamma'_{1,\text{med}} W_1) - \text{var}(X)(\beta_{\text{short}} - \beta_{\text{med}})^2. \quad (11)$$

Proof of Lemma 1. Proof of (8)

We have that

$$\begin{aligned} &\gamma_{1,\text{med}} + (\beta_{\text{med}} - \beta_{\text{long}}) \pi_1 \\ &= \text{var}(W_1)^{-1} (\beta_{\text{med}} \text{cov}(W_1, X) + \text{var}(W_1) \gamma_{1,\text{med}} - \beta_{\text{med}} \text{cov}(W_1, X)) + B \text{var}(W_1)^{-1} \text{cov}(W_1, X) \\ &= \text{var}(W_1)^{-1} (\text{cov}(W_1, Y) - \beta_{\text{med}} \text{cov}(W_1, X)) + B \text{var}(W_1)^{-1} \text{cov}(W_1, X) \\ &= \text{var}(W_1)^{-1} (\text{cov}(W_1, Y) - \beta_{\text{long}} \text{cov}(W_1, X)) \\ &= \text{var}(W_1)^{-1} (\text{cov}(W_1, \beta_{\text{long}} X + \gamma'_{1,\text{long}} W_1 + \gamma_{2,\text{long}} W_2 + Y^{\perp X, W}) - \beta_{\text{long}} \text{cov}(W_1, X)) \\ &= \gamma_{1,\text{long}}. \end{aligned}$$

The second equality follows from

$$\begin{aligned} \text{cov}(W_1, Y) &= \text{cov}(W_1, \mathbb{E}(Y \mid 1, X, W_1) + Y^{\perp X, W_1}) \\ &= \text{cov}(W_1, \beta_{\text{med}} X + \gamma'_{1,\text{med}} W_1) \\ &= \beta_{\text{med}} \text{cov}(W_1, X) + \gamma'_{1,\text{med}} \text{var}(W_1). \end{aligned}$$

The final equality follows from A4.

Proof of (9)

By definition of β_{med} and the FWL Theorem,

$$\begin{aligned} \beta_{\text{med}} \text{var}(X^{\perp W_1}) &= \text{cov}(Y, X^{\perp W_1}) \\ &= \text{cov}(\beta_{\text{long}} X + \gamma'_{1,\text{long}} W_1 + \gamma_{2,\text{long}} W_2 + Y^{\perp X, W}, X^{\perp W_1}) \\ &= \beta_{\text{long}} \text{var}(X^{\perp W_1}) + \gamma_{2,\text{long}} \text{cov}(W_2, X^{\perp W_1}). \end{aligned}$$

To finish the proof, we note that $\text{cov}(W_2, X^{\perp W_1}) = \text{cov}(W_2, X - \pi'_1 W_1) = \text{cov}(W_2, \pi'_1 W_1 + \pi_2 W_2 + X^{\perp W} - \pi'_1 W_1) = \pi_2$ by A5.

Proof of (10)

By definition of β_{short} ,

$$\begin{aligned}\beta_{\text{short}} \text{var}(X) &= \text{cov}(Y, X) \\ &= \text{cov}(\beta_{\text{med}}X + \gamma'_{1,\text{med}}W_1 + Y^{\perp X, W_1}, X) \\ &= \beta_{\text{med}} \text{var}(X) + \text{cov}(X, \gamma'_{1,\text{med}}W_1).\end{aligned}$$

Proof of (11)

By the definitions of R_{med}^2 and R_{short}^2 ,

$$\begin{aligned}(R_{\text{med}}^2 - R_{\text{short}}^2) \text{var}(Y) + \text{var}(X)(\beta_{\text{short}} - \beta_{\text{med}})^2 \\ &= (\text{var}(X\beta_{\text{med}} + \gamma'_{1,\text{med}}W_1) - \text{var}(X\beta_{\text{short}})) + \text{var}(X)(\beta_{\text{short}} - \beta_{\text{med}})^2 \\ &= \text{var}(X)(\beta_{\text{med}}^2 - \beta_{\text{short}}^2) + 2\beta_{\text{med}} \text{cov}(X, \gamma'_{1,\text{med}}W_1) + \text{var}(\gamma'_{1,\text{med}}W_1) + \text{var}(X)(\beta_{\text{short}} - \beta_{\text{med}})^2 \\ &= \text{var}(X)(\beta_{\text{med}}^2 - \beta_{\text{short}}^2 + 2\beta_{\text{med}}(\beta_{\text{short}} - \beta_{\text{med}}) + (\beta_{\text{short}} - \beta_{\text{med}})^2) + \text{var}(\gamma'_{1,\text{med}}W_1) \\ &= \text{var}(\gamma'_{1,\text{med}}W_1).\end{aligned}$$

The third equality follows from equation (10). □

Lemma 2 (Oster's 3rd Equation). Suppose assumptions A1–A5 hold. Then

$$(R_{\text{long}}^2 - R_{\text{med}}^2) \text{var}(Y) \delta \text{cov}(X, \gamma'_{1,\text{long}}W_1) = B \text{var}(X^{\perp W_1}) (\text{var}(\gamma'_{1,\text{long}}W_1) - B \delta \text{cov}(X, \gamma'_{1,\text{long}}W_1)).$$

Proof of Lemma 2. By definition,

$$\begin{aligned}R_{\text{long}}^2 \text{var}(Y) \\ &= \text{var}(\beta_{\text{long}}X + \gamma'_{1,\text{long}}W_1 + \gamma_{2,\text{long}}W_2) \\ &= \beta_{\text{long}}^2 \text{var}(X) + \text{var}(\gamma'_{1,\text{long}}W_1) + \text{var}(\gamma_{2,\text{long}}W_2) \\ &\quad + 2 \text{cov}(\beta_{\text{long}}X, \gamma'_{1,\text{long}}W_1) + 2 \text{cov}(\beta_{\text{long}}X, \gamma_{2,\text{long}}W_2) + 2 \text{cov}(\gamma'_{1,\text{long}}W_1, \gamma_{2,\text{long}}W_2) \\ &= \beta_{\text{long}}^2 \text{var}(X) + \text{var}(\gamma'_{1,\text{long}}W_1) + \text{var}(\gamma_{2,\text{long}}W_2) \\ &\quad + 2\beta_{\text{long}} \text{cov}(X, \gamma'_{1,\text{long}}W_1) + 2\beta_{\text{long}} \text{cov}(X, \gamma_{2,\text{long}}W_2) \\ &= \beta_{\text{long}}^2 \text{var}(X) + \text{var}(\gamma'_{1,\text{long}}W_1) + \text{var}(\gamma_{2,\text{long}}W_2) + 2\beta_{\text{long}} \text{cov}(X, \gamma'_{1,\text{long}}W_1) + 2\beta_{\text{long}}B \text{var}(X^{\perp W_1}).\end{aligned}$$

The third equality follows from assumption A4. The fourth follows equation (9). We can show that

$$\begin{aligned}R_{\text{med}}^2 \text{var}(Y) \\ &= \text{var}(\beta_{\text{med}}X + \gamma'_{1,\text{med}}W_1) \\ &= \beta_{\text{med}}^2 \text{var}(X) + \text{var}(\gamma'_{1,\text{med}}W_1) + 2\beta_{\text{med}} \text{cov}(X, \gamma'_{1,\text{med}}W_1) \\ &= (\beta_{\text{long}} + B)^2 \text{var}(X) + \text{var}((\gamma_{1,\text{long}} - B\pi_1)'W_1) + 2(\beta_{\text{long}} + B) \text{cov}(X, (\gamma_{1,\text{long}} - B\pi_1)'W_1) \\ &= (\beta_{\text{long}} + B)^2 \text{var}(X) + \text{var}(\gamma'_{1,\text{long}}W_1) + B^2 (\text{var}(X) - \text{var}(X^{\perp W_1})) \\ &\quad - 2B \text{cov}(X, \gamma'_{1,\text{long}}W_1) + 2(\beta_{\text{long}} + B) \left[\text{cov}(X, \gamma'_{1,\text{long}}W_1) + B (\text{var}(X^{\perp W_1}) - \text{var}(X)) \right] \\ &= \beta_{\text{long}}^2 \text{var}(X) + B^2 \text{var}(X) + 2\beta_{\text{long}}B \text{var}(X) + \text{var}(\gamma'_{1,\text{long}}W_1) + B^2 (\text{var}(X) - \text{var}(X^{\perp W_1})) \\ &\quad - 2B \text{cov}(X, \gamma'_{1,\text{long}}W_1) + 2(\beta_{\text{long}} + B) \left[\text{cov}(X, \gamma'_{1,\text{long}}W_1) + B (\text{var}(X^{\perp W_1}) - \text{var}(X)) \right] \\ &= \beta_{\text{long}}^2 \text{var}(X) + \text{var}(\gamma'_{1,\text{long}}W_1) + (2\beta_{\text{long}}B + B^2) \text{var}(X^{\perp W_1}) + 2\beta_{\text{long}} \text{cov}(X, \gamma'_{1,\text{long}}W_1).\end{aligned}$$

The third equality follows from equation (8). Subtracting this from the R-squared long equation gives

$$\begin{aligned}
& (R_{\text{long}}^2 - R_{\text{med}}^2) \text{var}(Y) \\
&= \beta_{\text{long}}^2 \text{var}(X) + \text{var}(\gamma'_{1,\text{long}} W_1) + \text{var}(\gamma_{2,\text{long}} W_2) + 2\beta_{\text{long}} \text{cov}(X, \gamma'_{1,\text{long}} W_1) + 2\beta_{\text{long}} B \text{var}(X^{\perp W_1}) \\
&\quad - \beta_{\text{long}}^2 \text{var}(X) - \text{var}(\gamma'_{1,\text{long}} W_1) - (2\beta_{\text{long}} B + B^2) \text{var}(X^{\perp W_1}) - 2\beta_{\text{long}} \text{cov}(X, \gamma'_{1,\text{long}} W_1) \\
&= \text{var}(\gamma_{2,\text{long}} W_2) - B^2 \text{var}(X^{\perp W_1}).
\end{aligned}$$

Then, multiply both sides of this equation by $\delta \text{cov}(X, \gamma'_{1,\text{long}} W_1)$:

$$\begin{aligned}
& (R_{\text{long}}^2 - R_{\text{med}}^2) \text{var}(Y) \delta \text{cov}(X, \gamma'_{1,\text{long}} W_1) \\
&= \delta \text{cov}(X, \gamma'_{1,\text{long}} W_1) \text{var}(\gamma_{2,\text{long}} W_2) - B^2 \text{var}(X^{\perp W_1}) \delta \text{cov}(X, \gamma'_{1,\text{long}} W_1) \\
&= \text{cov}(X, \gamma_{2,\text{long}} W_2) \text{var}(\gamma'_{1,\text{long}} W_1) - B^2 \text{var}(X^{\perp W_1}) \delta \text{cov}(X, \gamma'_{1,\text{long}} W_1) \\
&= \pi_2 \gamma_{2,\text{long}} \text{var}(\gamma'_{1,\text{long}} W_1) - B^2 \text{var}(X^{\perp W_1}) \delta \text{cov}(X, \gamma'_{1,\text{long}} W_1) \\
&= B \text{var}(X^{\perp W_1}) \text{var}(\gamma'_{1,\text{long}} W_1) - B^2 \text{var}(X^{\perp W_1}) \delta \text{cov}(X, \gamma'_{1,\text{long}} W_1).
\end{aligned}$$

The second equality follows from A2, and the fourth from equation (9). \square

Proof of Theorem 2. We split this proof into two parts.

Part 1: First we show that $\beta_{\text{long}} \in \mathcal{B}_I(\delta, R_{\text{long}}^2)$. To prove this, it suffices to show that $B = \beta_{\text{med}} - \beta_{\text{long}}$ satisfies $f(B, \delta, R_{\text{long}}^2) = 0$ and that $\beta_{\text{long}} \notin \mathcal{B}_{\text{A3fail}}$. By Lemma 2 we can write

$$(R_{\text{long}}^2 - R_{\text{med}}^2) \text{var}(Y) \delta \text{cov}(X, \gamma'_{1,\text{long}} W_1) = B \text{var}(X^{\perp W_1}) (\text{var}(\gamma'_{1,\text{long}} W_1) - B \delta \text{cov}(X, \gamma'_{1,\text{long}} W_1)).$$

Substituting $\gamma_{1,\text{long}} = \gamma_{1,\text{med}} + B\pi_1$, as shown in the Lemma 1, we obtain the equation

$$\begin{aligned}
& \delta(R_{\text{long}}^2 - R_{\text{med}}^2) \text{var}(Y) (\text{cov}(X, \gamma'_{1,\text{med}} W_1) + B \text{var}(\pi'_1 W_1)) \\
&= B \text{var}(X^{\perp W_1}) \left[\text{var}(\gamma'_{1,\text{med}} W_1) + 2B \text{cov}(X, \gamma'_{1,\text{med}} W_1) \right. \\
&\quad \left. + B^2 \text{var}(\pi'_1 W_1) - B \delta (\text{cov}(X, \gamma'_{1,\text{med}} W_1) + B \text{var}(\pi'_1 W_1)) \right]
\end{aligned}$$

where we used that $\text{cov}(X, \pi'_1 W_1) = \text{var}(\pi'_1 W_1)$, which follows from equation (7) and Assumption A4.

Collecting terms, this equation is equivalent to $f(B, \delta, R_{\text{long}}^2) = 0$ as defined above. Recall that $\gamma_{1,\text{long}} = \gamma_{1,\text{med}} + B\pi_1$. By A3, $\gamma_{1,\text{long}} \neq 0$ and thus $\beta_{\text{long}} = \beta_{\text{med}} - B \notin \mathcal{B}_{\text{A3fail}}$. Therefore, $\beta_{\text{long}} \in \mathcal{B}_I(\delta, R_{\text{long}}^2)$.

Part 2: Next we show that $\mathcal{B}_I(\delta, R_{\text{long}}^2)$ is sharp. Let $b \in \mathcal{B}_I(\delta, R_{\text{long}}^2)$. Let $B = \beta_{\text{med}} - b$. This implies that $f(B, \delta, R_{\text{long}}^2) = 0$ and $b \notin \mathcal{B}_{\text{A3fail}}$. Fix the distribution of the observables (Y, X, W_1) and assume that $\text{var}(Y, X, W_1)$ is positive definite.

Let

$$\begin{aligned}
g_1 &= \gamma_{1,\text{med}} + B\pi_1 \\
g_2 &= ((R_{\text{long}}^2 - R_{\text{med}}^2) \text{var}(Y) + B^2 \text{var}(X^{\perp W_1}))^{1/2} \\
p_1 &= \text{var}(W_1)^{-1} \text{cov}(W_1, X) \\
p_2 &= \frac{B \text{var}(X^{\perp W_1})}{((R_{\text{long}}^2 - R_{\text{med}}^2) \text{var}(Y) + B^2 \text{var}(X^{\perp W_1}))^{1/2}}.
\end{aligned}$$

Note that g_2 is well defined and strictly positive by $R_{\text{long}}^2 > R_{\text{med}}^2$ and $\text{var}(Y) > 0$, which follows from A1. p_2 is well defined by $g_2 > 0$. $g_1 \neq 0$ by $b \notin \mathcal{B}_{\text{A3fail}}$. Also note that $p_1 = \pi_1$ by our assumptions.

Under these choices, we have

$$\begin{aligned}
b \operatorname{var}(X) + g'_1 \operatorname{var}(W_1) p_1 + p_2 g_2 &= (\beta_{\text{med}} - B) \operatorname{var}(X) + (\gamma_{1,\text{med}} + B\pi_1)' \operatorname{var}(W_1) \pi_1 + B \operatorname{var}(X^{\perp W_1}) \\
&= \beta_{\text{med}} \operatorname{var}(X) + \gamma'_{1,\text{med}} \operatorname{var}(W_1) \pi_1 \\
&= \beta_{\text{med}} \operatorname{var}(X) + (\operatorname{cov}(Y, W_1) - \beta_{\text{med}} \operatorname{cov}(X, W_1)) \operatorname{var}(W_1)^{-1} \operatorname{var}(W_1) \pi_1 \\
&= \beta_{\text{med}} \operatorname{var}(X^{\perp W_1}) + \operatorname{cov}(Y, W_1)' \pi_1 \\
&= \operatorname{cov}(Y, X)
\end{aligned} \tag{12}$$

and

$$\begin{aligned}
bp'_1 \operatorname{var}(W_1) + g'_1 \operatorname{var}(W_1) &= (\beta_{\text{med}} - B) \pi'_1 \operatorname{var}(W_1) + (\gamma_{1,\text{med}} + B\pi_1)' \operatorname{var}(W_1) \\
&= \beta_{\text{med}} \pi'_1 \operatorname{var}(W_1) + (\operatorname{cov}(Y, W_1) - \beta_{\text{med}} \operatorname{cov}(X, W_1)) \operatorname{var}(W_1)^{-1} \operatorname{var}(W_1) \\
&= \operatorname{cov}(Y, W_1).
\end{aligned} \tag{13}$$

Our goal is to construct a joint distribution for (Y, X, W_1, W_2) that is consistent with the observed distribution of (Y, X, W_1) , that is such that $\operatorname{var}(Y, X, W_1, W_2)$ is finite, $\operatorname{var}(X, W_1, W_2)$ is positive definite, A3 holds, consistent with the value of R_{long}^2 , A2 holds, and such that projection equations (1) and (7) hold.

To help define the joint distribution of (Y, X, W_1, W_2) , let

$$\Sigma = \begin{pmatrix} \operatorname{var}(Y) & \operatorname{cov}(Y, X) & \operatorname{cov}(Y, W_1) & bp_2 + g_2 \\ \operatorname{cov}(Y, X) & \operatorname{var}(X) & \operatorname{cov}(X, W_1) & p_2 \\ \operatorname{cov}(W_1, Y) & \operatorname{cov}(W_1, X) & \operatorname{var}(W_1) & 0 \\ bp_2 + g_2 & p_2 & 0 & 1 \end{pmatrix}.$$

The top left $(d_1 + 2) \times (d_1 + 2)$ block is positive definite by assumption. By Sylvester's criterion, we can then check that this matrix is positive semi-definite by verifying its determinant is nonnegative. Before doing so, we first verify that $\operatorname{var}(X^{\perp W_1}) > p_2^2$:

$$\begin{aligned}
\operatorname{var}(X^{\perp W_1}) - p_2^2 &= \operatorname{var}(X^{\perp W_1}) - \frac{B^2 \operatorname{var}(X^{\perp W_1})^2}{(R_{\text{long}}^2 - R_{\text{med}}^2) \operatorname{var}(Y) + B^2 \operatorname{var}(X^{\perp W_1})} \\
&= \operatorname{var}(X^{\perp W_1}) \frac{(R_{\text{long}}^2 - R_{\text{med}}^2) \operatorname{var}(Y)}{(R_{\text{long}}^2 - R_{\text{med}}^2) \operatorname{var}(Y) + B^2 \operatorname{var}(X^{\perp W_1})} \\
&> 0.
\end{aligned}$$

The inequality follows from the positive-definiteness of $\operatorname{var}(Y, X, W_1)$, $R_{\text{long}}^2 > R_{\text{med}}^2$, and by $\operatorname{var}(Y) > 0$.

This inequality allows us to show that the matrix

$$\Sigma' = \begin{pmatrix} \operatorname{var}(X) & \operatorname{cov}(X, W_1) & p_2 \\ \operatorname{cov}(W_1, X) & \operatorname{var}(W_1) & 0 \\ p_2 & 0 & 1 \end{pmatrix}$$

is invertible via the partitioned inverse matrix formula:

$$\begin{aligned}
&\begin{pmatrix} \operatorname{var}(X) & \operatorname{cov}(X, W_1) & p_2 \\ \operatorname{cov}(W_1, X) & \operatorname{var}(W_1) & 0 \\ p_2 & 0 & 1 \end{pmatrix}^{-1} \\
&= (\operatorname{var}(X) - \operatorname{cov}(X, W_1) \operatorname{var}(W_1)^{-1} \operatorname{cov}(W_1, X) - p_2^2)^{-1} \\
&\cdot \begin{pmatrix} 1 & & - (p'_1 \ p_2) \\ \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} & (\operatorname{var}(X) - \operatorname{cov}(X, W_1) \operatorname{var}(W_1)^{-1} \operatorname{cov}(W_1, X) - p_2^2) & \begin{pmatrix} \operatorname{var}(W_1)^{-1} & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \begin{pmatrix} p'_1 & p_2 \end{pmatrix} \end{pmatrix}
\end{aligned}$$

$$= (\text{var}(X^{\perp W_1}) - p_2^2)^{-1} \begin{pmatrix} 1 & -p'_1 & -p_2 \\ -p_1 & \text{var}(W_1)^{-1} (\text{var}(X^{\perp W_1}) - p_2^2) + p_1 p'_1 & p_1 p_2 \\ -p_2 & p_2 p'_1 & \text{var}(X^{\perp W_1}) \end{pmatrix}.$$

Since Σ' is invertible, we can use the partitioned determinant formula to compute the determinant of Σ .

$$\begin{aligned} & \det(\Sigma) \\ &= \det \begin{pmatrix} \text{var}(Y) & \text{cov}(Y, X) & \text{cov}(Y, W_1) & bp_2 + g_2 \\ \text{cov}(Y, X) & \text{var}(X) & \text{cov}(X, W_1) & p_2 \\ \text{cov}(W_1, Y) & \text{cov}(W_1, X) & \text{var}(W_1) & 0 \\ bp_2 + g_2 & p_2 & 0 & 1 \end{pmatrix} \\ &= 1 \cdot \det \begin{pmatrix} \text{var}(Y) - (\text{cov}(Y, X) & \text{cov}(Y, W_1) & bp_2 + g_2) \begin{pmatrix} \text{var}(X) & \text{cov}(X, W_1) & p_2 \\ \text{cov}(W_1, X) & \text{var}(W_1) & 0 \\ p_2 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \text{cov}(Y, X) \\ \text{cov}(W_1, Y) \\ bp_2 + g_2 \end{pmatrix} \end{pmatrix} \\ &= \text{var}(Y) - (\text{cov}(Y, X) & \text{cov}(Y, W_1) & bp_2 + g_2) (\text{var}(X^{\perp W_1}) - p_2^2)^{-1} \\ &\quad \cdot \begin{pmatrix} 1 & -p'_1 & -p_2 \\ -p_1 & \text{var}(W_1)^{-1} (\text{var}(X^{\perp W_1}) - p_2^2) + p_1 p'_1 & p_1 p_2 \\ -p_2 & p_2 p'_1 & \text{var}(X^{\perp W_1}) \end{pmatrix} \begin{pmatrix} \text{cov}(Y, X) \\ \text{cov}(W_1, Y) \\ bp_2 + g_2 \end{pmatrix}. \end{aligned}$$

Using equations (12)–(13), we can write

$$\begin{aligned} & \begin{pmatrix} 1 & -p'_1 & -p_2 \\ -p_1 & \text{var}(W_1)^{-1} (\text{var}(X^{\perp W_1}) - p_2^2) + p_1 p'_1 & p_1 p_2 \\ -p_2 & p_2 p'_1 & \text{var}(X^{\perp W_1}) \end{pmatrix} \begin{pmatrix} \text{cov}(Y, X) \\ \text{cov}(W_1, Y) \\ bp_2 + g_2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -p'_1 & -p_2 \\ -p_1 & \text{var}(W_1)^{-1} (\text{var}(X^{\perp W_1}) - p_2^2) + p_1 p'_1 & p_1 p_2 \\ -p_2 & p_2 p'_1 & \text{var}(X^{\perp W_1}) \end{pmatrix} \begin{pmatrix} b \text{var}(X) + g'_1 \text{var}(W_1)p_1 + p_2 g_2 \\ b \text{var}(W_1)p_1 + \text{var}(W_1)g_1 \\ bp_2 + g_2 \end{pmatrix} \\ &= \begin{pmatrix} b \text{var}(X) + g'_1 \text{var}(W_1)p_1 + p_2 g_2 - b \text{var}(p'_1 W_1) - p'_1 \text{var}(W_1)g_1 - bp_2^2 - g_2 p_2 \\ -p_1(\beta_{\text{long}} \text{var}(X) + g'_1 \text{var}(W_1)p_1 + p_2 g_2) + (\text{var}(X^{\perp W_1}) - p_2^2)(bp_1 + g_1) \\ \quad + bp_1 \text{var}(p'_1 W_1) + p_1 p'_1 \text{var}(W_1)g_1 + p_1 p_2 (bp_2 + g_2) \\ -p_2(b \text{var}(X) + g'_1 \text{var}(W_1)p_1 + p_2 g_2) + p_2 b \text{var}(p'_1 W_1) + p_2 p'_1 \text{var}(W_1)g_1 + \text{var}(X^{\perp W_1})(bp_2 + g_2) \end{pmatrix} \\ &= \begin{pmatrix} b(\text{var}(X^{\perp W_1}) - p_2^2) \\ (\text{var}(X^{\perp W_1}) - p_2^2)g_1 \\ -p_2(p_2 g_2) + \text{var}(X^{\perp W_1})(g_2) \end{pmatrix} \\ &= (\text{var}(X^{\perp W_1}) - p_2^2) \begin{pmatrix} b \\ g_1 \\ g_2 \end{pmatrix}. \end{aligned}$$

Therefore,

$$\begin{aligned} & \text{var}(Y) - (\text{cov}(Y, X) & \text{cov}(Y, W_1) & bp_2 + g_2) (\text{var}(X^{\perp W_1}) - p_2^2)^{-1} \\ & \cdot \begin{pmatrix} 1 & -p'_1 & -p_2 \\ -p_1 & \text{var}(W_1)^{-1} (\text{var}(X^{\perp W_1}) - p_2^2) + p_1 p'_1 & p_1 p_2 \\ -p_2 & p_2 p'_1 & \text{var}(X^{\perp W_1}) \end{pmatrix} \begin{pmatrix} \text{cov}(Y, X) \\ \text{cov}(Y, W_1) \\ bp_2 + g_2 \end{pmatrix} \\ &= \text{var}(Y) - (b \text{var}(X) + g'_1 \text{var}(W_1)p_1 + p_2 g_2 & b \text{var}(W_1)p_1 + \text{var}(W_1)g_1 & bp_2 + g_2) \begin{pmatrix} b \\ g_1 \\ g_2 \end{pmatrix} \end{aligned}$$

$$= \text{var}(Y) - b^2 \text{var}(X) - 2bg'_1 \text{var}(W_1)p_1 - 2bp_2g_2 - \text{var}(g'_1W_1) - g_2^2.$$

We now substitute the choices for (b, g_1, p_1, p_2, g_2) and obtain

$$\begin{aligned} \det(\Sigma) &= \text{var}(Y) - (\beta_{\text{med}} - B)^2 \text{var}(X) - 2(\beta_{\text{med}} - B)(\gamma_{1,\text{med}} + B\pi_1)' \text{var}(W_1)\pi_1 - 2(\beta_{\text{med}} - B)B \text{var}(X^{\perp W_1}) \\ &\quad - \text{var}((\gamma_{1,\text{med}} + B\pi_1)'W_1) - (R_{\text{long}}^2 - R_{\text{med}}^2) \text{var}(Y) - B^2 \text{var}(X^{\perp W_1}) \\ &= \text{var}(Y)(1 - R_{\text{long}}^2) - \beta_{\text{med}}^2 \text{var}(X) + 2\beta_{\text{med}}B \text{var}(X) - B^2 \text{var}(X) - 2\beta_{\text{med}}\gamma'_{1,\text{med}} \text{var}(W_1)\pi_1 \\ &\quad - 2\beta_{\text{med}}B \text{var}(\pi'_1W_1) + 2B\gamma'_{1,\text{med}} \text{var}(W_1)\pi_1 + 2B^2 \text{var}(W'_1\pi_1) - 2\beta_{\text{med}}B \text{var}(X^{\perp W_1}) + 2B^2 \text{var}(X^{\perp W_1}) \\ &\quad - \text{var}(\gamma'_{1,\text{med}}W_1) - 2B\pi'_1 \text{var}(W_1)\gamma_{1,\text{med}} - B^2 \text{var}(\pi'_1W_1) + R_{\text{med}}^2 \text{var}(Y) - B^2 \text{var}(X^{\perp W_1}) \\ &= \text{var}(Y)(1 - R_{\text{long}}^2) - \left(\beta_{\text{med}}^2 \text{var}(X) + 2\beta_{\text{med}}\gamma'_{1,\text{med}} \text{var}(W_1)\pi_1 + \text{var}(\gamma'_{1,\text{med}}W_1) \right) + R_{\text{med}}^2 \text{var}(Y) \\ &= \text{var}(Y)(1 - R_{\text{long}}^2) - R_{\text{med}}^2 \text{var}(Y) + R_{\text{med}}^2 \text{var}(Y) \\ &= \text{var}(Y)(1 - R_{\text{long}}^2) \\ &\geq 0 \end{aligned}$$

by $R_{\text{long}}^2 \leq 1$. From this, we conclude that Σ is positive semidefinite.

To construct W_2 , let

$$\begin{pmatrix} \widetilde{W}_1 \\ \widetilde{X} \\ \widetilde{Y} \end{pmatrix} = \text{Chol}(\text{var}(W_1, X, Y)^{-1})' \begin{pmatrix} W_1 \\ X \\ Y \end{pmatrix}.$$

$\text{Chol}(\cdot)$ is the lower-triangular Cholesky decomposition of a positive-definite matrix. Let \widetilde{W}_2 be a unit variance random variable that is uncorrelated with $(\widetilde{W}_1, \widetilde{X}, \widetilde{Y})$ and let

$$\begin{pmatrix} W_2 \\ W_1 \\ X \\ Y \end{pmatrix} = \text{Chol}(\Sigma)' \begin{pmatrix} \widetilde{W}_2 \\ \widetilde{W}_1 \\ \widetilde{X} \\ \widetilde{Y} \end{pmatrix}.$$

This vector has variance matrix Σ and leaves (W_1, X, Y) unchanged due to the triangular structure of $\text{Chol}(\Sigma)$ and $\text{Chol}(\text{var}(W_1, X, Y)^{-1})$.

Therefore, this joint distribution of (Y, X, W_1, W_2) is consistent with the observed distribution of (Y, X, W_1) and has a finite variance matrix equal to Σ . $\text{var}(X, W_1, W_2) = \Sigma'$ is positive definite since it is a positive semidefinite matrix satisfying $\det(\Sigma') > 0$. Therefore, assumption A1 holds.

We now show this joint distribution and (b, g_1, p_2, g_2) are consistent with assumption A3. In what follows, let $(\beta, \gamma_{1,\text{long}}, \gamma_{2,\text{long}}, \pi_1, \pi_2) = (b, g_1, g_2, p_1, p_2)$. Then $\gamma_{2,\text{long}} \neq 0$ since g_2 was shown to be strictly positive. Similarly, $\gamma_{1,\text{long}} = \gamma_{1,\text{med}} + B\pi_1 \neq 0$ by $\beta_{\text{med}} - B \notin \mathcal{B}_{\text{A3fail}}$. Therefore, A3 is satisfied.

We now show this joint distribution and $(\beta, \gamma_{1,\text{long}}, \gamma_{2,\text{long}}, \pi_1, \pi_2)$ are consistent with the definition of R_{long}^2 and assumption A2.

For R_{long}^2 , we have that

$$\begin{aligned} &\beta^2 \text{var}(X) + \text{var}(\gamma'_{1,\text{long}}W_1) + \text{var}(\gamma_{2,\text{long}}W_2) + 2\beta \text{cov}(X, \gamma'_{1,\text{long}}W_1) \\ &\quad + 2\beta \text{cov}(X, \gamma_{2,\text{long}}W_2) + 2\text{cov}(\gamma'_{1,\text{long}}W_1, \gamma_{2,\text{long}}W_2) \\ &= \beta^2 \text{var}(X) + 2\beta\gamma'_{1,\text{long}} \text{var}(W_1)\pi_1 + 2\beta\pi_2\gamma_{2,\text{long}} + \text{var}(\gamma'_{1,\text{long}}W_1) + \gamma_{2,\text{long}}^2 \\ &= R_{\text{long}}^2 \text{var}(Y) \end{aligned}$$

as shown above in the proof of positive semi-definiteness of Σ .

For δ , using A3 we can write

$$\begin{aligned}
& \text{var}(\gamma'_{1,\text{long}} W_1) \text{var}(\gamma_{2,\text{long}} W_2) \left(\frac{\text{cov}(X, \gamma_{2,\text{long}} W_2)}{\text{var}(\gamma_{2,\text{long}} W_2)} - \delta \cdot \frac{\text{cov}(X, \gamma'_{1,\text{long}} W_1)}{\text{var}(\gamma'_{1,\text{long}} W_1)} \right) \\
&= \pi_2 \gamma_{2,\text{long}} \text{var}(\gamma'_{1,\text{long}} W_1) - \delta \text{cov}(X, \gamma'_{1,\text{long}} W_1) \gamma_{2,\text{long}}^2 \\
&= B \text{var}(X^{\perp W_1}) \text{var}((\gamma_{1,\text{med}} + B\pi_1)' W_1) - \delta \text{cov}(X, (\gamma_{1,\text{med}} + B\pi_1)' W_1) (\gamma_{2,\text{long}}^2) \\
&= B \text{var}(X^{\perp W_1}) (\text{var}(\gamma'_{1,\text{med}} W_1) + 2B \text{cov}(X, \gamma'_{1,\text{med}} W_1) + B^2 \text{var}(\pi'_1 W_1)) \\
&\quad - \delta ((R_{\text{long}}^2 - R_{\text{med}}^2) \text{var}(Y) + B^2 \text{var}(X^{\perp W_1})) (\text{cov}(X, \gamma'_{1,\text{med}} W_1) + B \text{var}(\pi'_1 W_1)) \\
&= -f_0(B) - \delta f_1(B, R_{\text{long}}^2).
\end{aligned}$$

By $f(B, \delta, R_{\text{long}}^2) = f_0(B) + \delta f_1(B, R_{\text{long}}^2) = 0$ we have that

$$\delta \cdot \frac{\text{cov}(X, \gamma'_{1,\text{long}} W_1)}{\text{var}(\gamma'_{1,\text{long}} W_1)} = \frac{\text{cov}(X, \gamma_{2,\text{long}} W_2)}{\text{var}(\gamma_{2,\text{long}} W_2)}$$

and A2 holds.

To finish the sharpness proof, we verify that this joint distribution is consistent with equations (1) and (7).

To verify equation (1), we show that

$$\text{cov} \left(Y - bX - g'_1 W_1 - g_2 W_2, \begin{pmatrix} X \\ W_1 \\ W_2 \end{pmatrix} \right) = 0.$$

To see this, note that

$$\begin{aligned}
\text{cov}(Y - bX - g'_1 W_1 - g_2 W_2, X) &= \text{cov}(Y, X) - b \text{var}(X) - g'_1 \text{var}(W_1) p_1 - g_2 p_2 \\
&= 0 \\
\text{cov}(Y - bX - g'_1 W_1 - g_2 W_2, W_1) &= \text{cov}(Y, W_1) - b \text{cov}(X, W_1) - g'_1 \text{var}(W_1) \\
&= 0
\end{aligned}$$

from equations (12) and (13) above. We can directly show that

$$\begin{aligned}
\text{cov}(Y - bX - g'_1 W_1 - g_2 W_2, W_2) &= b p_2 + g_2 - b p_2 - g_2 \text{var}(W_2) \\
&= 0
\end{aligned}$$

since $\text{var}(W_2) = 1$ by assumption A5.

To verify equation (7), we see that

$$\begin{aligned}
\text{cov} \left(X - p'_1 W_1 - p_2 W_2, \begin{pmatrix} W_1 \\ W_2 \end{pmatrix} \right) &= \begin{pmatrix} \text{cov}(X, W_1) - p'_1 \text{var}(W_1) \\ \text{cov}(X, W_2) - p_2 \text{var}(W_2) \end{pmatrix} \\
&= \begin{pmatrix} p'_1 \text{var}(W_1) - p'_1 \text{var}(W_1) \\ p_2 - p_2 \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ 0 \end{pmatrix}.
\end{aligned}$$

□

Proof of Proposition 2. Let $A \equiv \mathbb{L}(Y \mid 1, X, W_1, W_2)$. Note that $R_{\text{long}}^2 \text{var}(Y) = \text{var}(A)$ and that $R_{\text{med}}^2 \text{var}(Y) = \text{var}(\mathbb{L}(Y \mid 1, X, W_1))$. By iterated projections, $\mathbb{L}(A \mid 1, X, W_1) = \mathbb{L}(Y \mid 1, X, W_1)$. Therefore,

$$\text{var}(A - \mathbb{L}(Y \mid 1, X, W_1))$$

$$\begin{aligned}
&= \text{var}(A) - 2 \text{cov}(A, \mathbb{L}(Y | 1, X, W_1)) + \text{var}(\mathbb{L}(Y | 1, X, W_1)) \\
&= \text{var}(A) - 2 \text{cov}(\mathbb{L}(A | 1, X, W_1) + A^{\perp X, W_1}, \mathbb{L}(Y | 1, X, W_1)) + \text{var}(\mathbb{L}(Y | 1, X, W_1)) \\
&= \text{var}(A) - 2 \text{var}(\mathbb{L}(Y | 1, X, W_1)) + \text{var}(\mathbb{L}(Y | 1, X, W_1)) \\
&= \text{var}(A) - \text{var}(\mathbb{L}(Y | 1, X, W_1)) \\
&= (R_{\text{long}}^2 - R_{\text{med}}^2) \text{var}(Y).
\end{aligned}$$

Therefore, $R_{\text{med}}^2 = R_{\text{long}}^2$ implies that $A = \mathbb{L}(Y | 1, X, W_1)$ with probability 1. Rearranging this yields

$$(\beta_{\text{long}} - \beta_{\text{med}})X + (\gamma_{1,\text{long}} - \gamma_{1,\text{med}})'W_1 + \gamma_{2,\text{long}}W_2 = 0$$

with probability 1. By A1, (X, W_1, W_2) are not perfectly collinear and therefore $\beta_{\text{long}} = \beta_{\text{med}}$. \square

Proof of Corollary 1. When $\delta = 0$,

$$f(B, \delta, R_{\text{long}}^2) = f_0(B) = -B \text{var}(X^{\perp W_1}) \text{var}((\gamma_{1,\text{med}} + B\pi_1)'W_1),$$

a cubic polynomial. It has up to three roots because $\text{var}(X^{\perp W_1}) > 0$ by A1. $B = 0$ is a solution to the cubic equation $f_0(B) = 0$. Other solutions must satisfy $\text{var}((\gamma_{1,\text{med}} + B\pi_1)'W_1) = 0$, or $\gamma_{1,\text{med}} + B\pi_1 = 0$ by $\text{var}(W_1)$ being positive definite. Under assumptions A1–A5 and by Theorem 2, $\beta_{\text{long}} \notin \mathcal{B}_{\text{A3fail}}$ so these other solutions do not correspond to elements of $\mathcal{B}_I(0, R_{\text{long}}^2)$. The solution $B = 0$, corresponding to $\beta_{\text{long}} = \beta_{\text{med}}$, is part of the identified set provided that $\beta_{\text{med}} \notin \mathcal{B}_{\text{A3fail}}$, or that $\gamma_{1,\text{med}} + (\beta_{\text{med}} - \beta_{\text{med}})\pi_1 \neq 0$. This is directly assumed in the statement of the corollary, so $\mathcal{B}_I(0, R_{\text{long}}^2) = \{\beta_{\text{med}}\}$. \square

Proof of Theorem 3. If $\beta_{\text{long}} = \beta_{\text{hypo}}$, then $\beta_{\text{hypo}} \in \mathcal{B}_I(\delta, R_{\text{long}}^2)$. Then we must have

$$f_0(\beta_{\text{med}} - \beta_{\text{hypo}}) + \delta f_1(\beta_{\text{med}} - \beta_{\text{hypo}}, R_{\text{long}}^2) = 0.$$

Because of this, $f_1(\beta_{\text{med}} - \beta_{\text{hypo}}, R_{\text{long}}^2) = 0$ only if $f_0(\beta_{\text{med}} - \beta_{\text{hypo}}) = 0$. As shown in the proof of Corollary 1, this is the case only when $\beta_{\text{hypo}} = \beta_{\text{med}}$, in which case

$$f_1(0, R_{\text{long}}^2) = (R_{\text{long}}^2 - R_{\text{med}}^2) \text{var}(Y) \text{cov}(X, \gamma'_{1,\text{med}}W_1) \neq 0$$

since $\beta_{\text{short}} \neq \beta_{\text{med}}$. Therefore, $f_1(\beta_{\text{med}} - \beta_{\text{hypo}}, R_{\text{long}}^2) \neq 0$, and

$$\delta = \frac{-f_0(\beta_{\text{med}} - \beta_{\text{hypo}})}{f_1(\beta_{\text{med}} - \beta_{\text{hypo}}, R_{\text{long}}^2)}$$

is identified.

Finally, by definition, $\{\delta \in \mathbb{R} : \beta_{\text{hypo}} \in \mathcal{B}_I(\delta, R_{\text{long}}^2)\}$ is a singleton equal to

$$\frac{-f_0(\beta_{\text{med}} - \beta_{\text{hypo}})}{f_1(\beta_{\text{med}} - \beta_{\text{hypo}}, R_{\text{long}}^2)}.$$

Hence the exact value breakdown point equals the absolute value of this singleton. \square

Section 4

Proof of Theorem 4. We start this proof by noting that

$$\text{var}(\pi'_1 W_1) = \pi'_1 \text{var}(W_1) \pi_1 = \text{cov}(X, W_1) \text{var}(W_1)^{-1} \text{cov}(W_1, X) > 0$$

because $\text{var}(W_1)^{-1}$ is positive definite by A1, and because $\text{cov}(X, W_1) \neq 0$ by $\text{cov}(X, \gamma'_{1,\text{med}}W_1) \neq 0$.

We also note that $\pi_1 \neq 0$ implies the set $\mathcal{B}_{\text{A3fail}}$ is a singleton or empty.

Let $B_n \uparrow +\infty$ and

$$\delta_n = \frac{-f_0(B_n)}{f_1(B_n, R_{\text{long}}^2)}.$$

By construction, $\beta_{\text{med}} - B_n \in \mathcal{B}_I(\delta_n, R_{\text{long}}^2) \cup \mathcal{B}_{\text{A3fail}}$. Since $\mathcal{B}_{\text{A3fail}}$ contains at most one element, we let $\beta_{\text{med}} - B_n \in \mathcal{B}_I(\delta_n, R_{\text{long}}^2)$ without loss of generality by removing at most one element from this sequence.

Then

$$\begin{aligned} \delta_n &= \left(B_n^3 \text{var}(\pi_1' W_1) \text{var}(X^{\perp W_1}) + 2B_n^2 \text{var}(X^{\perp W_1}) \text{cov}(X, \gamma_{1,\text{med}}' W_1) + B_n \text{var}(X^{\perp W_1}) \text{var}(\gamma_{1,\text{med}}' W_1) \right) \\ &\quad \Bigg/ \left(B_n^3 \text{var}(X^{\perp W_1}) \text{var}(\pi_1' W_1) + B_n^2 \text{var}(X^{\perp W_1}) \text{cov}(X, \gamma_{1,\text{med}}' W_1) \right. \\ &\quad \left. + B_n(R_{\text{long}}^2 - R_{\text{med}}^2) \text{var}(Y) \text{var}(\pi_1' W_1) + (R_{\text{long}}^2 - R_{\text{med}}^2) \text{cov}(X, W_1' \gamma_{1,\text{med}}) \text{var}(Y) \right) \\ &\rightarrow 1 \end{aligned}$$

as $n \rightarrow \infty$ since $\text{var}(X^{\perp W_1}) \text{var}(\pi_1' W_1) > 0$ multiplies the cubic term in both the numerator and denominator. Note that $\text{var}(X^{\perp W_1}) > 0$ by A1. Similarly, we have that

$$\tilde{\delta}_n \equiv \frac{-f_0(-B_n)}{f_1(-B_n, R_{\text{long}}^2)} \rightarrow 1$$

as $n \rightarrow \infty$. Note that $\beta_{\text{med}} + B_n \in \mathcal{B}_I(\tilde{\delta}_n, R_{\text{long}}^2)$ by construction and by the same argument used to show $\beta_{\text{med}} - B_n \notin \mathcal{B}_{\text{A3fail}}$.

Moreover, for n sufficiently large we have that

$$\text{sign}(\delta_n - 1) = \text{sign}(\text{cov}(X, \gamma_{1,\text{med}}' W_1)).$$

We can see this from

$$\begin{aligned} \delta_n - 1 &= f_1(B_n, R_{\text{long}}^2)^{-1} (-f_0(B_n) - f_1(B_n, R_{\text{long}}^2)) \\ &= f_1(B_n, R_{\text{long}}^2)^{-1} \left(B_n^3 \text{var}(\pi_1' W_1) \text{var}(X^{\perp W_1}) + 2B_n^2 \text{var}(X^{\perp W_1}) \text{cov}(X, \gamma_{1,\text{med}}' W_1) \right. \\ &\quad \left. + B_n \text{var}(X^{\perp W_1}) \text{var}(\gamma_{1,\text{med}}' W_1) \right) \\ &\quad - f_1(B_n, R_{\text{long}}^2)^{-1} \left(B_n^3 \text{var}(X^{\perp W_1}) \text{var}(\pi_1' W_1) + B_n^2 \text{var}(X^{\perp W_1}) \text{cov}(X, \gamma_{1,\text{med}}' W_1) \right. \\ &\quad \left. + B_n(R_{\text{long}}^2 - R_{\text{med}}^2) \text{var}(Y) \text{var}(\pi_1' W_1) + (R_{\text{long}}^2 - R_{\text{med}}^2) \text{cov}(X, W_1' \gamma_{1,\text{med}}) \text{var}(Y) \right) \\ &= f_1(B_n, R_{\text{long}}^2)^{-1} B_n^2 \left(\text{var}(X^{\perp W_1}) \text{cov}(X, \gamma_{1,\text{med}}' W_1) + O(B_n^{-1}) \right), \end{aligned}$$

and from $\text{var}(X^{\perp W_1}) > 0$, $B_n^2 > 0$, $B_n^{-1} \rightarrow 0$, and $f_1(B_n, R_{\text{long}}^2) > 0$ as $n \rightarrow \infty$.

Likewise, for n sufficiently large we have that

$$\text{sign}(\tilde{\delta}_n - 1) = -\text{sign}(\text{cov}(X, \gamma_{1,\text{med}}' W_1))$$

since

$$-(f_0(-B_n) + f_1(-B_n, R_{\text{long}}^2)) = B_n^2 (\text{var}(X^{\perp W_1}) \text{cov}(X, \gamma_{1,\text{med}}' W_1) + O(B_n^{-1})),$$

and since $f_1(-B_n, R_{\text{long}}^2) < 0$ as $n \rightarrow \infty$.

Therefore, if $\text{cov}(X, \gamma_{1,\text{med}}' W_1) > 0$, for sufficiently large n the sequence δ_n is such that $\delta_n > 1$, $\delta_n \rightarrow 1$, and

$$\min \mathcal{B}_I(\delta_n, R_{\text{long}}^2) \leq \beta_{\text{med}} - B_n \rightarrow -\infty$$

as $n \rightarrow \infty$. Also let N be such that $\beta_{\text{med}} - B_n < 0$ and for all $n \geq N$.

Likewise, for sufficiently large n the sequence $\tilde{\delta}_n$ is such that $\tilde{\delta}_n < 1$, $\tilde{\delta}_n \rightarrow 1$, and

$$\max \mathcal{B}_I(\tilde{\delta}_n, R_{\text{long}}^2) \geq \beta_{\text{med}} + B_n \rightarrow +\infty$$

as $n \rightarrow \infty$. Also let \tilde{N} be such that $\beta_{\text{med}} + B_n > 0$ and $\tilde{\delta}_n \leq 1$ for all $n \geq \tilde{N}$.

If $\beta_{\text{med}} > 0$, then

$$\begin{aligned} \delta^{\text{bp,sign}}(R_{\text{long}}^2) &= \delta^{\text{bp},>}(R_{\text{long}}^2) \\ &= \sup\{|\delta| : \delta \in \mathbb{R}, b > 0 \text{ for all } b \in \mathcal{B}_I(\delta, R_{\text{long}}^2)\} \\ &\leq \delta_n \end{aligned}$$

for any $n \geq \tilde{N}$. The inequality follows from $\mathcal{B}_I(\delta_n, R_{\text{long}}^2)$ containing a negative number for all $n \geq N$. Taking $n \rightarrow \infty$ on both sides of the inequality yields $\delta^{\text{bp,sign}}(R_{\text{long}}^2) \leq 1$.

If $\beta_{\text{med}} < 0$, then

$$\begin{aligned} \delta^{\text{bp,sign}}(R_{\text{long}}^2) &= \delta^{\text{bp},<}(R_{\text{long}}^2) \\ &= \sup\{|\delta| : \delta \in \mathbb{R}, b < 0 \text{ for all } b \in \mathcal{B}_I(\delta, R_{\text{long}}^2)\} \\ &\leq \delta_{\tilde{N}} \\ &\leq 1, \end{aligned}$$

which follows from $\mathcal{B}(\delta_{\tilde{N}}, R_{\text{long}}^2)$ containing a positive value, and from $\tilde{\delta}_{\tilde{N}} \leq 1$ by its definition. An analogous argument shows that this conclusion holds if $\text{cov}(X, \gamma'_{1,\text{med}} W_1) < 0$ as well. \square

Proof of Corollary 2. By Theorem 2, $\beta_{\text{long}} \in \mathcal{B}_I(\delta, R_{\text{long}}^2)$. By A6, $\beta_{\text{long}} \in [\beta_{\text{med}} - M, \beta_{\text{med}} + M]$. Combining these yields that β_{long} lies in the intersection of these sets. Sharpness follows from the sharpness of $\mathcal{B}_I(\delta, R_{\text{long}}^2)$, which was established in Theorem 2, and from $\beta_{\text{long}} \in [\beta_{\text{med}} - M, \beta_{\text{med}} + M]$, which implies A6. \square

Proof of Corollary 3. This follows directly from Theorem 2 and Corollary 2. \square

Section 5

Proof of Proposition 3.

1. (\Rightarrow) As shown in Lemma 1, $\gamma_{1,\text{long}} = \gamma_{1,\text{med}} + (\beta_{\text{med}} - \beta_{\text{long}})\pi_1$ under assumptions A1 and A4. Therefore, if A8 holds, then $\gamma_{1,\text{med}} = \gamma_{1,\text{long}} - (\beta_{\text{med}} - \beta_{\text{long}})\pi_1 = (C - (\beta_{\text{med}} - \beta_{\text{long}}))\pi_1 = C_{\text{med}} \cdot \pi_1$ with $C_{\text{med}} = C - (\beta_{\text{med}} - \beta_{\text{long}})$.
2. (\Leftarrow) If there is a constant $C_{\text{med}} \in \mathbb{R}$ such that $\gamma_{1,\text{med}} = C_{\text{med}} \cdot \pi_1$, then we immediately see that $\gamma_{1,\text{long}} = (C_{\text{med}} + (\beta_{\text{med}} - \beta_{\text{long}}))\pi_1$ so A8 holds with $C = C_{\text{med}} + (\beta_{\text{med}} - \beta_{\text{long}})$. \square

Proof of Proposition 4. From the expression for $f(B, \delta, R_{\text{long}}^2)$, we note that $f(B, 1, R_{\text{long}}^2)$ is quadratic in B and is equal to

$$\begin{aligned} f(B, 1, R_{\text{long}}^2) &= f_0(B) + f_1(B, R_{\text{long}}^2) \\ &= -B^2 \text{var}(X^{\perp W_1}) \text{cov}(X, \gamma'_{1,\text{med}} W_1) + B((R_{\text{long}}^2 - R_{\text{med}}^2) \text{var}(Y) \text{var}(\pi'_1 W_1) - \text{var}(X^{\perp W_1}) \text{var}(\gamma'_{1,\text{med}} W_1)) \\ &\quad + (R_{\text{long}}^2 - R_{\text{med}}^2) \text{cov}(X, W'_1 \gamma_{1,\text{med}}) \text{var}(Y). \end{aligned}$$

The discriminant of this quadratic polynomial is

$$\Delta \equiv ((R_{\text{long}}^2 - R_{\text{med}}^2) \text{var}(Y) \text{var}(\pi'_1 W_1) - \text{var}(X^{\perp W_1}) \text{var}(\gamma'_{1,\text{med}} W_1))^2$$

$$\begin{aligned}
& + 4 \operatorname{var}(X^{\perp W_1}) \operatorname{cov}(X, \gamma'_{1,\text{med}} W_1)^2 (R_{\text{long}}^2 - R_{\text{med}}^2) \operatorname{var}(Y) \\
& = ((R_{\text{long}}^2 - R_{\text{med}}^2) \operatorname{var}(Y) \operatorname{var}(\pi'_1 W_1))^2 + (\operatorname{var}(X^{\perp W_1}) \operatorname{var}(\gamma'_{1,\text{med}} W_1))^2 \\
& \quad - 2(R_{\text{long}}^2 - R_{\text{med}}^2) \operatorname{var}(Y) \operatorname{var}(\pi'_1 W_1) \operatorname{var}(X^{\perp W_1}) \operatorname{var}(\gamma'_{1,\text{med}} W_1) \\
& \quad + 4 \operatorname{var}(X^{\perp W_1}) \operatorname{cov}(X, \gamma'_{1,\text{med}} W_1)^2 (R_{\text{long}}^2 - R_{\text{med}}^2) \operatorname{var}(Y).
\end{aligned}$$

To simplify this expression, we show that $\operatorname{cov}(X, \gamma'_{1,\text{med}} W_1)^2 = \operatorname{var}(\pi'_1 W_1) \operatorname{var}(\gamma'_{1,\text{med}} W_1)$. To see this, note that A8 and Proposition 3 implies that $\gamma_{1,\text{med}} = C_{\text{med}} \pi_1$. Therefore

$$\begin{aligned}
\operatorname{cov}(X, \gamma'_{1,\text{med}} W_1)^2 - \operatorname{var}(\pi'_1 W_1) \operatorname{var}(\gamma'_{1,\text{med}} W_1) &= C_{\text{med}}^2 \operatorname{cov}(X, \pi'_1 W_1)^2 - C_{\text{med}}^2 \operatorname{var}(\pi'_1 W_1) \operatorname{var}(\pi'_1 W_1) \\
&= C_{\text{med}}^2 (\operatorname{cov}(\pi'_1 W_1 + \pi_2 W_2 + X^{\perp W}, \pi'_1 W_1)^2 - \operatorname{var}(\pi'_1 W_1)^2) \\
&= C_{\text{med}}^2 (\operatorname{var}(\pi'_1 W_1)^2 - \operatorname{var}(\pi'_1 W_1)^2) \\
&= 0.
\end{aligned}$$

The fourth equality follows from A4. Therefore,

$$\operatorname{cov}(X, \gamma'_{1,\text{med}} W_1)^2 = \operatorname{var}(\pi'_1 W_1) \operatorname{var}(\gamma'_{1,\text{med}} W_1) \quad (14)$$

which implies that the discriminant can be written as

$$\begin{aligned}
\Delta &= ((R_{\text{long}}^2 - R_{\text{med}}^2) \operatorname{var}(Y) \operatorname{var}(\pi'_1 W_1))^2 + (\operatorname{var}(X^{\perp W_1}) \operatorname{var}(\gamma'_{1,\text{med}} W_1))^2 \\
&\quad + 2(R_{\text{long}}^2 - R_{\text{med}}^2) \operatorname{var}(Y) \operatorname{var}(\pi'_1 W_1) \operatorname{var}(X^{\perp W_1}) \operatorname{var}(\gamma'_{1,\text{med}} W_1) \\
&= ((R_{\text{long}}^2 - R_{\text{med}}^2) \operatorname{var}(Y) \operatorname{var}(\pi'_1 W_1) + \operatorname{var}(X^{\perp W_1}) \operatorname{var}(\gamma'_{1,\text{med}} W_1))^2.
\end{aligned}$$

Therefore, the quadratic equation $f(B, 1, R_{\text{long}}^2) = 0$ has the following two solutions

$$\begin{aligned}
& \left((R_{\text{long}}^2 - R_{\text{med}}^2) \operatorname{var}(Y) \operatorname{var}(\pi'_1 W_1) - \operatorname{var}(X^{\perp W_1}) \operatorname{var}(\gamma'_{1,\text{med}} W_1) \right. \\
& \quad \left. \pm \sqrt{((R_{\text{long}}^2 - R_{\text{med}}^2) \operatorname{var}(Y) \operatorname{var}(\pi'_1 W_1) + \operatorname{var}(X^{\perp W_1}) \operatorname{var}(\gamma'_{1,\text{med}} W_1))^2} \right) / 2 \operatorname{var}(X^{\perp W_1}) \operatorname{cov}(X, \gamma'_{1,\text{med}} W_1) \\
&= \left((R_{\text{long}}^2 - R_{\text{med}}^2) \operatorname{var}(Y) \operatorname{var}(\pi'_1 W_1) - \operatorname{var}(X^{\perp W_1}) \operatorname{var}(\gamma'_{1,\text{med}} W_1) \right. \\
& \quad \left. \pm ((R_{\text{long}}^2 - R_{\text{med}}^2) \operatorname{var}(Y) \operatorname{var}(\pi'_1 W_1) + \operatorname{var}(X^{\perp W_1}) \operatorname{var}(\gamma'_{1,\text{med}} W_1)) \right) / 2 \operatorname{var}(X^{\perp W_1}) \operatorname{cov}(X, \gamma'_{1,\text{med}} W_1) \\
&= \{B_1, B_2\}
\end{aligned}$$

where

$$B_1 = \frac{(R_{\text{long}}^2 - R_{\text{med}}^2) \operatorname{var}(Y) \operatorname{var}(\pi'_1 W_1)}{\operatorname{var}(X^{\perp W_1}) \operatorname{cov}(X, \gamma'_{1,\text{med}} W_1)} \quad \text{and} \quad B_2 = -\frac{\operatorname{var}(\gamma'_{1,\text{med}} W_1)}{\operatorname{cov}(X, \gamma'_{1,\text{med}} W_1)}.$$

These are well defined by $\operatorname{cov}(X, \gamma'_{1,\text{med}} W_1) = \operatorname{var}(X)(\beta_{\text{short}} - \beta_{\text{med}}) \neq 0$. They are distinct by $\Delta > 0$. We now simplify the expressions for B_1 and B_2 . Combining equations (10), (11), and (14), we can write

$$\begin{aligned}
\operatorname{var}(\gamma'_{1,\text{med}} W_1) &= (R_{\text{med}}^2 - R_{\text{short}}^2) \operatorname{var}(Y) + \operatorname{var}(X)(\beta_{\text{short}} - \beta_{\text{med}})^2 \\
&= (R_{\text{med}}^2 - R_{\text{short}}^2) \operatorname{var}(Y) + \frac{\operatorname{cov}(X, \gamma'_{1,\text{med}} W_1)^2}{\operatorname{var}(X)} \\
&= (R_{\text{med}}^2 - R_{\text{short}}^2) \operatorname{var}(Y) + \frac{\operatorname{var}(\pi'_1 W_1) \operatorname{var}(\gamma'_{1,\text{med}} W_1)}{\operatorname{var}(X)}.
\end{aligned}$$

Solving the previous equation for $\text{var}(\gamma'_{1,\text{med}}W_1)$ yields

$$\begin{aligned}\text{var}(\gamma'_{1,\text{med}}W_1) &= (R_{\text{med}}^2 - R_{\text{short}}^2) \text{var}(Y) \left(1 - \frac{\text{var}(\pi'_1 W_1)}{\text{var}(X)}\right)^{-1} \\ &= (R_{\text{med}}^2 - R_{\text{short}}^2) \text{var}(Y) \frac{\text{var}(X)}{\text{var}(X) - \text{var}(\pi'_1 W_1)} \\ &= (R_{\text{med}}^2 - R_{\text{short}}^2) \text{var}(Y) \frac{\text{var}(X)}{\text{var}(X^\perp W_1)}.\end{aligned}$$

This equality can be used to simplify B_1 as follows:

$$\begin{aligned}B_1 &= \frac{(R_{\text{long}}^2 - R_{\text{med}}^2) \text{var}(Y) \text{var}(\pi'_1 W_1)}{\text{var}(X^\perp W_1) \text{cov}(X, \gamma'_{1,\text{med}}W_1)} \\ &= \frac{(R_{\text{long}}^2 - R_{\text{med}}^2) \text{var}(Y) \text{cov}(X, \gamma'_{1,\text{med}}W_1)}{\text{var}(X^\perp W_1) \text{var}(\gamma'_{1,\text{med}}W_1)} \\ &= \frac{(R_{\text{long}}^2 - R_{\text{med}}^2) \text{var}(Y) (\beta_{\text{short}} - \beta_{\text{med}}) \text{var}(X)}{\text{var}(X^\perp W_1) (R_{\text{med}}^2 - R_{\text{short}}^2) \text{var}(Y) \frac{\text{var}(X)}{\text{var}(X^\perp W_1)}} \\ &= (\beta_{\text{short}} - \beta_{\text{med}}) \frac{R_{\text{long}}^2 - R_{\text{med}}^2}{R_{\text{med}}^2 - R_{\text{short}}^2}.\end{aligned}$$

The second equality follows from equation (14), and the third follows from our expression for $\text{var}(\gamma'_{1,\text{med}}W_1)$.

Meanwhile, we can show that $\beta_{\text{med}} - B_2 \in \mathcal{B}_{\text{A3fail}}$ and thus $\beta_{\text{med}} - B_2 \notin \mathcal{B}_I(1, R_{\text{long}}^2)$. To see this, note that

$$\begin{aligned}&\text{var}((\gamma_{1,\text{med}} + B_2 \pi_1)' W_1) \\ &= \text{var}\left(\gamma'_{1,\text{med}}W_1 - \frac{\text{var}(\gamma'_{1,\text{med}}W_1)}{\text{cov}(X, \gamma'_{1,\text{med}}W_1)} \pi'_1 W_1\right) \\ &= \text{var}(\gamma'_{1,\text{med}}W_1) - 2 \frac{\text{var}(\gamma'_{1,\text{med}}W_1)}{\text{cov}(X, \gamma'_{1,\text{med}}W_1)} \text{cov}(\gamma'_{1,\text{med}}W_1, \pi'_1 W_1) + \frac{\text{var}(\gamma'_{1,\text{med}}W_1)^2}{\text{cov}(X, \gamma'_{1,\text{med}}W_1)^2} \text{var}(\pi'_1 W_1) \\ &= \text{var}(\gamma'_{1,\text{med}}W_1) - 2 \frac{\text{var}(\gamma'_{1,\text{med}}W_1)}{\text{cov}(X, \gamma'_{1,\text{med}}W_1)} \text{cov}(\gamma'_{1,\text{med}}W_1, X) + \frac{\text{var}(\gamma'_{1,\text{med}}W_1)^2}{\text{cov}(X, \gamma'_{1,\text{med}}W_1)^2} \text{var}(\pi'_1 W_1) \\ &= \text{var}(\gamma'_{1,\text{med}}W_1) - 2 \text{var}(\gamma'_{1,\text{med}}W_1) + \frac{\text{var}(\gamma'_{1,\text{med}}W_1)^2 \text{var}(\pi'_1 W_1)}{\text{var}(\gamma'_{1,\text{med}}W_1) \text{var}(\pi'_1 W_1)} \\ &= 0.\end{aligned}$$

The fourth equality follows from equation (14) and from $\text{cov}(X, \gamma'_{1,\text{med}}W_1) \neq 0$. Therefore, $\gamma_{1,\text{med}} + B_2 \pi_1 = 0$, which implies that $\beta_{\text{med}} - B_2 \notin \mathcal{B}_I(1, R_{\text{long}}^2)$. Since $\pi_1 \neq 0$ and $\gamma_{1,\text{med}} \neq 0$, which follows from $\beta_{\text{short}} \neq \beta_{\text{med}}$, the set $\mathcal{B}_{\text{A3fail}}$ is a singleton. Since B_1 and B_2 are distinct, $\beta_{\text{med}} - B_1 \notin \mathcal{B}_{\text{A3fail}} = \{\beta_{\text{med}} - B_2\}$. By Theorem 2, we conclude that $\beta_{\text{med}} - B_1$ is the unique element of the identified set. \square

B List of Papers Used in Section 7 Survey

ABALUCK, J., M. CACERES BRAVO, P. HULL, AND A. STARC (2021): “Mortality effects and choice across private health insurance plans,” *Quarterly Journal of Economics*, 136, 1557–1610.

ABDULKADIROGLU, A., P. A. PATHAK, J. SCHELLENBERG, AND C. R. WALTERS (2020): “Do parents value school effectiveness?” *American Economic Review*, 110, 1502–39.

- ALLCOTT, H., R. DIAMOND, J.-P. DUBÉ, J. HANDBURY, I. RAHKOVSKY, AND M. SCHNELL (2019a): “Food deserts and the causes of nutritional inequality,” *Quarterly Journal of Economics*, 134, 1793–1844.
- ALLCOTT, H., B. B. LOCKWOOD, AND D. TAUBINSKY (2019b): “Regressive sin taxes, with an application to the optimal soda tax,” *Quarterly Journal of Economics*, 134, 1557–1626.
- ARBATLI, C. E., Q. H. ASHRAF, O. GALOR, AND M. KLEMP (2020): “Diversity and conflict,” *Econometrica*, 88, 727–797.
- BAZZI, S., M. FISZBEIN, AND M. GEBRESILASSE (2020): “Frontier culture: The roots and persistence of “rugged individualism” in the United States,” *Econometrica*, 88, 2329–2368.
- BERTRAND, M., R. BURGESS, A. CHAWLA, AND G. XU (2020): “The glittering prizes: Career incentives and bureaucrat performance,” *Review of Economic Studies*, 87, 626–655.
- BLADER, S., C. GARTENBERG, AND A. PRAT (2020): “The contingent effect of management practices,” *Review of Economic Studies*, 87, 721–749.
- BLOUIN, A. AND R. MACCHIAVELLO (2019): “Strategic default in the international coffee market,” *Quarterly Journal of Economics*, 134, 895–951.
- CHEN, Y. (2021): “Team-specific human capital and team performance: Evidence from doctors,” *American Economic Review*, 111, 3923–62.
- COUTTENIER, M., V. PETRENCU, D. ROHNER, AND M. THOENIG (2019): “The violent legacy of conflict: evidence on asylum seekers, crime, and public policy in Switzerland,” *American Economic Review*, 109, 4378–4425.
- DIPPEL, C. AND S. HEBLICH (2021): “Leadership in Social Movements: Evidence from the” Forty-Eighters” in the Civil War,” *American Economic Review*, 111, 472–505.
- ENKE, B. (2019): “Kinship, cooperation, and the evolution of moral systems,” *Quarterly Journal of Economics*, 134, 953–1019.
- (2020): “Moral values and voting,” *Journal of Political Economy*, 128, 3679–3729.
- EUGSTER, B. AND R. PARCHET (2019): “Culture and taxes,” *Journal of Political Economy*, 127, 296–337.
- FINKELSTEIN, A., M. GENTZKOW, AND H. WILLIAMS (2021): “Place-based drivers of mortality: Evidence from migration,” *American Economic Review*, 111, 2697–2735.
- GAVAZZA, A., M. NARDOTTO, AND T. VALLETTI (2019): “Internet and politics: Evidence from UK local elections and local government policies,” *Review of Economic Studies*, 86, 2092–2135.
- GOLDSMITH-PINKHAM, P., I. SORKIN, AND H. SWIFT (2020): “Bartik instruments: What, when, why, and how,” *American Economic Review*, 110, 2586–2624.
- GREGG, A. G. (2020): “Factory productivity and the concession system of incorporation in late imperial Russia, 1894–1908,” *American Economic Review*, 110, 401–27.
- GROSFELD, I., S. O. SAKALLI, AND E. ZHURAVSKAYA (2020): “Middleman minorities and ethnic violence: anti-Jewish pogroms in the Russian empire,” *Review of Economic Studies*, 87, 289–342.
- GROSJEAN, P. AND R. KHATTAR (2019): “It’s raining men! Hallelujah? The long-run consequences of male-biased sex ratios,” *Review of Economic Studies*, 86, 723–754.
- GURIEV, S., N. MELNIKOV, AND E. ZHURAVSKAYA (2021): “3G internet and confidence in government,” *Quarterly Journal of Economics*, 136, 2533–2613.
- HELDLING, L. (2021): “The origins of violence in Rwanda,” *Review of Economic Studies*, 88, 730–763.

- HOFFMAN, M. AND S. TADELIS (2021): “People management skills, employee attrition, and manager rewards: An empirical analysis,” *Journal of Political Economy*, 129, 243–285.
- MURPHY, R. AND F. WEINHARDT (2020): “Top of the class: The importance of ordinal rank,” *Review of Economic Studies*, 87, 2777–2826.
- SCHINDLER, D. AND M. WESTCOTT (2021): “Shocking racial attitudes: Black GIs in Europe,” *Review of Economic Studies*, 88, 489–520.
- SQUICCIARINI, M. P. (2020): “Devotion and development: Religiosity, education, and economic progress in nineteenth-century France,” *American Economic Review*, 110, 3454–91.
- TABELLINI, M. (2020): “Gifts of the immigrants, woes of the natives: Lessons from the age of mass migration,” *Review of Economic Studies*, 87, 454–486.
- VALENCIA CAICEDO, F. (2019): “The mission: Human capital transmission, economic persistence, and culture in South America,” *Quarterly Journal of Economics*, 134, 507–556.
- VERNER, E. AND G. GYÖNGYÖSI (2020): “Household debt revaluation and the real economy: Evidence from a foreign currency debt crisis,” *American Economic Review*, 110, 2667–2702.
- WEIGEL, J. L. (2020): “The participation dividend of taxation: How citizens in Congo engage more with the state when it tries to tax them,” *Quarterly Journal of Economics*, 135, 1849–1903.