Multiple Randomization Designs: Estimation and Inference with Interference

Lorenzo Masoero[1], Suhas Vijaykumar[1], Thomas S. Richardson[3], James McQueen[1], Ido Rosen[1], Brian Burdick[4], Pat Bajari[4], and Guido Imbens[2]

[1]Amazon, US
[2]Corresponding author, Graduate School of Business and Department of Economics, Stanford University, US
[3]Department of Statistics, University of Washington, US
[4]Work done while at Amazon, US

December 2, 2025

Abstract

Completely randomized experiments, originally developed by Fisher and Neyman in the 1930s, are still widely used in practice, even in online experimentation. However, such designs are of limited value for answering standard questions in marketplaces, where multiple populations of agents interact strategically, leading to complex patterns of spillover effects. In this paper, we derive the finite-sample properties of tractable estimators for "Simple Multiple Randomization Designs" (SMRDs), a new class of experimental designs which account for complex spillover effects in randomized experiments. Our derivations are obtained under a natural and general form of cross-unit interference, which we call "local interference." We discuss the estimation of main effects, direct effects, and spillovers, and present associated central limit theorems.

Keywords: Experimental Design, Randomization Inference, Spillovers, Marketplaces

1 Introduction

Randomized experiments, introduced in the 1920s [Neyman, 1923, Fisher, 1937], are an indispensable tool for estimating causal effects across many disciplines. For example, the Food and Drug Administration in the United States requires such experiments as part of the drug approval process. Recently, online experimentation has also become an integral part of product development in the private sector. Gupta et al. [2019] list some online businesses that collectively run hundreds of thousands of experiments annually.

Modern experimental contexts differ markedly from those that inspired early experimental designs: experiments are carried out in marketplaces, often online, where multiple populations of units interact strategically (*e.g.*, buyers and sellers; riders and drivers; renters and property managers; viewers, content creators and advertisers). A challenge posed by these settings is that cross-unit interactions often lead to interference or spillovers. In our running example of buyers and sellers, the treatment assigned to one unit in a population (*e.g.*, a seller) might affect the outcome for a different unit of the same population (another seller). If present, this interference invalidates conventional analyses of standard experimental designs.

We study the finite sample properties of tractable estimators for a new class of experimental designs, "Multiple Randomization Designs" (MRDs for short), which are tailored for experimentation in marketplaces [Bajari et al., 2023, Johari et al., 2022]. The distinguishing feature of these designs is that they involve multiple populations of units: treatment assignments and outcomes are measured at the level of a tuple of units, one from each population (*e.g.*, the impact of providing additional information on a buyer's past expenditure to a seller, measured at the buyer-seller level). The experimental designs correspond to distributions of assignments for these tuples of units, *e.g.,* over the buyer-seller pairs.

In the leading case we consider, a "Simple MRD" or SMRD, a subset of buyers is selected at random, and a subset of sellers is selected at random, and only the buyer-seller pairs where both the buyer and the seller was selected are exposed to the binary treatment. This two-level randomization serves to isolate and measure interference between units, thus distinguishing it from classical designs with multi-level randomization (*e.g.*, Latin square and split-plot designs).

This paper provides the first formal analysis of MRDs, showing that they may be used to (i) test for the presence of spillovers, (ii) estimate and conduct inference for the overall treatment effect in the presence of a large class of spillovers, and (iii) obtain—even without interference or spillovers—more precise estimates of the average causal effect than standard, single-sided randomization designs.

Our work contributes to the rapidly growing literature on causal inference under interference [Hong and Raudenbush, 2006, 2008, Hudgens and Halloran, 2008, Rosenbaum, 2007, Aronow, 2012, VanderWeele et al., 2014, Ogburn and VanderWeele, 2014, Athey et al., 2018, Ugander et al., 2013, Blake and Coey, 2014, Basse et al., 2019]. Recent research has focused on experimental design in settings with complex spillovers, differing mainly in the settings they consider and the corresponding assumptions placed on cross-unit interference. Some work considers cases where spillovers between units are mediated by low-dimensional measures, such as prices in a marketplace or shares of treated units in a peer group [Wager and Xu, 2021, Munro et al., 2021, Aronow and Samii, 2017]. Another line of work focuses on the role of clustering to mitigate interference, *e.g.*, Viviano et al. [2023]. A separate approach models interference in terms of a bipartite graph between units and treatment sites (*e.g.*, advertisers bidding on the same keywords, as in Zigler and Papadogeorgou, 2021, and Harshaw et al., 2022). Others consider crossover or switchback designs in dynamic contexts where treatments vary over time and have lasting effects [Cox and Reid, 2000, Bojinov et al., 2020, Xiong et al., 2023, Shi and Ye, 2023]. Finally, some work has modeled spatial or network spillovers in order to improve precision in survey experiments [Verbitsky-Savitz and Raudenbush, 2009].

Multiple Randomization Designs were informally introduced by Bajari et al. [2023] and Johari et al. [2022]. The key feature of MRDs is the presence of two or more populations, *e.g.* buyers and sellers, where interventions can be assigned and outcomes measured at the level of the buyer-seller pair. We provide exact characterizations of the design-based variance, together with corresponding variance estimators, and central limit theorems that allow for inference under these designs.

On the surface, MRDs share common features with Latin squares [Welch, 1937] and split-plot designs [Fisher, 1928, Zhao et al., 2018, Zhao and Ding, 2022], but they are fundamentally quite different. In all three cases, the experimental units are organized in a matrix or clustered structure. However, they differ in important ways: for example, Latin square designs are aimed at reducing variance through balance of the location of experimental units in a geographic space, whereas MRDs address interference and spillovers between experimental units. Meanwhile, although split-plot designs have been used to study spillovers (e.g., Hudgens and Halloran 2008, Zhao and Ding 2022), they consider units grouped into clusters as opposed to a two-dimensional array.

Multiple randomization allows us to account for interference in ways not possible with completely randomized experiments, but in doing so they complicate estimation and inference. Challenges arise from the intrinsic dependence structure in the assignment process across the two populations: buyers and sellers in our generic example. We address these using a randomization-based approach, where we take the potential outcomes under different treatment regimes as fixed. We exactly characterize the finite-sample variances of the proposed estimators with respect to the random design. We also propose conservative variance estimators, similar to those available for conventional randomized experiments. Finally, we prove design-based central limit theorems, extending the recent results of Li and Ding [2017a], Shi and Ding [2022a] for single population experiments to our setting with multiple-population experiments, under appropriate side assumptions.

Most similar to our work is Johari et al. [2022], who studied how spillover effects caused by interference can lead to bias in standard experimental designs, and analyzed a special case of the MRDs we consider in this paper. Johari et al. [2022] produce a dynamic, stochastic model of a two-sided marketplace with cross-unit interference. Following a detailed analysis of the model, they use it to illustrate the favorable properties of SMRDs in comparison to standard experimental designs.

2 Experiments in Marketplaces: Interference

We start by introducing a framework for randomized experiments in marketplaces with multiple populations of agents. We use the two-population buyer-seller (or customer-product) case as our generic example, but we emphasize that the ideas we present apply to other settings and extend to higher-order unit tuples, *e.g.*, subscriber-creator-advertiser, customer-restaurant-driver or passenger-airline-travel agent. Interference or spillover effects arise naturally in these settings: treatment of one unit can impact the outcomes of other units, invalidating assumptions that serve as the basis for analyzing standard experiments. An example of the treatment is the presentation of additional information (*e.g.,* in the form of more detailed reviews) shown to buyer i when viewing products from seller j.

In our generic buyer-seller example, one of the populations consists of I buyers, indexed by $i \in [I] := \{1, \ldots, I\}$. The buyers interact with members of the second population, consisting of J sellers indexed by $j \in [J] := \{1, \ldots, J\}$. Over a fixed period of time, say a week or a month, we measure for each buyer-seller pair an outcome metric of engagement Y_{ij} (*e.g.*, the amount of money paid by buyer i to seller j). The experimenter performs an intervention at the level of the buyer-seller pair (i, j), via the randomized treatment assignment $W_{ij} \in \{\mathrm{C}, \mathrm{T}\}$. Critically, the treatment might not be offered to all buyers who interact with a particular seller, nor to all sellers for any given buyer. Let $\mathbf{W} \in \{\mathrm{C}, \mathrm{T}\}^{I \times J}$ denote a random $I \times J$ matrix of treatment assignments with typical element $W_{ij} \in \{\mathrm{C}, \mathrm{T}\}$, and $\mathbf{w}$ a realization of this matrix.

We adopt the potential outcome framework [see e.g., Fisher, 1937, Neyman, 1923/1990, Imbens and Rubin, 2015]: for each value $\mathbf{w}$ of the $I \times J$ assignment matrix, $y_{ij}(\mathbf{w})$ is the corresponding potential outcome for unit (i, j), which is non-stochastic. An example assignment matrix is shown in (1), where rows identify five buyers and columns identify six sellers. Colors highlight four sets of experimental units (buyer-seller pairs), instead of the usual two. Three of these groups (pink, blue, and yellow) are assigned to the control treatment, and are differentiated by the fraction of "neighboring" buyer-seller pairs—units in the same row or column—which are assigned to treatment. For example, pairs (i, j) colored in pink are exposed to control, with 4/4 of units (i', j), $i' \neq i$ in the same column also exposed to control, and 5/6 of units (i, j'), $j' \neq j$ in the same row exposed to treatment. Similarly, yellow pairs are also exposed to control, again with 4/4 of units in the same column also exposed to control, but only 2/6 of units in the same row exposed to treatment. Prior to intervention, all four groups are comparable due to randomization. After the intervention, even the three groups of control units need not be comparable: their outcomes might differ systematically due to spillovers.

$$
\begin{array}{rl}
 & \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{Buyers} : i \\
\text{Sellers} : j \rightarrow & \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \quad \downarrow \\
\mathbf{W} = & \begin{pmatrix} \mathrm{C} & \mathrm{C} & \mathrm{T} & \mathrm{C} & \mathrm{T} & \mathrm{C} \\ \mathrm{C} & \mathrm{C} & \mathrm{T} & \mathrm{C} & \mathrm{T} & \mathrm{C} \\ \mathrm{T} & \mathrm{T} & \mathrm{T} & \mathrm{C} & \mathrm{T} & \mathrm{C} \\ \mathrm{C} & \mathrm{C} & \mathrm{T} & \mathrm{C} & \mathrm{T} & \mathrm{C} \\ \mathrm{T} & \mathrm{T} & \mathrm{T} & \mathrm{C} & \mathrm{T} & \mathrm{C} \end{pmatrix} \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array}
\end{array} \tag{1}
$$

Consider the buyer-seller example in which for pairs of buyers and sellers assigned to the treatment group the buyer gets to see more information about the seller or product in the form of additional reviews. Buyer 1 gets the additional information when interacting with sellers 3 and 5, but not when interacting with sellers 1, 2, 4 and 6. If the information is generally helpful, this may lead buyer 1 to switch engagement from sellers 1, 2, 4 and 6 to sellers 3 and 5, a common form of spillover. Sellers 3 and 5 are in the treatment group for all buyers. If the information raises the engagement with those sellers relative to, say sellers 4 and 6 who are always in the control group, this may lead sellers 3 and 5 to change other behaviors, such as their marketing strategy, leading to a different type of spillover.

Formally, spillovers are present whenever potential outcomes $y_{ij}(\mathbf{w})$ and $y_{ij}(\mathbf{w}')$ differ for assignments $\mathbf{w}$ and $\mathbf{w}'$ where the treatment for the pair (i, j) is identical, $w_{ij} = w'_{ij}$, but some other elements of the assignment matrices $\mathbf{w}$ and $\mathbf{w}'$ differ. Obtaining unbiased estimates of causal effects in the presence of spillovers is challenging: classical causal analyses typically impose strong assumptions that rule out any form of cross-unit interference (*e.g.*, the stable unit value assumption or SUTVA, Rubin, 1974).

We now introduce different assumptions on the potential outcomes, leading to different structures for the interference. We later discuss in section 3 how alternative forms of interference can be effectively addressed using specific experimental designs. The simplest possibility is to rule out *any* type of interference (a version of SUTVA where the experimental unit is given by a buyer-seller pair).

Assumption 2.1 (Strong No-Interference)**.** *Potential outcomes satisfy the strong no-interference assumption if* $y_{ij}(\mathbf{w}) = y_{ij}(\mathbf{w}')$*, for all* (i, j) *such that* $w_{ij} = w'_{ij}$*.*

Under assumption 2.1, a natural approach is to randomize all pairs, subject to treatment balance within buyers and sellers. This generally allows for more efficient estimation than designs which randomize only buyers or only sellers.

A natural way to weaken assumption 2.1 is to allow the outcome for a given buyer-seller pair to additionally depend on the treatment assignments involving the same buyer but different sellers (but not to depend on the assignments received by other buyers). Let $\mathbf{w}, \mathbf{w}'$ be assignment matrices where the treatment for the pair (i, j) coincides, so $w_{ij} = w'_{ij}$, but there is a seller j' for which $w_{ij'} \neq w'_{ij'}$. Under this type of interference, it may be that $y_{ij}(\mathbf{w}) \neq y_{ij}(\mathbf{w}')$. However, for any assignment $\mathbf{w}''$ with $w''_{ij'} = w_{ij'}, \forall j' \in [J]$, $y_{ij}(\mathbf{w}) = y_{ij}(\mathbf{w}'')$. We formalize this form of interference in assumption 2.2.

Assumption 2.2 (No-Interference for Buyers)**.** *Potential outcomes satisfy the no-interference for buyers assumption if* $y_{ij}(\mathbf{w}) = y_{ij}(\mathbf{w}')$ *for all* (i, j) *such that* $w_{ij'} = w'_{ij'}$ *for all* $j' \in [J]$*.*

Under assumption 2.2, changing one or more of the treatment assignments for a different buyer i' does not change the outcomes for buyer-seller pair (i, j). But, changing one or more of the treatments for a different seller j' may affect the outcome y_{ij}. Under this assumption a buyer-randomized experiment, corresponding to the matrix assignment later introduced in eq. (3), is a natural strategy. Similarly, a seller-randomized experiment is natural if we expect the following "no-interference for sellers" assumption to hold.

Assumption 2.3 (No-Interference for Sellers)**.** *Potential outcomes satisfy the no-interference for sellers assumption if* $y_{ij}(\mathbf{w}) = y_{ij}(\mathbf{w}')$ *for all* (i, j) *such that* $w_{i'j} = w'_{i'j}$ *for all* $i' \in [I]$*.*

Next, we consider an assumption first introduced in Bajari et al. [2023] that allows for some forms of interference across both buyers and sellers. This is a key assumption in our paper. It attempts to balance competing interests: allowing for a substantial degree of interference and at the same time imposing enough structure so that questions of interest are answerable.

Assumption 2.4 (Local Interference)**.** *Potential outcomes satisfy the local interference assumption if* $y_{ij}(\mathbf{w}) = y_{ij}(\mathbf{w}')$*, for any pair* (i, j)*, such that (a) the assignments for the pair* (i, j) *coincide,* $w_{ij} = w'_{ij}$*, (b) the fraction of treated sellers for buyer* i *coincide under* $\mathbf{w}$ *and* $\mathbf{w}'$*, and (c) the fraction of treated buyers for seller* j *coincide under* $\mathbf{w}$ *and* $\mathbf{w}'$*.*

Consider the following two assignment matrices $\mathbf{w}, \mathbf{w}'$:

$$\mathbf{w} = \begin{pmatrix} \text{C} & \text{T} & \text{C} & \text{C} & \text{C} \\ \text{T} & \text{C} & \text{T} & \text{C} & \text{T} \\ \text{T} & \text{C} & \text{T} & \text{T} & \text{T} \\ \text{C} & \text{C} & \text{C} & \text{C} & \text{C} \end{pmatrix}, \ \mathbf{w}' = \begin{pmatrix} \text{C} & \text{T} & \text{T} & \text{C} & \text{C} \\ \text{C} & \text{T} & \text{C} & \text{C} & \text{C} \\ \text{T} & \text{C} & \text{T} & \text{T} & \text{T} \\ \text{C} & \text{C} & \text{C} & \text{T} & \text{T} \end{pmatrix}.$$

Under local interference, the outcome for buyer-seller pair $(3, 3)$ must be identical for the assignment matrices $\mathbf{w}$ and $\mathbf{w}'$ (that is, $y_{33}(\mathbf{w}) = y_{33}(\mathbf{w}')$), because (a) the (3,3) elements of $\mathbf{w}$ and $\mathbf{w}'$ are identical, and (b) the third columns of the assignment matrices (given in purple) have the same fraction of treated pairs (1/2), and (c) the third rows of the assignment matrices (also given in purple) have the same fraction of treated pairs (4/5).

Although obviously weaker than assumption 2.1 which rules out all interference, and more flexible than assumption 2.2 which rules out interference between buyers while allowing for interference within sellers, local interference does still substantially restrict the possible forms of interference between units. In particular, for a given unit pair (i, j) only $I + J - 1$ of the total IJ unit-level assignments defining $\mathbf{w}$ are relevant to the realized outcome: those of pairs (i, j') and (i', j). Further, the unit-level outcome is a function of only three sufficient statistics: the unit's own treatment assignment (w_{ij}), and the averages of the (same) row and column to which the pair belongs ($\sum_{i'} w_{i'j}/I$, and $\sum_{j'} w_{ij'}/J$).

Similar forms of interference were previously proposed by Manski [2013] (cf. "anonymous interactions") and Hudgens and Halloran [2008] (cf. "stratified interference"). Despite its simplicity, we believe that this assumption is a natural starting point for approximating many types of interference that arise due to strategic behavior in a two-sided market. To illustrate, we now provide a simple example of a two-sided marketplace in which—at Nash equilibrium—potential outcomes exhibit both buyer and seller interference, and satisfy local interference. Later we show that under some designs, including the leading Simple MRD, local interference has no testable implications. We also show in section 6 that more complex MRDs do lead to testable implications on the conditional expectations (over treatment assignments) of the outcomes.

Example 2.5. Consider a two-sided platform where content creators $i \in [I]$ and advertisers $j \in [J]$ interact. Each content creator i produces corresponding content with score q_i^c, and each advertiser places ads with corresponding advertisement quality q_j^a. In this model, each creator-advertiser pair generates revenue y_{ij}. In the absence of any intervention, revenue generated by (i, j) is given by

$$y_{ij} = m_{ij}\{q_i^c + q_j^a\},$$

where the (fixed) scalar factor $m_{ij} \in \mathbb{R}$ reflects the compatibility between i and j (e.g., footwear ads might have higher compatibility with content produced by a creator focusing on sports). Creators and advertisers are compensated by the platform according to a contract: for each pair (i, j), creator i is compensated $r_i^c y_{ij}$ and advertiser j is compensated $r_j^a y_{ij}$, and the platform keeps $(1 - r_i^c - r_j^a)y_{ij}$; the platform negotiates r_i^c, r_j^a with each creator and advertiser. In practice, generating high-quality content requires costly effort. In particular, we suppose that both creators and advertisers maximize their total compensation minus the cost of effort:

$$U_i^c = \left(\sum_{j=1}^{J} r_i^c y_{ij}\right) - \frac{(q_i^c)^2}{2}, \qquad \text{and} \qquad U_j^a = \left(\sum_{i=1}^{I} r_j^a y_{ij}\right) - \frac{(q_j^a)^2}{2}.$$

In the static Nash equilibrium, each creator and advertiser solves the maximization problem treating the other agents' inputs q_i^c, q_j^a as fixed and known. This leads to the equilibrium actions

$$q_i^c = \sum_{j=1}^{J} r_i^c y_{ij}, \quad \text{and} \quad q_j^a = \sum_{i=1}^{I} r_j^a y_{ij}.$$

The platform hosting the content creators and advertisers tests the impact of a subsidy via a binary intervention $\mathbf{w}$ affecting the revenue as follows:

$$y_{ij}(\mathbf{w}) = (m_{ij} + \eta w_{ij})\{q_i^c(\mathbf{w}) + q_j^a(\mathbf{w})\}.$$

Here, for $\eta \in \mathbb{R}$, the factor $\eta w_{ij} \in \{0, \eta\}$ represents an extra incentive paid by the platform (η is the incentive, and w_{ij} is a binary treatment variable). Notice that each agent's incentives depends on the average treatment status of their interactions. This influences their action, which creates precisely a local interference structure. At Nash equilibrium, the revenue y_{ij} and profit π_{ij} both satisfy local interference; they are given by

$$\begin{aligned} y_{ij}(\mathbf{w}) &= (m_{ij} + \eta w_{ij}) \left[J r_i^c \left(\bar{m}_i^c + \eta \bar{w}_i^c\right) + I r_j^a \left(\bar{m}_j^a + \eta \bar{w}_j^a\right)\right], \\ \pi_{ij}(\mathbf{w}) &= \{(1 - r_i^c - r_j^a) m_{ij} - (r_i^c + r_j^a)\eta w_{ij})\} \left[J r_i^c \left(\bar{m}_i^c + \eta \bar{w}_i^c)\right) + I r_j^a \left(\bar{m}_j^a + \eta \bar{w}_j^a\right)\right], \end{aligned} \tag{2}$$

where $\bar{w}_i^c = \frac{1}{J}\sum_{j'=1}^{J} w_{ij'}$, and $\bar{m}_i^c = \frac{1}{J}\sum_{j'=1}^{J} m_{ij'}$ and $\bar{m}_j^a$, $\bar{w}_j^a$ are defined symmetrically.

Example 2.5 shows a two-sided-marketplace with strategic agents in which agents' equilibrium actions lead potential outcomes (revenue or profits) to satisfy local interference (as in eq. (2)). Local interference arises somewhat naturally, as it assumes the outcome of an interaction between two agents will depend non-parametrically on the interaction-level treatment, as well as both agents' cumulative exposure to treatment. More generally, local interference may be viewed as a natural, tractable first approximation to the complex spillover effects arising in a two-sided marketplace. In section 5, we simulate the above example to show that agents' strategic responses can lead to large spillover effects, which are neglected by traditional designs. In this way, our results are closely related to but distinct from the work of Munro et al. [2021] on treatment effects in market equilibrium: for example, the above Nash equilibrium in a finite marketplace is not captured by that work. It is also related to the works of Harshaw et al. [2022] and Aronow and Samii [2017] in that potential outcomes depend on low-dimensional measures of "exposure," though distinct in that we place agents on both sides—as opposed to one side —of the bipartite network.

3 Multiple Randomization Designs

Multiple Randomization Designs (MRDs) are a generalization of standard A/B tests to allow for spillover effects common in marketplaces [Bajari et al., 2023, Johari et al., 2022].

These designs can provably detect and measure spillover effects of the type introduced in section 2, as we will discuss in section 4. Let $\mathbb{W}$ denote the set of 2^{IJ} values that the random binary assignment matrix $\mathbf{W}$ can take. We now formally define MRDs.

Definition 3.1 (Multiple Randomization Designs)**.** *A Multiple Randomization Design (MRD) is a probability distribution over* $\mathbb{W}$, $p : \mathbb{W} \mapsto [0, 1)$, *such that (i)* $p(\cdot)$ *is row and column exchangeable, and (ii) there exists* $\overline{\overline{w}} \in (0, 1)$ *such that for any* $\boldsymbol{w} = (w_{ij}) \in \{0, 1\}^{I \times J}$ *in the support of* p,

$$\frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} \mathbf{1}(w_{ij} = \mathrm{T}) = \overline{\overline{w}}.$$

Note that a probability distribution $p(\cdot)$ over matrices $\mathbf{w}$ is said to be row (or column) exchangeable if, under $p(\cdot)$, any two assignments which differ by a permutation of the rows (or columns) are assigned the same probability. By imposing exchangeability of $p(\cdot)$ through definition 3.1(i) we rule out the possibility of degenerate experiments in which a single value $\mathbf{w}$ has probability one. Condition 3.1(ii) ensures that all assignments with positive probability have the same fraction $\overline{\overline{w}}$ of treated buyer-seller pairs. It is not strictly necessary, but it helps us to derive exact finite-sample results in section 4, clarifying what can be learned without large sample approximations.

Given an assignment matrix $\mathbf{w}$, for each buyer i let $\overline{w}_i^{\mathrm{B}}$ be the fraction of sellers j for which (i, j) received the treatment, and let $\overline{w}_j^{\mathrm{S}}$ be the symmetric quantity for seller j:

$$\overline{w}_i^{\mathrm{B}} := \sum_{j=1}^{J} \frac{\mathbf{1}(w_{ij} = \mathrm{T})}{J}, \qquad \text{and} \qquad \overline{w}_j^{\mathrm{S}} := \sum_{i=1}^{I} \frac{\mathbf{1}(w_{ij} = \mathrm{T})}{I}.$$

Definition 3.1 implies that $\overline{\overline{w}} = \sum_i \overline{w}_i^{\mathrm{B}}/I = \sum_j \overline{w}_j^{\mathrm{S}}/J$. A key feature of an MRD is that it allows both buyers and sellers to be exposed to different treatments within the same experiment. We refer to the presence of such variation in the assignment as *inhomogeneity* of the buyer or seller experience.

Definition 3.2 (Homogeneous and Inhomogeneous Experiences)**.** *Assignment* $\mathbf{w}$ *induces a homogeneous experience for buyer* i *if* $\overline{w}_i^{\mathrm{B}} \in \{0, 1\}$, *and an inhomogeneous experience for buyer* i *if* $\overline{w}_i^{\mathrm{B}} \in (0, 1)$. *Similarly, it induces a homogeneous experience for seller* j *if* $\overline{w}_j^{\mathrm{S}} \in \{0, 1\}$ *and an inhomogeneous experience for seller* j *if* $\overline{w}_j^{\mathrm{S}} \in (0, 1)$.

In assignment matrix (1), sellers 3, 4, 5 and 6 have a homogeneous experience while sellers 1 and 2 and all buyers have an inhomogeneous experience. Inhomogeneous experiences are at the heart of spillover concerns in our set-up. Suppose that the treatment corresponds to offering more information to some buyer-seller pairs. Buyers with an inhomogeneous experience may shift their engagement from sellers in the control group to sellers in the treatment group, without changing their overall engagement or expenditure.

Next, we showcase the flexibility of MRDs by defining three classes of experimental designs that fit within the general Definition 3.1. These three classes do not exhaust the possibilities, but make specific points: they show that MRDs (i) encompass standard experimental designs, (section 3.1), (ii) can increase efficiency (section 3.2) and (iii) most importantly, in certain cases can answer questions that standard designs cannot answer, as we discuss in section 3.3. We conclude the section by discussing connections between these designs and the local interference assumption introduced in section 2.

3.1 Single Randomization Designs

A Single Randomization Design (SRD) is an MRD where each buyer or seller has a homogeneous experience with probability one: i.e. a buyer experiment ($\overline{w}_i^{\mathrm{B}} \in \{0,1\}$ and $\overline{w}_j^{\mathrm{S}} = \overline{\overline{w}}$), or a seller experiment ($\overline{w}_i^{\mathrm{B}} = \overline{\overline{w}}$ and $\overline{w}_j^{\mathrm{S}} \in \{0,1\}$). A buyer experiment is a simple buyer-randomized A/B test, where assignment matrices are of the form of (3), with identical columns and constant rows:

$$\mathbf{w} = \begin{pmatrix} \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} \\ \mathrm{T} & \mathrm{T} & \mathrm{T} & \mathrm{T} & \mathrm{T} & \mathrm{T} & \mathrm{T} & \mathrm{T} \\ \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} \\ \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} \end{pmatrix}. \tag{3}$$

Here buyers $1,3,4$ are in the control group, and buyer 2 is in treatment. All buyers here have a homogeneous experience, whereas none of the sellers have a homogeneous experience.

3.2 Crossover Designs

In contrast to standard (buyer or seller) experiments, MRDs include experiments in which neither all buyers nor all sellers have homogeneous experiences. The simplest such an MRD is one in which all interactions (i,j) are randomly assigned. This design is widely used in settings where the second dimension is time, and where such designs have been referred to as rotation experiments [Cochran, 1939], crossover experiments [Brown Jr, 1980], or switchback experiments [Bojinov et al., 2020], although it is not limited to settings where time is one of the dimensions. An example is given in assignment matrix (4):

$$\mathbf{w} = \begin{array}{c} \text{Time Period} \\ \text{Individual Unit} \downarrow \hookrightarrow \end{array} \begin{array}{c|cccccccc|} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & \mathrm{T} & \mathrm{T} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{T} & \mathrm{C} & \mathrm{T} \\ 2 & \mathrm{T} & \mathrm{T} & \mathrm{T} & \mathrm{C} & \mathrm{T} & \mathrm{C} & \mathrm{C} & \mathrm{C} \\ 3 & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{T} & \mathrm{T} & \mathrm{T} & \mathrm{C} & \mathrm{T} \\ 4 & \mathrm{C} & \mathrm{C} & \mathrm{T} & \mathrm{C} & \mathrm{C} & \mathrm{T} & \mathrm{T} & \mathrm{T} \\ 5 & \mathrm{C} & \mathrm{T} & \mathrm{C} & \mathrm{T} & \mathrm{T} & \mathrm{C} & \mathrm{T} & \mathrm{C} \\ 6 & \mathrm{T} & \mathrm{C} & \mathrm{T} & \mathrm{T} & \mathrm{C} & \mathrm{C} & \mathrm{T} & \mathrm{C} \end{array}. \tag{4}$$

In assignment matrix (4) we consider a balanced design, where each unit is in the treatment group for four periods, and in every period exactly three units are in the treatment group. It is particularly attractive in settings where strong no-interference is reasonable (assumption 2.1), where, under additional assumptions on the potential outcomes, it can be shown to improve efficiency [Masoero et al., 2023].

Remark 3.3. A related experimental design is that of staggered adoption, or the "stepped wedge" design. There, units are assigned to the treatment at different points in time, but once assigned to the treatment they never exit. See Athey and Imbens [2022], Hemming et al. [2015] for analyses of these experiments, and Xiong et al. [2023] for optimal design.

3.3 Simple Multiple Randomization Designs

The next design we consider introduces systematic variation in $\overline{w}_i^{\mathrm{B}}$ over buyers and variation in $\overline{w}_j^{\mathrm{S}}$ over sellers. Such variation allows for the detection of spillovers, as well as for estimation of their magnitude. To accomplish this goal, we randomize buyers and sellers separately: we select at random I_T buyers, with $1 < I_T < I - 1$ and assign them $W_i^{\mathrm{B}} = 1$. For the remaining buyers, $W_i^{\mathrm{B}} = 0$, so that we have a buyer-assignment random vector $\vec{\boldsymbol{W}}^{\mathrm{B}} \in \{0,1\}^I$ with $\sum_i W_i^{\mathrm{B}} = I_T$. Symmetrically, we select J_T sellers at random, with $1 < J_T < J - 1$ and assign them $W_j^{\mathrm{S}} = 1$. The remaining sellers are assigned $W_j^{\mathrm{S}} = 0$, yielding a seller-assignment random vector $\vec{\boldsymbol{W}}^{\mathrm{S}} \in \{0,1\}^J$ with $\sum_j W_j^{\mathrm{S}} = J_T$. Then the assignment for the pair (i,j) is a function of the buyer and seller assignments W_i^{B} and W_j^{S}.

Definition 3.4 (Simple Multiple Randomization Designs)**.** *Given a population of I buyers and J sellers, a Simple Multiple Randomization Design (SMRD) is an MRD in which, for fixed proportions $\mathrm{p}^{\mathrm{B}} = I_T/I \in (0,1)$ and $\mathrm{p}^{\mathrm{S}} = J_T/J \in (0,1)$, we randomly assign to each buyer $W_i^{\mathrm{B}} \in \{0,1\}$ such that $\sum_i W_i^{\mathrm{B}} = I_T$, and independently randomly assign each seller $W_j^{\mathrm{S}} \in \{0,1\}$ such that $\sum_j W_j^{\mathrm{S}} = J_T$. The pair (i,j) is exposed to treatment via*

$$W_{ij} = \begin{cases} \mathrm{T} & \text{if } \min\left(w^{\mathrm{B}}, w^{\mathrm{S}}\right) = 1, \\ \mathrm{C} & \text{otherwise.} \end{cases} \tag{5}$$

While SMRDs do not have the richness of the full class of MRDs, they contain many of the insights that apply to the general case. This special case of MRDs has also been discussed in Johari et al. [2022], where the focus is on the bias of the difference in means estimator for the average treatment effect. See also Bajari et al. [2023], Li et al. [2021].

An assignment example for an SMRD is given in matrix (6), where the buyer-assignment vector $\vec{\boldsymbol{w}}^{\mathrm{B}} = [0,0,1,1]$ and seller-assignment vector $\vec{\boldsymbol{w}}^{\mathrm{S}} = [0,0,0,0,1,1,1,1]$ lead to:

$$\mathbf{w} = \begin{pmatrix} \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} \\ \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} \\ \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{T} & \mathrm{T} & \mathrm{T} & \mathrm{T} \\ \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{T} & \mathrm{T} & \mathrm{T} & \mathrm{T} \end{pmatrix}. \tag{6}$$

In these SMRDs, the pairs of binary values $(w_i^{\mathrm{B}}, w_j^{\mathrm{S}})$ induce four assignment types of buyer-seller pairs (each type identified by a different color in the assignment matrix (6)):

$$\gamma_{ij} = \begin{cases} \mathrm{cc} & \text{if } w_i^{\mathrm{B}} = 0, w_j^{\mathrm{S}} = 0 \text{ (so } w_{ij} = 0), \\ \mathrm{ib} & \text{if } w_i^{\mathrm{B}} = 1, w_j^{\mathrm{S}} = 0 \text{ (so } w_{ij} = 0), \\ \mathrm{is} & \text{if } w_i^{\mathrm{B}} = 0, w_j^{\mathrm{S}} = 1 \text{ (so } w_{ij} = 0), \\ \mathrm{tr} & \text{if } w_i^{\mathrm{B}} = 1, w_j^{\mathrm{S}} = 1 \text{ (so } w_{ij} = 1). \end{cases} \tag{7}$$

Here, cc is "homogeneous control", ib "inhomogeneous buyer control", is "inhomogeneous seller control", and tr "treated". Consistent with eq. (5), $w_{ij} = \mathrm{T}$ if $\gamma_{ij} = \mathrm{tr}$ and $w_{ij} = \mathrm{C}$ otherwise. The values w_i^{B} and w_j^{S} can be inferred from the assignment matrix $\mathbf{w}$, hence the type can be inferred from the assignment matrix, $\gamma_{ij} = \gamma_{ij}(\mathbf{w})$. These assignment types play an important role under the local interference assumption (2.4), as highlighted in lemma 3.5.

Lemma 3.5. *For* $\mathbf{w}, \mathbf{w}'$ *consistent with an SMRD and assuming that potential outcomes satisfy local interference (assumption 2.4), potential outcomes can be written as a function of the assignment types only: for* $\mathbf{w}, \mathbf{w}'$ *it holds that*

$$\gamma_{ij}(\mathbf{w}) = \gamma_{ij}(\mathbf{w}') \Rightarrow y_{ij}(\mathbf{w}) = y_{ij}(\mathbf{w}').$$

This simplification, where potential outcomes depend only on a function of their original argument, is related to the exposure mapping concept in Aronow and Samii [2017].

Of the four groups of buyer-seller pairs induced by an SMRD—all of which are comparable prior to treatment due to physical randomization—types cc, ib, is are all exposed to control. Having multiple sets of pairs which are (i) comparable prior to treatment, (ii) all exposed to the same treatment (control) and (iii) not comparable post-treatment, gives SMRDs the ability to detect interference. This ability is based on comparisons of average outcomes for these three groups in which pairs are all exposed to control. Under a simple buyer or seller experiment, where only two types are present, and only one is exposed to the control treatment, spillovers could not be detected.

An interesting feature of the SMRD is that the local interference assumption is not testable here: differences between expected outcomes for the comparison groups can always be rationalized in a way that is consistent with local interference. We observe outcomes for four types of pairs, $\gamma_{ij} \in \{\text{cc}, \text{ib}, \text{is}, \text{tr}\}$. The local interference assumption does not restrict the distribution of the outcomes for these four types. In contrast the no-interference for buyers assumption, assumption 2.2, does have testable implications in settings with a large number of buyers and sellers: it would imply that the distribution of outcomes for the cc pairs is the same as the distribution of outcomes for the ib pairs.

4 Estimation and Inference for SMRDs

We now describe methods that make the experimental designs introduced in the previous section practically useful by enabling statistical inference. Specifically, we provide five results. First, we introduce estimands and estimators for causal effects in the presence of local interference for the SMRD (section 4.1). Second, we show the proposed estimators are unbiased (section 4.2). Third, we characterize the exact finite sample variance of these estimators (section 4.3). Fourth, we derive, in the tradition of the causal inference literature, conservative estimators for their variances (section 4.4). Finally, we provide central limit theorems that allow for the construction of confidence intervals (section 4.5). Proofs are deferred to the appendix. While seemingly standard, our results require a non-trivial amount of technical complexity due to the fact that randomization acts jointly on the multiple dimensions through which potential outcomes are indexed.

In what follows, for a given type γ, we let I_γ (J_γ) denote the number of buyers i (sellers j) for which there is at least one pair (i, j) such that $\gamma_{ij} = \gamma$. For example, in the assignment of eq. (6), $I_\gamma = 2$ and $J_\gamma = 4$ for all $\gamma \in \{\text{cc}, \text{ib}, \text{is}, \text{tr}\}$. This is because the first two buyershave pairs exposed to cc, is (so that $I_\text{C} = I_\text{is} = 2$) and the last two have pairs exposed to ib, tr (so that $I_\text{ib} = I_\text{tr} = 2$). A symmetric argument holds for sellers. Moreover, whenever we consider an SMRD for which local interference holds, we leverage lemma 3.5 and — with some abuse of notation — write $y_{ij}(\gamma)$ instead of $y_{ij}(\mathbf{w})$.

4.1 Causal Estimands and Spillover Effects

Under the local interference assumption 2.4, lemma 3.5 proves that the potential outcomes y_{ij} are indexed by type $\gamma_{ij} \in \{\text{cc}, \text{ib}, \text{is}, \text{tr}\}$. Define the population averages by type:

$$\bar{\bar{y}}_\gamma := \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} y_{ij}(\gamma), \text{ for } \gamma \in \{\text{cc}, \text{ib}, \text{is}, \text{tr}\}. \tag{8}$$

For $\vec{\boldsymbol{\beta}} = [\beta_{\text{cc}}, \beta_{\text{ib}}, \beta_{\text{is}}, \beta_{\text{tr}}]^\top$, we consider causal estimands that can be written as linear combinations of the $\bar{\bar{y}}_\gamma$ defined in eq. (8):

$$\tau(\vec{\boldsymbol{\beta}}) := \beta_{\text{cc}}\bar{\bar{y}}_{\text{cc}} + \beta_{\text{ib}}\bar{\bar{y}}_{\text{ib}} + \beta_{\text{is}}\bar{\bar{y}}_{\text{is}} + \beta_{\text{tr}}\bar{\bar{y}}_{\text{tr}}. \tag{9}$$

This class of estimands includes many interesting quantities that shed light on the direct effect of the treatment, the spillover effects on untreated units stemming from applying treatment to other pairs, and the total effect. For example, $\vec{\boldsymbol{\beta}}_{\text{ATE}} := [-1, 0, 0, 1]^\top$ corresponds to $\tau_{\text{ATE}} := \tau(\vec{\boldsymbol{\beta}}_{\text{ATE}}) = \bar{\bar{y}}_{\text{tr}} - \bar{\bar{y}}_{\text{cc}}$, which is the average treatment effect of assigning both buyer i and seller j to treatment versus both being assigned to control under an SMRD design. Like all other estimands in our setting, τ_{ATE} is implicitly parametrized by the fractions $p^{\text{B}} \in (0, 1)$ of treated buyers, $p^{\text{S}} \in (0, 1)$ of treated sellers. For $\vec{\boldsymbol{\beta}}^{\text{B}}_{\text{spill}} := [-1, 1, 0, 0]^\top$, $\tau^{\text{B}}_{\text{spill}} := \tau(\vec{\boldsymbol{\beta}}^{\text{B}}_{\text{spill}}) = \bar{\bar{y}}_{\text{ib}} - \bar{\bar{y}}_{\text{cc}}$ measures a "buyer"-spillover effect. If there are no spillovers within buyers (assumption 2.2), this average causal effect is equal to zero. Thus, the estimated counterpart of this estimand sheds light on the presence of buyer spillovers. Similarly, for $\vec{\boldsymbol{\beta}}^{\text{S}}_{\text{spill}} := [-1, 0, 1, 0]^\top$, $\tau^{\text{B}}_{\text{spill}} := \bar{\bar{y}}_{\text{is}} - \bar{\bar{y}}_{\text{cc}}$ measures a "seller"-spillover effect. $\vec{\boldsymbol{\beta}}_{\text{direct}} := [1, -1, -1, 1]^\top$, which induces the effect τ_{direct}, is a measure of something closer to the direct effect of the treatment, removing the spillover effects.

To elaborate and be more precise about the value of these estimands for decision making, note that within the class of SMRD's indexed by the probabilities p^{B} and p^{S}, the population averages $\bar{\bar{y}}_\gamma$ depend on the values of these probabilities, other than $\bar{\bar{y}}_{\text{cc}}$. A natural object of interest for a decision maker is the average effect of switching from no exposure to all buyer/seller pairs exposed. This can be written as

$$\bar{\bar{y}}_{\text{tr}}(p^{\text{B}} = 1, p^{\text{S}} = 1) - \bar{\bar{y}}_{\text{cc}}(p^{\text{B}} = 0, p^{\text{S}} = 0).$$

This cannot be estimated directly from an SMRD experiment with a single pair of values $(p^{\text{B}}, p^{\text{S}})$, as it requires extrapolation to $p^{\text{S}} = 1$ and $p^{\text{B}} = 1$. Either doing an experiment with p^{B} and p^{S} close enough to one or carrying out a more complex experiment with variation in p^{S} and p^{B} would facilitate this. A second goal for the decision maker may be to assess the magnitude of the spillovers relative to direct effects. SMRD experimentation lowers precision relative to completely randomized experiments, and if one finds the the spillovers are modest, one may not need to be concerned about the spillovers in future experimentation.

Note that our analysis is richer than that presented in Johari et al. [2022], where the focus is only on the estimand defined as the average outcome for the treated, $\bar{\bar{y}}_{\text{tr}}$, and the average outcome for all pairs exposed to the control group, not adjusting for any spillovers.

4.2 Unbiased Estimators for the Causal Effects

In what follows, we use capital letters to denote stochastic counterparts of the corresponding population quantities. In particular, we use Γ_{ij} to denote the random "type" assigned to pair (i, j) in the context of an SMRD. Define the realized counterpart of the population average of the buyer-seller pairs by type introduced in eq. (8):

$$\widehat{\overline{\overline{Y}}}_\gamma := \frac{1}{I_\gamma J_\gamma} \sum_{i=1}^{I} \sum_{j=1}^{J} y_{ij}(\gamma) \mathbf{1}(\Gamma_{ij} = \gamma), \tag{10}$$

lemma 4.1 shows that in an SMRD under assumption 2.4, eq. (10) provides an unbiased estimator of the corresponding population average $\overline{\overline{y}}_\gamma$ defined in eq. (8).

Lemma 4.1. *Consider an SMRD in which local interference (assumption 2.4) holds. The plug-in estimators in eq. (10) satisfy*

$$\mathbb{E}\left[\widehat{\overline{\overline{Y}}}_\gamma\right] = \overline{\overline{y}}_\gamma, \ \forall\, \gamma \in \{\text{cc}, \text{ib}, \text{is}, \text{tr}\}.$$

In light of lemma 4.1, a direct application of the linearity of the expectation implies that simple plug-in estimators of causal effects $\tau(\vec{\boldsymbol{\beta}})$ of the form of eq. (9) are unbiased.

Theorem 4.2. *Consider an SMRD where assumption 2.4 holds. The plug-in estimators* $\hat{\tau}(\vec{\boldsymbol{\beta}}) = \beta_{\text{cc}} \widehat{\overline{\overline{Y}}}_{\text{cc}} + \beta_{\text{ib}} \widehat{\overline{\overline{Y}}}_{\text{ib}} + \beta_{\text{is}} \widehat{\overline{\overline{Y}}}_{\text{is}} + \beta_{\text{tr}} \widehat{\overline{\overline{Y}}}_{\text{tr}}$ *for* $\tau(\vec{\boldsymbol{\beta}})$ *defined in eq. (9) satisfy*

$$\mathbb{E}\left[\hat{\tau}(\vec{\boldsymbol{\beta}})\right] = \tau(\vec{\boldsymbol{\beta}}). \tag{11}$$

4.3 Variances of Linear Estimators

We now characterize the variances of linear estimators $\hat{\tau}(\vec{\boldsymbol{\beta}})$ (theorem 4.3) and provide conservative estimates for their variances (theorem 4.5). Our results generalize classic results for SRDs, but their derivation is more complex because of the double summation over buyers and sellers, and requires additional notation. Define the (population) average outcome for each buyer and each seller, for a given type γ:

$$\overline{y}_i^{\text{B}}(\gamma) := \frac{1}{J} \sum_{j=1}^{J} y_{ij}(\gamma), \qquad \text{and} \qquad \overline{y}_j^{\text{S}}(\gamma) := \frac{1}{I} \sum_{i=1}^{I} y_{ij}(\gamma). \tag{12}$$

Define the deviations from population averages for buyer i, seller j, and interactions (i, j):

$$\delta_i^{\text{B}}(\gamma) := \overline{y}_i^{\text{B}}(\gamma) - \overline{\overline{y}}_\gamma, \qquad \delta_j^{\text{S}}(\gamma) := \overline{y}_j^{\text{S}}(\gamma) - \overline{\overline{y}}_\gamma,$$

and

$$\delta_{ij}^{\text{BS}}(\gamma) := y_{ij}(\gamma) - \overline{y}_i^{\text{B}}(\gamma) - \overline{y}_j^{\text{S}}(\gamma) + \overline{\overline{y}}_\gamma.$$

Next define the population variances for each type at the buyer, seller, and interaction level:

$$\sigma_\gamma^{\mathrm{B}} := \frac{\sum_{i=1}^{I}\left[\delta_i^{\mathrm{B}}(\gamma)\right]^2}{I},\ \sigma_\gamma^{\mathrm{S}} := \frac{\sum_{j=1}^{J}\left[\delta_j^{\mathrm{S}}(\gamma)\right]^2}{J},\ \sigma_\gamma^{\mathrm{BS}} := \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\left[\delta_{ij}^{\mathrm{BS}}(\gamma)\right]^2}{IJ}.$$

We additionally define for all $\gamma, \gamma' \in \{\mathrm{cc}, \mathrm{ib}, \mathrm{is}, \mathrm{tr}\}$ the following quantities, which can be interpreted as the average square deviation from the mean at the buyer, seller, and interaction level:

$$\begin{aligned} \xi_{\gamma,\gamma'}^{\mathrm{B}} &:= \sum_{i=1}^{I} \frac{\left[\delta_i^{\mathrm{B}}(\gamma) - \delta_i^{\mathrm{B}}(\gamma')\right]^2}{I}, \qquad \xi_{\gamma,\gamma'}^{\mathrm{S}} := \sum_{j=1}^{J} \frac{\left[\delta_j^{\mathrm{S}}(\gamma) - \delta_j^{\mathrm{S}}(\gamma')\right]^2}{J}, \\ \xi_{\gamma,\gamma'}^{\mathrm{BS}} &:= \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}\left[\delta_{ij}^{\mathrm{BS}}(\gamma) - \delta_{ij}^{\mathrm{BS}}(\gamma')\right]^2. \end{aligned} \tag{13}$$

Last, define for $\gamma \in \{\mathrm{cc}, \mathrm{ib}, \mathrm{is}, \mathrm{tr}\}$ the weights

$$\alpha_\gamma^{\mathrm{B}} := \frac{1}{I-1}\frac{I-I_\gamma}{I_\gamma} \qquad \text{and} \qquad \alpha_\gamma^{\mathrm{S}} := \frac{1}{J-1}\frac{J-J_\gamma}{J_\gamma}. \tag{14}$$

Let

$$\nu_{\gamma,\gamma'}^{\mathrm{B}} := \begin{cases} \alpha_\gamma^{\mathrm{B}}/2 \text{ if } \gamma = \gamma',\ \text{or } (\gamma,\gamma') \in \{(\mathrm{cc},\mathrm{is}), (\mathrm{is},\mathrm{cc}), (\mathrm{ib},\mathrm{tr}), (\mathrm{tr},\mathrm{ib})\} \\ -1/(2(I-1)) \text{ otherwise,} \end{cases}$$

and

$$\nu_{\gamma,\gamma'}^{\mathrm{S}} := \begin{cases} \alpha_\gamma^{\mathrm{S}}/2 \text{ if } \gamma = \gamma' \text{ or } (\gamma,\gamma') \in \{(\mathrm{cc},\mathrm{ib}), (\mathrm{ib},\mathrm{cc}), (\mathrm{is},\mathrm{tr}), (\mathrm{tr},\mathrm{is})\} \\ -1/(2(J-1)) \text{ otherwise.} \end{cases}$$

We now characterize variances and covariances of all the estimators of the sample average defined in eq. (10).

Theorem 4.3. *For an SMRD where assumption 2.4 holds, and for all* γ, γ'*,*

$$\mathrm{Cov}\left[\hat{\overline{\overline{Y}}}_\gamma, \hat{\overline{\overline{Y}}}_{\gamma'}\right] = \nu_{\gamma,\gamma'}^{\mathrm{B}}\zeta_{\gamma,\gamma'}^{\mathrm{B}} + \nu_{\gamma,\gamma'}^{\mathrm{S}}\zeta_{\gamma,\gamma'}^{\mathrm{S}} + \nu_{\gamma,\gamma'}^{\mathrm{B}}\nu_{\gamma,\gamma'}^{\mathrm{S}}\zeta_{\gamma,\gamma'}^{\mathrm{BS}},$$

where for $x \in \{\mathrm{B}, \mathrm{S}, \mathrm{BS}\}$*,* $\zeta_{\gamma,\gamma'}^{x} := \sigma_\gamma^x + \sigma_{\gamma'}^x - \xi_{\gamma,\gamma'}^x$*.*

Variances for the type estimator $\hat{\overline{\overline{Y}}}_\gamma$ are obtained using the formula above whenever $\gamma' = \gamma$. Exact variances of estimators $\hat{\tau}(\vec{\boldsymbol{\beta}})$ can be directly obtained by noting that $\hat{\tau}(\vec{\boldsymbol{\beta}})$ is a linear estimator, for which the following decomposition holds:

$$\mathrm{Var}(aX + bY) = a^2\,\mathrm{Cov}(X,X) + b^2\,\mathrm{Cov}(Y,Y) + 2ab\,\mathrm{Cov}(X,Y).$$

4.4 Variance Estimation

We now present unbiased estimators for the variance of the sample average of potential outcomes defined in eq. (10); these are given in theorem 4.4. We then give lower and upper bounds on the variance of the linear estimators $\hat{\tau}(\vec{\boldsymbol{\beta}})$ in theorem 4.5.

Towards this goal, we proceed to define the sample counterparts of the population quantities introduced in section 4.3. Given a randomly drawn SMRD assignment matrix $\mathbf{W} \in \mathbb{W}$, inducing corresponding types $\mathbf{\Gamma}$, let $\mathcal{I}_\gamma := \{i \in [I] \text{ s.t. } \Gamma_{ij} = \gamma \text{ for some } j\}$ with size $|\mathcal{I}_\gamma| = I_\gamma$ and $\mathcal{J}_\gamma := \{j \in [J] \text{ s.t. } \Gamma_{ij} = \gamma \text{ for some } i\}$ with size $|\mathcal{J}_\gamma| = J_\gamma$. From eq. (7), each $i \in [I]$ belongs to exactly two sets $\mathcal{I}_\gamma$: if $W_i^{\mathrm{B}} = 0$, $i \in \mathcal{I}_{\mathrm{cc}}$ and $i \in \mathcal{I}_{\mathrm{is}}$. If $W_i^{\mathrm{B}} = 1$, $i \in \mathcal{I}_{\mathrm{ib}}$ and $i \in \mathcal{I}_{\mathrm{tr}}$. Symmetrically, each $j \in [J]$ belongs in exactly two sets $\mathcal{J}_\gamma$: if $W_j^{\mathrm{S}} = 0$, $j \in \mathcal{J}_{\mathrm{cc}}$ and $j \in \mathcal{J}_{\mathrm{ib}}$, and if $W_j^{\mathrm{S}} = 1$, $j \in \mathcal{J}_{\mathrm{is}}$ and $i \in \mathcal{J}_{\mathrm{tr}}$. For $i \in \mathcal{I}_\gamma, j \in \mathcal{J}_\gamma$ the sample counterparts $\widehat{\overline{Y}}_i^{\mathrm{B}}(\gamma)$ of $\overline{y}_i^{\mathrm{B}}(\gamma)$ and $\widehat{\overline{Y}}_j^{\mathrm{S}}(\gamma)$ of $\overline{y}_j^{\mathrm{S}}(\gamma)$ are:

$$\widehat{\overline{Y}}_i^{\mathrm{B}}(\gamma) := \frac{1}{J_\gamma} \sum_{j \in \mathcal{J}_\gamma} y_{ij}(\gamma), \qquad \widehat{\overline{Y}}_j^{\mathrm{S}}(\gamma) := \frac{1}{I_\gamma} \sum_{i \in \mathcal{I}_\gamma} y_{ij}(\gamma).$$

We define estimator counterparts $\widehat{\Sigma}_\gamma^{\mathrm{B}}$ for $\sigma_\gamma^{\mathrm{B}}$ (buyers) and $\widehat{\Sigma}_\gamma^{\mathrm{S}}$ for $\sigma_\gamma^{\mathrm{S}}$ (sellers):

$$\widehat{\Sigma}_\gamma^{\mathrm{B}} := \frac{1}{I_\gamma} \sum_{i \in \mathcal{I}_\gamma} \left[\widehat{\overline{Y}}_i^{\mathrm{B}}(\gamma) - \widehat{\overline{\overline{Y}}}_\gamma \right]^2, \quad \widehat{\Sigma}_\gamma^{\mathrm{S}} := \sum_{j \in \mathcal{J}_\gamma} \frac{1}{J_\gamma} \left[\widehat{\overline{Y}}_j^{\mathrm{S}}(\gamma) - \widehat{\overline{\overline{Y}}}_\gamma \right]^2.$$

For the interactions, we define the estimator counterpart $\widehat{\Sigma}_\gamma^{\mathrm{BS}}$ for $\sigma_\gamma^{\mathrm{BS}}$:

$$\widehat{\Sigma}_\gamma^{\mathrm{BS}} := \sum_{i \in \mathcal{I}_\gamma, j \in \mathcal{J}_\gamma} \frac{\left(y_{i,j}(\gamma) - \widehat{\overline{Y}}_i^{\mathrm{B}}(\gamma) - \widehat{\overline{Y}}_j^{\mathrm{S}}(\gamma) + \widehat{\overline{\overline{Y}}}_\gamma \right)^2}{I_\gamma J_\gamma}.$$

Theorem 4.4. *For an SMRD where assumption 2.4 holds, for all* $\gamma \in \{\mathrm{cc}, \mathrm{ib}, \mathrm{is}, \mathrm{tr}\}$,

$$\mathbb{E}\left[\widehat{\Sigma}(\gamma)\right] = \mathrm{Var}\left(\widehat{\overline{\overline{Y}}}_\gamma\right), \quad \textit{where}$$

$$\begin{aligned} \hat{\Sigma}(\gamma) := {} & \frac{\alpha_\gamma^{\mathrm{B}} \widehat{\Sigma}_\gamma^{\mathrm{B}} + \alpha_\gamma^{\mathrm{S}} \widehat{\Sigma}_\gamma^{\mathrm{S}} + \alpha_\gamma^{\mathrm{B}} \alpha_\gamma^{\mathrm{S}} \widehat{\Sigma}_\gamma^{\mathrm{BS}}}{1 - \alpha_\gamma^{\mathrm{B}} - \alpha_\gamma^{\mathrm{S}} + \alpha_\gamma^{\mathrm{B}} \alpha_\gamma^{\mathrm{S}}} \\ & - \frac{\frac{\alpha_\gamma^{\mathrm{B}}}{(1-\alpha_\gamma^{\mathrm{B}})}}{I_\gamma J_\gamma} \sum_{i \in \mathcal{I}_\gamma, j \in \mathcal{J}_\gamma} \frac{\left(y_{i,j}(\gamma) - \widehat{\overline{Y}}_i^{\mathrm{B}}(\gamma) \right)^2}{\frac{(J-1)(J_\gamma - 1)}{(J - J_\gamma)}} - \frac{\frac{\alpha_\gamma^{\mathrm{S}}}{(1-\alpha_\gamma^{\mathrm{S}})}}{I_\gamma J_\gamma} \sum_{i \in \mathcal{I}_\gamma, j \in \mathcal{J}_\gamma} \frac{\left(y_{i,j}(\gamma) - \widehat{\overline{Y}}_j^{\mathrm{S}}(\gamma) \right)^2}{\frac{(I-1)(I_\gamma - 1)}{(I - I_\gamma)}}. \end{aligned}$$

Young's inequality yields a conservative estimator for the variance of $\hat{\tau}(\vec{\boldsymbol{\beta}})$:

$$\widehat{\mathrm{Var}}\left(\hat{\tau}^{\mathrm{hi}}(\vec{\boldsymbol{\beta}})\right) = \sum_{\gamma \in \{\mathrm{cc}, \mathrm{ib}, \mathrm{is}, \mathrm{tr}\}} \beta_\gamma^2 \hat{\Sigma}_\gamma + \sum_{\gamma \neq \gamma'} \beta_\gamma \beta_{\gamma'} \left(\hat{\Sigma}_\gamma + \hat{\Sigma}_{\gamma'} \right) \tag{15}$$

This result mirrors the case of SRDs [Neyman, 1923/1990]. We provide the result for $\widehat{\tau}_{\mathrm{spill}}^{\mathrm{B}}$ in theorem 4.5. See lemma A.18 in the Appendix for the case of a generic $\hat{\tau}(\vec{\boldsymbol{\beta}})$.

Theorem 4.5. *Under the assumptions of theorem 4.4 a conservative estimator of* $\mathrm{Var}(\widehat{\tau}_{\mathrm{spill}}^{\mathrm{B}})$

is:

$$\widehat{\mathrm{Var}}^{\mathrm{hi}}(\hat{\tau}^{\mathrm{B}}_{\mathrm{spill}}) := 2\left(\widehat{\Sigma}(\mathrm{ib}) + \widehat{\Sigma}(\mathrm{cc})\right).$$

$\widehat{\mathrm{Var}}^{\mathrm{hi}}(\hat{\tau}^{\mathrm{B}}_{\mathrm{spill}})$ *is conservative in the usual sense that* $\mathbb{E}\left[\widehat{\mathrm{Var}}^{\mathrm{hi}}(\hat{\tau}^{\mathrm{B}}_{\mathrm{spill}})\right] \geq \mathrm{Var}(\hat{\tau}^{\mathrm{B}}_{\mathrm{spill}})$.

We emphasize that, while it is possible to provide an unbiased estimator for the variance of $\hat{\overline{\overline{Y}}}_\gamma$ (theorem 4.4), one *cannot* provide an unbiased estimator for the covariance of $\hat{\overline{\overline{Y}}}_\gamma$ and $\hat{\overline{\overline{Y}}}_{\gamma'}$ for $\gamma \neq \gamma'$ without stronger assumptions on the potential outcomes. The same phenomenon occurs for conventional randomized experiments. This is because the terms $\xi^x_{\gamma,\gamma'}$ introduced in eq. (13) depend on covariances of potential outcomes for the same buyer-seller pair, which cannot be identified from the observed data.

It is however possible to show that the variance estimator $\hat{\Sigma}(\gamma)$ converges to the true underlying variance Σ_γ, i.e. $\mathrm{Var}\left(\hat{\overline{\overline{Y}}}_\gamma\right)$, under relatively weak assumptions. By the continuous mapping theorem, this implies convergence of the general estimator $\widehat{\mathrm{Var}}^{\mathrm{hi}}[\hat{\tau}(\vec{\boldsymbol{\beta}})]$ to its (conservative) limit. A stronger version of this result was communicated to us by Sudijono et al. [2025]; a proof is given in appendix A.4 for completeness.

We now state the result. To do so, we introduce two additional assumptions which will also be used in section 4.5 to derive a central limit theorem.

Assumption 4.6. *Consider an SMRD with I buyers and J sellers, and assume that the local interference assumption assumption 2.4 holds. We impose the following regularity conditions.*

(a) Balance: for all $\gamma \in \{\mathrm{cc}, \mathrm{ib}, \mathrm{is}, \mathrm{tr}\}$, a valid assignment is characterized by fixed I_γ and J_γ, with $I/I_\gamma, J/J_\gamma \leq C_1$.

(b) Boundedness: for all buyers and seller interactions (i, j) and all types γ, $|y_{ij}(\gamma)| \leq C_2$.

Theorem 4.7. *Let $\hat{\tau}(\vec{\boldsymbol{\beta}})$ be the linear estimator given in Equation (11), and let $\widehat{\mathrm{Var}}^{\mathrm{hi}}[\hat{\tau}(\vec{\boldsymbol{\beta}})]$ be its conservative variance estimator given in theorem 4.4. Then, in any sequence of SMRDs satisfying assumption 4.6 and in which $(I^{-2} + J^{-2})/\,\mathbb{E}\{\widehat{\mathrm{Var}}^{\mathrm{hi}}[\hat{\tau}(\vec{\boldsymbol{\beta}})]\} \to 0$, we have*

$$\frac{\widehat{\mathrm{Var}}^{\mathrm{hi}}[\hat{\tau}(\vec{\boldsymbol{\beta}})]}{\mathbb{E}\left\{\widehat{\mathrm{Var}}^{\mathrm{hi}}[\hat{\tau}(\vec{\boldsymbol{\beta}})]\right\}} = 1 + o_p(1).$$

4.5 Finite Population Central Limit Theorem

We conclude this section by providing a quantitative central limit theorem for the estimators introduced in section 4. Notably, we do not assume that the observed units are drawn from an underlying "super-population," nor do we consider a sequence of experiments. Instead, our approach quantifies the distribution of our estimates using only the randomness of the design, in terms of well-defined properties of the finite population. This approach allows us to limit assumptions imposed on the potential outcomes. Our contribution can be seen as an extension to the multi-population setting of recent advances in the causal inference literature, and in particular of the works of Li and Ding [2017a] and Shi and Ding [2022a] for single-sided experiments. Our setting presents additional technical challenges, as the outcomes exhibit a complex dependence structure. Theorem 4.8 serves as the basis for statistical inference in the context of multiple randomization designs.

Theorem 4.8. *Consider an SMRD where assumption 2.4 and assumption 4.6 hold. Then we have*

$$\sup_{t\in\mathbb{R}}\left|\mathbb{P}\left\{\frac{\hat{\tau}(\vec{\boldsymbol{\beta}})-\tau(\vec{\boldsymbol{\beta}})}{\sqrt{\operatorname{Var}\left[\hat{\tau}(\vec{\boldsymbol{\beta}})\right]}}\le t\right\}-\Phi(t)\right|\le C\Delta^{\frac{1}{3}}\log\left(\frac{C}{\Delta}\right), \tag{16}$$

with $\Delta := \frac{C_1^2C_2(I^{-1}+J^{-1})}{\operatorname{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})\}^{\frac{1}{2}}/\|\vec{\boldsymbol{\beta}}\|}$ *and where* Φ *denotes the standard normal cumulative density function (CDF), and* $C>0$ *is a universal constant.*[1]

Remark 4.9 (Boundedness and sparsity)**.** In addition to ruling out heavy-tailed potential outcome distributions, an important limitation of theorem 4.8 is *sparsity*, when a large fraction of unit potential outcomes $y_{ij}(\gamma)$, or their differences $y_{ij}(\gamma)-y_{ij}(\gamma')$, are zero. Sparsity can also cause problems in CLTs for conventional randomized experiments such as the ones cited above. It may be especially relevant in our setting, however, where units correspond to pairwise interactions between large populations.

Since $\hat{\tau}(\vec{\boldsymbol{\beta}})$ is linear in observed outcomes Y_{ij}, the quantity $C_{\vec{\boldsymbol{\beta}}} = C_2/\{\operatorname{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})\}^{\frac{1}{2}}/\|\vec{\boldsymbol{\beta}}\|\}$ appearing in our bound (16) is invariant to re-scaling observations (*e.g.*, to ensure non-degeneracy of $\tau(\vec{\boldsymbol{\beta}})$). Theorem 4.8 requires that $C_{\vec{\boldsymbol{\beta}}}$ be small in comparison to $\{I^{-1}+J^{-1}\}^{-1}$ for I and J large, allowing a limited degree of sparsity. For example, if potential outcomes are binary, if I and J are of the same order, and if half of the rows and columns have a fraction 2μ of non-zero entries (while the rest are all zero), then our result requires μ to be much larger than $I^{-1/2}$. Generalizing theorem 4.8 to better accommodate heavy-tailed and sparse potential outcomes is an important direction for future work.

We articulate our proof in three main steps, described in detail in appendix B. First, we prove that if we fix the assignment of one of the two populations (*e.g.*, sellers), an analogous version of the results proved by Li and Ding [2017a] and Shi and Ding [2022a] holds for the multi-population setting, where the parameters of the CLT are indexed by the seller assignment (appendix B.1). Second, we show that with high probability, these fixed parameters are either themselves normally distributed, or else are close to their expected value (appendix B.2). Last, we combine these results to prove a CLT for simple double randomized experiments (appendix B.3).

The main challenge in proving our result is that separate randomization of the two populations creates two-way dependence in the realized outcomes, complicating the application of standard techniques. Similar settings have been studied using Stein's method of exchangeable pairs, although the proofs are quite complex [Zhao et al., 1997]. Interestingly, the proof of theorem 4.8 treats the two populations asymmetrically, although the final bound is symmetric in I and J.

[1] We are very grateful to Sudijono et al. [2025], who communicated an important idea that led to the correction of an error in the proof of Theorem 4.8.

Finally, we comment on the application of theorem 4.8 in practice. It is natural to replace the variance $\mathrm{Var}[\hat{\tau}(\vec{\boldsymbol{\beta}})]$ by its estimated upper bound $\widehat{\mathrm{Var}}^{\mathrm{hi}}[\hat{\tau}(\vec{\boldsymbol{\beta}})]$. Roughly speaking, the Studentized statistic $\hat{z}_\tau = \{\hat{\tau}(\vec{\boldsymbol{\beta}}) - \tau\}/\widehat{\mathrm{Var}}^{\mathrm{hi}}[\hat{\tau}(\vec{\boldsymbol{\beta}})]^{1/2}$ will be approximately normally distributed with variance at most 1 provided the denominator converges, which follows by theorem 4.7. One can then test one- and two-sided hypotheses on $\tau(\vec{\boldsymbol{\beta}})$ by comparing $\hat{z}_\tau$ to standard normal critical values. We empirically verify normality of the Studentized statistic and illustrate the resulting tests with synthetic data in section 5.

5 Simulations

We now verify the results of section 4 for SMRDs under local interference. Our simulations follow the model of strategic agents in a two-sided marketplace introduced in example 2.5, which naturally produces local interference. Additional experiments from a simple additive Gaussian model satisfying local interference are provided in appendix C. `Python` code to replicate all our simulations is available at https://github.com/lorenzomasoero/MultipleRandomizationDesigns.

The simulations following example 2.5 also illustrate the practical value of SMRDs. In the underlying model, higher quality ads increase the incentive to produce high quality content, and vice-versa. This is an example of *strategic complementarity*, a prominent and well-studied feature of many real-world marketplaces [Milgrom and Roberts, 1990]. In this model, it leads to significant positive spillovers for both advertisers and creators. These spillovers are neatly captured by the MRD, but cause conventional, single randomized experiments to underestimate the treatment effect, possibly leading to sub-optimal policies.

To empirically validate the results presented in section 4, we first instantiate the model from example 2.5 by fixing the incentive level $\eta = 5\%$ and drawing independent and identically distributed parameters $m_{ij} \sim \mathrm{Exp}(1)$, $r_i^c, r_j^a \sim \mathrm{Unif}([0, 1/5])$ across advertisers $i \in [I]$ and creators $j \in [J]$ (notice: here creators and advertisers have roles analogous to that of buyers and sellers in the discussion of sections 2 to 4). In our simulation, we let $I = 200$ and $J = 150$. Taking these parameters—which determine the fixed population—as given, we fix the treatment group size $I_T = 100$ and $J_T = 80$. We then sample treatment assignment matrices $\mathbf{W}$ at random from the SMRD $\mathbb{W}$, which determine realized equilibrium outcomes $Y_{ij}(\mathbf{W})$ to be the platform's profit following eq. (2).

Since local interference (assumption 2.4) is satisfied, $Y_{ij}(\mathbf{W})$ depends only upon the type $\Gamma_{ij} \in \{\mathrm{cc}, \mathrm{ib}, \mathrm{is}, \mathrm{tr}\}$ of unit (i, j), conditional upon the parameters I_T, J_T, and the fixed population. Each assignment $\mathbf{W}$ then corresponds to an observed matrix of $I \times J$ realized potential outcomes. We use the collection of outcomes from 10,000 independent re-randomizations to empirically verify the properties of the proposed estimators. Figure 1 reports the histogram of the values attained by $\widehat{\overline{\overline{Y}}}_{\mathrm{cc}}$ (left) and $\widehat{\Sigma}_{\mathrm{cc}}$ (right) across the 10,000 Monte Carlo replicates. As follows from lemma 4.1, $\widehat{\overline{\overline{Y}}}_{\mathrm{cc}}$ is centered at the true population value $\overline{\overline{y}}_\gamma$, and, using theorem 4.8, under mild conditions $\widehat{\overline{\overline{Y}}}_\gamma$ is approximately normally distributed. Moreover, the distance between the 2.5% and 97.5% quantiles of the distribution of the type estimator (red vertical lines) is close to the length of the 95% confidence interval around the population value $\overline{\overline{y}}_\gamma$, formed by using the true variance of $\widehat{\overline{\overline{Y}}}_\gamma$. In the right panel

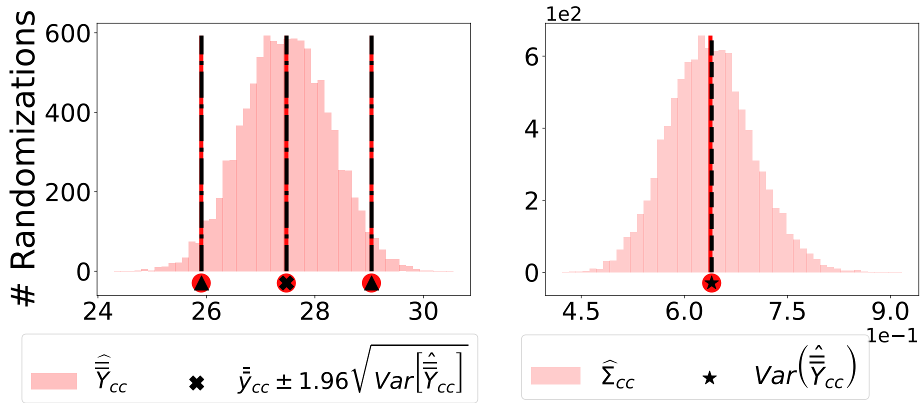

Figure 1: Distribution of $\widehat{\overline{\overline{Y}}}_{\text{cc}}$ (left) and of the variance estimator $\widehat{\Sigma}_{\text{cc}}$ (right). Black lines correspond to the population quantities $\overline{\overline{y}}_{\text{cc}}$, $\text{Var}\left(\widehat{\overline{\overline{Y}}}_{\text{cc}}\right)$.

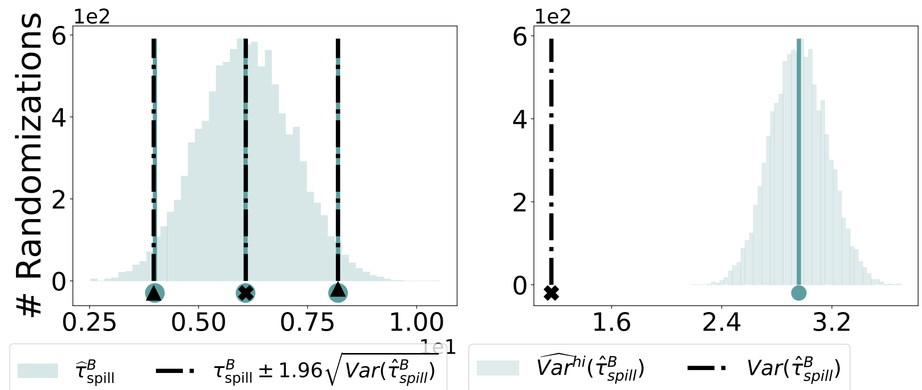

Figure 2: Distribution of the estimator for the spillover effect $\hat{\tau}^{B}_{\text{spill}}$ (left) and corresponding variance estimator $\widehat{\text{Var}}^{\text{hi}}(\hat{\tau}^{B}_{\text{spill}})$ (right). Black lines correspond to the population quantities.

of fig. 1, we show that $\widehat{\Sigma}_{\text{cc}}$ is an unbiased estimator for the variance of the type estimator, as proved in theorem 4.4. Analogous results hold for ib, is, tr.

We focus on the spillover effect $\tau^{\text{B}}_{\text{spill}}$ in fig. 2: the left panel shows the distribution of the unbiased estimator $\widehat{\tau}^{\text{B}}_{\text{spill}}$ (theorem 4.2). $\widehat{\tau}^{\text{B}}_{\text{spill}}$ is Gaussian (as shown in theorem 4.8), and conservative confidence intervals can be derived. The right panel contains the distribution of the upper bound $\widehat{\text{Var}}^{\text{hi}}(\widehat{\tau}^{\text{B}}_{\text{spill}})$ for the variance $\text{Var}(\widehat{\tau}^{\text{B}}_{\text{spill}})$ (theorem 4.5). Additional plots and implementation details are provided in appendix C.

Under mild conditions laid out in theorem 4.8 and the following discussion, one can practically test for the presence of positive spillover effects by constructing the Studentized statistic $\hat{z}_0 := \widehat{\tau}^{\text{B}}_{\text{spill}} / \{\widehat{\text{Var}}^{\text{hi}}(\widehat{\tau}^{\text{B}}_{\text{spill}})\}^{1/2}$ and comparing it to standard normal critical values. For our model, the conservative test $\hat{z}_0$ rejects the null hypothesis of no effect 99.5% of the time (Type-II error is 0.5%), showing substantial power to detect positive spillovers.

Finally, we compare MRDs to the standard practice of single randomization—randomizing exactly 50% of creators $i \in [I]$ into treatment, and treating all of their interactions as in eq. (3), and then using the standard difference-in-means estimator $\hat{\tau}_{\text{SRD}}$. In our model, such

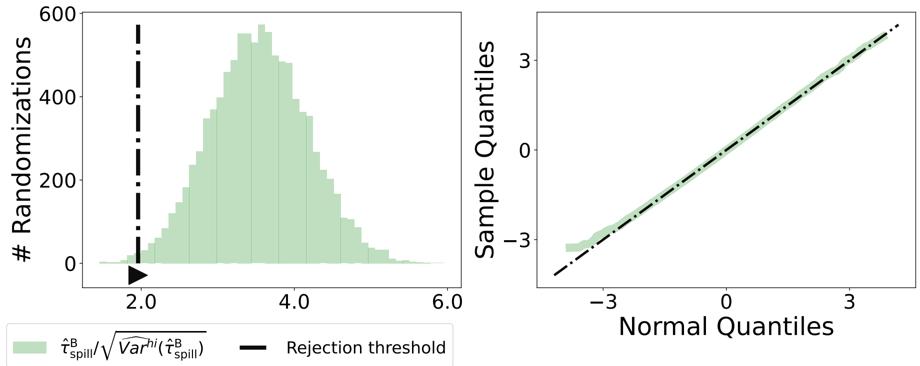

Figure 3: Left: distribution of the Studentized statistic $\widehat{\tau}^{\mathrm{B}}_{\mathrm{spill}}/\{\widehat{\mathrm{Var}}^{\mathrm{hi}}(\widehat{\tau}^{\mathrm{B}}_{\mathrm{spill}})\}^{1/2}$ and resulting conservative two-sided tests of the null hypothesis of no effect $H_0 = \{\tau^{\mathrm{B}}_{\mathrm{spill}} = 0\}$ (statistics to the right of the black line reject H_0). Right: QQ plot comparing the Studentized statistic to a Gaussian law with the same mean and variance.

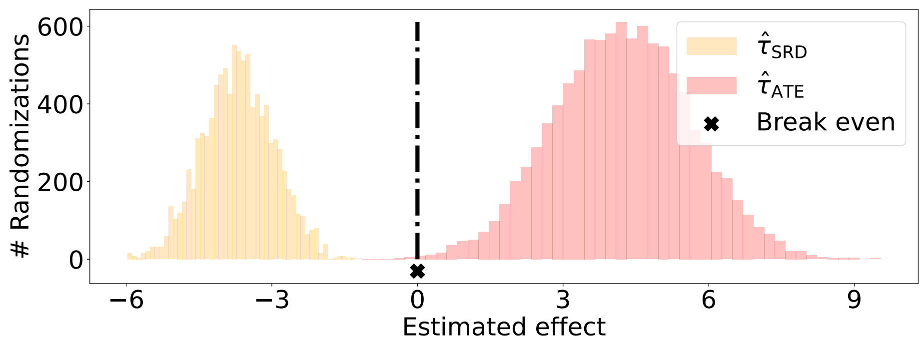

Figure 4: Distribution of the standard difference-in-means estimator $\hat{\tau}_{\mathrm{SRD}}$ across 10,000 single randomized experiments (yellow), compared to the distribution of $\hat{\tau}_{\mathrm{ATE}}$ in as many SDRDs (red). The estimators are produced by re-drawing different randomization designs for the same underlying finite population, with potential outcomes given by Example 2.5.

an estimator neglects positive spillovers mediated by advertisers' strategic responses. By comparing the distribution of the difference-in-means estimator $\hat{\tau}_{\mathrm{SRD}}$ under the standard creator-randomized design given by eq. (3) to the distribution of $\hat{\tau}_{\mathrm{ATE}}$ under the SMRD, we illustrate in fig. 4 that in our model the standard design usually produces the incorrect sign of the platform's profit relative to that which would be obtained by treating the whole population, while the SMRD usually produces the correct sign.

6 Extensions and future work

The designs discussed in section 3 are a few of many possible designs that fit into the MRD framework. While we have focused in detail on the "Simple" MRD case, many other designs fit the MRD paradigm—including clustered experiments, experiments involving three or more populations, etc. These generalizations also include time-randomized experiments: e.g., recently Masoero et al. [2023] used the MRD framework to show that under certain assumptions on the potential outcomes, switchback designs based upon the MRD framework can lead to more efficient estimates of causal effects. MRDs have also been used in practice in the context of online marketplaces, to quantify the direct and indirect effects of certain interventions; see, e.g., Masoero et al. [2024], Zhu et al. [2024], Bright et al. [2024].

Additionally, as highlighted in the discussion following assumption 2.4, we emphasize that the local interference assumption is only a starting point from which to rigorously study causal inference with MRDs. We envision that future work will study how MRDs can be used in conjunction with more complicated interference structures. Characterizing minimal restrictions on interference under which similar, design-based inference results can be derived is an open question beyond the scope of this paper.

To illustrate the richness of our framework, we conclude by describing four additional designs which fit within the MRD setting. First, instead of partitioning buyers and sellers into two groups each, we can assign them to a finite number of groups, with the assignment a function of this finer partition. This allows to generate more variation in $\overline{w}_i^{\mathrm{B}}$ and $\overline{w}_j^{\mathrm{S}}$ and in turn to build models for the dependence of the potential outcomes on the share of treated buyers and sellers that will permit more credible extrapolation to full exposure to treatment or control. As a simple example, we could endow each buyer i and seller j with scalar scores w_i^{B} and w_j^{S} (as opposed to binary values), and let the treatment assignment be defined by a modified version of eq. (5), *e.g.*, $f(w_i^{\mathrm{B}}, w_j^{\mathrm{S}}) = \mathbf{1}(w_i^{\mathrm{B}} + w_j^{\mathrm{S}}) > \kappa$ for a given threshold κ (*e.g.*, $\kappa = 0.5$ in 17).

$$
\begin{array}{l}
\text{Seller} \rightarrow \\
\text{Score}
\end{array}
\begin{array}{cccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
0 & 0 & 0.2 & 0.2 & 0.4 & 0.4 & 0.6 & 0.6
\end{array}
\quad \text{Buyer Score}
$$

$$
\mathbf{w} = \begin{pmatrix}
C & C & C & C & C & C & T & T \\
C & C & C & C & C & C & T & T \\
C & C & C & C & T & T & T & T \\
C & C & C & C & T & T & T & T \\
C & C & T & T & T & T & T & T \\
C & C & T & T & T & T & T & T
\end{pmatrix}
\begin{array}{cc}
1 & 0 \\
2 & 0 \\
3 & 0.2 \\
4 & 0.2 \\
5 & 0.4 \\
6 & 0.4
\end{array}
\tag{17}
$$

Second, one can first partition one of the groups (*e.g.*, sellers) into two random groups (A, B), and run a buyer experiment for one group and a seller experiment for the other.

$$
\begin{array}{l}
\text{Seller} \rightarrow \\
\text{Exp} \rightarrow
\end{array}
\begin{array}{cccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
A & A & A & A & A & B & B & B
\end{array}
\quad \text{Buyer}
$$

$$
\mathbf{w} = \begin{pmatrix}
C & C & C & C & C & T & C & T \\
C & C & C & C & C & T & C & T \\
T & T & T & T & T & T & C & T \\
C & C & C & C & C & T & C & T \\
T & T & T & T & T & T & C & T
\end{pmatrix}
\begin{array}{c}
1 \\ 2 \\ 3 \\ 4 \\ 5
\end{array}
$$

Third, when one wants to do a seller-clustered experiment, one may partition the buyer population into two groups, A and B, and then run a seller clustered experiment in one group and a regular seller experiment in the second group. This would allow the researchers to infer within the context of a single experiment the within-cluster spillovers, as well as get estimates of the overall average effect.

$$
\begin{array}{rcccccccc}
\text{Seller} \rightarrow & 1 & 2 & 3 & 4 & 5 & 6 & \text{Buyer} & \text{Exp.} \\
\text{Cluster} \rightarrow & \text{I} & \text{I} & \text{II} & \text{II} & \text{III} & \text{III} & & \\
\mathbf{w} = & {\color{red}C} & {\color{blue}C} & T & T & {\color{red}C} & {\color{blue}C} & 1 & A \\
& {\color{red}C} & {\color{blue}C} & T & T & {\color{red}C} & {\color{blue}C} & 2 & A \\
& {\color{red}C} & {\color{blue}C} & T & T & {\color{red}C} & {\color{blue}C} & 3 & A \\
& {\color{red}C} & {\color{blue}C} & T & T & {\color{red}C} & {\color{blue}C} & 4 & A \\
& T & {\color{brown}C} & T & T & T & {\color{brown}C} & 5 & B \\
& T & {\color{brown}C} & T & T & T & {\color{brown}C} & 6 & B \\
& T & {\color{brown}C} & T & T & T & {\color{brown}C} & 7 & B \\
& T & {\color{brown}C} & T & T & T & {\color{brown}C} & 8 & B
\end{array}
$$

Fourth, we can consider designs where the local interference assumption is testable.

$$
\begin{array}{rccccccccc}
\text{Seller} \rightarrow & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & \text{Buyer} \\
& {\color{red}C} & {\color{red}C} & {\color{blue}C} & {\color{blue}C} & {\color{blue}C} & {\color{blue}C} & {\color{blue}C} & {\color{blue}C} & 1 \\
& {\color{red}C} & {\color{red}C} & {\color{blue}C} & {\color{blue}C} & {\color{blue}C} & {\color{blue}C} & {\color{blue}C} & {\color{blue}C} & 2 \\
& T & T & {\color{brown}C} & {\color{brown}C} & {\color{brown}C} & {\color{brown}C} & {\color{brown}C} & {\color{brown}C} & 3 \\
\mathbf{w} = & T & T & {\color{brown}C} & {\color{brown}C} & {\color{brown}C} & {\color{brown}C} & {\color{brown}C} & {\color{brown}C} & 4 \\
& {\color{green}C} & {\color{green}C} & T & T & T & T & T & T & 5 \\
& {\color{green}C} & {\color{green}C} & T & T & T & T & T & T & 6
\end{array}
$$

Consider the red C and the blue C. In both cases they correspond to buyers who are in the control group for all sellers, and in both cases they correspond to sellers who are in the treatment group for 1/3 of the buyers. However, sellers in the red C pairs are in the treatment group for buyers who are very rarely in the treatment group, whereas the sellers in the blue C pairs are in the treatment group for buyers who are often in the treatment group. When local interference holds, that should not matter, but if local interference is violated, it may matter.

A Proofs for Multiple Randomization Designs

We here prove the results presented in Section 4. We consider conjunctive SMRDs (as per Definition 3.4) where local interference holds (assumption 2.4), with a total of I buyers, J sellers, and $I \times J$ units. All buyers and sellers are endowed with random variables $W_i^{\mathrm{B}}, W_j^{\mathrm{S}} \in \{0,1\}$, so that $I > I_{\mathrm{T}} > 1$ and $J > J_{\mathrm{T}} > 1$, where $I_{\mathrm{T}} := \sum_i W_i^{\mathrm{B}}$, $J_{\mathrm{T}} := \sum_j W_j^{\mathrm{S}}$.

Lemma A.1 (Lemma 3.5)**.** *Under local interference (Assumption 2.4), potential outcomes can be written as a function of the assignment types only: for* $\mathbf{w}, \mathbf{w}' \in \{0,1\}^{I\times J}$ *it holds that*

$$\gamma_{ij}(\mathbf{w}) = \gamma_{ij}(\mathbf{w}') \Rightarrow y_{ij}(\mathbf{w}) = y_{ij}(\mathbf{w}').$$

Proof. Under Assumption 2.4, for any (i,j) and any pair of assignment matrices $\boldsymbol{w}, \boldsymbol{w}' \in \{0,1\}^{I\times J}$ $y_{ij}(\boldsymbol{w}) = y_{ij}(\boldsymbol{w}')$ whenever (a) $w_{ij} = w'_{ij}$, (b) the fraction of treated sellers for buyer i coincides in $\boldsymbol{w}, \boldsymbol{w}'$ and (c) the fraction of treated buyers for seller j coincides in $\boldsymbol{w}, \boldsymbol{w}'$. If (a), (b) and (c) hold, it must be the case that $\gamma_{ij}(\boldsymbol{w}) = \gamma_{ij}(\boldsymbol{w}')$, yielding the thesis. □

A.1 Useful definitions

Recall the definitions of the average outcomes for each buyer and each seller:

$$\overline{y}_i^{\mathrm{B}}(\gamma) := \frac{1}{J}\sum_{j=1}^{J} y_{ij}(\gamma), \quad \overline{y}_j^{\mathrm{S}}(\gamma) := \frac{1}{I}\sum_{i=1}^{I} y_{ij}(\gamma) \quad \text{and} \quad \overline{\overline{y}}_\gamma := \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J} y_{ij}(\gamma).$$

For each type $\gamma \in \{\mathrm{cc}, \mathrm{ib}, \mathrm{is}, \mathrm{tr}\}$, buyer i and seller j, define the following deviations:

$$\delta_i^{\mathrm{B}}(\gamma) := \overline{y}_i^{\mathrm{B}}(\gamma) - \overline{\overline{y}}_\gamma, \quad \delta_j^{\mathrm{S}}(\gamma) := \overline{y}_j^{\mathrm{S}}(\gamma) - \overline{\overline{y}}_\gamma, \quad \delta_{ij}^{\mathrm{BS}}(\gamma) := y_{ij}(\gamma) - \overline{y}_i^{\mathrm{B}}(\gamma) - \overline{y}_j^{\mathrm{S}}(\gamma) + \overline{\overline{y}}_\gamma.$$

By definition, the sum of these deviations is equal to zero:

$$\sum_{i=1}^{I} \delta_i^{\mathrm{B}}(\gamma) = 0, \quad \sum_{i=1}^{I} \delta_{ij}^{\mathrm{BS}}(\gamma) = 0, \quad \sum_{j=1}^{J} \delta_j^{\mathrm{S}}(\gamma) = 0, \quad \sum_{j=1}^{J} \delta_{ij}^{\mathrm{BS}}(\gamma) = 0.$$

We decompose $y_{ij}(\gamma)$ as

$$y_{ij}(\gamma) = \overline{\overline{y}}_\gamma + \delta_i^{\mathrm{B}}(\gamma) + \delta_j^{\mathrm{S}}(\gamma) + \delta_{ij}^{\mathrm{BS}}(\gamma).$$

Last, for $\gamma \in \{\mathrm{cc}, \mathrm{ib}, \mathrm{is}, \mathrm{tr}\}$ we let I_γ be the number of buyers eligible for type γ and J_γ be the number of sellers eligible for type γ. Define $I_C := I - I_{\mathrm{T}}$ and $J_C := J - J_{\mathrm{T}}$, then $I_{\mathrm{cc}} = I_{\mathrm{C}}$, $J_{\mathrm{cc}} = J_{\mathrm{C}}$, $I_{\mathrm{ib}} = I_{\mathrm{T}}$, $J_{\mathrm{ib}} = J_{\mathrm{C}}$, $I_{\mathrm{is}} = I_{\mathrm{C}}$, $J_{\mathrm{is}} = J_{\mathrm{T}}$, $I_{\mathrm{tr}} = I_{\mathrm{T}}$, $J_{\mathrm{tr}} = J_{\mathrm{T}}$.

A.2 Linear representation of the type estimators

Recall from Definition 3.4 that W_i^{B} and W_j^{S} are random variables which determine whether buyer i and seller j are eligible to be exposed to the treatment.

Lemma A.2. *The (doubly averaged) sample mean estimator* $\widehat{\overline{\overline{Y}}}_\gamma$ *can be decomposed as*

$$\begin{aligned}
\widehat{\overline{\overline{Y}}}_{\mathrm{tr}} &= \overline{\overline{y}}_{\mathrm{tr}} + \frac{1}{I_{\mathrm{T}}}\sum_{i=1}^{I} W_i^{\mathrm{B}}\delta_i^{\mathrm{B}}(\mathrm{tr}) + \frac{1}{J_{\mathrm{T}}}\sum_{j=1}^{J} W_j^{\mathrm{S}}\delta_j^{\mathrm{S}}(\mathrm{tr}) + \frac{1}{I_{\mathrm{T}}J_{\mathrm{T}}}\sum_{i=1}^{I}\sum_{j=1}^{J} W_i^{\mathrm{B}}W_j^{\mathrm{S}}\delta_{ij}^{\mathrm{BS}}(\mathrm{tr}),\\
\widehat{\overline{\overline{Y}}}_{\mathrm{ib}} &= \overline{\overline{y}}_{\mathrm{ib}} + \frac{1}{I_{\mathrm{T}}}\sum_{i=1}^{I} W_i^{\mathrm{B}}\delta_i^{\mathrm{B}}(\mathrm{ib}) + \frac{1}{J_{\mathrm{C}}}\sum_{j=1}^{J} (1-W_j^{\mathrm{S}})\delta_j^{\mathrm{S}}(\mathrm{ib}) + \frac{1}{I_{\mathrm{T}}J_{\mathrm{C}}}\sum_{i=1}^{I}\sum_{j=1}^{J} W_i^{\mathrm{B}}(1-W_j^{\mathrm{S}})\delta_{ij}^{\mathrm{BS}}(\mathrm{ib}),\\
\widehat{\overline{\overline{Y}}}_{\mathrm{is}} &= \overline{\overline{y}}_{\mathrm{is}} + \frac{1}{I_{\mathrm{C}}}\sum_{i=1}^{I} (1-W_i^{\mathrm{B}})\delta_i^{\mathrm{B}}(\mathrm{is}) + \frac{1}{J_{\mathrm{T}}}\sum_{j=1}^{J} W_j^{\mathrm{S}}\delta_j^{\mathrm{S}}(\mathrm{is}) + \frac{1}{I_{\mathrm{C}}J_{\mathrm{T}}}\sum_{i=1}^{I}\sum_{j=1}^{J} (1-W_i^{\mathrm{B}})W_j^{\mathrm{S}}\delta_{ij}^{\mathrm{BS}}(\mathrm{is}),\\
\widehat{\overline{\overline{Y}}}_{\mathrm{cc}} &= \overline{\overline{y}}_{\mathrm{cc}} + \frac{1}{I_{\mathrm{C}}}\sum_{i=1}^{I} (1-W_i^{\mathrm{B}})\delta_i^{\mathrm{B}}(\mathrm{cc}) + \frac{1}{J_{\mathrm{C}}}\sum_{j=1}^{J} (1-W_j^{\mathrm{S}})\delta_j^{\mathrm{S}}(\mathrm{cc})\\
&\quad + \frac{1}{I_{\mathrm{C}}J_{\mathrm{C}}}\sum_{i=1}^{I}\sum_{j=1}^{J} (1-W_i^{\mathrm{B}})(1-W_j^{\mathrm{S}})\delta_{ij}^{\mathrm{BS}}(\mathrm{cc}).
\end{aligned} \tag{A.1}$$

Proof of Lemma A.2. Consider the case of $\widehat{\overline{\overline{Y}}}_{\mathrm{tr}}$: leveraging the decomposition of $y_{ij}(\mathrm{tr})$,

$$\begin{aligned}
\widehat{\overline{\overline{Y}}}_{\mathrm{tr}} &= \frac{1}{I_{\mathrm{T}}J_{\mathrm{T}}}\sum_{i=1}^{I}\sum_{j=1}^{J} W_i^{\mathrm{B}}W_j^{\mathrm{S}}y_{ij}(\mathrm{tr}) = \frac{1}{I_{\mathrm{T}}J_{\mathrm{T}}}\sum_{i=1}^{I}\sum_{j=1}^{J} W_i^{\mathrm{B}}W_j^{\mathrm{S}}\left(\overline{\overline{y}}_{\mathrm{tr}} + \delta_i^{\mathrm{B}}(\mathrm{tr}) + \delta_j^{\mathrm{S}}(\mathrm{tr}) + \delta_{ij}^{\mathrm{BS}}(\mathrm{tr})\right)\\
&= \overline{\overline{y}}_{\mathrm{tr}} + \frac{1}{I_{\mathrm{T}}}\sum_{i=1}^{I} W_i^{\mathrm{B}}\delta_i^{\mathrm{B}}(\mathrm{tr}) + \frac{1}{J_{\mathrm{T}}}\sum_{j=1}^{J} W_j^{\mathrm{S}}\delta_j^{\mathrm{S}}(\mathrm{tr}) + \frac{1}{I_{\mathrm{T}}J_{\mathrm{T}}}\sum_{i=1}^{I}\sum_{j=1}^{J} W_i^{\mathrm{B}}W_j^{\mathrm{S}}\delta_{ij}^{\mathrm{BS}}(\mathrm{tr}).
\end{aligned}$$

Results for $\gamma \neq \mathrm{tr}$ are similar and are omitted. ☐

A.3 Moment characterization

We use lemma A.2 to re-write the estimator $\widehat{\overline{\overline{Y}}}_\gamma$ of $\overline{\overline{y}}_\gamma$ as a linear combination of the random labels W_i^{B}, W_j^{S} with non-stochastic coefficients. We use this to derive the first two moments of $(\widehat{\overline{\overline{Y}}}_\gamma, \widehat{\overline{\overline{Y}}}_{\gamma'})$ under the SMRD design. To do so, we define the demeaned treatment $D_i^{\mathrm{B}} = W_i^{\mathrm{B}} - I_{\mathrm{T}}/I$, and $D_j^{\mathrm{S}} = W_j^{\mathrm{S}} - J_{\mathrm{T}}/J$.

Lemma A.3. *For* $i \neq i' \in [I]$, $\mathbb{E}\left[D_i^{\mathrm{B}}\right] = 0$, $\mathrm{Var}\left(D_i^{\mathrm{B}}\right) = \frac{I_{\mathrm{C}}I_{\mathrm{T}}}{I^2}$, $\mathrm{Cov}(D_i^{\mathrm{B}}, D_{i'}^{\mathrm{B}}) = -\frac{I_{\mathrm{C}}I_{\mathrm{T}}}{I^2(I-1)}$. *For* $j, j' \in [J]$, $j \neq j'$, $\mathbb{E}[D_j^{\mathrm{S}}] = 0$, $\mathrm{Var}(D_j^{\mathrm{S}}) = \frac{J_{\mathrm{C}}J_{\mathrm{T}}}{J^2}$, $\mathrm{Cov}(D_j^{\mathrm{S}}, D_{j'}^{\mathrm{S}}) = -\frac{J_{\mathrm{C}}J_{\mathrm{T}}}{J^2(J-1)}$. *Finally, because* D_i^{B} *and* D_j^{S} *are independent, we have* $\mathrm{Cov}(D_i^{\mathrm{B}}, D_j^{\mathrm{S}}) = 0$, $\forall\, i, j$.

Proof of Lemma A.3. W_i^{B} is a Bernoulli random variable with bias given by $p^{\mathrm{B}} = I_{\mathrm{T}}/I$, hence $\mathbb{E}\left[D_i^{\mathrm{B}}\right]=0$. Moreover, $\mathrm{Var}\left(D_i^{\mathrm{B}}\right)=\mathrm{Var}\left(W_i^{\mathrm{B}}\right)=\frac{I_{\mathrm{T}}}{I}\left(1-\frac{I_{\mathrm{T}}}{I}\right)=\frac{I_{\mathrm{C}}I_{\mathrm{T}}}{I^2}$. Last,

$$\mathrm{Cov}(D_i^{\mathrm{B}},D_{i'}^{\mathrm{B}})=\mathbb{E}[W_i^{\mathrm{B}}W_{i'}^{\mathrm{B}}]-\mathbb{E}[W_i^{\mathrm{B}}]\,\mathbb{E}[W_{i'}^{\mathrm{B}}]=\frac{I_{\mathrm{T}}}{I}\frac{I_{\mathrm{T}}-1}{I-1}-\frac{I_{\mathrm{T}}^2}{I^2}=-\frac{I_{\mathrm{C}}I_{\mathrm{T}}}{I^2(I-1)}.$$

Corresponding proofs for D_j^{S} are analogous and omitted. □

Note that the covariance between D_i^{B} and $D_{i'}^{\mathrm{B}}$ for $i\neq i'$ differs from zero because we fix the number of selected buyers at I_{T}, rather than tossing a coin for each buyer. Fixing the number of selected buyers is important for getting exact finite sample results for the variances. Define the average residuals by assignment type, for $\gamma\in\{\mathrm{cc},\mathrm{ib},\mathrm{is},\mathrm{tr}\}$:

$$\overline{\varepsilon}_{\gamma}^{\mathrm{B}}=\frac{1}{I_{\mathrm{T}}}\sum_{i=1}^{I}D_i^{\mathrm{B}}\delta_i^{\mathrm{B}}(\gamma),\quad \overline{\varepsilon}_{\gamma}^{\mathrm{S}}=\frac{1}{J_{\mathrm{T}}}\sum_{j=1}^{J}D_j^{\mathrm{S}}\delta_j^{\mathrm{S}}(\gamma),\quad \overline{\overline{\varepsilon}}_{\gamma}^{\mathrm{BS}}=\frac{1}{I_{\mathrm{T}}J_{\mathrm{T}}}\sum_{i=1}^{I}\sum_{j=1}^{J}D_i^{\mathrm{B}}D_j^{\mathrm{S}}\delta_{ij}^{\mathrm{BS}}(\gamma).$$

These representations allow us to split the averages of observed values $\widehat{\overline{\overline{Y}}}_{\gamma}$ into deterministic and stochastic components.

Lemma A.4. *(a) The sample estimates $\widehat{\overline{\overline{Y}}}_{\gamma}$, $\gamma\in\{\mathrm{cc},\mathrm{ib},\mathrm{is},\mathrm{tr}\}$ can be written as the sums of four terms:*

$$\widehat{\overline{\overline{Y}}}_{\gamma}=\overline{\overline{y}}_{\gamma}+\overline{\varepsilon}_{\gamma}^{\mathrm{B}}+\overline{\varepsilon}_{\gamma}^{\mathrm{S}}+\overline{\overline{\varepsilon}}_{\gamma}^{\mathrm{BS}},$$

(b) $\forall\,\gamma\in\{\mathrm{cc},\mathrm{ib},\mathrm{is},\mathrm{tr}\}$, the ϵ in the decomposition above are mean-zero error terms:

$$\mathbb{E}\left[\overline{\varepsilon}_{\gamma}^{\mathrm{B}}\right]=\mathbb{E}\left[\overline{\varepsilon}_{\gamma}^{\mathrm{S}}\right]=\mathbb{E}\left[\overline{\overline{\varepsilon}}_{\gamma}^{\mathrm{BS}}\right]=0.$$

(c) For all $\gamma\neq\gamma'\in\{\mathrm{cc},\mathrm{ib},\mathrm{is},\mathrm{tr}\}$, the error terms above are uncorrelated:

$$\mathrm{Cov}\left(\overline{\varepsilon}_{\gamma}^{\mathrm{B}},\overline{\varepsilon}_{\gamma'}^{\mathrm{S}}\right)=\mathrm{Cov}\left(\overline{\varepsilon}_{\gamma}^{\mathrm{B}},\overline{\overline{\varepsilon}}_{\gamma'}^{\mathrm{BS}}\right)=\mathrm{Cov}\left(\overline{\varepsilon}_{\gamma}^{\mathrm{S}},\overline{\overline{\varepsilon}}_{\gamma'}^{\mathrm{BS}}\right)=0.$$

Before proving this lemma, let us just provide an intuition about the decomposition of the four averages $\widehat{\overline{\overline{Y}}}_{\mathrm{cc}}$, $\widehat{\overline{\overline{Y}}}_{\mathrm{ib}}$, $\widehat{\overline{\overline{Y}}}_{\mathrm{is}}$, and $\widehat{\overline{\overline{Y}}}_{\mathrm{tr}}$ described above, as this is a key step to obtaining the variance of the estimator for the average treatment effect. In particular, looking at (i), the first term $\overline{\overline{y}}_{\gamma}$ is deterministic (the unweighted average of potential outcomes over all pairs (i,j), not depending on the assignment). The other three terms, $\overline{\varepsilon}_{\gamma}^{\mathrm{B}}$, $\overline{\varepsilon}_{\gamma}^{\mathrm{S}}$, and $\overline{\overline{\varepsilon}}_{\gamma}^{\mathrm{BS}}$, are mutually uncorrelated stochastic terms with expectation equal to zero. The variances of the four averages will depend on the variances of the three stochastic terms, and the covariances will depend on the covariances of the corresponding stochastic terms, *e.g.,* the covariance of $\overline{\varepsilon}_{\mathrm{tr}}^{\mathrm{B}}$ and $\overline{\varepsilon}_{\mathrm{ib}}^{\mathrm{B}}$, or the covariance of $\overline{\overline{\varepsilon}}_{\mathrm{cc}}^{\mathrm{BS}}$ and $\overline{\overline{\varepsilon}}_{\mathrm{is}}^{\mathrm{BS}}$.

Proof of Lemma A.4. For part (a) consider $\widehat{\overline{\overline{Y}}}_{\mathrm{tr}}$. Now consider for the treated type the average of the observed outcomes, decomposed as in Lemma A.2:

$$\widehat{\overline{\overline{Y}}}_{\mathrm{tr}}=\overline{\overline{y}}_{\mathrm{tr}}+\frac{1}{I_{\mathrm{T}}}\sum_{i=1}^{I}W_i^{\mathrm{B}}\delta_i^{\mathrm{B}}(\mathrm{tr})+\frac{1}{J_{\mathrm{T}}}\sum_{j=1}^{J}W_j^{\mathrm{S}}\delta_j^{\mathrm{S}}(\mathrm{tr})+\frac{1}{I_{\mathrm{T}}J_{\mathrm{T}}}\sum_{i=1}^{I}\sum_{j=1}^{J}W_i^{\mathrm{B}}W_j^{\mathrm{S}}\delta_{ij}^{\mathrm{BS}}(\mathrm{tr}).$$

Via Lemma A.3, substituting $D_i^{\mathrm{B}} + I_{\mathrm{T}}/I$ for W_i^{B} and $D_j^{\mathrm{S}} + J_{\mathrm{T}}/J$ for W_j^{S}, we can write

$$\begin{aligned}\widehat{\overline{\overline{Y}}}_{\mathrm{tr}} &= \overline{\overline{y}}_{\mathrm{tr}} + \frac{1}{I_{\mathrm{T}}}\sum_{i=1}^{I}\left(D_i^{\mathrm{B}} + \frac{I_{\mathrm{T}}}{I}\right)\delta_i^{\mathrm{B}}(\mathrm{tr}) + \frac{1}{J_{\mathrm{T}}}\sum_{j=1}^{J}\left(D_j^{\mathrm{S}} + \frac{J_{\mathrm{T}}}{J}\right)\delta_j^{\mathrm{S}}(\mathrm{tr}) \\ &+ \frac{1}{I_{\mathrm{T}}J_{\mathrm{T}}}\sum_{i=1}^{I}\sum_{j=1}^{J}\left(D_i^{\mathrm{B}} + \frac{I_{\mathrm{T}}}{I}\right)\left(D_j^{\mathrm{S}} + \frac{J_{\mathrm{T}}}{J}\right)\delta_{ij}^{\mathrm{BS}}(\mathrm{tr}).\end{aligned}$$

By definition, $\delta_{ij}^{\mathrm{BS}}(\mathrm{tr})$, $\delta_i^{\mathrm{B}}(\mathrm{tr})$ and $\delta_j^{\mathrm{S}}(\mathrm{tr})$ sum to zero. Hence the equation above simplifies to

$$\widehat{\overline{\overline{Y}}}_{\mathrm{tr}} = \overline{\overline{y}}_{\mathrm{tr}} + \sum_{i=1}^{I}\frac{D_i^{\mathrm{B}}\delta_i^{\mathrm{B}}(\mathrm{tr})}{I_{\mathrm{T}}} + \sum_{j=1}^{J}\frac{D_j^{\mathrm{S}}\delta_j^{\mathrm{S}}(\mathrm{tr})}{J_{\mathrm{T}}} + \sum_{i=1}^{I}\sum_{j=1}^{J}\frac{D_i^{\mathrm{B}}D_j^{\mathrm{S}}\delta_{ij}^{\mathrm{BS}}(\mathrm{tr})}{I_{\mathrm{T}}J_{\mathrm{T}}} = \overline{\overline{y}}_{\mathrm{tr}} + \overline{\varepsilon}_{\mathrm{tr}}^{\mathrm{B}} + \overline{\varepsilon}_{\mathrm{tr}}^{\mathrm{S}} + \overline{\overline{\varepsilon}}_{\mathrm{tr}}^{\mathrm{BS}}.$$

This concludes the proof of the first part of (a). The proofs of the other parts of (a) follow the same argument and are omitted. Given part (a), (b) follows immediately because D_i^{B} and D_j^{S} have expectation equal to zero. The same holds for the covariances in (c). □

Unbiasedness results in Lemma 4.1 and Theorem 4.2 follow directly from Lemma A.4.

Lemma A.5 (Lemma 4.1 in the main paper)**.** *Consider a SMRD in which Assumption 2.4 holds. The plug-in estimators in Equation* (10) *satisfy*

$$\mathbb{E}\left[\widehat{\overline{\overline{Y}}}_{\gamma}\right] = \overline{\overline{y}}_{\gamma}, \ \forall\, \gamma \in \{\mathrm{cc}, \mathrm{ib}, \mathrm{is}, \mathrm{tr}\}.$$

Proof of Lemma 4.1. Apply Lemma A.4, and linearity of the expectation operator. □

Theorem A.6 (Already Theorem 4.2 in the main paper)**.** *Consider a SMRD where Assumption 2.4 holds. The plug-in estimators $\hat{\tau}(\vec{\boldsymbol{\beta}})$ for $\tau(\vec{\boldsymbol{\beta}})$ defined in Equation* (9) *satisfy*

$$\mathbb{E}\left[\hat{\tau}(\vec{\boldsymbol{\beta}})\right] = \tau(\vec{\boldsymbol{\beta}}), \quad \textit{with} \quad \hat{\tau}(\vec{\boldsymbol{\beta}}) := \beta_{\mathrm{cc}}\widehat{\overline{\overline{Y}}}_{\mathrm{cc}} + \beta_{\mathrm{ib}}\widehat{\overline{\overline{Y}}}_{\mathrm{ib}} + \beta_{\mathrm{is}}\widehat{\overline{\overline{Y}}}_{\mathrm{is}} + \beta_{\mathrm{tr}}\widehat{\overline{\overline{Y}}}_{\mathrm{tr}}.$$

Proof of Theorem 4.2. Apply Lemma 4.1, and linearity of the expectation operator. □

We now move to the variance characterization. For $\gamma \in \{\mathrm{cc}, \mathrm{ib}, \mathrm{is}, \mathrm{tr}\}$, recall the definitions of the population variances of $\delta_i^{\mathrm{B}}(\gamma)$ and $\delta_j^{\mathrm{S}}(\gamma)$ given in Section 4:

$$\sigma_{\gamma}^{\mathrm{B}} := \sum_{i=1}^{I}\frac{\left(\delta_i^{\mathrm{B}}(\gamma)\right)^2}{I}, \quad \sigma_{\gamma}^{\mathrm{S}} := \sum_{j=1}^{J}\frac{\left(\delta_j^{\mathrm{S}}(\gamma)\right)^2}{J}, \quad \sigma_{\gamma}^{\mathrm{BS}} := \sum_{i=1}^{I}\sum_{j=1}^{J}\frac{\left(\delta_{ij}^{\mathrm{BS}}(\gamma)\right)^2}{IJ}.$$

Lemma A.7. *For $\gamma \in \{\mathrm{cc}, \mathrm{ib}, \mathrm{is}, \mathrm{tr}\}$, the variance of $\widehat{\overline{\overline{Y}}}_{\gamma}$ is:*

$$\begin{aligned}\mathrm{Var}_{\gamma} := \mathrm{Var}\left(\widehat{\overline{\overline{Y}}}_{\gamma}\right) &= \frac{I - I_{\gamma}}{I_{\gamma}}\frac{1}{I-1}\sigma_{\gamma}^{\mathrm{B}} + \frac{J - J_{\gamma}}{J_{\gamma}}\frac{1}{J-1}\sigma_{\gamma}^{\mathrm{S}} + \frac{I - I_{\gamma}}{I_{\gamma}}\frac{1}{I-1}\frac{J - J_{\gamma}}{J_{\gamma}}\frac{1}{J-1}\sigma_{\gamma}^{\mathrm{BS}} \\ &= \alpha_{\gamma}^{\mathrm{B}}\sigma_{\gamma}^{\mathrm{B}} + \alpha_{\gamma}^{\mathrm{S}}\sigma_{\gamma}^{\mathrm{S}} + \alpha_{\gamma}^{\mathrm{B}}\alpha_{\gamma}^{\mathrm{S}}\sigma_{\gamma}^{\mathrm{BS}},\end{aligned}$$

where $\alpha_{\gamma}^{\mathrm{B}}$ and $\alpha_{\gamma}^{\mathrm{S}}$ where defined in eq. (14) *in the main text.*

Proof of Lemma A.7. We consider $\gamma = \text{tr}$, (i.e., $I_\gamma = I_\text{T}$, $J_\gamma = J_\text{T}$). For $\gamma = \text{tr}$, $I_\text{C} = I - I_\gamma$ and $J_\text{C} = J - J_\gamma$. We show the three following equalities hold:

$$\text{Var}\left(\bar{\varepsilon}^{\text{B}}_{\text{tr}}\right) = \frac{I_\text{C}}{I_\text{T}}\frac{1}{I-1}\sigma^{\text{B}}_{\text{tr}}, \quad \text{Var}\left(\bar{\varepsilon}^{\text{S}}_{\text{tr}}\right) = \frac{J_\text{C}}{J_\text{T}}\frac{1}{J-1}\sigma^{\text{S}}_{\text{tr}}, \quad \text{Var}\left(\bar{\bar{\varepsilon}}^{\text{BS}}_{\text{tr}}\right) = \frac{I_\text{C}}{I_\text{T}}\frac{1}{I-1}\frac{J_\text{C}}{J_\text{T}}\frac{1}{J-1}\sigma^{\text{BS}}_{\text{tr}}. \tag{A.2}$$

Because Lemma A.4 implies that $\text{Var}_{\text{tr}} = \text{Var}\left(\hat{\bar{\bar{Y}}}_{\text{tr}}\right) = \text{Var}\left(\bar{\varepsilon}^{\text{B}}_{\text{tr}}\right) + \text{Var}\left(\bar{\varepsilon}^{\text{S}}_{\text{tr}}\right) + \text{Var}\left(\bar{\bar{\varepsilon}}^{\text{BS}}_{\text{tr}}\right)$, showing the three equalities in eq. (A.2) yields the thesis.

$$\begin{aligned}
\text{Var}\left(\bar{\varepsilon}^{\text{B}}_{\text{tr}}\right) &= \mathbb{E}\left[\left(\frac{1}{I_\text{T}}\sum_{i=1}^{I} D^{\text{B}}_i \delta^{\text{B}}_i(\text{tr})\right)^2\right] = \frac{1}{I_\text{T}^2}\mathbb{E}\left[\sum_{i=1}^{I}\sum_{i'=1}^{I} D^{\text{B}}_i D^{\text{B}}_{i'} \delta^{\text{B}}_i(\text{tr})\delta^{\text{B}}_{i'}(\text{tr})\right] \\
&= \frac{1}{I_\text{T}^2}\sum_{i=1}^{I}\sum_{i'=1}^{I}\mathbb{E}\left[D^{\text{B}}_i D^{\text{B}}_{i'}\right]\delta^{\text{B}}_i(\text{tr})\delta^{\text{B}}_{i'}(\text{tr}) \\
&= \frac{1}{I_\text{T}^2}\sum_{i=1}^{I}\mathbb{E}\left[(D^{\text{B}}_i)^2\right]\delta^{\text{B}}_i(\text{tr}) + \frac{1}{I_\text{T}^2}\sum_{i=1}^{I}\sum_{i'\neq i}\mathbb{E}\left[D^{\text{B}}_i D^{\text{B}}_{i'}\right]\delta^{\text{B}}_i(\text{tr})\delta^{\text{B}}_{i'}(\text{tr}) \\
&= \frac{1}{I_\text{T}^2}\sum_{i=1}^{I}\frac{I_\text{C}I_\text{T}}{I^2}\left(\delta^{\text{B}}_i(\text{tr})\right)^2 - \frac{1}{I_\text{T}^2}\sum_{i=1}^{I}\sum_{i'\neq i}\frac{I_\text{T}I_\text{C}}{I^2(I-1)}\delta^{\text{B}}_i(\text{tr})\delta^{\text{B}}_{i'}(\text{tr}) \\
&= \frac{1}{I_\text{T}^2}\sum_{i=1}^{I}\frac{I_\text{C}I_\text{T}}{I^2}\left(\delta^{\text{B}}_i(\text{tr})\right)^2 - \frac{1}{I_\text{T}^2}\sum_{i=1}^{I}\sum_{i'=1}^{I}\frac{I_\text{T}I_\text{C}}{I^2(I-1)}\delta^{\text{B}}_i(\text{tr})\delta^{\text{B}}_{i'}(\text{tr}) + \frac{1}{I_\text{T}^2}\sum_{i=1}^{I}\frac{I_\text{T}I_\text{C}}{I^2(I-1)}(\delta^{\text{B}}_i(\text{tr}))^2.
\end{aligned}$$

Because $\sum_i \delta^{\text{B}}_i(\text{tr}) = 0$, the term above involving the double sum is equal to zero:

$$\text{Var}\left(\bar{\varepsilon}^{\text{B}}_{\text{tr}}\right) = \frac{1}{I_\text{T}^2}\frac{I_\text{T}I_\text{C}}{I^2(I-1)}\sum_{i=1}^{I}(\delta^{\text{B}}_i(\text{tr}))^2 + \frac{1}{I_\text{T}^2}\frac{I_\text{T}I_\text{C}}{I^2}\sum_{i=1}^{I}\left(\delta^{\text{B}}_i(\text{tr})\right)^2 = \frac{I_\text{C}}{I_\text{T}}\frac{1}{I-1}\sigma^{\text{B}}_{\text{tr}}.$$

The second equality in eq. (A.2) is proved analogously. For the last equality in eq. (A.2),

$$\begin{aligned}
\text{Var}^{\text{BS}}_{\text{tr}} := \text{Var}\left(\bar{\bar{\varepsilon}}^{\text{BS}}_{\text{tr}}\right) &= \text{Var}\left(\frac{1}{I_\text{T}J_\text{T}}\sum_{i=1}^{I}\sum_{j=1}^{J} D^{\text{B}}_i D^{\text{S}}_j \delta^{\text{BS}}_{ij}(\text{tr})\right) \\
&= \mathbb{E}\left[\frac{1}{I_\text{T}^2 J_\text{T}^2}\sum_{i=1}^{I}\sum_{i'=1}^{I}\sum_{j=1}^{J}\sum_{j'=1}^{J} D^{\text{B}}_i D^{\text{B}}_{i'} D^{\text{S}}_j D^{\text{S}}_{j'} \delta^{\text{BS}}_{ij}(\text{tr})\delta^{\text{BS}}_{i',j'}(\text{tr})\right].
\end{aligned}$$

By independence of D^{B}_i and D^{S}_j, this is equal to

$$\text{Var}^{\text{BS}}_{\text{tr}} = \frac{1}{I_\text{T}^2 J_\text{T}^2}\sum_{i=1}^{I}\sum_{i'=1}^{I}\mathbb{E}\left[D^{\text{B}}_i D^{\text{B}}_{i'}\right]\sum_{j=1}^{J}\sum_{j'=1}^{J}\mathbb{E}\left[D^{\text{S}}_j D^{\text{S}}_{j'}\right]\delta^{\text{BS}}_{ij}(\text{tr})\delta^{\text{BS}}_{i',j'}(\text{tr}).$$

Now we expand the four-way sum above, noting that it is either the case that (a) : $i = i'$

and $j = j'$, (b) : $i = i'$ and $j \neq j'$, (c) : $i \neq i'$ and $j = j'$ or (d) : $i \neq i'$ and $j \neq j'$.

$$
\begin{aligned}
\mathrm{Var}^{\mathrm{BS}}_{\mathrm{tr}} &\overset{(a)}{=} \frac{1}{I_{\mathrm{T}}^2 J_{\mathrm{T}}^2} \sum_{i=1}^{I} \sum_{j=1}^{J} \mathbb{E}[(D_i^{\mathrm{B}})^2]\, \mathbb{E}[(D_j^{\mathrm{S}})^2] \left(\delta_{ij}^{\mathrm{BS}}(\mathrm{tr})\right)^2 \\
&\overset{(b)}{+} \frac{1}{I_{\mathrm{T}}^2 J_{\mathrm{T}}^2} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{j' \neq j}^{J} \mathbb{E}[(D_i^{\mathrm{B}})^2]\, \mathbb{E}[D_j^{\mathrm{S}} D_{j'}^{\mathrm{S}}] \delta_{ij}^{\mathrm{BS}}(\mathrm{tr}) \delta_{i,j'}^{\mathrm{BS}}(\mathrm{tr}) \\
&\overset{(c)}{+} \frac{1}{I_{\mathrm{T}}^2 J_{\mathrm{T}}^2} \sum_{i=1}^{I} \sum_{i' \neq i}^{I} \sum_{j=1}^{J} \mathbb{E}[D_i^{\mathrm{B}} D_{i'}^{\mathrm{B}}]\, \mathbb{E}[(D_j^{\mathrm{S}})^2] \delta_{ij}^{\mathrm{BS}}(\mathrm{tr}) \delta_{i',j}^{\mathrm{BS}}(\mathrm{tr}) \\
&\overset{(d)}{+} \frac{1}{I_{\mathrm{T}}^2 J_{\mathrm{T}}^2} \sum_{i=1}^{I} \sum_{i' \neq i}^{I} \sum_{j=1}^{J} \sum_{j' \neq j}^{J} \mathbb{E}[D_i^{\mathrm{B}} D_{i'}^{\mathrm{B}}]\, \mathbb{E}[D_j^{\mathrm{S}} D_{j'}^{\mathrm{S}}] \delta_{ij}^{\mathrm{BS}}(\mathrm{tr}) \delta_{i',j'}^{\mathrm{BS}}(\mathrm{tr}).
\end{aligned}
$$

Now we "complete" each of the last "incomplete" sums (b), (c), (d). For (b):

$$
\begin{aligned}
\sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{j' \neq j}^{J} \frac{\mathbb{E}[(D_i^{\mathrm{B}})^2]\, \mathbb{E}[D_j^{\mathrm{S}} D_{j'}^{\mathrm{S}}]}{I_{\mathrm{T}}^2 J_{\mathrm{T}}^2} \delta_{ij}^{\mathrm{BS}}(\mathrm{tr}) \delta_{i,j'}^{\mathrm{BS}}(\mathrm{tr}) &= -\frac{\left(\frac{I_{\mathrm{T}} I_{\mathrm{C}}}{I^2}\right)\left(\frac{J_{\mathrm{T}} J_{\mathrm{C}}}{J^2(J-1)}\right)}{I_{\mathrm{T}}^2 J_{\mathrm{T}}^2} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{j' \neq j}^{J} \delta_{ij}^{\mathrm{BS}}(\mathrm{tr}) \delta_{ij'}^{\mathrm{BS}}(\mathrm{tr}) \\
&= -\frac{\left(\frac{I_{\mathrm{T}} I_{\mathrm{C}}}{I^2}\right)\left(\frac{J_{\mathrm{T}} J_{\mathrm{C}}}{J^2(J-1)}\right)}{I_{\mathrm{T}}^2 J_{\mathrm{T}}^2} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{j'=1}^{J} \delta_{ij}^{\mathrm{BS}}(\mathrm{tr}) \delta_{ij'}^{\mathrm{BS}}(\mathrm{tr}) \\
&+ \frac{\left(\frac{I_{\mathrm{T}} I_{\mathrm{C}}}{I^2}\right)\left(\frac{J_{\mathrm{T}} J_{\mathrm{C}}}{J^2(J-1)}\right)}{I_{\mathrm{T}}^2 J_{\mathrm{T}}^2} \sum_{i=1}^{I} \sum_{j=1}^{J} (\delta_{ij}^{\mathrm{BS}}(\mathrm{tr}))^2 \\
&= \frac{1}{I_{\mathrm{T}}^2 J_{\mathrm{T}}^2} \left(\frac{I_{\mathrm{T}} I_{\mathrm{C}}}{I^2}\right)\left(\frac{J_{\mathrm{T}} J_{\mathrm{C}}}{J^2(J-1)}\right) \sum_{i=1}^{I} \sum_{j=1}^{J} (\delta_{ij}^{\mathrm{BS}}(\mathrm{tr}))^2,
\end{aligned}
$$

where we observe that $\sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{j'=1}^{J} \delta_{ij}^{\mathrm{BS}}(\mathrm{tr}) \delta_{ij'}^{\mathrm{BS}}(\mathrm{tr}) = 0$. A similar derivation allows us to "complete" (c), yielding:

$$
\sum_{i=1}^{I} \sum_{i' \neq i}^{I} \sum_{j=1}^{J} \frac{\mathbb{E}[D_i^{\mathrm{B}} D_{i'}^{\mathrm{B}}]\, \mathbb{E}[(D_j^{\mathrm{S}})^2] \delta_{ij}^{\mathrm{BS}}(\mathrm{tr}) \delta_{i'j}^{\mathrm{BS}}(\mathrm{tr})}{I_{\mathrm{T}}^2 J_{\mathrm{T}}^2} = \frac{\left(\frac{I_{\mathrm{T}} I_{\mathrm{C}}}{I^2(I-1)}\right)\left(\frac{J_{\mathrm{T}} J_{\mathrm{C}}}{J^2}\right)}{I_{\mathrm{T}}^2 J_{\mathrm{T}}^2} \sum_{i=1}^{I} \sum_{j=1}^{J} (\delta_{ij}^{\mathrm{BS}}(\mathrm{tr}))^2.
$$

Last, for (d),

$$
\sum_{i=1}^{I} \sum_{i' \neq i}^{I} \sum_{j=1}^{J} \sum_{j \neq j'}^{J} \frac{\mathbb{E}[D_i^{\mathrm{B}} D_{i'}^{\mathrm{B}}]\, \mathbb{E}[D_j^{\mathrm{S}} D_{j'}^{\mathrm{S}}] \delta_{ij}^{\mathrm{BS}}(\mathrm{tr}) \delta_{i'j}^{\mathrm{BS}}(\mathrm{tr})}{I_{\mathrm{T}}^2 J_{\mathrm{T}}^2} = \frac{\left(\frac{I_{\mathrm{T}} I_{\mathrm{C}}}{I^2(I-1)}\right)\left(\frac{J_{\mathrm{T}} J_{\mathrm{C}}}{J^2(J-1)}\right)}{I_{\mathrm{T}}^2 J_{\mathrm{T}}^2} \sum_{i=1}^{I} \sum_{j=1}^{J} (\delta_{ij}^{\mathrm{BS}}(\mathrm{tr}))^2.
$$

Plugging these back in $\mathrm{Var}_{\mathrm{tr}}^{\mathrm{BS}}$,

$$\begin{aligned}\mathrm{Var}_{\mathrm{tr}}^{\mathrm{BS}} &= \frac{1}{I_{\mathrm{T}}^2 J_{\mathrm{T}}^2} \frac{I_{\mathrm{C}} I_{\mathrm{T}} J_{\mathrm{C}} J_{\mathrm{T}}}{I^2 J^2} \left[1 + \frac{1}{I-1} + \frac{1}{J-1} + \frac{1}{(I-1)(J-1)}\right] \sum_{i=1}^{I} \sum_{j=1}^{J} (\delta_{ij}^{\mathrm{BS}}(\mathrm{tr}))^2 \\ &= \frac{1}{I_{\mathrm{T}}^2 J_{\mathrm{T}}^2} \frac{I_{\mathrm{C}} I_{\mathrm{T}} J_{\mathrm{C}} J_{\mathrm{T}}}{I^2 J^2} \left[\frac{IJ}{(I-1)(J-1)}\right] \sum_{i=1}^{I} \sum_{j=1}^{J} (\delta_{ij}^{\mathrm{BS}}(\mathrm{tr}))^2 = \frac{I_{\mathrm{C}}}{I_{\mathrm{T}}} \frac{1}{I-1} \frac{J_{\mathrm{C}}}{J_{\mathrm{T}}} \frac{1}{J-1} \sigma_{\mathrm{tr}}^{\mathrm{BS}}.\end{aligned}$$

□

In order to characterize the variance of the spillover effects, we need to characterize the covariance between the estimators $\widehat{\overline{\overline{Y}}}_{\gamma}, \widehat{\overline{\overline{Y}}}_{\gamma'}$, for $\gamma, \gamma' \in \{\mathrm{cc}, \mathrm{ib}, \mathrm{is}, \mathrm{tr}\}$. Recall the definitions provided in Section 4: for all $\gamma \neq \gamma' \in \{\mathrm{cc}, \mathrm{ib}, \mathrm{is}, \mathrm{tr}\}$ for buyers and the sellers

$$\xi_{\gamma,\gamma'}^{\mathrm{B}} := \sum_{i=1}^{I} \frac{\left(\delta_i^{\mathrm{B}}(\gamma) - \delta_i^{\mathrm{B}}(\gamma')\right)^2}{I}, \quad \xi_{\gamma,\gamma'}^{\mathrm{S}} := \sum_{j=1}^{J} \frac{\left(\delta_j^{\mathrm{S}}(\gamma) - \delta_j^{\mathrm{S}}(\gamma')\right)^2}{J}, \quad \xi_{\gamma,\gamma'}^{\mathrm{BS}} := \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{\left(\delta_{ij}^{\mathrm{BS}}(\gamma) - \delta_{ij}^{\mathrm{BS}}(\gamma')\right)^2}{IJ}.$$

Lemma A.8. *For $\gamma \neq \gamma' \in \{\mathrm{cc}, \mathrm{ib}, \mathrm{is}, \mathrm{tr}\}$, covariances of type estimators are*

$$\begin{aligned}\mathrm{Cov}_{\mathrm{tr},\mathrm{ib}} &:= \mathrm{Cov}\left(\widehat{\overline{\overline{Y}}}_{\mathrm{tr}}, \widehat{\overline{\overline{Y}}}_{\mathrm{ib}}\right) \\ &= \frac{I_{\mathrm{C}}}{2I_{\mathrm{T}}(I-1)} \left(\sigma_{\mathrm{tr}}^{\mathrm{B}} + \sigma_{\mathrm{ib}}^{\mathrm{B}} - \xi_{\mathrm{tr},\mathrm{ib}}^{\mathrm{B}}\right) - \frac{1}{2(J-1)} \left(\sigma_{\mathrm{tr}}^{\mathrm{S}} + \sigma_{\mathrm{ib}}^{\mathrm{S}} - \xi_{\mathrm{tr},\mathrm{ib}}^{\mathrm{S}}\right) \\ &\quad - \frac{I_{\mathrm{C}}}{2I_{\mathrm{T}}(I-1)(J-1)} \left(\sigma_{\mathrm{tr}}^{\mathrm{BS}} + \sigma_{\mathrm{ib}}^{\mathrm{BS}} - \xi_{\mathrm{tr},\mathrm{ib}}^{\mathrm{BS}}\right).\end{aligned}$$

Similarly,

$$\begin{aligned}\mathrm{Cov}_{\mathrm{tr},\mathrm{is}} &:= \mathrm{Cov}\left(\widehat{\overline{\overline{Y}}}_{\mathrm{tr}}, \widehat{\overline{\overline{Y}}}_{\mathrm{is}}\right) \\ &= -\frac{1}{2(I-1)} \left(\sigma_{\mathrm{tr}}^{\mathrm{B}} + \sigma_{\mathrm{is}}^{\mathrm{B}} - \xi_{\mathrm{tr},\mathrm{is}}^{\mathrm{B}}\right) + \frac{J_{\mathrm{C}}}{2J_{\mathrm{T}}(J-1)} \left(\sigma_{\mathrm{tr}}^{\mathrm{S}} + \sigma_{\mathrm{is}}^{\mathrm{S}} - \xi_{\mathrm{tr},\mathrm{is}}^{\mathrm{S}}\right) \\ &\quad - \frac{J_{\mathrm{C}}}{2IJ_{\mathrm{T}}(J-1)} \left(\sigma_{\mathrm{tr}}^{\mathrm{BS}} + \sigma_{\mathrm{is}}^{\mathrm{BS}} - \xi_{\mathrm{tr},\mathrm{is}}^{\mathrm{BS}}\right),\end{aligned}$$

$$\begin{aligned}\mathrm{Cov}_{\mathrm{tr},\mathrm{cc}} &:= \mathrm{Cov}\left(\widehat{\overline{\overline{Y}}}_{\mathrm{tr}}, \widehat{\overline{\overline{Y}}}_{\mathrm{cc}}\right) \\ &= -\frac{1}{2(I-1)} \left(\sigma_{\mathrm{tr}}^{\mathrm{B}} + \sigma_{\mathrm{cc}}^{\mathrm{B}} - \xi_{\mathrm{tr},\mathrm{cc}}^{\mathrm{B}}\right) - \frac{1}{2(J-1)} \left(\sigma_{\mathrm{tr}}^{\mathrm{S}} + \sigma_{\mathrm{cc}}^{\mathrm{S}} - \xi_{\mathrm{tr},\mathrm{cc}}^{\mathrm{S}}\right) \\ &\quad + \frac{1}{2(I-1)(J-1)} \left(\sigma_{\mathrm{tr}}^{\mathrm{BS}} + \sigma_{\mathrm{cc}}^{\mathrm{BS}} - \xi_{\mathrm{tr},\mathrm{cc}}^{\mathrm{BS}}\right),\end{aligned}$$

$$
\begin{aligned}
\mathrm{Cov}_{\mathrm{ib},\mathrm{is}} &:= \mathrm{Cov}\left(\widehat{\overline{\overline{Y}}}_{\mathrm{ib}}, \widehat{\overline{\overline{Y}}}_{\mathrm{is}}\right) \\
&= -\frac{1}{2(I-1)}\left(\sigma^{\mathrm{B}}_{\mathrm{ib}} + \sigma^{\mathrm{B}}_{\mathrm{is}} - \xi^{\mathrm{B}}_{\mathrm{ib},\mathrm{is}}\right) - \frac{1}{2(J-1)}\left(\sigma^{\mathrm{S}}_{\mathrm{ib}} + \sigma^{\mathrm{S}}_{\mathrm{is}} - \xi^{\mathrm{S}}_{\mathrm{ib},\mathrm{is}}\right) \\
&+ \frac{1}{2(I-1)(J-1)}\left(\sigma^{\mathrm{BS}}_{\mathrm{ib}} + \sigma^{\mathrm{BS}}_{\mathrm{is}} - \xi^{\mathrm{BS}}_{\mathrm{ib},\mathrm{is}}\right),
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Cov}_{\mathrm{ib},\mathrm{cc}} &:= \mathrm{Cov}\left(\widehat{\overline{\overline{Y}}}_{\mathrm{ib}}, \widehat{\overline{\overline{Y}}}_{\mathrm{cc}}\right) \\
&= -\frac{1}{2(I-1)}\left(\sigma^{\mathrm{B}}_{\mathrm{ib}} + \sigma^{\mathrm{B}}_{\mathrm{cc}} - \xi^{\mathrm{B}}_{\mathrm{ib},\mathrm{cc}}\right) - \frac{J_{\mathrm{C}}}{2J_{\mathrm{T}}(J-1)}\left(\sigma^{\mathrm{S}}_{\mathrm{ib}} + \sigma^{\mathrm{S}}_{\mathrm{cc}} - \xi^{\mathrm{S}}_{\mathrm{ib},\mathrm{cc}}\right) \\
&- \frac{J_{\mathrm{C}}}{2(I-1)J_{\mathrm{T}}(J-1)}\left(\sigma^{\mathrm{BS}}_{\mathrm{ib}} + \sigma^{\mathrm{BS}}_{\mathrm{cc}} - \xi^{\mathrm{BS}}_{\mathrm{ib},\mathrm{cc}}\right),
\end{aligned}
$$

and last

$$
\begin{aligned}
\mathrm{Cov}_{\mathrm{is},\mathrm{cc}} &:= \mathrm{Cov}\left(\widehat{\overline{\overline{Y}}}_{\mathrm{is}}, \widehat{\overline{\overline{Y}}}_{\mathrm{cc}}\right) \\
&= \frac{I_{\mathrm{C}}}{2I_{\mathrm{T}}(I-1)}\left(\sigma^{\mathrm{B}}_{\mathrm{is}} + \sigma^{\mathrm{B}}_{\mathrm{cc}} - \xi^{\mathrm{B}}_{\mathrm{is},\mathrm{cc}}\right) - \frac{1}{2(J-1)}\left(\sigma^{\mathrm{S}}_{\mathrm{is}} + \sigma^{\mathrm{S}}_{\mathrm{cc}} - \xi^{\mathrm{S}}_{\mathrm{is},\mathrm{cc}}\right) \\
&- \frac{I_{\mathrm{C}}}{2I_{\mathrm{T}}(I-1)(J-1)}\left(\sigma^{\mathrm{BS}}_{\mathrm{is}} + \sigma^{\mathrm{BS}}_{\mathrm{cc}} - \xi^{\mathrm{BS}}_{\mathrm{is},\mathrm{cc}}\right).
\end{aligned}
$$

Proof of Lemma A.8. We show the three following equalities:

$$
\mathrm{Cov}^{\mathrm{B}}_{\mathrm{tr},\mathrm{ib}} := \mathrm{Cov}\left(\overline{\varepsilon}^{\mathrm{B}}_{\mathrm{tr}}, \overline{\varepsilon}^{\mathrm{B}}_{\mathrm{ib}}\right) = \frac{I_{\mathrm{C}}}{2I_{\mathrm{T}}(I-1)}\left(\sigma^{\mathrm{B}}_{\mathrm{tr}} + \sigma^{\mathrm{B}}_{\mathrm{ib}} - \xi^{\mathrm{B}}_{\mathrm{tr},\mathrm{ib}}\right), \tag{A.3}
$$

$$
\mathrm{Cov}^{\mathrm{S}}_{\mathrm{tr},\mathrm{ib}} := \mathrm{Cov}\left(\overline{\varepsilon}^{\mathrm{S}}_{\mathrm{tr}}, \overline{\varepsilon}^{\mathrm{S}}_{\mathrm{ib}}\right) = \frac{1}{2(J-1)}\left(\sigma^{\mathrm{S}}_{\mathrm{tr}} + \sigma^{\mathrm{S}}_{\mathrm{ib}} - \xi^{\mathrm{S}}_{\mathrm{tr},\mathrm{ib}}\right), \tag{A.4}
$$

and

$$
\mathrm{Cov}^{\mathrm{BS}}_{\mathrm{tr},\mathrm{ib}} := \mathrm{Cov}\left(\overline{\overline{\varepsilon}}^{\mathrm{BS}}_{\mathrm{tr}}, \overline{\overline{\varepsilon}}^{\mathrm{BS}}_{\mathrm{ib}}\right) = \frac{I_{\mathrm{C}}}{2I_{\mathrm{T}}(I-1)(J-1)}\left(\sigma^{\mathrm{BS}}_{\mathrm{tr}} + \sigma^{\mathrm{BS}}_{\mathrm{ib}} - \xi^{\mathrm{BS}}_{\mathrm{tr},\mathrm{ib}}\right). \tag{A.5}
$$

In combination with the fact that

$$
\mathrm{Cov}\left(\widehat{\overline{\overline{Y}}}_{\mathrm{tr}}, \widehat{\overline{\overline{Y}}}_{\mathrm{ib}}\right) = \mathrm{Cov}\left(\overline{\varepsilon}^{\mathrm{B}}_{\mathrm{tr}}, \overline{\varepsilon}^{\mathrm{B}}_{\mathrm{ib}}\right) - \mathrm{Cov}\left(\overline{\varepsilon}^{\mathrm{S}}_{\mathrm{tr}}, \overline{\varepsilon}^{\mathrm{S}}_{\mathrm{ib}}\right) - \mathrm{Cov}\left(\overline{\overline{\varepsilon}}^{\mathrm{BS}}_{\mathrm{tr}}, \overline{\overline{\varepsilon}}^{\mathrm{BS}}_{\mathrm{ib}}\right),
$$

this proves the first result.

First (A.3):

$$\begin{aligned}
\text{Cov}^{\text{B}}_{\text{tr},\text{ib}} &= \mathbb{E}\left[\left(\frac{1}{I_{\text{T}}}\sum_{i=1}^{I} D_i^{\text{B}}\delta_i^{\text{B}}(\text{tr})\right)\left(\frac{1}{I_{\text{T}}}\sum_{i=1}^{I} D_i^{\text{B}}\delta_i^{\text{B}}(\text{ib})\right)\right] = \mathbb{E}\left[\frac{1}{I_{\text{T}}^2}\sum_{i=1}^{I}\sum_{i'=1}^{I} D_i^{\text{B}} D_{i'}^{\text{B}}\delta_i^{\text{B}}(\text{tr})\delta_{i'}^{\text{B}}(\text{ib})\right] \\
&= \frac{1}{I_{\text{T}}^2}\sum_{i=1}^{I}\sum_{i'=1}^{I}\mathbb{E}\left[D_i^{\text{B}} D_{i'}^{\text{B}}\right]\delta_i^{\text{B}}(\text{tr})\delta_{i'}^{\text{B}}(\text{ib}) \\
&= -\frac{1}{I_{\text{T}}^2}\sum_{i=1}^{I}\sum_{i'=1}^{I}\frac{I_{\text{C}} I_{\text{T}}}{I^2(I-1)}\delta_i^{\text{B}}(\text{tr})\delta_{i'}^{\text{B}}(\text{ib}) + \frac{1}{I_{\text{T}}^2}\sum_{i=1}^{I}\left(\frac{I_{\text{C}} I_{\text{T}}}{I^2(I-1)} + \frac{I_{\text{C}} I_{\text{T}}}{I^2}\right)\delta_i^{\text{B}}(\text{tr})\delta_i^{\text{B}}(\text{ib}).
\end{aligned}$$

Because $\sum_i \delta_i^{\text{B}}(\text{tr}) = 0$ the first term is equal to zero. Thus,

$$\text{Cov}^{\text{B}}_{\text{tr},\text{ib}} = \frac{I_{\text{C}}}{I_{\text{T}} I}\left(\frac{1}{I-1}\sum_{i=1}^{I}\delta_i^{\text{B}}(\text{tr})\delta_i^{\text{B}}(\text{ib})\right).$$

Because

$$\begin{aligned}
\sigma^{\text{B}}_{\text{tr},\text{ib}} &= \frac{1}{I}\sum_{i=1}^{I}\left(\delta_i^{\text{B}}(\text{tr}) - \delta_i^{\text{B}}(\text{ib})\right)^2 \\
&= \frac{1}{I}\sum_{i=1}^{I}\left(\delta_i^{\text{B}}(\text{tr})\right)^2 + \frac{1}{I}\sum_{i=1}^{I}\left(\delta_i^{\text{B}}(\text{ib})\right)^2 - \frac{2}{I}\sum_{i=1}^{I}\left(\delta_i^{\text{B}}(\text{tr})\delta_i^{\text{B}}(\text{ib})\right)^2 \\
&= \sigma^{\text{B}}_{\text{tr}} + \sigma^{\text{B}}_{\text{ib}} - \frac{2I_{\text{T}}(I-1)}{I_{\text{C}}}\text{Cov}^{\text{B}}_{\text{tr},\text{ib}},
\end{aligned}$$

we have

$$\text{Cov}^{\text{B}}_{\text{tr},\text{ib}} = \frac{I_{\text{C}}}{2I_{\text{T}}(I-1)}\left(\sigma^{\text{B}}_{\text{tr}} + \sigma^{\text{B}}_{\text{ib}} - \xi^{\text{B}}_{\text{tr},\text{ib}}\right).$$

This completes the proof of (A.3). Similarly, to prove (A.4), we have

$$\begin{aligned}
\text{Cov}^{\text{S}}_{\text{tr},\text{ib}} &= \mathbb{E}\left[\left(\frac{1}{J_{\text{T}}}\sum_{j=1}^{J} D_j^{\text{S}}\delta_j^{\text{S}}(\text{tr})\right)\left(\frac{1}{J_{\text{C}}}\sum_{j=1}^{J} D_j^{\text{S}}\delta_j^{\text{S}}(\text{ib})\right)\right] = \mathbb{E}\left[\frac{1}{J_{\text{T}} J_{\text{C}}}\sum_{j=1}^{J}\sum_{j'=1}^{J} D_i^{\text{S}} D_{j'}^{\text{S}}\delta_j^{\text{S}}(\text{tr})\delta_{j'}^{\text{S}}(\text{ib})\right] \\
&= \frac{1}{J_{\text{T}} J_{\text{C}}}\sum_{j=1}^{J}\sum_{j'=1}^{J}\mathbb{E}\left[D_j^{\text{S}} D_{j'}^{\text{S}}\right]\delta_j^{\text{S}}(\text{tr})\delta_{j'}^{\text{S}}(\text{ib}) \\
&= -\frac{1}{J_{\text{T}} J_{\text{C}}}\sum_{j=1}^{J}\sum_{j'=1}^{J}\frac{J_{\text{C}} J_{\text{T}}}{J^2(J-1)}\delta_j^{\text{S}}(\text{tr})\delta_{j'}^{\text{S}}(\text{ib}) + \frac{1}{J_{\text{T}} J_{\text{C}}}\sum_{j=1}^{J}\left(\frac{J_{\text{C}} J_{\text{T}}}{J^2(J-1)} + \frac{J_{\text{C}} J_{\text{T}}}{J^2}\right)\delta_j^{\text{S}}(\text{tr})\delta_j^{\text{S}}(\text{ib}) \\
&= \frac{1}{J_{\text{T}} J_{\text{C}}}\sum_{j=1}^{J}\left(\frac{J_{\text{C}} J_{\text{T}}}{J^2(J-1)} + \frac{J_{\text{C}} J_{\text{T}}}{J^2}\right)\delta_j^{\text{S}}(\text{tr})\delta_j^{\text{S}}(\text{ib}) = \frac{1}{J}\left(\frac{1}{J-1}\sum_{j=1}^{J}\delta_j^{\text{S}}(\text{tr})\delta_j^{\text{S}}(\text{ib})\right).
\end{aligned}$$

Because

$$\begin{aligned}\xi^{\mathrm{S}}_{\mathrm{tr,ib}} &= \frac{1}{J}\sum_{j=1}^{J}\left(\delta^{\mathrm{S}}_{j}(\mathrm{tr}) - \delta^{\mathrm{S}}_{j}(\mathrm{ib})\right)^2 \\ &= \frac{1}{J}\sum_{j=1}^{J}\left(\delta^{\mathrm{S}}_{j}(\mathrm{tr})\right)^2 + \frac{1}{J}\sum_{j=1}^{J}\left(\delta^{\mathrm{S}}_{j}(\mathrm{ib})\right)^2 - \frac{2}{J}\sum_{j=1}^{J}\left(\delta^{\mathrm{S}}_{j}(\mathrm{tr})\delta^{\mathrm{S}}_{j}(\mathrm{ib})\right)^2 \\ &= \sigma^{\mathrm{S}}_{\mathrm{tr}} + \sigma^{\mathrm{S}}_{\mathrm{ib}} + 2(J-1)\,\mathrm{Cov}^{\mathrm{S}}_{\mathrm{tr,ib}},\end{aligned}$$

it follows that

$$\mathrm{Cov}^{\mathrm{S}}_{\mathrm{tr,ib}} = \frac{1}{2(J-1)}\left(\sigma^{\mathrm{S}}_{\mathrm{tr}} + \sigma^{\mathrm{S}}_{\mathrm{ib}} - \xi^{\mathrm{S}}_{\mathrm{tr,ib}}\right).$$

This finishes the proof of (A.4). Third, consider (A.5):

$$\mathrm{Cov}^{\mathrm{BS}}_{\mathrm{tr,ib}} = \mathbb{E}\left[\frac{1}{I_{\mathrm{T}}^2 J_{\mathrm{T}} J_{\mathrm{C}}}\sum_{i,i'=1}^{I}\sum_{j,j'=1}^{J} D^{\mathrm{B}}_{i} D^{\mathrm{S}}_{j} D^{\mathrm{B}}_{i'} D^{\mathrm{S}}_{j'} \delta^{\mathrm{BS}}_{ij}(\mathrm{tr})\delta^{\mathrm{BS}}_{i'j'}(\mathrm{ib})\right].$$

By independence of D^{B}_{i} and D^{S}_{j}, this is equal to

$$\mathrm{Cov}^{\mathrm{BS}}_{\mathrm{tr,ib}} = \frac{1}{I_{\mathrm{T}}^2 J_{\mathrm{C}} J_{\mathrm{T}}}\sum_{i,i'=1}^{I}\sum_{j,j'=1}^{J}\mathbb{E}\left[D^{\mathrm{B}}_{i} D^{\mathrm{B}}_{i'}\right]\mathbb{E}\left[D^{\mathrm{S}}_{i} D^{\mathrm{S}}_{j'}\right]\delta^{\mathrm{BS}}_{ij}(\mathrm{tr})\delta^{\mathrm{BS}}_{i'j'}(\mathrm{ib}).$$

Using the covariances and variances for D^{B}_{i} and $D^{\mathrm{B}}_{i'}$ and for D^{S}_{j} and $D^{\mathrm{S}}_{j'}$ this is equal to

$$\begin{aligned}\mathrm{Cov}^{\mathrm{BS}}_{\mathrm{tr,ib}} = &\frac{1}{I_{\mathrm{T}}^2 J_{\mathrm{C}} J_{\mathrm{T}}}\sum_{i=1}^{I}\sum_{i'=1}^{I}\sum_{j=1}^{J}\sum_{j'=1}^{J}\frac{I_{\mathrm{C}} I_{\mathrm{T}}}{I^2(I-1)}\frac{J_{\mathrm{C}} J_{\mathrm{T}}}{J^2(J-1)}\delta^{\mathrm{BS}}_{ij}(\mathrm{tr})\delta^{\mathrm{BS}}_{i'j'}(\mathrm{ib}) \\ &- \frac{1}{I_{\mathrm{T}}^2 J_{\mathrm{C}} J_{\mathrm{T}}}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{j'=1}^{J}\frac{I_{\mathrm{C}} I_{\mathrm{T}}}{I^2(I-1)}\frac{J_{\mathrm{C}} J_{\mathrm{T}}}{J(J-1)}\delta^{\mathrm{BS}}_{ij}(\mathrm{tr})\delta^{\mathrm{BS}}_{ij'}(\mathrm{ib}) \\ &- \frac{1}{I_{\mathrm{T}}^2 J_{\mathrm{C}} J_{\mathrm{T}}}\sum_{i=1}^{I}\sum_{i'=1}^{I}\sum_{j=1}^{J}\frac{I_{\mathrm{C}} I_{\mathrm{T}}}{I(I-1)}\frac{J_{\mathrm{C}} J_{\mathrm{T}}}{J^2(J-1)}\delta^{\mathrm{BS}}_{ij}(\mathrm{tr})\delta^{\mathrm{BS}}_{i'j}(\mathrm{ib}) \\ &+ \frac{1}{I_{\mathrm{T}}^2 J_{\mathrm{C}} J_{\mathrm{T}}}\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{I_{\mathrm{C}} I_{\mathrm{T}}}{I(I-1)}\frac{J_{\mathrm{C}} J_{\mathrm{T}}}{J(J-1)}\delta^{\mathrm{BS}}_{ij}(\mathrm{tr})\delta^{\mathrm{BS}}_{ij}(\mathrm{ib}).\end{aligned}$$

Because $\sum_i \sum_j \delta^{\mathrm{BS}}_{ij}(\gamma) = 0$, the first three terms are equal to zero, and so

$$\begin{aligned}\mathrm{Cov}^{\mathrm{BS}}_{\mathrm{tr,ib}} &= \frac{1}{I_{\mathrm{T}}^2 J_{\mathrm{C}} J_{\mathrm{T}}}\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{I_{\mathrm{C}} I_{\mathrm{T}}}{I(I-1)}\frac{J_{\mathrm{C}} J_{\mathrm{T}}}{J(J-1)}\delta^{\mathrm{BS}}_{ij}(\mathrm{tr})\delta^{\mathrm{BS}}_{ij}(\mathrm{ib}) \\ &= \frac{I_{\mathrm{C}}}{I_{\mathrm{T}} I J}\left(\frac{1}{(I-1)(J-1)}\sum_{i=1}^{I}\sum_{j=1}^{J}\delta^{\mathrm{BS}}_{ij}(\mathrm{tr})\delta^{\mathrm{BS}}_{ij}(\mathrm{ib})\right).\end{aligned}$$

Because

$$
\begin{aligned}
\xi^{\mathrm{BS}}_{\mathrm{tr},\mathrm{ib}} &= \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}\left(\delta^{\mathrm{BS}}_{ij}(\mathrm{tr})-\delta^{\mathrm{BS}}_{ij}(\mathrm{ib})\right)^2 \\
&= \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}\left(\delta^{\mathrm{BS}}_{ij}(\mathrm{tr})\right)^2 + \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}\left(\delta^{\mathrm{BS}}_{ij}(\mathrm{ib})\right)^2 \\
&- \frac{2}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}\delta^{\mathrm{BS}}_{ij}(\mathrm{tr})\delta^{\mathrm{BS}}_{ij}(\mathrm{ib}) \\
&= \sigma^{\mathrm{BS}}_{\mathrm{tr}} + \sigma^{\mathrm{BS}}_{\mathrm{ib}} - \frac{2I_{\mathrm{T}}(I-1)(J-1)}{I_{\mathrm{C}}}\mathrm{Cov}^{\mathrm{BS}}_{\mathrm{tr},\mathrm{ib}},
\end{aligned}
$$

it follows that

$$
\mathrm{Cov}^{\mathrm{BS}}_{\mathrm{tr},\mathrm{ib}} = \frac{I_{\mathrm{C}}}{2I_{\mathrm{T}}(I-1)(J-1)}\left(\sigma^{\mathrm{BS}}_{\mathrm{tr}} + \sigma^{\mathrm{BS}}_{\mathrm{ib}} - \xi^{\mathrm{BS}}_{\mathrm{tr},\mathrm{ib}}\right).
$$

This finishes the proof of (A.5). The proofs for the other pairwise comparisons follow the same pattern and are omitted. □

Theorem A.9 (Theorem 4.3 in the main paper)**.** *For a SMRD where Assumption 2.4 holds,*

$$
\mathrm{Cov}\left[\widehat{\overline{\overline{Y}}}_{\gamma}, \widehat{\overline{\overline{Y}}}_{\gamma'}\right] = \nu^{\mathrm{B}}_{\gamma,\gamma'}\zeta^{\mathrm{B}}_{\gamma,\gamma'} + \nu^{\mathrm{S}}_{\gamma,\gamma'}\zeta^{\mathrm{S}}_{\gamma,\gamma'} + 2\nu^{\mathrm{B}}_{\gamma,\gamma'}\nu^{\mathrm{S}}_{\gamma,\gamma'}\zeta^{\mathrm{BS}}_{\gamma,\gamma'}, \tag{A.6}
$$

where for $x \in \{\mathrm{B},\mathrm{S},\mathrm{BS}\}$*,* $\zeta^{x}_{\gamma,\gamma'} := \sigma^{x}_{\gamma} + \sigma^{x}_{\gamma'} - \xi^{x}_{\gamma,\gamma'}$ *and*

$$
\nu^{\mathrm{B}}_{\gamma,\gamma'} := \begin{cases} \alpha^{\mathrm{B}}_{\gamma}/2 \text{ if } \gamma=\gamma', \text{ or } (\gamma,\gamma') \in \{(\mathrm{cc},\mathrm{is}),(\mathrm{is},\mathrm{cc}),(\mathrm{ib},\mathrm{tr}),(\mathrm{tr},\mathrm{ib})\} \\ -1/(2(I-1)) \text{ otherwise,} \end{cases}
$$

and

$$
\nu^{\mathrm{S}}_{\gamma,\gamma'} := \begin{cases} \alpha^{\mathrm{S}}_{\gamma}/2 \text{ if } \gamma=\gamma' \text{ or } (\gamma,\gamma') \in \{(\mathrm{cc},\mathrm{ib}),(\mathrm{ib},\mathrm{cc}),(\mathrm{is},\mathrm{tr}),(\mathrm{tr},\mathrm{is})\} \\ -1/(2(J-1)) \text{ otherwise.} \end{cases}
$$

Proof of Theorem 4.3. Lemma A.7 (for $\gamma = \gamma'$) and Lemma A.8 (for $\gamma \neq \gamma'$) prove this result. We spell these cases out and verify that the expressions derived in lemmas A.7 and A.8 match with the compact representation provided in eq. (A.6).

- if $\gamma = \gamma'$, using Lemma A.7 and the definitions of $\alpha^{\mathrm{B}}_{\gamma}$ and $\alpha^{\mathrm{S}}_{\gamma}$,

$$
\begin{aligned}
\mathrm{Cov}\left[\widehat{\overline{\overline{Y}}}_{\gamma}, \widehat{\overline{\overline{Y}}}_{\gamma}\right] = \mathrm{Var}\left(\widehat{\overline{\overline{Y}}}_{\gamma}\right) &= \frac{I-I_{\gamma}}{I_{\gamma}}\frac{1}{I-1}\sigma^{\mathrm{B}}_{\gamma} + \frac{J-J_{\gamma}}{J_{\gamma}}\frac{1}{J-1}\sigma^{\mathrm{S}}_{\gamma} + \frac{I-I_{\gamma}}{I_{\gamma}}\frac{1}{I-1}\frac{J-J_{\gamma}}{J_{\gamma}}\frac{1}{J-1}\sigma^{\mathrm{BS}}_{\gamma} \\
&= \alpha^{\mathrm{B}}_{\gamma}\sigma^{\mathrm{B}}_{\gamma} + \alpha^{\mathrm{S}}_{\gamma}\sigma^{\mathrm{S}}_{\gamma} + \alpha^{\mathrm{B}}_{\gamma}\alpha^{\mathrm{S}}_{\gamma}\sigma^{\mathrm{BS}}_{\gamma}.
\end{aligned}
$$

  We verify that eq. (A.6) is correct by spelling out $\nu^{\mathrm{B}}_{\gamma,\gamma}$, $\nu^{\mathrm{S}}_{\gamma,\gamma}$, $\zeta^{\mathrm{B}}_{\gamma,\gamma}$, $\zeta^{\mathrm{S}}_{\gamma,\gamma}$ — and check that we get the same result as above:

$$\begin{aligned}\mathrm{Cov}\left[\widehat{\widehat{\overline{Y}}}_{\gamma}, \widehat{\widehat{\overline{Y}}}_{\gamma'}\right] &= \nu^{\mathrm{B}}_{\gamma,\gamma'}\zeta^{\mathrm{B}}_{\gamma,\gamma'} + \nu^{\mathrm{S}}_{\gamma,\gamma'}\zeta^{\mathrm{S}}_{\gamma,\gamma'} + 2\nu^{\mathrm{B}}_{\gamma,\gamma'}\nu^{\mathrm{S}}_{\gamma,\gamma'}\zeta^{\mathrm{BS}}_{\gamma,\gamma'} \\ &= \frac{\alpha^{\mathrm{B}}_{\gamma}}{2}(2\sigma^{\mathrm{B}}_{\gamma}) + \frac{\alpha^{\mathrm{S}}_{\gamma}}{2}(2\sigma^{\mathrm{S}}_{\gamma}) + 2\frac{\alpha^{\mathrm{B}}_{\gamma}}{2}\frac{\alpha^{\mathrm{S}}_{\gamma}}{2}(2\sigma^{\mathrm{BS}}_{\gamma}) \\ &= \alpha^{\mathrm{B}}_{\gamma}\sigma^{\mathrm{B}}_{\gamma} + \alpha^{\mathrm{S}}_{\gamma}\sigma^{\mathrm{S}}_{\gamma} + \alpha^{\mathrm{B}}_{\gamma}\alpha^{\mathrm{S}}_{\gamma}\sigma^{\mathrm{BS}}_{\gamma}.\end{aligned}$$

- if $\gamma \neq \gamma'$, use Lemma A.8, and consider any of the treatment pairs (e.g., (ib, is)):

$$\begin{aligned}\mathrm{Cov}\left[\widehat{\widehat{\overline{Y}}}_{\mathrm{ib}}, \widehat{\widehat{\overline{Y}}}_{\mathrm{is}}\right] &= -\frac{1}{2(I-1)}\left(\sigma^{\mathrm{B}}_{\mathrm{ib}} + \sigma^{\mathrm{B}}_{\mathrm{is}} - \xi^{\mathrm{B}}_{\mathrm{ib,is}}\right) - \frac{1}{2(J-1)}\left(\sigma^{\mathrm{S}}_{\mathrm{ib}} + \sigma^{\mathrm{S}}_{\mathrm{is}} - \xi^{\mathrm{S}}_{\mathrm{ib,is}}\right) \\ &+ \frac{1}{2(I-1)(J-1)}\left(\sigma^{\mathrm{BS}}_{\mathrm{ib}} + \sigma^{\mathrm{BS}}_{\mathrm{is}} - \xi^{\mathrm{BS}}_{\mathrm{ib,is}}\right) \\ &= \nu^{\mathrm{S}}_{\gamma,\gamma'}\zeta^{\mathrm{B}}_{\gamma,\gamma'} + \nu^{\mathrm{S}}_{\gamma,\gamma'}\zeta^{\mathrm{S}}_{\gamma,\gamma'} + 2\nu^{\mathrm{B}}_{\gamma}\nu^{\mathrm{S}}_{\gamma'}\zeta^{\mathrm{BS}}_{\gamma,\gamma'},\end{aligned}$$

which matches the compact representation.

□

To present our results on estimates of the variance, we first review a classic result for variances of a simple two-arms experiment, when a single population is present.

Lemma A.10. *Let y_i, $i = 1, \ldots, I$ be a population of I units with (non-random) potential outcomes $y_i(\mathrm{cc})$ (if unit i is in the control group) and $y_i(\mathrm{tr})$ (if unit i is in the treatment group). Let the treatment group be identified by the index set $\mathcal{I}_{\mathrm{tr}} = \{i_1, \ldots, i_{I_{\mathrm{tr}}}\} \subset \{1, \ldots, I\}$, of size $|\mathcal{I}_{\mathrm{tr}}| = I_{\mathrm{tr}}$, with $2 \leq I_{\mathrm{tr}} \leq I - 2$. Let $\mathcal{I}_{\mathrm{cc}} = \{1, \ldots, I\} \setminus \mathcal{I}_{\mathrm{tr}}$ be the index set of the $I_{\mathrm{cc}} := I - I_{\mathrm{tr}}$ units assigned to the control group. For $\gamma \in \{\mathrm{cc}, \mathrm{tr}\}$, let*

$$\bar{y}_{\gamma} = \frac{1}{I}\sum_{i=1}^{I} y_i(\gamma) \quad \text{and} \quad s_{\gamma} = \frac{1}{I}\sum_{i=1}^{I}\left(y_i(\gamma) - \bar{y}_{\gamma}\right)^2.$$

be the mean and variance of the potential outcomes in the population. Define the corresponding plug-in estimates for these to be

$$\widehat{\overline{Y}}_{\gamma} = \frac{1}{I_{\gamma}}\sum_{i \in \mathcal{I}_{\gamma}} y_i(\gamma), \quad \text{and} \quad \hat{S}_{\gamma} = \frac{1}{I_{\gamma}}\sum_{i \in \mathcal{I}_{\gamma}}\left(y_i(\gamma) - \widehat{\overline{Y}}_{\gamma}\right)^2.$$

Then it holds

$$\mathbb{E}\left[\widehat{\overline{Y}}_{\gamma}\right] = \bar{y}_{\gamma}, \quad \text{and} \quad \mathrm{Var}\left(\widehat{\overline{Y}}_{\gamma}\right) = \frac{I - I_{\gamma}}{I_{\gamma}}\frac{1}{I-1}s_{\gamma}, \quad \text{and} \quad \mathbb{E}\left(\hat{S}_{\gamma}\right) = \frac{I_{\gamma}-1}{I_{\gamma}}\frac{I}{I-1}s_{\gamma}.$$

I.e., $\widehat{\overline{Y}}_{\gamma}$ is an unbiased estimate of the population mean $\bar{y}_{\gamma}$. We can obtain an unbiased estimate of the variance of this estimator by reweighing $\hat{S}_{\gamma}$:

$$\widehat{\mathrm{Var}}\left(\widehat{\overline{Y}}_{\gamma}\right) := \frac{I - I_{\gamma}}{I_{\gamma}-1}\frac{1}{I}\hat{S}_{\gamma} \quad \text{satisfies} \quad \mathbb{E}\left[\widehat{\mathrm{Var}}\left(\widehat{\overline{Y}}_{\gamma}\right)\right] = \mathrm{Var}\left(\widehat{\overline{Y}}_{\gamma}\right). \tag{A.7}$$

Proof of Lemma A.10. See *e.g.* Cochran [1977, Theorems 2.1, 2.2, 2.4]. □

A.4 Variance estimation in SMRDs: proofs

Here we provide lower and upper bounds on the variance of causal effects in SMRDs (theorems 4.4 and 4.5). For a SMRD in which local interference holds, given an assignment matrix $\mathbf{w}$, denote by $\mathcal{I}_\gamma \subseteq \{1, \dots, I\}$ the subset of buyers' indices for which there exists at least one seller j such that unit (i, j) has type γ: $\mathcal{I}_\gamma := \{i \in \{1, \dots, I\} \ : \ \gamma_{ij} = \gamma \text{ for some } j\}$. Symmetrically, let $\mathcal{J}_\gamma \subseteq \{1, \dots, J\}$ the subset of sellers' indices for which there exists at least one buyer i such that unit (i, j) has type γ. Consistent with appendix A.1, $I_\gamma = |\mathcal{I}_\gamma|$ and $J_\gamma = |\mathcal{J}_\gamma|$ denote the sizes of these index sets. Exactly $I_\gamma J_\gamma$ units are assigned type γ. Define now, the (nonrandom) row and column *partial* mean of the matrix of potential outcomes: for a given row i, the average over a fixed index set of columns $\mathcal{J}_\gamma \subseteq [J]$ — symmetrically, for a given column j, the average over a fixed set of rows $\mathcal{I}_\gamma \subseteq [I]$:

$$\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_\gamma}(\gamma) = \frac{1}{J_\gamma} \sum_{j \in \mathcal{J}_\gamma} y_{i,j}(\gamma) \quad \text{and} \quad \overline{y}^{\mathrm{S}}_{\mathcal{I}_\gamma,j}(\gamma) = \frac{1}{I_\gamma} \sum_{i \in \mathcal{I}_\gamma} y_{i,j}(\gamma).$$

For a given SMRD, with (random) assignment matrix $\mathbf{W}$ and characterized by (random) index sets $\mathcal{I}_\gamma, \mathcal{J}_\gamma$ for each $\gamma \in \{\mathrm{cc}, \mathrm{ib}, \mathrm{is}, \mathrm{tr}\}$, $i \in \mathcal{I}_\gamma, j \in \mathcal{J}_\gamma$, define the random average over the columns selected by the set $\mathcal{J}_\gamma$ (or the rows selected by $\mathcal{I}_\gamma$):

$$\widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma) := \frac{1}{J_\gamma} \sum_{j \in \mathcal{J}_\gamma} y_{i,j}(\gamma), \quad \text{and} \quad \widehat{\overline{Y}}^{\mathrm{S}}_j(\gamma) := \frac{1}{I_\gamma} \sum_{i \in \mathcal{I}_\gamma} y_{i,j}(\gamma).$$

Remark The quantities $\widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma)$ and $\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_\gamma}(\gamma)$ are *both* averages over J_γ elements of the i-th row of the matrix of potential outcomes $Y(\gamma)$. However, $\widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma)$ is random: it is an estimator resulting from the random selection of J_γ distinct columns, whereas $\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_\gamma}(\gamma)$ is a fixed population value, obtained by averaging over the fixed J_γ distinct indices $\{j_1, \dots, j_{J_\gamma}\} = \mathcal{J}_\gamma$. Define the sample "plug-in" counterparts of the population quantities $\sigma^{\mathrm{B}}_\gamma, \sigma^{\mathrm{S}}_\gamma$ and $\sigma^{\mathrm{BS}}_\gamma$:

$$\hat{\Sigma}^{\mathrm{B}}_\gamma = \frac{1}{I_\gamma} \sum_{i \in \mathcal{I}_\gamma} \left(\widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma) - \widehat{\overline{\overline{Y}}}_\gamma \right)^2, \quad \hat{\Sigma}^{\mathrm{S}}_\gamma = \frac{1}{J_\gamma} \sum_{j \in \mathcal{J}_\gamma} \left(\widehat{\overline{Y}}^{\mathrm{S}}_j(\gamma) - \widehat{\overline{\overline{Y}}}_\gamma \right)^2,$$

and

$$\hat{\Sigma}^{\mathrm{BS}}_\gamma := \frac{1}{I_\gamma J_\gamma} \sum_{i \in \mathcal{I}_\gamma} \sum_{j \in \mathcal{J}_\gamma} \left(y_{i,j}(\gamma) - \widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma) - \widehat{\overline{Y}}^{\mathrm{S}}_j(\gamma) + \widehat{\overline{\overline{Y}}}_\gamma \right)^2.$$

$\hat{\Sigma}^{\mathrm{B}}_\gamma, \hat{\Sigma}^{\mathrm{S}}_\gamma, \hat{\Sigma}^{\mathrm{BS}}_\gamma$ are stochastic and depend on the (random) assignment $\mathbf{W}$ through the index sets $\mathcal{I}_\gamma, \mathcal{J}_\gamma$. Last, define the variances of partial averages over subsets $\mathcal{J}_\gamma$ and $\mathcal{I}_\gamma$

$$\eta^{\mathrm{B}}_\gamma := \frac{\sum_{\mathcal{J}_\gamma} \sum_i \left\{ \overline{y}^{\mathrm{B}}_{i,\mathcal{J}_\gamma}(\gamma) - \overline{y}^{\mathrm{B}}_i(\gamma) \right\}^2}{I \binom{J}{J_\gamma}}, \quad \eta^{\mathrm{S}}_\gamma := \frac{\sum_{\mathcal{I}_\gamma} \sum_j \left\{ \overline{y}^{\mathrm{S}}_{\mathcal{I}_\gamma,j}(\gamma) - \overline{y}^{\mathrm{S}}_j(\gamma) \right\}^2}{J \binom{I}{I_\gamma}}. \tag{A.8}$$

Notice that the sums in eq. (A.8) are over all subsets of J_γ disjoint indices in $[J]$ (η^{B}_γ) or I_γ disjoint indices in $[I]$ (η^{S}_γ). In lemmas A.12, A.13 and A.15, we analyze the expectation of each term $\hat{\Sigma}^{\mathrm{B}}_\gamma, \hat{\Sigma}^{\mathrm{S}}_\gamma, \hat{\Sigma}^{\mathrm{BS}}_\gamma$ separately. First, we state a useful result in lemma A.11.

Lemma A.11. *Let*

$$\chi^{2,\mathrm{B}}_\gamma := \mathbb{E}\left[\sum_{i\in\mathcal{I}_\gamma}\left(\widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma) - \overline{\overline{y}}_\gamma\right)^2\right]. \tag{A.9}$$

It holds $\chi^{2,\mathrm{B}}_\gamma = I_\gamma\left(\sigma^{\mathrm{B}}_\gamma + \eta^{\mathrm{B}}_\gamma\right)$, *where* η^{B}_γ *was defined in eq.* (A.8).

Proof of lemma A.11. Consider $\chi^{2,\mathrm{B}}_\gamma$ as defined in eq. (A.9), where the expectation is taken with respect to the random assignment matrices $\mathbf{W}$. Under (simple) double randomization, every assignment matrix $\mathbf{W}$ supported on $\mathbb{W}$ is equivalently characterized by the index sets $\mathcal{I}_\gamma, \mathcal{J}_\gamma$, for $\gamma \in \{\mathrm{cc}, \mathrm{ib}, \mathrm{is}, \mathrm{tr}\}$. That is, to each $\mathbf{W}$, there is one and only one collection of index sets $\mathcal{I}_\gamma, \mathcal{J}_\gamma$ for $\gamma \in \{\mathrm{cc}, \mathrm{ib}, \mathrm{is}, \mathrm{tr}\}$, and viceversa. Notice that there are exactly $\binom{I}{I_\gamma}\binom{J}{J_\gamma}$ such assignments. Each assignment can be determined by forming index set $\mathcal{I}_\gamma$ by selecting at random I_γ rows and index set $\mathcal{J}_\gamma$ by selecting at random J_γ columns. Every row $i \in \{1, \ldots, I\}$ appears in exactly $\binom{I-1}{I_\gamma-1}$ index sets $\mathcal{I}_\gamma$. Hence,

$$\chi^{2,\mathrm{B}}_\gamma = \frac{\binom{I-1}{I_\gamma-1}}{\binom{I}{I_\gamma}\binom{J}{J_\gamma}}\sum_{i=1}^{I}\sum_{\mathcal{J}_\gamma}\left\{\left(\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_\gamma}(\gamma) - \overline{\overline{y}}_\gamma\right)^2\right\} = \frac{I_\gamma}{I\binom{J}{J_\gamma}}\sum_{i=1}^{I}\sum_{\mathcal{J}_\gamma}\left\{\left(\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_\gamma}(\gamma) - \overline{\overline{y}}_\gamma\right)^2\right\}, \tag{A.10}$$

where the second sum is over all $\binom{J}{J_\gamma}$ subsets $\mathcal{J}_\gamma$ of J_γ distinct indices in $\{1, \ldots, J\}$.

We further decompose $\chi^{2,\mathrm{B}}_\gamma$: fix a row i and disjoint indices $\mathcal{J}_\gamma = \{j_1, \ldots, j_{J_\gamma}\} \subseteq [J]$:

$$\begin{aligned}
\{\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_\gamma}(\gamma) - \overline{\overline{y}}_\gamma\}^2 &= \left\{\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_\gamma}(\gamma) - \overline{y}^{\mathrm{B}}_i(\gamma) + \overline{y}^{\mathrm{B}}_i(\gamma) - \overline{\overline{y}}_\gamma\right\}^2 \\
&= \left\{\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_\gamma}(\gamma) - \overline{y}^{\mathrm{B}}_i(\gamma)\right\}^2 + \left\{\overline{y}^{\mathrm{B}}_i(\gamma) - \overline{\overline{y}}_\gamma\right\}^2 + 2\left\{\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_\gamma}(\gamma) - \overline{y}^{\mathrm{B}}_i(\gamma)\right\}\left\{\overline{y}^{\mathrm{B}}_i(\gamma) - \overline{\overline{y}}_\gamma\right\}.
\end{aligned}$$

Summing over all choices $\mathcal{J}_\gamma$ of J_γ disjoint indices in the set $\{1, \ldots, J\}$,

$$\sum_{\mathcal{J}_\gamma}\left\{\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_\gamma}(\gamma) - \overline{\overline{y}}_\gamma\right\}^2 = \sum_{\mathcal{J}_\gamma}\left\{\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_\gamma}(\gamma) - \overline{y}^{\mathrm{B}}_i\right\}^2 + \binom{J}{J_\gamma}\left\{\overline{y}^{\mathrm{B}}_i(\gamma) - \overline{\overline{y}}_\gamma\right\}^2,$$

using $\sum_{\mathcal{J}_\gamma}\left\{\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_\gamma}(\gamma) - \overline{y}^{\mathrm{B}}_i(\gamma)\right\} = 0$, since $\overline{y}^{\mathrm{B}}_i(\gamma) = \frac{\sum_{\mathcal{J}_\gamma}\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_\gamma}}{\binom{J}{J_\gamma}}$. Summing over buyers:

$$\begin{aligned}
\sum_{i=1}^{I}\sum_{\mathcal{J}_\gamma}\left\{\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_\gamma}(\gamma) - \overline{\overline{y}}_\gamma\right\}^2 &= \sum_{\mathcal{J}_\gamma}\sum_{i=1}^{I}\left\{\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_\gamma}(\gamma) - \overline{y}^{\mathrm{B}}_i(\gamma)\right\}^2 + \binom{J}{J_\gamma}\sum_i\left\{\overline{y}^{\mathrm{B}}_i(\gamma) - \overline{\overline{y}}_\gamma\right\}^2 \\
&= I\binom{J}{J_\gamma}\left(\eta^{\mathrm{B}}_\gamma + \sigma^{\mathrm{B}}_\gamma\right),
\end{aligned}$$

where $\eta^{\mathrm{B}}_\gamma := \frac{1}{I}\binom{J}{J_\gamma}^{-1}\sum_{\mathcal{J}_\gamma}\sum_i\left\{\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_\gamma}(\gamma)-\overline{y}^{\mathrm{B}}_i(\gamma)\right\}^2$. Hence, plugging this in eq. (A.10),

$$\chi^{2,\mathrm{B}}_\gamma=\sum_{i=1}^{I}\sum_{\mathcal{J}_\gamma}\frac{\left\{\left(\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_\gamma}(\gamma)-\overline{\overline{y}}_\gamma\right)^2\right\}}{\frac{I}{I_\gamma}\binom{J}{J_\gamma}}=\frac{\left[I\binom{J}{J_\gamma}\left(\eta^{\mathrm{B}}_\gamma+\sigma^{\mathrm{B}}_\gamma\right)\right]}{\frac{I}{I_\gamma}\binom{J}{J_\gamma}}=I_\gamma\left(\sigma^{\mathrm{B}}_\gamma+\eta^{\mathrm{B}}_\gamma\right). \tag{A.11}$$

□

Lemma A.12. *It holds*

$$\mathbb{E}\left[\hat{\Sigma}^{\mathrm{B}}_\gamma\right]=\sigma^{\mathrm{B}}_\gamma-\operatorname{Var}\left(\widehat{\overline{\overline{Y}}}_\gamma\right)+\eta^{\mathrm{B}}_\gamma. \tag{A.12}$$

Proof of lemma A.12.

$$\begin{aligned}\mathbb{E}\left[\hat{\Sigma}^{\mathrm{B}}_\gamma\right]&=\mathbb{E}\left[\frac{1}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma}\left(\widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma)-\widehat{\overline{\overline{Y}}}_\gamma\right)^2\right]=\frac{1}{I_\gamma}\mathbb{E}\left[\sum_{i\in\mathcal{I}_\gamma}\left\{\left[\widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma)-\overline{\overline{y}}_\gamma\right]-\left[\widehat{\overline{\overline{Y}}}_\gamma-\overline{\overline{y}}_\gamma\right]\right\}^2\right]\\&=\frac{1}{I_\gamma}\mathbb{E}\left[\sum_{i\in\mathcal{I}_\gamma}\left[\widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma)-\overline{\overline{y}}_\gamma\right]^2+\sum_{i\in\mathcal{I}_\gamma}\left[\widehat{\overline{\overline{Y}}}_\gamma-\overline{\overline{y}}_\gamma\right]^2-2\sum_{i\in\mathcal{I}_\gamma}\left(\widehat{\overline{\overline{Y}}}_\gamma-\overline{\overline{y}}_\gamma\right)\left(\widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma)-\overline{\overline{y}}_\gamma\right)\right].\end{aligned}$$

$$\begin{aligned}\mathbb{E}\left[\hat{\Sigma}^{\mathrm{B}}_\gamma\right]&=\frac{1}{I_\gamma}\left\{\mathbb{E}\left[\sum_{i\in\mathcal{I}_\gamma}\left\{\widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma)-\overline{\overline{y}}_\gamma\right\}^2+I_\gamma\left\{\widehat{\overline{\overline{Y}}}_\gamma-\overline{\overline{y}}_\gamma\right\}^2-2\left\{\widehat{\overline{\overline{Y}}}_\gamma-\overline{\overline{y}}_\gamma\right\}I_\gamma\left\{\widehat{\overline{\overline{Y}}}_\gamma-\overline{\overline{y}}_\gamma\right\}\right]\right\}\\&=\frac{1}{I_\gamma}\mathbb{E}\left[\sum_{i\in\mathcal{I}_\gamma}\left\{\widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma)-\overline{\overline{y}}_\gamma\right\}^2\right]-\mathbb{E}\left[\left\{\widehat{\overline{\overline{Y}}}_\gamma-\overline{\overline{y}}_\gamma\right\}^2\right].\end{aligned}$$

Hence, we write

$$\mathbb{E}\left[\hat{\Sigma}^{\mathrm{B}}_\gamma\right]=\frac{1}{I_\gamma}\chi^{2,\mathrm{B}}_\gamma-\operatorname{Var}\left(\widehat{\overline{\overline{Y}}}_\gamma\right), \tag{A.13}$$

where we used the definition of $\chi^{2,\mathrm{B}}_\gamma$ given in eq. (A.9). Now plugging in eq. (A.11) in eq. (A.13) it follows

$$\mathbb{E}\left[\hat{\Sigma}^{\mathrm{B}}_\gamma\right]=\frac{1}{I_\gamma}\chi^{2,\mathrm{B}}_\gamma-\operatorname{Var}\left(\widehat{\overline{\overline{Y}}}_\gamma\right)=\sigma^{\mathrm{B}}_\gamma-\operatorname{Var}\left(\widehat{\overline{\overline{Y}}}_\gamma\right)+\eta^{\mathrm{B}}_\gamma.$$

□

Lemma A.13. *Recall η^{S}_γ defined in eq. (A.8). It holds*

$$\mathbb{E}\left[\hat{\Sigma}^{\mathrm{S}}_\gamma\right]=\sigma^{\mathrm{S}}_\gamma-\operatorname{Var}\left(\widehat{\overline{\overline{Y}}}_\gamma\right)+\eta^{\mathrm{S}}_\gamma, \tag{A.14}$$

Proof of lemma A.13. The proof is identical to lemma A.12, where we let $\chi^{2,\mathrm{S}}_\gamma$ be the column counterpart to eq. (A.11), $\chi_\gamma^{2,\mathrm{S}} := \mathbb{E}\left[\sum_{j\in\mathcal{J}_\gamma}\left(\widehat{\overline{Y}}{}^{\mathrm{S}}_j(\gamma)-\overline{\overline{y}}_\gamma\right)^2\right]$, where, by the same argument of lemma A.11, it holds

$$\chi_\gamma^{2,\mathrm{S}} = J_\gamma\sigma_\gamma^{\mathrm{S}} + \frac{J_\gamma}{J}\binom{I}{I_\gamma}^{-1}\sum_{\mathcal{I}_\gamma}\sum_{j=1}^{J}(\overline{y}^{\mathrm{S}}_{\mathcal{I}_\gamma,j}-\widehat{\overline{Y}}{}^{\mathrm{S}}_j)^2 = J_\gamma\left(\sigma_\gamma^{\mathrm{S}}+\eta_\gamma^{\mathrm{S}}\right), \tag{A.15}$$

in which we sum over all $\binom{I}{I_\gamma}$ index sets $\mathcal{I}_\gamma$ of I_γ disjoint indices in $\{1,\ldots,I\}$. □

We now characterize $\hat{\Sigma}_\gamma^{\mathrm{BS}}$. We first state a useful decomposition for matrices.

Lemma A.14. *Let $\boldsymbol{x}\in\mathbb{R}^{I\times J}$ be a matrix, and $\bar{\bar{x}} := (IJ)^{-1}\sum_{i,j}x_{i,j}$ be the grand mean of the matrix, where averaging is uniform across entries. Let $\bar{x}_i^{\mathrm{B}} := J^{-1}\sum_j x_{i,j}$ and $\bar{x}_j^{\mathrm{S}} := I^{-1}\sum_i x_{i,j}$ be the average of the i-th row and of the j-th column respectively. It holds*

$$\sum_{i,j}(x_{i,j}-\bar{\bar{x}})^2 = J\sum_i\left(\bar{x}_i^{\mathrm{B}}-\bar{\bar{x}}\right)^2 + J\sum_j\left(\bar{x}_j^{\mathrm{S}}-\bar{\bar{x}}\right)^2 + \sum_{i,j}\left(x_{i,j}-\bar{x}_i^{\mathrm{B}}-\bar{x}_j^{\mathrm{S}}+\bar{\bar{x}}\right)^2.$$

Proof of lemma A.14.

$$\begin{aligned}\sum_{i,j}(x_{i,j}-\bar{\bar{x}})^2 &= \sum_{i,j}\left(x_{i,j}\pm\bar{x}_i^{\mathrm{B}}\pm\bar{\bar{x}}\pm\bar{x}_j^{\mathrm{S}}\pm\bar{\bar{x}}+\bar{\bar{x}}\right)^2\\ &= \sum_{i,j}\left\{\left(\bar{x}_i^{\mathrm{B}}-\bar{\bar{x}}\right)+\left(\bar{x}_s^{\mathrm{S}}-\bar{\bar{x}}\right)+\left(x_{i,j}-\bar{x}_i^{\mathrm{B}}-\bar{x}_j^{\mathrm{S}}+\bar{\bar{x}}\right)\right\}^2\\ &= \sum_{i,j}\left(\bar{x}_i^{\mathrm{B}}-\bar{\bar{x}}\right)^2+\sum_{i,j}\left(\bar{x}_j^{\mathrm{S}}-\bar{\bar{x}}\right)^2+\sum_{i,j}\left(x_{i,j}-\bar{x}_i^{\mathrm{B}}-\bar{x}_j^{\mathrm{S}}+\bar{\bar{x}}\right)^2,\end{aligned}$$

where we have noted that all the cross terms in the square cancel since

$$\sum_{i,j}\left(\bar{x}_i^{\mathrm{B}}-\bar{\bar{x}}\right)=0,\quad \sum_{i,j}\left(\bar{x}_j^{\mathrm{S}}-\bar{\bar{x}}\right)=0,\quad \sum_{i,j}\left(x_{i,j}-\bar{x}_i^{\mathrm{B}}-\bar{x}_j^{\mathrm{S}}+\bar{\bar{x}}\right)=0.$$

Hence,

$$\sum_{i,j}(x_{i,j}-\bar{\bar{x}})^2 = J\sum_i\left(\bar{x}_i^{\mathrm{B}}-\bar{\bar{x}}\right)^2 + I\sum_j\left(\bar{x}_j^{\mathrm{S}}-\bar{\bar{x}}\right)^2 + \sum_{i,j}\left(x_{i,j}-\bar{x}_i^{\mathrm{B}}-\bar{x}_j^{\mathrm{S}}+\bar{\bar{x}}\right)^2.$$

□

For our matrix of potential outcomes $\boldsymbol{y}(\gamma)$, direct application of lemma A.14 gives us

$$\begin{aligned}\frac{1}{IJ}\sum_{i,j}\left[y_{i,j}(\gamma)-\overline{\overline{y}}_\gamma\right]^2 &= \frac{1}{I}\sum_i\left\{\overline{y}_i^{\mathrm{B}}(\gamma)-\overline{\overline{y}}_\gamma\right\}^2+\frac{1}{J}\sum_j\left\{\overline{y}_j^{\mathrm{S}}(\gamma)-\overline{\overline{y}}_\gamma\right\}^2\\ &\quad+\frac{1}{IJ}\sum_{i,j}\left\{y_{i,j}(\gamma)-\overline{y}_i^{\mathrm{B}}(\gamma)-\overline{y}_j^{\mathrm{S}}(\gamma)+\overline{\overline{y}}_\gamma\right\}^2=\sigma_\gamma^{\mathrm{B}}+\sigma_\gamma^{\mathrm{S}}+\sigma_\gamma^{\mathrm{BS}}.\end{aligned} \tag{A.16}$$

We now analyze the expectation of the crossed term $\hat{\Sigma}_\gamma^{\mathrm{BS}}$.

Lemma A.15. *It holds*

$$\mathbb{E}\left[\hat{\Sigma}^{\mathrm{BS}}_{\gamma}\right] = \sigma^{\mathrm{BS}}_{\gamma} + \mathrm{Var}\left(\widehat{\overline{\overline{Y}}}_{\gamma}\right) - \eta^{\mathrm{B}}_{\gamma} - \eta^{\mathrm{S}}_{\gamma}.$$

Proof of lemma A.15.

$$\hat{\Sigma}^{\mathrm{BS}}_{\gamma} = \frac{1}{I_\gamma J_\gamma}\sum_{i\in\mathcal{I}_\gamma}\sum_{j\in\mathcal{J}_\gamma}\left\{y_{i,j}(\gamma) - \widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma) - \widehat{\overline{Y}}^{\mathrm{S}}_j(\gamma) + \widehat{\overline{\overline{Y}}}_{\gamma}\right\}^2.$$

Expanding the square,

$$\begin{aligned}
\hat{\Sigma}^{\mathrm{BS}}_{\gamma} &= \frac{1}{I_\gamma J_\gamma}\sum_{i\in\mathcal{I}_\gamma}\sum_{j\in\mathcal{J}_\gamma}\left\{\left(y_{i,j}(\gamma) - \overline{\overline{y}}_\gamma\right) - \left(\widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma) - \overline{\overline{y}}_\gamma\right) - \left(\widehat{\overline{Y}}^{\mathrm{S}}_j(\gamma) - \overline{\overline{y}}_\gamma\right) + \left(\widehat{\overline{\overline{Y}}}_{\gamma} - \overline{\overline{y}}_\gamma\right)\right\}^2 \\
&= \frac{1}{I_\gamma J_\gamma}\sum_{i\in\mathcal{I}_\gamma}\sum_{j\in\mathcal{J}_\gamma}\left(y_{i,j}(\gamma) - \overline{\overline{y}}_\gamma\right)^2 + \frac{1}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma}\left(\widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma) - \overline{\overline{y}}_\gamma\right)^2 \\
&+ \frac{1}{J_\gamma}\sum_{j\in\mathcal{J}_\gamma}\left(\widehat{\overline{Y}}^{\mathrm{S}}_j(\gamma) - \overline{\overline{y}}_\gamma\right)^2 + \left(\widehat{\overline{\overline{Y}}}_{\gamma} - \overline{\overline{y}}_\gamma\right)^2 \\
&- \frac{2}{I_\gamma J_\gamma}\sum_{i\in\mathcal{I}_\gamma}\left(\widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma) - \overline{\overline{y}}_\gamma\right)\sum_{j\in\mathcal{J}_\gamma}\left(y_{i,j}(\gamma) - \overline{\overline{y}}_\gamma\right) - \frac{2}{I_\gamma J_\gamma}\sum_{j\in\mathcal{J}_\gamma}\left(\widehat{\overline{Y}}^{\mathrm{S}}_j(\gamma) - \overline{\overline{y}}_\gamma\right)\sum_{i\in\mathcal{I}_\gamma}\left(y_{i,j}(\gamma) - \overline{\overline{y}}_\gamma\right) \\
&+ \frac{2}{I_\gamma J_\gamma}\left(\widehat{\overline{\overline{Y}}}_{\gamma} - \overline{\overline{y}}_\gamma\right)\sum_{i\in\mathcal{I}_\gamma}\sum_{j\in\mathcal{J}_\gamma}\left(y_{i,j}(\gamma) - \overline{\overline{y}}_\gamma\right) + \frac{2}{I_\gamma J_\gamma}\sum_{i\in\mathcal{I}_\gamma}\left(\widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma) - \overline{\overline{y}}_\gamma\right)\sum_{j\in\mathcal{J}_\gamma}\left(\widehat{\overline{Y}}^{\mathrm{S}}_j(\gamma) - \overline{\overline{y}}_\gamma\right) \\
&- \frac{2}{I_\gamma J_\gamma}\left(\widehat{\overline{\overline{Y}}}_{\gamma} - \overline{\overline{y}}_\gamma\right)J_\gamma\sum_{i\in\mathcal{I}_\gamma}\left(\widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma) - \overline{\overline{y}}_\gamma\right) - \frac{2}{I_\gamma J_\gamma}\left(\widehat{\overline{\overline{Y}}}_{\gamma} - \overline{\overline{y}}_\gamma\right)I_\gamma\sum_{j\in\mathcal{J}_\gamma}\left(\widehat{\overline{Y}}^{\mathrm{S}}_j(\gamma) - \overline{\overline{y}}_\gamma\right) \\
&= \frac{1}{I_\gamma J_\gamma}\sum_{i\in\mathcal{I}_\gamma}\sum_{j\in\mathcal{J}_\gamma}\left(y_{i,j}(\gamma) - \overline{\overline{y}}_\gamma\right)^2 + \frac{1}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma}\left(\widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma) - \overline{\overline{y}}_\gamma\right)^2 \\
&+ \frac{1}{J_\gamma}\sum_{j\in\mathcal{J}_\gamma}\left(\widehat{\overline{Y}}^{\mathrm{S}}_j(\gamma) - \overline{\overline{y}}_\gamma\right) + \left(\widehat{\overline{\overline{Y}}}_{\gamma} - \overline{\overline{y}}_\gamma\right)^2 - \frac{2}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma}\left(\widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma) - \overline{\overline{y}}_\gamma\right)^2 - \frac{2}{J_\gamma}\sum_{j\in\mathcal{J}_\gamma}\left(\widehat{\overline{Y}}^{\mathrm{S}}_j(\gamma) - \overline{\overline{y}}_\gamma\right),
\end{aligned}$$

so

$$\begin{aligned}
\hat{\Sigma}^{\mathrm{BS}}_{\gamma} = &\frac{1}{I_\gamma J_\gamma}\sum_{i\in\mathcal{I}_\gamma}\sum_{j\in\mathcal{J}_\gamma}\left(y_{i,j}(\gamma) - \overline{\overline{y}}_\gamma\right)^2 - \frac{1}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma}\left(\widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma) - \overline{\overline{y}}_\gamma\right)^2 \\
&- \frac{1}{J_\gamma}\sum_{j\in\mathcal{J}_\gamma}\left(\widehat{\overline{Y}}^{\mathrm{S}}_j(\gamma) - \overline{\overline{y}}_\gamma\right)^2 + \left(\widehat{\overline{\overline{Y}}}_{\gamma} - \overline{\overline{y}}_\gamma\right)^2.
\end{aligned}$$

Under the expectation operator,

$$\begin{aligned}\mathbb{E}\left[\hat{\Sigma}^{\mathrm{BS}}_\gamma\right] &= \frac{1}{I_\gamma J_\gamma}\mathbb{E}\left[\sum_{i\in\mathcal{I}_\gamma}\sum_{j\in\mathcal{J}_\gamma}\left\{y_{i,j}(\gamma)-\overline{\overline{y}}_\gamma\right\}^2\right] - \frac{1}{I_\gamma}\mathbb{E}\left[\sum_{i\in\mathcal{I}_\gamma}\left(\widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma)-\overline{\overline{y}}_\gamma\right)^2\right]\\ &\quad - \frac{1}{J_\gamma}\mathbb{E}\left[\sum_{j\in\mathcal{J}_\gamma}\left(\widehat{\overline{Y}}^{\mathrm{S}}_j(\gamma)-\overline{\overline{y}}_\gamma\right)\right] + \mathrm{Var}\left(\widehat{\overline{\overline{Y}}}_\gamma\right)\\ &= \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}\left\{y_{i,j}(\gamma)-\overline{\overline{y}}_\gamma\right\}^2 - \frac{1}{I_\gamma}\chi^{2,\mathrm{B}}(\gamma) - \frac{1}{J_\gamma}\chi^{2,\mathrm{S}}(\gamma) + \mathrm{Var}\left(\widehat{\overline{\overline{Y}}}_\gamma\right).\end{aligned}$$

Now, leveraging eq. (A.16) for the first summation, eq. (A.11) for the second summation, and eq. (A.15) for the third summation,

$$\begin{aligned}\mathbb{E}\left[\hat{\Sigma}^{\mathrm{BS}}_\gamma\right] &= \sigma^{\mathrm{B}}_\gamma + \sigma^{\mathrm{S}}_\gamma + \sigma^{\mathrm{BS}}_\gamma - \left[\sigma^{\mathrm{B}}_\gamma + \eta^{\mathrm{B}}_\gamma\right] - \left[\sigma^{\mathrm{S}}_\gamma + \eta^{\mathrm{S}}_\gamma\right] + \mathrm{Var}\left(\widehat{\overline{\overline{Y}}}_\gamma\right)\\ &= \sigma^{\mathrm{BS}}_\gamma + \mathrm{Var}\left(\widehat{\overline{\overline{Y}}}_\gamma\right) - \eta^{\mathrm{B}}_\gamma - \eta^{\mathrm{S}}_\gamma. \end{aligned} \tag{A.17}$$

□

We now use the characterizations eqs. (A.12), (A.14) and (A.17), to define an unbiased estimator for $\mathrm{Var}\left(\widehat{\overline{\overline{Y}}}_\gamma\right)$, as stated in theorem 4.4.

Theorem A.16 (Already theorem 4.4 in the main paper)**.** *For a SMRD where assumption 2.4 holds, for all* $\gamma \in \{\mathrm{cc}, \mathrm{ib}, \mathrm{is}, \mathrm{tr}\}$,

$$\mathbb{E}\left[\widehat{\Sigma}_\gamma\right] = \mathrm{Var}\left(\widehat{\overline{\overline{Y}}}_\gamma\right),$$

where

$$\begin{aligned}\hat{\Sigma}_\gamma &:= \frac{\alpha^{\mathrm{B}}_\gamma\hat{\Sigma}^{\mathrm{B}}_\gamma + \alpha^{\mathrm{S}}_\gamma\hat{\Sigma}^{\mathrm{S}}_\gamma + (\alpha^{\mathrm{B}}_\gamma\alpha^{\mathrm{S}}_\gamma)\hat{\Sigma}^{\mathrm{BS}}_\gamma}{1-\alpha^{\mathrm{B}}_\gamma-\alpha^{\mathrm{S}}_\gamma+\alpha^{\mathrm{B}}_\gamma\alpha^{\mathrm{S}}_\gamma}\\ &\quad - \frac{\alpha^{\mathrm{B}}_\gamma}{1-\alpha^{\mathrm{B}}_\gamma}\frac{J-J_\gamma}{(J-1)(J_\gamma-1)}\frac{1}{J_\gamma I_\gamma}\sum_{i\in\mathcal{I}_\gamma}\sum_{j\in\mathcal{J}_\gamma}\left(y_{i,j}(\gamma)-\widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma)\right)^2\\ &\quad - \frac{\alpha^{\mathrm{S}}_\gamma}{1-\alpha^{\mathrm{S}}_\gamma}\frac{I-I_\gamma}{(I-1)(I_\gamma-1)}\frac{1}{I_\gamma J_\gamma}\sum_{j\in\mathcal{J}_\gamma}\sum_{i\in\mathcal{I}_\gamma}\left(y_{i,j}(\gamma)-\widehat{\overline{Y}}^{\mathrm{S}}_j(\gamma)\right)^2,\end{aligned}$$

and where we have used the previously defined (non-random) coefficients $\alpha^{\mathrm{B}}_\gamma$ *and* $\alpha^{\mathrm{S}}_\gamma$.

Proof of theorem 4.4 and theorem A.16. Given $\alpha^{\mathrm{B}}_\gamma, \alpha^{\mathrm{S}}_\gamma$, lemma A.7 allows us to write

$$\mathrm{Var}\left(\widehat{\overline{\overline{Y}}}_\gamma\right) = \alpha^{\mathrm{B}}_\gamma\sigma^{\mathrm{B}}_\gamma + \alpha^{\mathrm{S}}_\gamma\sigma^{\mathrm{S}}_\gamma + \alpha^{\mathrm{B}}_\gamma\alpha^{\mathrm{S}}_\gamma\sigma^{\mathrm{BS}}_\gamma.$$

Define

$$\hat{G}_\gamma = \alpha_\gamma^{\mathrm{B}}\hat{\Sigma}_\gamma^{\mathrm{B}} + \alpha_\gamma^{\mathrm{S}}\hat{\Sigma}_\gamma^{\mathrm{S}} + \alpha_\gamma^{\mathrm{B}}\alpha_\gamma^{\mathrm{S}}\hat{\Sigma}_\gamma^{\mathrm{BS}},$$

and apply the expectation operator, leveraging the results in lemmas A.12, A.13 and A.15,

$$\begin{aligned}\mathbb{E}\left[\hat{G}_\gamma\right] &= \alpha_\gamma^{\mathrm{B}}\,\mathbb{E}\left[\hat{\Sigma}_\gamma^{\mathrm{B}}\right] + \alpha_\gamma^{\mathrm{S}}\,\mathbb{E}\left[\hat{\Sigma}_\gamma^{\mathrm{S}}\right] + \alpha_\gamma^{\mathrm{B}}\alpha_\gamma^{\mathrm{S}}\,\mathbb{E}\left[\hat{\Sigma}_\gamma^{\mathrm{BS}}\right] \\ &= \alpha_\gamma^{\mathrm{B}}\left(\sigma_\gamma^{\mathrm{B}} - \mathrm{Var}\left(\widehat{\overline{\overline{Y}}}_\gamma\right) + \eta_\gamma^{\mathrm{B}}\right) + \alpha_\gamma^{\mathrm{S}}\left(\sigma_\gamma^{\mathrm{S}} - \mathrm{Var}\left(\widehat{\overline{\overline{Y}}}_\gamma\right) + \eta_\gamma^{\mathrm{S}}\right) \\ &\quad + \alpha_\gamma^{\mathrm{B}}\alpha_\gamma^{\mathrm{S}}\left(\sigma_\gamma^{\mathrm{BS}} + \mathrm{Var}(\widehat{\overline{\overline{Y}}}_\gamma) - \eta_\gamma^{\mathrm{B}} - \eta_\gamma^{\mathrm{S}}\right).\end{aligned}$$

Rearranging,

$$\mathbb{E}\left[\hat{G}_\gamma\right] = \mathrm{Var}\left(\widehat{\overline{\overline{Y}}}_\gamma\right)\left\{1 - \alpha_\gamma^{\mathrm{B}} - \alpha_\gamma^{\mathrm{S}} + \alpha_\gamma^{\mathrm{B}}\alpha_\gamma^{\mathrm{S}}\right\} + \alpha_\gamma^{\mathrm{B}}\{1 - \alpha_\gamma^{\mathrm{S}}\}\eta_\gamma^{\mathrm{B}} + \alpha_\gamma^{\mathrm{S}}\left\{1 - \alpha_\gamma^{\mathrm{B}}\right\}\eta_\gamma^{\mathrm{S}}.$$

and, observing that

$$\frac{x(1-y)}{1-x-y+xy} = \frac{x(1-y)}{(1-x)(1-y)} = \frac{x}{1-x},$$

and rescaling the quantity above,

$$\frac{\mathbb{E}\left[\hat{G}_\gamma\right]}{1 - \alpha_\gamma^{\mathrm{B}} - \alpha_\gamma^{\mathrm{S}} + \alpha_\gamma^{\mathrm{B}}\alpha_\gamma^{\mathrm{S}}} = \mathrm{Var}\left(\widehat{\overline{\overline{Y}}}_\gamma\right) + \frac{\alpha_\gamma^{\mathrm{B}}}{1-\alpha_\gamma^{\mathrm{B}}}\eta_\gamma^{\mathrm{B}} + \frac{\alpha_\gamma^{\mathrm{S}}}{1-\alpha_\gamma^{\mathrm{S}}}\eta_\gamma^{\mathrm{S}}.$$

We now leverage standard results to obtain unbiased estimates for $\eta_\gamma^{\mathrm{B}}, \eta_\gamma^{\mathrm{S}}$. First, the variance of the row-mean estimate follows from lemma A.10:

$$\mathrm{Var}\left(\widehat{\overline{Y}}_i^{\mathrm{B}}(\gamma)\right) = \mathbb{E}\left[\left(\widehat{\overline{Y}}_i^{\mathrm{B}}(\gamma) - \overline{y}_i^{\mathrm{B}}(\gamma)\right)^2\right] = \frac{\left[\frac{1}{J-1}\sum_{j=1}^{J}\left\{y_{i,j}(\gamma) - \overline{y}_i^{\mathrm{B}}(\gamma)\right\}^2\right]}{\frac{J_\gamma J}{J-J_\gamma}}, \tag{A.18}$$

where eq. (A.18) is implied by standard results in sampling theory: in a SMRD we can see each row i as its own population with mean $\overline{y}_i^{\mathrm{B}}$ and corresponding estimate $\widehat{\overline{Y}}_i^{\mathrm{B}}$. Then, for those rows which feature at least two columns of type γ, we can provide an unbiased estimate of the variance term in eq. (A.18). Define the sample estimate

$$\widehat{\mathrm{Var}}\left(\widehat{\overline{Y}}_i^{\mathrm{B}}(\gamma)\right) := \frac{J-J_\gamma}{J_\gamma}\frac{1}{J}\left[\frac{1}{J_\gamma - 1}\sum_{j\in\mathcal{J}_\gamma}\left\{y_{i,j}(\gamma) - \widehat{\overline{Y}}_i^{\mathrm{B}}(\gamma)\right\}^2\right].$$

From eq. (A.7),

$$\mathbb{E}\left[\frac{1}{J_\gamma - 1}\sum_{j\in\mathcal{J}_\gamma}\left\{y_{i,j}(\gamma) - \widehat{\overline{Y}}_i^{\mathrm{B}}(\gamma)\right\}^2\right] = \frac{1}{J-1}\sum_{j=1}^{J}\left\{y_{i,j}(\gamma) - \overline{y}_i^{\mathrm{B}}(\gamma)\right\}^2,$$

which directly implies that

$$\mathbb{E}\left[\widehat{\text{Var}}\left(\widehat{\overline{Y}}_i^{\text{B}}(\gamma)\right)\right] = \text{Var}\left(\widehat{\overline{Y}}_i^{\text{B}}(\gamma)\right).$$

Averaging these estimates over the rows,

$$\widehat{\eta}_\gamma^{\text{B}} = \frac{1}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma}\widehat{\text{Var}}\left(\widehat{\overline{Y}}_i^{\text{B}}(\gamma)\right),$$

satisfying

$$\mathbb{E}\left[\widehat{\eta}_\gamma^{\text{B}}\right] = \mathbb{E}\left[\frac{1}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma}\widehat{\text{Var}}\left(\widehat{\overline{Y}}_i^{\text{B}}(\gamma)\right)\right] = \frac{1}{I}\binom{J_\gamma}{J}^{-1}\sum_{\mathcal{J}_\gamma}\sum_{i=1}^{I}\left\{\widehat{\overline{Y}}_i^{\text{B}}(\gamma) - \overline{y}_i^{\text{B}}(\gamma)\right\}^2 =: \eta_\gamma^{\text{B}}.$$

Symmetrically for the sellers,

$$\text{Var}\left(\widehat{\overline{Y}}_j^{\text{S}}(\gamma)\right) = \mathbb{E}\left[\left(\widehat{\overline{Y}}_j^{\text{S}}(\gamma) - \overline{y}_j^{\text{S}}(\gamma)\right)^2\right] = \frac{I-I_\gamma}{I_\gamma}\frac{1}{I}\frac{1}{I-1}\sum_{i=1}^{I}(y_{i,j}(\gamma) - \overline{y}_j^{\text{S}}(\gamma))^2,$$

then

$$\widehat{\text{Var}}\left(\widehat{\overline{Y}}_j^{\text{S}}(\gamma)\right) := \frac{I-I_\gamma}{I_\gamma}\frac{1}{I_\gamma - 1}\frac{1}{I}\sum_{i\in\mathcal{I}_\gamma}\left(y_{i,j}(\gamma) - \overline{y}_j^{\text{S}}(\gamma)\right)^2.$$

It holds

$$\mathbb{E}\left[\widehat{\text{Var}}\left(\widehat{\overline{Y}}_j^{\text{S}}(\gamma)\right)\right] = \text{Var}\left(\widehat{\overline{Y}}_j^{\text{S}}(\gamma)\right).$$

Average these estimates over the columns,

$$\widehat{\eta}_\gamma^{\text{S}} = \frac{1}{J_\gamma}\sum_{j\in\mathcal{J}_\gamma}\widehat{\text{Var}}\left(\widehat{\overline{Y}}_j^{\text{S}}(\gamma)\right), \quad \text{satisfying} \quad \mathbb{E}\left[\widehat{\eta}_\gamma^{\text{S}}\right] = \eta_\gamma^{\text{S}}.$$

Therefore,

$$\hat{\Sigma}_\gamma = \hat{G}_\gamma - \frac{\alpha_\gamma^{\text{B}}}{1-\alpha_\gamma^{\text{B}}}\widehat{\eta}_\gamma^{\text{B}} - \frac{\alpha_\gamma^{\text{S}}}{1-\alpha_\gamma^{\text{S}}}\widehat{\eta}_\gamma^{\text{S}} \quad \text{satisfies} \quad \mathbb{E}\left[\hat{\Sigma}_\gamma\right] = \text{Var}\left(\widehat{\overline{\overline{Y}}}_\gamma\right).$$

□

Theorem A.17 (Already theorem 4.5 in the main paper). *Under the assumptions of theorem 4.4 a conservative estimator for* $\text{Var}(\widehat{\tau}_{\text{spill}}^{\text{B}})$ *is:*

$$\widehat{\text{Var}}^{\text{hi}}(\widehat{\tau}_{\text{spill}}^{\text{B}}) := 2\left\{\widehat{\Sigma}_{\text{ib}} + \widehat{\Sigma}_{\text{cc}}\right\}.$$

Proof of theorem 4.5. Recall that $\widehat{\tau}^{\mathrm{B}}_{\mathrm{spill}} := \widehat{\overline{\overline{Y}}}_{\mathrm{ib}} - \widehat{\overline{\overline{Y}}}_{\mathrm{cc}}$, so that

$$\mathrm{Var}\left[\widehat{\tau}^{\mathrm{B}}_{\mathrm{spill}}\right] = \mathrm{Var}\left[\widehat{\overline{\overline{Y}}}_{\mathrm{ib}}\right] + \mathrm{Var}\left[\widehat{\overline{\overline{Y}}}_{\mathrm{cc}}\right] - 2\,\mathrm{Cov}(\mathrm{ib}, \mathrm{cc}). \tag{A.19}$$

We have unbiased estimators $\hat{\Sigma}_{\mathrm{ib}}$ for $\mathrm{Var}(\widehat{\overline{\overline{Y}}}_{\mathrm{ib}})$ and $\hat{\Sigma}_{\mathrm{cc}}$ for $\mathrm{Var}(\widehat{\overline{\overline{Y}}}_{\mathrm{cc}})$. To obtain a conservative variance estimator for $\widehat{\tau}^{\mathrm{B}}_{\mathrm{spill}}$, it remains for us to find a conservative estimator for the covariance term $\mathrm{Cov}(\mathrm{ib}, \mathrm{cc})$. Letting $\hat{\Delta}_\gamma := \widehat{\overline{\overline{Y}}}_\gamma - \overline{\overline{y}}_\gamma$, we use Young's (AM-GM) inequality as follows:

$$\mathrm{Cov}(\mathrm{ib}, \mathrm{cc}) \leq |\,\mathbb{E}[\hat{\Delta}_{\mathrm{ib}}\hat{\Delta}_{\mathrm{cc}}]| \leq \mathbb{E}[|\hat{\Delta}_{\mathrm{ib}}||\hat{\Delta}_{\mathrm{cc}}|] \leq \frac{\mathbb{E}[\hat{\Delta}^2_{\mathrm{ib}}]}{2}\frac{\mathbb{E}[\hat{\Delta}^2_{\mathrm{cc}}]}{2} = \frac{\Sigma_{\mathrm{ib}}}{2} + \frac{\Sigma_{\mathrm{cc}}}{2}. \tag{A.20}$$

Therefore, we can obtain from the sample a conservative estimator for the covariance via

$$\mathbb{E}\left[\frac{\hat{\Sigma}_{\mathrm{ib}}}{2} + \frac{\hat{\Sigma}_{\mathrm{cc}}}{2}\right] \geq |\,\mathrm{Cov}(\mathrm{ib}, \mathrm{cc})|. \tag{A.21}$$

Finally, defining

$$\widehat{\mathrm{Var}}^{\mathrm{hi}}(\widehat{\tau}^{\mathrm{B}}_{\mathrm{spill}}) := \widehat{\Sigma}(\mathrm{ib}) + \widehat{\Sigma}(\mathrm{cc}) + 2\left(\frac{\hat{\Sigma}_{\mathrm{ib}}}{2} + \frac{\hat{\Sigma}_{\mathrm{cc}}}{2}\right) = 2\left(\widehat{\Sigma}(\mathrm{ib}) + \widehat{\Sigma}(\mathrm{cc})\right)$$

and applying the expectation operator to each term we prove our thesis:

$$\mathbb{E}\left[\widehat{\mathrm{Var}}^{\mathrm{hi}}(\widehat{\tau}^{\mathrm{B}}_{\mathrm{spill}})\right] \geq \mathrm{Var}\left[\widehat{\overline{\overline{Y}}}_{\mathrm{ib}}\right] + \mathrm{Var}\left[\widehat{\overline{\overline{Y}}}_{\mathrm{cc}}\right] - 2\,\mathrm{Cov}(\mathrm{ib}, \mathrm{cc}) = \mathrm{Var}(\widehat{\tau}^{\mathrm{B}}_{\mathrm{spill}}).$$

□

Lemma A.18. *Let $\tau(\vec{\boldsymbol{\beta}})$ be as in eq. (9), and $\hat{\tau}(\vec{\boldsymbol{\beta}})$ be its estimator counterpart as per eq. (11). It holds $\mathbb{E}\left[\hat{\tau}(\vec{\boldsymbol{\beta}})\right] = \tau(\vec{\boldsymbol{\beta}})$. Extending theorem 4.5 yields a conservative estimator of $\mathrm{Var}\left(\hat{\tau}(\vec{\boldsymbol{\beta}})\right)$ via $\widehat{\mathrm{Var}}^{\mathrm{hi}}\left(\hat{\tau}(\vec{\boldsymbol{\beta}})\right) = \sum_{\gamma\in\{\mathrm{cc},\mathrm{ib},\mathrm{is},\mathrm{tr}\}} \beta^2_\gamma \hat{\Sigma}_\gamma + \sum_{\gamma\neq\gamma'} \beta_\gamma\beta_{\gamma'}\left(\hat{\Sigma}_\gamma + \hat{\Sigma}_{\gamma'}\right)$.*

Proof. Unbiasedness of $\hat{\tau}(\vec{\boldsymbol{\beta}})$ follows directly from linearity of the expectation and lemma 4.1. The variance of $\hat{\tau}(\vec{\boldsymbol{\beta}})$ es given by:

$$\mathrm{Var}\left(\hat{\tau}(\vec{\boldsymbol{\beta}})\right) = \sum_{\gamma\in\{\mathrm{cc},\mathrm{ib},\mathrm{is},\mathrm{tr}\}} \beta^2_\gamma\,\mathrm{Var}\left(\widehat{\overline{\overline{Y}}}_\gamma\right) + 2\sum_{\gamma\neq\gamma'} \beta_\gamma\beta_{\gamma'}\,\mathrm{Cov}\left(\widehat{\overline{\overline{Y}}}_\gamma, \widehat{\overline{\overline{Y}}}_{\gamma'}\right). \tag{A.22}$$

Plug-in estimates $\hat{\Sigma}_\gamma$ are unbiased for $\mathrm{Var}\left(\widehat{\overline{\overline{Y}}}_\gamma\right)$ as per theorem 4.4. The covariance terms can be bounded as in eq. (A.21) — $|\,\mathrm{Cov}(\gamma, \gamma')| \leq \frac{1}{2}\,\mathbb{E}[\hat{\Sigma}_\gamma + \hat{\Sigma}_{\gamma'}]$, yielding the result. □

A.5 Probability limit

Theorem A.19. *Consider any sequence of SMRDs in which $I, J \uparrow \infty$, where the local interference assumption holds, and which satisfy assumption 4.6. Let $\hat{\tau}(\vec{\boldsymbol{\beta}})$ be the linear estimator introduced in theorem 4.2, and let $\widehat{\mathrm{Var}}^{\mathrm{hi}}\left(\hat{\tau}(\vec{\boldsymbol{\beta}})\right)$ be its conservative variance estimator given in theorem 4.5. Then, if $I^{-2} + J^{-2} = o\left(\mathbb{E}\left\{\widehat{\mathrm{Var}}[\hat{\tau}(\vec{\boldsymbol{\beta}})]\right\}\right)$, we have*

$$\frac{\widehat{\mathrm{Var}}^{\mathrm{hi}}\left(\hat{\tau}(\vec{\boldsymbol{\beta}})\right)}{\mathbb{E}\left\{\widehat{\mathrm{Var}}^{\mathrm{hi}}\left(\hat{\tau}(\vec{\boldsymbol{\beta}})\right)\right\}} = 1 + o_p(1).$$

Proof. By the continuous mapping theorem, given the characterization of $\widehat{\mathrm{Var}}^{\mathrm{hi}}\left(\hat{\tau}(\vec{\boldsymbol{\beta}})\right)$ in lemma A.18, it suffices to consider the case where $\hat{\tau}(\vec{\boldsymbol{\beta}}) = \widehat{\overline{\overline{Y}}}_\gamma$, i.e. where $\vec{\boldsymbol{\beta}}$ is a standard basis vector in $\mathbb{R}^4$. In this case, theorem 4.4 shows that $\widehat{\mathrm{Var}}^{\mathrm{hi}}\left(\hat{\tau}(\vec{\boldsymbol{\beta}})\right) = \mathrm{Var}\left(\widehat{\overline{\overline{Y}}}_\gamma\right)$ is unbiased.

Given weights $\alpha^{\mathrm{B}}_\gamma$ and $\alpha^{\mathrm{S}}_\gamma$ defined in eq. (14), by theorem 4.3 (see also theorem A.9),

$$\Sigma_\gamma \coloneqq \mathrm{Var}\left(\widehat{\overline{\overline{Y}}}_\gamma\right) = \alpha^{\mathrm{B}}_\gamma \sigma^{\mathrm{B}}_\gamma + \alpha^{\mathrm{S}}_\gamma \sigma^{\mathrm{S}}_\gamma + \alpha^{\mathrm{B}}_\gamma \alpha^{\mathrm{S}}_\gamma \sigma^{\mathrm{BS}}_\gamma. \tag{A.23}$$

Using the facts that $\alpha^{\mathrm{B}}_\gamma = O(I^{-1})$ and $\alpha^{\mathrm{S}}_\gamma = O(J^{-1})$, we may note that as $I, J \uparrow \infty$, $1 - \alpha^{\mathrm{B}}_\gamma - \alpha^{\mathrm{S}}_\gamma + \alpha^{\mathrm{B}}_\gamma \alpha^{\mathrm{S}}_\gamma \sim 1, \quad 1 - \alpha^{\mathrm{S}}_\gamma \sim 1, \quad 1 - \alpha^{\mathrm{B}}_\gamma \sim 1$. Hence $\hat{\Sigma}_\gamma$ simplifies asymptotically as:

$$\hat{\Sigma}_\gamma \sim \underbrace{\alpha^{\mathrm{B}}_\gamma \widehat{\Sigma}^{\mathrm{B}}_\gamma}_{\text{(a.1)}} + \underbrace{\alpha^{\mathrm{S}}_\gamma \widehat{\Sigma}^{\mathrm{S}}_\gamma}_{\text{(a.2)}} + \underbrace{\alpha^{\mathrm{B}}_\gamma \alpha^{\mathrm{S}}_\gamma \widehat{\Sigma}^{\mathrm{BS}}_\gamma}_{\text{(b)}} - \underbrace{\frac{1}{J^2 I}\left[\frac{1}{I_\gamma}\sum_{i \in \mathcal{I}_\gamma} \hat{v}^{\mathrm{B}}_{\gamma,i}\right] - \frac{1}{I^2 J}\left[\frac{1}{J_\gamma}\sum_{j \in \mathcal{J}_\gamma} \hat{v}^{\mathrm{S}}_{\gamma,j}\right]}_{\text{(c)}}. \tag{A.24}$$

Here, we have adopted the notation

$$\hat{v}^{\mathrm{B}}_{\gamma,i} \coloneqq \frac{1}{J_\gamma}\sum_{j \in \mathcal{J}_\gamma}\left(y_{i,j}(\gamma) - \widehat{\overline{Y}}^{\mathrm{B}}_i(\gamma)\right)^2, \quad \text{and} \quad \hat{v}^{\mathrm{S}}_{\gamma,j} \coloneqq \frac{1}{I_\gamma}\sum_{i \in \mathcal{I}_\gamma}\left(y_{i,j}(\gamma) - \widehat{\overline{Y}}^{\mathrm{S}}_j(\gamma)\right)^2. \tag{A.25}$$

By boundedness assumption 4.6 (b), $|\hat{v}^{\mathrm{B}}_{\gamma,i}| \le 4C_2^2$ and $|\hat{v}^{\mathrm{S}}_{\gamma,j}| \le 4C_2^2$. We conclude immediately that (c) is $O(I^{-2}J^{-1} + I^{-1}J^{-2})$ almost surely. Thus, combining eqs. (A.23) and (A.24),

$$\frac{\hat{\Sigma}_\gamma}{\Sigma_\gamma} = 1 + \frac{\alpha^{\mathrm{B}}_\gamma(\widehat{\Sigma}^{\mathrm{B}}_\gamma - \sigma^{\mathrm{B}}_\gamma) + \alpha^{\mathrm{S}}_\gamma(\widehat{\Sigma}^{\mathrm{S}}_\gamma - \sigma^{\mathrm{S}}_\gamma) + \alpha^{\mathrm{B}}_\gamma \alpha^{\mathrm{S}}_\gamma(\widehat{\Sigma}^{\mathrm{BS}}_\gamma - \sigma^{\mathrm{BS}}_\gamma) + O(I^{-2}J^{-1} + I^{-1}J^{-2})}{\Sigma_\gamma}.$$

Then, to show the thesis it suffices for us to bound (a.1), (a.2) and (b) in eq. (A.24) above.

By Lemma A.24 along with the facts that $\Sigma_\gamma \ge \sigma^{\mathrm{B}}_\gamma$ and $\alpha^{\mathrm{B}}_\gamma = O(1/I)$,

$$\alpha^{\mathrm{B}}_\gamma\left(\widehat{\Sigma}^{\mathrm{B}}_\gamma - \sigma^{\mathrm{B}}_\gamma\right) = O_p\left(I^{-1}\sqrt{\Sigma_\gamma[I^{-1} + J^{-1}] + (IJ)^{-1}}\right).$$

Analogously, Lemma A.25 together with $\Sigma_\gamma \geq \sigma^{\mathrm{S}}_\gamma$ and $\alpha^{\mathrm{S}}_\gamma = O(1/J)$ gives

$$\alpha^{\mathrm{S}}_\gamma \left(\widehat{\Sigma}^{\mathrm{S}}_\gamma - \sigma^{\mathrm{S}}_\gamma\right) = O_p\left(J^{-1}\sqrt{\Sigma_\gamma[I^{-1}+J^{-1}]} + (IJ)^{-1}\right).$$

Lastly, Lemma A.26 together with $\Sigma_\gamma \geq \sigma^{\mathrm{S}}_\gamma + \sigma^{\mathrm{B}}_\gamma$ and $\alpha^{\mathrm{B}}_\gamma\alpha^{\mathrm{S}}_\gamma = O(I^{-1}J^{-1})$ gives

$$\alpha^{\mathrm{B}}_\gamma\alpha^{\mathrm{S}}_\gamma \left(\widehat{\Sigma}^{\mathrm{BS}}_\gamma - \sigma^{\mathrm{BS}}_\gamma\right) = O_p\left((IJ)^{-1}\sqrt{\Sigma_\gamma[I^{-1}+J^{-1}]} + (IJ)^{-1}[I^{-1}+J^{-1}]\right).$$

Omitting lower-order terms and simplifying fractions, we arrive at

$$\frac{\hat{\Sigma}_\gamma}{\Sigma_\gamma} = 1 + O_p\left(\sqrt{\frac{[I^{-1}+J^{-1}]^3}{\Sigma_\gamma}} + \frac{(IJ)^{-1}}{\Sigma_\gamma}\right) = 1 + o_p(1),$$

where the last equality holds because AM-GM ensures $(IJ)^{-1} \leq \frac{1}{2}[I^{-2}+J^{-2}]$. □

Given a parameter space U along with random variables $\{X_u : u \in U\}$ and real numbers $\{t_u : u \in U\}$, in this section we write $X_u = O_p^{\mathrm{fin}}(t_u)$ if $\{X_u/t_u : u \in U\}$ is tight: $\sup_{u\in U}\mathbb{P}(|X_u/t_u| > r) \downarrow 0$ as $r \uparrow \infty$ [Billingsley, 2008]. This immediately implies the usual, sequential definition: given a sequence of elements $u_n \in U$ such that $X_{u_n} = O_p^{\mathrm{fin}}(t_{u_n})$, it follows immediately that $X_{u_n} = O_p(t_{u_n})$ in the usual sense, meaning that the sequence X_{u_n}/t_{u_n} is tight: $\sup_n \mathbb{P}(|X_{u_n}/t_{u_n}| > r) \downarrow 0$ as $r \uparrow \infty$. We typically omit reference to the parameter space U as it will be clear from context.

Lemma A.20 (Single randomized convergence)**.** *Let $a_1, a_2, \ldots, a_I$ be bounded real numbers, $|a_i| \leq M$ for all $1 \leq i \leq I$. Write $\bar{a} = I^{-1}\sum_{i=1}^I a_i$ and $\sigma_a^2 = I^{-1}\sum_{i=1}^I (a_i - \bar{a})^2$. Then,*

$$\frac{1}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma}(a_i - \bar{a})^2 - \sigma_a^2 = O_p^{\mathrm{fin}}\left(\sqrt{M^2\sigma_a^2/I_\gamma}\right) \quad \text{and} \quad \frac{1}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma} a_i - \bar{a} = O_p^{\mathrm{fin}}\left(\sqrt{\sigma_a^2/I_\gamma}\right). \tag{A.26}$$

Proof. Equation (A.26) is taken from the proof of Li and Ding [2017b, Proposition 1]. In particular, the variance of both sums are bounded there, and Equation (A.26) then follows by Chebyshev's inequality. □

Lemma A.21. *Let $X_1, X_2, \ldots, X_n$ be random variables with the following tail bound property: for all $p \in (0,1)$ and all $i \in \{1, 2, ..., n\}$, $\mathbb{P}\left(|X_i| > D\sqrt{\log(2/p)}\right) \leq p$, where $D > 0$ is a constant. Then for a fixed probability $\eta \in (0,1)$,*

$$\mathbb{P}\left(\max_{1\leq i\leq n}|X_i| > D\sqrt{\log(2n/\eta)}\right) \leq \eta.$$

Proof. Set $p = \eta/n$, then

$$\mathbb{P}\left(|X_i| > D\sqrt{\log(n/\eta)}\right) \leq \eta/n,$$

$\forall i \in [n]$. By the union bound, for a fixed probability $\eta \in (0,1)$, $\mathbb{P}\left(\max_{1\le i\le n} |X_i| > D\sqrt{\log(2n/\eta)}\right) \le \eta$. ☐

In what follows, we consider a bounded array of real numbers $A = (a_{ij})_{i\in[I],j\in[J]}$ such that for all (i,j), $|a_{ij}| \le M$. For this array, we write:

$$\bar{\bar{a}} = (IJ)^{-1}\sum_{i=1}^{I}\sum_{j=1}^{J} a_{ij}, \quad \text{and} \quad \bar{a}_i = J^{-1}\sum_{j=1}^{J} a_{ij}, \quad \text{and} \quad \sigma_{\bar{a}}^2 = I^{-1}\sum_{i=1}^{I}(\bar{a}_i - \bar{\bar{a}})^2.$$

Lemma A.22. *Let $\mathcal{I}_\gamma$ be a random selection of $I_\gamma \in \{2,\dots,I-2\}$ indices (and symmetrically $\mathcal{J}_\gamma$ a selection of a random selection of $J_\gamma \in \{2,\dots,J-2\}$ indices). It holds:*

$$\frac{1}{I_\gamma J_\gamma}\sum_{i\in\mathcal{I}_\gamma}\sum_{j\in\mathcal{J}_\gamma}(a_{ij} - \bar{\bar{a}}) = O_p^{\text{fin}}\left(M\left[\sqrt{\frac{\sigma_{\bar{a}}^2}{I_\gamma}} + \sqrt{\frac{\log I}{I_\gamma J_\gamma}}\right]\right). \tag{A.27}$$

Proof. We write the left-hand side of eq. (A.27) as

$$\frac{1}{I_\gamma J_\gamma}\sum_{i\in\mathcal{I}_\gamma}\sum_{j\in\mathcal{J}_\gamma}(a_{ij} - \bar{\bar{a}}) = \frac{1}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma}\left[\left(\frac{1}{J_\gamma}\sum_{j\in\mathcal{J}_\gamma} a_{ij} - \bar{a}_i\right) + (\bar{a}_i - \bar{\bar{a}})\right] =: \frac{1}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma}\left[\epsilon_i + (\bar{a}_i - \bar{\bar{a}})\right]. \tag{A.28}$$

Here we have defined $\epsilon_i := \frac{1}{J_\gamma}\sum_{j\in\mathcal{J}_\gamma} a_{ij} - \bar{a}_i$, which depends only on the seller randomization. The second summand on the right-hand side of eq. (A.28) is $O_p^{\text{fin}}(\sqrt{\sigma_{\bar{a}}^2/I_\gamma})$ by eq. (A.26). For the first summand, we begin by bounding $\tilde{M} = \max_{i\in[I]} |\epsilon_i|$, which again depends only upon the seller randomization. By lemma B.6, which shows concentration of single-randomized sums, we have

$$\mathbb{P}\{|\epsilon_i| > CMJ_\gamma^{-1/2}\sqrt{\log(2/\eta)}\} \le \eta. \tag{A.29}$$

By lemma A.21, we then find that

$$\mathbb{P}\left\{\tilde{M} > CMJ_\gamma^{-1/2}\sqrt{\log(2I/\eta)}\right\} \le \eta. \tag{A.30}$$

Let $\mathcal{A}_1 = \left\{\left|\frac{1}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma}\epsilon_i\right| > C\tilde{M}I_\gamma^{-1/2}\sqrt{\log(2/\eta)}\right\}$, and $\mathcal{A}_2 := \{\tilde{M} > CMJ_\gamma^{-1/2}\sqrt{\log(2I/\eta)}\}$, for $\eta > 0$. Then, (i) Lemma B.6 (conditional on seller randomization) implies $\mathbb{P}\{\mathcal{A}_1\} \le \eta$, and (ii) eq. (A.30) implies $\mathbb{P}\{\mathcal{A}_2\} \le \eta$. By the union bound, $\mathbb{P}(\mathcal{A}_1^c \cup \mathcal{A}_1^c) \le \mathbb{P}(\mathcal{A}_1^c) + \mathbb{P}(\mathcal{A}_1^c) \le 2\eta$. Both bounds hold simultaneously with probability at least $1 - 2\eta$. When both hold:

$$\begin{aligned}
\left|\frac{1}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma}\epsilon_i\right| &\le C\tilde{M}I_\gamma^{-1/2}\sqrt{\log(2/\eta)} \\
&\le C\left(CMJ_\gamma^{-1/2}\sqrt{\log(2I/\eta)}\right)I_\gamma^{-1/2}\sqrt{\log(2/\eta)} \\
&= C^2M(I_\gamma J_\gamma)^{-1/2}\sqrt{\log(2I/\eta)\cdot\log(2/\eta)}
\end{aligned}$$

For any fixed η, $\sqrt{\log(2/\eta)} = O(1)$ and $\sqrt{\log(2I/\eta)} = O(\sqrt{\log I})$, giving us

$$\frac{1}{I_\gamma J_\gamma}\sum_{i\in\mathcal{I}_\gamma}\sum_{j\in\mathcal{J}_\gamma}(a_{ij}-\bar{\bar{a}}) = \frac{1}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma}\epsilon_i + \frac{1}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma}\bar{a}_i - \bar{\bar{a}} = O_p^{\text{fin}}\left(M\cdot\sqrt{\frac{\log I}{I_\gamma J_\gamma}} + \sqrt{\frac{\sigma_{\bar{a}}^2}{I_\gamma}}\right).$$

□

Lemma A.23. *Under the same assumptions of lemma A.22,*

$$\frac{1}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma}\left(\frac{1}{J_\gamma}\sum_{j\in\mathcal{J}_\gamma}a_{ij}-\bar{a}_i\right)^2 = O_p^{\text{fin}}\left(\frac{M^2}{J_\gamma}\right). \tag{A.31}$$

Proof. We use facts about Orlicz norms, collected in definition B.11. To prove eq. (A.31), note that eq. (A.29) implies $\|\epsilon_i\|_{\psi_2} = O(MJ_\gamma^{-1/2})$. So $\|\epsilon_i^2\|_{\psi_1} = \|\epsilon_i\|_{\psi_2}^2 = O(M^2J_\gamma^{-1})$. Given any buyer assignment via $\mathcal{I}_\gamma$, we have $\|I_\gamma^{-1}\sum_{i\in\mathcal{I}_\gamma}\epsilon_i^2\|_{\psi_1} \le I_\gamma^{-1}\sum_{i\in\mathcal{I}_\gamma}\|\epsilon_i^2\|_{\psi_1} = O(M^2J_\gamma^{-1})$ by Jensen's inequality. Thus, conditional upon any seller assignment, the left-hand side of (A.31) is $O_p^{\text{fin}}(M^2J_\gamma^{-1})$, so its marginal distribution is also $O_p^{\text{fin}}(M^2J_\gamma^{-1})$. □

Bounding key terms

Lemma A.24. *Consider a sequence of SMRDs with sample sizes $I, J \uparrow \infty$ where assumption 2.4 and assumption 4.6 hold. Then,*

$$\hat{\Sigma}_\gamma^{\text{B}} = \sigma_\gamma^{\text{B}} + O_p\left(\sqrt{\sigma_\gamma^{\text{B}}[I_\gamma^{-1}+J_\gamma^{-1}]} + J_\gamma^{-1}\right). \tag{A.32}$$

Proof. We decompose $\hat{\Sigma}_\gamma^{\text{B}}$ as in Lemma A.12,

$$\hat{\Sigma}_\gamma^{\text{B}} = \frac{1}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma}\left(\widehat{\overline{Y}}_i^{\text{B}}(\gamma) \pm \bar{\bar{y}}_\gamma - \widehat{\overline{\overline{Y}}}_\gamma\right)^2 = \underbrace{\frac{1}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma}\left[\widehat{\overline{Y}}_i^{\text{B}}(\gamma) - \bar{\bar{y}}_\gamma\right]^2}_{\text{(a.1.1)}} - \underbrace{\left[\widehat{\overline{\overline{Y}}}_\gamma - \bar{\bar{y}}_\gamma\right]^2}_{\text{(a.1.2)}}.$$

We analyze (a.1.1) and (a.1.2) separately.

Bounding (a.1.1): The term (a.1.1) can be decomposed as

$$\begin{aligned}\text{(a.1.1)} &= \frac{1}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma}\left[\widehat{\overline{Y}}_i^{\text{B}}(\gamma) \pm \bar{y}_i^{\text{B}}(\gamma) - \bar{\bar{y}}_\gamma\right]^2 \\ &= \frac{1}{I_\gamma}\left[\sum_{i\in\mathcal{I}_\gamma}\left[\delta_i^{\text{B}}(\gamma)\right]^2 + 2\sum_{i\in\mathcal{I}_\gamma}\delta_i^{\text{B}}(\gamma)\epsilon_i^{\text{B}} + \sum_{i\in\mathcal{I}_\gamma}\left(\epsilon_i^{\text{B}}(\gamma)\right)^2\right],\end{aligned} \tag{A.33}$$

where $\epsilon_i^{\mathrm{B}}(\gamma) = \widehat{\overline{Y}}_i^{\mathrm{B}}(\gamma) - \overline{y}_i^{\mathrm{B}}(\gamma)$ represents the sampling error in row i. By eq. (A.26) in Lemma A.20 and eq. (A.31) in Lemma A.22 , with $a_{ij} = y_{ij}(\gamma)$ and $M = C_2$, we have

$$\frac{1}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma}\left[\delta_i^{\mathrm{B}}(\gamma)\right]^2 = \sigma_\gamma^{\mathrm{B}} + O_p\left(\sqrt{C_2^2\sigma_\gamma^{\mathrm{B}}/I_\gamma}\right), \quad \frac{1}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma}\left(\epsilon_i^{\mathrm{B}}(\gamma)\right)^2 = O_p(J_\gamma^{-1}). \tag{A.34}$$

By Cauchy-Schwarz, $\frac{1}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma}\left[\delta_i^{\mathrm{B}}(\gamma)\right]\left(\epsilon_i^{\mathrm{B}}\right) \leq \left(\frac{1}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma}\left[\delta_i^{\mathrm{B}}(\gamma)\right]^2\right)^{1/2}\left(\frac{1}{I_\gamma}\sum_{i\in\mathcal{I}_\gamma}\left(\epsilon_i^{\mathrm{B}}(\gamma)\right)^2\right)^{1/2}$, bounding the cross term in eq. (A.33). Substituting and absorbing lower order terms

$$(\mathsf{a.1.1}) = \sigma_\gamma^{\mathrm{B}} + O_p\left(\sqrt{\sigma_\gamma^{\mathrm{B}}[I_\gamma^{-1} + J_\gamma^{-1}]} + 1/J_\gamma\right) \tag{A.35}$$

Bounding (a.1.2)**:** Applying eq. (A.27), with $(a_{ij}) = [y_{ij}(\gamma)]$ and $M = C_2$, we have

$$\widehat{\overline{\overline{Y}}}_\gamma - \overline{\overline{y}}_\gamma = O_p\left(\sqrt{\frac{\sigma_\gamma^{\mathrm{B}}}{I_\gamma}} + \sqrt{\frac{\log I}{I_\gamma J_\gamma}}\right). \tag{A.36}$$

Taking the square of eq. (A.36), combining with (a.1.2), suppressing the dependence upon C_2, removing lower-order terms and combining with eq. (A.35), we obtain eq. (A.32). □

Lemma A.25. *Under the assumptions of lemma A.24,*

$$\hat{\Sigma}_\gamma^{\mathrm{S}} = \sigma_\gamma^{\mathrm{S}} + O_p\left(\sqrt{\sigma_\gamma^{\mathrm{S}}[I_\gamma^{-1} + J_\gamma^{-1}]} + I_\gamma^{-1}\right). \tag{A.37}$$

Proof. Symmetric to the above. □

Lemma A.26. *Under the assumptions of lemma A.24,*

$$\widehat{\Sigma}_\gamma^{\mathrm{BS}} = \sigma_\gamma^{\mathrm{BS}} + O_p\left(\sqrt{[\sigma_\gamma^{\mathrm{S}} + \sigma_\gamma^{\mathrm{B}}][I_\gamma^{-1} + J_\gamma^{-1}]} + [I_\gamma^{-1} + J_\gamma^{-1}]\right)$$

Proof. From the decomposition provided in lemma A.15, we have:

$$\widehat{\Sigma}_\gamma^{\mathrm{BS}} = \overbrace{\sum_{i\in\mathcal{I}_\gamma}\sum_{j\in\mathcal{J}_\gamma}\frac{\left(y_{i,j}(\gamma) - \overline{\overline{y}}_\gamma\right)^2}{I_\gamma J_\gamma}}^{(\mathsf{b.1})} - \overbrace{\sum_{i\in\mathcal{I}_\gamma}\frac{\left[\widehat{\overline{Y}}_i^{\mathrm{B}}(\gamma) - \overline{\overline{y}}_\gamma\right]^2}{I_\gamma}}^{(\mathsf{b.2})} - \overbrace{\sum_{j\in\mathcal{J}_\gamma}\frac{\left(\widehat{\overline{Y}}_j^{\mathrm{S}}(\gamma) - \overline{\overline{y}}_\gamma\right)^2}{J_\gamma}}^{(\mathsf{b.3})} + \overbrace{\left(\widehat{\overline{\overline{Y}}}_\gamma - \overline{\overline{y}}_\gamma\right)^2}^{(\mathsf{b.4})}.$$

We will analyze each term separately and establish their convergence properties.

Bounding (b.1) Applying eq. (A.31) with $a_{ij} = \delta_{ij}(\gamma)^2$ and $M = C_2^2$, we have

$$(\mathsf{b.1}) = \frac{1}{I_\gamma J_\gamma}\sum_{i\in\mathcal{I}_\gamma}\sum_{j\in\mathcal{J}_\gamma}\delta_{ij}(\gamma)^2 = \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}\delta_{ij}(\gamma)^2 + O_p\left(\sqrt{\frac{\sigma_\gamma^{\mathrm{B}}}{I_\gamma}} + \sqrt{\frac{\log I}{I_\gamma J_\gamma}}\right).$$

By eq. (A.17), we know: $\frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}\delta_{ij}(\gamma)^2 = \sigma_\gamma^{\mathrm{B}} + \sigma_\gamma^{\mathrm{S}} + \sigma_\gamma^{\mathrm{BS}}$, hence

$$(\mathsf{b.1}) = \sigma_\gamma^{\mathrm{B}} + \sigma_\gamma^{\mathrm{S}} + \sigma_\gamma^{\mathrm{BS}} + O_p\left(\sqrt{\frac{\sigma_\gamma^{\mathrm{B}}}{I_\gamma}} + \sqrt{\frac{\log I}{I_\gamma J_\gamma}}\right). \tag{A.38}$$

Bounding (b.2)**:** From the analysis of (a.1.1) in eq. (A.35), we know that

$$(\mathsf{b.2}) = \sigma_\gamma^{\mathrm{B}} + O_p\left(\sqrt{\sigma_\gamma^{\mathrm{B}}[I_\gamma^{-1} + J_\gamma^{-1}]} + 1/J_\gamma\right).$$

Bounding (b.3)**:** Symmetrical to (b.2), applying the same analysis to columns instead of rows,

$$(\mathsf{b.3}) = \sigma_\gamma^{\mathrm{S}} + O_p\left(\sqrt{\sigma_\gamma^{\mathrm{S}}[I_\gamma^{-1} + J_\gamma^{-1}]} + 1/I_\gamma\right).$$

Bounding (b.4)**:** Taking the square of eq. (A.36) and using $(a+b)^2 \le 2(a^2+b^2)$ (by am-gm),

$$(\mathsf{b.4}) = \left(\widehat{\overline{\overline{Y}}}_\gamma - \overline{\overline{y}}_\gamma\right)^2 = O_p\left(\frac{\sigma_\gamma^{\mathrm{B}}}{I_\gamma} + \frac{\log I}{I_\gamma J_\gamma}\right).$$

Combining (b) **terms.** Using $\sqrt{a} + \sqrt{b} \le \sqrt{a+b}$ and removing lower-order terms,

$$\widehat{\Sigma}_\gamma^{\mathrm{BS}} = \sigma_\gamma^{\mathrm{BS}} + O_p\left(\sqrt{[\sigma_\gamma^{\mathrm{S}} + \sigma_\gamma^{\mathrm{B}}][I_\gamma^{-1} + J_\gamma^{-1}]} + [I_\gamma^{-1} + J_\gamma^{-1}]\right).$$

□

B Proof of Theorem 4.8

In this section we prove Theorem 4.8. We consider an SDRD with two populations (buyers, sellers), and a binary treatment assignment at the (buyer-seller) pair level. A fixed proportion of buyers $p^{\mathrm{B}} := I_{\mathrm{T}}/I \in (0,1)$ are assigned at random $W_i^B = 1$, which makes them eligible for treatment. The remaining $I_{\mathrm{C}} = I - I_{\mathrm{T}}$ are assigned $W_i^B = 0$. Similarly, $p^{\mathrm{S}} := J_{\mathrm{T}}/J \in (0,1)$ of sellers are assigned $W_j^{\mathrm{S}} = 1$ (i.e., are eligible), while the remaining $J_{\mathrm{C}} = J - J_{\mathrm{T}}$ sellers are assigned $W_j^{\mathrm{S}} = 0$. Treatment is assigned via $W_{ij} = W_i^{\mathrm{B}} W_j^{\mathrm{S}}$.

Remarks on notation Recall that $[n] := \{1,\ldots,n\}$. Given a k-dimensional vector $\boldsymbol{a}$, $\|\boldsymbol{a}\|_2 = a_1^2 + \ldots + a_k^2$ denotes its 2-norm and $\|A\|_{op} = \max_{\|x\|_2=1}\|Ax\|_2$ its operator norm. We often use I_0 in place of I_{C} and I_1 in place of I_{T} (symmetrically, J_0 for J_{C} and J_1 for J_{T}) whenever it is more natural to do so. Last, $C, C', C'', \ldots$ denote absolute positive constants whose value may change from line to line. Under local interference (Assumption 2.4), as per Lemma 3.5, each buyer-seller pair (i,j) has only 4 potential outcomes: $Y_{i,j} = Y_{i,j}(\gamma)$, where $\gamma \in \{\mathrm{cc}, \mathrm{ib}, \mathrm{is}, \mathrm{tr}\}$. We denote with $\gamma_{i,j}$ the type of the pair (i,j), as per Equation (7).

Goal of the proof For a fixed size of the two populations, $\boldsymbol{N} = (I, J)$ and for $\gamma \in \{\text{cc}, \text{ib}, \text{is}, \text{tr}\}$, we aim to prove joint normality of linear combinations of the random variables

$$\widehat{\overline{\overline{Y}}}_{\gamma} = \widehat{\overline{\overline{Y}}}_{\gamma,\boldsymbol{N}} = \frac{1}{N_\gamma} \sum_{i=1}^{I} \sum_{j=1}^{J} Y_{ij}(\gamma) \mathbf{1}\{\gamma_{i,j} = \gamma\},$$

where $N_\gamma = \sum_{i,j} 1(\gamma_{ij} = \gamma)$. We write

$$\widehat{\boldsymbol{\tau}} = \left[\widehat{\overline{\overline{Y}}}_{\text{cc}}, \widehat{\overline{\overline{Y}}}_{\text{ib}}, \widehat{\overline{\overline{Y}}}_{\text{is}}, \widehat{\overline{\overline{Y}}}_{\text{tr}}\right]^\top \equiv [\hat{\tau}_{\text{cc}}, \hat{\tau}_{\text{ib}}, \hat{\tau}_{\text{is}}, \hat{\tau}_{\text{tr}}]^\top \tag{B.1}$$

to denote the (random) vector of group averages, and $\boldsymbol{\tau}$ to denote its population counterpart,

$$\boldsymbol{\tau} = \left[\overline{\overline{y}}_{\text{cc}}, \overline{\overline{y}}_{\text{ib}}, \overline{\overline{y}}_{\text{is}}, \overline{\overline{y}}_{\text{tr}}\right]^\top \equiv [\tau(\text{cc}), \tau(\text{ib}), \tau(\text{is}), \tau(\text{tr})]^\top . \tag{B.2}$$

Roughly, our proof technique is as follows:

Step 1 In Appendix B.1 we show that under fixed sellers' assignments W_j^{S}, for $j \in [J]$, standard results of Li and Ding [2017b], Shi and Ding [2022b] extend to SMRDs: a "conditional" CLT for $\widehat{\boldsymbol{\tau}}$ holds, with the limiting distribution parameterized by the sellers' assignments.

Step 2 In Appendix B.2 we prove that when considering the random assignment of sellers, the mean of the limiting distribution in Step 1 is itself normally distributed. Meanwhile, its variance is close to a fixed, deterministic value, independent of both assignments.

Step 3 Last, we combine these show in Appendix B.3 that the marginal distribution of $\hat{\tau}(\vec{\boldsymbol{\beta}})$ is also approximately Gaussian.

B.1 Step 1: a conditional CLT

We now show that conditional upon the seller assignments (W_j^S), we can derive central limit theorems for the MRD estimators in Section 4 that mirror those known for estimators in standard, single randomized experiments [Li and Ding, 2017b, Shi and Ding, 2022b]. Let Π denote a uniform random permutation of the seller indices $[J]$, i.e. a map $\Pi : [J] \to [J]$ such that $\{\Pi(1), \ldots, \Pi(J)\} = [J]$. Without loss of generality, we can suppose treatment labels are generated according to $W_j^S = \mathbb{1}\{\Pi^{-1}(j) > J_0\}$. We proceed in this section by conditioning on a particular realization $\Pi = \pi$.

For a fixed permutation π, let $\mathcal{J}_0^\pi := \{\pi(1), \ldots, \pi(J_0)\} \subset [J]$ be the set of sellers with $W_j^{\text{S}} = 0$ and let $\mathcal{J}_1^\pi := \{\pi(J_0 + 1), \ldots, \pi(J)\} = [J] \setminus \mathcal{J}_0^\pi$ be the seller indices with $W_j^{\text{S}} = 1$. Conditional upon $\Pi = \pi$, each buyer i has the following "realizable" potential outcomes:

- $\overline{y}^{\text{B}}_{i,\mathcal{J}_0^\pi}(\text{cc}) = \frac{1}{J_0} \sum_{j=1}^{J_0} Y_{i,\pi(j)}(\text{cc})$ and $\overline{y}^{\text{B}}_{i,\mathcal{J}_0^\pi}(\text{ib}) = \frac{1}{J_0} \sum_{j=1}^{J_0} Y_{i,\pi(j)}(\text{ib})$, which average the unit-level potential outcomes of interactions (i, j) for sellers with $W_j^{\text{S}} = 0$;
- $\overline{y}^{\text{B}}_{i,\mathcal{J}_1^\pi}(\text{is}) = \frac{1}{J_1} \sum_{j=J_0+1}^{J} Y_{i,\pi(j)}(\text{is})$ and $\overline{y}^{\text{B}}_{i,\mathcal{J}_1^\pi}(\text{tr}) = \frac{1}{J_1} \sum_{j=J_0+1}^{J} Y_{i,\pi(j)}(\text{tr})$, which average the unit-level potential outcomes of interactions (i, j) for sellers with $W_j^{\text{S}} = 1$.

We can then view our SDRD as a standard randomized experiment with I units, where each buyer i can be thought of as having potential outcomes corresponding to the above:

$$\overline{\boldsymbol{y}}^{\mathrm{B}}_{i,\pi}(0) = \left(\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_0^\pi}(\mathrm{cc}) \quad 0 \quad \overline{y}^{\mathrm{B}}_{i,\mathcal{J}_1^\pi}(\mathrm{is}) \quad 0\right)^\top, \ \overline{\boldsymbol{y}}^{\mathrm{B}}_{i,\pi}(1) = \left(0 \quad \overline{y}^{\mathrm{B}}_{i,\mathcal{J}_0^\pi}(\mathrm{ib}) \quad 0 \quad \overline{y}^{\mathrm{B}}_{i,\mathcal{J}_1^\pi}(\mathrm{tr})\right)^\top. \tag{B.3}$$

Notice that we have two potential outcomes for each buyer (since each buyer can either be assigned $W_i^{\mathrm{B}} = 0$ or $W_i^{\mathrm{B}} = 1$), and these potential outcomes are vectors in $\mathbb{R}^4$ (there is one potential outcome for each type $\gamma \in \{\mathrm{cc}, \mathrm{ib}, \mathrm{is}, \mathrm{tr}\}$). The population averages of these vectors are defined as:

$$\overline{\overline{\boldsymbol{y}}}_{\bullet,\pi}(0) = \left(\tfrac{1}{I}\textstyle\sum_{i=1}^I \overline{y}^{\mathrm{B}}_{i,\mathcal{J}_0^\pi}(\mathrm{cc}) \quad 0 \quad \tfrac{1}{I}\textstyle\sum_{i=1}^I \overline{y}^{\mathrm{B}}_{i,\mathcal{J}_1^\pi}(\mathrm{is}) \quad 0\right)^\top =: \left(\overline{\overline{y}}_{\bullet,\mathcal{J}_0^\pi}(\mathrm{cc}) \quad 0 \quad \overline{\overline{y}}_{\bullet,\mathcal{J}_1^\pi}(\mathrm{is}) \quad 0\right)^\top;$$
$$\overline{\overline{\boldsymbol{y}}}_{\bullet,\pi}(1) = \left(0 \quad \tfrac{1}{I}\textstyle\sum_{i=1}^I \overline{y}^{\mathrm{B}}_{i,\mathcal{J}_0^\pi}(\mathrm{ib}) \quad 0 \quad \tfrac{1}{I}\textstyle\sum_{i=1}^I \overline{y}^{\mathrm{B}}_{i,\mathcal{J}_1^\pi}(\mathrm{tr})\right)^\top =: \left(0 \quad \overline{\overline{y}}_{\bullet,\mathcal{J}_0^\pi}(\mathrm{ib}) \quad 0 \quad \overline{\overline{y}}_{\bullet,\mathcal{J}_1^\pi}(\mathrm{tr})\right)^\top.$$

We further define the difference between the outcome at the unit level (eq. (B.3)) and the mean across all units (previous display) at the buyer-level treatment $q = 0, 1$:

$$\dot{\overline{\boldsymbol{y}}}^{\mathrm{B}}_{i,\pi}(q) = \overline{\boldsymbol{y}}^{\mathrm{B}}_{i,\pi}(q) - \overline{\overline{\boldsymbol{y}}}_{\bullet,\pi}(q) \in \mathbb{R}^4. \tag{B.4}$$

Following Li and Ding [2017b] we define the buyer-level vector of treatment effects $\boldsymbol{\tau}_i^\pi$

$$\boldsymbol{\tau}_i^\pi = \left[\overline{\boldsymbol{y}}^{\mathrm{B}}_{i,\pi}(0) + \overline{\boldsymbol{y}}^{\mathrm{B}}_{i,\pi}(1)\right] = \left(\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_0^\pi}(\mathrm{cc}) \quad \overline{y}^{\mathrm{B}}_{i,\mathcal{J}_0^\pi}(\mathrm{ib}) \quad \overline{y}^{\mathrm{B}}_{i,\mathcal{J}_1^\pi}(\mathrm{is}) \quad \overline{y}^{\mathrm{B}}_{i,\mathcal{J}_1^\pi}(\mathrm{tr})\right)^\top.$$

In turn, define the π-conditional population average across all buyers:

$$\boldsymbol{\tau}^\pi = \frac{1}{I}\sum_{i=1}^I \boldsymbol{\tau}_i^\pi = \overline{\overline{\boldsymbol{y}}}_{\bullet,\pi}(0) + \overline{\overline{\boldsymbol{y}}}_{\bullet,\pi}(1) = \left(\overline{\overline{y}}_{\bullet,\mathcal{J}_0^\pi}(\mathrm{cc}) \quad \overline{\overline{y}}_{\bullet,\mathcal{J}_0^\pi}(\mathrm{ib}) \quad \overline{\overline{y}}_{\bullet,\mathcal{J}_1^\pi}(\mathrm{is}) \quad \overline{\overline{y}}_{\bullet,\mathcal{J}_1^\pi}(\mathrm{tr})\right)^\top. \tag{B.5}$$

We define the centered counterpart of $\boldsymbol{\tau}_i^\pi$:

$$\dot{\boldsymbol{\tau}}_i^\pi = \boldsymbol{\tau}_i^\pi - \boldsymbol{\tau}^\pi = \begin{pmatrix} \overline{y}^{\mathrm{B}}_{i,\mathcal{J}_0^\pi}(\mathrm{cc}) - \overline{\overline{y}}_{\bullet,\mathcal{J}_0^\pi}(\mathrm{cc}) \\ \overline{y}^{\mathrm{B}}_{i,\mathcal{J}_0^\pi}(\mathrm{ib}) - \overline{\overline{y}}_{\bullet,\mathcal{J}_0^\pi}(\mathrm{ib}) \\ \overline{y}^{\mathrm{B}}_{i,\mathcal{J}_1^\pi}(\mathrm{is}) - \overline{\overline{y}}_{\bullet,\mathcal{J}_1^\pi}(\mathrm{is}) \\ \overline{y}^{\mathrm{B}}_{i,\mathcal{J}_1^\pi}(\mathrm{tr}) - \overline{\overline{y}}_{\bullet,\mathcal{J}_1^\pi}(\mathrm{tr}) \end{pmatrix}. \tag{B.6}$$

Given a (random) assignment of buyers, $W_i^{\mathrm{B}} \in \{0, 1\}$ for $i = 1, \ldots, I$, the natural sample counterpart of $\boldsymbol{\tau}^\pi$ is $\widehat{\boldsymbol{\tau}}^\pi$, where we replace each coordinate with the sample mean across units i for which the type was observed, $\widehat{\overline{\overline{Y}}}_{\bullet,\mathcal{J}_\gamma^\pi}(\gamma)$:

$$\widehat{\boldsymbol{\tau}}^\pi = \begin{pmatrix} \frac{1}{I_0}\sum_{i:W_i^{\mathrm{B}}=0}^{I_0} \overline{y}^{\mathrm{B}}_{i,\mathcal{J}_0^\pi}(\mathrm{cc}) \\ \frac{1}{I_1}\sum_{i:W_i^{\mathrm{B}}=1}^{I_1} \overline{y}^{\mathrm{B}}_{i,\mathcal{J}_0^\pi}(\mathrm{ib}) \\ \frac{1}{I_0}\sum_{i:W_i^{\mathrm{B}}=0}^{I_0} \overline{y}^{\mathrm{B}}_{i,\mathcal{J}_1^\pi}(\mathrm{is}) \\ \frac{1}{I_1}\sum_{i:W_i^{\mathrm{B}}=1}^{I_1} \overline{y}^{\mathrm{B}}_{i,\mathcal{J}_1^\pi}(\mathrm{tr}) \end{pmatrix} = \begin{pmatrix} \widehat{\overline{\overline{Y}}}_{\bullet,\mathcal{J}_0^\pi}(\mathrm{cc}) \\ \widehat{\overline{\overline{Y}}}_{\bullet,\mathcal{J}_0^\pi}(\mathrm{ib}) \\ \widehat{\overline{\overline{Y}}}_{\bullet,\mathcal{J}_1^\pi}(\mathrm{is}) \\ \widehat{\overline{\overline{Y}}}_{\bullet,\mathcal{J}_1^\pi}(\mathrm{tr}) \end{pmatrix}. \tag{B.7}$$

The randomness in $\widehat{\boldsymbol{\tau}}^\pi$ only stems from the assignment of the I buyers via W_i^{B}. With this characterization in place, we can extend Li and Ding [2017b, Theorem 3] to our estimator

$\widehat{\boldsymbol{\tau}}^\pi$. First, define for $q, r \in \{0,1\}$ the finite population cross-covariance

$$S^\pi_{q,r} := \frac{1}{I-1}\sum_{i=1}^{I}\left\{\overline{\boldsymbol{y}}^{\mathrm{B}}_{i,\pi}(q) - \overline{\overline{\boldsymbol{y}}}_{\bullet,\pi}(q)\right\}\left\{\overline{\boldsymbol{y}}^{\mathrm{B}}_{i,\pi}(r) - \overline{\overline{\boldsymbol{y}}}_{\bullet,\pi}(r)\right\}^\top = \frac{1}{I-1}\sum_{i=1}^{I}\dot{\overline{\boldsymbol{y}}}^{\mathrm{B}}_{i,\pi}(q)\dot{\overline{\boldsymbol{y}}}^{\mathrm{B}}_{i,\pi}(r)^\top,$$

and the finite population covariance of the individual effects

$$S^2_{\widehat{\boldsymbol{\tau}}^\pi} := \frac{1}{I-1}\sum_{i=1}^{I}\left\{\boldsymbol{\tau}^\pi_i - \boldsymbol{\tau}^\pi\right\}\left\{\boldsymbol{\tau}^\pi_i - \boldsymbol{\tau}^\pi\right\}^\top. \tag{B.8}$$

Theorem B.1 (Theorem 3 in Li and Ding [2017b])**.** *Consider an SDRD under the local interference assumption. Conditionally on the sellers' assignments via $\pi : [J] \to [J]$, the experiment is equivalent to a single-randomized experiment with I units and $Q = 2$ treatments indexed by $W_i^{\mathrm{B}} \in \{0,1\}$ and potential outcomes $\overline{\boldsymbol{y}}^{\mathrm{B}}_{i,\pi}(W_i^{\mathrm{B}}) \in \mathbb{R}^4$. The estimator $\widehat{\boldsymbol{\tau}}^\pi$ is unbiased for $\boldsymbol{\tau}^\pi$:*

$$\mathbb{E}\left[\widehat{\boldsymbol{\tau}}^\pi\right] = \boldsymbol{\tau}^\pi.$$

The covariance of $\widehat{\boldsymbol{\tau}}^\pi$ is given by

$$V^\pi := \mathrm{Cov}\left\{\widehat{\boldsymbol{\tau}}^\pi\right\} = \sum_{q=1}^{Q}\frac{1}{I_q}S^\pi_{q,q} - \frac{1}{I}S^2_{\boldsymbol{\tau}^\pi},$$

where

$$\frac{S^\pi_{0,0}}{I_0} + \frac{S^\pi_{1,1}}{I_1} = \sum_{i=1}^{I}\begin{pmatrix} \frac{\left(\dot{\tau}_i^\pi(\mathrm{cc})\right)^2}{(I-1)I_0} & 0 & \frac{\left(\dot{\tau}_i^\pi(\mathrm{cc})\dot{\tau}_i^\pi(\mathrm{is})\right)}{(I-1)I_0} & 0 \\ 0 & \frac{\left(\dot{\tau}_i^\pi(\mathrm{ib})\right)^2}{(I-1)I_1} & 0 & \frac{\left(\dot{\tau}_i^\pi(\mathrm{ib})\dot{\tau}_i^\pi(\mathrm{tr})\right)}{(I-1)I_1} \\ \frac{\left(\dot{\tau}_i^\pi(\mathrm{cc})\dot{\tau}_i^\pi(\mathrm{is})\right)}{(I-1)I_0} & 0 & \frac{\left(\dot{\tau}_i^\pi(\mathrm{is})\right)^2}{(I-1)I_0} & 0 \\ 0 & \frac{\left(\dot{\tau}_i^\pi(\mathrm{ib})\dot{\tau}_i^\pi(\mathrm{tr})\right)}{(I-1)I_1} & 0 & \frac{\left(\dot{\tau}_i^\pi(\mathrm{tr})\right)^2}{(I-1)I_1} \end{pmatrix},$$

and

$$\frac{S^2_{\boldsymbol{\tau}^\pi}}{I} = \sum_{i=1}^{I}\begin{pmatrix} \frac{\left(\dot{\tau}_i^\pi(\mathrm{cc})\right)^2}{(I-1)I} & \frac{\left(\dot{\tau}_i^\pi(\mathrm{cc})\dot{\tau}_i^\pi(\mathrm{ib})\right)}{(I-1)I} & \frac{\left(\dot{\tau}_i^\pi(\mathrm{cc})\dot{\tau}_i^\pi(\mathrm{is})\right)}{(I-1)I} & \frac{\left(\dot{\tau}_i^\pi(\mathrm{cc})\dot{\tau}_i^\pi(\mathrm{tr})\right)}{(I-1)I} \\ \frac{\left(\dot{\tau}_i^\pi(\mathrm{cc})\dot{\tau}_i^\pi(\mathrm{ib})\right)}{(I-1)I} & \frac{\left(\dot{\tau}_i^\pi(\mathrm{ib})\right)^2}{(I-1)I} & \frac{\sum_{i=1}^{I}\left(\dot{\tau}_i^\pi(\mathrm{ib})\dot{\tau}_i^\pi(\mathrm{is})\right)}{(I-1)I} & \frac{\left(\dot{\tau}_i^\pi(\mathrm{ib})\dot{\tau}_i^\pi(\mathrm{tr})\right)}{(I-1)I} \\ \frac{\left(\dot{\tau}_i^\pi(\mathrm{cc})\dot{\tau}_i^\pi(\mathrm{is})\right)}{(I-1)I} & \frac{\left(\dot{\tau}_i^\pi(\mathrm{ib})\dot{\tau}_i^\pi(\mathrm{is})\right)}{(I-1)I} & \frac{\left(\dot{\tau}_i^\pi(\mathrm{is})\right)^2}{(I-1)I} & \frac{\left(\dot{\tau}_i^\pi(\mathrm{is})\dot{\tau}_i^\pi(\mathrm{tr})\right)}{(I-1)I} \\ \frac{\left(\dot{\tau}_i^\pi(\mathrm{cc})\dot{\tau}_i^\pi(\mathrm{tr})\right)}{(I-1)I} & \frac{\left(\dot{\tau}_i^\pi(\mathrm{ib})\dot{\tau}_i^\pi(\mathrm{tr})\right)}{(I-1)I} & \frac{\left(\dot{\tau}_i^\pi(\mathrm{is})\dot{\tau}_i^\pi(\mathrm{tr})\right)}{(I-1)I} & \frac{\left(\dot{\tau}_i^\pi(\mathrm{tr})\right)^2}{(I-1)I} \end{pmatrix}.$$

Hence,

$$\mathrm{Cov}\left\{\widehat{\boldsymbol{\tau}}^\pi\right\} = \sum_{i=1}^{I}\begin{pmatrix} \frac{\frac{I_1}{I_0}\left(\dot{\tau}_i^\pi(\mathrm{cc})\right)^2}{(I-1)I} & -\frac{\left(\dot{\tau}_i^\pi(\mathrm{cc})\dot{\tau}_i^\pi(\mathrm{ib})\right)}{(I-1)I} & \frac{\frac{I_1}{I_0}\left(\dot{\tau}_i^\pi(\mathrm{cc})\dot{\tau}_i^\pi(\mathrm{is})\right)}{(I-1)I} & -\frac{\left(\dot{\tau}_i^\pi(\mathrm{cc})\dot{\tau}_i^\pi(\mathrm{tr})\right)}{(I-1)I} \\ -\frac{\left(\dot{\tau}_i^\pi(\mathrm{cc})\dot{\tau}_i^\pi(\mathrm{ib})\right)}{(I-1)I} & \frac{\frac{I_0}{I_1}\left(\dot{\tau}_i^\pi(\mathrm{ib})\right)^2}{(I-1)I} & -\frac{\left(\dot{\tau}_i^\pi(\mathrm{ib})\dot{\tau}_i^\pi(\mathrm{is})\right)}{(I-1)I} & \frac{\frac{I_0}{I_1}\left(\dot{\tau}_i^\pi(\mathrm{ib})\dot{\tau}_i^\pi(\mathrm{tr})\right)}{(I-1)I} \\ \frac{\frac{I_1}{I_0}\left(\dot{\tau}_i^\pi(\mathrm{cc})\dot{\tau}_i^\pi(\mathrm{is})\right)}{(I-1)I} & -\frac{\left(\dot{\tau}_i^\pi(\mathrm{ib})\dot{\tau}_i^\pi(\mathrm{is})\right)}{(I-1)I} & \frac{\frac{I_1}{I_0}\left(\dot{\tau}_i^\pi(\mathrm{is})\right)^2}{(I-1)I} & -\frac{\left(\dot{\tau}_i^\pi(\mathrm{is})\dot{\tau}_i^\pi(\mathrm{tr})\right)}{(I-1)I} \\ -\frac{\left(\dot{\tau}_i^\pi(\mathrm{cc})\dot{\tau}_i^\pi(\mathrm{tr})\right)}{(I-1)I} & \frac{\frac{I_0}{I_1}\left(\dot{\tau}_i^\pi(\mathrm{ib})\dot{\tau}_i^\pi(\mathrm{tr})\right)}{(I-1)I} & -\frac{\left(\dot{\tau}_i^\pi(\mathrm{is})\dot{\tau}_i^\pi(\mathrm{tr})\right)}{(I-1)I} & \frac{\frac{I_0}{I_1}\left(\dot{\tau}_i^\pi(\mathrm{tr})\right)^2}{(I-1)I} \end{pmatrix}. \tag{B.9}$$

Proof. The proof is given in Theorem 3 in Li and Ding [2017b]. □

Theorem 4 in Li and Ding [2017b] provides a CLT which relies on the existence of an asymptotic limit for V^π. Since V^π is random in our context and depends upon the finite population, we instead derive a Berry-Esseen type result following Shi and Ding [2022b]. In what follows, $(V^{\pi})^{\frac{1}{2}}$ is defined as the symmetric square root of V^{π}, and $(V^{\pi})^{-\frac{1}{2}}$ is its pseudoinverse. In particular, we need not assume that V^{π} has full rank.

Theorem B.2 (Theorem S4 in Shi and Ding [2022b])**.** *Let $V^{\pi} := \mathrm{Cov}\{\widehat{\boldsymbol{\tau}}^{\pi}\}$ as characterized in (B.9). Then, there exists a universal constant C such that for all $\boldsymbol{\alpha} \in \mathbb{R}^4$ with $\|\boldsymbol{\alpha}\|_2 = 1$,*

$$\sup_{t\in\mathbb{R}} \left| \mathbb{P}\left\{ \boldsymbol{\alpha}^{\top} (V^{\pi})^{-\frac{1}{2}} (\widehat{\boldsymbol{\tau}}^{\pi} - \boldsymbol{\tau}^{\pi}) > t \right\} - \Phi(t) \right| \leq C \max_{i\in[I]} \max_{q\in\{0,1\}} \frac{|\boldsymbol{\alpha}^{\top} (V^{\pi})^{-\frac{1}{2}} \dot{\overline{\boldsymbol{y}}}^{\mathrm{B}}_{i,\pi}(q)|}{I_q}. \tag{B.10}$$

Proof. See Theorem S4 of Shi and Ding [2022b]. □

Theorem B.2 provides a Berry-Esseen bound for $\widehat{\boldsymbol{\tau}}^{\pi}$, where the upper bound depends on both $\dot{\overline{\boldsymbol{y}}}^{\mathrm{B}}_{i,\pi}$ and V^{π} (see the right-hand side of eq. (B.10)). Now, using assumption 4.6(b) of bounded potential outcomes, we state a slightly different form Theorem B.2 where the bound does not depend on $\dot{\overline{\boldsymbol{y}}}^{\mathrm{B}}_{i,\pi}$. We use the notation introduced in Equation (11), so that $\hat{\tau}^{\pi}(\vec{\boldsymbol{\beta}}) = \sum_{\gamma\in\{\mathrm{cc},\mathrm{ib},\mathrm{is},\mathrm{tr}\}} \hat{\tau}^{\pi}_{\gamma}\beta_{\gamma}$.

Lemma B.3. *Under the same setting of Theorem B.2 and further assuming bounded potential outcomes as per assumption 4.6(b):*

$$\sup_{t\in\mathbb{R}} \left| \mathbb{P}\left\{ \frac{\widehat{\tau}^{\pi}(\vec{\boldsymbol{\beta}}) - \tau^{\pi}(\vec{\boldsymbol{\beta}})}{\sqrt{\mathrm{Var}\{\hat{\tau}^{\pi}(\vec{\boldsymbol{\beta}})\}}} > t \right\} - \Phi(t) \right| \leq C \frac{C_2}{\min\{I_0, I_1\}} \frac{\|\vec{\boldsymbol{\beta}}\|_2}{\sqrt{\mathrm{Var}\{\hat{\tau}^{\pi}(\vec{\boldsymbol{\beta}})\}}}. \tag{B.11}$$

Proof. We first consider the case in which V^{π} is invertible. This case contains the main ideas and is technically simpler than the general case.
Invertible case. Let $\boldsymbol{\alpha} = (V^{\pi})^{\frac{1}{2}}\vec{\boldsymbol{\beta}} / \|(V^{\pi})^{\frac{1}{2}}\vec{\boldsymbol{\beta}}\|_2$, so that $\|\boldsymbol{\alpha}\|_2 = 1$ by construction. Plugging this choice of $\boldsymbol{\alpha}$ in Equation (B.10), Theorem B.2,

$$\sup_{t\in\mathbb{R}} \left| \mathbb{P}\left\{ \frac{\vec{\boldsymbol{\beta}}^{\top} (\widehat{\boldsymbol{\tau}}^{\pi} - \boldsymbol{\tau}^{\pi})}{\|(V^{\pi})^{\frac{1}{2}}\vec{\boldsymbol{\beta}}\|_2} > t \right\} - \Phi(t) \right| \leq \frac{C}{\|(V^{\pi})^{\frac{1}{2}}\vec{\boldsymbol{\beta}}\|_2} \max_{i\in[I]} \max_{q\in\{0,1\}} \frac{|\vec{\boldsymbol{\beta}}^{\top} \dot{\overline{\boldsymbol{y}}}^{\mathrm{B}}_{i,\pi}(q)|}{I_q}.$$

Applying the Cauchy-Schwarz inequality on the right hand side yields

$$\sup_{t\in\mathbb{R}} \left| \mathbb{P}\left\{ \frac{\vec{\boldsymbol{\beta}}^{\top} (\widehat{\boldsymbol{\tau}}^{\pi} - \boldsymbol{\tau}^{\pi})}{\|(V^{\pi})^{\frac{1}{2}}\vec{\boldsymbol{\beta}}\|_2} > t \right\} - \Phi(t) \right| \leq \frac{C}{\|(V^{\pi})^{\frac{1}{2}}\vec{\boldsymbol{\beta}}\|_2} \|\vec{\boldsymbol{\beta}}\|_2 \max_{i\in[I]} \max_{q\in\{0,1\}} \frac{\|\dot{\overline{\boldsymbol{y}}}^{\mathrm{B}}_{i,\pi}(q)\|_2}{I_q}.$$

Last, since by assumption 4.6 (b), each entry of $\dot{\overline{\boldsymbol{y}}}^{\mathrm{B}}_{i,\pi}(q)$ has absolute value at most $2C_2$, and since there are exactly 2 non-zero entries in each $\dot{\overline{\boldsymbol{y}}}^{\mathrm{B}}_{i,\pi}(q)$ (cf Equation (B.3)), we conclude that $\max_i \max_q \|\dot{\overline{\boldsymbol{y}}}^{\mathrm{B}}_{i,\pi}(q)\|_2 \leq \sqrt{2 \times (2C_2)^2} = \sqrt{8}C_2$. Plugging this in, and noting that

$\text{Var}\{\hat{\tau}^\pi(\vec{\boldsymbol{\beta}})\} = \vec{\boldsymbol{\beta}}^\top V^\pi \vec{\boldsymbol{\beta}}$, so that $\sqrt{\text{Var}\{\hat{\tau}^\pi(\vec{\boldsymbol{\beta}})\}} = \|(V^\pi)^{1/2}\vec{\boldsymbol{\beta}}\|_2$ yields the thesis:

$$\sup_{t\in\mathbb{R}} \left| \mathbb{P}\left\{ \frac{\hat{\tau}^\pi(\vec{\boldsymbol{\beta}}) - \tau^\pi(\vec{\boldsymbol{\beta}})}{\sqrt{\text{Var}\{\hat{\tau}^\pi(\vec{\boldsymbol{\beta}})\}}} > t \right\} - \Phi(t) \right| = \sup_{t\in\mathbb{R}} \left| \mathbb{P}\left\{ \frac{\vec{\boldsymbol{\beta}}^\top (\hat{\boldsymbol{\tau}}^\pi - \boldsymbol{\tau}^\pi)}{\|(V^\pi)^{\frac{1}{2}}\vec{\boldsymbol{\beta}}\|_2} > t \right\} - \Phi(t) \right|$$
$$\leq \frac{\sqrt{8}CC_2}{\min\{I_0, I_1\}} \frac{\|\vec{\boldsymbol{\beta}}\|_2}{\sqrt{\text{Var}\{\hat{\tau}^\pi(\vec{\boldsymbol{\beta}})\}}}.$$

Non-invertible case. In case V^π is not invertible, eq. (B.10) in Theorem B.2 instead gives

$$\sup_{t\in\mathbb{R}} \left| \mathbb{P}\left\{ \frac{\vec{\boldsymbol{\beta}}^\top (\hat{\boldsymbol{\tau}}^\pi - \boldsymbol{\tau}^\pi)}{\|(V^\pi)^{\frac{1}{2}}\vec{\boldsymbol{\beta}}\|_2} > t \right\} - \Phi(t) \right| \leq \frac{C}{\|(V^\pi)^{\frac{1}{2}}\vec{\boldsymbol{\beta}}\|_2} \max_{i\in[I]} \max_{q\in\{0,1\}} \frac{|\vec{\boldsymbol{\beta}}^\top (V^\pi)^{\frac{1}{2}} (V^\pi)^{-\frac{1}{2}} \dot{\bar{\boldsymbol{y}}}^{\text{B}}_{i,\pi}(q)|}{I_q}.$$

Now we use Cauchy-Schwarz and the operator norm inequality to bound the righthand side,

$$\begin{aligned} |\vec{\boldsymbol{\beta}}^\top (V^\pi)^{\frac{1}{2}} (V^\pi)^{-\frac{1}{2}} \dot{\bar{\boldsymbol{y}}}^{\text{B}}_{i,\pi}(q)| &\leq \|\vec{\boldsymbol{\beta}}^\top\|_2 \|(V^\pi)^{\frac{1}{2}} (V^\pi)^{-\frac{1}{2}} \dot{\bar{\boldsymbol{y}}}^{\text{B}}_{i,\pi}(q)\|_2 \\ &\leq \|\vec{\boldsymbol{\beta}}^\top\|_2 \|(V^\pi)^{\frac{1}{2}} (V^\pi)^{-\frac{1}{2}}\|_{op} \|\dot{\bar{\boldsymbol{y}}}^{\text{B}}_{i,\pi}(q)\|_2 \\ &\leq \|\vec{\boldsymbol{\beta}}^\top\|_2 \|\dot{\bar{\boldsymbol{y}}}^{\text{B}}_{i,\pi}(q)\|_2, \end{aligned}$$

where in the last step we use the fact that $\|(V^\pi)^{\frac{1}{2}} (V^\pi)^{-\frac{1}{2}}\|_{op} \leq 1$. Thus,

$$\sup_{t\in\mathbb{R}} \left| \mathbb{P}\left\{ \frac{\vec{\boldsymbol{\beta}}^\top (\hat{\boldsymbol{\tau}}^\pi - \boldsymbol{\tau}^\pi)}{\|(V^\pi)^{\frac{1}{2}}\vec{\boldsymbol{\beta}}\|_2} > t \right\} - \Phi(t) \right| \leq \frac{C}{\|(V^\pi)^{\frac{1}{2}}\vec{\boldsymbol{\beta}}\|_2} \|\vec{\boldsymbol{\beta}}^\top\|_2 \max_{i\in[I]} \max_{q\in\{0,1\}} \frac{\|\dot{\bar{\boldsymbol{y}}}^{\text{B}}_{i,\pi}(q)\|_2}{I_q}.$$

The proof then proceeds as in the invertible case. □

This concludes the first section.

B.2 Analysis of conditional mean and covariance

As before, let $\Pi : [J] \to [J]$ denote a permutation chosen uniformly at random. In this section we characterize the distribution of the Π-conditional mean vector $\boldsymbol{\tau}^\Pi$, and the Π-conditional covariance matrix $V^\Pi := \text{Cov}(\hat{\boldsymbol{\tau}}^\Pi)$ introduced in Appendix B.1. This allows us to transfer the results Appendix B.1 —which depend on the particular seller assignment π—to the general case of random seller assignment Π.

We start by recalling standard results on concentration of random permutations in Appendix B.2.1. We use these results to characterize the conditional mean and covariance. We characterize the mean $\mathbb{E}_\Pi\left[\boldsymbol{\tau}^\Pi\right]$ in Appendix B.2.2 and prove concentration of $\boldsymbol{\tau}^\Pi$ around $\mathbb{E}_\Pi\left[\boldsymbol{\tau}^\Pi\right]$ in Appendix B.2.3 . We then show that $\boldsymbol{\tau}^\Pi$ is approximately normal in Appendix B.2.4. We characterize $\mathbb{E}_\Pi[V^\Pi]$ in Appendix B.2.5 and show that V^Π concentrates around $\mathbb{E}_\Pi[V^\Pi]$ in Appendix B.2.6. Finally in Appendix B.2.7, we use this concentration to express the conditional CLT (lemma B.3) in a more convenient form.

B.2.1 Useful results on concentration for random permutations

We first provide some notation. Let $\mathfrak{P}^J$ be the set of permutations of $[J]$. Given two permutations $\pi_1, \pi_2 \in \mathfrak{P}^J$, let $\delta(\pi_1, \pi_2)$ be their convex distance:

$$\delta(\pi_1, \pi_2) = \sup_{\|a\|_2=1} \sum_{j=1}^{J} |a_j| \mathbb{1}\{\pi_1(j) \neq \pi_2(j)\}. \tag{B.12}$$

Moreover, given a set $S \subset \mathfrak{P}^J$ and a permutation $\pi \in \mathfrak{P}^J$, with some slight abuse of notation, we let $\delta(\pi, S) = \inf_{s\in S} \delta(\pi, s)$, i.e. the distance of π to $S \subset \mathfrak{P}^J$ is the distance to the nearest point in S. To establish concentration of $\boldsymbol{\tau}^{\Pi}$ around $\mathbb{E}_{\Pi}\left[\boldsymbol{\tau}^{\Pi}\right]$ (and similarly that the covariance V^{Π} concentrates around $\mathbb{E}_{\Pi}[V^{\Pi}]$), we will use an isoperimetric inequality for uniform random permutations, along with a well-known corollary. In particular, we will reduce the problem of establishing concentration for the conditional mean and variance to that of establishing concentration for suitable L-Lipschitz continuous functions of Π. Towards that goal, in what follows we let $X : \mathfrak{P}^J \to \mathbb{R}$ denote an L-Lipschitz continuous function with respect to the distance δ defined in eq. (B.12): there exists some $L > 0$ for which $\forall\, \pi_1, \pi_2 \in \mathfrak{P}^J$

$$|X(\pi_1) - X(\pi_2)| \leq L\delta(\pi_1, \pi_2). \tag{B.13}$$

Lemma B.4 (Talagrand [1995], Theorem 5.1)**.** *Let $\Pi : [J] \to [J]$ be a permutation chosen uniformly at random in $\mathfrak{P}^J$. Then for a set $S \subseteq \mathfrak{P}^J$,*

$$\mathbb{P}(\Pi \in S)\, \mathbb{E}\left[\exp\left\{\frac{\delta(\Pi, S)^2}{16}\right\}\right] \leq 1,$$

where we recall that the distance to the set S is defined as $\delta(\pi, S) = \inf_{s\in S} \delta(\pi, s)$.

Proof. See Talagrand [1995, Theorem 5.1]. □

Lemma B.4 has the following well-known corollary.

Corollary B.5 (Concentration for random permutations)**.** *Suppose that $X : \mathfrak{P}^J \to \mathbb{R}$ is L-Lispchitz continuous as per Equation (B.13). Let $\Pi \in \mathfrak{P}^J$ be chosen uniformly at random. Then, for $t > 0$,*

$$\mathbb{P}\{|X(\Pi) - \mathbb{E}[X(\Pi)]| > 8Lt\} \leq 2e^{-t^2/8}.$$

Proof. Let $\nu \in \mathbb{R}$ be the median of $X(\Pi)$ when $\Pi \sim \mathrm{Unif}(\mathfrak{P}^J)$, i.e.

$$\nu = \{\inf z \in \mathbb{R} \;:\; \mathbb{P}[X(\Pi) \leq z] \geq 1/2\}.$$

Let $S := \left\{\pi \in \mathfrak{P}^J \,\middle|\, X(\pi) \leq \nu\right\}$. By Markov's inequality, Lemma B.4, and the fact that $\mathbb{P}(\Pi \in S) \geq 1/2$,

$$\begin{aligned}
\mathbb{P}\left(L\delta(\Pi, S) \geq s\right) &= \mathbb{P}\left(\exp\left\{\frac{\delta(\Pi, S)^2}{16}\right\} \geq \exp\left\{\frac{s^2}{16L^2}\right\}\right) \\
&\leq \mathbb{E}\left[\exp\left\{\frac{\delta(\Pi, S)^2}{16}\right\}\right] \exp\left\{-\frac{s^2}{16L^2}\right\} \\
&\leq 2\exp\left\{-\frac{s^2}{16L^2}\right\}.
\end{aligned} \tag{B.14}$$

L-Lipschitz continuity of X with respect to δ implies $|X(\Pi) - \nu| \leq L\delta(\Pi, S)$. Using Equation (B.14) we then can bound the deviations of $X(\Pi)$ from its median:

$$\mathbb{P}(X(\Pi) - \nu \geq s) \leq \mathbb{P}(L\delta(\Pi, S) \geq s) \leq 2\exp\left\{-\frac{s^2}{16L^2}\right\}, \tag{B.15}$$

and symmetrically,

$$\mathbb{P}(X(\Pi) - \nu \leq -s) \leq \mathbb{P}(-L\delta(\Pi, S) \leq -s) \leq 2\exp\left\{-\frac{s^2}{16L^2}\right\}. \tag{B.16}$$

Finally, we transfer this to concentration around the mean of $X(\Pi)$. To avoid confusion, and with an exception to our general notation, let $\boldsymbol{\pi} \approx 3.14$ denote the universal constant.

$$\begin{aligned}\mathbb{E}[X(\Pi) - \nu] \leq \mathbb{E}\left[(X(\Pi) - \nu)\, 1\{X(\Pi) > \nu\}\right] &= \int_0^\infty \mathbb{P}\left(X(\Pi) - \nu > t\right) \mathrm{d}t \\ &\leq \int_0^\infty 2e^{-t^2/(16L^2)} \mathrm{d}t = \sqrt{16\boldsymbol{\pi}}L,\end{aligned} \tag{B.17}$$

where in the first equality we have used the fact that $(X(\Pi) - \nu)\, 1\{X(\Pi) > \nu\} \geq 0$ is a non-negative random variable (for which the tail probability formula of its expected value holds), and in the last inequality we have used Equation (B.15); the integral is computed by noting it coincides with that of a scaled Gaussian density. Symmetrically, $\mathbb{E}[\nu - X(\Pi)] \leq \sqrt{16\boldsymbol{\pi}}L$. The two combined yield an upper and lower bound on the mean $\mathbb{E}[X(\Pi)]$ in terms of the median and the Lipschitz constant:

$$\mathbb{E}[X(\Pi)] \leq \nu + \sqrt{16\boldsymbol{\pi}}L \quad \text{and} \quad \mathbb{E}[X(\Pi)] \geq \nu - \sqrt{16\boldsymbol{\pi}}L. \tag{B.18}$$

Using the lower bound on $\mathbb{E}[X(\Pi)]$ in Equation (B.18) we obtain

$$\begin{aligned}\mathbb{P}\{X(\Pi) - \mathbb{E}[X(\Pi)] > t\} &\leq \mathbb{P}\{X(\Pi) - (\nu - \sqrt{16\boldsymbol{\pi}}L) > t\} \\ &= \mathbb{P}\{X(\Pi) - \nu > t - \sqrt{16\boldsymbol{\pi}}L\},\end{aligned} \tag{B.19}$$

and now applying Equation (B.16)

$$\mathbb{P}\{X(\Pi) - \mathbb{E}[X(\Pi)] > t + \sqrt{16\boldsymbol{\pi}}L\} \leq 2\exp\left\{-\frac{t^2}{16L^2}\right\}. \tag{B.20}$$

Symmetrically, we use the upper bound on $\mathbb{E}[X(\Pi)]$ in Equation (B.18) to obtain

$$\mathbb{P}\{X(\Pi) - \mathbb{E}[X(\Pi)] < -t - \sqrt{16\boldsymbol{\pi}}L\} \leq 2\exp\left\{-\frac{t^2}{16L^2}\right\}. \tag{B.21}$$

Hence, combining Equations (B.20) and (B.21) via a union bound and choosing $t = \sqrt{16\boldsymbol{\pi}}Lz$,

$$\mathbb{P}\{|X(\Pi) - \mathbb{E}[X(\Pi)]| > (z+1)\sqrt{16\boldsymbol{\pi}}L\} \leq 4e^{-\boldsymbol{\pi}z^2} \leq 4e^{-z^2}.$$

Finally, note that for any $u = z + 1 \geq 0$, hence for $u \geq 2$, we may rewrite this as

$$\mathbb{P}\{|X(\Pi) - \mathbb{E}[X(\Pi)]| > u\sqrt{16\boldsymbol{\pi}}L\} \leq 4e^{-(u-1)^2} \leq 4e^{-u^2/4},$$

since $u^2/(u-1)^2 \leq 4$ for $u \geq 2$. Meanwhile for $0 < u \leq 2$ the bound is larger than 1, hence it holds trivially. Similarly, we simplify $4e^{-u^2/4} \vee 1 \leq 2e^{-u^2/8} \vee 1$ and note that $\sqrt{16\boldsymbol{\pi}} \leq 8$. □

Finally, we apply the result to our context. The following lemma allows us to show concentration of sums of potential outcomes under simple random sampling.

Lemma B.6. *For some $M > 0$, let $\boldsymbol{b} = [b_1, \ldots, b_J]^\top \in [-M, M]^J$ be a vector of scalars, and let $\mathcal{J}_\gamma$ be one of the sets $\{\Pi(1), \Pi(2), \ldots, \Pi(J_0)\}$ (if $\gamma \in \{\text{cc}, \text{ib}\}$) or $\{\Pi(J_0+1), \Pi(J_0+2), \ldots, \Pi(J)\}$ (if $\gamma \in \{\text{is}, \text{tr}\}$), so that $|\mathcal{J}_\gamma| = J_\gamma$. Put $X_{\boldsymbol{b}}(\Pi) = \sum_{j \in \mathcal{J}_\gamma} b_{\Pi(j)}$, for $\Pi \sim \text{Unif}(\mathfrak{P}^J)$. Then we have the bound*

$$\mathbb{P}\left\{|X_{\boldsymbol{b}}(\Pi) - \mathbb{E}\left[X_{\boldsymbol{b}}(\Pi)\right]| > (8M\sqrt{J_\gamma})t\right\} \leq 2e^{-t^2/8}. \tag{B.22}$$

Proof. Without loss of generality, we will consider the case that $\mathcal{J}_\gamma = \{\Pi(1), \Pi(2), \ldots, \Pi(J_0)\}$; the other case is symmetric. Put $a_j = J_0^{-1/2}$ for $1 \leq j \leq J_0$ and $a_j = 0$ for $j > J_0$, and note that $\|\mathbf{a}\|_2 = 1$. We will apply Corollary B.5 using the weights $\mathbf{a}$. For any two permutations $\pi_1, \pi_2 \in \mathfrak{P}^J$,

$$\begin{aligned}
|X_{\boldsymbol{b}}(\pi_1) - X_{\boldsymbol{b}}(\pi_2)| = \left|\sum_{j=1}^{J_0} b_{\pi_1(j)} - b_{\pi_2(j)}\right| &\leq \sum_{j=1}^{J_0} |b_{\pi_2(j)}| \mathbb{1}\{\pi_1(j) \neq \pi_2(j)\} \\
&= \sum_{j=1}^{J} \sqrt{J_0}|a_j||b_{\pi_2(j)}| \mathbb{1}\{\pi_1(j) \neq \pi_2(j)\} \\
&\leq \sqrt{J_0}M \sum_{j=1}^{J} |a_j| \mathbb{1}\{\pi_1(j) \neq \pi_2(j)\} \\
&\leq \sqrt{J_0}M\delta(\pi_1, \pi_2).
\end{aligned}$$

These steps follow by the triangle inequality, by our choice of a_j, by the fact $|b_j| \leq M$, and by the definition of δ given in Equation (B.12). The inequality (B.22) then follows by Corollary B.5, as we have just shown that $X_{\boldsymbol{b}}$ is L-Lipschitz with respect to the convex distance $\delta(\pi_1, \pi_2)$, with $L = \sqrt{J_0}M$. □

Finally, we state a technical lemma which will help us apply lemma B.6 to expressions which depend on potential outcomes $y_{ij}(\gamma)$ for multiple types γ.

Lemma B.7. *Under assumption 4.6(a), for any $\gamma \in \{\text{cc}, \text{ib}, \text{is}, \text{tr}\}$ and any $b_1(\gamma), \ldots, b_J(\gamma) \in [-M, M]$, there exists a collection of numbers $\tilde{b}_1(\gamma), \ldots, \tilde{b}_{J_0}(\gamma)$ with absolute value at most $2C_1M$, such that for any $\Pi \in \mathfrak{P}^J$,*

$$\frac{1}{J_\gamma} \sum_{j \in \mathcal{J}_\gamma^\Pi} b_j(\gamma) = \frac{1}{J_0} \sum_{j=1}^{J_0} \tilde{b}_{\Pi(j)}(\gamma).$$

Note, the left-hand side is a sum over J_γ terms.The right-hand side is a sum over J_0 terms, irrespective of γ.

Proof. We proceed by cases, first considering $\gamma \in \{\text{cc}, \text{ib}\}$ and then $\gamma \in \{\text{is}, \text{tr}\}$. Let $\Pi \in \mathfrak{P}^J$ be arbitrary. Recall that by construction, we have $\mathcal{J}_\gamma^\Pi = \{\Pi(1), \ldots, \Pi(J_0)\}$ if $\gamma \in \{\text{cc}, \text{ib}\}$ and $\mathcal{J}_\gamma^\Pi = \{\Pi(J_0+1), \ldots, \Pi(J)\}$ if $\gamma \in \{\text{is}, \text{tr}\}$. Then, if $\gamma \in \{\text{cc}, \text{ib}\}$:

$$\frac{1}{J_\gamma} \sum_{j \in \mathcal{J}_\gamma^\Pi} b_j(\gamma) = \frac{1}{J_0} \sum_{j=1}^{J_0} b_{\Pi(j)}(\gamma).$$

The claim then directly holds by taking $\tilde{b}_{\Pi(j)}(\gamma) = b_{\Pi(j)}(\gamma)$ for all j.

On the other hand, if $\gamma \in \{\text{is}, \text{tr}\}$, letting $\bar{b}(\gamma) := \sum_{j=1}^J b_j(\gamma)/J$, we have

$$\begin{aligned}\frac{1}{J_\gamma} \sum_{j \in \mathcal{J}_\gamma^\Pi} b_j(\gamma) &= \frac{1}{J_1} \sum_{j=J_0+1}^{J} b_{\Pi(j)}(\gamma) = \frac{1}{J_1} \sum_{j=1}^{J} b_{\Pi(j)}(\gamma) \left[1 - \mathbb{1}(j \leq J_0)\right] = \frac{J}{J_1}\bar{b}(\gamma) - \frac{1}{J_1} \sum_{j=1}^{J_0} b_{\Pi(j)}(\gamma) \\ &= \frac{1}{J_0} \sum_{j=1}^{J_0} \left(\frac{J}{J_1}\bar{b}(\gamma) - \frac{J_0}{J_1} b_{\Pi(j)}(\gamma) \right).\end{aligned}$$

We then take $\tilde{b}_j(\gamma) = (J/J_1)\bar{b}(\gamma) - (J_0/J_1) b_{\Pi(j)}(\gamma)$; in either case, $|\tilde{b}_j(\gamma)| \leq 2C_1 M$. □

B.2.2 Computing the expectation of the mean

Recall the definition of $\boldsymbol{\tau}^\pi$, given in Equation (B.5). In what follows, with a slight abuse of notation, we let $\tau^\pi(\gamma)$ be the entry of $\boldsymbol{\tau}^\pi$ referring to type $\gamma \in \{\text{cc}, \text{ib}, \text{is}, \text{tr}\}$. We let $\tau^\Pi(\gamma)$ represent the same quantity, now indexed by a random $\Pi \sim \text{Unif}(\mathfrak{P}^J)$.

We note that the expectation of $\tau^\Pi(\gamma)$ coincides with the population mean $\overline{\overline{y}}_\gamma$:

$$\begin{aligned}\mathbb{E}_\Pi\left[\tau^\Pi(\gamma)\right] &= \mathbb{E}_\Pi\left[\frac{1}{I}\sum_{i=1}^{I} \frac{1}{J_\gamma} \sum_{j=1}^{J_\gamma} y_{i,\Pi(j)}(\gamma)\right] = \frac{1}{IJ_\gamma} \sum_{i=1}^{I} \mathbb{E}_\Pi\left[\sum_{j=1}^{J_\gamma} y_{i,\Pi(j)}(\gamma)\right] \\ &= \frac{1}{IJ_\gamma} \sum_{i=1}^{I} \frac{\binom{J-1}{J_\gamma - 1}}{\binom{J}{J_\gamma}} \sum_{j=1}^{J} y_{i,j}(\gamma) = \frac{1}{IJ_\gamma} \sum_{i=1}^{I} \frac{J_\gamma}{J} \sum_{j=1}^{J} y_{i,j}(\gamma) = \overline{\overline{y}}_\gamma.\end{aligned}$$

By linearity of the expectation operator, for any $\vec{\boldsymbol{\beta}}$ it also holds that

$$\mathbb{E}_\Pi\left[\tau^\Pi(\vec{\boldsymbol{\beta}})\right] = \tau(\vec{\boldsymbol{\beta}}). \tag{B.23}$$

B.2.3 Concentration of the mean

We now use lemmas B.6 and B.7 to show concentration of $\tau^\Pi(\vec{\boldsymbol{\beta}})$ around its expectation.

Lemma B.8. *Let Π be a uniform random permutation of $[J]$. Then under assumption 4.6, which imposes a balanced experiment with bounded potential outcomes, it holds*

$$\mathbb{P}\left(|\tau^\Pi(\vec{\boldsymbol{\beta}}) - \tau(\vec{\boldsymbol{\beta}})| > \left[32 J_0^{-1/2} C_2 C_1\right] \|\vec{\boldsymbol{\beta}}\|_2 t\right) \leq 2e^{-t^2/8}. \tag{B.24}$$

Proof. Equation (B.24) is a bound around deviations of $\tau^{\Pi}(\vec{\boldsymbol{\beta}})$ around its mean, since $\mathbb{E}[\tau^{\Pi}(\vec{\boldsymbol{\beta}})] = \tau(\vec{\boldsymbol{\beta}})$ as per eq. (B.23). To show eq. (B.24) we will first prove that $\tau^{\Pi}(\vec{\boldsymbol{\beta}}) = J_0^{-1}\sum_{j=1}^{J_0} b_{\Pi(j)}$ for some suitably bounded numbers $(b_1, \dots, b_J)$ via lemma B.7, and then conclude using lemma B.6. From the definition of $\tau^{\Pi}(\vec{\boldsymbol{\beta}})$,

$$\begin{aligned}\tau^{\Pi}(\vec{\boldsymbol{\beta}}) &= \sum_{\gamma\in\{\text{cc},\text{ib},\text{is},\text{tr}\}} \beta_\gamma \frac{1}{I}\frac{1}{J_\gamma}\sum_{i=1}^{I}\sum_{j\in\mathcal{J}_\gamma^{\Pi}} y_{i,j} = \sum_{\gamma\in\{\text{cc},\text{ib},\text{is},\text{tr}\}} \beta_\gamma \frac{1}{J_\gamma}\sum_{j\in\mathcal{J}_\gamma^{\Pi}}\left(\frac{1}{I}\sum_{i=1}^{I} y_{i,j}(\gamma)\right) \\ &= \sum_{\gamma\in\{\text{cc},\text{ib},\text{is},\text{tr}\}} \beta_\gamma \frac{1}{J_\gamma}\sum_{j\in\mathcal{J}_\gamma^{\Pi}} \overline{y}_j^{\mathrm{S}}(\gamma).\end{aligned}$$

Because of the boundedness assumption (b) in assumption 4.6, $|\overline{y}_j^{\mathrm{S}}(\gamma)| \le C_2$ for each $j \in [J]$. By lemma B.7, we may find numbers $|\tilde{b}_j(\gamma)| \le 2C_1C_2$ which allow us to rewrite $\tau^{\Pi}(\vec{\boldsymbol{\beta}})$ as

$$\tau^{\Pi}(\vec{\boldsymbol{\beta}}) = \sum_{\gamma\in\{\text{cc},\text{ib},\text{is},\text{tr}\}} \beta_\gamma \frac{1}{J_0}\sum_{j=1}^{J_0}\tilde{b}_{\Pi(j)}(\gamma) = \frac{1}{J_0}\sum_{j=1}^{J_0}\left\{\sum_{\gamma\in\{\text{cc},\text{ib},\text{is},\text{tr}\}}\beta_\gamma\tilde{b}_{\Pi(j)}(\gamma)\right\}.$$

By Hölder's inequality and the fact that $\|v\|_1 \le \sqrt{4}\|v\|_2$ for $v \in \mathbb{R}^4$, we can further bound the bracketed terms in the equation above as

$$\left|\sum_{\gamma\in\{\text{cc},\text{ib},\text{is},\text{tr}\}}\beta_\gamma\tilde{b}_j(\gamma)\right| \le \|\vec{\boldsymbol{\beta}}\|_1 \max_\gamma |\tilde{b}_j(\gamma)| \le \|\vec{\boldsymbol{\beta}}\|_1 2C_1C_2 \le 4C_1C_2\|\vec{\boldsymbol{\beta}}\|_2.$$

Then by considering the bounded vector $\boldsymbol{b} = [b_1, \dots, b_J]$ in which each $b_j = \sum_{\gamma\in\{\text{cc},\text{ib},\text{is},\text{tr}\}}\beta_\gamma\tilde{b}_j(\gamma) \le 4C_1C_2\|\vec{\boldsymbol{\beta}}\|_2$, we can apply lemma B.6 to $\tau^{\Pi}(\vec{\boldsymbol{\beta}})$ (in turn, eq. (B.24) follows):

$$\tau^{\Pi}(\vec{\boldsymbol{\beta}}) = \frac{1}{J_0}\sum_{j=1}^{J_0} b_j \le 4C_1C_2\|\vec{\boldsymbol{\beta}}\|_2.$$

□

B.2.4 A CLT for the conditional mean

Finally, we note an unconditional normal approximation for $\boldsymbol{\tau}^{\Pi} = \mathbb{E}[\boldsymbol{\tau} \mid \Pi]$ which mirrors Lemma B.3 above. It follows from the observation that $\boldsymbol{\tau}^{\Pi}$ is the standard mean estimator corresponding to a completely randomized experiment in which J_1 out of J units are treated, with the following vector-valued potential outcomes:

$$\overline{\boldsymbol{y}}_j^{\mathrm{S}}(0) = \begin{pmatrix}\overline{y}_j^{\mathrm{S}}(\text{cc}) & 0 & \overline{y}_j^{\mathrm{S}}(\text{is}) & 0\end{pmatrix}^\top; \quad \overline{\boldsymbol{y}}_j^{\mathrm{S}}(1) = \begin{pmatrix}0 & \overline{y}_j^{\mathrm{S}}(\text{ib}) & 0 & \overline{y}_j^{\mathrm{S}}(\text{tr})\end{pmatrix}^\top.$$

Lemma B.9. *Assuming bounded potential outcomes as per assumption 4.6(b):*

$$\sup_{t\in\mathbb{R}} \left| \mathbb{P}\left\{ \frac{\tau^{\Pi}(\vec{\boldsymbol{\beta}}) - \tau(\vec{\boldsymbol{\beta}})}{\sqrt{\mathrm{Var}\{\tau^{\Pi}(\vec{\boldsymbol{\beta}})\}}} \le t \right\} - \Phi(t) \right| \le \frac{CC_2}{\min\{J_0, J_1\}} \frac{\|\vec{\boldsymbol{\beta}}\|_2}{\sqrt{\mathrm{Var}\{\tau^{\Pi}(\vec{\boldsymbol{\beta}})\}}}. \tag{B.25}$$

Proof. Identical to Lemma B.3. □

In general, the fluctuations of the conditional mean $\tau^{\Pi}(\vec{\boldsymbol{\beta}})$ may not be negligible; this may occur, e.g., if the number of sellers is small. Lemma B.9 shows that $\tau^{\Pi}(\vec{\boldsymbol{\beta}})$ is itself approximately Gaussian, so that we can derive the CLT for $\hat{\tau}(\vec{\boldsymbol{\beta}})$ despite this possibility. Our use of lemma B.9 in this capacity was initially suggested by Sudijono et al. [2025].

B.2.5 Computing the expectation of the covariance

Mirroring Appendix B.2.2 we now compute $\mathbb{E}_{\Pi}\left[\mathrm{Cov}\left\{\widehat{\boldsymbol{\tau}}^{\Pi}\right\}\right]$ — the *expectation* of $\mathrm{Cov}\left\{\widehat{\boldsymbol{\tau}}^{\Pi}\right\}$ over the uniform measure on the space of permutations $\mathfrak{P}^J$ of $[J]$.

Lemma B.10. *For* $\gamma, \gamma' \in \{\mathrm{cc}, \mathrm{ib}, \mathrm{is}, \mathrm{tr}\}$ *and* $i, i' \in [I]$ *define*

$$\rho_{i,i'}^{\gamma,\gamma'} := \mathbb{E}_{\Pi}\left[\overline{y}^{\mathrm{B}}_{i,\mathcal{J}^{\Pi}_{\gamma}}(\gamma)\overline{y}^{\mathrm{B}}_{i',\mathcal{J}^{\Pi}_{\gamma'}}(\gamma')\right].$$

Now, we characterize the entry $\mathrm{Cov}\left\{\widehat{\boldsymbol{\tau}}^{\Pi}\right\}_{\gamma,\gamma'}$ *associated with types* γ, γ'*:*

$$\mathbb{E}_{\Pi}\left[\mathrm{Cov}\left\{\widehat{\boldsymbol{\tau}}^{\Pi}\right\}\right]_{\gamma,\gamma'} = \kappa_{\gamma,\gamma'}\left\{\sum_{i=1}^{I}\rho_{i,i}^{\gamma,\gamma'} - \frac{1}{I}\sum_{i=1}^{I}\sum_{i'=1}^{I}\rho_{i,i'}^{\gamma,\gamma'}\right\},$$

where

$$\kappa_{\gamma,\gamma'} = \begin{cases} 1 & \text{if } \gamma = \gamma' \\ \frac{I-I_\gamma}{I_\gamma}\frac{1}{I(I-1)} & \text{if } \gamma \neq \gamma' \wedge \mathcal{J}^{\Pi}_{\gamma} = \mathcal{J}^{\Pi}_{\gamma'} \text{ (e.g., } \gamma = \mathrm{cc}, \gamma' = \mathrm{is}) \\ -\frac{I-I_\gamma}{I_\gamma}\frac{1}{I(I-1)} & \text{if } \gamma \neq \gamma' \wedge \mathcal{J}^{\Pi}_{\gamma} \neq \mathcal{J}^{\Pi}_{\gamma'} \text{ (e.g., } \gamma = \mathrm{cc}, \gamma' = \mathrm{ib}). \end{cases}$$

Proof. The main contribution of this proof is just to make the coeficients $\rho_{i,i'}^{\gamma,\gamma'}$ explicit; to do so, we specialize them into three different cases: (i) $\gamma = \gamma'$, (ii) $\gamma \neq \gamma'$ and $\mathcal{J}^{\Pi}_{\gamma} = \mathcal{J}^{\Pi}_{\gamma'}$, and (iii) $\gamma \neq \gamma'$ and $\mathcal{J}^{\Pi}_{\gamma} \neq \mathcal{J}^{\Pi}_{\gamma'}$.

(i) $\gamma = \gamma'$. We have

$$\rho_{i,i'}^{\gamma,\gamma} = \mathbb{E}\left[\overline{y}^{\mathrm{B}}_{i,\mathcal{J}^{\Pi}_{\gamma}}(\gamma)\overline{y}^{\mathrm{B}}_{i',\mathcal{J}^{\Pi}_{\gamma}}(\gamma)\right] = \left(\frac{1}{J_\gamma}\right)^2 \mathbb{E}_{\Pi}\left[\sum_{j,j'\in\mathcal{J}^{\Pi}_{\gamma}} y_{i,j}(\gamma)y_{i',j'}(\gamma)\right].$$

Observing that among the total $\binom{J}{J_\gamma}$ selection of indices $\mathcal{J}_\gamma^\Pi$, exactly $\binom{J-1}{J_\gamma-1}$ of these index sets contain index j and exactly $\binom{J-2}{J_\gamma-2}$ of these index sets contain the pair (j,j') for $j \neq j$,

$$\begin{aligned}\rho_{i,i'}^{\gamma,\gamma} &= \left(\frac{1}{J_\gamma}\right)^2 \left[\frac{J_\gamma}{J}\sum_{j=1}^{J} y_{i,j}(\gamma)^2 + \frac{J_\gamma(J_\gamma-1)}{J(J-1)}\sum_{j=1}^{J}\sum_{j'\neq j} y_{i,j}(\gamma)y_{i,j'}(\gamma)\right] \\ &= \frac{1}{J_\gamma J}\left[\sum_{j=1}^{J} y_{i,j}(\gamma)y_{i',j}(\gamma) + \frac{J_\gamma-1}{J-1}\sum_{j=1}^{J}\sum_{j'\neq j} y_{i,j}(\gamma)y_{i',j'}(\gamma)\right]. \end{aligned} \tag{B.26}$$

(ii) $\gamma \neq \gamma'$ and $\mathcal{J}_\gamma^\Pi = \mathcal{J}_{\gamma'}^\Pi$. The derivation is analogous to (i), and just requires swapping the argument of the second column-wise mean $\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_\gamma^\Pi}(\gamma)$ with γ', i.e. consider $\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_\gamma^\Pi}(\gamma')$; because $\mathcal{J}_\gamma^\Pi = \mathcal{J}_{\gamma'}^\Pi$, this change is only in the argument and not in the set indexing this mean:

$$\rho_{i,i'}^{\gamma,\gamma'} = \mathbb{E}\left[\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_\gamma^\Pi}(\gamma)\overline{y}^{\mathrm{B}}_{i',\mathcal{J}_\gamma^\Pi}(\gamma')\right] = \left(\frac{1}{J_\gamma}\right)^2 \mathbb{E}_\Pi\left[\sum_{j,j'\in\mathcal{J}_\gamma^\Pi} y_{i,j}(\gamma)y_{i',j'}(\gamma')\right],$$

and now observing that among the total $\binom{J}{J_\gamma}$ selection of indices $\mathcal{J}_\gamma^\Pi$, exactly $\binom{J-1}{J_\gamma-1}$ of these index sets contain index j and exactly $\binom{J-2}{J_\gamma-2}$ of these index sets contain the pair (j,j') for $j \neq j$,

$$\begin{aligned}\rho_{i,i'}^{\gamma,\gamma'} &= \left(\frac{1}{J_\gamma}\right)^2 \left[\frac{J_\gamma}{J}\sum_{j=1}^{J} y_{i,j}(\gamma)y_{i',j}(\gamma') + \frac{J_\gamma(J_\gamma-1)}{J(J-1)}\sum_{j=1}^{J}\sum_{j'\neq j} y_{i,j}(\gamma)y_{i',j'}(\gamma')\right] \\ &= \frac{1}{J_\gamma J}\left[\sum_{j=1}^{J} y_{i,j}(\gamma)y_{i',j}(\gamma') + \frac{J_\gamma-1}{J-1}\sum_{j=1}^{J}\sum_{j'\neq j} y_{i,j}(\gamma)y_{i',j'}(\gamma')\right]. \end{aligned} \tag{B.27}$$

(iii) $\gamma \neq \gamma'$ and $\mathcal{J}_\gamma \neq \mathcal{J}_{\gamma'}$. The derivation is analogous to (i), and just requires swapping the argument and index set of the second column-wise mean $\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_\gamma^\Pi}(\gamma)$ with γ', i.e. consider $\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_{\gamma'}^\Pi}(\gamma')$; because $\mathcal{J}_\gamma^\Pi \neq \mathcal{J}_{\gamma'}^\Pi$, this change is affecting both the argument and the index set defining this mean:

$$\rho_{i,i'}^{\gamma,\gamma'} = \mathbb{E}\left[\overline{y}^{\mathrm{B}}_{i,\mathcal{J}_\gamma^\Pi}(\gamma)\overline{y}^{\mathrm{B}}_{i',\mathcal{J}_{\gamma'}^\Pi}(\gamma')\right] = \left(\frac{1}{J_\gamma J_{\gamma'}}\right) \mathbb{E}_\Pi\left[\sum_{j\in\mathcal{J}_\gamma}\sum_{j'\in\mathcal{J}_{\gamma'}} y_{i,j}(\gamma)y_{i',j'}(\gamma')\right]$$

and now observing that among the total $\binom{J}{J_\gamma}$ selection of indices, there are exactly $\binom{J-2}{J_\gamma-2}$ selections such that $j, j' \in \mathcal{J}_\gamma^\Pi$, and exactly $\binom{J-2}{J-J_\gamma-2}$ selection of indices such that $j, j' \notin \mathcal{J}_\gamma^\Pi$, then there are exactly $\binom{J}{J_0} - \binom{J-2}{J_0-2} - \binom{J-2}{J_1-2}$ of the total $\binom{J}{J_0}$ such that j and j' do not both belong to $\mathcal{J}_\gamma^\Pi$

$$\rho_{i,i'}^{\gamma,\gamma'} = \left(\frac{1}{J_\gamma J_{\gamma'}}\right)\left(1 - \frac{J_0(J_0-1)}{J(J-1)} - \frac{J_1(J_1-1)}{J(J-1)}\right)\sum_{j=1}^{J}\sum_{j'\neq j} y_{i,j}(\gamma) y_{i',j'}(\gamma'). \tag{B.28}$$

□

B.2.6 Concentration of the covariance

Recall the definition of the covariance matrix $S^2_{\hat{\boldsymbol{\tau}}^\pi}$ given in Equation (B.8):

$$S^2_{\hat{\boldsymbol{\tau}}^\pi} := \frac{1}{I-1}\sum_{i=1}^{I}\{\boldsymbol{\tau}_i^\pi - \boldsymbol{\tau}^\pi\}\{\boldsymbol{\tau}_i^\pi - \boldsymbol{\tau}^\pi\}^\top,$$

and let $S_\pi(\gamma,\gamma') := S^2_{\hat{\boldsymbol{\tau}}^\pi}(\gamma,\gamma')$ be its entry associated with types γ, γ', where we drop the dependence on $\hat{\boldsymbol{\tau}}$ for the ease of notation. Moreover, we have by eq. (B.9),

$$\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})|\Pi = \pi\} = \mathrm{Var}\{\hat{\tau}^\pi(\vec{\boldsymbol{\beta}})\} = \sum_{\gamma,\gamma'\in\{\mathrm{cc},\mathrm{ib},\mathrm{is},\mathrm{tr}\}} \beta_\gamma\beta_{\gamma'}\frac{A(\gamma,\gamma')}{I(I-1)}\sum_{i=1}^{I}\dot{\tau}_i^\pi(\gamma)\dot{\tau}_i^\pi(\gamma'), \tag{B.29}$$

with $\dot{\tau}_i^\pi(\gamma)$ defined as an entry in the vector $\dot{\boldsymbol{\tau}}_i^\pi$ of eq. (B.6), and

$$A = \begin{pmatrix} A(\mathrm{cc},\mathrm{cc}) & A(\mathrm{cc},\mathrm{ib}) & A(\mathrm{cc},\mathrm{is}) & A(\mathrm{cc},\mathrm{tr}) \\ A(\mathrm{ib},\mathrm{cc}) & A(\mathrm{ib},\mathrm{ib}) & A(\mathrm{ib},\mathrm{is}) & A(\mathrm{ib},\mathrm{tr}) \\ A(\mathrm{is},\mathrm{cc}) & A(\mathrm{is},\mathrm{ib}) & A(\mathrm{is},\mathrm{is}) & A(\mathrm{is},\mathrm{tr}) \\ A(\mathrm{tr},\mathrm{cc}) & A(\mathrm{tr},\mathrm{ib}) & A(\mathrm{tr},\mathrm{is}) & A(\mathrm{tr},\mathrm{tr}) \end{pmatrix} = \{A(\gamma,\gamma')\}_{\gamma,\gamma'\in\{\mathrm{cc},\mathrm{ib},\mathrm{is},\mathrm{tr}\}} \in \mathbb{R}^{4\times 4} \tag{B.30}$$

a matrix with entries $A(\gamma,\gamma')$ indexed by types γ, γ', similar to $S^2_{\hat{\boldsymbol{\tau}}^\pi}$. Explicitly, they are

$$\dot{\tau}_i^\pi(\gamma) = \frac{1}{J_\gamma}\sum_{j\in\mathcal{J}_\gamma^\pi}[y_{i,j}(\gamma) - \overline{\overline{y}}_{\bullet,\pi}(\gamma)] \qquad \text{and} \qquad A = \begin{pmatrix} I_1/I_0 & -1 & I_1/I_0 & -1 \\ -1 & I_1/I_0 & -1 & I_1/I_0 \\ I_1/I_0 & -1 & I_1/I_0 & -1 \\ -1 & I_1/I_0 & -1 & I_1/I_0 \end{pmatrix}.$$

Our arguments will also make use of certain facts about Orlicz norms, which are collected in the following definition. We refer to Vershynin [2018, Chapter 2] for proofs.

Definition B.11. *For a real random variable X, and $k \geq 1$, we define $\|X\|_{\psi_p}$ as the smallest $t > 0$ such that $\mathbb{E}[\exp\{(X/t)^p\}] \leq 2$, provided that such a t exists (otherwise, it is ∞). For X, Y real, Borel random variables:*

1. $\|aX + bY\|_{\psi_k} \leq a\|X\|_{\psi_k} + b\|Y\|_{\psi_k}$ *for* $k = 1, 2$;
2. $\|X - \mathbb{E}[X]\|_{\psi_k} \leq 2\|X\|_{\psi_k}$;
3. *if* $\mathbb{P}(|X| > t) \leq 2e^{-t^2/s_0^2}$ *then* $\|X\|_{\psi_2} \leq Cs_0$
4. (a) *if* $\|X\|_{\psi_2} \leq s_1$ *then* $\mathbb{P}(|X| > t) \leq 2e^{-t^2/Cs_1^2}$, *so* $\mathbb{P}\{|X| \leq s_1\sqrt{C\log(2/\eta)}\} \geq 1 - \eta$;

(b) *if* $\|X\|_{\psi_1} \le s_2$ *then* $\mathbb{P}(|X| > t) \le 2e^{-t/Cs_2}$, *so* $\mathbb{P}\{|X| \le Cs_2 \log(2/\eta)\} \ge 1-\eta$,

5. $\|X^2\|_{\psi_1} \le \|X\|_{\psi_2}^2$.

Lemma B.12. *Under assumption 4.6, for a sufficiently large universal constant C, we have with probability at least $1-\eta$*

$$\left|\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})|\Pi\} - \mathbb{E}\left[\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})|\Pi\}\right]\right| \le \frac{CC_1^4C_2^2\|\vec{\boldsymbol{\beta}}\|_2^2}{J(I-1)}\log(4/\eta) + \frac{CC_1^2C_2\|\vec{\boldsymbol{\beta}}\|_2}{\sqrt{J(I-1)}}\sqrt{\mathbb{E}[\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})|\Pi\}]\log(4/\eta)}.$$

Proof. Throughout the proof, let C be a sufficiently large universal constant. Using Lemma B.7 and the fact that potential outcomes are bounded as per assumption 4.6 (b), we can rewrite

$$\dot{\tau}_i^{\Pi}(\gamma) = \frac{1}{J_\gamma}\sum_{j\in\mathcal{J}_\gamma^{\Pi}}[y_{i,j}(\gamma) - \overline{\overline{y}}_{\bullet,\Pi}(\gamma)] = \frac{1}{J_0}\sum_{j=1}^{J_0}\tilde{b}_{i,\Pi(j)}(\gamma),$$

for $|\tilde{b}_{i,j}(\gamma)| \le 4C_1C_2$, since $|y_{i,j}(\gamma) - \overline{\overline{y}}_{\bullet,\Pi}(\gamma)| \le 2C_2$. By using eq. (B.29), which allows us to express $\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})|\Pi = \pi\}$ as a sum over $\dot{\tau}_i^{\Pi}(\gamma)$, the decomposition above leads us to

$$\mathrm{Var}\{\hat{\tau}^{\Pi}(\vec{\boldsymbol{\beta}})\} = \sum_{\gamma,\gamma'\in\{\mathrm{cc},\mathrm{ib},\mathrm{is},\mathrm{tr}\}}\beta_\gamma\beta_{\gamma'}\frac{A(\gamma,\gamma')}{I(I-1)}\sum_{i=1}^{I}\left(\frac{1}{J_0}\sum_{j=1}^{J_0}\tilde{b}_{i,\Pi(j)}(\gamma)\right)\left(\frac{1}{J_0}\sum_{j=1}^{J_0}\tilde{b}_{i,\Pi(j)}(\gamma')\right).$$

To re-write this as a sum of squares, we use the eigenvector decomposition for the matrix A in eq. (B.30): $A = 2(I_1/I_0 - 1)uu^\top + 2(I_1/I_0+1)vv^\top$ for $u = (1,1,1,1)^\top$ and $v = (-1,1,-1,1)^\top$. Therefore we can write the (γ,γ')-th entry of A as $A(\gamma,\gamma') = A_1(\gamma)A_1(\gamma') + A_2(\gamma)A_2(\gamma')$ for $|A_1(\gamma)|, |A_2(\gamma)| \le \sqrt{2(C_1+1)}$. Thus, $\mathrm{Var}\{\hat{\tau}^{\Pi}(\vec{\boldsymbol{\beta}})\} = V_1 + V_2$ where for $k \in \{1,2\}$

$$\begin{aligned} V_k &= \sum_{\gamma,\gamma'\in\{\mathrm{cc},\mathrm{ib},\mathrm{is},\mathrm{tr}\}}\beta_\gamma\beta_{\gamma'}\frac{A_k(\gamma)A_k(\gamma')}{I(I-1)}\sum_{i=1}^{I}\left(\frac{1}{J_0}\sum_{j=1}^{J_0}\tilde{b}_{i,\Pi(j)}(\gamma)\right)\left(\frac{1}{J_0}\sum_{j=1}^{J_0}\tilde{b}_{i,\Pi(j)}(\gamma')\right) \\ &= \frac{1}{I(I-1)}\sum_{i=1}^{I}\left(\frac{1}{J_0}\sum_{j=1}^{J_0}\sum_{\gamma}\beta_\gamma A_k(\gamma)\tilde{b}_{i,\Pi(j)}(\gamma)\right)\left(\frac{1}{J_0}\sum_{j=1}^{J_0}\sum_{\gamma'}\beta_{\gamma'}A_k(\gamma')\tilde{b}_{i,\Pi(j)}(\gamma')\right). \end{aligned}$$

Hence, defining $X_{i,k} := \frac{1}{J_0}\sum_{j=1}^{J_0}\left(\sum_\gamma \beta_\gamma A_k(\gamma)\tilde{b}_{i,\Pi(j)}(\gamma)\right)$, $V_k = \frac{1}{I(I-1)}\sum_{i=1}^{I}X_{i,k}^2$. Next, we can use Cauchy-Schwarz and our previous bounds on A_k and $\tilde{b}_{i,j}$ to bound these terms constituting $X_{i,k}$s in parentheses as $|\sum_\gamma \beta_\gamma A_k(\gamma)\tilde{b}_{i,j}(\gamma)| \le C\|\vec{\boldsymbol{\beta}}\|_2 C_1^{3/2}C_2$ for $k = 1,2$. This gives us the trivial bound $|X_{i,k}| \le \sqrt{32}\|\vec{\boldsymbol{\beta}}\|_2(C_1+1)^{3/2}C_2$, so

$$\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})|\Pi\} = \sum_{k\in\{1,2\}}\frac{1}{I(I-1)}\sum_{i=1}^{I}X_{i,k}^2 \le \frac{64\|\vec{\boldsymbol{\beta}}\|_2^2(C_1+1)^3C_2^2}{I-1}; \tag{B.31}$$

clearly, the right-hand side of eq. (B.31) also bounds $\mathbb{E}[\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})|\Pi\}]$.

Thus, applying Lemma B.6 with $b_{i,j} = \sum_\gamma \beta_\gamma A_k(\gamma)\tilde{b}_{i,j}(\gamma)$, we find that for a large enough universal constant C, we have for all $i \in [I]$ and $k = 1, 2$:

$$\|X_{i,k} - \mathbb{E}[X_{i,k}]\|_{\psi_2} \leq C C_1^{3/2} C_2 \|\vec{\boldsymbol{\beta}}\|_2 J_0^{-1/2}. \tag{B.32}$$

Finally, by linearity of expectation and the identity

$$X^2 - \mathbb{E}[X^2] = \{(X - \mathbb{E}[X])^2 - \mathbb{E}[(X - \mathbb{E}[X])^2]\} + 2(X - \mathbb{E}[X])\,\mathbb{E}[X]$$

which is seen by expanding $X^2 = \{\mathbb{E}[X] + (X - \mathbb{E}[X])\}^2$ and its expectation, we have

$$\begin{aligned}
&|\,\mathrm{Var}\{\hat{\tau}^{\Pi}(\vec{\boldsymbol{\beta}})\} - \mathbb{E}[\mathrm{Var}\{\hat{\tau}^{\Pi}(\vec{\boldsymbol{\beta}})\}]| \\
&= |V_1 + V_2 - \mathbb{E}[V_1 + V_2]| \\
&= \left| \frac{1}{I(I-1)} \sum_{i=1}^{I} \sum_{k \in \{1,2\}} (X_{i,k} - \mathbb{E}[X_{i,k}])^2 - \mathbb{E}[(X_{i,k} - \mathbb{E}[X_{i,k}])^2] + 2(X_{i,k} - \mathbb{E}[X_{i,k}])\,\mathbb{E}[X_{i,k}] \right|
\end{aligned}$$

$$\leq \underbrace{\left| \frac{1}{I(I-1)} \sum_{i=1}^{I} \sum_{k\in\{1,2\}} (X_{i,k} - \mathbb{E}[X_{i,k}])^2 - \mathbb{E}[(X_{i,k} - \mathbb{E}[X_{i,k}])^2] \right|}_{s_1}$$

$$+ \underbrace{\left| \frac{1}{I(I-1)} \sum_{i=1}^{I} \sum_{k\in\{1,2\}} 2(X_{i,k} - \mathbb{E}[X_{i,k}])\, \mathbb{E}[X_{i,k}] \right|}_{s_2}$$

For the first summand s_1, we have by the triangle inequality

$$\begin{aligned}
\|s_1\|_{\psi_1} &= \left| \frac{1}{I(I-1)} \sum_{i=1}^{I} \sum_{k\in\{1,2\}} (X_{i,k} - \mathbb{E}[X_{i,k}])^2 - \mathbb{E}[(X_{i,k} - \mathbb{E}[X_{i,k}])^2] \right| \\
&\leq \frac{1}{I(I-1)} \sum_{i=1}^{I} \sum_{k\in\{1,2\}} \|(X_{i,k} - \mathbb{E}[X_{i,k}])^2 - \mathbb{E}[(X_{i,k} - \mathbb{E}[X_{i,k}])^2]\|_{\psi_1}.
\end{aligned}$$

By definition B.11 (ii) and (vi), $\|U^2 - \mathbb{E}[U^2]\|_{\psi_1} \leq 2\|U^2\|_{\psi_1} \leq 2\|U\|_{\psi_2}^2$, we can last apply eq. (B.32) to obtain

$$\|s_1\|_{\psi_1} \leq \frac{2}{I(I-1)} \sum_{i=1}^{I} \sum_{k\in\{1,2\}} \|X_{i,k} - \mathbb{E}[X_{i,k}]\|_{\psi_2}^2 \leq \frac{CC_1^3 C_2^2 \|\vec{\boldsymbol{\beta}}\|_2^2}{J_0(I-1)}.$$

For the second summand s_2, we similarly have by the triangle inequality and the same bound used above (eq. (B.32)):

$$\begin{aligned}
\|s_2\|_{\psi_2} &= \left| \frac{1}{I(I-1)} \sum_{i=1}^{I} \sum_{k\in\{1,2\}} 2(X_{i,k} - \mathbb{E}[X_{i,k}])\, \mathbb{E}[X_{i,k}] \right| \\
&\leq \frac{1}{I(I-1)} \sum_{i=1}^{I} \sum_{k\in\{1,2\}} 2\|X_{i,k} - \mathbb{E}[X_{i,k}]\|_{\psi_2} |\, \mathbb{E}[X_{i,k}]| \\
&\leq \frac{2CC_1^{3/2} C_2 \|\vec{\boldsymbol{\beta}}\|_2}{\sqrt{J_0}(I-1)} \left(\frac{1}{2I} \sum_{i=1}^{I} \sum_{k\in\{1,2\}} |\mathbb{E}[X_{i,k}]| \right).
\end{aligned}$$

By two applications of Jensen's inequality to the term in parentheses, this is

$$\begin{aligned}
&\leq \frac{2CC_1^{3/2} C_2 \|\vec{\boldsymbol{\beta}}\|_2}{\sqrt{J_0}(I-1)} \left(\frac{1}{2I} \sum_{i=1}^{I} \sum_{k\in\{1,2\}} \sqrt{\mathbb{E}[X_{i,k}^2]} \right) \\
&\leq \frac{2CC_1^{3/2} C_2 \|\vec{\boldsymbol{\beta}}\|_2}{\sqrt{J_0}(I-1)} \left(\sqrt{\frac{1}{2I} \sum_{i=1}^{I} \sum_{k\in\{1,2\}} \mathbb{E}[X_{i,k}^2]} \right)
\end{aligned}$$

$$= \frac{\sqrt{2}CC_1^{3/2}C_2\|\vec{\boldsymbol{\beta}}\|_2}{\sqrt{J_0(I-1)}} \left(\sqrt{\frac{1}{I(I-1)} \sum_{i=1}^{I} \sum_{k \in \{1,2\}} \mathbb{E}[X_{i,k}^2]} \right)$$
$$= \frac{\sqrt{2}CC_1^{3/2}C_2\|\vec{\boldsymbol{\beta}}\|_2}{\sqrt{J_0(I-1)}} \sqrt{\mathbb{E}[\mathrm{Var}\{\hat{\tau}^{\Pi}(\vec{\boldsymbol{\beta}})\}]},$$

It follows from definition B.11 (iv) and (v) that, for a universal constant C' possibly larger than C, the following events each have probability at least $1-\eta$

$$s_1 \le \frac{C'C_1^3C_2^2\|\vec{\boldsymbol{\beta}}\|_2^2}{J_0(I-1)} \log(2/\eta)$$
$$s_2 \le \frac{C'C_1^{3/2}C_2\|\vec{\boldsymbol{\beta}}\|_2}{\sqrt{J_0(I-1)}} \sqrt{\mathbb{E}[\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})|\Pi\}]} \log(2/\eta)^{1/2}.$$

Thus, replacing η by $\eta/2$ and using a union bound, it holds with probability $1-\eta$ that

$$\begin{aligned} |\mathrm{Var}\{\hat{\tau}^{\Pi}(\vec{\boldsymbol{\beta}})\} - \mathbb{E}[\mathrm{Var}\{\hat{\tau}^{\Pi}(\vec{\boldsymbol{\beta}})\}]| &\le s_1 + s_2 \\ &\le \frac{C'C_1^3C_2^2\|\vec{\boldsymbol{\beta}}\|_2^2}{J_0(I-1)} \log(4/\eta) \\ &\quad + \frac{C'C_1^{3/2}C_2\|\vec{\boldsymbol{\beta}}\|_2}{\sqrt{J_0(I-1)}} \sqrt{\mathbb{E}[\mathrm{Var}\{\hat{\tau}^{\Pi}(\vec{\boldsymbol{\beta}})\}] \log(4/\eta)}, \end{aligned}$$

which, after noting $J_0 \ge J/C_1$ and simplifying, gives us the claimed inequality. □

B.2.7 Simplifying the conditional CLT

We now simplify Lemma B.3, which, as stated, involves normalization by the random conditional variance $\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})|\Pi\}$. In Lemma B.17, we show that $\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})|\Pi\}$ can be replaced by the deterministic quantity $\mathbb{E}[\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})|\Pi\}]$, which simplifies the analysis in Appendix B.3 to follow. We first state three helper lemmas, namely lemmas B.13 to B.15.

Lemma B.13 (Lemma 2.1 of Chernozhukov et al. [2016]). *Let X, Y be real, Borel random variables. Suppose that $\mathbb{P}(|X-Y|>\nu) \le \eta$. Then*

$$\sup_{t\in\mathbb{R}} |\mathbb{P}(X \le t) - \mathbb{P}(Y \le t)| \le \eta + \sup_{t\in\mathbb{R}} \mathbb{P}(|Y-t| \le \nu).$$

If $Y \sim N(0,1)$ is standard normal then the RHS of the bound above simplifies to $\eta + 2\nu$.

Proof. The first statement is exactly Lemma 2.1 in Chernozhukov et al. [2016]; the second claim follows as the density of a standard Gaussian random variable is bounded by 1. □

Lemma B.14. *Let X, Y be real, Borel random variables. Let $\Phi_\sigma(t) = \Phi(t/\sigma)$ be the CDF of a zero-mean Gaussian with variance σ^2.*

If $\mathbb{P}(|X-Y|>\nu) \le \eta$ and $\sup_{t\in\mathbb{R}} |\mathbb{P}(Y \le t) - \Phi_\sigma(t)| \le \epsilon$, then

$$\sup_{t\in\mathbb{R}} |\mathbb{P}(X \le t) - \Phi_\sigma(t)| \le \eta + 3\epsilon + 2\nu/\sigma.$$

Proof. Given any $t \in \mathbb{R}$ our assumptions imply $\mathbb{P}(Y \le t) \le \Phi_\sigma(t) + \epsilon$. Using continuity of Φ_σ, they also imply $\mathbb{P}(Y < t) = \lim_{u\uparrow t} \mathbb{P}(Y \le u) \ge \lim_{u\uparrow t} \Phi_\sigma(u) - \epsilon = \Phi_\sigma(t) - \epsilon$. Thus,

$$\begin{aligned}\sup_{t\in\mathbb{R}} \mathbb{P}(|Y-t| \le \nu) &= \sup_{t\in\mathbb{R}} \{\mathbb{P}(Y \le t+\nu) - \mathbb{P}(Y < t-\nu)\} \\ &\le \sup_{t\in\mathbb{R}} \{\Phi_\sigma(t+\nu) + \epsilon - [\Phi_\sigma(t-\nu) - \epsilon]\} \\ &= 2\epsilon + \sup_{t\in\mathbb{R}} \left\{ \int_{t-\nu}^{t+\nu} \Phi'_\sigma(u)\mathrm{d}u \right\} \le 2\epsilon + 2\nu/\sigma,\end{aligned}$$

where we have used $|\Phi'_\sigma| \le 1/\sigma$ in the last step. We then use Lemma B.13 and the triangle inequality for the norm $\|F\|_\infty = \sup_{t\in\mathbb{R}} |F(t)|$ to conclude:

$$\begin{aligned}\sup_{t\in\mathbb{R}} |\mathbb{P}(X \le t) - \Phi(t)| &\le \sup_{t\in\mathbb{R}} |\mathbb{P}(X \le t) - \mathbb{P}(Y \le t)| + \sup_{t\in\mathbb{R}} |\mathbb{P}(Y \le t) - \Phi(t)| \\ &\le \eta + \sup_{t\in\mathbb{R}} \mathbb{P}(|Y-t| \le \nu) + \epsilon \le 3\epsilon + 2\nu/\sigma + \eta.\end{aligned}$$

□

Lemma B.15. *Let Φ be the standard Gaussian CDF. For any $\eta \in (0,1)$*

$$\sup_{s\in\mathbb{R}} \left| \Phi\left(\frac{s-\mu_1}{\sigma_1}\right) - \Phi\left(\frac{s-\mu_2}{\sigma_2}\right) \right| \le \eta + \frac{|\mu_1-\mu_2|}{\sigma_2} + \frac{|\sigma_1-\sigma_2|}{\sigma_2}\sqrt{2\log(e/\eta)}. \tag{B.33}$$

Proof. Note that by making the substitution $t = \sigma_2 s + \mu_2$ we have

$$\sup_{s\in\mathbb{R}} \left| \Phi\left(\frac{s-\mu_1}{\sigma_1}\right) - \Phi\left(\frac{s-\mu_2}{\sigma_2}\right) \right| = \sup_{t\in\mathbb{R}} \left| \Phi\left(\frac{t-(\mu_1-\mu_2)/\sigma_2}{\sigma_1/\sigma_2}\right) - \Phi(t) \right|$$

Let Y be standard normal. Define $X = (\sigma_1/\sigma_2)Y + (\mu_1-\mu_2)/\sigma_2$. By construction, then,

$$|Y - X| \le \sigma_2^{-1}|\mu_1-\mu_2| + |Y||\sigma_1/\sigma_2 - 1|.$$

By the Gaussian concentration inequality $\mathbb{P}(|Y| > t) = 2\Phi(-t) \le 2e^{-t^2/2}$ where the last inequality holds for $t \ge 1$, we can choose $t = \sqrt{2\log(e/\eta)} \ge 1$ to deduce that with probability at least $1-\eta$ we have $|Y| \le \sqrt{2\log(e/\eta)}$. On this event, hence with probability $1-\eta$,

$$|Y - X| \le \frac{|\mu_1-\mu_2|}{\sigma_2} + \frac{|\sigma_1-\sigma_2|}{\sigma_2}\sqrt{2\log(e/\eta)}. \tag{B.34}$$

The proof then follows immediately by applying Lemma B.13. □

Lemma B.16. *Let V be a discrete, real random variable. For a real, Borel random variable X, a cumulative distribution function F, and $\varepsilon : \mathbb{R} \to \mathbb{R}$, suppose that*

$$\sup_{t\in\mathbb{R}} |\mathbb{P}(X \le t \mid V) - F(t)| \le \varepsilon(V).$$

Then, if $U = u(V)$ is a $\sigma(V)$-measurable random variable, it also holds that

$$|\mathbb{P}(X \le U \mid V) - F(U)| \le \varepsilon(V).$$

Proof. For any fixed v in the support of V and $u = u(v)$ we have

$$|\mathbb{P}(X \le u \mid V = v) - F(u)| \le \sup_{t \in \mathbb{R}} |\mathbb{P}(X \le u \mid V = v) - F(t)| \le \varepsilon(v).$$

□

Lemma B.17. *Under assumptions (a) and (b), it holds with probability* $1 - \eta$ *that*

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P} \left\{ \frac{\widehat{\tau}^{\Pi}(\vec{\boldsymbol{\beta}}) - \tau^{\Pi}(\vec{\boldsymbol{\beta}})}{\sqrt{\mathbb{E}[\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})|\Pi\}]}} \le t \middle| \Pi \right\} - \Phi(t) \right| \le \eta + \frac{CC_1^2 C_2 \|\vec{\boldsymbol{\beta}}\|_2 (I^{-1} + J^{-1})}{\sqrt{\mathbb{E}[\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})|\Pi\}]}} \log(C/\eta) \tag{B.35}$$

Proof. Put $\sigma_\Pi^2 = \mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})|\Pi\}$ and $\sigma_2^2 = \mathbb{E}[\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})|\Pi\}]$. Lemma B.3 states that

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P} \left\{ \sigma_\Pi^{-1} [\widehat{\tau}^{\Pi}(\vec{\boldsymbol{\beta}}) - \tau^{\Pi}(\vec{\boldsymbol{\beta}})] \le t \middle| \Pi \right\} - \Phi(t) \right| \le \frac{CC_2 \|\vec{\boldsymbol{\beta}}\|_2}{\min\{I_0, I_1\}} \frac{1}{\sigma_\Pi} =: \frac{B_1}{\sigma_\Pi}$$

Π is discrete and $U_t = \sigma_\Pi^{-1} t$ is $\sigma(\Pi)$-measurable, and $\widehat{\tau}^{\Pi}(\vec{\boldsymbol{\beta}}) - \tau^{\Pi}(\vec{\boldsymbol{\beta}}) \le t \iff \sigma_\Pi^{-1}[\widehat{\tau}^{\Pi}(\vec{\boldsymbol{\beta}}) - \tau^{\Pi}(\vec{\boldsymbol{\beta}})] \le U_t$. Thus, combining the previous display with Lemma B.16 gives

$$\left| \mathbb{P} \left\{ \sigma_\Pi^{-1} [\widehat{\tau}^{\Pi}(\vec{\boldsymbol{\beta}}) - \tau^{\Pi}(\vec{\boldsymbol{\beta}})] \le U_t \middle| \Pi \right\} - \Phi(U_t) \right| = \left| \mathbb{P} \left\{ \widehat{\tau}^{\Pi}(\vec{\boldsymbol{\beta}}) - \tau^{\Pi}(\vec{\boldsymbol{\beta}}) \le t \middle| \Pi \right\} - \Phi(U_t) \right| \le \sigma_\Pi^{-1} B_1.$$

for any $t \in \mathbb{R}$. By the triangle inequality for $\|f\|_\infty = \sup_{t \in \mathbb{R}} |f(t)|$, the above, and using lemma B.15 to bound $\sup_{t \in \mathbb{R}} \left| \Phi(\sigma_\Pi^{-1} t) - \Phi(\sigma_2^{-1} t) \right|$, it holds for $\eta \in (0, 1)$ that

$$\begin{aligned}
&\sup_{t \in \mathbb{R}} \left| \mathbb{P} \left\{ [\widehat{\tau}^{\Pi}(\vec{\boldsymbol{\beta}}) - \tau^{\Pi}(\vec{\boldsymbol{\beta}})] \le t \middle| \Pi \right\} - \Phi(\sigma_2^{-1} t) \right| \\
&\qquad \le \sup_{t \in \mathbb{R}} \left| \mathbb{P} \left\{ [\widehat{\tau}^{\Pi}(\vec{\boldsymbol{\beta}}) - \tau^{\Pi}(\vec{\boldsymbol{\beta}})] \le t \middle| \Pi \right\} - \Phi(\sigma_\Pi^{-1} t) \right| + \sup_{t \in \mathbb{R}} \left| \Phi(\sigma_\Pi^{-1} t) - \Phi(\sigma_2^{-1} t) \right| \\
&\qquad \le \sigma_\Pi^{-1} B_1 + \sup_{t \in \mathbb{R}} \left| \Phi(\sigma_\Pi^{-1} t) - \Phi(\sigma_2^{-1} t) \right| \\
&\qquad \le \sigma_\Pi^{-1} B_1 + \eta + \left| \frac{\sigma_\Pi - \sigma_2}{\sigma_2} \right| \sqrt{2 \log(e/\eta)}.
\end{aligned}$$

We now manipulate the bound above to remove its dependence on the random quantity σ_Π. We start by considering the first term, $\sigma_\Pi^{-1} B_1$.

First term, case $\sigma_\Pi \ge \frac{1}{2}\sigma_2$: If $\sigma_\Pi \ge \frac{1}{2}\sigma_2$ then $\sigma_\Pi^{-1} B_1 \le 2\sigma_2^{-1} B_1$, implying the bound

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P} \left\{ \widehat{\tau}^{\Pi}(\vec{\boldsymbol{\beta}}) - \tau^{\Pi}(\vec{\boldsymbol{\beta}}) \le t \middle| \Pi \right\} - \Phi(\sigma_2^{-1} t) \right| \le \frac{2B_1}{\sigma_2} + \eta + 2 \left| \frac{\sigma_\Pi - \sigma_2}{\sigma_2} \right| \sqrt{2 \log(e/\eta)}. \tag{B.36}$$

First term, case $\sigma_\Pi < \frac{1}{2}\sigma_2$: On the other hand, if $\sigma_\Pi < \frac{1}{2}\sigma_2$, then $|\sigma_2 - \sigma_\Pi| \ge \frac{1}{2}\sigma_2$, so the third term satisfies $2\sigma_2^{-1}|\sigma_\Pi - \sigma_2| \sqrt{2 \log(e/\eta)} \ge 1$ and the above becomes trivial. We conclude that eq. (B.36) holds.

Third term: Next, we handle the third term $\sigma_2^{-1}|\sigma_\Pi - \sigma_2|$. Combining the inequality $|x-1| \le |x-1||x+1| = |x^2-1|$ for $x = \sigma_\Pi/\sigma_2 > 0$ with Lemma B.12 gives us that with with probability $1-\eta'$,

$$\left|\frac{\sigma_\Pi - \sigma_2}{\sigma_2}\right| \le \left|\frac{\sigma_\Pi^2 - \sigma_2^2}{\sigma_2^2}\right| \le \frac{B_2^2 \log(4/\eta')}{\sigma_2^2} + \frac{B_2\sigma_2\sqrt{\log(4/\eta')}}{\sigma_2^2}; \quad B_2 := \frac{CC_1^2 C_2 \|\vec{\boldsymbol{\beta}}\|_2}{\sqrt{J(I-1)}}$$

The first term is the square of the second term, and that the bound becomes trivial if either exceeds 1, so we may assume that the second term is larger and deduce that with probability $1-\eta'$

$$\sup_{t\in\mathbb{R}} \left|\mathbb{P}\left\{\hat{\tau}^\Pi(\vec{\boldsymbol{\beta}}) - \tau^\Pi(\vec{\boldsymbol{\beta}}) \le t \middle| \Pi\right\} - \Phi(\sigma_2^{-1}t)\right| \le \frac{2B_1}{\sigma_2} + \eta + \frac{B_2\sqrt{8\log(e/\eta)\log(4/\eta')}}{\sigma_2}.$$

Finally, we make the substitution $t' = \sigma_2 t$ above, rearrange, simplify $\min\{I_0, I_1\} \ge I/C_1 \ge I/C_1^3$ in the definition of B_1 and $(IJ)^{-1/2} \le (I^{-1} + J^{-1})/2$ in B_2, and finally take $\eta' = \eta$ to deduce the claimed inequality. □

B.3 Final result

In this subsection we combine the results of appendix B.1 and appendix B.2 and finally state and prove the CLT presented in Theorem 4.8.

B.3.1 Combining CLTs

For any permutation Π, we have the decomposition

$$\hat{\tau}(\vec{\boldsymbol{\beta}}) - \tau(\vec{\boldsymbol{\beta}}) = \{\hat{\tau}(\vec{\boldsymbol{\beta}}) - \tau^\Pi(\vec{\boldsymbol{\beta}})\} + \{\tau^\Pi(\vec{\boldsymbol{\beta}}) - \tau(\vec{\boldsymbol{\beta}})\}. \tag{B.37}$$

Lemmas B.9 and B.17 yield the two following Gaussian approximations

$$\frac{\{\hat{\tau}(\vec{\boldsymbol{\beta}}) - \tau^\Pi(\vec{\boldsymbol{\beta}})\}}{\mathbb{E}[\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})|\Pi\}]^{-1/2}} \overset{\mathrm{d}}{\approx} N(0,1), \quad \text{and} \quad \frac{\{\tau^\Pi(\vec{\boldsymbol{\beta}}) - \tau(\vec{\boldsymbol{\beta}})\}}{\mathrm{Var}\{\tau^\Pi(\vec{\boldsymbol{\beta}})\}^{-1/2}} \overset{\mathrm{d}}{\approx} N(0,1).$$

The rate of convergence in both cases is a function of sample sizes I and J and was characterized in lemmas B.9 and B.17. We use the decomposition in eq. (B.37) to combine these approximations to recover a Gaussian approximation of $\hat{\tau}(\vec{\boldsymbol{\beta}}) - \tau(\vec{\boldsymbol{\beta}})$ (the left-hand side of eq. (B.37)). This is accomplished with the following technical lemma.

Lemma B.18. *Let V be a discrete random variable, and X be different $\sigma(V)$-measurable random variable. Let Y be a real valued random variable. For F, G conditional distribution functions with bounded densities F' and G', if for $\eta \in (0,1)$ and some numbers $\Delta(\eta), \Delta' > 0$,*

$$\mathbb{P}\left\{\sup_{t\in\mathbb{R}} |\mathbb{P}(Y \le t \mid V) - F(t)| \le \Delta(\eta)\right\} \ge 1-\eta, \quad \textit{and} \sup_{t\in\mathbb{R}} |\mathbb{P}(X \le t) - G(t)| \le \Delta' \tag{B.38}$$

then, for $Z \sim F$ and $W \sim G$ independent of each other and of (X, Y, V), we have

$$\sup_{t\in\mathbb{R}} |\mathbb{P}\{X + Y \le t\} - \mathbb{P}\{X + Z \le t\}| \le \eta + \Delta(\eta) \tag{B.39}$$

$$\sup_{t\in\mathbb{R}} |\mathbb{P}\{X + Z \le t\} - \mathbb{P}\{W + Z \le t\}| \le \Delta' \tag{B.40}$$

$$\sup_{t\in\mathbb{R}} |\mathbb{P}\{X + Y \le t\} - \mathbb{P}\{W + Z \le t\}| \le \eta + \Delta(\eta) + \Delta'. \tag{B.41}$$

Proof. Let $t \in \mathbb{R}$ be given. We have

$$\begin{aligned}
|\mathbb{P}\{X + Y \le t\} - \mathbb{E}[F(t - X)]| &= \Big|\mathbb{E}\left[\mathbb{P}(Y \le t - X \mid V) - F(t - X)\right]\Big| \\
&\le \mathbb{E}\left[\big|\mathbb{P}(Y \le t - X \mid V) - F(t - X)\big|\right] \\
&\le \mathbb{E}\left[\sup_{u\in\mathbb{R}} \big|\mathbb{P}(Y \le u \mid V) - F(u)\big|\right], \qquad \text{(B.42)}
\end{aligned}$$

where we have used Jensen's inequality and then Lemma B.16 (as X is $\sigma(V)$-measurable). Define now the event $\mathcal{E}_\eta := \{|\mathbb{P}(Y \le u \mid V) - F(u)| \le \Delta(\eta)\}$. We use $\mathcal{E}_\eta$ to bound the argument of the expectation in eq. (B.42) as follows: (i) on $\mathcal{E}_\eta$, by its definition,

$$\sup_{u\in\mathbb{R}} \big|\mathbb{P}(Y \le u \mid V) - F(u)\big| \le \Delta(\eta);$$

(ii) on $\mathcal{E}_\eta^c$, $\sup_{u\in\mathbb{R}} \big|\mathbb{P}(Y \le u \mid V) - F(u)\big| \le 1$, since it is an absolute difference of probabilities. Hence, we conclude that

$$|\mathbb{P}\{X + Y \le t\} - \mathbb{E}[F(t - X)]| \le \mathbb{E}\left[\mathbb{1}\{\mathcal{E}_\eta^c\} + \mathbb{1}\{\mathcal{E}_\eta\}\Delta(\eta)\right] \le \eta + \Delta(\eta), \tag{B.43}$$

where the very last inequality follows from noting that by eq. (B.38), $\mathbb{P}(\mathcal{E}_\eta^c) \le \eta$.

Next, we can simplify $\mathbb{E}[F(t - X)]$ using the fundamental theorem of calculus, Fubini's theorem, and the convolution formula for sums of independent random variables as follows:

$$\begin{aligned}
\mathbb{E}[F(t - X)] = \mathbb{E}\left[\int_{-\infty}^{t-X} F'(u)\mathrm{d}u\right] &= \mathbb{E}\left[\int_{-\infty}^{\infty} F'(u)\mathbb{1}\{u \le t - X\}\mathrm{d}u\right] \\
&= \int_{-\infty}^{\infty} F'(u)\mathbb{P}\{X \le t - u\}\mathrm{d}u \qquad \text{(B.44)} \\
&= \mathbb{P}(X + Z \le t). \qquad \text{(B.45)}
\end{aligned}$$

Since $t \in \mathbb{R}$ was arbitrary, substituting eq. (B.45) in eq. (B.43) proves eq. (B.39).

Further, expanding from eq. (B.44) allows us to write:

$$\begin{aligned}
\mathbb{E}[F(t - X)] &= \int_{-\infty}^{\infty} F'(u)\mathbb{P}\{X \le t - u\}\mathrm{d}u \\
&= \int_{-\infty}^{\infty} F'(u)G(t - u)\mathrm{d}u + \int_{-\infty}^{\infty} F'(u)[\mathbb{P}\{X \le t - u\} - G(t - u)]\mathrm{d}u \\
&= \mathbb{P}(Z + W \le t) + \int_{-\infty}^{\infty} F'(u)[\mathbb{P}\{X \le t - u\} - G(t - u)]\mathrm{d}u.
\end{aligned}$$

It follows that we may bound

$$\begin{aligned}\left|\mathbb{E}[F(t-X)]-\mathbb{P}(Z+W\le t)\right| &= \left|\int_{-\infty}^{\infty} F'(u)[\mathbb{P}\{X\le t-u\}-G(t-u)]\mathrm{d}u\right| \\ &\le \int_{-\infty}^{\infty} F'(u)\,|\mathbb{P}\{X\le t-u\}-G(t-u)\mathrm{d}u| \\ \left|\mathbb{P}(X+Z\le t)-\mathbb{P}(Z+W\le t)\right| &\le \int_{-\infty}^{\infty} F'(u)\Delta'\mathrm{d}u=\Delta', \end{aligned} \tag{B.46}$$

where we have plugged in the equality $\mathbb{E}[F(t-X)]=\mathbb{P}(X+Z\le t)$ from eq. (B.45). Because $t\in\mathbb{R}$ was arbitrary, this proves eq. (B.40).

Finally, we use the triangle inequality and eqs. (B.43) and (B.46) to deduce:

$$\begin{aligned}|\mathbb{P}\{X+Y\le t\}-\mathbb{P}\{Z+W\le t\}| &\le |\mathbb{P}\{X+Y\le t\}-\mathbb{E}[F(t-X)]| \\ &\quad + |\mathbb{E}[F(t-X)]-\mathbb{P}\{Z+W\le t\}| \\ &\le \eta+\Delta(\eta)+\Delta'\end{aligned}$$

After noting that $t\in\mathbb{R}$ was arbitrary, this proves eq. (B.41). □

B.3.2 Final bound

Finally we prove Theorem 4.8. Next, we give the main result, which shows how we combine the normal approximations in Lemmas B.9 and B.17 with Lemma B.18.

Lemma B.19. *For some $\Delta_1>0$, suppose that*

$$\sup_{t\in\mathbb{R}}\left|\mathbb{P}\left\{\frac{\tau^{\Pi}(\vec{\boldsymbol{\beta}})-\tau(\vec{\boldsymbol{\beta}})}{\sqrt{\mathrm{Var}\{\tau^{\Pi}(\vec{\boldsymbol{\beta}})\}}}\le t\right\}-\Phi(t)\right|\le\frac{\Delta_1}{\sqrt{\mathrm{Var}\{\tau^{\Pi}(\vec{\boldsymbol{\beta}})\}}} \tag{B.47}$$

and that with probability at least $1-\eta$,

$$\sup_{t\in\mathbb{R}}\left|\mathbb{P}\left\{\frac{\hat{\tau}(\vec{\boldsymbol{\beta}})-\tau^{\Pi}(\vec{\boldsymbol{\beta}})}{\sqrt{\mathbb{E}[\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})|\Pi\}]}}\le t\,\middle|\,\Pi\right\}-\Phi(t)\right|\le\eta+\frac{\Delta_2\log(C/\eta)}{\sqrt{\mathbb{E}[\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})|\Pi\}]}}. \tag{B.48}$$

Then, with $\xi(C,t)=Ct\log(C/t)$ and some $\Delta_2>0$, we have

$$\sup_{u\in\mathbb{R}}\left|\mathbb{P}\left\{\frac{\hat{\tau}(\vec{\boldsymbol{\beta}})-\tau(\vec{\boldsymbol{\beta}})}{\sqrt{\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})\}}}\le u\right\}-\Phi(u)\right|\le\xi\left(C,\frac{(\Delta_1+\Delta_2)^{1/3}}{\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})\}^{1/6}}\right). \tag{B.49}$$

Before we prove Lemma B.19, we show how it implies Theorem 4.8.

Corollary B.20 (Theorem 4.8 in the main paper)**.** *Under Assumptions (a) and (b), it* *holds for universal constants $C, C' > 0$ that*

$$\sup_{t\in\mathbb{R}} \left| \mathbb{P}\left\{ \frac{\widehat{\tau}(\vec{\boldsymbol{\beta}}) - \tau(\vec{\boldsymbol{\beta}})}{\sqrt{\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})\}}} \le t \right\} - \Phi(t) \right| \le C'\Delta^{1/3}\log(C'/\Delta); \quad \Delta = \left(\frac{CC_1^2C_2\|\vec{\boldsymbol{\beta}}\|_2(I^{-1}+J^{-1})}{\sqrt{\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})\}}} \right).$$

Proof of Corollary B.20. By Lemma B.9 and Lemma B.17, eqs. (B.47) and (B.48) hold with

$$\Delta_1 = CC_1C_2J^{-1}; \quad \Delta_2 = CC_1^2C_2\|\vec{\boldsymbol{\beta}}\|_2(I^{-1}+J^{-1})$$

Thus, using $C_1^2 \ge C_1$ since $C_1 \ge 1$ by definition, eq. (B.49) holds with $(\Delta_1+\Delta_2) = CC_1^2C_2\|\vec{\boldsymbol{\beta}}\|_2(I^{-1}+J^{-1})$. Thus, by Lemma B.19

$$\sup_{t\in\mathbb{R}} \left| \mathbb{P}\left\{ \frac{\widehat{\tau}(\vec{\boldsymbol{\beta}}) - \tau(\vec{\boldsymbol{\beta}})}{\sqrt{\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})\}}} \le t \right\} - \Phi(t) \right| \le C\Delta^{1/3}\log(C/\Delta).$$

□

Proof of Lemma B.19. We introduce the shorthand $\sigma_1^2 = \mathrm{Var}\{\tau^{\Pi}(\vec{\boldsymbol{\beta}})\}$, $\sigma_2^2 = \mathbb{E}[\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})|\Pi\}]$, and $\sigma^2 = \mathrm{Var}\,\hat{\tau}(\vec{\boldsymbol{\beta}})$, so in particular $\sigma_1^2+\sigma_2^2 = \sigma^2$, and we write $\Phi_s(t) = \Phi(t/s)$ for the Gaussian CDF with scale s. Finally, note that we may assume $(\Delta_1+\Delta_2)/\sigma \le 1$, or else the final bound becomes trivially true.

After substituting $u = (\sigma/\sigma_2)t$ and rearranging, eq. (B.48) gives that

$$\sup_{t\in\mathbb{R}} \left| \mathbb{P}\left\{ \hat{\tau}(\vec{\boldsymbol{\beta}}) - \tau^{\Pi}(\vec{\boldsymbol{\beta}}) \le t|\Pi \right\} - \Phi_{\sigma_2}(t) \right| \le \eta + (\Delta_2/\sigma_2)\log(1/\eta) \tag{B.50}$$

with probability $1-\eta$. Substituting $u = (\sigma/\sigma_1)t$ in eq. (B.47) similarly gives

$$\sup_{t\in\mathbb{R}} \left| \mathbb{P}\left\{ \tau^{\Pi}(\vec{\boldsymbol{\beta}}) - \tau(\vec{\boldsymbol{\beta}}) \le t \right\} - \Phi_{\sigma_1}(t) \right| \le \Delta_1/\sigma_1. \tag{B.51}$$

Now, let $g_1 \sim N(0,\sigma_1^2)$ and $g_2 \sim N(0,\sigma_2^2)$ be independent Gaussian random variables, which are also independent of the random assignment. Applying Lemma B.18 with $X = \tau^{\Pi}(\vec{\boldsymbol{\beta}}) - \tau(\vec{\boldsymbol{\beta}})$, $Y = \hat{\tau}(\vec{\boldsymbol{\beta}}) - \tau^{\Pi}(\vec{\boldsymbol{\beta}})$, $V = \Pi$, $W = g_1$ and $Z = g_2$ gives the bounds

$$\sup_{t\in\mathbb{R}} \left| \mathbb{P}\left\{ \hat{\tau}(\vec{\boldsymbol{\beta}}) - \tau^{\Pi}(\vec{\boldsymbol{\beta}}) \le t \right\} - \mathbb{P}\{\tau^{\Pi}(\vec{\boldsymbol{\beta}}) - \tau(\vec{\boldsymbol{\beta}}) + g_2 \le t\} \right| \le 2\eta + \frac{\Delta_2\log(C/\eta)}{\sigma_2}, \tag{B.52}$$

$$\sup_{t\in\mathbb{R}} \left| \mathbb{P}\left\{ \tau^{\Pi}(\vec{\boldsymbol{\beta}}) - \tau(\vec{\boldsymbol{\beta}}) + g_2 \le t \right\} - \Phi_{\sigma}(t) \right| \le \Delta_1/\sigma_1, \tag{B.53}$$

$$\sup_{t\in\mathbb{R}} \left| \mathbb{P}\left\{ \hat{\tau}(\vec{\boldsymbol{\beta}}) - \tau(\vec{\boldsymbol{\beta}}) \le t \right\} - \Phi_{\sigma}(t) \right| \le 2\eta + (\Delta_1/\sigma_1) + \frac{\Delta_2\log(C/\eta)}{\sigma_2}, \tag{B.54}$$

where we used $W + Z = g_1 + g_2 \sim N(0,\sigma^2)$ since g_1, g_2 are independent and $\sigma_1^2+\sigma_2^2 = \sigma^2$. We then consider cases, first assuming that $\Delta_1/\sigma_1 \le \sigma_1/\sigma$ and $\Delta_2/\sigma_2 \le \sigma_2/\sigma$; otherwise we will show that the proof simplifies.

Case 1, $\Delta_1/\sigma_1 \le \sigma_1/\sigma$ and $\Delta_2/\sigma_2 \le \sigma_2/\sigma$. In this case, we start from eq. (B.54). Our assumption that $\Delta_1/\sigma_1 \le \sigma_1/\sigma$ and $\Delta_2/\sigma_2 \le \sigma_2/\sigma$ implies $\frac{\Delta_1}{\sigma} = \frac{\Delta_1}{\sigma_1}\frac{\sigma_1}{\sigma} \ge \frac{\Delta_1^2}{\sigma_1^2}$ and $\frac{\Delta_1}{\sigma} = \frac{\Delta_1}{\sigma_1}\frac{\sigma_1}{\sigma} \ge \frac{\Delta_1^2}{\sigma_1^2}$ Plugging this into the above, putting $\Delta_{1+2} := \Delta_1 + \Delta_2$, and using $\sqrt{a}+\sqrt{b} \le 2\sqrt{a+b}$, we get

$$\sup_{t\in\mathbb{R}} \left|\mathbb{P}\left\{\hat{\tau}(\vec{\boldsymbol{\beta}}) - \tau(\vec{\boldsymbol{\beta}}) \le t\right\} - \Phi_\sigma(t)\right| \le 2\eta + \frac{\sqrt{\Delta_1}+\sqrt{\Delta_2}\log(C/\eta)}{\sqrt{\sigma}} \le 2\eta + 2\sqrt{\frac{\Delta_{1+2}}{\sigma}}\log(C/\eta).$$

Since we may assume $\Delta_{1+2}/\sigma \le 1$ or else the final bound is trivial, we may plug in $\eta = \Delta_{1+2}/\sigma$ to obtain the simplified bound

$$\begin{aligned}\sup_{t\in\mathbb{R}} \left|\mathbb{P}\left\{\hat{\tau}(\vec{\boldsymbol{\beta}}) - \tau(\vec{\boldsymbol{\beta}}) \le t\right\} - \Phi_\sigma(t)\right| &\le C'(\Delta_{1+2}/\sigma)^{1/2}\log\{C'/(\Delta_{1+2}/\sigma)\} \\ &\le C'(\Delta_{1+2}/\sigma)^{1/3}\log\{C'/(\Delta_{1+2}/\sigma)\}.\end{aligned}$$

This is precisely our claim, after taking $u = \sigma t$.

Case 2: $\Delta_2/\sigma_2 > \sigma_2/\sigma$. Multiplying both sides by σ_2/σ gives $\Delta_2/\sigma > \sigma_2^2/\sigma^2$. Moreover, we may assume that $\Delta_1/\sigma_1 \le \sigma_1/\sigma$, since otherwise the same reasoning gives $\Delta_1/\sigma > \sigma_1^2/\sigma^2$, implying $\Delta_{1+2}/\sigma = (\Delta_1+\Delta_2)/\sigma > (\sigma_1^2+\sigma_2^2)/\sigma^2 = 1$, in which case the bound is trivial. Multiplying both sides of the inequality $\Delta_1/\sigma_1 \le \sigma_1/\sigma$ by Δ_1/σ_1 gives $(\Delta_1/\sigma_1)^2 \le \Delta_1/\sigma$. To summarize, we may assume

$$\sigma_2/\sigma < \sqrt{\Delta_2/\sigma}; \quad \Delta_1/\sigma_1 \le \sqrt{\Delta_1/\sigma}. \tag{B.55}$$

By definition, $\mathbb{E}\{[\hat{\tau}(\vec{\boldsymbol{\beta}}) - \tau^{\Pi}(\vec{\boldsymbol{\beta}})]^2\} = \mathbb{E}[\mathbb{E}\{[\hat{\tau}(\vec{\boldsymbol{\beta}}) - \tau^{\Pi}(\vec{\boldsymbol{\beta}})]^2|\Pi\}] = \mathbb{E}[\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}})|\Pi\}] = \sigma_2^2$. By Chebyshev's inequality, using $\mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}}) - \tau^{\Pi}(\vec{\boldsymbol{\beta}}) - g_2\} = \mathrm{Var}\{\hat{\tau}(\vec{\boldsymbol{\beta}}) - \tau^{\Pi}(\vec{\boldsymbol{\beta}})\} + \mathrm{Var}\{g_2\} = 2\sigma_2^2$ due to independence of g_2, it holds with probability at least $1-\eta$ that

$$|\{\tau^{\Pi}(\vec{\boldsymbol{\beta}}) - \tau(\vec{\boldsymbol{\beta}}) + g_2\} - \{\hat{\tau}(\vec{\boldsymbol{\beta}}) - \tau(\vec{\boldsymbol{\beta}})\}| = |\hat{\tau}(\vec{\boldsymbol{\beta}}) - \tau^{\Pi}(\vec{\boldsymbol{\beta}}) - g_2| \le \sigma_2\sqrt{\frac{2}{\eta}}.$$

Applying Lemma B.14 with the above bound and eq. (B.53), we have

$$\sup_{t\in\mathbb{R}} \left|\mathbb{P}\left\{\hat{\tau}(\vec{\boldsymbol{\beta}}) - \tau(\vec{\boldsymbol{\beta}}) \le t\right\} - \Phi_\sigma(t)\right| \le C\left\{\eta + (\Delta_1/\sigma_1) + \frac{\sigma_2}{\sigma}\sqrt{\frac{2}{\eta}}\right\}$$

By eq. (B.55) and $\Delta_1, \Delta_2 \le \Delta_{1+2}$, this simplifies to

$$\le C\left\{\eta + \sqrt{\Delta_{1+2}/\sigma} + \sqrt{\Delta_{1+2}/\sigma}\sqrt{\frac{2}{\eta}}\right\}.$$

Plugging in $\eta = (\Delta_{1+2}/\sigma)^{1/3}$, which we may assume is at most 1, this is

$$\begin{aligned}&\le C'(\Delta_{1+2}/\sigma)^{1/3} \\ &\le C''(\Delta_{1+2}/\sigma)^{1/3}\log\{C''/(\Delta_{1+2}/\sigma)\}.\end{aligned}$$

Case 3: $\Delta_1/\sigma_1 > \sigma_1/\sigma$. This is completely analogous to Case 2, with (Δ_1, σ_1) swapped with (Δ_2, σ_2), and eq. (B.47) replaced by eq. (B.48). By a symmetric argument, we may assume

$$\sigma_1/\sigma < \sqrt{\Delta_1/\sigma}; \quad \Delta_2/\sigma_2 \le \sqrt{\Delta_2/\sigma}. \tag{B.56}$$

By definition, $\mathrm{Var}\{\tau^{\Pi}(\vec{\boldsymbol{\beta}}) - \tau(\vec{\boldsymbol{\beta}})\} = \mathrm{Var}\{\tau^{\Pi}(\vec{\boldsymbol{\beta}})\} = \sigma_1^2$, and $\mathrm{Var}\{\tau^{\Pi}(\vec{\boldsymbol{\beta}}) - \tau(\vec{\boldsymbol{\beta}}) - g_1\} = 2\sigma_1^2$ by independence of g_1, so by Chebyshev's inequality,

$$|\{\tau^{\Pi}(\vec{\boldsymbol{\beta}}) - \tau(\vec{\boldsymbol{\beta}}) + g_2\} - \{g_1 + g_2\}| = |\tau^{\Pi}(\vec{\boldsymbol{\beta}}) - \tau(\vec{\boldsymbol{\beta}}) - g_1| \le \sigma_1 \sqrt{\frac{2}{\eta}}.$$

Applying Lemma B.14 with the above bound and eq. (B.52), we have

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P}\left\{ \hat{\tau}(\vec{\boldsymbol{\beta}}) - \tau(\vec{\boldsymbol{\beta}}) + g_2 \le t \right\} - \Phi_\sigma(t) \right| \le C \left\{ \eta + (\Delta_2/\sigma_2) \log(C/\eta) + \frac{\sigma_1}{\sigma} \sqrt{\frac{2}{\eta}} \right\}.$$

By eq. (B.56) and $\Delta_1, \Delta_2 \le \Delta_{1+2}$, this simplifies to

$$\le C \left\{ \eta + \sqrt{\Delta_{1+2}/\sigma} \log(C/\eta) + \sqrt{\Delta_{1+2}/\sigma} \sqrt{\frac{2}{\eta}} \right\}.$$

Plugging in $\eta = (\Delta_{1+2}/\sigma)^{1/3}$, which we may assume is at most 1, this is

$$\le C'(\Delta_{1+2}/\sigma)^{1/3} \log\{C'/(\Delta_{1+2}/\sigma)\}.$$

□

C Additional simulations

We show simulations for results of section 4 for SMRDs under local interference. Fix $1 \le I_T \le I - 1, 1 \le J_T \le J - 1$, and let $P_{\boldsymbol{\Gamma}}$ be the distribution over the matrix of types $\boldsymbol{\Gamma}$ induced by sampling $\mathbf{W}$ from a SMRD as per Equation (7). We draw data via:

$$\boldsymbol{\Gamma} \sim P_{\boldsymbol{\Gamma}}(\cdot), \quad \text{and} \quad Y_{ij} \mid \boldsymbol{\Gamma} \sim F_{\gamma_{ij}}(\cdot). \tag{C.1}$$

Here, potential outcomes are distributed as follows:

$$Y_{ij} \mid \Gamma_{ij} \overset{ind}{\sim} \begin{cases} F_0(\cdot) & \text{if } \gamma = \text{cc}, \\ F_0(\cdot) + F_{\mathrm{B}}(\cdot) & \text{if } \gamma = \text{ib}, \\ F_0(\cdot) + F_{\mathrm{S}}(\cdot) & \text{if } \gamma = \text{is}, \\ F_1(\cdot) + F_{\mathrm{B}}(\cdot) + F_{\mathrm{S}}(\cdot) & \text{if } \gamma = \text{tr}. \end{cases} \tag{C.2}$$

F_ℓ are distributions, $\ell \in \{0, \mathrm{B}, \mathrm{S}, 1\}$. By construction, data drawn from Equation (C.2) satisfies the local interference assumption (2.4). In our illustration the F_ℓ are Gaussian, although this is not required — indeed, we do not need to impose any parametric assumption on the specification Equation (C.2) for our simulations to be consistent with the theory proved in section 4.2. We set $p_0 := 1, p_1 := 1$, and the proportions of treated buyers and sellers in the MRD be $p_{\mathrm{B}} := I_{\mathrm{T}}/I$ and $p_{\mathrm{S}} := J_{\mathrm{T}}/J$, and $F_\ell(\cdot) = \mathcal{N}(p_\ell \mu_\ell, \sigma_\ell^2)$, $\ell \in \{0, \mathrm{B}, \mathrm{S}, 1\}$. We set $I = 200$, $J = 150$, $p_{\mathrm{B}} = 0.45$, $p_{\mathrm{S}} = 0.55$ and $\mu_0 = 3, \mu_B = -1, \mu_S = -1, \mu_1 = 6$ and $\sigma_x = 1 \; \forall \; x \in \{0, B, S, 1\}$.

To assess validity of the results presented in Section 4, we draw matrices $\boldsymbol{Y}(\gamma) = [Y_{ij}(\gamma)]$ of $I \times J$ fixed potential outcomes $\forall \gamma \in \{\mathrm{cc}, \mathrm{ib}, \mathrm{is}, \mathrm{tr}\}$ via Equation (C.2). We sample 10,000 assignment matrices $\mathbf{W}$ i.i.d. at random from the SMRD $\mathbb{W}$ (equivalently, we sample types $\mathbf{\Gamma}$ from $P_{\mathbf{\Gamma}}$ in Equation (C.1)). Each assignment corresponds to a matrix of types and hence which potential outcomes are observed. To each assignment corresponds an observed matrix of $I \times J$ realized potential outcomes. We use the collection of outcomes from the 10,000 re-randomizations to empirically verify the properties of the proposed estimators.

For the type estimator defined in Equation (10) we check that $\widehat{\overline{\overline{Y}}}_\gamma$ is an unbiased estimate of $\overline{\overline{y}}_\gamma$ (Lemma 4.1) and that $\widehat{\Sigma}_\gamma$ is an unbiased estimator of the variance of the type estimator (Theorem 4.4). Figure 5 reports the histogram of the values attained by $\widehat{\overline{\overline{Y}}}_{\mathrm{cc}}$ across the 10,000 Monte Carlo replicates. From Equation (C.2) (and, under mild assumptions, from the CLT), the type estimator is normally distributed, and from Lemma 4.1, it is centered at the true population value $\overline{\overline{y}}_\gamma$. Moreover, the distance between the 2.5% and 97.5% quantiles of the distribution of the type estimator is close to the length of our 95% confidence interval. In the right panel, we show that $\widehat{\Sigma}_{\mathrm{cc}}$ is an unbiased estimator for the variance of the type estimator, as proved in Theorem 4.4. Analogous results hold for $\mathrm{ib}, \mathrm{is}, \mathrm{tr}$.

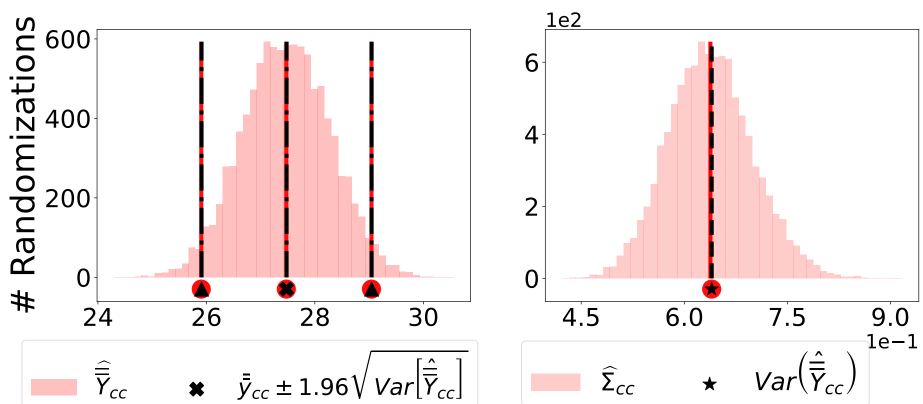

Figure 5: Distribution of $\widehat{\overline{\overline{Y}}}_{\mathrm{cc}}$ (left) and of the variance estimator $\widehat{\Sigma}_{\mathrm{cc}}$ (right). Black lines are plotted in correspondence of the population quantities $\overline{\overline{y}}_{\mathrm{cc}}$, $\mathrm{Var}\left(\widehat{\overline{\overline{Y}}}_{\mathrm{cc}}\right)$.

Figure 6 focuses on the spillover effect $\tau^{\mathrm{B}}_{\mathrm{spill}}$: the left panel shows the distribution of the unbiased estimator $\widehat{\tau}^{\mathrm{B}}_{\mathrm{spill}}$ (Theorem 4.2). $\widehat{\tau}^{\mathrm{B}}_{\mathrm{spill}}$ is a linear combination of Gaussians, and usual confidence intervals can be derived. The right panel contains the distribution of the upper bound $\widehat{\mathrm{Var}}^{\mathrm{hi}}(\widehat{\tau}^{\mathrm{B}}_{\mathrm{spill}})$ for the variance $\mathrm{Var}(\widehat{\tau}^{\mathrm{B}}_{\mathrm{spill}})$ (Theorem 4.5).

C.1 Figures for the average-type and spillover effects

For each $\gamma \in \{\mathrm{cc}, \mathrm{ib}, \mathrm{is}, \mathrm{tr}\}$, we report properties of $\widehat{\overline{\overline{Y}}}_\gamma$ similar to fig. 5 in figs. 7 to 9. In figs. 10 to 12 we provide plots for estimators of spillover effects.

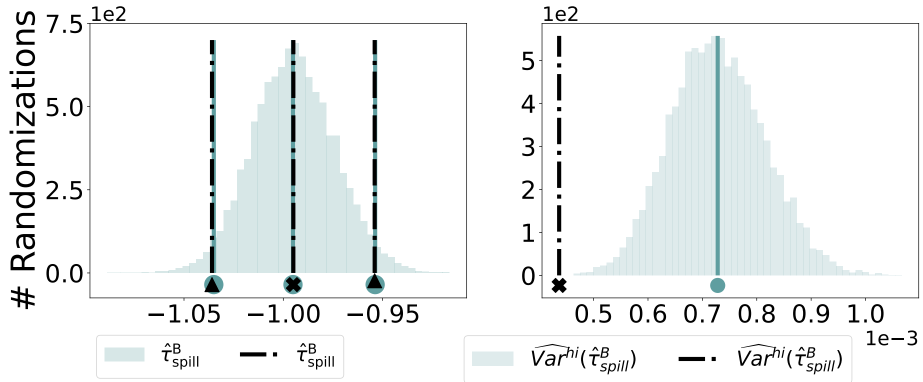

Figure 6: Distribution of the estimator for the spillover effect $\hat{\tau}^B_{\text{spill}}$ (left) and corresponding variance estimator $\widehat{\text{Var}}^{\text{hi}}(\hat{\tau}^B_{\text{spill}})$ (right). Black lines correspond to the population quantities.

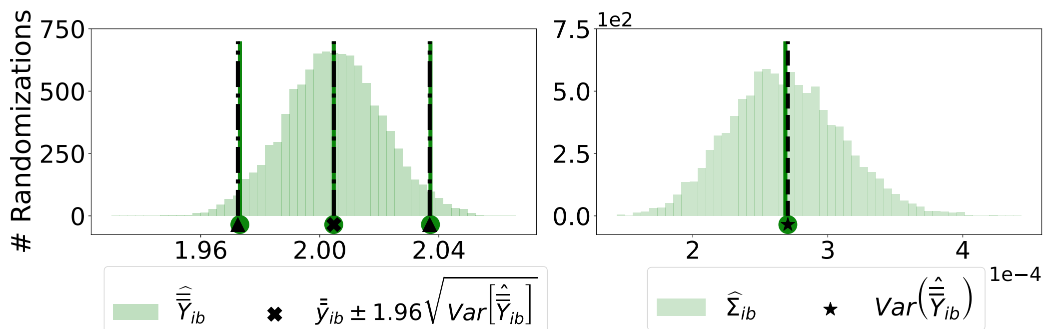

Figure 7: Same as Figure 1, now for ib.

C.2 Testing under the null hypothesis

Similar to section 5, we here provide additional results where we show that, under the null hypothesis of no effect, we can use our derived variance formulae to construct valid test statistics. We consider again the case of $I = 200$ and $J = 150$, and let $\mu_0 = \mu_1 = 3$, $\sigma_0 = \sigma_1 = 1$. We let $\mu_{\text{S}} = \mu_{\text{B}} = 0$ and $\sigma_{\text{S}} = \sigma_{\text{B}} = 0$, leading to potential outcomes $Y_{i,j}(\gamma) = Y_{i,j}(\gamma')$ for $\gamma, \gamma' \in \{\text{cc}, \text{ib}, \text{is}\}$ and $Y_{ij}(\text{cc}) \overset{d}{=} Y_{ij}(\text{tr})$. We run 10,000 Monte Carlo draws, keeping the underlying potential outcomes fixed and randomizing over the assignments $\mathbf{W}$ using $I_1 = 90$ and $J_1 = 85$. We report results from this simulation in fig. 13.

We also compute the test statistics $\hat{t} = \hat{\tau}/\sqrt{\text{Var}(\hat{\tau})}$, where in the denominator we use the true (unknown) variance of the estimator, given in eq. (A.22). Under this enforced null, we test the null hypothesis that a pair of types γ, γ' is associated with no (average) effects on the outcome Y, $H_0^{(\gamma,\gamma')} = \{\tau(\gamma, \gamma') = 0\}$ using the test statistic $\hat{t}$ defined above, and leveraging the normality of the CLT derived in theorem 4.8. The corresponding p-values obtained by a standard two sided t-test are uniformly distributed, as expected (fig. 14).

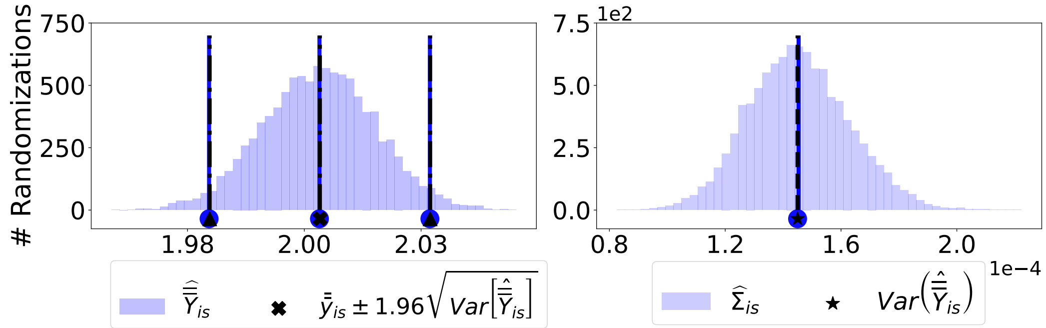

Figure 8: Same as Figure 1, now for is

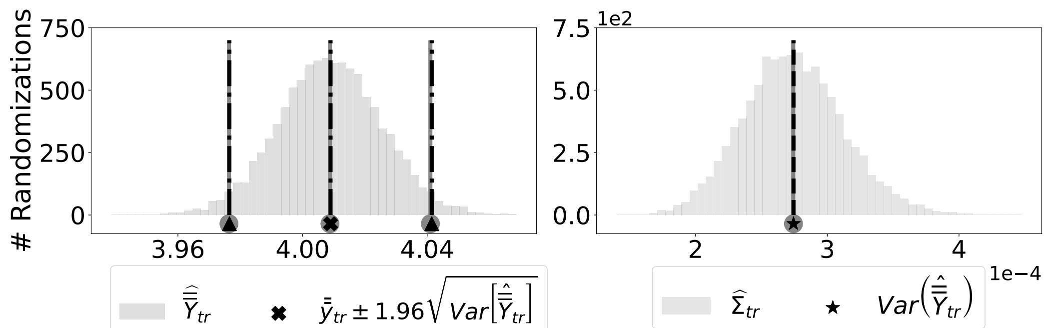

Figure 9: Same as Figure 1, now for tr

References

P. M. Aronow. A general method for detecting interference between units in randomized experiments. *Sociological Methods & Research*, 41(1):3–16, 2012.

P. M. Aronow and C. Samii. Estimating average causal effects under general interference, with application to a social network experiment. *The Annals of Applied Statistics*, 11(4): 1912–1947, 2017.

S. Athey and G. W. Imbens. Design-based analysis in difference-in-differences settings with staggered adoption. *Journal of Econometrics*, 226(1):62–79, 2022.

S. Athey, D. Eckles, and G. W. Imbens. Exact p-values for network interference. *Journal of the American Statistical Association*, 113(521):230–240, 2018.

P. Bajari, B. Burdick, G. Imbens, L. Masoero, J. McQueen, T. Richardson, and I. Rosen. Experimental design in marketplaces. *Statistical Science*, 2023.

G. W. Basse, A. Feller, and P. Toulis. Randomization tests of causal effects under interference. *Biometrika*, 106(2):487–494, 2019.

P. Billingsley. *Probability and measure.* John Wiley & Sons, 2008.

T. Blake and D. Coey. Why marketplace experimentation is harder than it seems: The role of test-control interference. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 567–582, 2014.

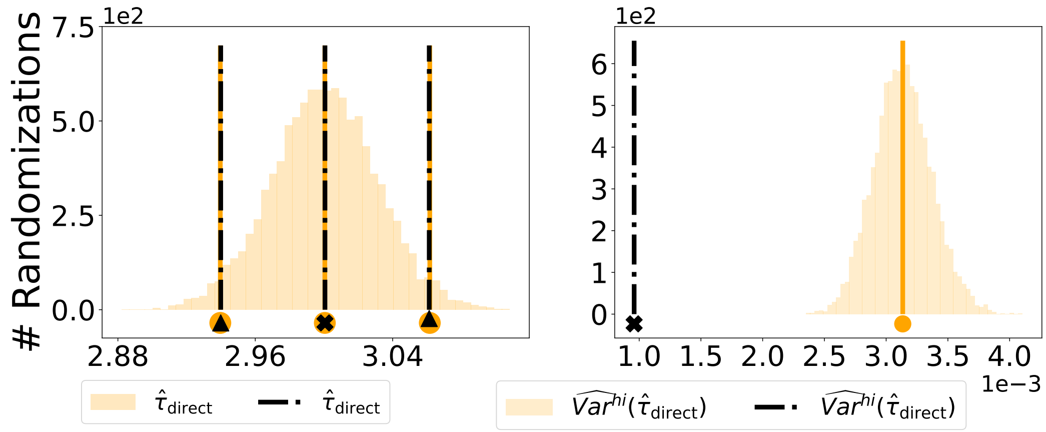

Figure 10: Same as Figure 6, now for $\hat{\tau}_{\text{direct}}$.

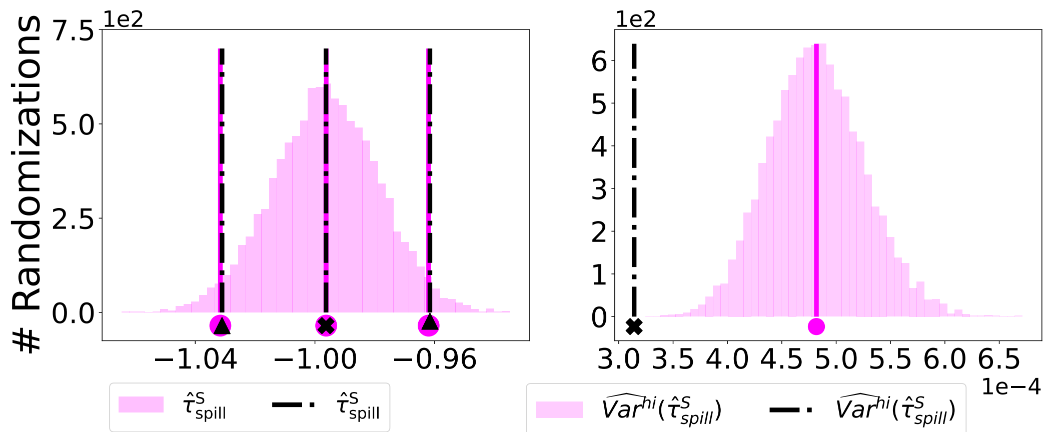

Figure 11: Same as Figure 6, now for $\hat{\tau}^{S}_{\text{spill}}$.

I. Bojinov, D. Simchi-Levi, and J. Zhao. Design and analysis of switchback experiments. *Available at SSRN 3684168*, 2020.

I. Bright, A. Delarue, and I. Lobel. Reducing marketplace interference bias via shadow prices. *Management Science*, 2024.

B. W. Brown Jr. The crossover experiment for clinical trials. *Biometrics*, pages 69–79, 1980.

V. Chernozhukov, D. Chetverikov, and K. Kato. Empirical and multiplier bootstraps for suprema of empirical processes of increasing complexity, and related gaussian couplings. *Stochastic Processes and their Applications*, 126(12):3632–3651, 2016.

W. Cochran. Long-term agricultural experiments. *Supplement to the Journal of the Royal Statistical Society*, 6(2):104–148, 1939.

W. G. Cochran. *Sampling Techniques*. John Wiley & Sons, New York, third edition, September 1977.

D. R. Cox and N. Reid. *The theory of the design of experiments*. Chapman and Hall/CRC, 2000.

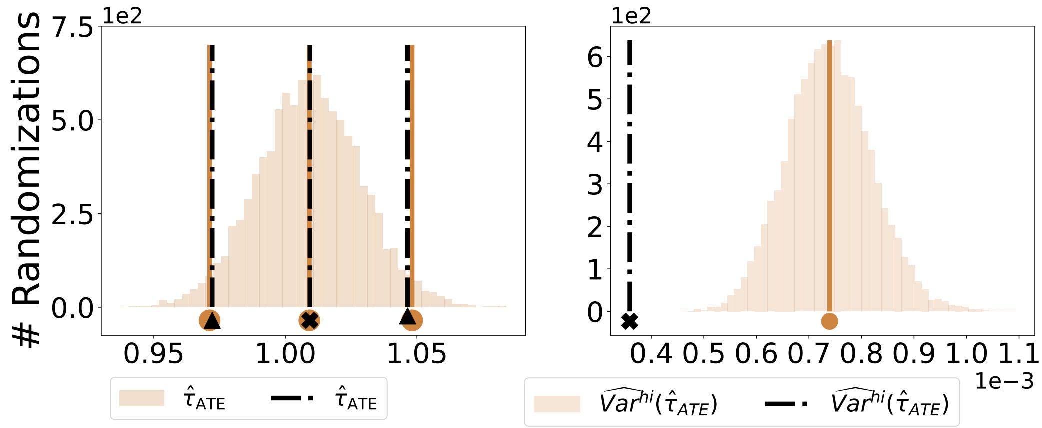

Figure 12: Same as Figure 6, now for $\hat{\tau}_{\text{ATE}}$.

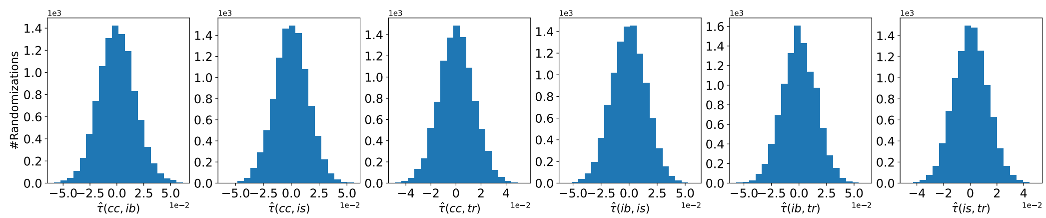

Figure 13: Empirical distribution of $\hat{\tau}$ over 10,000 Monte-Carlo draws

R. A. Fisher. *Statistical methods for research workers*. Number 5. Oliver and Boyd, 1928.

R. A. Fisher. *The design of experiments*. Oliver And Boyd; Edinburgh; London, 1937.

S. Gupta, R. Kohavi, D. Tang, Y. Xu, R. Andersen, E. Bakshy, N. Cardin, S. Chandran, N. Chen, D. Coey, et al. Top challenges from the first practical online controlled experiments summit. *ACM SIGKDD Explorations Newsletter*, 21(1):20–35, 2019.

C. Harshaw, F. Sävje, D. Eisenstat, V. Mirrokni, and J. Pouget-Abadie. Design and analysis of bipartite experiments under a linear exposure-response model. *Proceedings of the 23rd ACM Conference on Economics and Computation*, page 606, 2022. URL https://dl.acm.org/doi/10.1145/3490486.3538269.

K. Hemming, T. P. Haines, P. J. Chilton, A. J. Girling, and R. J. Lilford. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. *BMJ*, 350, 2015.

G. Hong and S. W. Raudenbush. Evaluating kindergarten retention policy: A case study of causal inference for multilevel observational data. *Journal of the American Statistical Association*, 101(475):901–910, 2006.

G. Hong and S. W. Raudenbush. Causal inference for time-varying instructional treatments. *Journal of Educational and Behavioral Statistics*, 33(3):333–362, 2008.

M. G. Hudgens and M. E. Halloran. Toward causal inference with interference. *Journal of the American Statistical Association*, 103(482):832–842, 2008.

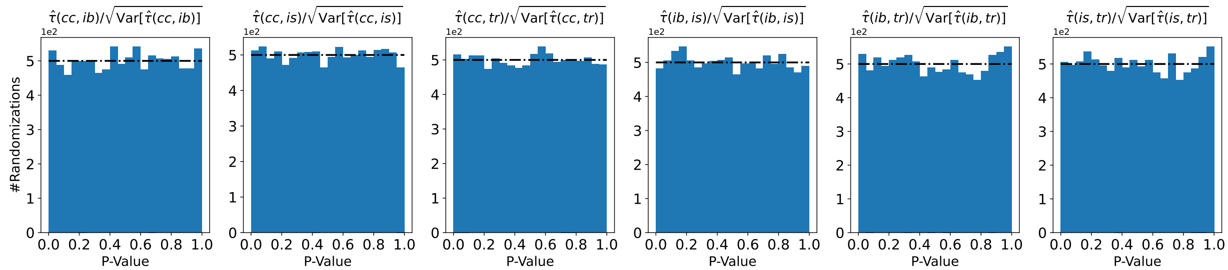

Figure 14: Empirical distribution of the p-values for the test of the null hypothesis $H_0^{(\gamma,\gamma')}$ that the means are identical in two different types γ, γ'.

G. W. Imbens and D. B. Rubin. *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge University Press, 2015.

R. Johari, H. Li, I. Liskovich, and G. Y. Weintraub. Experimental design in two-sided platforms: An analysis of bias. *Management Science*, 2022.

H. Li, G. Zhao, R. Johari, and G. Y. Weintraub. Interference, bias, and variance in two-sided marketplace experimentation: Guidance for platforms. *arXiv preprint arXiv:2104.12222*, 2021.

X. Li and P. Ding. General forms of finite population central limit theorems with applications to causal inference. *Journal of the American Statistical Association*, 112(520):1759–1769, 2017a.

X. Li and P. Ding. General forms of finite population central limit theorems with applications to causal inference. *Journal of the American Statistical Association*, 112(520):1759–1769, 2017b.

C. F. Manski. Identification of treatment response with social interactions. *The Econometrics Journal*, 16(1):S1–S23, 2013.

L. Masoero, G. Imbens, T. Richardson, J. McQueen, S. Vijaykumar, and I. Rosen. Efficient switchback experiments via multiple randomization designs. *Code@MIT*, 2023.

L. Masoero, G. Imbens, S. Vijaykumar, and S. Hut. Measuring direct and indirect impacts in a multi-sided marketplace: Evidence from a clustered multiple randomization experiment. *Code@MIT*, 2024.

P. Milgrom and J. Roberts. Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica: Journal of the Econometric Society*, pages 1255–1277, 1990.

E. Munro, S. Wager, and K. Xu. Treatment effects in market equilibrium. *arXiv preprint arXiv:2109.11647*, 2021.

J. Neyman. Sur les applications de la théorie des probabilités aux experiences agricoles: Essai des principes. *Roczniki Nauk Rolniczych*, 10:1–51, 1923.

J. Neyman. On the application of probability theory to agricultural experiments. Essay on principles. Section 9. *Statistical Science*, 5(4):465–472, 1923/1990.

E. L. Ogburn and T. J. VanderWeele. Causal diagrams for interference. *Statistical science*, 29(4):559–578, 2014.

P. R. Rosenbaum. Interference between units in randomized experiments. *Journal of the American Statistical Association*, 102(477):191–200, 2007. ISSN 01621459.

D. B. Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66(5):688, 1974.

D. Shi and T. Ye. Behavioral carry-over effect and power consideration in crossover trials. *arXiv preprint arXiv:2302.01246*, 2023.

L. Shi and P. Ding. Berry–Esseen bounds for design-based causal inference with possibly diverging treatment levels and varying group sizes. *arXiv preprint arXiv:2209.12345*, 2022a.

L. Shi and P. Ding. Berry–Esseen bounds for design-based causal inference with possibly diverging treatment levels and varying group sizes. *arXiv preprint arXiv:2209.12345*, 2022b.

T. Sudijono, L. Lei, L. Masoero, S. Vijaykumar, G. Imbens, and J. McQueen. Regression adjustments for double randomization in two-sided marketplaces, January 2025. Forthcoming.

M. Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. *Publications Mathématiques de l'Institut des Hautes Etudes Scientifiques*, 81:73–205, 1995.

J. Ugander, B. Karrer, L. Backstrom, and J. Kleinberg. Graph cluster randomization: Network exposure to multiple universes. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 329–337, 2013.

T. J. VanderWeele, E. J. T. Tchetgen, and M. E. Halloran. Interference and sensitivity analysis. *Statistical science: a review journal of the Institute of Mathematical Statistics*, 29(4):687, 2014.

N. Verbitsky-Savitz and S. W. Raudenbush. Exploiting spatial dependence to improve measurement of neighborhood social processes. *Sociological Methodology*, 39(1):151–183, 2009. doi: https://doi.org/10.1111/j.1467-9531.2009.01221.x.

R. Vershynin. *High-dimensional probability: An introduction with applications in data science*, volume 47. Cambridge university press, 2018.

D. Viviano, L. Lei, G. Imbens, B. Karrer, O. Schrijvers, and L. Shi. Causal clustering: design of cluster experiments under network interference. *arXiv preprint arXiv:2310.14983*, 2023.

S. Wager and K. Xu. Experimenting in equilibrium. *Management Science*, 2021.

B. L. Welch. On the z-test in randomized blocks and Latin squares. *Biometrika*, 29(1/2): 21–52, 1937.

R. Xiong, S. Athey, M. Bayati, and G. Imbens. Optimal experimental design for staggered rollouts. *Management Science*, 2023.

A. Zhao and P. Ding. Reconciling design-based and model-based causal inferences for split-plot experiments. *The Annals of Statistics*, 50(2):1170–1192, 2022.

A. Zhao, P. Ding, R. Mukerjee, and T. Dasgupta. Randomization-based causal inference from split-plot designs. *The Annals of Statistics*, 46(5):1876 – 1903, 2018. doi: 10.1214/17-AOS1605. URL https://doi.org/10.1214/17-AOS1605.

L. Zhao, Z. Bai, C.-C. Chao, and W.-Q. Liang. Error bound in a central limit theorem of double-indexed permutation statistics. *The Annals of Statistics*, 25(5):2210–2227, 1997.

Z. Zhu, Z. Cai, L. Zheng, and N. Si. Seller-side experiments under interference induced by feedback loops in two-sided platforms. *arXiv preprint arXiv:2401.15811*, 2024.

C. M. Zigler and G. Papadogeorgou. Bipartite causal inference with interference. *Statistical science: a review journal of the Institute of Mathematical Statistics*, 36(1):109, 2021.