SHARPENING THE ROSENBAUM SENSITIVITY BOUNDS TO ADDRESS CONCERNS ABOUT INTERACTIONS BETWEEN OBSERVED AND UNOBSERVED COVARIATES

Siyu Heng and Dylan S. Small

University of Pennsylvania

Abstract: In observational studies, it is typically unrealistic to assume that treatments are assigned randomly, even conditional on adjusting for all observed covariates. Therefore, a sensitivity analysis is often needed to examine how hidden biases due to unobserved covariates affect inferences on treatment effects. In matched observational studies, where each treated unit is matched to one or multiple untreated controls for observed covariates, the Rosenbaum bounds sensitivity analysis is one of the most popular sensitivity analysis models. We show that in the presence of interactions between observed and unobserved covariates, directly applying the Rosenbaum bounds almost inevitably exaggerates the report of sensitivity of causal conclusions to hidden bias. We give sharper odds ratio bounds to fix this deficiency. We illustrate our new method by studying the effect of a anger/hostility tendency on the risk of experiencing heart problems.

Key words and phrases: Causal inference, gene-environment interaction, interaction terms, matching, observational studies, sensitivity analysis.

1. Introduction

In a randomized experiment, units are assigned randomly to the treatment or control groups. In contrast, in an observational study, treatments are not assigned randomly to units. As such, differences between the outcomes of the treated and control groups can be a biased estimate of the true treatment effect because of baseline differences between the two groups. Baseline differences that can be captured by observed covariates can often be removed by model-based adjustments or matching. Among these methods, matching has been used extensively as a nonparametric way of adjusting for the observed covariates in observational studies: each treated unit is matched to one or several controls (i.e., untreated units) on baseline observed covariates such that the treated units and controls are similar in measured confounders, as they would be under a randomized experiment, and the comparisons are made within these matched sets (Rubin (1973);

Corresponding author: Dylan S. Small, Department of Statistics, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104, USA. E-mail: dsmall@wharton.upenn.edu.

Rosenbaum (2002, 2010); Hansen (2004); Stuart (2010); Zubizarreta, Cerdá and Rosenbaum (2013); Pimentel et al. (2015); Zubizarreta and Keele (2017)).

However, some important baseline covariates may be unobserved, which means the treatment assignments may not be random within each matched set. A sensitivity analysis asks how a departure from the random assignment of a treatment affects the causal conclusion drawn from a primary analysis that assumes the treatment is assigned randomly, conditional on the observed covariates. Among various sensitivity analysis models in matched observational studies, the Rosenbaum bounds sensitivity analysis (Rosenbaum (1987, 2002)) is one of the most popular. The analysis introduces a uniform sensitivity parameter $\Gamma \geq 1$ bounding the ratio of the odds of treatment within each matched set: the more Γ departs from one, the more the treatment assignment potentially departs from random assignment in each matched set. Then, researchers typically consider the "worst-case" p-value, which is defined as the largest p-value, given the sensitivity parameter Γ, over all possible arrangements of unobserved covariates (i.e., unmeasured confounders) (Rosenbaum (2002)). For examples of studies that use the Rosenbaum bounds sensitivity analysis, see Normand et al. (2001), Rosenbaum (2002), Rosenbaum (2004), Heller, Rosenbaum and Small (2009), Silber et al. (2009), Stuart and Hanna (2013), Zubizarreta, Cerdá and Rosenbaum (2013), Hsu et al. (2015), Zubizarreta et al. (2016), Ertefaie, Small and Rosenbaum (2018), Karmakar, French and Small (2019), Zhao (2019), Cohen, Olson and Fogarty (2020), Fogarty (2020), and Zhang et al. (2021). Many other sensitivity analysis models also build on the Rosenbaum bounds sensitivity analysis (e.g., Gastwirth, Krieger and Rosenbaum, 1998; Ichino, Mealli and Nannicini, 2008; Rosenbaum and Silber, 2009; Nattino and Lu, 2018; Fogarty and Hasegawa, 2019).

Here, we show that in the presence of any interactions between the observed and unobserved covariates in a logit model of the treatment assignment probability, the Rosenbaum bounds are almost inevitably loose for some of the matched sets. Interactions between observed and unobserved covariates commonly exist in observational studies. One such setting is the extensively studied "gene–environment interaction" (G×E), where two different genotypes respond to environmental variations in different ways (Ottman, 1996; Caspi et al., 2002). In many studies, such genotypes are not identified, measured, or publicly available, and should be considered as unobserved covariates that interact with observed environmental covariates (Pérusse and Bouchard (1999)). Directly applying the Rosenbaum bounds sensitivity analysis in such settings can greatly exaggerate the sensitivity of the causal conclusion to hidden bias. To perform a more informative and less conservative sensitivity analysis in matched studies, we give sharper odds

ratio bounds when there is concern about a possible interaction between an observed covariate and an unobserved covariate. We apply our new result to study the causal effect of anger/hostility tendency on the risk of experiencing heart disease. The data used for this work are publicly available at `https://www.ssc.wisc.edu/wlsresearch/`. The code and codebook used for this work are available at `https://github.com/siyuheng/Sharpening-the-Rosenbaum-Bounds`.

2. A Brief Review of the Rosenbaum Bounds Sensitivity Analysis

We briefly review the classical framework for the Rosenbaum bounds sensitivity analysis for a matched observational study in which each treated unit is matched to one or more controls (Rosenbaum (2002)). There are I matched sets $i = 1, \ldots, I$, and matched set i contains n_i $(n_i \geq 2)$ units, yielding $N = \sum_{i=1}^{I} n_i$ units in total. In each matched set, one unit receives the treatment and the others receive the control. Let $Z_{ij} = 1$ if unit j in matched set i received the treatment, otherwise let $Z_{ij} = 0$. Therefore, we have $\sum_{j=1}^{n_i} Z_{ij} = 1$, for all i. Let $\mathbf{x}_{ij} = (x_{ij(1)}, \ldots, x_{ij(K)})^T$ denote the K observed covariates, and let u_{ij} be an unobserved covariate of unit j in matched set i. The sets are matched for the observed covariates, but not for the unobserved covariate. Therefore, $\mathbf{x}_{ij} = \mathbf{x}_{ij'}$ for all i, j, and j', but possibly $u_{ij} \neq u_{ij'}$ if $j \neq j'$ (Rosenbaum (2002)). Denote the common observed covariates for units in matched sets i as $\mathbf{x}_i = (x_{i(1)}, \ldots, x_{i(K)})^T$, where $\mathbf{x}_i = \mathbf{x}_{ij} = \mathbf{x}_{ij'}$, for all i, j, j'. Under the potential outcomes framework, if unit j in matched set i received the treatment (i.e., $Z_{ij} = 1$), we observe the potential outcome r_{Tij}; otherwise (i.e., $Z_{ij} = 0$), we observe the potential outcome r_{Cij} (Neyman (1990); Rubin (1974)). Therefore, the observed outcome for each i, j is $R_{ij} = Z_{ij} r_{Tij} + (1 - Z_{ij}) r_{Cij}$. Denote the collection of observed outcomes as $\mathbf{R} = (R_{11}, \ldots, R_{In_I})^T$ and the collection of unobserved covariates as $\mathbf{u} = (u_{11}, \ldots, u_{In_I})^T$. Write $\mathcal{F} = \{(r_{Tij}, r_{Cij}, \mathbf{x}_{ij}, u_{ij}), \ i = 1, \ldots, I, \ j = 1, \ldots, n_i\}$, and let $\mathcal{Z}$ be the set of all possible values of $\mathbf{Z} = (Z_{11}, \ldots, Z_{In_I})^T$, where $\mathbf{Z} \in \mathcal{Z}$ if and only if $\sum_{j=1}^{n_i} Z_{ij} = 1$, for all i. Let $|A|$ denote the number of elements of a finite set A, and define the indicator function $\mathbf{1}\{A\} = 1$ if A is true, and $\mathbf{1}\{A\} = 0$ otherwise. Let $a \gg b$ denote that a is much greater than b.

Fisher's sharp null hypothesis of no treatment effect asserts that $H_0 : r_{Tij} = r_{Cij}$, for all i, j. In a randomized experiment, we can assume that $\mathrm{pr}(\mathbf{Z} = \mathbf{z} \mid \mathcal{F}, \mathcal{Z}) = 1/|\mathcal{Z}| = 1/(\prod_{i=1}^{I} n_i)$ for all $\mathbf{z} \in \mathcal{Z}$. In a stratified randomized experiment, with one unit in each matched set being assigned randomly to the treatment, the significance level of a test statistic T being greater than or equal

to the observed value t can be computed as

$$\mathrm{pr}(T \geq t \mid \mathcal{F}, \mathcal{Z}) = \sum_{\mathbf{z} \in \mathcal{Z}} \mathbf{1}\{T(\mathbf{z}, \mathbf{R}) \geq t\} \, \mathrm{pr}(\mathbf{Z} = \mathbf{z} \mid \mathcal{F}, \mathcal{Z})$$
$$= \frac{|\{\mathbf{z} \in \mathcal{Z} : T(\mathbf{z}, \mathbf{R}) \geq t\}|}{|\mathcal{Z}|}.$$

In an observational study, it is often unrealistic to assume that a treatment is assigned randomly, even within a matched set of units with the same observed covariates, owing to the possible presence of an unobserved covariate. A sensitivity analysis is therefore needed to determine how departures from a random assignment of the treatment affect the causal conclusions drawn from a primary analysis that assumes the treatment is assigned randomly within each matched set. Let $\pi_{ij} = P(Z_{ij} = 1 \mid \mathcal{F})$ denote the probability that unit j in matched set i will receive the treatment. The Rosenbaum bounds sensitivity analysis considers that two units ij and ij' in the same matched set i, with the same observed covariates $\mathbf{x}_{ij} = \mathbf{x}_{ij'} = \mathbf{x}_i$, may differ in their odds of receiving the treatment by at most a factor of $\Gamma \geq 1$:

$$\Gamma^{-1} \leq \frac{\pi_{ij}(1 - \pi_{ij'})}{\pi_{ij'}(1 - \pi_{ij})} \leq \Gamma, \quad \text{for all } i \in \{1, \dots, I\} \text{ and } j, j' \in \{1, \dots, n_i\}. \quad (2.1)$$

Constraint (2.1) is also known as the Rosenbaum bounds (DiPrete and Gangl (2004)). It is clear that the more Γ departs from one, the more the treatment assignment potentially departs from a random assignment. In the Rosenbaum bounds sensitivity analysis, researchers are interested in the "worst-case" (i.e., the largest possible) p-value reported by a test statistic T, given its observed value t under constraint (2.1) (Rosenbaum (2002)). In practice, researchers gradually increase the sensitivity parameter Γ, report the "worst-case" p-value under each Γ, and find the largest Γ such that the "worst-case" p-value exceeds the prespecified level α. Such a changepoint Γ is called a "sensitivity value" and informs the magnitude of the potential hidden bias required to alter the causal conclusion (Zhao (2019)).

For example, in a paired study where $n_i = 2$ for all i, a commonly used family of test statistics are sign-score statistics, including McNemar's test and Wilcoxon's signed rank test. Their general form is $T_{\mathrm{ss}} = \sum_{i=1}^{I} d_i \sum_{j=1}^{2} c_{ij} Z_{ij}$, where both $d_i \geq 0$ and $c_{ij} \in \{0, 1\}$ are functions of $\mathbf{R}$, and so are fixed under H_0. When each R_{ij} is binary, setting $d_i = 1$ and $c_{ij} = R_{ij}$ gives McNemar's test. For $i = 1, \dots, I$, define $\overline{T}_{\Gamma,i}$ as independent random variables taking the d_i with

probability p_i^+ and the value zero with probability $1 - p_i^+$, where

$$p_i^+ = \begin{cases} 0 & \text{if } c_{i1} = c_{i2} = 0, \\ 1 & \text{if } c_{i1} = c_{i2} = 1, \\ \dfrac{\Gamma}{1 + \Gamma} & \text{if } c_{i1} \neq c_{i2}. \end{cases}$$

As shown in Section 4.3 in Rosenbaum (2002), under the Rosenbaum bounds (2.1), for all t and $\Gamma \geq 1$, we have $\mathrm{pr}(T_{\mathrm{ss}} \geq t \mid \mathcal{F}, \mathcal{Z}) \leq \mathrm{pr}(\sum_{i=1}^{I} \overline{T}_{\Gamma,i} \geq t \mid \mathcal{F}, \mathcal{Z})$. That is, in the Rosenbaum bounds sensitivity analysis, the "worst-case" p-value under Γ reported by T_{ss} given t is $\mathrm{pr}(\sum_{i=1}^{I} \overline{T}_{\Gamma,i} \geq t \mid \mathcal{F}, \mathcal{Z})$. Assuming no interactions between observed and unobserved covariates, the upper bound $\mathrm{pr}(\sum_{i=1}^{I} \overline{T}_{\Gamma,i} \geq t \mid \mathcal{F}, \mathcal{Z})$ is sharp in the sense that it can be achieved for some arrangements of unobserved covariates; see Section 4.3 in Rosenbaum (2002). However, as we show in Theorem 1 in Section 3, in the presence of interactions between observed and unobserved covariates, the upper bound $\mathrm{pr}(\sum_{i=1}^{I} \overline{T}_{\Gamma,i} \geq t \mid \mathcal{F}, \mathcal{Z})$ induced from the Rosenbaum bounds (2.1) is in general loose, in the sense that it cannot be achieved for any arrangements of unobserved covariates. The result in Theorem 1 also holds for general matching regimes (including pair matching, matching with multiple controls, and full matching).

3. The Rosenbaum Bounds are Loose in the Presence of X–U Interactions

The Rosenbaum bounds (2.1) are an odds ratio bound imposed on all matched sets that does not explicitly involve the observed covariates $\mathbf{x}_{ij}$ and a hypothesized unobserved covariate u_{ij}; it is natural to consider how it can be derived from a model expressed in terms of $\mathbf{x}_{ij}$ and u_{ij} for the treatment assignment probability π_{ij} (Rosenbaum (2002)). Since the first paper on the Rosenbaum bounds sensitivity analysis (Rosenbaum (1987)), considering a logit form linking π_{ij} to $\mathbf{x}_{ij}$ and u_{ij} with no interactions between $\mathbf{x}_{ij}$ and u_{ij} has been a routine way of interpreting the Rosenbaum bounds (2.1) and has been applied in numerous studies (Rosenbaum (2002); DiPrete and Gangl (2004)):

$$\log\left(\frac{\pi_{ij}}{1 - \pi_{ij}}\right) = g(\mathbf{x}_{ij}) + \gamma u_{ij}, \quad u_{ij} \in [0, 1], \tag{3.1}$$

where $\gamma \in \mathbb{R}$ is unknown, and $g(\cdot)$ is an arbitrary unknown function of $\mathbf{x}_{ij}$. Note that the constraint $u_{ij} \in [0, 1]$ is no more restrictive than assuming a bounded support of u_{ij}, and is only imposed to make the scalar γ more interpretable

(Rosenbaum (1987, 2002)). For example, if the original unobserved covariate $\widetilde{u}_{ij} \in [0, \xi]$, for some $0 < \xi < +\infty$, we just need to consider a normalized unobserved covariate $u_{ij} = \xi^{-1}\widetilde{u}_{ij} \in [0, 1]$ and consider the logit model of π_{ij} in terms of $\mathbf{x}_{ij}$ and $u_{ij} = \xi^{-1}\widetilde{u}_{ij} \in [0, 1]$. It is then straightforward to show that the Rosenbaum bounds (2.1) can be implied from (3.1) with $\Gamma = \exp(|\gamma|)$ (Rosenbaum (2002); DiPrete and Gangl (2004)). Note that the Rosenbaum bounds (2.1) are a type of nonparametric odds ratio bounds that can be implied from more general models on π_{ij} than model (3.1), with an appropriate equation linking Γ and the proposed model on π_{ij}. In the previous literature, model (3.1) is used extensively as a working model for interpreting the Rosenbaum bounds, owing to its simplicity and clarity. However, as we show in the rest of this section, only considering model (3.1) and ignoring potential interactions between observed and unobserved covariates can be harmful in a sensitivity analysis.

In this section, we instead consider a more general model of π_{ij} in terms of $\mathbf{x}_{ij}$ and u_{ij}, allowing for any possible additive two-way interactions between each $\mathbf{x}_{ij(k)}$ and u_{ij} (X–U interactions):

$$\log\left(\frac{\pi_{ij}}{1 - \pi_{ij}}\right) = g(\mathbf{x}_{ij}) + \beta^T \mathbf{x}_{ij} \times u_{ij} + \gamma u_{ij}, \quad u_{ij} \in [0, 1], \qquad (3.2)$$

where $\beta^T \in \mathbb{R}^K$ and $\gamma \in \mathbb{R}$ are unknown, and $g(\cdot)$ is an arbitrary unknown function of $\mathbf{x}_{ij}$. Similarly to the arguments under model (3.1), the constraint $u_{ij} \in [0, 1]$ is no more restrictive than assuming a bounded support of u_{ij}, and is only imposed to make β^T and γ more interpretable. When $\beta^T = \mathbf{0}$ (i.e., no X–U interactions), model (3.2) reduces to the original model (3.1) that motivated the Rosenbaum bounds sensitivity analysis (Rosenbaum (1987)). Under (3.2), according to the definition of Γ in the Rosenbaum bounds (2.1), the following equation linking the prespecified sensitivity parameter Γ and the unknown parameters (β^T, γ) can be obtained:

$$\begin{aligned} \Gamma &= \max_{i,j,j'} \frac{\pi_{ij}(1 - \pi_{ij'})}{\pi_{ij'}(1 - \pi_{ij})} \quad \text{subject to } \mathbf{x}_{ij} = \mathbf{x}_{ij'} \text{ and } u_{ij}, u_{ij'} \in [0, 1] \text{ for all } i, j, j' \\ &= \max_{i=1,\dots,I} \exp(|\beta^T \mathbf{x}_i + \gamma|). \end{aligned} \qquad (3.3)$$

See the Appendix for a derivation of equation (3.3). Note that when $\beta^T = \mathbf{0}$, equation (3.3) reduces to the commonly used equation $\Gamma = \exp(|\gamma|)$ obtained under model (3.1). A key insight from equation (3.3) is that, in the presence of X–U interactions, setting the sensitivity parameter Γ not only incorporates our prior belief on the unknown structural parameters (β^T, γ), but also information

about the matched observed covariates $\mathbf{x}_i$, for $i = 1, \ldots, I$. The following result claims that the Rosenbaum bounds (2.1) will almost inevitably be conservative if there are any interactions between observed and unobserved covariates.

Theorem 1. *Consider the sensitivity parameter Γ defined in the Rosenbaum bounds (2.1). Let $\Gamma > 1$, and suppose that there exist two matched sets i_1 and i_2 such that $\mathbf{x}_{i_1} \neq \pm \mathbf{x}_{i_2}$. Then, we have under model (3.2), there exist some $\mathbf{x}^* \in \mathbb{R}^K$ and a subset E of $\mathbb{R}^K$ of Lebesgue measure zero, such that for any $\beta^T \neq \mathbf{0}$ (i.e., if there exist any interaction terms between $\mathbf{x}_{ij}$ and u_{ij}) and $\beta^T \notin E$, the Rosenbaum bounds (2.1) are loose for any matched set i with $\mathbf{x}_i \neq \pm \mathbf{x}^*$, in the sense that for any matched set i with $\mathbf{x}_i \neq \pm \mathbf{x}^*$, there exists some $\Upsilon_i < \Gamma$ such that*

$$\Upsilon_i^{-1} \leq \frac{\pi_{ij}(1 - \pi_{ij'})}{\pi_{ij'}(1 - \pi_{ij})} \leq \Upsilon_i, \quad \text{for all } j, j'.$$

Proofs of all theorems and corollaries presented here are provided in the Appendix. We consider a simple example to illustrate the principle of Theorem 1.

Example 1. Suppose there is only one observed covariate $x_{ij} \in \{0, 1\}$, and also an unobserved covariate $u_{ij} \in [0, 1]$. Under model (3.2), we have $\log(\pi_{ij}/(1 - \pi_{ij})) = g(x_{ij}) + \beta x_{ij} u_{ij} + \gamma u_{ij}$. According to (3.3), we have $\Gamma = \max\{\exp(|\gamma|), \exp(|\beta + \gamma|)\}$. It is clear that if $\beta \neq 0$ or -2γ, we have $\exp(|\gamma|) \neq \exp(|\beta + \gamma|)$. If $\Gamma = \exp(|\gamma|) > \exp(|\beta + \gamma|)$, then the Rosenbaum bounds (2.1) are loose for any matched set i with the common observed covariate $x_i = 1$. That is, for all matched sets i with $x_i = 1$, we have

$$\Gamma^{-1} < \exp(-|\beta + \gamma|) \leq \frac{\pi_{ij}(1 - \pi_{ij'})}{\pi_{ij'}(1 - \pi_{ij})} \leq \exp(|\beta + \gamma|) < \Gamma,$$
$$\text{for all } j, j' \text{ and } u_{ij}, u_{ij'} \in [0, 1].$$

Similarly, if $\Gamma = \exp(|\beta + \gamma|) > \exp(|\gamma|)$, the Rosenbaum bounds (2.1) are loose for any matched set i with $x_i = 0$. Therefore, when $\beta \neq 0$, unless $\beta \in \{-2\gamma\}$ (a subset of $\mathbb{R}$ of Lebesgue measure zero), the Rosenbaum bounds are loose for either all matched sets i with $x_i = 0$ or all matched sets i with $x_i = 1$.

4. Sharper Odds Ratio Bounds Accounting for X–U Interactions

In this section, we give new odds ratio bounds that are sharper than the Rosenbaum bounds (2.1) when a researcher is concerned about the possible interaction between a particular observed covariate, say, the kth component $x_{(k)}$ of the observed covariates vector $\mathbf{x}$, and the unobserved covariate u. We consider a sub-model of (3.2) that allows for a possible interaction term linking $x_{(k)}$ and u:

$$\text{logit}(\pi_{ij}) = \log\left(\frac{\pi_{ij}}{1 - \pi_{ij}}\right) = g(\mathbf{x}_{ij}) + \widetilde{\beta}\,\widetilde{x}_{ij(k)}u_{ij} + \gamma u_{ij}, \quad u_{ij} \in [0, 1], \quad (4.1)$$

where $\widetilde{\beta}, \gamma \in \mathbb{R}$ are unknown, and $g(\cdot)$ is an unknown function of $\mathbf{x}_{ij}$. Again, the unobserved covariate $u_{ij} \in [0, 1]$ is normalized to make $\widetilde{\beta}$ and γ more interpretable. Each $\widetilde{x}_{ij(k)} = (x_{ij(k)} - \min_{i,j} x_{ij(k)})/(\max_{i,j} x_{ij(k)} - \min_{i,j} x_{ij(k)}) \in [0, 1]$ is also normalized to make $\widetilde{\beta}$ more interpretable. Note that when $\widetilde{\beta} = 0$, model (4.1) reduces to the original model assuming no interaction terms that motivated the Rosenbaum bounds sensitivity analysis (Rosenbaum (1987)). In addition to the sensitivity parameter Γ defined in (2.1) that quantifies the magnitude of the largest possible bias over all matched sets, when $\Gamma > 1$, we introduce another prespecified sensitivity parameter λ under model (4.1) as

$$\lambda = \frac{(\partial \text{logit}(\pi_{ij})/\partial u_{ij}) \mid x_{ij(k)} = \max_{i,j} x_{ij(k)}}{(\partial \text{logit}(\pi_{ij})/\partial u_{ij}) \mid x_{ij(k)} = \min_{i,j} x_{ij(k)}} = \frac{\widetilde{\beta} + \gamma}{\gamma}, \quad \gamma \neq 0. \quad (4.2)$$

The sensitivity parameter λ quantifies how distinct the effects of u can be on the treatment assignment probability under the largest and smallest possible values of $x_{(k)}$. Note that when $\widetilde{\beta} = 0$ (i.e., no interaction between $x_{(k)}$ and u), we have $\lambda = 1$. Let $\widetilde{x}_{i(k)}$ denote the normalized common covariate $x_{i(k)}$ for matched set i; therefore, $\widetilde{x}_{i(k)} = \widetilde{x}_{ij(k)} = \widetilde{x}_{ij'(k)}$, for all j, j'. Then, we have the following sharper odds ratio bounds.

Theorem 2. *Consider the sensitivity parameter Γ defined in the Rosenbaum bounds (2.1) with $\Gamma > 1$. Under model (4.1), which allows for a possible interaction between the observed covariate $x_{(k)}$ and the normalized unobserved covariate u, consider the sensitivity parameter λ defined in (4.2). Then, we have*

$$\Gamma_{\lambda,i}^{-1} \leq \frac{\pi_{ij}(1 - \pi_{ij'})}{\pi_{ij'}(1 - \pi_{ij})} \leq \Gamma_{\lambda,i} \quad \text{for all } i \in \{1, \ldots, I\} \text{ and } j, j' \in \{1, \ldots, n_i\}, \quad (4.3)$$

where

$$\Gamma_{\lambda,i} = \begin{cases} \Gamma^{|(\lambda-1)\widetilde{x}_{i(k)}+1|} & \text{if } |\lambda| \leq 1, \\ \Gamma^{|(1-\lambda^{-1})\widetilde{x}_{i(k)}+\lambda^{-1}|} & \text{if } |\lambda| > 1. \end{cases}$$

We have $1 \leq \Gamma_{\lambda,i} \leq \Gamma$, for all i, and the equality $\Gamma_{\lambda,i} = \Gamma$ holds for matched set i if and only if at least one of the following three conditions holds: (a) $\lambda = 1$; (b) $|\lambda| \leq 1$ and $x_{i(k)} = \min_i x_{i(k)}$; (c) $|\lambda| \geq 1$ and $x_{i(k)} = \max_i x_{i(k)}$. The bounds (4.3) are sharp in the sense that for all i, j, j', there exist $u_{ij}, u_{ij'} \in [0, 1]$ such that $\{\pi_{ij}(1 - \pi_{ij'})\}/\{\pi_{ij'}(1 - \pi_{ij})\} = \Gamma_{\lambda,i}$.

A key feature of the sharper odds ratio bounds in Theorem 2 is that they in-

Table 1. $\Gamma_{\lambda,i}$ in the sharper odds ratio bounds proposed in Theorem 2 under various Γ, λ, and $\widetilde{x}_{i(k)} \in [0,1]$. Note that $\Gamma_{\lambda,i} = \Gamma$ when $\lambda = 1$.

$\Gamma = 2$	$\lambda = \frac{1}{8}$	$\lambda = \frac{1}{4}$	$\lambda = \frac{1}{2}$	$\lambda = 1$	$\lambda = 2$	$\lambda = 4$	$\lambda = 8$
$\widetilde{x}_{i(k)} = 0$	2.00	2.00	2.00	2.00	1.41	1.19	1.09
$\widetilde{x}_{i(k)} = \frac{1}{5}$	1.77	1.80	1.87	2.00	1.52	1.32	1.23
$\widetilde{x}_{i(k)} = \frac{2}{5}$	1.57	1.62	1.74	2.00	1.62	1.46	1.39
$\widetilde{x}_{i(k)} = \frac{3}{5}$	1.39	1.46	1.62	2.00	1.74	1.62	1.57
$\widetilde{x}_{i(k)} = \frac{4}{5}$	1.23	1.32	1.52	2.00	1.87	1.80	1.77
$\widetilde{x}_{i(k)} = 1$	1.09	1.19	1.41	2.00	2.00	2.00	2.00
$\Gamma = 3$	$\lambda = \frac{1}{8}$	$\lambda = \frac{1}{4}$	$\lambda = \frac{1}{2}$	$\lambda = 1$	$\lambda = 2$	$\lambda = 4$	$\lambda = 8$
$\widetilde{x}_{i(k)} = 0$	3.00	3.00	3.00	3.00	1.73	1.32	1.15
$\widetilde{x}_{i(k)} = \frac{1}{5}$	2.48	2.54	2.69	3.00	1.93	1.55	1.39
$\widetilde{x}_{i(k)} = \frac{2}{5}$	2.04	2.16	2.41	3.00	2.16	1.83	1.69
$\widetilde{x}_{i(k)} = \frac{3}{5}$	1.69	1.83	2.16	3.00	2.41	2.16	2.04
$\widetilde{x}_{i(k)} = \frac{4}{5}$	1.39	1.55	1.93	3.00	2.69	2.54	2.48
$\widetilde{x}_{i(k)} = 1$	1.15	1.32	1.73	3.00	3.00	3.00	3.00

corporate the information of the observed covariates among the matched samples, which is ignored by the Rosenbaum bounds (2.1). In Table 1, we present numerical illustrations of the gap between the sensitivity parameter Γ in the Rosenbaum bounds (2.1) and $\Gamma_{\lambda,i}$ in the sharper odds ratio bounds proposed in Theorem 2.

If the observed covariate $x_{(k)} \in \{0,1\}$ is a binary (dummy) variable, model (4.1) reduces to

$$\text{logit}(\pi_{ij}) = g(\mathbf{x}_{ij}) + \widetilde{\beta}\, x_{ij(k)}u_{ij} + \gamma u_{ij}, \quad x_{ij(k)} \in \{0,1\}, \quad u_{ij} \in [0,1], \qquad (4.4)$$

and the sensitivity parameter λ as defined in (4.2) reduces to

$$\lambda = \frac{(\partial \text{logit}(\pi_{ij})/\partial u_{ij}) \mid x_{ij(k)} = 1}{(\partial \text{logit}(\pi_{ij})/\partial u_{ij}) \mid x_{ij(k)} = 0} = \frac{\widetilde{\beta} + \gamma}{\gamma}, \quad \gamma \neq 0.$$

That is, the sensitivity parameter λ is simply the ratio of the effect of u on the logit of the treatment assignment probability (denoted as $\partial \text{logit}/\partial u$) conditional on $x_{(k)} = 1$ to that conditional on $x_{(k)} = 0$. Theorem 2 implies the following sharper odds ratio bounds when $x_{(k)}$ is binary.

Corollary 1. *Under the same setting as that in Theorem 2, if the observed covariate $x_{(k)} \in \{0,1\}$ is a binary (dummy) variable, we have:*

1. *If $|\lambda| = 1$, then the Rosenbaum bounds (2.1) are sharp for all matched sets,*

in the sense that for all i, j, j', there exist some $u_{ij}, u_{ij'} \in [0, 1]$ such that $\{\pi_{ij}(1 - \pi_{ij'})\}/\{\pi_{ij'}(1 - \pi_{ij})\} = \Gamma$ or Γ^{-1}.

2. *If $|\lambda| < 1$, then the Rosenbaum bounds (2.1) are sharp for all matched sets i with $x_{i(k)} = 0$. However, for all matched sets i with $x_{i(k)} = 1$, the Rosenbaum bounds (2.1) can be improved as follows: for all i, j, j' with $x_{i(k)} = 1$, we have $\Gamma^{-|\lambda|} \leq \{\pi_{ij}(1 - \pi_{ij'})\}/\{\pi_{ij'}(1 - \pi_{ij})\} \leq \Gamma^{|\lambda|}$.*

3. *If $|\lambda| > 1$, then the Rosenbaum bounds (2.1) are sharp for all matched sets i with $x_{i(k)} = 1$. However, for all matched sets i with $x_{i(k)} = 0$, the Rosenbaum bounds (2.1) can be improved as follows: for all i, j, j' with $x_{i(k)} = 0$, we have $\Gamma^{-1/|\lambda|} \leq \{\pi_{ij}(1 - \pi_{ij'})\}/\{\pi_{ij'}(1 - \pi_{ij})\} \leq \Gamma^{1/|\lambda|}$.*

Corollary 1 implies that in the binary covariate case, the sign of the sensitivity parameter λ does not matter in a sensitivity analysis. It also implies that in this case, the more $|\lambda|$ departs from one, the less the treatment assignments can potentially depart from random assignments within some matched sets. When $|\lambda| = 1$, the bounds in Corollary 1 reduce to the Rosenbaum bounds (2.1).

Remark 1. Note that even if some observed covariates are not exactly matched for some matched sets, the sharper odds ratio bounds proposed in Theorem 2 and Corollary 1 can still be used as approximate sensitivity bounds, as long as the following three conditions hold: 1) the function $g(\mathbf{x}_{ij})$ in model (4.1) is a continuous function; 2) the observed covariates are very close among individuals within each matched set i (i.e., $\mathbf{x}_{ij} \approx \mathbf{x}_{ij'}$, for all i, j, j'); and 3) the observed covariates $x_{(k)}$ that can interact with the unobserved covariate are equal among individuals within each matched set (i.e., $x_{ij(k)} = x_{ij'(k)} = x_{i(k)}$, for all i, j, j'). Then, the following calculation, combined with the arguments in the proof of Theorem 2 in the Appendix, justify the validity of the approximation:

$$
\begin{aligned}
\frac{\pi_{ij}(1 - \pi_{ij'})}{\pi_{ij'}(1 - \pi_{ij})} &= \frac{\exp\{g(\mathbf{x}_{ij}) + \widetilde{\beta}\,\widetilde{x}_{ij(k)}u_{ij} + \gamma u_{ij}\}}{\exp\{g(\mathbf{x}_{ij'}) + \widetilde{\beta}\,\widetilde{x}_{ij'(k)}u_{ij'} + \gamma u_{ij'}\}} \\
&= \frac{\exp\{g(\mathbf{x}_{ij})\}}{\exp\{g(\mathbf{x}_{ij'})\}} \frac{\exp\{\widetilde{\beta}\,\widetilde{x}_{ij(k)}u_{ij} + \gamma u_{ij}\}}{\exp\{\widetilde{\beta}\,\widetilde{x}_{ij'(k)}u_{ij'} + \gamma u_{ij'}\}} \\
&\approx \exp\{(\widetilde{\beta}\,\widetilde{x}_{i(k)} + \gamma)(u_{ij} - u_{ij'})\} \\
&\quad (\text{because } \mathbf{x}_{ij} \approx \mathbf{x}_{ij'} \text{ and } \widetilde{x}_{ij(k)} = \widetilde{x}_{ij'(k)} = \widetilde{x}_{i(k)}).
\end{aligned}
$$

5. Implementation of the Proposed Sharper Odds Ratio Bounds

As we have shown in Theorem 2 in Section 4, considering potential X–U interactions can lead to sharper odds ratio bounds than the Rosenbaum bounds (2.1). By implementing the sharper bounds proposed in Theorem 2, researchers can perform a less conservative sensitivity analysis than when applying the Rosenbaum bounds directly. The sharper bounds proposed in Theorem 2 can be embedded directly in many previous results based on a Rosenbaum bounds sensitivity analysis: in many cases, researchers just need to follow the procedure of the Rosenbaum sensitivity analysis, replacing all sensitivity parameters Γ with $\Gamma_{\lambda,i}$ proposed in Theorem 2 in each matched set. For example, the following result shows how applying Theorem 2 to perform a sensitivity analysis with a sign-score statistic T_{ss} can result in a less conservative "worst-case" p-value than the one reported by directly applying the Rosenbaum bounds sensitivity analysis with T_{ss}.

Corollary 2. *Let $T_{ss} = \sum_{i=1}^{I} d_i \sum_{j=1}^{2} c_{ij} Z_{ij}$ be a sign-score statistic as introduced in Section 2. Consider testing Fisher's sharp null of no treatment effect H_0, and the sensitivity parameters Γ defined in the Rosenbaum bounds (2.1) and λ defined in (4.2) under model (4.1). Define $\Gamma_{\lambda,i}$ as in Theorem 2 and $\overline{T}_{\Gamma,i}$ as in Section 2. For $i = 1, \ldots, I$, define $\widetilde{T}_{\Gamma,\lambda,i}$ to be independent random variables taking the d_i with probability $\widetilde{p}_{\lambda,i}$ and the value zero with probability $1 - \widetilde{p}_{\lambda,i}$, where*

$$\widetilde{p}_{\lambda,i} = \begin{cases} 0 & \text{if } c_{i1} = c_{i2} = 0, \\ 1 & \text{if } c_{i1} = c_{i2} = 1, \\ \dfrac{\Gamma_{\lambda,i}}{1 + \Gamma_{\lambda,i}} & \text{if } c_{i1} \neq c_{i2}. \end{cases}$$

Then, for all t and any fixed $\Gamma > 1$ and $\lambda \in \mathbb{R}$, we have $pr(T_{ss} \geq t \mid \mathcal{F}, \mathcal{Z}) \leq pr(\sum_{i=1}^{I} \widetilde{T}_{\Gamma,\lambda,i} \geq t \mid \mathcal{F}, \mathcal{Z})$, for any $\mathbf{u} \in [0,1]^N$, and we have $pr(\sum_{i=1}^{I} \widetilde{T}_{\Gamma,\lambda,i} \geq t \mid \mathcal{F}, \mathcal{Z}) \leq pr(\sum_{i=1}^{I} \overline{T}_{\Gamma,i} \geq t \mid \mathcal{F}, \mathcal{Z})$. The upper bound $pr(\sum_{i=1}^{I} \widetilde{T}_{\Gamma,\lambda,i} \geq t \mid \mathcal{F}, \mathcal{Z})$ is sharp in the sense that it can be achieved for some $\mathbf{u}$.

6. Illustration: The Effect of Anger/Hostility Tendency on Heart Problems

Type-A behavior is characterized by hostility, intense ambition, a competitive "drive," a constant preoccupation with deadlines, and a sense of time urgency (Rosenman et al. (1976)). Early research data suggested that type-A behavior was related to heart problems, but the original findings have not been supported by subsequent research (Myrtek (2001)). Some researchers have turned their fo-

cus to whether tending to be angry and hostile—one of the specific aspects of a type-A personality—could cause heart problems (Chida and Steptoe (2009)). To study this, we consider data among males from the Wisconsin Longitudinal Study, a long-term study of a random sample of individuals who graduated from Wisconsin high schools in 1957 (Herd, Carr and Roan (2014)). We define a binary indicator of tending to be angry/hostile (i.e., treated), taking the value one if the respondent said on the 1992–1993 survey (when respondents were approximately 53) that in the last week there were three or more days on which he felt angry or hostile for several hours, and zero (i.e., control) if there were no such days in the last week. We compare high anger/hostility tendency (i.e., treatment) to low anger/hostility tendency (i.e., control) and exclude middle levels. We take the outcome (heart problem indicator) to be one if the respondent reported having had a heart attack, coronary heart disease, or other heart problems in the 2003–2005 survey, and zero otherwise. We pair match each treated individual with a control on the following cardiovascular disease risk factors (Kawachi et al. (1996)): age, educational attainment, body mass indicator, drinking alcohol or not, smoking regularly or not, and childhood maltreatment indicator. The childhood maltreatment indicator is one if the respondent reported any childhood physical or sexual abuse, and zero otherwise. Childhood maltreatment has been found to be associated with both anger/hostility tendency and heart problems (Carver et al. (2014); Korkeila et al. (2010)), and therefore is a confounder that needs to be controlled for. We discarded all records with missing outcomes or covariates, and used optimal matching (Rosenbaum (2010)) to match each treated with a control for the six baseline observed covariates, leaving 54 matched pairs. The absolute standardized differences (i.e., difference in means divided by the pooled standard deviation) between the treated and control groups are less than 0.1 for all six baseline observed covariates, indicating good overall balance (Rosenbaum (2010)). The smoking indicator and childhood maltreatment indicator are exactly matched between the treated unit and control within each matched pair. Although the other four covariates, that is, age, educational attainment, body mass indicator, and drinking alcohol or not, are not exactly matched between the treated unit and its matched control for all pairs, they are closely matched for most pairs, because the correlations for these four covariates equal 0.92, 0.98, 0.95, and 0.86, respectively. Therefore, according to Remark 1, in our study, the sharper odds ratio bounds developed in Section 4 can be applied as approximate sensitivity bounds in our sensitivity analysis.

Another covariate we are concerned about as a confounder is the genotype monoamine oxidase A (MAOA), which has been found to be associated with both

Table 2. The "worst-case" p-values reported by McNemar's test under various Γ and λ. When $|\lambda| = 1$, they are the same as those reported by the Rosenbaum bounds sensitivity analysis.

	$|\lambda| = \frac{1}{8}$	$|\lambda| = \frac{1}{4}$	$|\lambda| = \frac{1}{2}$	$|\lambda| = 1$	$|\lambda| = 2$	$|\lambda| = 4$	$|\lambda| = 8$
$\Gamma = 1.31$	0.037	0.039	0.042	**0.050**	0.033	0.027	0.024
$\Gamma = 1.37$	0.043	0.045	**0.050**	0.060	0.038	0.030	0.026
$\Gamma = 1.42$	0.048	**0.050**	0.056	0.068	0.042	0.032	0.028
$\Gamma = 1.44$	**0.050**	0.052	0.058	0.072	0.043	0.033	0.028
$\Gamma = 1.52$	0.058	0.061	0.069	0.087	**0.050**	0.036	0.031
$\Gamma = 1.81$	0.090	0.098	0.114	0.150	0.076	**0.050**	0.040
$\Gamma = 2.11$	0.127	0.140	0.166	0.223	0.105	0.065	**0.050**
Sensitivity value	1.44	1.42	1.37	1.31	1.52	1.81	2.11

aggressive behavior and heart disease (McDermott et al. (2009); Kaludercic et al. (2011)). The genetic data of the Wisconsin Longitudinal Study are not publicly available, therefore we treat the MAOA genotype as an unobserved covariate. We denote the unobserved MAOA genotype indicator (i.e., u in model (4.1)) as one if the individual has a low-activity MAOA genotype (MAOA-L), and zero if it is a high-activity MAOA genotype (MAOA-H). According to a controlled experiment done in McDermott et al. (2009), individuals with MAOA-L are more likely to show aggression, suggesting $\gamma > 0$ in model (4.1). Childhood maltreatment has been shown to significantly interact with the MAOA genotype to confer risk for aggressive behavior: maltreated children with MAOA-L are more likely to develop violent behavior or show hostility (Caspi et al. (2002); Byrd and Manuck (2014)), suggesting that the coefficient of the interaction term $\widetilde{\beta}$ in model (4.1) is greater than zero, and that the sensitivity parameter $\lambda = (\widetilde{\beta} + \gamma)/\gamma > 1$. While setting a precise range for λ needs further empirical study, some related studies suggest that $\lambda \gg 1$. For example, according to Figure 2A in Caspi et al. (2002), among severely maltreated (during childhood) males, the logit of probability of conduct disorder among those with MAOA-L is much greater than that among those with MAOA-H. In contrast, among non-maltreated males, these two logits are extremely close. Therefore, if we treat the conduct disorder indicator as a proxy for the anger/hostility tendency indicator, the results from Caspi et al. (2002) suggest that $\partial \text{logit}/\partial u$ if maltreated is much greater than that if non-maltreated (i.e., $\widetilde{\beta} + \gamma \gg \gamma$), implying $\lambda \gg 1$.

We use Corollary 2 to calculate the "worst-case" p-values $\text{pr}(\sum_{i=1}^{I} \widetilde{T}_{\Gamma,\lambda,i} \geq t \mid \mathcal{F}, \mathcal{Z})$ reported by McNemar's test under various Γ and λ, where λ quantifies the possible interaction between the childhood maltreatment indicator and the

MAOA genotype; see Table 2. We also report corresponding sensitivity values under various λ. Note that when $|\lambda| = 1$, the "worst-case" p-values are the same as those reported by the Rosenbaum bounds sensitivity analysis. As discussed above, we are particularly concerned about the cases with $\lambda > 1$. From Table 2, we can see that by applying the sharper odds ratio bounds developed in Section 4, the "worst-case" p-values are much less conservative than those reported by applying the Rosenbaum bounds sensitivity analysis directly, especially when λ is much greater than one (i.e., there is a significant X–U interaction), making a sensitivity analysis significantly more insensitive to hidden bias caused by the potential unobserved covariate. Therefore, for this particular data set, directly applying the Rosenbaum bounds sensitivity analysis can only detect a significant treatment effect up to a moderate magnitude of hidden bias (i.e., $\Gamma = 1.31$). In contrast, applying our sharper odds ratio bounds to perform a sensitivity analysis allows the researcher to detect a significant treatment effect when there is a significant X–U interaction, say, $\lambda \geq 2$, up to a significantly larger magnitude of hidden bias, $\Gamma = 1.52$. A bias of $\Gamma = 1.5$ is nontrivial because it corresponds to an unobserved covariate that doubles the odds of treatment and increases the odds of a positive treated-minus-control difference in observed outcomes by a factor of four (Rosenbaum and Silber (2009)).

7. Discussion

We here provide practical guidance for empirical researchers on when and how our new odds ratio bounds should be used when conducting a Rosenbaum-type sensitivity analysis in matched studies. On the one hand, if a researcher has some prior knowledge about which direction or to what extent the effects of the concerned unobserved covariate should vary on the treatment assignment probability with different values of the related observed covariate (i.e., a plausible range of the sensitivity parameter λ defined in (4.2)), we strongly recommend that, in addition to a sensitivity analysis using the traditional Rosenbaum bounds (2.1) (i.e., setting $\lambda = 1$), she or he report the results of a sensitivity analysis under a plausible range of λ to better incorporate the expert knowledge, thus making the sensitivity analysis more informative and less conservative, as shown in Section 6. On the other hand, even if there is no evidence of any X–U interactions or credible information on the range of λ for the concerned X–U interaction term, an empirical researcher could still benefit from using our new methods. Suppose a researcher conducted the Rosenbaum bounds sensitivity analysis and found that the "worst-case" p-values $> \alpha = 0.05$, even under Γ close to one; that is, the

sensitivity value is small. Instead of rushing to claim that the causal conclusion is inevitably sensitive to hidden bias, the researcher can diagnose the reasons for sensitivity by selecting candidate X–U interaction terms, and checking the corresponding "worst-case" p-values and sensitivity values under various λ using our new odds ratio bound. If the sensitivity values are always small for a reasonably wide range of λ, then she or he can confirm that the causal conclusion should indeed be sensitive to hidden bias, regardless of potential X–U interactions. If instead the sensitivity value becomes substantially larger as λ departs from one, then this implies that the previous finding that the causal conclusion is sensitive to hidden bias could be due to ignoring the possible X–U interactions, in which case, the researcher can further investigate the possibility of actual X–U interactions to report the sensitivity analysis in a more comprehensive way.

There are limitations to the new odds ratio bound introduced in this work. First, it is only applicable for two-way X–U interactions. For example, if there is an additional three-way interaction term $\widetilde{x}_{ij(k)}\widetilde{x}_{ij(k')}u_{ij}$ in the treatment assignment probability model (4.1) for some $k \neq k'$, then the sensitivity parameter λ defined in (4.2) cannot fully capture how the effects of the unobserved covariate u_{ij} vary on the treatment assignment probability π_{ij} with different values of the two observed covariates $\widetilde{x}_{ij(k)}$ and $\widetilde{x}_{ij(k')}$. Second, when the observed covariate $\widetilde{x}_{ij(k)}$ in the concerned X–U interaction term is not binary, our new odds ratio bound is not applicable if the interaction term in model (4.1) is instead $f(\widetilde{x}_{ij(k)})u_{ij}$, where $f(\widetilde{x}_{ij(k)})$ is nonlinear in $\widetilde{x}_{ij(k)}$. Third, our new odds ratio bound cannot directly handle multiple X–U interaction terms, for example, when there are two interaction terms $\widetilde{x}_{ij(k)}u_{ij}$ and $\widetilde{x}_{ij(k')}u_{ij}$ in model (4.1). Although, in principle, this type of problem can be solved by introducing additional sensitivity parameters into the odds ratio bounds, doing so can make a sensitivity analysis complicated and difficult to interpret. Fourth, in the presence of X–U interactions, after computing the sensitivity value, it requires further study to clarify how large the sensitivity value needs to be to make a study insensitive to hidden bias. Assuming no X–U interactions, Hsu and Small (2013) proposed a strategy that calibrates the values of the sensitivity parameters in matched observational studies to the observed covariates to help empirical researchers set plausible ranges of values for the sensitivity parameters. It would be helpful to investigate how to extend Hsu and Small (2013) approach to the settings allowing X–U interactions. Fifth, in this work, we test only Fisher's sharp null hypothesis and have not discussed testing other types of null hypotheses, such as Neyman's weak null hypothesis. In pair-matched observational studies, a promising direction for extending the proposed method to allow testing Neyman's weak

null hypothesis is to investigate whether the sharper bounds proposed in Theorem 2 can be incorporated into the studentized sensitivity analysis framework for the sample average treatment effect developed in Fogarty (2020). Despite these limitations, this work shows how investigating the confounding mechanism more carefully in a matched observational study can make a sensitivity analysis more informative and comprehensive. It might be fruitful for future research to explore how other structural constraints besides the one we explored here might be used.

Acknowledgments

Appendix

A derivation of equation (3.3)

According to the definition of Γ in the Rosenbaum bounds (2.1),

$$\Gamma = \max_{i,j,j'} \frac{\pi_{ij}(1 - \pi_{ij'})}{\pi_{ij'}(1 - \pi_{ij})} \quad \text{subject to } \mathbf{x}_{ij} = \mathbf{x}_{ij'} \text{ and } u_{ij}, u_{ij'} \in [0,1] \text{ for all } i, j, j'.$$

So we have under model (3.2),

$$
\begin{aligned}
\Gamma &= \max_{i,j,j'} \max_{u_{ij}, u_{ij'} \in [0,1]} \frac{\pi_{ij}(1 - \pi_{ij'})}{\pi_{ij'}(1 - \pi_{ij})} \\
&= \max_{i,j,j'} \max_{u_{ij}, u_{ij'} \in [0,1]} \frac{\exp\{g(\mathbf{x}_{ij}) + \beta^T \mathbf{x}_{ij} \times u_{ij} + \gamma u_{ij}\}}{\exp\{g(\mathbf{x}_{ij'}) + \beta^T \mathbf{x}_{ij'} \times u_{ij'} + \gamma u_{ij'}\}} \\
&= \max_{i,j,j'} \max_{u_{ij}, u_{ij'} \in [0,1]} \exp\{(\beta^T \mathbf{x}_i + \gamma)(u_{ij} - u_{ij'})\} \quad (\text{because } \mathbf{x}_{ij} = \mathbf{x}_{ij'} = \mathbf{x}_i) \\
&= \max_{i,j,j'} \exp(|\beta^T \mathbf{x}_i + \gamma|) \\
&= \max_{i=1,\dots,I} \exp(|\beta^T \mathbf{x}_i + \gamma|).
\end{aligned}
$$

Therefore the desired equation holds.

Proof of Theorem 1

Proof. Let $\mathbf{x}^* \in \{\mathbf{x}_1, \ldots, \mathbf{x}_I\}$ be an observed covariate vector such that $\Gamma = \max_{i=1,\ldots,I} \exp(|\beta^T \mathbf{x}_i + \gamma|) = \exp(|\beta^T \mathbf{x}^* + \gamma|) > 1$. Since there exist two matched sets i_1 and i_2 such that $\mathbf{x}_{i_1} \neq \pm \mathbf{x}_{i_2}$, we have $\{\mathbf{x}_1, \ldots, \mathbf{x}_I\} \setminus \{\mathbf{x}^*, -\mathbf{x}^*\} \neq \emptyset$. For any matched set i such that $\mathbf{x}_i \neq \pm \mathbf{x}^*$, define the set

$$E_i = \left\{ \beta^T \in \mathbb{R}^K : \ \exp(|\beta^T \mathbf{x}_i + \gamma|) = \exp(|\beta^T \mathbf{x}^* + \gamma|) \ \text{and} \ \ \beta^T \neq \mathbf{0} \right\}.$$

Since $\mathbf{x}_i \neq \pm \mathbf{x}^*$, for any $\gamma \in \mathbb{R}$ we have

$$E_i = \left\{ \beta^T \in \mathbb{R}^K : \ \beta^T (\mathbf{x}_i - \mathbf{x}^*) = 0 \ \text{and} \ \ \beta^T \neq \mathbf{0} \right\}$$
$$\cup \left\{ \beta^T \in \mathbb{R}^K : \ \beta^T (\mathbf{x}_i + \mathbf{x}^*) + 2\gamma = 0 \ \text{and} \ \ \beta^T \neq \mathbf{0} \right\}$$

is a subset of $\mathbb{R}^K$ of Lebesgue measure zero. Let $E = \cup_{i: \ \mathbf{x}_i \neq \pm \mathbf{x}^*} E_i$, then E is also a subset of $\mathbb{R}^K$ of Lebesgue measure zero. For any matched set i such that $\mathbf{x}_i \neq \pm \mathbf{x}^*$, when $\beta^T \neq \mathbf{0}$ and $\beta^T \notin E$, note that

$$\begin{aligned}
\frac{\pi_{ij}(1 - \pi_{ij'})}{\pi_{ij'}(1 - \pi_{ij})} &= \frac{\exp\{g(\mathbf{x}_{ij}) + \beta^T \mathbf{x}_{ij} \times u_{ij} + \gamma u_{ij}\}}{\exp\{g(\mathbf{x}_{ij'}) + \beta^T \mathbf{x}_{ij'} \times u_{ij'} + \gamma u_{ij'}\}} \\
&= \exp\{(\beta^T \mathbf{x}_i + \gamma)(u_{ij} - u_{ij'})\} \quad (\text{because } \mathbf{x}_{ij} = \mathbf{x}_{ij'} = \mathbf{x}_i) \\
&\leq \exp(|\beta^T \mathbf{x}_i + \gamma|) \quad (\text{because } u_{ij}, u_{ij'} \in [0,1]) \\
&< \exp(|\beta^T \mathbf{x}^* + \gamma|) \quad (\text{by the definition of } E) \\
&= \Gamma,
\end{aligned}$$

therefore setting $\Upsilon_i = \exp(|\beta^T \mathbf{x}_i + \gamma|)$ we have, by symmetry,

$$\Gamma^{-1} < \exp(-|\beta^T \mathbf{x}_i + \gamma|) \leq \frac{\pi_{ij}(1 - \pi_{ij'})}{\pi_{ij'}(1 - \pi_{ij})} \leq \exp(|\beta^T \mathbf{x}_i + \gamma|) < \Gamma, \ \ \text{for all } j, j'.$$

So the desired result follows.

Proof of Theorem 2

Proof. Under model (4.1), according to the definition of Γ in the Rosenbaum bounds (2.1), we have

$$\Gamma = \max_{i,j,j'} \frac{\pi_{ij}(1 - \pi_{ij'})}{\pi_{ij'}(1 - \pi_{ij})} \quad \text{subject to } \mathbf{x}_{ij} = \mathbf{x}_{ij'} \text{ and } u_{ij}, u_{ij'} \in [0,1] \text{ for all } i, j, j',$$

where

$$\frac{\pi_{ij}(1 - \pi_{ij'})}{\pi_{ij'}(1 - \pi_{ij})} = \frac{\exp\{g(\mathbf{x}_{ij}) + \widetilde{\beta}\,\widetilde{x}_{ij(k)}u_{ij} + \gamma u_{ij}\}}{\exp\{g(\mathbf{x}_{ij'}) + \widetilde{\beta}\,\widetilde{x}_{ij'(k)}u_{ij'} + \gamma u_{ij'}\}}$$

$$= \exp\{(\widetilde{\beta}\,\widetilde{x}_{i(k)} + \gamma)(u_{ij} - u_{ij'})\}$$

$$\text{(because } \mathbf{x}_{ij} = \mathbf{x}_{ij'} = \mathbf{x}_i \text{ and } \widetilde{x}_{ij(k)} = \widetilde{x}_{ij'(k)} = \widetilde{x}_{i(k)})$$

$$\leq \exp(|\widetilde{\beta}\,\widetilde{x}_{i(k)} + \gamma|) \quad \text{(because } u_{ij}, u_{ij'} \in [0, 1]).$$

Therefore, we have

$$\Gamma = \max_{i,j,j'} \frac{\pi_{ij}(1 - \pi_{ij'})}{\pi_{ij'}(1 - \pi_{ij})} \quad \text{subject to } \mathbf{x}_{ij} = \mathbf{x}_{ij'} \text{ and } u_{ij}, u_{ij'} \in [0, 1] \text{ for all } i, j, j'$$

$$= \max_{i=1,\ldots,I} \exp(|\widetilde{\beta}\,\widetilde{x}_{i(k)} + \gamma|)$$

$$= \max\{\exp(|\gamma|), \exp(|\widetilde{\beta} + \gamma|)\}$$

$$\left(\text{because } \widetilde{x}_{i(k)} = \frac{x_{i(k)} - \min_i x_{i(k)}}{\max_i x_{i(k)} - \min_i x_{i(k)}} \in [0, 1]\right).$$

- Case 1: $|\lambda| = |(\widetilde{\beta} + \gamma)/\gamma| \leq 1$. In this case we have $\Gamma = \max\{\exp(|\gamma|), \exp(|\widetilde{\beta} + \gamma|)\} = \exp(|\gamma|)$. Therefore, we have

$$\frac{\pi_{ij}(1 - \pi_{ij'})}{\pi_{ij'}(1 - \pi_{ij})} \leq \exp(|\widetilde{\beta}\,\widetilde{x}_{i(k)} + \gamma|)$$

$$= \exp\left\{|\gamma| \times \left|\left(\frac{\widetilde{\beta} + \gamma}{\gamma} - 1\right)\widetilde{x}_{i(k)} + 1\right|\right\}$$

$$= \Gamma^{|(\lambda-1)\widetilde{x}_{i(k)}+1|} \quad \left(\text{because } \Gamma = \exp(|\gamma|) \text{ and } \lambda = \frac{\widetilde{\beta} + \gamma}{\gamma}\right).$$

Therefore, by symmetry we have

$$\Gamma^{-|(\lambda-1)\widetilde{x}_{i(k)}+1|} \leq \frac{\pi_{ij}(1 - \pi_{ij'})}{\pi_{ij'}(1 - \pi_{ij})} \leq \Gamma^{|(\lambda-1)\widetilde{x}_{i(k)}+1|}, \quad \text{for all } i, j, j',$$

and the above bounds are sharp in the sense that the upper bound can be achieved when $u_{ij} - u_{ij'} = \text{sign}(\widetilde{\beta}\,\widetilde{x}_{i(k)} + \gamma)$, and the lower bound can be achieved when $u_{ij} - u_{ij'} = -\text{sign}(\widetilde{\beta}\,\widetilde{x}_{i(k)} + \gamma)$, where we let $\text{sign}(x)$ equal 1 if $x > 0$, equal -1 if $x < 0$, and equal 0 if $x = 0$. Since $\widetilde{x}_{i(k)} \in [0, 1]$ and $|\lambda| \leq 1$, we have $\Gamma^{|(\lambda-1)\widetilde{x}_{i(k)}+1|} \leq \Gamma$ and the equality holds if and only if at least one of the following three conditions holds: (a) $\lambda = 1$; (b) $x_{i(k)} = \min_i x_{i(k)}$ (i.e., $\widetilde{x}_{i(k)} = 0$); (c) $\lambda = -1$ and $x_{i(k)} = \max_i x_{i(k)}$ (i.e., $\widetilde{x}_{i(k)} = 1$).

- Case 2: $|\lambda| = |(\widetilde{\beta} + \gamma)/\gamma| > 1$. In this case we have $\Gamma = \max\{\exp(|\gamma|), \exp(|\widetilde{\beta} + \gamma|)\} = \exp(|\widetilde{\beta} + \gamma|)$. So we have

$$\frac{\pi_{ij}(1 - \pi_{ij'})}{\pi_{ij'}(1 - \pi_{ij})} \leq \exp(|\widetilde{\beta}\, \widetilde{x}_{i(k)} + \gamma|)$$

$$= \exp\left\{|\widetilde{\beta} + \gamma| \times \left|\left(1 - \frac{\gamma}{\widetilde{\beta} + \gamma}\right)\widetilde{x}_{i(k)} + \frac{\gamma}{\widetilde{\beta} + \gamma}\right|\right\}$$

$$= \Gamma^{|(1 - \lambda^{-1})\widetilde{x}_{i(k)} + \lambda^{-1}|}$$

$$\left(\text{because } \Gamma = \exp(|\widetilde{\beta} + \gamma|) \text{ and } \lambda = \frac{\widetilde{\beta} + \gamma}{\gamma}\right).$$

Therefore, by symmetry we have

$$\Gamma^{-|(1 - \lambda^{-1})\widetilde{x}_{i(k)} + \lambda^{-1}|} \leq \frac{\pi_{ij}(1 - \pi_{ij'})}{\pi_{ij'}(1 - \pi_{ij})} \leq \Gamma^{|(1 - \lambda^{-1})\widetilde{x}_{i(k)} + \lambda^{-1}|}, \quad \text{for all } i, j, j',$$

and the above bounds are sharp, which is similar to the argument in Case 1. Since $\widetilde{x}_{i(k)} \in [0, 1]$ and $|\lambda| > 1$, we have $\Gamma^{|(1 - \lambda^{-1})\widetilde{x}_{i(k)} + \lambda^{-1}|} \leq \Gamma$ and the equality holds if and only if $x_{i(k)} = \max_i x_{i(k)}$ (i.e., $\widetilde{x}_{i(k)} = 1$).

The desired result follows from combining the arguments in Case 1 and Case 2.

Proof of Corollary 1

Proof. Consider the $\Gamma_{\lambda,i}$ defined in Theorem 2.

- Case 1: $|\lambda| = 1$. In this case, we have $\Gamma_{\lambda,i} = \Gamma^{|(\lambda-1)\widetilde{x}_{i(k)}+1|} = \Gamma$ for $x_{i(k)} \in \{0, 1\}$.

- Case 2: $|\lambda| < 1$. In this case, if $x_{i(k)} = 0$, we have $\Gamma_{\lambda,i} = \Gamma^{|(\lambda-1)\widetilde{x}_{i(k)}+1|} = \Gamma$. If $x_{i(k)} = 1$, we have $\Gamma_{\lambda,i} = \Gamma^{|(\lambda-1)\widetilde{x}_{i(k)}+1|} = \Gamma^{|\lambda|}$.

- Case 3: $|\lambda| > 1$. In this case, if $x_{i(k)} = 1$, we have $\Gamma_{\lambda,i} = \Gamma^{|(1-\lambda^{-1})\widetilde{x}_{i(k)}+\lambda^{-1}|} = \Gamma$. If $x_{i(k)} = 0$, we have $\Gamma_{\lambda,i} = \Gamma^{|(1-\lambda^{-1})\widetilde{x}_{i(k)}+\lambda^{-1}|} = \Gamma^{1/|\lambda|}$.

Then the desired result follows immediately from applying Theorem 2.

Proof of Corollary 2

Proof. The proof follows from a direct adjustment of the proof of Proposition 13 in Rosenbaum (2002). For each fixed $\widetilde{\beta}, \gamma, u_{ij}, i = 1, \ldots, I$ and $j = 1, 2$, the test statistic T_{ss} is the sum of I independent random variables, where the ith variable equals d_i with probability

$$p_i = \frac{c_{i1} \exp\{(\widetilde{\beta}\ \widetilde{x}_{i(k)} + \gamma)(u_{i1} - u_{i2})\} + c_{i2}}{1 + \exp\{(\widetilde{\beta}\ \widetilde{x}_{i(k)} + \gamma)(u_{i1} - u_{i2})\}},$$

and equals 0 with probability $1 - p_i$. Note that from the proof of Theorem 2, we have $\Gamma_{\lambda,i} = \exp\{|\widetilde{\beta}\ \widetilde{x}_{i(k)} + \gamma|\}$. Following the proof of Proposition 13 in Rosenbaum (2002), the upper bound distribution $\mathrm{pr}(\sum_{i=1}^{I} \widetilde{T}_{\Gamma,\lambda,i} \geq t \mid \mathcal{F}, \mathcal{Z})$ is the distribution of T_{ss} when $u_{ij} = c_{ij}$ if $\widetilde{\beta}\ \widetilde{x}_{i(k)} + \gamma \geq 0$ and $u_{ij} = 1 - c_{ij}$ if $\widetilde{\beta}\ \widetilde{x}_{i(k)} + \gamma < 0$, resulting in the desired

$$\widetilde{p}_{\lambda,i} = \begin{cases} 0 & \text{if } c_{i1} = c_{i2} = 0, \\ 1 & \text{if } c_{i1} = c_{i2} = 1, \\ \dfrac{\exp\{|\widetilde{\beta}\ \widetilde{x}_{i(k)} + \gamma|\}}{1 + \exp\{|\widetilde{\beta}\ \widetilde{x}_{i(k)} + \gamma|\}} & \text{if } c_{i1} \neq c_{i2}. \end{cases}$$

$$= \begin{cases} 0 & \text{if } c_{i1} = c_{i2} = 0, \\ 1 & \text{if } c_{i1} = c_{i2} = 1, \\ \dfrac{\Gamma_{\lambda,i}}{1 + \Gamma_{\lambda,i}} & \text{if } c_{i1} \neq c_{i2}. \end{cases}$$

Applying Theorem 2, the inequality $\mathrm{pr}(\sum_{i=1}^{I} \widetilde{T}_{\Gamma,\lambda,i} \geq t \mid \mathcal{F}, \mathcal{Z}) \leq \mathrm{pr}(\sum_{i=1}^{I} \overline{T}_{\Gamma,i} \geq t \mid \mathcal{F}, \mathcal{Z})$ holds for all t, $\Gamma > 1$ and $\lambda \in \mathbb{R}$.

References

Byrd, A. L. and Manuck, S. B. (2014). MAOA, childhood maltreatment, and antisocial behavior: Meta-analysis of a gene-environment interaction. *Biological Psychiatry* **75**, 9–17.

Carver, C. S., Johnson, S. L., McCullough, M. E., Forster, D. E. and Joormann, J. (2014). Adulthood personality correlates of childhood adversity. *Frontiers in Psychology* **5**, 1357.

Caspi, A., McClay, J., Moffitt, T. E., Mill, J., Martin, J., Craig, I. W. et al. (2002). Role of genotype in the cycle of violence in maltreated children. *Science* **297**, 851–854.

Chida, Y. and Steptoe, A. (2009). The association of anger and hostility with future coronary heart disease: A meta-analytic review of prospective evidence. *Journal of the American College of Cardiology* **53**, 936–946.

Cohen, P. L., Olson, M. A. and Fogarty, C. B. (2020). Multivariate one-sided testing in matched observational studies as an adversarial game. *Biometrika* **107**, 809–825.

DiPrete, T. A. and Gangl, M. (2004). 7. Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments. *Sociological Methodology* **34**, 271–310.

Ertefaie, A., Small, D. S. and Rosenbaum, P. R. (2018). Quantitative evaluation of the trade-off of strengthened instruments and sample size in observational studies. *Journal of the American Statistical Association* **113**, 1122–1134.

Fogarty, C. B. (2020). Studentized sensitivity analysis for the sample average treatment effect in paired observational studies. *Journal of the American Statistical Association* **115**, 1518–1530.

Fogarty, C. B. and Hasegawa, R. B. (2019). Extended sensitivity analysis for heterogeneous unmeasured confounding with an application to sibling studies of returns to education. *The Annals of Applied Statistics* **13**, 767–796.

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

Hansen, B. B. (2004). Full matching in an observational study of coaching for the sat. *Journal of the American Statistical Association* **99**, 609–618.

Heller, R., Rosenbaum, P. R. and Small, D. S. (2009). Split samples and design sensitivity in observational studies. *Journal of the American Statistical Association* **104**, 1090–1101.

Herd, P., Carr, D. and Roan, C. (2014). Cohort profile: Wisconsin longitudinal study (WLS). *International Journal of Epidemiology* **43**, 34–41.

Hsu, J. Y. and Small, D. S. (2013). Calibrating sensitivity analyses to observed covariates in observational studies. *Biometrics* **69**, 803–811.

Hsu, J. Y., Zubizarreta, J. R., Small, D. S. and Rosenbaum, P. R. (2015). Strong control of the familywise error rate in observational studies that discover effect modification by exploratory methods. *Biometrika* **102**, 767–782.

Ichino, A., Mealli, F. and Nannicini, T. (2008). From temporary help jobs to permanent employment: What can we learn from matching estimators and their sensitivity? *Journal of Applied Econometrics* **23**, 305–327.

Kaludercic, N., Carpi, A., Menabò, R., Di Lisa, F. and Paolocci, N. (2011). Monoamine oxidases (MAO) in the pathogenesis of heart failure and ischemia/reperfusion injury. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **1813**, 1323–1332.

Karmakar, B., French, B. and Small, D. S. (2019). Integrating the evidence from evidence factors in observational studies. *Biometrika* **106**, 353–367.

Kawachi, I., Sparrow, D., Spiro III, A., Vokonas, P. and Weiss, S. T. (1996). A prospective study of anger and coronary heart disease: The normative aging study. *Circulation* **94**, 2090–2095.

Korkeila, J., Vahtera, J., Korkeila, K., Kivimäki, M., Sumanen, M., Koskenvuo, K. et al. (2010). Childhood adversities as predictors of incident coronary heart disease and cerebrovascular disease. *Heart* **96**, 298–303.

McDermott, R., Tingley, D., Cowden, J., Frazzetto, G. and Johnson, D. D. (2009). Monoamine oxidase a gene (MAOA) predicts behavioral aggression following provocation. *Proceedings of the National Academy of Sciences* **106**, 2118–2123.

Myrtek, M. (2001). Meta-analyses of prospective studies on coronary heart disease, type a personality, and hostility. *International Journal of Cardiology* **79**, 245–251.

Nattino, G. and Lu, B. (2018). Model assisted sensitivity analyses for hidden bias with binary outcomes. *Biometrics* **74**, 1141–1149.

Neyman, J. S. (1990). On the application of probability theory to agricultural experiments. Essay on principles. Section 9. (Translated and edited by D. M. Dabrowska and T. P. Speed). *Statistical Science* **5**, 465–480.

Normand, S.-L. T., Landrum, M. B., Guadagnoli, E., Ayanian, J. Z., Ryan, T. J., Cleary, P. D. et al. (2001). Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: A matched analysis using propensity scores. *Journal of Clinical Epidemiology* **54**, 387–398.

Ottman, R. (1996). Gene–environment interaction: Definitions and study designs. *Preventive Medicine* **25**, 764.

Pérusse, L. and Bouchard, C. (1999). Genotype-environment interaction in human obesity/

discussion. *Nutrition Reviews* **57**, S31.

Pimentel, S. D., Kelz, R. R., Silber, J. H. and Rosenbaum, P. R. (2015). Large, sparse optimal matching with refined covariate balance in an observational study of the health outcomes produced by new surgeons. *Journal of the American Statistical Association* **110**, 515–527.

Rosenbaum, P. R. (1987). Sensitivity analysis for certain permutation inferences in matched observational studies. *Biometrika* **74**, 13–26.

Rosenbaum, P. R. (2002). *Observational Studies.* Springer.

Rosenbaum, P. R. (2004). Design sensitivity in observational studies. *Biometrika* **91**, 153–164.

Rosenbaum, P. R. (2010). *Design of Observational Studies.* Springer.

Rosenbaum, P. R. and Silber, J. H. (2009). Amplification of sensitivity analysis in matched observational studies. *Journal of the American Statistical Association* **104**, 1398–1405.

Rosenman, R. H., Brand, R. J., Sholtz, R. I. and Friedman, M. (1976). Multivariate prediction of coronary heart disease during 8.5 year follow-up in the western collaborative group study. *American Journal of Cardiology* **37**, 903–910.

Rubin, D. B. (1973). Matching to remove bias in observational studies. *Biometrics* 159–183.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* **66**, 688–701.

Silber, J. H., Lorch, S. A., Rosenbaum, P. R., Medoff-Cooper, B., Bakewell-Sachs, S., Millman, A. et al. (2009). Time to send the preemie home? Additional maturity at discharge and subsequent health care costs and outcomes. *Health Services Research* **44**, 444–463.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science* **25**, 1–21.

Stuart, E. A. and Hanna, D. B. (2013). Commentary: Should epidemiologists be more sensitive to design sensitivity? *Epidemiology* **24**, 88–89.

Zhang, B., Weiss, J., Small, D. S. and Zhao, Q. (2021). Selecting and ranking individualized treatment rules with unmeasured confounding. *Journal of the American Statistical Association* **116**, 295–308.

Zhao, Q. (2019). On sensitivity value of pair-matched observational studies. *Journal of the American Statistical Association* **114**, 713–722.

Zubizarreta, J., Lorch, S., Marshall, G., D'Apremont, I. and Tapia, J. (2016). Effect of prophylactic cpap in very low birth weight infants in South America. *Journal of Perinatology* **36**, 629–634.

Zubizarreta, J. R., Cerdá, M. and Rosenbaum, P. R. (2013). Effect of the 2010 chilean earthquake on posttraumatic stress reducing sensitivity to unmeasured bias through study design. *Epidemiology (Cambridge, Mass.)* **24**, 79–87.

Zubizarreta, J. R. and Keele, L. (2017). Optimal multilevel matching in clustered observational studies: A case study of the effectiveness of private schools under a large-scale voucher system. *Journal of the American Statistical Association* **112**, 547–560.

Siyu Heng

Graduate Group in Applied Mathematics and Computational Science, School of Arts and Sciences, University of Pennsylvania, Philadelphia, PA 19104, USA.

E-mail: siyuheng@sas.upenn.edu

Dylan S. Small

Department of Statistics, The Wharton School, University of Pennsylvania, Philadelphia, PA

19104, USA.

E-mail: dsmall@wharton.upenn.edu