Supplementary Material for "Non-ignorable Attrition in Pairwise Randomized Experiments"

Kentaro Fukumoto
Gakushuin University, Tokyo, Japan.
E-mail: Kentaro.Fukumoto@gakushuin.ac.jp

1. INTRODUCTION

The principal goal of the Supplementary Material is to prove propositions and comments thereof in the main manuscript. For that purpose, I prepare many lemmas. I also (re-)introduce all notations so that the Supplementary Material are self-contained. Note that notation is different between the main manuscript and the Supplementary Material. In the main manuscript, I use conventional notation so that readers can easily understand the paper. If I used the same notation in the Supplementary Material, however, I'm afraid the Supplementary Material would be much longer and harder to read. In addition, the Supplementary Material presents propositions, sometimes with their (tedious) closed forms. The proposition and (sub-)section numbers correspond to those in the main manuscript up to the third section, though the equation numbers do not. The fourth section showcases an application. The fifth section compares this study with Imai & Jiang (2018). The sixth section concludes.

Pairwise randomized experiments is also called matched-pair design or study (Donner & Klar 2000, Imai et al. 2009), pair-matched study (Hayward et al. 2006), paired randomization or (randomized) experiments (Abadie & Imbens 2008; Glennerster & Takavarasha 2013, 158–160; Imbens & Rubin 2015, 53–54; Snedecor & Cochran 1980, ch. 6), or randomized paired (comparison) design (Box et al. 2005, 81–91; Imbens & Rubin 2015, 52, 219). The design follows a golden rule, "block what you can and randomize what you cannot" (Box et al. 2005, 93).

A major problem of pairwise randomized experiments is attrition; outcome values of some units are sometimes missing (Donner & Klar 2000, p. 40; Hayes & Moulton 2009, pp. 72–74; Glennerster & Takavarasha 2013, p. 159). This problem arises more frequently than generally acknowledged. For instance, nurses depart from protocol requirements (Family Heart Study Group 1994), care homes are unable to provide regular data on outcomes (Hayward et al. 2006), no residents meet eligibility criteria for enrollment in a nursing home (Loeb et al. 2006), elections are uncontested (Panagopoulos & Green 2008), schools withdraw from experiments after randomization but before outcome measurement(Angrist & Lavy 2009), and subjects are tired of responding to follow-up survey (Enos 2014).

Moreover, attrition might be non-ignorable (or, in different terminology, not missing at random (Little & Rubin 2002, p. 12) or not missing independent of a potential outcome (Gerber & Green 2012, ch. 7, esp., p. 219)). For example, in the treated group of a remedial education program, as compared with the control group, low-achieving children are more likely to do well and remain in the program, but high-achieving children tend to leave the school because they are unhappy that they must study with low-achieving students (Glennerster & Takavarasha 2013, pp. 307–309). In this case, attrition is associated with test scores.

Examples of studies which use the UDE are Loeb et al. (2006), Angrist & Lavy (2009), and Panagopoulos & Green (2008). Application examples of the PDE are Family Heart Study Group (1994), Hayward et al. (2006), and Enos (2014).

As Dunning (2011) and Gerber & Green (2012) warn and I will also formalize , the PDE can "lead[s] to bias when attrition is a function of potential outcomes" (243). Though King et al. (2007) argue that the PDE is unbiased as long as units are missing "for a reason related to one or more of the variables we matched on" (490), this manuscript shows otherwise. Additionally, while Imai et al. (2009) claim that the PDE "retain[s] the benefits of randomization ... regardless of the missing data mechanism" (44), this manuscript demonstrates that the missing data mechanism matters for the properties of the PDE.

2. FINITE SAMPLE

2.1. Setting

2.1.1. Notation

Realized Values. Suppose that there are $n(\geq 2)$ pairs and each pair is composed of two units. Unit $i \in \{1, 2\}$ in pair $j \in \{1, 2, \ldots, n\}$ is denoted by unit ij. We assign either treatment or control to unit ij. Denote

- Y_{ij}: the outcome of unit ij
- R_{ij}: the response of unit ij

$$R_{ij} = \begin{cases} 1 \text{ if } Y_{ij} \text{ is observed,} \\ 0 \text{ if } Y_{ij} \text{ is missing.} \end{cases}$$

- X^T_{ij}: the treatment indicator of unit ij

$$X^T_{ij} = \begin{cases} 1 \text{ when treatment is assigned to unit } ij, \\ 0 \text{ when control is assigned to unit } ij. \end{cases}$$

- X^C_{ij}: the control indicator of unit ij

$$X^C_{ij} = \begin{cases} 0 \text{ when treatment is assigned to unit } ij, \\ 1 \text{ when control is assigned to unit } ij. \end{cases}$$

Stochastic variables are denoted by upper-case letters. Note also that superscript T means not transpose but treatment. Moreover, the generic assignment indicator is denoted by X^A_{ij} where $A \in \{T, C\}$. It immediately follows

$$\sum_A X^A_{ij} = 1, \tag{1}$$

where $\sum_A \equiv \sum_{A \in \{T,C\}}$.

The pairwise randomized experiments (or the matched-pair design) implies that

$$\sum_i X^A_{ij} = 1, \tag{2}$$

where $\sum_i \equiv \sum_{i=1}^2$, namely, for every j, either $X^T_{1j} = X^C_{2j} = 1$ (and, thus, according to Equation1, $X^T_{2j} = X^C_{1j} = 0$) or $X^T_{1j} = X^C_{2j} = 0$ (and thus $X^T_{2j} = X^C_{1j} = 1$).

Potential Values. The stable unit treatment value assumption (Imbens & Rubin 2015) is applied not only to Y_{ij} but also to R_{ij}. Define

- y^T_{ij}: the potential outcome of unit ij if treatment were assigned to unit ij.
- y^C_{ij}: the potential outcome of unit ij if control were assigned to unit ij.
- r^T_{ij}: the potential response of unit ij if treatment were assigned to unit ij.

$$r^T_{ij} = \begin{cases} 1 \text{ if } Y_{ij} = y^T_{ij} \text{ is observed with treatment assigned to unit } ij, \\ 0 \text{ if } Y_{ij} = y^T_{ij} \text{ is missing with treatment assigned to unit } ij. \end{cases}$$

- r_{ij}^C: the potential response of unit ij if control were assigned to unit ij.

$$r_{ij}^C = \begin{cases} 1 \text{ if } Y_{ij} = y_{ij}^C \text{ is observed with control assigned to unit } ij, \\ 0 \text{ if } Y_{ij} = y_{ij}^C \text{ is missing with control assigned to unit } ij. \end{cases}$$

(Fixed values are denoted by lower-case letters.) It immediately follows

$$\begin{aligned} Y_{ij} &= \sum_A y_{ij}^A X_{ij}^A, \\ R_{ij} &= \sum_A r_{ij}^A X_{ij}^A. \end{aligned} \tag{3}$$

Let Q_{ij} be the generic quantity of unit ij. Its vector is denoted by by the corresponding bold face:

$$\boldsymbol{Q} \equiv \{Q_{ij}\}_{11}^{2n} \equiv (Q_{11}, Q_{21}, Q_{12}, Q_{22}, \ldots, Q_{1n}, Q_{2n}).$$

Let q the generic quantity which is constant irrespective of $\boldsymbol{X}$. When $Q_{ij} = q$ is constant across units, we denote

$$\boldsymbol{Q} = \boldsymbol{q} \equiv \{q\}_{11}^{2n}.$$

For unit ij, denote the value of the other unit of the same pair (i.e., pair mate) by Q_{-ij}. Specifically,

$$Q_{-ij} \equiv \begin{cases} Q_{2j} \text{ when } i = 1, \\ Q_{1j} \text{ when } i = 2. \end{cases}$$

Denote

$$\begin{aligned} \boldsymbol{Q}_{-i} &\equiv \{Q_{-ij}\}_{11}^{2n}, \\ \sum_h \boldsymbol{Q}^{(h)} &\equiv \left\{ \sum_h Q_{ij}^{(h)} \right\}_{11}^{2n}, \\ \prod_h \boldsymbol{Q}^{(h)} &\equiv \left\{ \prod_h Q_{ij}^{(h)} \right\}_{11}^{2n}, \\ \boldsymbol{Q}^h &\equiv \prod_{h'=1}^{h} \boldsymbol{Q}, \\ q\boldsymbol{Q} &\equiv \boldsymbol{q}\boldsymbol{Q}, \\ -\boldsymbol{Q} &\equiv (-1)\boldsymbol{Q}, \\ \boldsymbol{Q}^{(1)} - \boldsymbol{Q}^{(2)} &\equiv \boldsymbol{Q}^{(1)} + (-\boldsymbol{Q}^{(2)}), \end{aligned}$$

where h is the generic counter, $\sum_h \equiv \sum_{h=1}^{h_{\max}}$. Note that $\boldsymbol{Q}^{(1)}\boldsymbol{Q}^{(2)}$ denote element-wise multiplication, neither inner nor outer product of two vectors.

Denote the dummy variable vector space by

$$\mathbb{U} \equiv \{\boldsymbol{U} | U_{ij} \in \{0, 1\}\}.$$

For $\boldsymbol{Q}$ and the generic weight $\boldsymbol{Z}$, where $Z_{ij} \geq 0$, when $\boldsymbol{Z} \neq \boldsymbol{0}$, the weighted mean operator is defined as

$$E(\boldsymbol{Q}|\boldsymbol{Z}) \equiv \frac{\sum_j \sum_i Z_{ij} Q_{ij}}{\sum_j \sum_i Z_{ij}}, \tag{4}$$

where $\sum_j \equiv \sum_{j=1}^{n}$, and, by abusing notation,

$$E(\boldsymbol{Q}|\boldsymbol{0}) \equiv 0. \tag{5}$$

LEMMA 1 (ARITHMETIC OF VECTORS). *(1)*

$$\boldsymbol{Q}^{(1)} + \boldsymbol{Q}^{(2)} = \boldsymbol{Q}^{(2)} + \boldsymbol{Q}^{(1)}$$

(2)

$$\boldsymbol{Q}^{(1)}\boldsymbol{Q}^{(2)} = \boldsymbol{Q}^{(2)}\boldsymbol{Q}^{(1)}$$

(3)

$$(\boldsymbol{Q}^{(1)} \pm \boldsymbol{Q}^{(2)})\boldsymbol{Q}^{(3)} = \boldsymbol{Q}^{(1)}\boldsymbol{Q}^{(3)} \pm \boldsymbol{Q}^{(2)}\boldsymbol{Q}^{(3)}$$

(4)

$$\sum_i Q_{-ij} = \sum_i Q_{ij}$$

(5) For $\boldsymbol{U} \in \mathbb{U}$*,*

$$\begin{aligned} \boldsymbol{U}^2 &= \boldsymbol{U} \\ \mathbf{1} - \boldsymbol{U} &\in \mathbb{U} \\ \boldsymbol{U}_{-i} &\in \mathbb{U}. \end{aligned}$$

(6) When $\boldsymbol{U}^{(h)} \in \mathbb{U}$*,*

$$\prod_h \boldsymbol{U}^{(h)} \in \mathbb{U}.$$

PROOF. (1)–(3) Obvious.
(4)

$$\begin{aligned} \sum_i Q_{-ij} &= Q_{2j} + Q_{1j} \\ &= \sum_i Q_{ij} \end{aligned}$$

(5) It is obvious that $\boldsymbol{U}_{-i} \in \mathbb{U}$. Since $U_{ij} \in \{0, 1\}$,

$$\begin{aligned} U_{ij}^2 &= U_{ij} \\ 1 - U_{ij} &\in \{0, 1\} \end{aligned}$$

The desired results immediately follow.
(6) Obvious.

□

Since Lemma 1 (1)–(3) is obvious, I do not mention them even when I invoke them in the following proof.

LEMMA 2 (GENERIC ASSIGNMENT INDICATOR). *(1)*

$$\boldsymbol{X}^A = \boldsymbol{X}_{-i}^{-A}$$

(2)

$$\prod_A \boldsymbol{X}^A = \mathbf{0}$$

(3)

$$\boldsymbol{X}^A \boldsymbol{X}_{-i}^A = \mathbf{0}$$

(4)

$$\boldsymbol{X}^A \boldsymbol{X}_{-i}^{-A} = \boldsymbol{X}^A$$

PROOF. (1)

$$\begin{aligned}\boldsymbol{X}^A &= \mathbf{1} - \boldsymbol{X}^{-A} \quad (\because \text{Equation 1}\\ &= \boldsymbol{X}_{-i}^{-A} \quad (\because \text{Equation 2})\end{aligned}$$

(2)

$$\begin{aligned}\boldsymbol{X}^A\boldsymbol{X}^{-A} &= \boldsymbol{X}^A(1 - \boldsymbol{X}^A) \quad (\because \text{Equation 1})\\ &= \boldsymbol{X}^A - (\boldsymbol{X}^A)^2\\ &= \boldsymbol{X}^A - \boldsymbol{X}^A \quad (\because \boldsymbol{X}^A \in \mathbb{U}, \text{Lemma 1 (5)})\\ &= \mathbf{0}\end{aligned}$$

(3)

$$\begin{aligned}\boldsymbol{X}^A\boldsymbol{X}_{-i}^{A} &= \boldsymbol{X}^A\boldsymbol{X}^{-A} \quad (\because \text{Lemma 2 (1)})\\ &= \mathbf{0} \quad (\because \text{Lemma 2 (2)})\end{aligned}$$

(4)

$$\begin{aligned}\boldsymbol{X}^A\boldsymbol{X}_{-i}^{-A} &= (\boldsymbol{X}^A)^2 \quad (\because \text{Lemma 2 (1)})\\ &= \boldsymbol{X}^A \quad (\because \boldsymbol{X}^A \in \mathbb{U}, \text{Lemma 1 (5)})\end{aligned}$$

□

LEMMA 3 (ARITHMETIC OF WEIGHTED MEAN). *(1)*

$$E\Big(\sum_h \boldsymbol{Q}^{(h)}\Big|\boldsymbol{Z}\Big) = \sum_h E(\boldsymbol{Q}^{(h)}|\boldsymbol{Z}).$$

(2)

$$E(q\boldsymbol{Q}|\boldsymbol{Z}) = qE(\boldsymbol{Q}|\boldsymbol{Z}).$$

(3) When $\boldsymbol{Z} \neq \mathbf{0}$*,*

$$E(\boldsymbol{q}|\boldsymbol{Z}) = q.$$

(4) When $\boldsymbol{Z}^{(1)}\boldsymbol{Z}^{(2)} = \boldsymbol{Z}^{(2)}$*,*

$$E(\boldsymbol{Z}^{(1)}\boldsymbol{Q}|\boldsymbol{Z}^{(2)}) = E(\boldsymbol{Q}|\boldsymbol{Z}^{(2)}).$$

(5) When $\boldsymbol{Z}^{(1)} \neq \mathbf{0}$

$$E(\boldsymbol{Q}|\boldsymbol{Z}^{(1)}) = \frac{\sum_j\sum_i Z_{ij}^{(1)}Z_{ij}^{(2)}}{\sum_j\sum_i Z_{ij}^{(1)}}E(\boldsymbol{Q}|\boldsymbol{Z}^{(1)}\boldsymbol{Z}^{(2)}) + \Big(1 - \frac{\sum_j\sum_i Z_{ij}^{(1)}Z_{ij}^{(2)}}{\sum_j\sum_i Z_{ij}^{(1)}}\Big)E\{\boldsymbol{Q}|\boldsymbol{Z}^{(1)}(\mathbf{1} - \boldsymbol{Z}^{(2)})\}.$$

PROOF. (1) When $\boldsymbol{Z} = \mathbf{0}$, this is obvious thanks to Equation 5. Otherwise,

$$\begin{aligned}E\Big(\sum_h \boldsymbol{Q}^{(h)}\Big|\boldsymbol{Z}\Big) &= \frac{\sum_j\sum_i Z_{ij}\sum_h Q_{ij}^{(h)}}{\sum_j\sum_i Z_{ij}} \quad (\because \text{Equation 4})\\ &= \sum_h \frac{\sum_j\sum_i Z_{ij}Q_{ij}^{(h)}}{\sum_j\sum_i Z_{ij}}\\ &= \sum_h E(\boldsymbol{Q}^{(h)}|\boldsymbol{Z}) \quad (\because \text{Equation 4})\end{aligned}$$

(2) When $\boldsymbol{Z} = \boldsymbol{0}$, this is obvious thanks to Equation 5. Otherwise,

$$\begin{aligned} E(q\boldsymbol{Q}|\boldsymbol{Z}) &= \frac{\sum_j \sum_i Z_{ij} q Q_{ij}}{\sum_j \sum_i Z_{ij}} \quad (\because \text{Equation 4}) \\ &= q \frac{\sum_j \sum_i Z_{ij} Q_{ij}}{\sum_j \sum_i Z_{ij}} \\ &= qE(\boldsymbol{Q}|\boldsymbol{Z}) \quad (\because \text{Equation 4}) \end{aligned}$$

(3)

$$\begin{aligned} E(\boldsymbol{q}|\boldsymbol{Z}) &= \frac{\sum_j \sum_i Z_{ij} q}{\sum_j \sum_i Z_{ij}} \quad (\because \text{Equation 4}) \\ &= q \frac{\sum_j \sum_i Z_{ij}}{\sum_j \sum_i Z_{ij}} \\ &= q. \end{aligned}$$

(4) When $\boldsymbol{Z}^{(2)} = \boldsymbol{0}$, this is obvious thanks to Equation 5. Otherwise,

$$\begin{aligned} E(\boldsymbol{Z}^{(1)}\boldsymbol{Q}|\boldsymbol{Z}^{(2)}) &= \frac{\sum_j \sum Z_{ij}^{(2)} Z_{ij}^{(1)} Q_{ij}}{\sum_j \sum Z_{ij}^{(2)}} \quad (\because \text{Equation 4}) \\ &= \frac{\sum_j \sum Z_{ij}^{(2)} Q_{ij}}{\sum_j \sum Z_{ij}^{(2)}} \quad (\because \boldsymbol{Z}^{(1)}\boldsymbol{Z}^{(2)} = \boldsymbol{Z}^{(2)}) \\ &= E(\boldsymbol{Q}|\boldsymbol{Z}^{(2)}) \quad (\because \text{Equation 4}) \end{aligned}$$

(5) When $\boldsymbol{Z}^{(1)}\boldsymbol{Z}^{(2)} \neq \boldsymbol{0}$ and $\boldsymbol{Z}^{(1)}(\boldsymbol{1} - \boldsymbol{Z}^{(2)}) \neq \boldsymbol{0}$,

$$\begin{aligned} & E(\boldsymbol{Q}|\boldsymbol{Z}^{(1)}) \\ &= \frac{\sum_j \sum_i Z_{ij}^{(1)} Q_{ij}}{\sum_j \sum_i Z_{ij}^{(1)}} \quad (\because \text{Equation 4}, \boldsymbol{Z}^{(1)} \neq \boldsymbol{0}) \\ &= \frac{\sum_j \sum_i Z_{ij}^{(1)} \{Z_{ij}^{(2)} + (1 - Z_{ij}^{(2)})\} Q_{ij}}{\sum_j \sum_i Z_{ij}^{(1)}} \\ &= \frac{\sum_j \sum_i Z_{ij}^{(1)} Z_{ij}^{(2)} Q_{ij}}{\sum_j \sum_i Z_{ij}^{(1)}} + \frac{\sum_j \sum_i Z_{ij}^{(1)} (1 - Z_{ij}^{(2)}) Q_{ij}}{\sum_j \sum_i Z_{ij}^{(1)}} \\ &= \frac{\sum_j \sum_i Z_{ij}^{(1)} Z_{ij}^{(2)}}{\sum_j \sum_i Z_{ij}^{(1)}} \frac{\sum_j \sum_i Z_{ij}^{(1)} Z_{ij}^{(2)} Q_{ij}}{\sum_j \sum_i Z_{ij}^{(1)} Z_{ij}^{(2)}} + \frac{\sum_j \sum_i Z_{ij}^{(1)} (1 - Z_{ij}^{(2)})}{\sum_j \sum_i Z_{ij}^{(1)}} \frac{\sum_j \sum_i Z_{ij}^{(1)} (1 - Z_{ij}^{(2)}) Q_{ij}}{\sum_j \sum_i Z_{ij}^{(1)} (1 - Z_{ij}^{(2)})} \\ & (\because \boldsymbol{Z}^{(1)}\boldsymbol{Z}^{(2)} \neq \boldsymbol{0}, \boldsymbol{Z}^{(1)}(\boldsymbol{1} - \boldsymbol{Z}^{(2)}) \neq \boldsymbol{0}) \\ &= \frac{\sum_j \sum_i Z_{ij}^{(1)} Z_{ij}^{(2)}}{\sum_j \sum_i Z_{ij}^{(1)}} E(\boldsymbol{Q}|\boldsymbol{Z}^{(1)}\boldsymbol{Z}^{(2)}) + \left(1 - \frac{\sum_j \sum_i Z_{ij}^{(1)} Z_{ij}^{(2)}}{\sum_j \sum_i Z_{ij}^{(1)}}\right) E\{\boldsymbol{Q}|\boldsymbol{Z}^{(1)}(\boldsymbol{1} - \boldsymbol{Z}^{(2)})\} \quad (\because \text{Equation 4}) \end{aligned}$$

When $\boldsymbol{Z}^{(1)}\boldsymbol{Z}^{(2)} = \boldsymbol{0}$, it follows that

$$\begin{aligned} \sum_j \sum_i Z_{ij}^{(1)} Z_{ij}^{(2)} &= 0 \\ \boldsymbol{Z}^{(1)}(\boldsymbol{1} - \boldsymbol{Z}^{(2)}) &= \boldsymbol{Z}^{(1)} - \boldsymbol{Z}^{(1)}\boldsymbol{Z}^{(2)} \quad (\because \text{Lemma 1 (3)}) \\ &= \boldsymbol{Z}^{(1)} \end{aligned}$$

Therefore,

$$\begin{aligned}
&\frac{\sum_j \sum_i Z_{ij}^{(1)} Z_{ij}^{(2)}}{\sum_j \sum_i Z_{ij}^{(1)}} E(\boldsymbol{Q}|\boldsymbol{Z}^{(1)}\boldsymbol{Z}^{(2)}) + \Big(1 - \frac{\sum_j \sum_i Z_{ij}^{(1)} Z_{ij}^{(2)}}{\sum_j \sum_i Z_{ij}^{(1)}}\Big) E\{\boldsymbol{Q}|\boldsymbol{Z}^{(1)}(\boldsymbol{1} - \boldsymbol{Z}^{(2)})\} \\
&= \frac{0}{\sum_j \sum_i Z_{ij}^{(1)}} E(\boldsymbol{Q}|\boldsymbol{0}) + \Big(1 - \frac{0}{\sum_j \sum_i Z_{ij}^{(1)}}\Big) E(\boldsymbol{Q}|\boldsymbol{Z}^{(1)}) \\
&= E(\boldsymbol{Q}|\boldsymbol{Z}^{(1)}).
\end{aligned}$$

The case of $\boldsymbol{Z}^{(1)}(\boldsymbol{1} - \boldsymbol{Z}^{(2)}) = \boldsymbol{0}$ is similar.

□

2.1.2. Estimation

The vector of the unit level treatment effects is defined as

$$\boldsymbol{\tau} \equiv \boldsymbol{y}^T - \boldsymbol{y}^C. \tag{6}$$

(Unobservable quantities are denoted by Greek letters.) Thus, the main estimand of this section, the finite sample average treatment effect, is defined as

$$\bar{\tau} \equiv E(\boldsymbol{\tau}|\boldsymbol{1}). \tag{7}$$

Denote the generic realized category indicator, which is a function of $\boldsymbol{X}$, by $\boldsymbol{K}_G \in \mathbb{U}$. Define

- $k_{G,ij}^T$: the generic potential category indicator of unit ij if treatment were assigned to unit ij.
- $k_{G,ij}^C$: the generic potential category indicator of unit ij if control were assigned to unit ij.

It immediately follows

$$\boldsymbol{K}_G = \sum_A \boldsymbol{k}_G^A \boldsymbol{X}^A. \tag{8}$$

Denote the number of units in the treatment status A which take a value of $K_{G,ij} = 1$ by

$$\begin{aligned}
N_G^A &\equiv N^A(\boldsymbol{K}_G) \\
&\equiv \sum_j \sum_i K_{G,ij} X_{ij}^A.
\end{aligned} \tag{9}$$

For $\boldsymbol{K}_G$, the generic ATE estimator is defined as

$$\begin{aligned}
\hat{\tau}_G &\equiv \hat{\tau}(\boldsymbol{K}_G) \\
&\equiv E(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^T) - E(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^C)
\end{aligned} \tag{10}$$

only if (but not always if) $N_G^A \geq 1$.

The full sample estimator is defined as

$$\hat{\tau}_F \equiv \hat{\tau}(\boldsymbol{K}_F) \tag{11}$$

only when $\boldsymbol{R} = \boldsymbol{1}$, where

$$\begin{aligned}
\boldsymbol{K}_F &= \boldsymbol{1} \\
\therefore \boldsymbol{k}_F^T &= \boldsymbol{k}_F^C \\
&\equiv \boldsymbol{k}_F \\
&= \boldsymbol{1} \\
N_F^T &\equiv N^T(\boldsymbol{K}_F) \\
= N_F^C &\equiv N^C(\boldsymbol{K}_F) \\
&\equiv N_F \\
&= n \quad (\because \text{Equations 1 and 9})
\end{aligned} \tag{12}$$

The unitwise deletion estimator (UDE) is defined as

$$\hat{\tau}_U \equiv \hat{\tau}(\boldsymbol{K}_U) \tag{13}$$

only when $N_U^A \equiv N^A(\boldsymbol{K}_U) \geq 1$, where

$$\begin{aligned} \boldsymbol{K}_U &= \boldsymbol{R} \\ \therefore \boldsymbol{k}_U^A &\equiv \boldsymbol{r}^A. \end{aligned} \tag{14}$$

Let

$$-A = \begin{cases} C \text{ when } A = T \\ T \text{ when } A = C. \end{cases}$$

The pairwise deletion estimator (PDE) is defined as

$$\hat{\tau}_P \equiv \hat{\tau}(\boldsymbol{K}_P), \tag{15}$$

only when $N_P \geq 1$, where

$$\begin{aligned} \boldsymbol{K}_P &= \boldsymbol{R}\boldsymbol{R}_{-i} \\ \therefore \boldsymbol{k}_P^A &\equiv \boldsymbol{r}^A \boldsymbol{r}_{-i}^{-A}, \\ N_P^T &\equiv N^T(\boldsymbol{K}_P) \\ = N_P^C &\equiv N^C(\boldsymbol{K}_P) \\ &\equiv N_P. \end{aligned} \tag{16}$$

The last equality follows because

$$\begin{aligned} N_P^A &= \sum_j \sum_i R_{ij} R_{-ij} X_{ij}^A \quad (\because \text{Equations 9 and 16}) \\ &= \sum_j \sum_i R_{-ij} R_{ij} X_{-ij}^{-A} \quad (\because \text{Lemma 2 (1)}) \\ &= \sum_j \sum_i R_{ij} R_{-ij} X_{ij}^{-A} \quad (\because \text{exchanging } i \text{ and } -i\text{, Lemma 1 (4)}) \\ &= N_P^{-A}. \end{aligned}$$

Obviously, $\hat{\tau}_F$ is not available in the situation of interest where some outcomes are missing ($\boldsymbol{R} \neq \boldsymbol{1}$). Rather, analysis of $\hat{\tau}_F$, whose results are already reported in Imai (2008) and Imbens & Rubin (2015, ch. 10) (without a few exceptions mentioned below), provides reference benchmarks against which this study compares properties of $\hat{\tau}_P$ and $\hat{\tau}_U$. I also define potential category indicators as $\boldsymbol{k}_F \equiv \boldsymbol{1}, \quad \boldsymbol{k}_U^T \equiv \boldsymbol{r}^T, \quad \boldsymbol{k}_U^C \equiv \boldsymbol{r}^C, \quad \boldsymbol{k}_P \equiv \boldsymbol{r}^T \boldsymbol{r}_{-i}^C$.

2.1.3. Decomposition

In order to clarify properties of estimators, we decompose potential outcomes into three components. We define the mean of potential outcome as

$$\mu^A \equiv E(\boldsymbol{y}^A|\boldsymbol{1}), \tag{17}$$

the between-pair deviation of potential outcome as

$$\begin{aligned} \boldsymbol{\beta}^A &\equiv \frac{1}{2}(\boldsymbol{y}^A + \boldsymbol{y}_{-i}^A) - \boldsymbol{\mu}^A \\ &= \frac{1}{2}(\boldsymbol{y}_{-i}^A + \boldsymbol{y}^A) - \boldsymbol{\mu}^A \\ &= \boldsymbol{\beta}_{-i}^A, \end{aligned} \tag{18}$$

and the within-pair deviation of potential outcome as

$$\begin{aligned}\boldsymbol{\omega}^A &\equiv \boldsymbol{y}^A - \frac{1}{2}(\boldsymbol{y}^A + \boldsymbol{y}^A_{-i}) \\ &= -\Big\{\boldsymbol{y}^A_{-i} - \frac{1}{2}(\boldsymbol{y}^A_{-i} + \boldsymbol{y}^A)\Big\} \\ &= -\boldsymbol{\omega}^A_{-i}.\end{aligned} \tag{19}$$

Generally, denote the sets of the generic between-pair and within-pair component of potential outcome by

$$\mathbb{B} \equiv \{\boldsymbol{\beta}|\boldsymbol{\beta} = \boldsymbol{\beta}_{-i}\} \tag{20}$$
$$\mathbb{W} \equiv \{\boldsymbol{\omega}|\boldsymbol{\omega} = -\boldsymbol{\omega}_{-i}\}, \tag{21}$$

respectively, whose elements are constant irrespective of $\boldsymbol{X}^A$ (because they are part of potential outcome). When $\boldsymbol{\beta} \in \mathbb{B}$, we denote $\beta_{\cdot j} \equiv \beta_{1j} = \beta_{2j}$.

Lemma 4 (Decomposition). *(1)*

$$\boldsymbol{y}^A = \boldsymbol{\mu}^A + \boldsymbol{\beta}^A + \boldsymbol{\omega}^A$$

(2)

$$\bar{\tau} = \mu^T - \mu^C$$

(3)

$$\boldsymbol{\tau} = \bar{\boldsymbol{\tau}} + \boldsymbol{\beta}^T - \boldsymbol{\beta}^C + \boldsymbol{\omega}^T - \boldsymbol{\omega}^C$$

Lemma 4 (3) means that the unit treatment effects can be heterogeneous.

Proof. (1)

$$\begin{aligned}\boldsymbol{y}^A &= \frac{1}{2}(\boldsymbol{y}^A + \boldsymbol{y}^A_{-i}) + \boldsymbol{\omega}^A \quad (\because \text{Equation 19}) \\ &= \boldsymbol{\mu}^A + \boldsymbol{\beta}^A + \boldsymbol{\omega}^A. \quad (\because \text{Equation 18})\end{aligned}$$

(2)

$$\begin{aligned}\bar{\tau} &= E\{\boldsymbol{\tau}|\mathbf{1}\} \quad (\because \text{Equation 7}) \\ &= E\{\boldsymbol{y}^T - \boldsymbol{y}^C|\mathbf{1}\} \quad (\because \text{Equation 6}) \\ &= \mu^T - \mu^C \quad (\because \text{Lemma 3, Equation 17})\end{aligned}$$

(3)

$$\begin{aligned}\boldsymbol{\tau} &= \boldsymbol{y}^T - \boldsymbol{y}^C \quad (\because \text{Equation 6}) \\ &= (\boldsymbol{\mu}^T + \boldsymbol{\beta}^T + \boldsymbol{\omega}^T) - (\boldsymbol{\mu}^C + \boldsymbol{\beta}^C + \boldsymbol{\omega}^C) \quad (\because \text{Lemma 4 (1)}) \\ &= \bar{\boldsymbol{\tau}} + \boldsymbol{\beta}^T - \boldsymbol{\beta}^C + \boldsymbol{\omega}^T - \boldsymbol{\omega}^C \quad (\because \text{Lemma 4 (2)})\end{aligned}$$

□

Lemma 5 (Average of Deviation). *(1) When $\boldsymbol{\omega} \in \mathbb{W}$,*

$$\sum_i \omega_{ij} = 0, \quad E(\boldsymbol{\omega}|\mathbf{1}) = 0.$$

(2)

$$\sum_j \beta^A_{\cdot j} = 0, \quad E(\boldsymbol{\beta}^A|\mathbf{1}) = 0.$$

PROOF. (1) It holds

$$\sum_i \omega_{ij} = \omega_{ij} + \omega_{-ij} \\ = \omega_{ij} - \omega_{ij} \quad (\because \boldsymbol{\omega} \in \mathbb{W}) \\ = 0.$$

It also holds

$$E(\boldsymbol{\omega}|\mathbf{1}) = \frac{\sum_j \sum_i \omega_{ij}}{\sum_j \sum_i 1} \\ = \frac{\sum_j 0}{2n} \\ = 0$$

(2) It holds

$$\sum_j \beta^A_{\cdot j} = \sum_j \left\{\frac{1}{2}(y^A_{ij} + y^A_{-ij}) - \mu^A\right\} \quad (\because \text{Equation 18}) \\ = \frac{1}{2}\sum_j \sum_i y^A_{ij} - \sum_j \mu^A \\ = \frac{1}{2} \cdot 2n\mu^A - n\mu^A \quad (\because \text{Equation 17}) \\ = 0.$$

It also holds

$$E(\boldsymbol{\beta}^A|\mathbf{1}) = \frac{\sum_j \sum_i \beta^A_{ij}}{\sum_j \sum_i 1} \\ = \frac{2\sum_j \beta^A_{\cdot j}}{2n} \\ = \frac{0}{n} \\ = 0$$

□

LEMMA 6 (ESTIMATION ERROR: FS). *(1)*

$$\hat{\tau}_F - \bar{\tau} = E(\boldsymbol{\omega}^T + \boldsymbol{\omega}^C|\boldsymbol{K}_F\boldsymbol{X}^T) \\ = -E(\boldsymbol{\omega}^T + \boldsymbol{\omega}^C|\boldsymbol{K}_F\boldsymbol{X}^C)$$

(2) When $N_P \geq 1$,

$$\hat{\tau}_P - \bar{\tau} = E(\boldsymbol{\beta}^T - \boldsymbol{\beta}^C + \boldsymbol{\omega}^T + \boldsymbol{\omega}^C|\boldsymbol{K}_P\boldsymbol{X}^T) \\ = E\{(\boldsymbol{\beta}^T - \boldsymbol{\beta}^C) - (\boldsymbol{\omega}^T + \boldsymbol{\omega}^C)|\boldsymbol{K}_P\boldsymbol{X}^C\}$$

(3) When $N^A_U \geq 1$,

$$\hat{\tau}_U - \bar{\tau} = E(\boldsymbol{\beta}^T + \boldsymbol{\omega}^T|\boldsymbol{K}_U\boldsymbol{X}^T) - E(\boldsymbol{\beta}^C + \boldsymbol{\omega}^C|\boldsymbol{K}_U\boldsymbol{X}^C).$$

PROOF. Note

$$\boldsymbol{Y}\boldsymbol{X}^A = \Big(\sum_{A'} \boldsymbol{y}^{A'}\boldsymbol{X}^{A'}\Big)\boldsymbol{X}^A \quad (\because \text{Equation 3}) \\ = \boldsymbol{y}^A\boldsymbol{X}^A \quad (\because \boldsymbol{X}^A \in \mathbb{U}, \text{Lemmas 1 (5) and 2 (2)}) \tag{22}$$

For $\boldsymbol{K}_G \in \mathbb{U}$,

$$\begin{aligned}
&E(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^A) \\
&= E(\boldsymbol{Y}\boldsymbol{X}^A|\boldsymbol{K}_G\boldsymbol{X}^A) \quad (\because \boldsymbol{X}^A \in \mathbb{U}, \text{Lemma 1 (5) and 3 (4)}) \\
&= E(\boldsymbol{y}^A\boldsymbol{X}^A|\boldsymbol{K}_G\boldsymbol{X}^A) \quad (\because \boldsymbol{X}^A \in \mathbb{U}, \text{Equation 22}) \\
&= E(\boldsymbol{y}^A|\boldsymbol{K}_G\boldsymbol{X}^A) \quad (\because \boldsymbol{X}^A \in \mathbb{U}, \text{Lemma 1 (5) and 3 (4) })
\end{aligned} \tag{23}$$

When $N_G^A \geq 1$,

$$\begin{aligned}
&\hat{\tau}_G - \bar{\tau} \\
&= E(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^T) - E(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^C) - \bar{\tau} \quad (\because N_G^A \geq 1, \text{Equation 10}) \\
&= E(\boldsymbol{y}^T|\boldsymbol{K}_G\boldsymbol{X}^T) - E(\boldsymbol{y}^C|\boldsymbol{K}_G\boldsymbol{X}^C) - \bar{\tau} \quad (\because \text{Equation 23}) \\
&= E(\boldsymbol{\mu}^T + \boldsymbol{\beta}^T + \boldsymbol{\omega}^T|\boldsymbol{K}_G\boldsymbol{X}^T) - E(\boldsymbol{\mu}^C + \boldsymbol{\beta}^C + \boldsymbol{\omega}^C|\boldsymbol{K}_G\boldsymbol{X}^C) - \bar{\tau} \quad (\because \text{Lemma 4 (1)}) \\
&= \mu^T + E(\boldsymbol{\beta}^T + \boldsymbol{\omega}^T|\boldsymbol{K}_G\boldsymbol{X}^T) - \mu^C - E(\boldsymbol{\beta}^C + \boldsymbol{\omega}^C|\boldsymbol{K}_G\boldsymbol{X}^C) - \bar{\tau} \quad (\because \text{Lemma 3 (1) and (2)}) \\
&= E(\boldsymbol{\beta}^T + \boldsymbol{\omega}^T|\boldsymbol{K}_G\boldsymbol{X}^T) - E(\boldsymbol{\beta}^C + \boldsymbol{\omega}^C|\boldsymbol{K}_G\boldsymbol{X}^C) \quad (\because \text{Lemma 4 (2)})
\end{aligned} \tag{24}$$

(3) Substitute $\boldsymbol{K}_G = \boldsymbol{K}_U$. When $N_U^A \geq 1$, it follows that $N_G^A \geq 1$ ($\because$ Equation 9) and, thus, Equation 24 is equivalent to the desired result where $\hat{\tau}_G = \hat{\tau}_U$.

(2) Note

$$\begin{aligned}
\boldsymbol{K}_P &= \boldsymbol{R}\boldsymbol{R}_{-i} \quad (\because \text{Equations 16}) \\
&= \boldsymbol{R}_{-i}\boldsymbol{R} \\
&= \boldsymbol{K}_{P,-i} \quad (\because \text{Equations 16})
\end{aligned} \tag{25}$$

It holds

$$\begin{aligned}
E(\boldsymbol{\beta}^A + \boldsymbol{\omega}^A|\boldsymbol{K}_P\boldsymbol{X}^A) &= E(\boldsymbol{\beta}_{-i}^A - \boldsymbol{\omega}_{-i}^A|\boldsymbol{K}_{P,-i}\boldsymbol{X}_{-i}^{-A}) \quad (\because \text{Equations 18, 19, and 25, Lemma 2 (1)}) \\
&= E(\boldsymbol{\beta}^A - \boldsymbol{\omega}^A|\boldsymbol{K}_P\boldsymbol{X}^{-A}) \quad (\because \text{exchanging } i \text{ and } -i)
\end{aligned} \tag{26}$$

Substitute $\boldsymbol{K}_G = \boldsymbol{K}_P$. When $N_P \geq 1$, it follows that $N_G^A \geq 1$ ($\because$ Equations 9 and 16) and, thus,

$$\begin{aligned}
\hat{\tau}_P - \bar{\tau} &= E(\boldsymbol{\beta}^T + \boldsymbol{\omega}^T|\boldsymbol{K}_P\boldsymbol{X}^T) - E(\boldsymbol{\beta}^C + \boldsymbol{\omega}^C|\boldsymbol{K}_P\boldsymbol{X}^C) \quad (\because \text{Equation 24}) \\
&= E(\boldsymbol{\beta}^T + \boldsymbol{\omega}^T|\boldsymbol{K}_P\boldsymbol{X}^T) - E(\boldsymbol{\beta}^C - \boldsymbol{\omega}^C|\boldsymbol{K}_P\boldsymbol{X}^T) \quad (\because \text{Equation 26}) \\
&= E(\boldsymbol{\beta}^T - \boldsymbol{\beta}^C + \boldsymbol{\omega}^T + \boldsymbol{\omega}^C|\boldsymbol{K}_P\boldsymbol{X}^T) \quad (\because \text{Lemma 3 (1)})
\end{aligned}$$

Similarly,

$$\hat{\tau}_P - \bar{\tau} = E\{(\boldsymbol{\beta}^T - \boldsymbol{\beta}^C) - (\boldsymbol{\omega}^T + \boldsymbol{\omega}^C)|\boldsymbol{K}_P\boldsymbol{X}^C\}.$$

(1)

$$\begin{aligned}
E(\boldsymbol{\beta}^A|\boldsymbol{K}_F\boldsymbol{X}^{A'}) &= \frac{\sum_j \sum_i K_{F,ij} X_{ij}^{A'} \beta_{ij}^A}{\sum_j \sum_i K_{F.ij} X_{ij}^{A'}} \quad (\because \boldsymbol{K}_F\boldsymbol{X}^{A'} \neq \boldsymbol{0}, \text{Equation 4}) \\
&= \frac{\sum_j 1 \cdot \beta_{\cdot j}^A \sum_i X_{ij}^{A'}}{\sum_j 1 \sum_i X_{ij}^{A'}} \quad (\because \text{Equations 12 and 18}) \\
&= \frac{\sum_j \beta_{\cdot j}^A \cdot 1}{\sum_j 1} \quad (\because \text{Equation 2}) \\
&= 0. \quad (\because \text{Lemma 5 (2)})
\end{aligned} \tag{27}$$

When we substitute $\boldsymbol{R} = \boldsymbol{1}$, it follows that $\boldsymbol{K}_P = \boldsymbol{K}_F, \hat{\tau}_P = \hat{\tau}_F, N_P = N_F = n \geq 2 > 1$, and thus

$$\begin{aligned}\hat{\tau}_F - \bar{\tau} &= E(\boldsymbol{\beta}^T - \boldsymbol{\beta}^C + \boldsymbol{\omega}^T + \boldsymbol{\omega}^C | \boldsymbol{K}_F \boldsymbol{X}^T) \quad (\because \text{Lemma 6 (2)}) \\ &= E(\boldsymbol{\omega}^T + \boldsymbol{\omega}^C | \boldsymbol{K}_F \boldsymbol{X}^T) \quad (\because \text{Equations 3 (1) and 27})\end{aligned}$$

Similarly,

$$\hat{\tau}_F - \bar{\tau} = -E(\boldsymbol{\omega}^T + \boldsymbol{\omega}^C | \boldsymbol{K}_F \boldsymbol{X}^C)$$

□

2.1.4. Assumption

Mandatory Assumption of Random Treatment Assignment. We suppose that $\boldsymbol{y}^T$, $\boldsymbol{y}^C$, $\boldsymbol{r}^T$, $\boldsymbol{r}^C$, $\boldsymbol{k}_G^T$, and $\boldsymbol{k}_G^C$ (or variables and vectors which are denoted by lower-case letters) are fixed but $\boldsymbol{X}^A$ (and, thus, $\boldsymbol{Y}$, $\boldsymbol{R}$, and $\boldsymbol{K}_G$ as well as variables and vectors which are denoted by upper-case letters) is stochastic. Let $\mathbb{Y}_{\max}$ be the set of values Y_{ij}'s can take (or $\mathbb{Y}_{\max} = \mathbb{Y}^*$, which will appear in the next section). We assume

- Ignorability of treatment assignment: for any $\boldsymbol{y}^{(1)}, \boldsymbol{y}^{(2)} \in \mathbb{Y}_{\max}$ and any $\boldsymbol{r}^{(1)}, \boldsymbol{r}^{(2)} \in \mathbb{U}$,

$$\Pr\{\boldsymbol{X}^A | \boldsymbol{y}^T = \boldsymbol{y}^{(1)}, \boldsymbol{y}^C = \boldsymbol{y}^{(2)}, \boldsymbol{r}^T = \boldsymbol{r}^{(1)}, \boldsymbol{r}^C = \boldsymbol{r}^{(2)}\} = \Pr(\boldsymbol{X}^A) \tag{28}$$

- Independence of treatment assignment: for any $i \in \{1, 2\}$,

$$\Pr(\boldsymbol{X}^A) = \prod_j \Pr(X_{ij}^A) \tag{29}$$

- Isoprobability of treatment assignment:

$$\Pr(X_{ij}^A = 1) = \Pr(X_{ij}^A = 0) = \frac{1}{2} \tag{30}$$

My conjecture is that the assumption of isoprobability of treatment assignment will not be essential. Let $\pi_{ij}^A \equiv \Pr(X_{ij}^A = 1)$. Instead of Equations 18 and 19, probably we only have to redefine

$$\boldsymbol{\beta}^A \equiv (\boldsymbol{\pi}^A \boldsymbol{y}^A + \boldsymbol{\pi}_{-i}^A \boldsymbol{y}_{-i}^A) - \boldsymbol{\mu}^A$$

and

$$\boldsymbol{\omega}_{ij}^A \equiv \boldsymbol{y}^A - (\boldsymbol{\pi}^A \boldsymbol{y}^A + \boldsymbol{\pi}_{-i}^A \boldsymbol{y}_{-i}^A)$$

so that we retain most results below.

Optional Assumption of Potential Response Match. For $\boldsymbol{K}_G \in \mathbb{U}$, we define the following condition:

Condition 1 (Matched Attrition: FS).

$$\boldsymbol{k}_G^A = \boldsymbol{k}_{G,-i}^A.$$

Under Condition 1, define

$$\begin{aligned}n_G^A &\equiv n(\boldsymbol{k}_G^A) \\ &\equiv \sum_j k_{G,\cdot j}^A\end{aligned} \tag{31}$$

which is constant irrespective of $\boldsymbol{X}^A$ because $k_{G,\cdot j}^A$'s are constant irrespective of $\boldsymbol{X}^A$.

While the above assumptions of random treatment assignment is mandatory, the following five assumptions of potential response (and outcome) are optional; we may invoke one or more of them in some lemmas and propositions below.

When $\boldsymbol{K}_G = \boldsymbol{K}_F$, Condition 1 always holds and $\mathbb{X}_{\text{def}}(\hat{\tau}_G) = \mathbb{X}_{\max}$ is equivalent to

Assumption 1 (No Attrition: FS).

$$\boldsymbol{r}^T = \boldsymbol{r}^C = \mathbf{1}.$$

When $\boldsymbol{K}_G = \boldsymbol{K}_U$, Condition 1 is equivalent to

Assumption 2 (Unitwise Matched Attrition: FS).

$$\boldsymbol{r}^T = \boldsymbol{r}^T_{-i},$$
$$\boldsymbol{r}^C = \boldsymbol{r}^C_{-i}.$$

When $\boldsymbol{K}_G = \boldsymbol{K}_P$, Condition 1 is equivalent to

Assumption 3 (Pairwise Matched Attrition: FS).

$$\boldsymbol{r}^T \boldsymbol{r}^C_{-i} = \boldsymbol{r}^T_{-i} \boldsymbol{r}^C.$$

Optional Assumption of Ignorable Potential Response. Define conditional empirical distribution by

$$P(\boldsymbol{Q} = q|\boldsymbol{Z} = z) \equiv \frac{\sum_j \sum_i I(Q_{ij} = q) I(Z_{ij} = z)}{\sum_j \sum_i I(Z_{ij} = z)}, \tag{32}$$

only when $\sum_j \sum_i I(Z_{ij} = z) \geq 1$, where

$$I(Q_{ij} = q) = \begin{cases} 1 \text{ when } Q_{ij} = q \\ 0 \text{ otherwise.} \end{cases}$$

In particular, for $\boldsymbol{U} \in \mathbb{U}$, it follows that $I(U_{ij} = 1) = U_{ij}$. Define (marginal) empirical distribution by

$$\begin{aligned} P(\boldsymbol{Q} = q) &\equiv P(\boldsymbol{Q} = q|\mathbf{1} = 1) \\ &= \frac{\sum_j \sum_i I(Q_{ij} = q)}{2n}, \end{aligned} \tag{33}$$

which can be always defined.

Let $\mathbb{Y}^A \subseteq \mathbb{Y}_{\max}$ be the sets of all values y^A_{ij} take, namely,

$$\mathbb{Y}^A = \{y^{(h)}|h \in \{1, \ldots, h_{\max}\}, \forall h, \exists ij, y^{(h)} = y^A_{ij}, \forall ij, \exists h, y^{(h)} = y^A_{ij}\}$$

We define the following condition as well:

Condition 2 (Ignorable Attrition: FS). *For any* $y^A \in \mathbb{Y}^A$,

$$P(\boldsymbol{k}^A_G = 1|\boldsymbol{y}^A = y^A) = P(\boldsymbol{k}^A_G = 1).$$

Note that $\sum_j \sum_i I(y^A_{ij} = y^A) \geq 1$ because $y^A \in \mathbb{Y}^A$.

In general, when $\boldsymbol{U} \in \mathbb{U}, u \in \{0, 1\}$, and $\sum_j \sum_i I(Z_{ij} = z) \geq 1$,

$$\begin{aligned} P(\boldsymbol{U} = 1 - u|\boldsymbol{Z} = z) &= \frac{\sum_j \sum_i I(U_{ij} = 1 - u) I(Z_{ij} = z)}{\sum_j \sum_i I(Z_{ij} = z)} \quad (\because \text{Equation 32}) \\ &= \frac{\sum_j \sum_i \{1 - I(U_{ij} = u)\} I(Z_{ij} = z)}{\sum_j \sum_i I(Z_{ij} = z)} \quad (\because \boldsymbol{U} \in \mathbb{U}, u \in \{0, 1\}) \\ &= 1 - \frac{\sum_j \sum_i I(U_{ij} = u) I(Z_{ij} = z)}{\sum_j \sum_i I(Z_{ij} = z)} \\ &= 1 - P(\boldsymbol{U} = u|\boldsymbol{Z} = z) \quad (\because \text{Equation 32}) \end{aligned} \tag{34}$$

Therefore, under Condition 2, for any $y^A \in \mathbb{Y}$,

$$\begin{aligned} P(\boldsymbol{k}_G^A = 0|\boldsymbol{y}^A = y^A) &= 1 - P(\boldsymbol{k}_G^A = 1|\boldsymbol{y}^A = y^A) \quad (\because \boldsymbol{K}_G \in \mathbb{U}, \text{Equation 34}) \\ &= 1 - P(\boldsymbol{k}_G^A = 1) \quad (\because \text{Condition 2}) \\ &= P(\boldsymbol{k}_G^A = 0) \quad (\because \boldsymbol{K}_G \in \mathbb{U}, \text{Equations 33 and 34}) \end{aligned}$$

When $\boldsymbol{K}_G = \boldsymbol{K}_F$, Condition 2 always holds.
When $\boldsymbol{K}_G = \boldsymbol{K}_U$, Condition 2 is equivalent to

Assumption 4 (Unitwise Ignorable Attrition : FS). *For any $y^T \in \mathbb{Y}^T$ and $y^C \in \mathbb{Y}^C$,*

$$\begin{aligned} P(\boldsymbol{r}^T = 1|\boldsymbol{y}^T = y^T) = P(\boldsymbol{r}^T = 1) \\ P(\boldsymbol{r}^C = 1|\boldsymbol{y}^C = y^C) = P(\boldsymbol{r}^C = 1). \end{aligned}$$

Note that it always hold

$$\begin{aligned} P(\boldsymbol{Q} = q) &= \frac{1}{2n}\sum_j\sum_i I(Q_{ij} = q) \quad (\because \text{Equation 33}) \\ &= \frac{1}{2n}\sum_j\sum_i I(Q_{-ij} = q) \quad (\because \text{Lemma 1 (4)}) \\ &= P(\boldsymbol{Q}_{-i} = q) \quad (\because \text{Equation 33}) \end{aligned}$$

Thus, when $\boldsymbol{K}_G = \boldsymbol{K}_P$, Condition 2 is equivalent to

Assumption 5 (Pairwise Ignorable Attrition : FS). *For any $y^T \in \mathbb{Y}^T$ and $y^C \in \mathbb{Y}^C$,*

$$\begin{aligned} & P(\boldsymbol{r}^T\boldsymbol{r}_{-i}^C = 1|\boldsymbol{y}^T = y^T) = P(\boldsymbol{r}^T\boldsymbol{r}_{-i}^C = 1) \\ = \ & P(\boldsymbol{r}_{-i}^T\boldsymbol{r}^C = 1|\boldsymbol{y}^C = y^C) = P(\boldsymbol{r}_{-i}^T\boldsymbol{r}^C = 1) \end{aligned}$$

Unfortunately, it is unusual that attrition is ignorable in either sense. For example, in the evaluation of a medication, attrition due to poor health including death may be increased in the control group (r_{ij}^C is more likely to be 0 as y_{ij}^C becomes smaller), but "attrition due to the fact that the subject feel[s] healthier and stop[s] complying with the experimental protocol may be increased in the treatment group" (r_{ij}^T is more likely to be 0 as y_{ij}^T becomes larger) (Duflo et al. 2008, p. 3943). Since the main object of this study is non-ignorable attrition, Assumptions 4 and 5 simply offer benchmarks against which this study compares cases of non-ignorable attrition.

Barnard et al. (2003, 304) assume (latent) ignorability where "potential outcomes are independent of missingness" (given observed covariates conditional on the compliance strata). Their assumption means that, in terms of my framework, for any $y^T \in \mathbb{Y}^T$, $y^C \in \mathbb{Y}^C$,and $r^{(1)}, r^{(2)} \in \{0, 1\}$,

$$P(\boldsymbol{r}^T = r^{(1)}, \boldsymbol{r}^C = r^{(2)}|\boldsymbol{y}^T = y^{T)}, \boldsymbol{y}^C = y^C) = P(\boldsymbol{r}^T = r^{(1)}, \boldsymbol{r}^C = r^{(2)})$$

which is stronger than Assumption 4 or 5.

Lemma 7 (Expectation Representation of Weighted Mean). *When $\boldsymbol{Z} \in \mathbb{U}, \boldsymbol{Z} \neq \boldsymbol{0}$,*

$$E(\boldsymbol{Q}|\boldsymbol{Z}) = \sum_q P(\boldsymbol{Q} = q|\boldsymbol{Z} = 1)q$$

where

$$\begin{aligned} &\sum_q \equiv \sum_{q\in\mathbb{Q}} \\ &\mathbb{Q} \supseteq \{q^{(h)}|h \in \{1, \ldots, h_{\max}\}, \forall h, \exists ij, q^{(h)} = Q_{ij}, \forall ij, \exists h, q^{(h)} = Q_{ij}\}. \end{aligned}$$

PROOF.

$$\begin{aligned}
E(\boldsymbol{Q}|\boldsymbol{Z}) &= \frac{\sum_j \sum_i Z_{ij} Q_{ij}}{\sum_j \sum_i Z_{ij}} \quad (\because \text{Equation 4, } \boldsymbol{Z} \neq \boldsymbol{0}) \\
&= \frac{\sum_j \sum_i I(Z_{ij}=1) Q_{ij}}{\sum_j \sum_i I(Z_{ij}=1)} \quad (\because \boldsymbol{Z} \in \mathbb{U}) \\
&= \frac{\sum_j \sum_i I(Z_{ij}=1) \sum_q I(Q_{ij}=q) q}{\sum_j \sum_i I(Z_{ij}=1)} \\
&= \sum_q \frac{\sum_j \sum_i I(Z_{ij}=1) I(Q_{ij}=q)}{\sum_j \sum_i I(Z_{ij}=1)} q \\
&= \sum_q P(\boldsymbol{Q}=q|\boldsymbol{Z}=1) q \quad (\because \text{Equation 32})
\end{aligned}$$

□

LEMMA 8 (REALIZED CATEGORY AND POTENTIAL RESPONSE).

$$\boldsymbol{K}_G \boldsymbol{X}^A = \boldsymbol{k}_G^A \boldsymbol{X}^A$$

PROOF.

$$\begin{aligned}
\boldsymbol{K}_G \boldsymbol{X}^A &= (\boldsymbol{k}_G^A \boldsymbol{X}^A + \boldsymbol{k}_G^{-A} \boldsymbol{X}^{-A}) \boldsymbol{X}^A \quad (\because \text{Equation 8}) \\
&= \boldsymbol{k}_G^A \boldsymbol{X}^A + \boldsymbol{k}_G^{-A} \cdot \boldsymbol{0} \quad (\because \boldsymbol{X}^A \in \mathbb{U}, \text{Lemma 1 (5) and 2 (2)}) \\
&= \boldsymbol{k}_G^A \boldsymbol{X}^A
\end{aligned}$$

□

LEMMA 9 (IMPLICATION OF ASSUMPTION). *(1) Under Assumption 3,*

$$\begin{aligned}
\boldsymbol{k}_P^T &= \boldsymbol{k}_P^C \\
&\equiv \boldsymbol{k}_P \\
N_P &= n(\boldsymbol{k}_P) \\
&\equiv n_P.
\end{aligned}$$

(2) Under Assumption 2, in addition to Lemma 9 (1),

$$\begin{aligned}
\prod_A \boldsymbol{k}_U^A &= \boldsymbol{k}_P \\
N_U^A &= n(\boldsymbol{k}_U^A) \\
&\equiv n_U^A
\end{aligned}$$

(3) Under Assumption 4, for any $y \in \mathbb{Y}$, $r \in \{0,1\}, \boldsymbol{r}^A \neq \boldsymbol{1} - \boldsymbol{r}$,

$$P(\boldsymbol{y}^A = y | \boldsymbol{r}^A = r) = P(\boldsymbol{y}^A = y)$$

(4) Under Assumption 5, for any $y \in \mathbb{Y}$, $r \in \{0,1\}, \boldsymbol{r}^A \boldsymbol{r}_{-i}^{-A} \neq \boldsymbol{1} - \boldsymbol{r}$,

$$P(\boldsymbol{y}^A = y | \boldsymbol{r}^A \boldsymbol{r}_{-i}^{-A} = r) = P(\boldsymbol{y}^A = y)$$

(5) Under Assumption 2, Assumption 3 holds.
(6) Under Assumption 1, Assumption 2 and 3 hold.

When $\boldsymbol{K}_G = \boldsymbol{K}_F$, Condition 1 always holds. Thus, I can define

$$\begin{aligned}
n_F &\equiv n(\boldsymbol{k}_F) \\
&= n \quad (\because \text{Equations 12 and 31})
\end{aligned} \tag{35}$$

PROOF. (1)–(2) Suppose that $\boldsymbol{K}_G \in \mathbb{U}$ satisfies Condition 1.

$$
\begin{aligned}
N_G^A &= \sum_j \sum_i K_{G,ij} X_{ij}^A \quad (\because \text{Equation 9}) \\
&= \sum_j \sum_i k_{G,ij}^A X_{ij}^A \quad (\because \text{Lemma 8}) \\
&= \sum_j k_{G,\cdot j}^A \sum_i X_{ij}^A \quad (\because \text{Condition 1}) \\
&= \sum_j k_{G,\cdot j}^A \quad (\because \text{Equation 2}) \\
&= n(\boldsymbol{k}_G^A) \quad (\because \text{Equation 31})
\end{aligned}
\tag{36}
$$

(1)

$$
\begin{aligned}
\boldsymbol{k}_P^A &= \boldsymbol{r}^A \boldsymbol{r}_{-i}^{-A} \quad (\because \text{Equation 16}) \\
&= \boldsymbol{r}_{-i}^A \boldsymbol{r}^{-A} \quad (\because \text{Assumption 3 (1)}) \\
&= \boldsymbol{k}_P^{-A} \quad (\because \text{Equation 16}) \\
&\equiv \boldsymbol{k}_P.
\end{aligned}
\tag{37}
$$

Substitute $\boldsymbol{K}_G = \boldsymbol{K}_P$. Under Assumption 3, it follows that Condition 1 holds. It follows that

$$
\begin{aligned}
N_P &= N(\boldsymbol{K}_P^A) \quad (\because \text{Equation 16}) \\
&= n(\boldsymbol{k}_P^A) \quad (\because \text{Equations 9, 31, and 36}) \\
&= n(\boldsymbol{k}_P) \quad (\because \text{Equation 37}) \\
&\equiv n_P.
\end{aligned}
$$

(2) Thanks to Lemma 9 (5), Lemma 9 (1) holds. It also holds

$$
\begin{aligned}
\prod_A \boldsymbol{k}_U^A &= \boldsymbol{r}^T \boldsymbol{r}^C \quad (\because \text{Equation 14}) \\
&= \boldsymbol{r}^T \boldsymbol{r}_{-i}^C \quad (\because \text{Assumption 2}) \\
&= \boldsymbol{k}_P \quad (\because \text{Equation 16, Lemma 9 (1)})
\end{aligned}
$$

Substitute $\boldsymbol{K}_G = \boldsymbol{K}_U$. Under Assumption 2, it follows that Condition 1 holds. According to Equation 36, it follows $N_U^A = n_U^A$.

(3)–(4) Denote $\sum_y \equiv \sum_{y \in \mathbb{Y}}$. Under Condition 2, $y \in \mathbb{Y}$, and $\boldsymbol{k}_G^A \neq \boldsymbol{0}$, it follows that, in the spirit of Bayes' Rule,

$$
\begin{aligned}
&P(\boldsymbol{y}^A = y | \boldsymbol{k}_G^A = 1) \\
&= \frac{\sum_j \sum_i k_{G,ij}^A I(y_{ij}^A = y)}{\sum_j \sum_i k_{G,ij}^A} \quad (\because \text{Equation 32}, \boldsymbol{k}_G^A \in \mathbb{U}, \boldsymbol{k}_G^A \neq \boldsymbol{0}, \sum_j \sum_i I(k_{G,ij}^A = 1) \geq 1) \\
&= \frac{\sum_j \sum_i k_{G,ij}^A I(y_{ij}^A = y)}{\sum_j \sum_i k_{G,ij}^A \sum_{y'} I(y_{ij}^A = y')} \quad (\because \sum_{y'} I(y_{ij}^A = y') = 1) \\
&= \frac{\sum_j \sum_i k_{G,ij}^A I(y_{ij}^A = y)}{\sum_{y'} \sum_j \sum_i k_{G,ij}^A I(y_{ij}^A = y')} \\
&= \Big[\Big\{ \sum_j \sum_i I(y_{ij}^A = y) \Big\} \frac{\sum_j \sum_i k_{G,ij}^A I(y_{ij}^A = y)}{\sum_j \sum_i I(y_{ij}^A = y)} \Big] \\
&\quad \div \sum_{y'} \Big[\Big\{ \sum_j \sum_i I(y_{ij}^A = y') \Big\} \frac{\sum_j \sum_i k_{G,ij}^A I(y_{ij}^A = y')}{\sum_j \sum_i I(y_{ij}^A = y')} \Big] \\
&\quad (\because y, y' \in \mathbb{Y}, \sum_j \sum_i I(y_{ij}^A = y) \geq 1, \sum_j \sum_i I(y_{ij}^A = y') \geq 1) \\
&= \frac{\sum_j \sum_i I(y_{ij}^A = y) P(\boldsymbol{k}_G^A = 1 | y_{ij}^A = y)}{\sum_{y'} \sum_j \sum_i I(y_{ij}^A = y') P(\boldsymbol{k}_G^A = 1 | y_{ij}^A = y')} \quad (\because \text{Equation 32}, \boldsymbol{k}_G^A \in \mathbb{U}) \\
&= \frac{\sum_j \sum_i I(y_{ij}^A = y) P(\boldsymbol{k}_G^A = 1)}{\sum_{y'} \sum_j \sum_i I(y_{ij}^A = y') P(\boldsymbol{k}_G^A = 1)} \quad (\because \text{Condition 2}) \\
&= \frac{P(\boldsymbol{k}_G^A = 1) \sum_j \sum_i I(y_{ij}^A = y)}{P(\boldsymbol{k}_G^A = 1) \sum_j \sum_i \sum_{y'} I(y_{ij}^A = y')} \\
&= \frac{\sum_j \sum_i 1 \cdot I(y_{ij}^A = y)}{\sum_j \sum_i 1} \\
&= P(\boldsymbol{y}^A = y). \quad (\because \text{Equation 33})
\end{aligned} \tag{38}
$$

When Condition 2 holds for $\boldsymbol{K}_G$, it also holds for $\boldsymbol{K}'_G \equiv \boldsymbol{1} - \boldsymbol{K}_G$ because

$$
\begin{aligned}
P(\boldsymbol{k}_G^{A\prime} = r | \boldsymbol{y}^A = y) &= P(\boldsymbol{1} - \boldsymbol{k}_G^A = r | \boldsymbol{y}^A = y) \quad (\because \boldsymbol{K}'_G = \boldsymbol{1} - \boldsymbol{K}_G) \\
&= P(\boldsymbol{k}_G^A = 1 - r | \boldsymbol{y}^A = y) \quad (\because \boldsymbol{K}_G \in \mathbb{U}) \\
&= P(\boldsymbol{k}_G^A = 1 - r) \quad (\because \text{Condition 2}, \ 1 - r \in \{0, 1\}) \\
&= P(\boldsymbol{1} - \boldsymbol{k}_G^A = r) \quad (\because \boldsymbol{K}_G \in \mathbb{U}) \\
&= P(\boldsymbol{k}_G^{A\prime} = r) \quad (\because \boldsymbol{K}'_G = \boldsymbol{1} - \boldsymbol{K}_G)
\end{aligned}
$$

Thus, when $\boldsymbol{k}_G^A \neq \boldsymbol{1}$,

$$
\begin{aligned}
P(\boldsymbol{y}^A = y | \boldsymbol{k}_G^A = 0) &= P(\boldsymbol{y}^A = y | \boldsymbol{1} - \boldsymbol{k}_G^A = 1) \quad (\because \boldsymbol{K}_G \in \mathbb{U}) \\
&= P(\boldsymbol{y}^A = y | \boldsymbol{k}_G^{A\prime} = 1) \quad (\because \boldsymbol{K}'_G \equiv \boldsymbol{1} - \boldsymbol{K}_G) \\
&= P(\boldsymbol{y}^A = y) \quad (\because \text{Condition 2, Equation 38})
\end{aligned} \tag{39}
$$

(3) Substitute $\boldsymbol{K}_G = \boldsymbol{K}_U$. Under Assumption 4, it follows that Condition 2 holds. Equations 38 and 39 are equivalent to the desired results.

(4) Substitute $\boldsymbol{K}_G = \boldsymbol{K}_P$. Under Assumption 5, it follows that Condition 2 holds. Equations 38 and 39 are equivalent to the desired results.

(5) and (6) Obvious.

□

Note that Frangakis & Rubin (1999, 369) make Lemma 9 (3) (conditoned on the latent compliance indicator) their assumption, rather than lemma derived from it. What matters is that, under (latent) ignorability, "[p]otential outcomes and associated potential non-response indicators are independent" (within each level of the latent compliance covariate).

2.2. Bias

For the generic function of the treatment assignment vector, $f(\boldsymbol{X}^A)$, the expectation operator is defined as

$$\mathbb{E}\{f(\boldsymbol{X}^A)\} \equiv \sum_{\boldsymbol{x}\in\mathbb{X}_{\text{def}}(f)} \Pr\{\boldsymbol{X}^A = \boldsymbol{x} | \boldsymbol{X}^A \in \mathbb{X}_{\text{def}}(f)\} f(\boldsymbol{x}) \tag{40}$$

only when the assignment space is not empty, that is,

$$\begin{aligned}\mathbb{X}_{\text{def}}(f) &\equiv \{\boldsymbol{x}|\boldsymbol{x}\in\mathbb{X}_{\max}, f(\boldsymbol{x}) \text{ can be defined}\}\\ &\neq \emptyset,\end{aligned}$$

where

$$\mathbb{X}_{\max} \equiv \Big\{\boldsymbol{x}\Big|\boldsymbol{x}\in\mathbb{U}, \sum_i x_{ij} = 1\Big\}.$$

Specifically, when $\mathbb{X}_{\text{def}}(f) = \mathbb{X}_{\max}$,

$$\begin{aligned}\mathbb{E}\{f(\boldsymbol{X}^A)\} &= \sum_{\boldsymbol{x}\in\mathbb{X}_{\max}} \Pr(\boldsymbol{X}^A = \boldsymbol{x}) f(\boldsymbol{x})\\ &= \sum_{x_{11}=0}^{1}\sum_{x_{12}=0}^{1}\cdots\sum_{x_{1n}=0}^{1} \Pr(\boldsymbol{X}^A = \boldsymbol{x}) f(\boldsymbol{x})\\ &= \sum_{x_{11}=0}^{1}\sum_{x_{12}=0}^{1}\cdots\sum_{x_{1n}=0}^{1} \Big\{\prod_j \Pr(X_{1j}^A = x_{1j})\Big\} f(\boldsymbol{x}) \quad (\because \text{Equation 29})\\ &= \sum_{x_{11}=0}^{1}\sum_{x_{12}=0}^{1}\cdots\sum_{x_{1n}=0}^{1} \Big(\frac{1}{2}\Big)^n f(\boldsymbol{x}) \quad (\because \text{Equation 30})\end{aligned} \tag{41}$$

When $\boldsymbol{X}^A = \boldsymbol{x}^{(0)}$ is realized and $f(\boldsymbol{x}^{(0)})$ can be defined, $\mathbb{X}_{\text{def}}(f) \neq \emptyset$ ($\because \boldsymbol{x}^{(0)} \in \mathbb{X}_{\text{def}}(f)$) and, therefore, $\mathbb{E}\{f(\boldsymbol{X}^A)\}$ can be defined.

LEMMA 10 (ARITHMETIC OF EXPECTATION). *(1) When $\mathbb{X}_{def}(\sum_h f^{(h)}) = \mathbb{X}_{def}(f^{(h')})$ for any $h' \in \{1, 2, \dots, h_{\max}\}$,*

$$\mathbb{E}\Big\{\sum_h f^{(h)}(\boldsymbol{X}^A)\Big\} = \sum_h \mathbb{E}\{f^{(h)}(\boldsymbol{X}^A)\}$$

(2)

$$\mathbb{E}\{qf(\boldsymbol{X}^A)\} = q\mathbb{E}\{f(\boldsymbol{X}^A)\}$$

PROOF. (1)

$$\begin{aligned}&\mathbb{E}\Big\{\sum_h f^{(h)}(\boldsymbol{X}^A)\Big\}\\ &= \sum_{\boldsymbol{x}\in\mathbb{X}_{\text{def}}(\sum_h f^{(h)})} \Pr\Big\{\boldsymbol{X}^A = \boldsymbol{x}\Big|\boldsymbol{X}^A \in \mathbb{X}_{\text{def}}\Big(\sum_h f^{(h)}\Big)\Big\} \sum_h f^{(h)}(\boldsymbol{x})\\ &= \sum_h \sum_{\boldsymbol{x}\in\mathbb{X}_{\text{def}}(f^{(h)})} \Pr\{\boldsymbol{X}^A = \boldsymbol{x}|\boldsymbol{X}^A \in \mathbb{X}_{\text{def}}(f^{(h)})\} f^{(h)}(\boldsymbol{x}) \quad \Big(\because \mathbb{X}_{\text{def}}\Big(\sum_h f^{(h)}\Big) = \mathbb{X}_{\text{def}}(f^{(h')})\Big)\\ &= \sum_h \mathbb{E}\{f^{(h)}(\boldsymbol{X}^A)\}\end{aligned}$$

(2) Note that $f'(\boldsymbol{X}^A) \equiv qf(\boldsymbol{X}^A)$ can be defined if and only if $f(\boldsymbol{X}^A)$ can be defined. Thus, $\mathbb{X}_{\text{def}}(f') = \mathbb{X}_{\text{def}}(f)$. It follows

$$\begin{aligned}
\mathbb{E}\{qf(\boldsymbol{X}^A)\} &= \sum_{\boldsymbol{x}\in\mathbb{X}_{\text{def}}(f')} \Pr\{\boldsymbol{X}^A = \boldsymbol{x} | \boldsymbol{X}^A \in \mathbb{X}_{\text{def}}(f')\} qf(\boldsymbol{x}) \\
&= q \sum_{\boldsymbol{x}\in\mathbb{X}_{\text{def}}(f)} \Pr\{\boldsymbol{X}^A = \boldsymbol{x} | \boldsymbol{X}^A \in \mathbb{X}_{\text{def}}(f)\} f(\boldsymbol{x}) \\
&\quad (\because q \text{ is constant irrespective of } \boldsymbol{X}^A, \mathbb{X}_{\text{def}}(f') = \mathbb{X}_{\text{def}}(f)) \\
&= q\mathbb{E}\{f(\boldsymbol{X}^A)\}
\end{aligned}$$

□

LEMMA 11 (FULL ASSIGNMENT SPACE). *(1)*

$$\mathbb{E}(X_{ij}^A) = \frac{1}{2}$$

(2) For $j' \neq j$ *and* $x \in \{0, 1\}$*,*

$$\mathbb{E}(X_{ij}^A | X_{i'j'}^A = x) = \frac{1}{2}$$

PROOF. (1)

$$\begin{aligned}
\mathbb{E}(X_{ij}^A) &= \sum_{x_{i1}=0}^{1} \sum_{x_{i2}=0}^{1} \cdots \sum_{x_{in}=0}^{1} \Big\{ \prod_{j'} \Pr(X_{ij'}^A = x_{ij'}) \Big\} x_{ij} \quad (\because \mathbb{X}_{\text{def}}(X_{ij}^A) = \mathbb{X}_{\max}, \text{ Equation 41}) \\
&= \sum_{x_{i1}=0}^{1} \cdots \sum_{x_{i(j''-1)}=0}^{1} \sum_{x_{i(j''+1)}=0}^{1} \cdots \sum_{x_{in}=0}^{1} \Big\{ \sum_{x_{ij''}=0}^{1} \Pr(X_{ij''}^A = x_{ij''}) \Big\} \Big\{ \prod_{j' \neq j''} \Pr(X_{ij'}^A = x_{ij'}) \Big\} x_{ij} \\
&\quad (\text{where } j'' \neq j) \\
&= \sum_{x_{i1}=0}^{1} \cdots \sum_{x_{i(j''-1)}=0}^{1} \sum_{x_{i(j''+1)}=0}^{1} \cdots \sum_{x_{in}=0}^{1} \Big\{ \prod_{j' \neq j''} \Pr(X_{ij'}^A = x_{ij'}) \Big\} x_{ij} \\
&\quad (\because \sum_{x_{ij''}=0}^{1} \Pr(X_{ij''}^A = x_{ij''}) = 1) \\
&= \sum_{x_{ij}=0}^{1} \Pr(X_{ij}^A = x_{ij}) x_{ij} \quad (\because \text{repeating for all } j'\text{'s except } j) \\
&= \Pr(X_{ij}^A = 1) \\
&= \frac{1}{2} \quad (\because \text{Equation 30})
\end{aligned}$$

(2) The third line of the above equalities is equal to $\mathbb{E}(X_{ij}^A | X_{i'j''}^A = x)$ where $x \in \{0, 1\}$. The desired result immediately follows.

□

LEMMA 12 (EXPECTATION OF WEIGHTED MEAN). *Suppose* $\boldsymbol{\omega} \in \mathbb{W}, \boldsymbol{\beta} \in \mathbb{B}$, *and* $\boldsymbol{K}_G \in \mathbb{U}$ *satisfies Condition 1.*

(1)

$$\begin{aligned}
\mathbb{E}\{E(\boldsymbol{\omega} | \boldsymbol{K}_G \boldsymbol{X}^A)\} &= \mathbb{E}\{E(\boldsymbol{\omega} | \boldsymbol{k}_G^A \boldsymbol{X}^A)\} \\
&= E(\boldsymbol{\omega} | \boldsymbol{k}_G^A) \\
&= 0.
\end{aligned}$$

(2)

$$\begin{aligned}\mathbb{E}\{E(\boldsymbol{\beta}|\boldsymbol{K}_G\boldsymbol{X}^A)\} &= \mathbb{E}\{E(\boldsymbol{\beta}|\boldsymbol{k}_G^A\boldsymbol{X}^A)\} \\ &= E(\boldsymbol{\beta}|\boldsymbol{k}_G^A\boldsymbol{X}^A) \\ &= E(\boldsymbol{\beta}|\boldsymbol{k}_G^A).\end{aligned}$$

When $\boldsymbol{k}_G^A \neq \boldsymbol{0}$,

$$\mathbb{E}\{E(\boldsymbol{\beta}|\boldsymbol{K}_G\boldsymbol{X}^A)\} = \frac{\sum_j k_{G,\cdot j}^A \beta_{\cdot j}}{\sum_j k_{G,\cdot j}^A}$$

Proof. (1) Because of Lemma 8,

$$\mathbb{E}\{E(\boldsymbol{\omega}|\boldsymbol{K}_G\boldsymbol{X}^A)\} = \mathbb{E}\{E(\boldsymbol{\omega}|\boldsymbol{k}_G^A\boldsymbol{X}^A)\}.$$

When $\boldsymbol{k}_G^A \neq \boldsymbol{0}$, it holds

$$\begin{aligned}&\mathbb{E}\{E(\boldsymbol{\omega}|\boldsymbol{k}_G^A\boldsymbol{X}^A)\} \\ &= \frac{1}{n_G^A}\mathbb{E}\Big(\sum_j k_{G,\cdot j}^A \sum_i X_{ij}^A\omega_{ij}\Big) \\ &\quad (\because \text{Condition 1, } \boldsymbol{k}_G^A\boldsymbol{X}^A \neq \boldsymbol{0}\text{, Equations 4, 9, and 36, Lemma 10 (2)}) \\ &= \frac{1}{n_G^A}\sum_j \mathbb{E}\Big(k_{G,\cdot j}^A \sum_i X_{ij}^A\omega_{ij}\Big) \\ &\quad (\because \text{Lemma 10 (1), } \mathbb{X}_{\text{def}}(\sum_j k_{G,\cdot j}^A \sum_i X_{ij}^A\omega_{ij}) = \mathbb{X}_{\text{def}}(k_{G,\cdot j}^A \sum_i X_{ij}^A\omega_{ij}) = \mathbb{X}_{\max}) \\ &= \frac{1}{n_G^A}\sum_j k_{G,\cdot j}^A \mathbb{E}\Big(\sum_i X_{ij}^A\omega_{ij}\Big) \quad (\because \text{Lemma 10 (2), } k_{G,\cdot j}^A \text{ is constant irrespective of } \boldsymbol{X}^A) \\ &= \frac{1}{n_G^A}\sum_j k_{G,\cdot j}^A \sum_i \mathbb{E}(X_{ij}^A\omega_{ij}) \quad (\because \text{Lemma 10 (1), } \mathbb{X}_{\text{def}}(\sum_i X_{ij}^A\omega_{ij}) = \mathbb{X}_{\text{def}}(X_{ij}^A\omega_{ij}) = \mathbb{X}_{\max}) \\ &= \frac{1}{n_G^A}\sum_j k_{G,\cdot j}^A \sum_i \mathbb{E}(X_{ij}^A)\omega_{ij} \quad (\because \text{Lemma 10 (2), } \omega_{ij} \text{ is constant irrespective of } \boldsymbol{X}^A) \\ &= \frac{1}{n_G^A}\sum_j k_{G,\cdot j}^A \sum_i \frac{1}{2}\omega_{ij} \quad (\because \text{Lemma 11 (1) }) \\ &= \frac{1}{n_G^A}\sum_j k_{G,\cdot j}^A \cdot 0 \quad (\because \text{Equation 21}) \\ &= 0.\end{aligned}$$

It also holds

$$\begin{aligned}E(\boldsymbol{\omega}|\boldsymbol{k}_G^A) &= \frac{\sum_j \sum_i k_{G,\cdot j}^A \omega_{ij}}{\sum_j \sum_i k_{G,\cdot j}^A} \quad (\because \text{Equation 4 and 36, Condition 1}) \\ &= \frac{\sum_j k_{G,\cdot j}^A \sum_i \omega_{ij}}{\sum_j k_{G,\cdot j}^A \sum_i 1} \\ &= \frac{\sum_j k_{G,\cdot j}^A \cdot 0}{2n_G^A} \quad (\because \text{Lemma 5 (1), Equation 36}) \\ &= 0.\end{aligned}$$

When $\boldsymbol{k}_G^A = \boldsymbol{0}$, the desired results follow thanks to Equation 5 and $\boldsymbol{k}_G^A\boldsymbol{X}^A = \boldsymbol{0}$.

(2) Because of Lemma 8,

$$\mathbb{E}\{E(\boldsymbol{\beta}|\boldsymbol{K}_G\boldsymbol{X}^A)\} = \mathbb{E}\{E(\boldsymbol{\beta}|\boldsymbol{k}_G^A\boldsymbol{X}^A)\}.$$

When $\boldsymbol{k}_G^A \neq \mathbf{0}$,

$$\begin{aligned}
E(\boldsymbol{\beta}|\boldsymbol{k}_G^A\boldsymbol{X}^A) &= \frac{\sum_j \sum_i k_{G,\cdot j}^A X_{ij}^A \beta_{\cdot j}}{\sum_j \sum_i k_{G,\cdot j}^A X_{ij}^A} \quad (\because \text{ Condition 1, } \boldsymbol{k}_G^A\boldsymbol{X}^A \neq \mathbf{0}\text{, Equation 4 and 36})\\
&= \frac{\sum_j k_{G,\cdot j}^A \beta_{\cdot j} \sum_i X_{ij}^A}{\sum_j k_{G,\cdot j}^A \sum_i X_{ij}^A}\\
&= \frac{\sum_j k_{G,\cdot j}^A \beta_{\cdot j}}{\sum_j k_{G,\cdot j}^A} \quad (\because \text{Equation 2})\\
&= \frac{\sum_j \sum_i k_{G,\cdot j}^A \beta_{\cdot j}}{\sum_j \sum_i k_{G,\cdot j}^A}\\
&= E(\boldsymbol{\beta}|\boldsymbol{k}_G^A) \quad (\because \text{Equation 4})
\end{aligned}$$

Taking expectation of both sides of the previous equation, we obtain

$$\begin{aligned}
\mathbb{E}\{E(\boldsymbol{\beta}|\boldsymbol{k}_G^A\boldsymbol{X}^A)\} &= \mathbb{E}\{E(\boldsymbol{\beta}|\boldsymbol{k}_G^A)\}\\
&= E(\boldsymbol{\beta}|\boldsymbol{k}_G^A) \quad (\because \text{Equation 10 (2)})
\end{aligned}$$

where the second equality follows because $\boldsymbol{\beta}$ and $\boldsymbol{k}_G^A$ is constant irrespective of $\boldsymbol{X}^A$ and, thus, so is $E(\boldsymbol{\beta}|\boldsymbol{k}_G^A)$.

When $\boldsymbol{k}_G^A = \mathbf{0}$,

$$\begin{aligned}
E(\boldsymbol{\beta}|\boldsymbol{k}_G^A) &= 0 \quad (\because \text{Equation 5})\\
\therefore \mathbb{E}\{E(\boldsymbol{\beta}|\boldsymbol{k}_G^A)\} &= 0 \quad (\because \text{Lemma 10})\\
\mathbb{E}\{E(\boldsymbol{\beta}|\boldsymbol{k}_G^A\boldsymbol{X}^A)\} &= 0 \quad (\because \boldsymbol{k}_G^A\boldsymbol{X}^A = \mathbf{0}, \text{Equation 5, Lemma 10})
\end{aligned}$$

□

LEMMA 13 (SUBSTITUTION). *Below, $\boldsymbol{k}_G^{TC}$ (Equation 103), n_G^{TC} (Equation 104), Equations 105 through 108 will be defined or proved in the next subsection.*

(1) When $\boldsymbol{K}_G = \boldsymbol{K}_F$ holds, it follows that Condition 1 ($\because$ Equation 12) holds and

$$\begin{aligned}
&\hat{\tau}_G = \hat{\tau}_F \quad (\because \text{Equations 10 and 11})\\
&\boldsymbol{k}_G^T = \boldsymbol{k}_G^C = \boldsymbol{k}_G^{TC} \equiv \boldsymbol{k}_G = \boldsymbol{k}_F \quad (\because \text{Equation 12})\\
&n_G^T = n_G^C = n_G^{TC} \equiv n_G = n_F = n \quad (\because \text{Equations 31, 35, and 104})
\end{aligned}$$

(2) When $\boldsymbol{K}_G = \boldsymbol{K}_P$ and Assumption 3 hold, it follows that Condition 1 ($\because$ Assumption 3) holds and

$$\begin{aligned}
&\hat{\tau}_G = \hat{\tau}_P \quad (\because \text{Equations 10 and 15})\\
&\boldsymbol{k}_G^T = \boldsymbol{k}_G^C = \boldsymbol{k}_G^{TC} \equiv \boldsymbol{k}_G = \boldsymbol{k}_P \quad (\because \text{Lemma 9 (1), Equation 107})\\
&n_G^T = n_G^C = n_G^{TC} \equiv n_G = n_P \quad (\because \text{Lemma 9 (1), Equations 31 and 108})
\end{aligned}$$

(3) When $\boldsymbol{K}_G = \boldsymbol{K}_U$ and Assumption 2 hold, it follows that Condition 1 ($\because$ Assumption 2)

holds and

$$\begin{aligned}
\hat{\tau}_G &= \hat{\tau}_U \quad (\because \text{ Equations 10 and 13})\\
\boldsymbol{k}_G^T &= \boldsymbol{k}_U^T \quad (\because \text{ by definition})\\
\boldsymbol{k}_G^C &= \boldsymbol{k}_U^C \quad (\because \text{ by definition})\\
n_G^T &= n_U^T \quad (\because \text{ Equations 31, Lemma 9 (2)})\\
n_G^C &= n_U^C \quad (\because \text{ Equations 31, Lemma 9 (2)})\\
\boldsymbol{k}_G^{TC} &= \boldsymbol{k}_U^{TC} = \boldsymbol{k}_P \quad (\because \text{ Equation 105})\\
n_G^{TC} &= n_U^{TC} = n_P \quad (\because \text{ Equation 106})
\end{aligned}$$

PROOF. As annotated in the lemma.

PROPOSITION 1 (BIAS OF ATE ESTIMATORS: FS). *(1) Under Assumption 1,*

$$\mathbb{E}(\hat{\tau}_F) - \bar{\tau} = 0.$$

(2) Under Assumption 3 and $n_P \geq 1$,

$$\mathbb{E}(\hat{\tau}_P) - \bar{\tau} = E(\boldsymbol{\beta}^T - \boldsymbol{\beta}^C | \boldsymbol{k}_P).$$

(3) Under Assumption 2 and $n_U^T, n_U^C \geq 1$,

$$\mathbb{E}(\hat{\tau}_U) - \bar{\tau} = E(\boldsymbol{\beta}^T | \boldsymbol{k}_U^T) - E(\boldsymbol{\beta}^C | \boldsymbol{k}_U^C).$$

(4) Under Assumptions 3 and 5, and $n_P \geq 1$,

$$\mathbb{E}(\hat{\tau}_P) - \bar{\tau} = 0.$$

(5) Under Assumptions 2 and 4, and $n_U^T, n_U^C \geq 1$,

$$\mathbb{E}(\hat{\tau}_U) - \bar{\tau} = 0.$$

PROOF. (1)–(3) Suppose that $\mathbb{X}_{\text{def}}(\hat{\tau}_G) = \mathbb{X}_{\text{max}}$. Thus, for any $\boldsymbol{X} \in \mathbb{X}_{\text{max}}$, it holds that $N_G^A \geq 1$. It follows

$$\begin{aligned}
\mathbb{E}(\hat{\tau}_G) - \bar{\tau} &= \mathbb{E}(\hat{\tau}_G - \bar{\tau}) \quad (\because \text{ Lemma 10})\\
&= \mathbb{E}\{E(\boldsymbol{\beta}^T + \boldsymbol{\omega}^T | \boldsymbol{K}_G \boldsymbol{X}^T) - E(\boldsymbol{\beta}^C + \boldsymbol{\omega}^C | \boldsymbol{K}_G \boldsymbol{X}^C)\} \quad (\because \text{ Equation 24}, N_G^A \geq 1, \\
&\quad \mathbb{X}_{\text{def}}(\hat{\tau}_G - \bar{\tau}) = \mathbb{X}_{\text{def}}\{E(\boldsymbol{\beta}^T + \boldsymbol{\omega}^T | \boldsymbol{K}_G \boldsymbol{X}^T) - E(\boldsymbol{\beta}^C + \boldsymbol{\omega}^C | \boldsymbol{K}_G \boldsymbol{X}^C)\} = \mathbb{X}_{\text{max}})
\end{aligned} \tag{42}$$

In addition, suppose that $\boldsymbol{K}_G \in \mathbb{U}$ satisfies Condition 1 as well. Applying Lemmas 3, 10, and 12 to Equation 42, we obtain

$$\mathbb{E}(\hat{\tau}_G) - \bar{\tau} = E(\boldsymbol{\beta}^T | \boldsymbol{k}_G^T) - E(\boldsymbol{\beta}^C | \boldsymbol{k}_G^C). \tag{43}$$

In particular, when $\boldsymbol{k}_G^T = \boldsymbol{k}_G^C \equiv \boldsymbol{k}_G$, Equation 43 leads to

$$E(\boldsymbol{\beta}^T | \boldsymbol{k}_G) - E(\boldsymbol{\beta}^C | \boldsymbol{k}_G) = E(\boldsymbol{\beta}^T - \boldsymbol{\beta}^C | \boldsymbol{k}_G). \quad (\because \text{ Lemma 3 (1)}) \tag{44}$$

(1) Under Assumption 1, it holds that $\mathbb{X}_{\text{def}}(\hat{\tau}_F) = \mathbb{X}_{\text{max}}$. When $\boldsymbol{K}_G = \boldsymbol{K}_F \in \mathbb{U}$, it follows that

$$\begin{aligned}
\mathbb{E}(\hat{\tau}_F) - \bar{\tau} &= E(\boldsymbol{\beta}^T - \boldsymbol{\beta}^C | \boldsymbol{k}_F) \quad (\because \text{ Lemma 13 (1), Equations 43 and 44})\\
&= 0 \quad (\because \text{ Lemmas 3 (1) and 5 (2), Equation 12})
\end{aligned}$$

(2) Under Assumption 3 and $N_P = n_P \geq 1$ ($\because$ Lemma 9 (1)), it holds that $\mathbb{X}_{\text{def}}(\hat{\tau}_P) = \mathbb{X}_{\text{max}}$. When $\boldsymbol{K}_G = \boldsymbol{K}_P \in \mathbb{U}$, according to Lemma 13 (2), it follows that Equations 43 and 44 are equivalent to the desired result.

(3) Under Assumption 2 and $N_U^A = n_U^A \geq 1$ ($\because$ Lemma 9 (2)), it holds that $\mathbb{X}_{\text{def}}(\hat{\tau}_U) = \mathbb{X}_{\max}$. When $\boldsymbol{K}_G = \boldsymbol{K}_U \in \mathbb{U}$, according to Lemma 13 (3), it follows that Equation 43 is equivalent to the desired result.

(4)–(5) Suppose that $\boldsymbol{K}_G \in \mathbb{U}$ satisfies Condition 2 . When $\boldsymbol{k}_G^A \neq \boldsymbol{0}$, it follows

$$
\begin{aligned}
E(\boldsymbol{y}^A|\boldsymbol{k}_G^A) &= \sum_y P(\boldsymbol{y}^A = y|\boldsymbol{k}_G^A = 1)y \quad (\because \text{Lemma 7, } \boldsymbol{k}_G^A \neq \boldsymbol{0}) \\
&= \sum_y P(\boldsymbol{y}^A = y|\boldsymbol{1} = 1)y \quad (\because \text{Equations 33 and 38}, \boldsymbol{k}_G^A \neq \boldsymbol{0}) \\
&= E(\boldsymbol{y}^A|\boldsymbol{1}) \quad (\because \text{Lemma 7}) \\
&= \mu^A. \quad (\because \text{Equation 17 })
\end{aligned} \tag{45}
$$

In addition, suppose that $\boldsymbol{K}_G$ satisfies Condition 1 as well. It follows

$$
\begin{aligned}
E(\boldsymbol{\beta}^A|\boldsymbol{k}_G^A) &= E(\boldsymbol{\mu}^A|\boldsymbol{k}_G^A) - \mu^A + E(\boldsymbol{\beta}^A|\boldsymbol{k}_G^A) + E(\boldsymbol{\omega}^A|\boldsymbol{k}_G^A) \quad (\because \text{Lemmas 3 (3) and 12 (1), } \boldsymbol{\omega}^A \in \mathbb{W}) \\
&= E(\boldsymbol{\mu}^A + \boldsymbol{\beta}^A + \boldsymbol{\omega}^A|\boldsymbol{k}_G^A) - \mu^A \quad (\because \text{Lemma 3 (1)}) \\
&= E(\boldsymbol{y}^A|\boldsymbol{k}_G^A) - \mu^A \quad (\because \text{Lemma 4}) \\
&= 0 \quad (\because \text{Equation 45})
\end{aligned} \tag{46}
$$

In addition, suppose that $\mathbb{X}_{\text{def}}(\hat{\tau}_G) = \mathbb{X}_{\max}$. It follows that

$$
\begin{aligned}
\mathbb{E}(\hat{\tau}_G) - \bar{\tau} &= E(\boldsymbol{\beta}^T|\boldsymbol{k}_G^T) - E(\boldsymbol{\beta}^C|\boldsymbol{k}_G^C) \quad (\because \text{Equation 43}) \\
&= 0. \quad (\because \text{Equation 46 })
\end{aligned} \tag{47}
$$

(4) Substitute $\boldsymbol{K}_G = \boldsymbol{K}_P$. Under Assumption 3 and 5, it follows that Conditions 1 and 2 hold. When $n_P \geq 1$, it holds that $\boldsymbol{k}_G^A \neq \boldsymbol{0}$ ($\because$ Lemma 14(4)) and $\mathbb{X}_{\text{def}}(\hat{\tau}_G) = \mathbb{X}_{\max}$. Equation 47 is equivalent to the desired result.

(5) Substitute $\boldsymbol{K}_G = \boldsymbol{K}_U$. Under Assumption 2 and 4, it follows that Conditions 1 and 2 hold. When $n_U^T, n_U^C \geq 1$, it holds that $\boldsymbol{k}_G^A \neq \boldsymbol{0}$ ($\because$ Lemma 14(4)) and $\mathbb{X}_{\text{def}}(\hat{\tau}_G) = \mathbb{X}_{\max}$. Equation 47 is equivalent to the desired result.

□

It is well known that in the case of no missing values, $\hat{\tau}_F$ is unbiased for $\bar{\tau}$ (e.g., Imai 2008, 4861). When attrition is matched (Assumption 2 or 3) but not ignorable (Assumption 4 or 5) , $\hat{\tau}_U$ *and* $\hat{\tau}_P$ are biased (except for knife-edge situations such as $E(\boldsymbol{\beta}^T - \boldsymbol{\beta}^C|\boldsymbol{k}_P) = 0$ or $E(\boldsymbol{\beta}^T|\boldsymbol{k}_U^T) - E(\boldsymbol{\beta}^C|\boldsymbol{k}_U^C) = 0$) . This is true even if the pair matching is perfectly effective in the sense that in every pair the realized outcome of one unit is the counter-factual outcome of the other unit, namely, $\boldsymbol{y}^T = \boldsymbol{y}_{-i}^T$ and $\boldsymbol{y}^C = \boldsymbol{y}_{-i}^C$. Interestingly, both advocates and critics of $\hat{\tau}_P$ seem to share misunderstanding of conditions in which $\hat{\tau}_P$ is unbiased. That is, they mistakenly think that matched attrition (Assumption 2 or 3) is a sufficient condition. According to King et al. (2007), proponents of $\hat{\tau}_P$, "if we lose a cluster [i.e., unit] for a reason related to one or more of the variables we matched on [i.e., Assumption 2], . . . we would be fully protected from bias due to any variable we were able to match on" (490). Dunning (2011), denouncing $\hat{\tau}_P$, argues similarly that, for $\hat{\tau}_P$ to be unbiased, "we have to assume that all units with the same values of the blocked covariate respond similarly to treatment assignment" (15), that is, Assumption 2. Both statements, however, are not true according to Proposition 1 (2). (Recall that Assumption 2 leads to Assumption 3.)

LATE of Observable Pairs. When $\boldsymbol{k}_P \neq \boldsymbol{0}$, define the local average treatment effect (LATE) of "always-reporting pairs" by

$$
\begin{aligned}
\bar{\tau}_P &\equiv E(\boldsymbol{\tau}|\boldsymbol{k}_P) \\
&= E(\bar{\boldsymbol{\tau}} + \boldsymbol{\beta}^T - \boldsymbol{\beta}^C + \boldsymbol{\omega}^T - \boldsymbol{\omega}^C|\boldsymbol{k}_P) \quad (\because \text{Lemma 4 (3)}) \\
&= \bar{\tau} + E(\boldsymbol{\beta}^T - \boldsymbol{\beta}^C|\boldsymbol{k}_P) \quad (\because \boldsymbol{k}_P \neq \boldsymbol{0}, \text{Lemmas 3 (1) and (3) and 12 (1)})
\end{aligned} \tag{48}
$$

It follows

$$\begin{aligned}\mathbb{E}(\hat{\tau}_P) - \bar{\tau}_P &= \{\mathbb{E}(\hat{\tau}_P) - \bar{\tau}\} - (\bar{\tau}_P - \bar{\tau}) \\ &= E(\boldsymbol{\beta}^T - \boldsymbol{\beta}^C | \boldsymbol{k}_P) - E(\boldsymbol{\beta}^T - \boldsymbol{\beta}^C | \boldsymbol{k}_P) \quad (\because \text{Equation 48}) \\ &= 0.\end{aligned}$$

Non-compliance. Now, we consider non-compliance cases briefly. Let $\boldsymbol{R}'$ compliance indicator. Define $\boldsymbol{r}^{A\prime}$ accordingly. Define treatment received as

$$\boldsymbol{X}^{A\prime} \equiv \boldsymbol{R}'\boldsymbol{X}^A + (\boldsymbol{1} - \boldsymbol{R}')\boldsymbol{X}^{-A} \tag{49}$$

and redefine realized outcome as

$$\begin{aligned}\boldsymbol{Y}' &\equiv \sum_A \boldsymbol{X}^{A\prime}\boldsymbol{y}^A \\ &= \sum_A \{\boldsymbol{R}'\boldsymbol{X}^A + (\boldsymbol{1} - \boldsymbol{R}')\boldsymbol{X}^{-A}\}\boldsymbol{y}^A \quad (\because \text{Equation 49}) \\ &= \sum_A \{\boldsymbol{r}^{A\prime}\boldsymbol{y}^A + (\boldsymbol{1} - \boldsymbol{r}^{A\prime})\boldsymbol{y}^{-A}\}\boldsymbol{X}^A \quad (\because \text{Lemma 8})\end{aligned} \tag{50}$$

Assume matched potential compliance ($\boldsymbol{r}^{A\prime} = \boldsymbol{r}_{-i}^{A\prime}$), which corresponds to Assumption 2. Denote a set of pairs which share compliance status by

$$\mathbb{J}_{R'}(g^T, g^C) \equiv \left\{ j \middle| \prod_A (r_{\cdot j}^{A\prime})^{g^A}(1 - r_{\cdot j}^{A\prime})^{1-g^A} = 1 \right\}.$$

It follows that $\mathbb{J}_{R'}(1,1)$, $\mathbb{J}_{R'}(1,0)$, $\mathbb{J}_{R'}(0,1)$, and $\mathbb{J}_{R'}(0,0)$ are sets of compliers, always takers, never takers, and defiers, respectively.

The general ATE estimator is defined as

$$\hat{\tau}'(\boldsymbol{K}_G) \equiv E(\boldsymbol{Y}'|\boldsymbol{K}_G\boldsymbol{X}^T) - E(\boldsymbol{Y}'|\boldsymbol{K}_G\boldsymbol{X}^C) \tag{51}$$

only when $N_G^A \geq 1$.

Suppose $\boldsymbol{K}_G \in \mathbb{U}$ satisfies Condition 1, $\boldsymbol{k}_G^T = \boldsymbol{k}_G^C \equiv \boldsymbol{k}_G \neq \boldsymbol{0}$, and monotonicity (or no defiers,

$\mathbb{J}_{R'}(0,0) = \emptyset$). It follows

$$\begin{aligned}
\hat{\tau}'(\boldsymbol{K}_G) &= \frac{\sum_j \sum_i k_{G,\cdot j} X_{ij}^T Y_{ij}'}{\sum_j \sum_i k_{G,\cdot j} X_{ij}^T} - \frac{\sum_j \sum_i k_{G,\cdot j} X_{ij}^C Y_{ij}'}{\sum_j \sum_i k_{G,\cdot j} X_{ij}^C} \\
&\quad (\because \text{Equations 4 and 51, Lemma 8, Condition 1, } \boldsymbol{k}_G^T = \boldsymbol{k}_G^C \equiv \boldsymbol{k}_G \neq \boldsymbol{0}, N_G = n_G \geq 1) \\
&= \frac{1}{n_G} \sum_j k_{G,\cdot j} \sum_i Y_{ij}' (X_{ij}^T - X_{ij}^C) \\
&= \frac{1}{n_G} \sum_j k_{G,\cdot j} \sum_i [\{r_{\cdot j}^{T\prime} y_{ij}^T + (1 - r_{\cdot j}^{T\prime}) y_{ij}^C\} X_{ij}^T - \{r_{\cdot j}^{C\prime} y_{ij}^C + (1 - r_{\cdot j}^{C\prime}) y_{ij}^T\} X_{ij}^C] \\
&\quad (\because \text{Lemma 1 (5), 2 (2), Equation 50}) \\
&= \frac{1}{n_G} \Big\{ \sum_{j \in \mathbb{J}_{R'}(1,1)} k_{G,\cdot j} \sum_i (y_{ij}^T X_{ij}^T - y_{ij}^C X_{ij}^C) + \sum_{j \in \mathbb{J}_{R'}(1,0)} k_{G,\cdot j} \sum_i (y_{ij}^T X_{ij}^T - y_{ij}^T X_{ij}^C) \\
&\quad + \sum_{j \in \mathbb{J}_{R'}(0,1)} k_{G,\cdot j} \sum_i (y_{ij}^C X_{ij}^T - y_{ij}^C X_{ij}^C) \Big\} \quad (\because \mathbb{J}_{R'}(0,0) = \emptyset) \\
&= \frac{1}{n_G} \Big\{ \sum_{j \in \mathbb{J}_{R'}(1,1)} k_{G,\cdot j} \sum_i (y_{ij}^T - y_{-ij}^C) X_{ij}^T + \sum_{j \in \mathbb{J}_{R'}(1,0)} k_{G,\cdot j} \sum_i (y_{ij}^T - y_{-ij}^T) X_{ij}^T \\
&\quad + \sum_{j \in \mathbb{J}_{R'}(0,1)} k_{G,\cdot j} \sum_i (y_{ij}^C - y_{-ij}^C) X_{ij}^T \Big\} \quad (\because \text{Lemmas 1 (4) and 2 (1)}) \\
&= \frac{1}{n_G} \Big[\sum_{j \in \mathbb{J}_{R'}(1,1)} k_{G,\cdot j} \sum_i \{(\mu^T + \beta_{\cdot j}^T + \omega_{ij}^T) - (\mu^C + \beta_{\cdot j}^C + \omega_{-ij}^C)\} X_{ij}^T \\
&\quad + \sum_{j \in \mathbb{J}_{R'}(1,0)} k_{G,\cdot j} \sum_i \{(\mu^T + \beta_{\cdot j}^T + \omega_{ij}^T) - (\mu^T + \beta_{\cdot j}^T + \omega_{-ij}^T)\} X_{ij}^T \\
&\quad + \sum_{j \in \mathbb{J}_{R'}(0,1)} k_{G,\cdot j} \sum_i \{(\mu^C + \beta_{\cdot j}^C + \omega_{ij}^C) - (\mu^C + \beta_{\cdot j}^C + \omega_{-ij}^C)\} X_{ij}^T \Big] \quad (\because \text{Lemma 4 (1)}) \\
&= \frac{1}{n_G} \Big\{ \sum_{j \in \mathbb{J}_{R'}(1,1)} k_{G,\cdot j} \sum_i (\tau_{ij} + 2\omega_{ij}^C) X_{ij}^T + \sum_{j \in \mathbb{J}_{R'}(1,0)} k_{G,\cdot j} \sum_i 2\omega_{ij}^T X_{ij}^T \\
&\quad + \sum_{j \in \mathbb{J}_{R'}(0,1)} k_{G,\cdot j} \sum_i 2\omega_{ij}^C X_{ij}^T \Big\} \quad (\because \text{Lemmas 4 (2) and (3) and 5 (1)})
\end{aligned} \tag{52}$$

Let $n_{R'}(\boldsymbol{k}_G) \equiv \sum_j r_{\cdot j}^{T\prime} r_{\cdot j}^{C\prime} k_{G,\cdot j}$. It follows

$$\begin{aligned}
\mathbb{E}\{\hat{\tau}'(\boldsymbol{K}_G)\} &= \frac{1}{n_G} \Big\{ \sum_{j \in \mathbb{J}_{R'}(1,1)} k_{G,\cdot j} \sum_i (\tau_{ij} + 2\omega_{ij}^C) \mathbb{E}(X_{ij}^T) + \sum_{j \in \mathbb{J}_{R'}(1,0)} k_{G,\cdot j} \sum_i 2\omega_{ij}^T \mathbb{E}(X_{ij}^T) \\
&\quad + \sum_{j \in \mathbb{J}_{R'}(0,1)} k_{G,\cdot j} \sum_i 2\omega_{ij}^C \mathbb{E}(X_{ij}^T) \Big\} \quad (\because \text{Equation 52, Lemma 10 (1)}) \\
&= \frac{1}{n_G} \sum_{j \in \mathbb{J}_{R'}(1,1)} k_{G,\cdot j} \sum_i \frac{1}{2} \tau_{ij} \quad (\because \text{Lemmas 5 (1) and 11 (1)}) \\
&= \frac{1}{2n_G} \sum_j \sum_i r_{\cdot j}^{T\prime} r_{\cdot j}^{C\prime} k_{G,\cdot j} \tau_{ij} \\
&= \frac{n_{R'}(\boldsymbol{k}_G)}{n_G} E(\boldsymbol{\tau} | \boldsymbol{r}^{T\prime} \boldsymbol{r}^{C\prime} \boldsymbol{k}_G).
\end{aligned} \tag{53}$$

Let $\boldsymbol{K}'_P \equiv \boldsymbol{R}'\boldsymbol{R}'_{-i}$. It follows

$$\begin{aligned}\boldsymbol{k}_P^{A\prime} &= \boldsymbol{r}^{A\prime}\boldsymbol{r}_{-i}^{-A\prime} \\ &= \boldsymbol{r}^{A\prime}\boldsymbol{r}^{-A\prime} \quad (\because \boldsymbol{r}^{A\prime} = \boldsymbol{r}_{-i}^{A\prime}) \\ &= \boldsymbol{k}_P^{-A\prime} \\ &\equiv \boldsymbol{k}'_P,\end{aligned} \tag{54}$$

which corresponds to Assumption 3. It also follows that

$$\begin{aligned}N'_P &\equiv N^A(\boldsymbol{K}'_P) \quad (\because \text{Equation 16}) \\ &= n(\boldsymbol{k}'_P) \quad (\because \text{Equation 54, Lemma 9 (1)}) \\ &\equiv n'_P.\end{aligned} \tag{55}$$

Specifically, the intention-to-treat estimator is denoted by

$$\begin{aligned}\hat{\tau}'_F &\equiv \hat{\tau}'(\boldsymbol{K}_F) \\ &= \frac{1}{n}\Big\{\sum_{j\in\mathbb{J}_{R'}(1,1)}\sum_i(\tau_{ij}+2\omega_{ij}^C)X_{ij}^T + \sum_{j\in\mathbb{J}_{R'}(1,0)}\sum_i 2\omega_{ij}^T X_{ij}^T + \sum_{j\in\mathbb{J}_{R'}(0,1)}\sum_i 2\omega_{ij}^C X_{ij}^T\Big\} \\ &\quad (\because \boldsymbol{K}_F \in \mathbb{U} \text{ satisfies Condition 1, } \boldsymbol{k}_F^T = \boldsymbol{k}_F^C \neq \boldsymbol{0}, \text{ Equations 12 and 52})\end{aligned} \tag{56}$$

and its expectation is equal to

$$\begin{aligned}\mathbb{E}(\hat{\tau}'_F) &= \frac{n_{R'}(\boldsymbol{k}_F)}{n_F}E(\boldsymbol{\tau}|\boldsymbol{r}^{T\prime}\boldsymbol{r}^{C\prime}\boldsymbol{k}_F) \quad (\because \text{Equation 53}) \\ &= \frac{n'_P}{n}E(\boldsymbol{\tau}|\boldsymbol{k}'_P) \quad (\because \boldsymbol{k}'_P = \boldsymbol{r}^{T\prime}\boldsymbol{r}^{C\prime}\boldsymbol{k}_F, n'_P = n_{R'}(\boldsymbol{k}_F), \text{ Equation 54})\end{aligned} \tag{57}$$

When $N'_P \geq 1$, the pairwise deletion estimator is defined as

$$\begin{aligned}\hat{\tau}'_P &\equiv \hat{\tau}'(\boldsymbol{K}'_P) \\ &= \frac{1}{n'_P}\sum_{j\in\mathbb{J}_{R'}(1,1)}\sum_i(\tau_{ij}+2\omega_{ij}^C)X_{ij}^T \quad (\because \boldsymbol{K}'_P \in \mathbb{U} \text{ satisfies Condition 1, Equations 52, 54,} \\ &\quad \text{and 55, } N'_P = n'_P \geq 1, \boldsymbol{k}'_P \neq \boldsymbol{0}, k_{P,\cdot j} = 1\forall j \in \mathbb{J}_{R'}(1,1), k_{P,\cdot j} = 0\forall j \in \mathbb{J}_{R'}(1,0)\cup\mathbb{J}_{R'}(0,1))\end{aligned} \tag{58}$$

and its expectation is equal to

$$\begin{aligned}\mathbb{E}(\hat{\tau}'_P) &= \frac{n_{R'}(\boldsymbol{k}_P)}{n'_P}E(\boldsymbol{\tau}|\boldsymbol{r}^{T\prime}\boldsymbol{r}^{C\prime}\boldsymbol{k}'_P) \quad (\because \text{Equation 53}) \\ &= E(\boldsymbol{\tau}|\boldsymbol{k}'_P) \quad (\because \boldsymbol{r}^{T\prime}\boldsymbol{r}^{C\prime}\boldsymbol{k}'_P = (\boldsymbol{k}'_P)^2 = \boldsymbol{k}'_P, n_{R'}(\boldsymbol{k}_P) = n'_P, \text{ Equation 54}) \\ &\equiv \bar{\tau}'_P,\end{aligned} \tag{59}$$

which is the local average treatment effect of compliers ($\{j|\boldsymbol{k}'_P = 1\} = \mathbb{J}_{R'}(1,1)$).

When $N'_P \geq 1$, denote the instrumental variable estimator by

$$\begin{aligned}\hat{\tau}'_{\text{IV}} &\equiv \frac{E(\boldsymbol{Y}'|\boldsymbol{X}^T) - E(\boldsymbol{Y}'|\boldsymbol{X}^C)}{E(\boldsymbol{X}^{T\prime}|\boldsymbol{X}^T) - E(\boldsymbol{X}^{T\prime}|\boldsymbol{X}^C)} \\ &= \left(\frac{N'_P}{n}\right)^{-1}\hat{\tau}'_F \quad (\because \text{Equations 12, 52, and 56})\end{aligned} \tag{60}$$

because

$$
\begin{aligned}
& E(\boldsymbol{X}^{T\prime}|\boldsymbol{X}^{T}) - E(\boldsymbol{X}^{T\prime}|\boldsymbol{X}^{C}) \\
= & \frac{1}{n}\sum_j \sum_i (X_{ij}^T - X_{ij}^C)\{R'_{ij}X_{ij}^T + (1-R'_{ij})X_{ij}^C\} \quad (\because \text{Equations 2, 4, and 49}) \\
= & \frac{1}{n}\sum_j \sum_i \{X_{ij}^T r_{ij}^{T\prime} - X_{ij}^C(1-r_{ij}^{C\prime})\} \quad (\because \text{Equations 1 (5) and 2 (2), Lemma 8}) \\
= & \frac{1}{n}\Big[\sum_j \sum_i X_{ij}^T r_{\cdot j}^{T\prime}\{r_{\cdot j}^{C\prime} + (1-r_{\cdot j}^{C\prime})\} - \sum_j \sum_i X_{ij}^T(1-r_{\cdot j}^{C\prime})\Big] \quad (\because \text{Lemmas 1 (4) and 2 (1)}) \\
= & \frac{1}{n}\Big\{\sum_j \sum_i X_{ij}^T r_{\cdot j}^{T\prime} r_{\cdot j}^{C\prime} - \sum_j \sum_i X_{ij}^T(1-r_{\cdot j}^{T\prime})(1-r_{\cdot j}^{C\prime})\Big\} \\
= & \frac{1}{n}\sum_j r_{\cdot j}^{T\prime} r_{\cdot j}^{C\prime} \quad (\because \mathbb{J}_{R'}(0,0)=\emptyset, \text{Equation 1}) \\
= & \frac{N'_P}{n} \quad (\because \text{Equations 54 and 55})
\end{aligned}
\tag{61}
$$

Therefore, when $n'_P \geq 1$,

$$
\begin{aligned}
\mathbb{E}(\hat{\tau}'_{IV}) &= \left(\frac{n'_P}{n}\right)^{-1}\mathbb{E}(\hat{\tau}'_F) \quad (\because \text{Equations 55 and 60, Lemma 10 (1)}) \\
&= \bar{\tau}'_P \quad (\because \text{Equations 57 and 59}) \\
&= \mathbb{E}(\hat{\tau}'_P) \quad (\because \text{Equation 59})
\end{aligned}
$$

This means that both $\hat{\tau}'_{IV}$ and $\hat{\tau}'_P$ are unbiased estimators for $\bar{\tau}'_P$. Note that, according to Equations 55, 56, 58, and 60,

$$
\hat{\tau}'_{IV} - \hat{\tau}'_P = \frac{1}{n'_P}\Big\{\sum_{j\in\mathbb{J}_{R'}(1,0)}\sum_i 2\omega_{ij}^T X_{ij}^T + \sum_{j\in\mathbb{J}_{R'}(0,1)}\sum_i 2\omega_{ij}^C X_{ij}^T\Big\},
$$

which is not equal to zero except knife-edge situation.

Complete Randomization. Only in this part, we assume complete randomization instead of pairwise randomization. That is, we keep Equation 28, while, instead of Equations 29 and 30, we assume isoprobability of treatment assignment: for any $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)} \in \mathbb{X}_{\max}^{\mathrm{CR}}$,

$$
\Pr^{\mathrm{CR}}(\boldsymbol{X}^A = \boldsymbol{x}^{(1)}) = \Pr^{\mathrm{CR}}(\boldsymbol{X}^A = \boldsymbol{x}^{(2)}), \tag{62}
$$

where

$$
\mathbb{X}_{\max}^{\mathrm{CR}} \equiv \Big\{\boldsymbol{x}\Big|\boldsymbol{x}\in\mathbb{U}, \sum_j\sum_i x_{ij} = n\Big\}.
$$

Expectation operator is denoted by

$$
\mathbb{E}^{\mathrm{CR}}\{f(\boldsymbol{X}^A)\} \equiv \sum_{\boldsymbol{x}\in\mathbb{X}_{\mathrm{def}}^{\mathrm{CR}}(f)} \Pr\{\boldsymbol{X}^A = \boldsymbol{x}|\boldsymbol{X}^A \in \mathbb{X}_{\mathrm{def}}^{\mathrm{CR}}(f)\}f(\boldsymbol{x})
$$

where

$$
\begin{aligned}
\mathbb{X}_{\mathrm{def}}^{\mathrm{CR}}(f) &\equiv \{\boldsymbol{x}|\boldsymbol{x}\in\mathbb{X}_{\max}^{\mathrm{CR}}, f(\boldsymbol{x}) \text{ can be defined}\} \\
&\neq \emptyset.
\end{aligned}
$$

We assume that $\boldsymbol{K}_G \in \mathbb{U}$ does NOT satisfy Condition 1 but $N_G^A \geq 1$ for any $\boldsymbol{X}^A \in \mathbb{X}_{\max}^{\mathrm{CR}}$, that is, $n_G^{A\cdot CR} > n$, where

$$
n_G^{A\cdot CR} = \sum_j\sum_i k_{G,ij}^A.
$$

There are ${}_{n_G^{A\cdot CR}}C_{N_G^A}$ ways of assigning treatment to N_G^A units from $n_G^{A\cdot CR}$ units of $\{ij|k_{G,ij}^A = 1\}$. Among them, there are ${}_{n_G^{A\cdot CR}-1}C_{N_G^A-1}$ ways of assigning treatment to unit $ij \in \{ij|k_{G,ij}^A = 1\}$ and $N_G^{A1} - 1$ units from $n_G^{A\cdot CR} - 1$ units of $\{i'j'|k_{G,i'j'}^A = 1\} \setminus \{ij\}$. Therefore, for $ij \in \{ij|k_{G,ij}^A = 1\}$, it follows

$$\begin{aligned}
\mathbb{E}^{\text{CR}}(X_{ij}^A|N_G^A) &= \Pr^{\text{CR}}(X_{ij}^A = 1|N_G^A) \\
&= \Pr^{\text{CR}}(X_{ij}^A = 1, N_G^A) \div \Pr(N_G^A) \\
&= \Pr^{\text{CR}}\Big(X_{ij}^A = 1, \sum_{i'j' \in \{i'j'|k_{G,i'j'}^A = 1\} \setminus \{ij\}} X_{i'j'}^A = N_G^A - 1\Big) \div \Pr(N_G^A) \\
&= {}_{n_G^{A\cdot CR}-1}C_{N_G^A-1} \div {}_{n_G^{A\cdot CR}}C_{N_G^A} \quad (\because \text{Equation 62}) \\
&= \frac{(n_G^{A\cdot CR} - 1)!}{\{(n_G^{A\cdot CR} - 1) - (N_G^A - 1)\}!(N_G^A - 1)!} \div \frac{n_G^{A\cdot CR}!}{(n_G^{A\cdot CR} - N_G^A)!N_G^A!} \quad (\because \text{Equation 192}) \\
&= \frac{N_G^A}{n_G^{A\cdot CR}}.
\end{aligned} \tag{63}$$

It follows

$$\begin{aligned}
\mathbb{E}^{\text{CR}}\{E(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^A)\} &= \mathbb{E}^{\text{CR}}\Big(\frac{1}{N_G^A}\sum_j\sum_i k_{G,ij}^A X_{ij}^A y_{ij}^A\Big) \quad (\because \text{Equations 4 and 9}) \\
&= \mathbb{E}^{\text{CR}}\Big\{\frac{1}{N_G^A}\mathbb{E}^{\text{CR}}\Big(\sum_{ij:k_{G,ij}^A=0} 0 \cdot X_{ij}^A y_{ij}^A + \sum_{ij:k_{G,ij}^A=1} 1 \cdot X_{ij}^A y_{1j}^A \Big| N_G^A\Big)\Big\} \\
&= \mathbb{E}^{\text{CR}}\Big\{\frac{1}{N_G^A}\sum_{ij:k_{G,ij}^A=1} \mathbb{E}^{\text{CR}}(X_{ij}^A|N_G^A) y_{ij}^A\Big\} \quad (\because \text{Lemma 10}) \\
&= \mathbb{E}^{\text{CR}}\Big(\frac{1}{N_G^A}\sum_{ij:k_{G,ij}^A=1} \frac{N_G^A}{n_G^{A\cdot CR}} y_{ij}^A\Big) \quad (\because \text{Equation 63}) \\
&= \frac{1}{n_G^{A\cdot CR}}\sum_{ij:k_{G,ij}^A=1} y_{ij}^A \\
&= E(\boldsymbol{y}^A|\boldsymbol{k}_G^A).
\end{aligned} \tag{64}$$

In addition, suppose that $\mathbb{X}_{\text{def}}^{\text{CR}}(\hat{\tau}_G) = \mathbb{X}_{\text{max}}^{\text{CR}}$ (thus, $N_G^A \geq 1$ for any A and $\boldsymbol{X}^A \in \mathbb{X}_{\text{max}}^{\text{CR}}$). It follows

$$\begin{aligned}
\mathbb{E}^{\text{CR}}(\hat{\tau}_G) &= \mathbb{E}^{\text{CR}}\{E(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^T) - E(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^C)\} \\
&\qquad (\because N_G^A \geq 1, \mathbb{X}_{\text{def}}^{\text{CR}}(\hat{\tau}_G) = \mathbb{X}_{\text{def}}^{\text{CR}}\{E(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^A) - E(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^C)\} = \mathbb{X}_{\text{max}}^{\text{CR}}) \\
&= E(\boldsymbol{y}^T|\boldsymbol{k}_G^T) - E(\boldsymbol{y}^C|\boldsymbol{k}_G^C) \quad (\because \text{Lemma 10 (1), Equation 64})
\end{aligned} \tag{65}$$

If we assume Condition 2 but not Condition 1, it follows

$$\begin{aligned}
\mathbb{E}^{\text{CR}}(\hat{\tau}_G) &= E(\boldsymbol{y}^T|\boldsymbol{k}_G^T) - E(\boldsymbol{y}^C|\boldsymbol{k}_G^C) \quad (\because \text{Equation 65}) \\
&= \mu^T - \mu^C \quad (\because \text{Condition 2, Equation 45}) \\
&= \bar{\tau} \quad (\because \text{Lemma 4 (2)}) \\
\therefore \mathbb{E}^{\text{CR}}(\hat{\tau}_G) - \bar{\tau} &= 0.
\end{aligned}$$

Or if we only assume $\boldsymbol{k}_G^T = \boldsymbol{k}_G^C \equiv \boldsymbol{k}_G$ but not Condition 1, it follows

$$\begin{aligned}
\mathbb{E}^{\text{CR}}(\hat{\tau}_G) &= E(\boldsymbol{y}^T|\boldsymbol{k}_G) - E(\boldsymbol{y}^C|\boldsymbol{k}_G) \quad (\because \text{Equation 65}, \boldsymbol{k}_U^T = \boldsymbol{k}_U^C \equiv \boldsymbol{k}_U) \\
&= E(\boldsymbol{\tau}|\boldsymbol{k}_G) \quad (\because \text{Lemma 3 (1), Equation 6})
\end{aligned}$$

which is the local average treatment effect of such units that $k_{G,ij} = 1$ or a principal effect where the corresponding principal stratum is $\{ij|k_{G,ij} = 1\}$.

No Attrition Match. When attrition is ignorable (Assumption 4 or 5) but not matched (i.e., Assumptions 2 and 3 do not hold), $\hat{\tau}_U$ and $\hat{\tau}_P$ are unbiased for $\bar{\tau}$ under *complete* randomization, as is well known, but biased for $\bar{\tau}$ ($=\bar{\tau}_P$) under *pairwise* randomization (except for knife-edge situations). An intuitive reason follows. Suppose that Assumption 2 does not hold and $k^T_{U,1j_1} = 1, k^T_{U,2j_1} = 0, k^T_{U,1j_2} = k^T_{U,2j_2} = 1$. Denote $N^T_{U,-j} \equiv \sum_{j' \neq j} \sum_i K_{U,ij'} X_{ij'}$. In calculating $\mathbb{E}\{E(\boldsymbol{Y}|\boldsymbol{K}_U\boldsymbol{X})\}$, under complete randomization, the weights for Y_{1j_1} and Y_{ij_2} are the same, $1/(\sum_j \sum_i k^T_{U,ij})$. Thus, it holds that $\mathbb{E}\{E(\boldsymbol{Y}|\boldsymbol{K}_U\boldsymbol{X})\} = E(\boldsymbol{y}^T|\boldsymbol{k}^T_U)$. Under pairwise randomization, however, the weight for Y_{1j_1} is $\mathbb{E}\{1/2(N^T_{U,-j_1} + 1)\}$ (note that either unit $1j_2$ or unit $2j_2$ is always observed), while the weight for Y_{ij_2} is the average between $\mathbb{E}\{1/2(N^T_{U,-j_1} + 1)\}$ (in the case of $X_{1j_1} = 1$, where the set of treated and observed units contains unit $1j_1$) and $\mathbb{E}(1/2N^T_{U,-j_1})$ (in the case of $X_{1j_1} = 0$, where the set of treated and observed units contains neither unit $1j_1$ (not treated) nor unit $2j_1$ (treated but not observed)). Since the weight for Y_{1j_1} is smaller than the weight for Y_{ij_2}, $\mathbb{E}\{E(\boldsymbol{Y}|\boldsymbol{K}_U\boldsymbol{X})\}$ is not equal to $E(\boldsymbol{y}^T|\boldsymbol{k}^T_U)$ any more.

Relatedly, if we assume $R_{ij}(1) = R_{ij}(0)$ for all i and j but not Assumption 2 and define $\bar{\tau}_U \equiv E\{Y_{ij}(1) - Y_{ij}(0) \mid R_{ij}(1) = R_{ij}(0) = 1\}$, $\hat{\tau}_U$ is unbiased for $\bar{\tau}_U$ under *complete* randomization but biased for $\bar{\tau}_U$ under *pairwise* randomization (except for knife-edge situations).

Below, we give more general argument. We assume pairwise randomization (Equations 28, 29, and 30, not 62) again. For $\boldsymbol{K}_G \in \mathbb{U}$ and $h \in \{0, 1, 2\}$, let

$$\begin{aligned} \mathbb{J}^{Ah}_G &\equiv \Big\{ j \Big| \sum_i k^A_{G,ij} = h \Big\} \\ N^{Ah}_G &\equiv \sum_{j \in \mathbb{J}^{Ah}_G} \sum_i K_{G,ij} X^A_{ij} \\ &= \sum_{j \in \mathbb{J}^{Ah}_G} \sum_i k^A_{G,ij} X^A_{ij}. \quad (\because \text{Lemma 8}) \end{aligned} \tag{66}$$

For $j \in \mathbb{J}^{A1}_G$, without loss of generality, suppose $k^A_{G,1j} = 1, k^A_{G,2j} = 0$ (or renumber i accordingly). Specifically,

$$\begin{aligned} N^{A0}_G &= \sum_{j \in \mathbb{J}^{A0}_G} \sum_i 0 \cdot X^A_{ij} \\ &= 0, \end{aligned} \tag{67}$$

and

$$\begin{aligned} N^{A1}_G &= \sum_{j \in \mathbb{J}^{A1}_G} (1 \cdot X^A_{1j} + 0 \cdot X^A_{2j}) \\ &= \sum_{j \in \mathbb{J}^{A1}_G} X^A_{1j}, \end{aligned} \tag{68}$$

and

$$\begin{aligned} N^{A2}_G &= \sum_{j \in \mathbb{J}^{A2}_G} \sum_i 1 \cdot X^A_{ij} \\ &= \sum_{j \in \mathbb{J}^{A2}_G} 1 \quad (\because \text{Equation 2}) \\ &= |\mathbb{J}^{A2}_G| \\ &\equiv n^{A2}_G. \end{aligned} \tag{69}$$

It follows

$$\begin{aligned}
N_G^A &= N_G^{A0} + N_G^{A1} + N_G^{A2} \\
&\qquad (\because \text{Equations 9 and 66, } \mathbb{J}_G^{Ah} \cap \mathbb{J}_G^{Ah'} = \emptyset \text{ for any } h \neq h', \bigcup_{h=0}^{2} \mathbb{J}_G^{Ah} = \{1, 2, \ldots, n\}) \\
&= N_G^{A1} + n_G^{A2}. \quad (\because \text{Equations 67, 68, and 69})
\end{aligned} \tag{70}$$

Note that $0 \leq N_G^{A1} \leq n_{G,\max}^{A1}$ where

$$\begin{aligned}
n_{G,\max}^{A1} &\equiv |\mathbb{J}_G^{A1}| \\
&= \sum_{j \in \mathbb{J}_G^{A1}} 1.
\end{aligned}$$

Assume that $\boldsymbol{K}_G$ does NOT satisfy Condition 1 (that is, for some A, $n_{G,\max}^{A1} \geq 1$). There are ${}_{n_{G,\max}^{A1}}C_{N_G^{A1}}$ ways of assigning treatment to unit $i = 1$ of N_G^{A1} pairs from $n_{G,\max}^{A1}$ pairs in $\mathbb{J}_G^{A1}$ (see also Equation 192). Among them, there are ${}_{n_{G,\max}^{A1}-1}C_{N_G^{A1}-1}$ ways of assigning treatment to unit $1j$, where $j \in \mathbb{J}_G^{A1}$, and unit $i = 1$ of $N_G^{A1} - 1$ pairs from $n_{G,\max}^{A1} - 1$ pairs in $\mathbb{J}_G^{A1} \setminus \{j\}$. Therefore, for $j \in \mathbb{J}_G^{A1}$, it follows

$$\begin{aligned}
\mathbb{E}(X_{ij}^A | N_G^A) &= \Pr(X_{ij}^A = 1 | N_G^A - n_G^{A2}) \quad (\because X_{ij}^A \in \{0, 1\}, n_G^{A2} \text{ is constant}) \\
&= \Pr(X_{ij}^A = 1 | N_G^{A1}) \quad (\because \text{Equation 70}) \\
&= \Pr(X_{ij}^A = 1, N_G^{A1}) \div \Pr(N_G^{A1}) \\
&= \Pr\left(X_{ij}^A = 1, \sum_{j' \in \mathbb{J}_G^{A1}, j' \neq j} X_{1j'}^A = N_G^{A1} - 1\right) \div \Pr(N_G^{A1}) \quad (\because \text{Equation 68}) \\
&= {}_{n_{G,\max}^{A1}-1}C_{N_G^{A1}-1} \div {}_{n_{G,\max}^{A1}}C_{N_G^{A1}} \quad (\because \text{Equations 29 and 30}) \\
&= \frac{(n_{G,\max}^{A1} - 1)!}{\{(n_{G,\max}^{A1} - 1) - (N_G^{A1} - 1)\}!(N_G^{A1} - 1)!} \div \frac{n_{G,\max}^{A1}!}{(n_{G,\max}^{A1} - N_G^{A1})!N_G^{A1}!} \quad (\because \text{Equation 192}) \\
&= \frac{N_G^{A1}}{n_{G,\max}^{A1}}.
\end{aligned} \tag{71}$$

For $j \in \mathbb{J}_G^{A2}$, it holds

$$\begin{aligned}
\mathbb{E}(X_{ij}^A | N_G^A) &= \Pr(X_{ij}^A = 1 | N_G^{A1}) \\
&= \Pr(X_{ij}^A = 1) \\
&= \frac{1}{2}. \quad (\because \text{Lemma 11 (1)})
\end{aligned} \tag{72}$$

In addition, assume $N_G^A \geq 1$ for any $\boldsymbol{X}^A \in \mathbb{X}_{\max}$ (that is, $n_G^{A2} \geq 1$). It follows

$$\begin{aligned}
&\mathbb{E}\{E(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^A)\}\\
&= \mathbb{E}\Big(\frac{1}{N_G^A}\sum_j\sum_i k_{G,ij}^A X_{ij}^A y_{ij}^A\Big) \quad (\because \text{Equations 4 and 9})\\
&= \mathbb{E}\Big[\frac{1}{N_G^A}\mathbb{E}\Big\{\sum_{j\in\mathbb{J}_G^{A0}}\sum_i 0\cdot X_{ij}^A y_{ij}^A + \sum_{j\in\mathbb{J}_G^{A1}}(1\cdot X_{1j}^A y_{1j}^A + 0\cdot X_{2j}^A y_{2j}^A) + \sum_{j\in\mathbb{J}_G^{A2}}\sum_i 1\cdot X_{ij}^A y_{ij}^A \Big| N_G^A\Big\}\Big]\\
&= \mathbb{E}\Big[\frac{1}{N_G^A}\Big\{\sum_{j\in\mathbb{J}_G^{A1}}\mathbb{E}(X_{1j}^A|N_G^A)y_{1j}^A + \sum_{j\in\mathbb{J}_G^{A2}}\sum_i \mathbb{E}(X_{ij}^A|N_G^A)y_{ij}^A\Big\}\Big] \quad (\because \text{Lemma 10})\\
&= \mathbb{E}\Big[\frac{1}{N_G^A}\Big\{\sum_{j\in\mathbb{J}_G^{A1}}\frac{N_G^{A1}}{n_{G,\max}^{A1}}y_{1j}^A + \sum_{j\in\mathbb{J}_G^{A2}}\sum_i \frac{1}{2}y_{ij}^A\Big\}\Big] \quad (\because \text{Equations 71 and 72})\\
&= \mathbb{E}\Big(\frac{N_G^{A1}}{N_G^A}\Big)\frac{1}{n_{G,\max}^{A1}}\sum_{j\in\mathbb{J}_G^{A1}} y_{1j}^A + \mathbb{E}\Big(\frac{1}{N_G^A}\Big)\frac{1}{2}\sum_{j\in\mathbb{J}_G^{A2}}\sum_i y_{ij}^A\\
&= \mathbb{E}\Big(\frac{N_G^{A1}}{N_G^{A1}+n_G^{A2}}\Big)E\{\boldsymbol{y}^A|\boldsymbol{k}_G^A(1-\boldsymbol{k}_{G,-i}^A)\} + \mathbb{E}\Big(\frac{n_G^{A2}}{N_G^{A1}+n_G^{A2}}\Big)E(\boldsymbol{y}^A|\boldsymbol{k}_G^A\boldsymbol{k}_{G,-i}^A).\\
&\quad (\because \text{Equations 4, 9, and 70})
\end{aligned} \tag{73}$$

For integers $\bar{n}, \bar{n}_{\max}$ and a real number p, where $\bar{n}_{\max} \geq \bar{n} \geq 0, 1 > p > 0$, we denote binomial distribution as

$$B(\bar{n}|\bar{n}_{\max}, p) \equiv {}_{\bar{n}_{\max}}C_{\bar{n}}\,p^{\bar{n}}(1-p)^{\bar{n}_{\max}-\bar{n}}$$

(see also Equation 192). According to Equations 29 and 30,

$$\begin{aligned}
N_G^{A1} &\sim B\Big(\bar{n}\Big|n_{G,\max}^{A1}, \frac{1}{2}\Big)\\
&= {}_{n_{G,\max}^{A1}}C_{\bar{n}}\Big(\frac{1}{2}\Big)^{n_{G,\max}^{A1}}.
\end{aligned} \tag{74}$$

Recall that expectation of binomial distribution is

$$\sum_{\bar{n}=0}^{\bar{n}_{\max}} B(\bar{n}|\bar{n}_{\max}, p)\bar{n} = p\bar{n}_{\max}. \tag{75}$$

Thus,

$$\begin{aligned}
n_{G,\mathrm{mean}}^{A1} &\equiv \mathbb{E}(N_G^{A1})\\
&= \frac{1}{2}n_{G,\max}^{A1}. \quad (\because \text{Equation 75})
\end{aligned} \tag{76}$$

In general, when $q^{(1)}, q^{(2)} > 0$, since $\frac{1}{q}$ is convex,

$$\frac{1}{q^{(1)}-q^{(2)}} + \frac{1}{q^{(1)}+q^{(2)}} > \frac{2}{q^{(1)}}. \tag{77}$$

For $q > 0$,

$$\mathbb{E}\Big(\frac{1}{N_G^{A1}+q}\Big) = \Big(\frac{1}{2}\Big)^{n_{G,\max}^{A1}}\sum_{\bar{n}=0}^{n_{G,\max}^{A1}}\frac{1}{\bar{n}+q}\,{}_{n_{G,\max}^{A1}}C_{\bar{n}} \quad (\because \text{Equation 74}) \tag{78}$$

When $n_{G,\max}^{A1}$ is an even number, let

$$h_{\max} \equiv \frac{n_{G,\max}^{A1}}{2}$$

and note that, for $h = \{1, 2, \ldots, h_{\max}\}$,

$$
\begin{aligned}
{}_{n^{A1}_{G,\max}}C_{n^{A1}_{G,\text{mean}}-h} &= \frac{n^{A1}_{G,\max}!}{(n^{A1}_{G,\text{mean}} - h)!\{n^{A1}_{G,\max} - (n^{A1}_{G,\text{mean}} - h)\}!} \quad (\because \text{Equation 192}) \\
&= \frac{n^{A1}_{G,\max}!}{(\frac{n^{A1}_{G,\max}}{2} - h)!\{n^{A1}_{G,\max} - (\frac{n^{A1}_{G,\max}}{2} - h)\}!} \quad (\because \text{Equation 76}) \\
&= \frac{n^{A1}_{G,\max}!}{\{n^{A1}_{G,\max} - (n^{A1}_{G,\text{mean}} + h)\}!(n^{A1}_{G,\text{mean}} + h)!} \quad (\because \text{Equation 76}) \\
&= {}_{n^{A1}_{G,\max}}C_{n^{A1}_{G,\text{mean}}+h} \quad (\because \text{Equation 192})
\end{aligned} \tag{79}
$$

Equation 78 leads to

$$
\begin{aligned}
&\left(\frac{1}{2}\right)^{n^{A1}_{G,\max}} \Big\{ \sum_{h=1}^{h_{\max}} \Big(\frac{1}{n^{A1}_{G,\text{mean}} - h + q} + \frac{1}{n^{A1}_{G,\text{mean}} + h + q} \Big) {}_{n^{A1}_{G,\max}}C_{n^{A1}_{G,\text{mean}}-h} \\
&+ \frac{1}{n^{A1}_{G,\text{mean}} + q} {}_{n^{A1}_{G,\max}}C_{n^{A1}_{G,\text{mean}}} \Big\} (\because \text{Equation 79}) \\
&> \left(\frac{1}{2}\right)^{n^{A1}_{G,\max}} \Big\{ \sum_{h=1}^{h_{\max}} \frac{2}{n^{A1}_{G,\text{mean}} + q} {}_{n^{A1}_{G,\max}}C_{n^{A1}_{G,\text{mean}}-h} + \frac{1}{n^{A1}_{G,\text{mean}} + q} {}_{n^{A1}_{G,\max}}C_{n^{A1}_{G,\text{mean}}} \Big\} \\
&\quad (\because \text{Equation 77 where } q^{(1)} = n^{A1}_{G,\text{mean}} + q, q^{(2)} = h) \\
&= \frac{1}{n^{A1}_{G,\text{mean}} + q} \sum_{\bar{n}=0}^{n^{A1}_{G,\max}} {}_{n^{A1}_{G,\max}}C_{\bar{n}} \left(\frac{1}{2}\right)^{n^{A1}_{G,\max}} \\
&= \frac{1}{n^{A1}_{G,\text{mean}} + q} \quad (\because \text{Equation 74 and axiom of probability})
\end{aligned} \tag{80}
$$

When $n^{A1}_{G,\max}$ is an odd number, let

$$
h_{\max} \equiv \frac{n^{A1}_{G,\max} + 1}{2}
$$

and note that, for $h = \{1, 2, \ldots, h_{\max}\}$ and $h' = h - 0.5$,

$$
{}_{n^{A1}_{G,\max}}C_{n^{A1}_{G,\text{mean}}-h'} = {}_{n^{A1}_{G,\max}}C_{n^{A1}_{G,\text{mean}}+h'}. \tag{81}
$$

Equation 78 leads to

$$
\begin{aligned}
&\left(\frac{1}{2}\right)^{n^{A1}_{G,\max}} \sum_{h=1}^{h_{\max}} \Big(\frac{1}{n^{A1}_{G,\text{mean}} - h' + q} + \frac{1}{n^{A1}_{G,\text{mean}} + h' + q} \Big) {}_{n^{A1}_{G,\max}}C_{n^{A1}_{G,\text{mean}}-h'} \quad (\because \text{Equation 81}) \\
&> \left(\frac{1}{2}\right)^{n^{A1}_{G,\max}} \sum_{h=1}^{h_{\max}} \frac{2}{n^{A1}_{G,\text{mean}} + q} {}_{n^{A1}_{G,\max}}C_{n^{A1}_{G,\text{mean}}-h'} \\
&= \frac{1}{n^{A1}_{G,\text{mean}} + q} \sum_{\bar{n}=0}^{n^{A1}_{G,\max}} {}_{n^{A1}_{G,\max}}C_{\bar{n}} \left(\frac{1}{2}\right)^{n^{A1}_{G,\max}} \\
&= \frac{1}{n^{A1}_{G,\text{mean}} + q}
\end{aligned} \tag{82}
$$

Therefore, whichever $n^{A1}_{G,\max}$ is an even or odd number, Equations 78, 80, and 82 lead to

$$
\mathbb{E}\Big(\frac{1}{N^{A1}_G + q}\Big) > \frac{1}{n^{A1}_{G,\text{mean}} + q}. \tag{83}
$$

Therefore,

$$\mathbb{E}\Big(\frac{n_G^{A2}}{N_G^{A1}+n_G^{A2}}\Big) > \frac{n_G^{A2}}{n_{G,\text{mean}}^{A1}+n_G^{A2}} \quad (\because \text{Lemma 10 (2), Equation 83 where } q = n_G^{A2} > 0) \tag{84}$$

and

$$\begin{aligned}\mathbb{E}\Big(\frac{N_G^{A1}}{N_G^{A1}+n_G^{A2}}\Big) &= \mathbb{E}\Big(1-\frac{n_G^{A2}}{N_G^{A1}+n_G^{A2}}\Big)\\ &< 1-\frac{n_G^{A2}}{n_{G,\text{mean}}^{A1}+n_G^{A2}} \quad (\because \text{Lemma 10, Equation 84})\\ &= \frac{n_{G,\text{mean}}^{A1}}{n_{G,\text{mean}}^{A1}+n_G^{A2}}.\end{aligned} \tag{85}$$

It holds

$$\begin{aligned}&E(\boldsymbol{y}^A|\boldsymbol{k}_G^A)\\ &= \frac{n_{G,\max}^{A1}}{n_{G,\max}^{A1}+2n_G^{A2}}E\{\boldsymbol{y}^A|\boldsymbol{k}_G^A(1-\boldsymbol{k}_{G,-i}^A)\} + \frac{2n_G^{A2}}{n_{G,\max}^{A1}+2n_G^{A2}}E(\boldsymbol{y}^A|\boldsymbol{k}_G^A\boldsymbol{k}_{G,-i}^A) \quad (\because \text{Lemma 3 (5)})\\ &= \frac{n_{G,\text{mean}}^{A1}}{n_{G,\text{mean}}^{A1}+n_G^{A2}}E\{\boldsymbol{y}^A|\boldsymbol{k}_G^A(1-\boldsymbol{k}_{G,-i}^A)\} + \frac{n_G^{A2}}{n_{G,\text{mean}}^{A1}+n_G^{A2}}E(\boldsymbol{y}^A|\boldsymbol{k}_G^A\boldsymbol{k}_{G,-i}^A). \quad (\because \text{Equation 76})\end{aligned} \tag{86}$$

According to Equations 73, 84, 85, and 86, it follows that $\mathbb{E}\{E(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^A)\}$ is not equal to $E(\boldsymbol{y}^A|\boldsymbol{k}_G^A)$ without knife-edge situation. Thus, Equation 43 does not hold; even if we assume Condition 2, it does not follow that $\hat{\tau}_G$ is unbiased for $\bar{\tau}$; if we assume $\boldsymbol{k}_G^T = \boldsymbol{k}_G^C \equiv \boldsymbol{k}_G$, it follows that $\hat{\tau}_G$ is not necessarily unbiased for $E(\boldsymbol{\tau}|\boldsymbol{k}_G)$ but for a particular principal effect:

$$\begin{aligned}\mathbb{E}(\hat{\tau}_G) &= \mathbb{E}\{E(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^T) - E(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^C)\}\\ &\quad (\because N_G^A \geq 1, \mathbb{X}_{\text{def}}(\hat{\tau}_G) = \mathbb{X}_{\text{def}}\{E(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^A) - E(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^C)\} = \mathbb{X}_{\max})\\ &= \Big[\mathbb{E}\Big(\frac{N_G^{\cdot 1}}{N_G^{\cdot 1}+n_G^{\cdot 2}}\Big)E\{\boldsymbol{y}^T|\boldsymbol{k}_G(1-\boldsymbol{k}_{G,-i})\} + \mathbb{E}\Big(\frac{n_G^{\cdot 2}}{N_G^{\cdot 1}+n_G^{\cdot 2}}\Big)E(\boldsymbol{y}^T|\boldsymbol{k}_G\boldsymbol{k}_{G,-i})\Big]\\ &\quad - \Big[\mathbb{E}\Big(\frac{N_G^{\cdot 1}}{N_G^{\cdot 1}+n_G^{\cdot 2}}\Big)E\{\boldsymbol{y}^C|\boldsymbol{k}_G(1-\boldsymbol{k}_{G,-i})\} + \mathbb{E}\Big(\frac{n_G^{\cdot 2}}{N_G^{\cdot 1}+n_G^{\cdot 2}}\Big)E(\boldsymbol{y}^C|\boldsymbol{k}_G\boldsymbol{k}_{G,-i})\Big]\\ &\quad (\because \text{Equation 73}, \boldsymbol{k}_G^T = \boldsymbol{k}_G^C \equiv \boldsymbol{k}_G)\\ &= \mathbb{E}\Big(\frac{N_G^{\cdot 1}}{N_G^{\cdot 1}+n_G^{\cdot 2}}\Big)E\{\boldsymbol{\tau}|\boldsymbol{k}_G(1-\boldsymbol{k}_{G,-i})\} + \mathbb{E}\Big(\frac{n_G^{\cdot 2}}{N_G^{\cdot 1}+n_G^{\cdot 2}}\Big)E(\boldsymbol{\tau}|\boldsymbol{k}_G\boldsymbol{k}_{G,-i})\\ &\quad (\because \text{Lemma 3 (1), Equation 6})\end{aligned} \tag{87}$$

where the corresponding principal stratum is $\{ij|k_{G,ij}(1-k_{G,-ij})=1\} \cup \{ij|k_{G,ij}k_{G,-ij}=1\}$; this is a local *weighted* average treatment effect.

Equations 84 and 85 imply that the weight for $j \in \mathbb{J}_G^{A1}$ is smaller than that for $j \in \mathbb{J}_G^{A2}$:

$$\mathbb{E}\Big(\frac{N_G^{A1}}{N_G^{A1}+n_G^{A2}}\Big)\frac{1}{n_{G,\text{mean}}^{A1}} < \frac{1}{n_{G,\text{mean}}^{A1}+n_G^{A2}} < \mathbb{E}\Big(\frac{n_G^{A2}}{N_G^{A1}+n_G^{A2}}\Big)\frac{1}{n_G^{A2}}.$$

Below, I will give an alternative, and more intuitive, reason. Denote

$$N_{G,-j}^A \equiv \sum_j \sum_i K_{G,ij}X_{ij}^A.$$

For $j \in \mathbb{J}_G^{A1}$, assume $N_{G,-j}^A \geq 1$ for any $\boldsymbol{X}^A \in \mathbb{X}_{\max}$. It follows

$$\begin{aligned}
&\mathbb{E}\{E(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^A)\}\\
&= \mathbb{E}\{E(\boldsymbol{y}^A|\boldsymbol{k}_G^A\boldsymbol{X}^A)|X_{1j}^A = 0\}\Pr(X_{1j}^A = 0) + \mathbb{E}\{E(\boldsymbol{y}^A|\boldsymbol{k}_G^A\boldsymbol{X}^A)|X_{1j}^A = 1\}\Pr(X_{1j}^A = 1)\\
&= \mathbb{E}\Big(\frac{1}{N_{G,-j}^A}\sum_{j'\neq j}\sum_i k_{G,ij'}^A X_{ij'}^A y_{ij'}^A\Big)\frac{1}{2} + \mathbb{E}\Big\{\frac{1}{N_{G,-j}^A+1}\Big(\sum_{j'\neq j}\sum_i k_{G,ij'}^A X_{ij'}^A y_{ij'}^A + y_{1j}^A\Big)\Big\}\frac{1}{2}\\
&\quad (\because \text{Equations 4 and 29, Lemma 11 (1)}, k_{G,1j}^A = 1, k_{G,2j}^A = 0, N_{G,-j}^A \geq 1)\\
&= \frac{1}{2}\mathbb{E}\Big[\mathbb{E}\Big\{\frac{1}{N_{G,-j}^A}\sum_{j'\neq j}\sum_i k_{G,ij'}^A X_{ij'}^A y_{ij'}^A + \frac{1}{N_{G,-j}^A+1}\Big(\sum_{j'\neq j}\sum_i k_{G,ij'}^A X_{ij'}^A y_{ij'}^A + y_{1j}^A\Big)\Big|N_{G,-j}^A\Big\}\Big]\\
&= \frac{1}{2}\mathbb{E}\Big\{\Big(\frac{1}{N_{G,-j}^A} + \frac{1}{N_{G,-j}^A+1}\Big)\sum_{j'\neq j}\sum_i k_{G,ij'}^A \mathbb{E}(X_{ij'}^A|N_{G,-j}^A) y_{ij'}^A + \frac{1}{N_{G,-j}^A+1}y_{1j}^A\Big\}\\
&= \sum_{j'\in\mathbb{J}_G^{A2}}\sum_i \frac{1}{2}\mathbb{E}\Big\{\Big(\frac{1}{N_{G,-j}^A} + \frac{1}{N_{G,-j}^A+1}\Big)\frac{1}{2}\Big\}1\cdot y_{ij'}^A\\
&\quad + \sum_{j'\in\mathbb{J}_G^{A1}\setminus\{j\}} \frac{1}{2}\mathbb{E}\Big\{\Big(\frac{1}{N_{G,-j}^A} + \frac{1}{N_{G,-j}^A+1}\Big)\frac{N_{G,-j}^A}{n_{G,\max}^{A1}-1}\Big\}(1\cdot y_{1j'}^A + 0\cdot y_{2j'}^A)\\
&\quad + \frac{1}{2}\mathbb{E}\Big(\frac{1}{N_{G,-j}^A+1}\Big)y_{1j}^A
\end{aligned}$$

Thus, the weight for $j \in \mathbb{J}_G^{A1}$ is $\frac{1}{2}\mathbb{E}\Big(\frac{1}{N_{G,-j}^A+1}\Big)$ and that for $j' \in \mathbb{J}_G^{A2}$ is $\frac{1}{2}\mathbb{E}\Big\{\Big(\frac{1}{N_{G,-j}^A} + \frac{1}{N_{G,-j}^A+1}\Big)\frac{1}{2}\Big\}$, while the difference is

$$\begin{aligned}
&\frac{1}{2}\mathbb{E}\Big\{\Big(\frac{1}{N_{G,-j}^A} + \frac{1}{N_{G,-j}^A+1}\Big)\frac{1}{2}\Big\} - \frac{1}{2}\mathbb{E}\Big(\frac{1}{N_{G,-j}^A+1}\Big)\\
&= \frac{1}{4}\mathbb{E}\Big(\frac{1}{N_{G,-j}^A} - \frac{1}{N_{G,-j}^A+1}\Big)\\
&= \frac{1}{4}\mathbb{E}\Big\{\frac{1}{N_{G,-j}^A(N_{G,-j}^A+1)}\Big\}\\
&> 0.
\end{aligned}$$

For instance, when $\boldsymbol{K}_G = \boldsymbol{K}_U$ and $A = T$, the denominator of the weight for $j \in \mathbb{J}_U^{T1}$ is the number of other treated and observed units ($N_{U,-j}^T$) and j itself ($+1$), though that for $j' \in \mathbb{J}_U^{T2}$ is the number of treated and observed units including j ($N_{U,-j}^T + 1$) in the case of $X_{ij}^T = 1$ (half cases) but the number of treated and observed units excluding j ($N_{U,-j}^T$) in the case of $X_{ij}^T = 0$ (the other half cases). Note that $N_{U,-j}^T$ does not depend on $X_{ij'}^T$. Obviously, the former weight is smaller than the latter.

2.3. Variance

Variance. The sample deviation is defined as

$$\boldsymbol{D}(\boldsymbol{Q}|\boldsymbol{Z}) \quad \equiv \quad \boldsymbol{Q} - \boldsymbol{E}(\boldsymbol{Q}|\boldsymbol{Z}). \tag{88}$$

where $\boldsymbol{E}(\boldsymbol{Q}|\boldsymbol{Z}) \equiv \{E(\boldsymbol{Q}|\boldsymbol{Z})\}_{11}^{2n}$. The weighted covariance and weighted variance operators are defined as

$$V(\boldsymbol{Q}^{(1)}, \boldsymbol{Q}^{(2)}|\boldsymbol{Z}) \equiv V_{h=1}^{2}(\boldsymbol{Q}^{(h)}|\boldsymbol{Z}) \equiv E\Big\{\prod_{h=1}^{2} \boldsymbol{D}(\boldsymbol{Q}^{(h)}|\boldsymbol{Z})\Big|\boldsymbol{Z}\Big\} \tag{89}$$

$$V^2(\boldsymbol{Q}|\boldsymbol{Z}) \equiv V_{h=1}^{2}(\boldsymbol{Q}|\boldsymbol{Z}). \tag{90}$$

It follows that

$$\begin{aligned} V_{h=1}^{2}(\boldsymbol{Q}^{(h)}|\boldsymbol{0}) &= E\Big\{\prod_{h=1}^{2} \boldsymbol{D}(\boldsymbol{Q}^{(h)}|\boldsymbol{0})\Big|\boldsymbol{0}\Big\} \quad (\because \text{Equation 89}) \\ &= 0. \quad (\because \text{Equation 5}) \end{aligned} \tag{91}$$

The population deviation, covariance operator, and variance operator are defined as

$$\mathbb{D}\{f(\boldsymbol{X}^A)\} \equiv f(\boldsymbol{X}^A) - \mathbb{E}\{f(\boldsymbol{X}^A)\} \tag{92}$$

$$\mathbb{V}\{f^{(1)}(\boldsymbol{X}^A), f^{(2)}(\boldsymbol{X}^A)\} \equiv \mathbb{V}_{h=1}^{2}\{f^{(h)}(\boldsymbol{X}^A)\} \equiv \mathbb{E}\Big[\prod_{h=1}^{2} \mathbb{D}\{f^{(h)}(\boldsymbol{X}^A)\}\Big] \tag{93}$$

$$\begin{aligned} \mathbb{V}^2\{f(\boldsymbol{X}^A))\} &\equiv \mathbb{V}_{h=1}^{2}\{f(\boldsymbol{X}^A)\} \\ &= \mathbb{E}[\{f(\boldsymbol{X}^A)\}^2] - [\mathbb{E}\{f(\boldsymbol{X}^A)\}]^2 \end{aligned} \tag{94}$$

<u>Number of Units.</u> For $\boldsymbol{K}_G \in \mathbb{U}$, let

$$\begin{aligned} \boldsymbol{K}_G^{SQ} &\equiv \boldsymbol{K}^{SQ}(\boldsymbol{K}_G) \\ &\equiv \boldsymbol{K}_G \boldsymbol{K}_{G,-i} \\ &= \boldsymbol{K}_{G,-i} \boldsymbol{K}_G \\ &= \boldsymbol{K}_{G,-i}^{SQ} \quad (\because \text{by definition}). \end{aligned} \tag{95}$$

We also define

$$\begin{aligned} N_G^{TC} &\equiv N^{TC}(\boldsymbol{K}_G) \\ &\equiv \sum_j \sum_i K_{G,ij} X_{ij}^A K_{G,-ij} X_{-ij}^{-A} \\ &= \sum_j \sum_i K_{G,ij}^{SQ} (X_{ij}^A)^2 \quad (\because \text{Equation 95, Lemma 2 (1)}) \\ &= \sum_j \sum_i K_{G,ij}^{SQ} X_{ij}^A \quad (\because \boldsymbol{X}^A \in \mathbb{U}, \text{Lemma 1 (5)}) \end{aligned} \tag{96}$$

This is constant irrespective of $A \in \{T, C\}$ because

$$\begin{aligned} \sum_j \sum_i K_{G,ij}^{SQ} X_{ij}^{-A} &= \sum_j \sum_i K_{G,-ij}^{SQ} X_{-ij}^{-A} \quad (\because \text{exchanging } i \text{ and } -i, \text{Lemma 1 (4)}) \\ &= \sum_j \sum_i K_{G,ij}^{SQ} X_{ij}^{A} \quad (\because \text{Equation 95, Lemma 2 (1)}) \end{aligned}$$

Specifically, when $\boldsymbol{K}_G = \boldsymbol{K}_F$,

$$\begin{aligned}
\boldsymbol{K}_F^{SQ} &\equiv \boldsymbol{K}^{SQ}(\boldsymbol{K}_F) \\
&= \boldsymbol{K}_F \boldsymbol{K}_{F,-i} \quad (\because \text{Equation 95}) \\
&= \boldsymbol{1} \cdot \boldsymbol{1}_{-i} \quad (\because \text{Equation 12}) \\
&= \boldsymbol{K}_F \quad (\because \text{Equation 12}) \qquad (97) \\
N_F^{TC} &\equiv \boldsymbol{N}^{TC}(\boldsymbol{K}_F) \\
&= \sum_j \sum_i K_{F,ij}^{SQ} X_{ij}^A \quad (\because \text{Equation 96}) \\
&= \sum_j \sum_i K_{F,ij} X_{ij}^A \quad (\because \text{Equation 97}) \\
&= N_F \quad (\because \text{Equations 9 and 12}) \qquad (98)
\end{aligned}$$

when $\boldsymbol{K}_G = \boldsymbol{K}_U$,

$$\begin{aligned}
\boldsymbol{K}_U^{SQ} &\equiv \boldsymbol{K}^{SQ}(\boldsymbol{K}_U) \\
&= \boldsymbol{K}_U \boldsymbol{K}_{U,-i} \quad (\because \text{Equation 95}) \\
&= \boldsymbol{R}\boldsymbol{R}_{-i} \quad (\because \text{Equation 14}) \\
&= \boldsymbol{K}_P \quad (\because \text{Equation 16}) \qquad (99) \\
N_U^{TC} &\equiv \boldsymbol{N}^{TC}(\boldsymbol{K}_U) \\
&= \sum_j \sum_i K_{U,ij}^{SQ} X_{ij}^A \quad (\because \text{Equation 96}) \\
&= \sum_j \sum_i K_{P,ij} X_{ij}^A \quad (\because \text{Equation 99}) \\
&= N_P \quad (\because \text{Equations 9 and 16}) \qquad (100)
\end{aligned}$$

and, when $\boldsymbol{K}_G = \boldsymbol{K}_P$,

$$\begin{aligned}
\boldsymbol{K}_P^{SQ} &\equiv \boldsymbol{K}^{SQ}(\boldsymbol{K}_P) \\
&= \boldsymbol{K}_P \boldsymbol{K}_{P,-i} \quad (\because \text{Equation 95}) \\
&= \boldsymbol{R}\boldsymbol{R}_{-i}\boldsymbol{R}_{-i}\boldsymbol{R} \quad (\because \text{Equation 16}) \\
&= \boldsymbol{R}\boldsymbol{R}_{-i} \quad (\because \boldsymbol{R} \in \mathbb{U}, \text{Lemma 1 (5)}) \\
&= \boldsymbol{K}_P \quad (\because \text{Equation 16}) \qquad (101) \\
N_P^{TC} &\equiv \boldsymbol{N}^{TC}(\boldsymbol{K}_P) \\
&= \sum_j \sum_i K_{P,ij}^{SQ} X_{ij}^A \quad (\because \text{Equation 96}) \\
&= \sum_j \sum_i K_{P,ij} X_{ij}^A \quad (\because \text{Equation 101}) \\
&= N_P \quad (\because \text{Equations 9 and 16}) \qquad (102)
\end{aligned}$$

Let

$$\boldsymbol{k}_G^{TC} \equiv \boldsymbol{k}^{TC}(\boldsymbol{K}_G) \equiv \boldsymbol{k}_G^T \boldsymbol{k}_{G,-i}^C \qquad (103)$$

and, under Condition 1,

$$n_G^{TC} \equiv n(\boldsymbol{k}_G^{TC}) \equiv \sum_j k_{G,\cdot j}^{TC}. \qquad (104)$$

Specifically, under Assumption 2,

$$\begin{aligned}
\boldsymbol{k}_U^{TC} &\equiv \boldsymbol{k}^{TC}(\boldsymbol{K}_U) \\
&= \boldsymbol{k}_U^T \boldsymbol{k}_{U,-i}^C \quad (\because \text{Equation 103}) \\
&= \boldsymbol{r}^T \boldsymbol{r}^C \quad (\because \text{Equation 14, Assumption 2}) \\
&= \boldsymbol{k}_P \quad (\because \text{Lemma 9 (2)})
\end{aligned} \qquad (105)$$

and

$$\begin{aligned} n_U^{TC} &\equiv n(\boldsymbol{k}_U^{TC}) \\ &= \sum_j k_{U,\cdot j}^{TC} \quad (\because \text{Equation 104}) \\ &= \sum_j k_{P,\cdot j} \quad (\because \text{Equation 105}) \\ &= n_P. \quad (\because \text{Lemma 9 (1), Equation 31}). \end{aligned} \tag{106}$$

Under Assumption 3,

$$\begin{aligned} \boldsymbol{k}_P^{TC} &\equiv \boldsymbol{k}^{TC}(\boldsymbol{K}_P) \\ &= \boldsymbol{k}_P^T \boldsymbol{k}_{P,-i}^C \quad (\because \text{Equation 103}) \\ &= \boldsymbol{k}_P \boldsymbol{k}_{P,-i} \quad (\because \text{Lemma 9 (1)}) \\ &= (\boldsymbol{k}_P)^2 \quad (\because \text{Assumption 3}) \\ &= \boldsymbol{k}_P \quad (\because \boldsymbol{k}_P \in \mathbb{U}, \text{Lemma 1 (5)}) \end{aligned} \tag{107}$$

and,

$$\begin{aligned} n_P^{TC} &\equiv n(\boldsymbol{k}_P^{TC}) \\ &= \sum_j k_{P,\cdot j}^{TC} \quad (\because \text{Equation 104}) \\ &= \sum_j k_{P,\cdot j} \quad (\because \text{Equation 107}) \\ &= n_P. \quad (\because \text{Lemma 9 (2)}) \end{aligned} \tag{108}$$

For $h \in \{1, 2\}$, let $\boldsymbol{k}_G^{(h)} \in \{\boldsymbol{k}_G^T, \boldsymbol{k}_G^C\}$. Under Condition 1, define

$$n_G^{(1)\cdot(2)} \equiv \sum_j \prod_{h=1}^{2} k_{G,\cdot j}^{(h)}, \tag{109}$$

which is constant irrespective of $\boldsymbol{X}^A$.

LEMMA 14 (NUMBER OF UNITS). *Suppose that* $\boldsymbol{K}_G \in \mathbb{U}$.
(1) When $\boldsymbol{K}_G = \boldsymbol{K}_{G,-i}$,

$$\begin{aligned} \boldsymbol{K}_G^{SQ} &= \boldsymbol{K}_G \\ N_G^T = N_G^C &= N_G^{TC} \equiv N_G. \end{aligned}$$

(2) Under Condition 1,

$$\begin{aligned} \prod_A \boldsymbol{k}_G^A &= \boldsymbol{k}_G^{TC} = \boldsymbol{k}_{G,-i}^{TC} \\ N_G^{TC} &= n_G^{TC} \\ n_G^{TC} &\leq n_G^A \end{aligned}$$

(3) For $h \in \{1, 2\}$, *let* $\boldsymbol{k}_G^{(h)} \in \{\boldsymbol{k}_G^T, \boldsymbol{k}_G^C\}$. *When* $\boldsymbol{k}_G^{(1)} = \boldsymbol{k}_G^{(2)} \equiv \boldsymbol{k}_G$,

$$\begin{aligned} \prod_{h=1}^{2} \boldsymbol{k}_G^{(h)} &= \boldsymbol{k}_G \\ n_G^{(1)} = n_G^{(2)} &= n_G^{(1)\cdot(2)} \equiv n_G. \end{aligned}$$

(4) Under Condition 1, if and only if $n_G^A = 0$, it holds $\boldsymbol{k}_G^A = \mathbf{0}$. If and only if $n_G^{TC} = 0$, it holds $\boldsymbol{k}_G^{TC} = \mathbf{0}$.

(5)

$$N_G^{TC} \leq N_G^A$$

PROOF. (1) Note

$$\begin{aligned} \boldsymbol{K}_G^{SQ} &= \boldsymbol{K}_G \boldsymbol{K}_{G,-i} \quad (\because \text{Equation 95}) \\ &= (\boldsymbol{K}_G)^2 \quad (\because \boldsymbol{K}_G = \boldsymbol{K}_{G,-i}) \\ &= \boldsymbol{K}_G \quad (\because \boldsymbol{K}_G \in \mathbb{U}, \text{Lemma 1 (5)}) \end{aligned} \tag{110}$$

It follows

$$\begin{aligned} N_G^A &= \sum_j \sum_i K_{G,ij} X_{ij}^A \quad (\because \text{Equation 9}) \\ &= \sum_j \sum_i K_{G,-ij} X_{-ij}^{-A} \quad (\because \boldsymbol{K}_G = \boldsymbol{K}_{G,-i}, \text{Lemma 2 (1)}) \\ &= \sum_j \sum_i K_{G,ij} X_{ij}^{-A} \quad (\because \text{exchanging } i \text{ and } -i, \text{ Lemma 1 (4)}) \\ &= N_G^{-A} \\ &\equiv N_G \end{aligned} \tag{111}$$

It holds that

$$\begin{aligned} N_G^{TC} &= \sum_j \sum_i K_{G,ij}^{SQ} X_{ij}^A \quad (\because \text{Equation 96}) \\ &= \sum_j \sum_i K_{G,ij} X_{ij}^A \quad (\because \text{Equation 110}) \\ &= N_G^A \quad (\because \text{Equation 9}) \\ &= N_G \quad (\because \text{Equation 111}) \end{aligned}$$

(2) It holds

$$\begin{aligned} \prod_A \boldsymbol{k}_G^A &= \boldsymbol{k}_G^T \boldsymbol{k}_{G,-i}^C \quad (\because \text{Condition 1}) \\ &= \boldsymbol{k}_G^{TC} \quad (\because \text{Equation 103}) \\ \prod_A \boldsymbol{k}_G^A &= \boldsymbol{k}_{G,-i}^T \boldsymbol{k}_G^C \quad (\because \text{Condition 1}) \\ &= \boldsymbol{k}_{G,-i}^{TC} \quad (\because \text{Equation 103}) \end{aligned}$$

and

$$\begin{aligned} N_G^{TC} &= \sum_j \sum_i K_{G,ij} X_{ij}^T K_{G,-ij} X_{-ij}^C \quad (\because \text{Equation 96}) \\ &= \sum_j \sum_i k_{G,ij}^T X_{ij}^T k_{G,-ij}^C X_{-ij}^C \quad (\because \text{Lemma 8}) \\ &= \sum_j k_{G,\cdot j}^T k_{G,\cdot j}^C \sum_i X_{ij}^T \quad (\because \text{Condition 1, Lemma 2 (4)}) \\ &= \sum_j \prod_A k_{G,\cdot j}^A \cdot 1 \quad (\because \text{Equation 2}) \\ &= n_G^{TC} \quad (\because \text{Equation 109}) \end{aligned} \tag{112}$$

and

$$\begin{aligned} k^{TC}_{G,ij} &= k^{T}_{G,ij} k^{C}_{G,-ij} \quad (\because \text{Equation 103}) \\ &\leq k^{A}_{G,ij} \quad (\because k^{A}_{G,ij} \in \{0,1\}) \\ \therefore n^{TC}_{G} &\leq n^{A}_{G} \quad (\because \text{Equations 31 and 104}) \end{aligned}$$

(3)

$$\begin{aligned} \prod_{h=1}^{2} \boldsymbol{k}^{(h)}_{G} &= (\boldsymbol{k}_G)^2 \quad (\because \boldsymbol{k}^{(1)}_{G} = \boldsymbol{k}^{(2)}_{G} = \boldsymbol{k}_G) \\ &= \boldsymbol{k}_G \quad (\because \boldsymbol{k}_G \in \mathbb{U}, \text{Lemma 1 (5)}) \end{aligned}$$

It is obvious that $n^{(1)}_{G} = n^{(2)}_{G} = n_G$.

$$\begin{aligned} n^{(1)\cdot(2)}_{G} &= \sum_j \prod_{h=1}^{2} k^{(h)}_{G,\cdot j} \quad (\because \text{Equation 109}) \\ &= \sum_j (k_{G,\cdot j})^2 \quad (\because \boldsymbol{k}^{(1)}_{G} = \boldsymbol{k}^{(2)}_{G} = \boldsymbol{k}_G) \\ &= \sum_j k_{G,\cdot j} \quad (\because \boldsymbol{k}_G \in \mathbb{U}, \text{Lemma 1 (5)}) \\ &= n_G \quad (\because \text{Equations 36}) \end{aligned} \tag{113}$$

(4) If $\boldsymbol{k}^{A}_{G} = \boldsymbol{0}$, it follows that $n^{A}_{G} = 0$ due to Equation 31.

Suppose that $\boldsymbol{k}^{A}_{G} \neq \boldsymbol{0}$. Since $\boldsymbol{K}_G \in \mathbb{U}$, there is j such that $k^{A}_{G,\cdot j} = 1$ (for, otherwise, $\boldsymbol{k}^{A}_{G} = \boldsymbol{0}$). Thus, according to Equation 31, $n^{A}_{G} > 0$.

Therefore, $\boldsymbol{k}^{A}_{G} = \boldsymbol{0}$ is equivalent to $n^{A}_{G} = 0$.

Similarly, we can show that $\boldsymbol{k}^{TC}_{G} = \boldsymbol{0}$ is equivalent to $n^{TC}_{G} = 0$.

(5)

$$\begin{aligned} K^{SQ}_{G,ij} &= K_{G,ij} K_{G,-ij} \quad (\because \text{Equation 103}) \\ &\leq K_{G,ij} \quad (\because K_{G,-ij} \in \{0,1\}) \\ \therefore N^{TC}_{G} &\leq N^{A}_{G} \quad (\because \text{Equations 9 and 96 }) \end{aligned}$$

□

LEMMA 15 (ARITHMETIC OF DEVIATION AND VARIANCE). *(1) When* $\boldsymbol{Z}^{(2)} \neq \boldsymbol{0}$,

$$\boldsymbol{D}\{\boldsymbol{D}(\boldsymbol{Q}|\boldsymbol{Z}^{(1)})|\boldsymbol{Z}^{(2)}\} = \boldsymbol{D}(\boldsymbol{Q}|\boldsymbol{Z}^{(2)})$$

(2)

$$V^{2}_{h=1}\{\boldsymbol{D}(\boldsymbol{Q}^{(h)}|\boldsymbol{Z})|\boldsymbol{Z}\} = V^{2}_{h=1}(\boldsymbol{Q}^{(h)}|\boldsymbol{Z})$$

(3) When $\boldsymbol{Z}^{(h)}\boldsymbol{U} = \boldsymbol{U} \in \mathbb{U}$ *for* $h \in \{1,2\}$,

$$V^{2}_{h=1}(\boldsymbol{Z}^{(h)}\boldsymbol{Q}^{(h)}|\boldsymbol{U}) = V^{2}_{h=1}(\boldsymbol{Q}^{(h)}|\boldsymbol{U})$$

(4)

$$V^{2}_{h=1}(\boldsymbol{q}^{(h)}_{const} + q^{(h)}_{multi}\boldsymbol{Q}^{(h)}|\boldsymbol{Z}) = \Big(\prod_{h=1}^{2} q^{(h)}_{multi}\Big) V^{2}_{h=1}(\boldsymbol{Q}^{(h)}|\boldsymbol{Z})$$

(5)

$$\boldsymbol{D}\Big(\sum_h \boldsymbol{Q}^{(h)}\Big|\boldsymbol{Z}\Big) = \sum_h \boldsymbol{D}(\boldsymbol{Q}^{(h)}|\boldsymbol{Z})$$

In particular,

$$\boldsymbol{D}(\boldsymbol{Q}^{(1)} \pm \boldsymbol{Q}^{(2)}|\boldsymbol{Z}) = \boldsymbol{D}(\boldsymbol{Q}^{(1)}|\boldsymbol{Z}) \pm \boldsymbol{D}(\boldsymbol{Q}^{(2)}|\boldsymbol{Z})$$

(6)

$$\sum_{H(1)=1}^{H_{\max}(1)} \sum_{H(2)=1}^{H_{\max}(2)} V_{h=1}^2(\boldsymbol{Q}^{(h,H(h))}|\boldsymbol{Z}) = V_{h=1}^2\Big(\sum_{H(h)=1}^{H_{\max}(h)} \boldsymbol{Q}^{(h,H(h))}\Big|\boldsymbol{Z}\Big)$$

In particular,

$$\sum_{h=1}^{2} V^2(\boldsymbol{Q}^{(h)}|\boldsymbol{Z}) \pm 2V_{h=1}^2(\boldsymbol{Q}^{(h)}|\boldsymbol{Z}) = V^2(\boldsymbol{Q}^{(1)} \pm \boldsymbol{Q}^{(2)}|\boldsymbol{Z})$$

(7)

$$V_{h=1}^2(\boldsymbol{Q}^{(h)}|\boldsymbol{Z}) = E\Big(\prod_{h=1}^{2} \boldsymbol{Q}^{(h)}\Big|\boldsymbol{Z}\Big) - \prod_{h=1}^{2} E(\boldsymbol{Q}^{(h)}|\boldsymbol{Z})$$

(8) When $\boldsymbol{K}_G = \boldsymbol{K}_{G,-i}$,

$$V(\boldsymbol{Q}, \boldsymbol{Q}_{-i}|\boldsymbol{K}_G\boldsymbol{X}^{A}) = V(\boldsymbol{Q}, \boldsymbol{Q}_{-i}|\boldsymbol{K}_G\boldsymbol{X}^{-A}).$$

These hold even if we replace $E(\cdot)$, $\boldsymbol{D}(\cdot)$, *and* $V(\cdot)$ *by* $\mathbb{E}(\cdot)$, $\mathbb{D}(\cdot)$, *and* $\mathbb{V}(\cdot)$, *repsectively.*

PROOF. (1)

$$\begin{aligned} E\{\boldsymbol{D}(\boldsymbol{Q}|\boldsymbol{Z}^{(1)})|\boldsymbol{Z}^{(2)}\} &= E\{\boldsymbol{Q} - \boldsymbol{E}(\boldsymbol{Q}|\boldsymbol{Z}^{(1)})|\boldsymbol{Z}^{(2)}\} \quad (\because \text{Equation 88}) \\ &= E(\boldsymbol{Q}|\boldsymbol{Z}^{(2)}) - E\{\boldsymbol{E}(\boldsymbol{Q}|\boldsymbol{Z}^{(1)})|\boldsymbol{Z}^{(2)}\} \quad (\because \text{Lemma 3 (1)}) \\ &= E(\boldsymbol{Q}|\boldsymbol{Z}^{(2)}) - E(\boldsymbol{Q}|\boldsymbol{Z}^{(1)}) \quad (\because \text{Lemma 3 (3)}, \boldsymbol{Z}^{(2)} \neq \boldsymbol{0}). \end{aligned} \tag{114}$$

Thus,

$$\begin{aligned} &\boldsymbol{D}\{\boldsymbol{D}(\boldsymbol{Q}|\boldsymbol{Z}^{(1)})|\boldsymbol{Z}^{(2)}\} \\ &= \boldsymbol{D}(\boldsymbol{Q}|\boldsymbol{Z}^{(1)}) - \boldsymbol{E}\{\boldsymbol{D}(\boldsymbol{Q}|\boldsymbol{Z}^{(1)})|\boldsymbol{Z}^{(2)}\} \quad (\because \text{Equation 88}) \\ &= \{\boldsymbol{Q} - \boldsymbol{E}(\boldsymbol{Q}|\boldsymbol{Z}^{(1)})\} - \{\boldsymbol{E}(\boldsymbol{Q}|\boldsymbol{Z}^{(2)}) - \boldsymbol{E}(\boldsymbol{Q}|\boldsymbol{Z}^{(1)})\} \quad (\because \text{Equations 88 and 114 }) \\ &= \boldsymbol{Q} - \boldsymbol{E}(\boldsymbol{Q}|\boldsymbol{Z}^{(2)}) \\ &= \boldsymbol{D}(\boldsymbol{Q}|\boldsymbol{Z}^{(2)}) \quad (\because \text{Equation 88}) \end{aligned}$$

(2) When $\boldsymbol{Z} = \boldsymbol{0}$, this is obvious thanks to Equations 5 and 88 (or Equation 91). Otherwise,

$$\begin{aligned} V_{h=1}^2\{\boldsymbol{D}(\boldsymbol{Q}^{(h)}|\boldsymbol{Z})|\boldsymbol{Z}\} &= E\Big[\prod_{h=1}^{2} \boldsymbol{D}\{\boldsymbol{D}(\boldsymbol{Q}^{(h)}|\boldsymbol{Z})|\boldsymbol{Z}\}\Big|\boldsymbol{Z}\Big] \quad (\because \text{Equation 89}) \\ &= E\Big\{\prod_{h=1}^{2} \boldsymbol{D}(\boldsymbol{Q}^{(h)}|\boldsymbol{Z})\Big|\boldsymbol{Z}\Big\} \quad (\because \text{Lemma 15 (1)}, \boldsymbol{Z} \neq \boldsymbol{0}) \\ &= V_{h=1}^2(\boldsymbol{Q}^{(h)}|\boldsymbol{Z}) \quad (\because \text{Equation 89}) \end{aligned}$$

(3)

$$
\begin{aligned}
&V^2_{h=1}(\boldsymbol{Z}^{(h)}\boldsymbol{Q}^{(h)}|\boldsymbol{U}) \\
&= E\Big\{\prod_{h=1}^{2}\boldsymbol{D}(\boldsymbol{Z}^{(h)}\boldsymbol{Q}^{(h)}|\boldsymbol{U})\Big|\boldsymbol{U}\Big\} \quad (\because \text{Equation 89}) \\
&= E\Big[\prod_{h=1}^{2}\{\boldsymbol{Z}^{(h)}\boldsymbol{Q}^{(h)} - \boldsymbol{E}(\boldsymbol{Z}^{(h)}\boldsymbol{Q}^{(h)}|\boldsymbol{U})\}\Big|\boldsymbol{U}\Big] \quad (\because \text{Equation 88}) \\
&= E\Big[\prod_{h=1}^{2}\{\boldsymbol{Z}^{(h)}\boldsymbol{Q}^{(h)} - \boldsymbol{E}(\boldsymbol{Q}^{(h)}|\boldsymbol{U})\}\Big|\boldsymbol{U}\Big] \quad (\because \text{Lemma 3 (4)}) \\
&= E\Big[\prod_{h=1}^{2}\boldsymbol{U}\{\boldsymbol{Z}^{(h)}\boldsymbol{Q}^{(h)} - \boldsymbol{E}(\boldsymbol{Q}^{(h)}|\boldsymbol{U})\}\Big|\boldsymbol{U}\Big] \quad (\because \boldsymbol{U}\in\mathbb{U}, \text{Lemmas 1 (5) and 3 (4)}) \\
&= E\Big[\prod_{h=1}^{2}\boldsymbol{U}\{\boldsymbol{Q}^{(h)} - \boldsymbol{E}(\boldsymbol{Q}^{(h)}|\boldsymbol{U})\}\Big|\boldsymbol{U}\Big] \quad (\because \boldsymbol{Z}^{(h)}\boldsymbol{U} = \boldsymbol{U}\text{for } h\in\{1,2\}) \\
&= E\Big[\prod_{h=1}^{2}\{\boldsymbol{Q}^{(h)} - \boldsymbol{E}(\boldsymbol{Q}^{(h)}|\boldsymbol{U})\}\Big|\boldsymbol{U}\Big] \quad (\because \boldsymbol{U}\in\mathbb{U}, \text{Lemmas 1 (5) and 3 (4)}) \\
&= E\Big\{\prod_{h=1}^{2}\boldsymbol{D}(\boldsymbol{Q}^{(h)}|\boldsymbol{U})\Big|\boldsymbol{U}\Big\} \quad (\because \text{Equation 88}) \\
&= V^2_{h=1}(\boldsymbol{Q}^{(h)}|\boldsymbol{U}) \quad (\because \text{Equation 89})
\end{aligned}
$$

(4) When $\boldsymbol{Z} = \boldsymbol{0}$, thanks to Equation 91, both sides of the equation are equal to each other (zero). Below, suppose $\boldsymbol{Z} \neq \boldsymbol{0}$. Note

$$
\begin{aligned}
&\boldsymbol{D}(\boldsymbol{q}_{\text{const}} + q_{\text{multi}}\boldsymbol{Q}|\boldsymbol{Z}) \\
&= \boldsymbol{q}_{\text{const}} + q_{\text{multi}}\boldsymbol{Q} - \boldsymbol{E}(\boldsymbol{q}_{\text{const}} + q_{\text{multi}}\boldsymbol{Q}|\boldsymbol{Z}) \quad (\because \text{Equation 88}) \\
&= \{\boldsymbol{q}_{\text{const}} + q_{\text{multi}}\boldsymbol{Q}\} - \{\boldsymbol{q}_{\text{const}} + q_{\text{multi}}\boldsymbol{E}(\boldsymbol{Q}|\boldsymbol{Z})\} \quad (\because \boldsymbol{Z}\neq\boldsymbol{0}, \text{Lemma 3 (1), (2), and (3)}) \\
&= q_{\text{multi}}\{\boldsymbol{Q} - \boldsymbol{E}(\boldsymbol{Q}|\boldsymbol{Z})\} \\
&= q_{\text{multi}}\boldsymbol{D}(\boldsymbol{Q}|\boldsymbol{Z}) \quad (\because \text{Equation 88})
\end{aligned}
\tag{115}
$$

Thus,

$$
\begin{aligned}
V^2_{h=1}(\boldsymbol{q}^{(h)}_{\text{const}} + q^{(h)}_{\text{multi}}\boldsymbol{Q}^{(h)}|\boldsymbol{Z}) &= E\Big\{\prod_{h=1}^{2}\boldsymbol{D}(\boldsymbol{q}^{(h)}_{\text{const}} + q^{(h)}_{\text{multi}}\boldsymbol{Q}^{(h)}|\boldsymbol{Z})\Big|\boldsymbol{Z}\Big\} \quad (\because \text{Equation 89}) \\
&= E\Big\{\prod_{h=1}^{2}q^{(h)}_{\text{multi}}\boldsymbol{D}(\boldsymbol{Q}^{(h)}|\boldsymbol{Z})\Big|\boldsymbol{Z}\Big\} \quad (\because \text{Equation 115}) \\
&= \Big(\prod_{h=1}^{2}q^{(h)}_{\text{multi}}\Big)E\Big\{\prod_{h=1}^{2}\boldsymbol{D}(\boldsymbol{Q}^{(h)}|\boldsymbol{Z})\Big|\boldsymbol{Z}\Big\} \quad (\because \text{Lemma 3 (2)}) \\
&= \Big(\prod_{h=1}^{2}q^{(h)}_{\text{multi}}\Big)V^2_{h=1}(\boldsymbol{Q}^{(h)}|\boldsymbol{Z}) \quad (\because \text{Equation 90})
\end{aligned}
$$

(5)

$$\begin{aligned} \boldsymbol{D}\Big(\sum_h \boldsymbol{Q}^{(h)}\Big|\boldsymbol{Z}\Big) &= \sum_h \boldsymbol{Q}^{(h)} - \boldsymbol{E}\Big(\sum_h \boldsymbol{Q}^{(h)}\Big|\boldsymbol{Z}\Big) \quad (\because \text{Equation 88}) \\ &= \sum_h \boldsymbol{Q}^{(h)} - \sum_h \boldsymbol{E}(\boldsymbol{Q}^{(h)}|\boldsymbol{Z}) \quad (\because \text{Lemma 3 (1)}) \\ &= \sum_h \boldsymbol{D}(\boldsymbol{Q}^{(h)}|\boldsymbol{Z}) \quad (\because \text{Equation 88}) \end{aligned}$$

In particular,

$$\begin{aligned} \boldsymbol{D}(\boldsymbol{Q}^{(1)} \pm \boldsymbol{Q}^{(2)}|\boldsymbol{Z}) &= (\boldsymbol{Q}^{(1)} \pm \boldsymbol{Q}^{(2)}) - \boldsymbol{E}(\boldsymbol{Q}^{(1)} \pm \boldsymbol{Q}^{(2)}|\boldsymbol{Z}) \quad (\because \text{Equation 88}) \\ &= (\boldsymbol{Q}^{(1)} \pm \boldsymbol{Q}^{(2)}) - \{\boldsymbol{E}(\boldsymbol{Q}^{(1)}|\boldsymbol{Z}) \pm \boldsymbol{E}(\boldsymbol{Q}^{(2)}|\boldsymbol{Z})\} \quad (\because \text{Lemma 3 (1)}) \\ &= \{\boldsymbol{Q}^{(1)} - \boldsymbol{E}(\boldsymbol{Q}^{(1)}|\boldsymbol{Z})\} \pm \{\boldsymbol{Q}^{(2)} - \boldsymbol{E}(\boldsymbol{Q}^{(2)}|\boldsymbol{Z})\} \\ &= \boldsymbol{D}(\boldsymbol{Q}^{(1)}|\boldsymbol{Z}) \pm \boldsymbol{D}(\boldsymbol{Q}^{(2)}|\boldsymbol{Z}) \quad (\because \text{Equation 88}) \end{aligned}$$

(6)

$$\begin{aligned} V^2_{h=1}\Big(\sum_{H(h)=1}^{H_{\max}(h)} \boldsymbol{Q}^{(h,H(h))}\Big|\boldsymbol{Z}\Big) &= E\Big\{\prod_{h=1}^{2} \boldsymbol{D}\Big(\sum_{H(h)=1}^{H_{\max}(h)} \boldsymbol{Q}^{(h,H(h))}\Big|\boldsymbol{Z}\Big)\Big|\boldsymbol{Z}\Big\} \quad (\because \text{Equation 90}) \\ &= E\Big\{\prod_{h=1}^{2} \sum_{H(h)=1}^{H_{\max}(h)} \boldsymbol{D}(\boldsymbol{Q}^{(h,H(h))}|\boldsymbol{Z})\Big|\boldsymbol{Z}\Big\} \quad (\because \text{Lemma 15 (5)}) \\ &= \sum_{H(1)=1}^{H_{\max}(1)} \sum_{H(2)=1}^{H_{\max}(2)} E\Big\{\prod_{h=1}^{2} \boldsymbol{D}(\boldsymbol{Q}^{(h,H(h))}|\boldsymbol{Z})\Big|\boldsymbol{Z}\Big\} \quad (\because \text{Lemmas 3 (1) and 16 (1)}) \\ &= \sum_{H(1)=1}^{H_{\max}(1)} \sum_{H(2)=1}^{H_{\max}(2)} V^2_{h=1}(\boldsymbol{Q}^{(h,H(h))}|\boldsymbol{Z}) \end{aligned}$$

In particular,

$$\begin{aligned} &\sum_{h=1}^{2} V^2(\boldsymbol{Q}^{(h)}|\boldsymbol{Z}) \pm 2V^2_{h=1}(\boldsymbol{Q}^{(h)}|\boldsymbol{Z}) \\ &= \sum_{h=1}^{2} E[\{\boldsymbol{D}(\boldsymbol{Q}^{(h)}|\boldsymbol{Z})\}^2|\boldsymbol{Z}] \pm 2E\Big\{\prod_{h=1}^{2} \boldsymbol{D}(\boldsymbol{Q}^{(h)}|\boldsymbol{Z})\Big|\boldsymbol{Z}\Big\} \quad (\because \text{Equation 90}) \\ &= E\Big[\sum_{h=1}^{2}\{\boldsymbol{D}(\boldsymbol{Q}^{(h)}|\boldsymbol{Z})\}^2 \pm 2\prod_{h=1}^{2} \boldsymbol{D}(\boldsymbol{Q}^{(h)}|\boldsymbol{Z})\Big|\boldsymbol{Z}\Big] \quad (\because \text{Lemma 3 (1)}) \\ &= E[\{\boldsymbol{D}(\boldsymbol{Q}^{(1)}|\boldsymbol{Z}) \pm \boldsymbol{D}(\boldsymbol{Q}^{(2)}|\boldsymbol{Z})\}^2|\boldsymbol{Z}] \\ &= E[\{\boldsymbol{D}(\boldsymbol{Q}^{(1)} \pm \boldsymbol{Q}^{(2)}|\boldsymbol{Z})\}^2|\boldsymbol{Z}] \quad (\because \text{Lemma 15 (5)}) \\ &= V^2(\boldsymbol{Q}^{(1)} \pm \boldsymbol{Q}^{(2)}|\boldsymbol{Z}) \quad (\because \text{Lemma 90}) \end{aligned}$$

(7)

$$
\begin{aligned}
& V_{h=1}^{2}(\boldsymbol{Q}^{(h)}|\boldsymbol{Z}) \\
&= E[\{\boldsymbol{Q}^{(1)} - \boldsymbol{E}(\boldsymbol{Q}^{(1)}|\boldsymbol{Z})\}\{\boldsymbol{Q}^{(2)} - \boldsymbol{E}(\boldsymbol{Q}^{(2)}|\boldsymbol{Z})\}|\boldsymbol{Z}] \quad (\because \text{Equation 90}) \\
&= E\{\boldsymbol{Q}^{(1)}\boldsymbol{Q}^{(2)} + \boldsymbol{E}(\boldsymbol{Q}^{(1)}|\boldsymbol{Z})\boldsymbol{E}(\boldsymbol{Q}^{(2)}|\boldsymbol{Z}) - \boldsymbol{Q}^{(1)}\boldsymbol{E}(\boldsymbol{Q}^{(2)}|\boldsymbol{Z}) - \boldsymbol{E}(\boldsymbol{Q}^{(1)}|\boldsymbol{Z})\boldsymbol{Q}^{(2)}|\boldsymbol{Z}\} \\
&= E(\boldsymbol{Q}^{(1)}\boldsymbol{Q}^{(2)}|\boldsymbol{Z}) + E\{\boldsymbol{E}(\boldsymbol{Q}^{(1)}|\boldsymbol{Z})\boldsymbol{E}(\boldsymbol{Q}^{(2)}|\boldsymbol{Z})|\boldsymbol{Z}\} - E\{\boldsymbol{Q}^{(1)}\boldsymbol{E}(\boldsymbol{Q}^{(2)}|\boldsymbol{Z})|\boldsymbol{Z}\} \\
&\quad - E\{\boldsymbol{E}(\boldsymbol{Q}^{(1)}|\boldsymbol{Z})\boldsymbol{E}\{\boldsymbol{Q}^{(2)}|\boldsymbol{Z})|\boldsymbol{Z}\} \quad (\because \text{Equation 3 (1)}) \\
&= E(\boldsymbol{Q}^{(1)}\boldsymbol{Q}^{(2)}|\boldsymbol{Z}) + E(\boldsymbol{Q}^{(1)}|\boldsymbol{Z})E(\boldsymbol{Q}^{(2)}|\boldsymbol{Z}) - E(\boldsymbol{Q}^{(1)}|\boldsymbol{Z})E(\boldsymbol{Q}^{(2)}|\boldsymbol{Z}) \\
&\quad - E(\boldsymbol{Q}^{(1)}|\boldsymbol{Z})E(\boldsymbol{Q}^{(2)}|\boldsymbol{Z}) \quad (\because \text{Equation 3 (2)}) \\
&= E\Big(\prod_{h=1}^{2}\boldsymbol{Q}^{(h)}\Big|\boldsymbol{Z}\Big) - \prod_{h=1}^{2}E(\boldsymbol{Q}^{(h)}|\boldsymbol{Z})
\end{aligned}
$$

(8) Note

$$
\begin{aligned}
V(\boldsymbol{Q}^{(1)}, \boldsymbol{Q}^{(2)}|\boldsymbol{Z}) &= E\{D(\boldsymbol{Q}^{(1)}|\boldsymbol{Z})D(\boldsymbol{Q}^{(2)}|\boldsymbol{Z})|\boldsymbol{Z}\} \quad (\because \text{Equation 90}) \\
&= E\{D(\boldsymbol{Q}^{(2)}|\boldsymbol{Z})D(\boldsymbol{Q}^{(1)}|\boldsymbol{Z})|\boldsymbol{Z}\} \\
&= V(\boldsymbol{Q}^{(2)}, \boldsymbol{Q}^{(1)}|\boldsymbol{Z}) \quad (\because \text{Equation 90})
\end{aligned} \tag{116}
$$

Thus,

$$
\begin{aligned}
& V(\boldsymbol{Q}, \pm\boldsymbol{Q}_{-i}|\boldsymbol{K}_G\boldsymbol{X}^A) \\
&= \pm V(\boldsymbol{Q}_{-i}, \boldsymbol{Q}|\boldsymbol{K}_{G,-i}\boldsymbol{X}_{-i}^{-A}) \quad (\because \boldsymbol{K}_G = \boldsymbol{K}_{G,-i}, \text{Equation 116, Lemmas 2 (1) and 15 (4)}) \\
&= V(\boldsymbol{Q}, \pm\boldsymbol{Q}_{-i}|\boldsymbol{K}_G\boldsymbol{X}^{-A}). \quad (\because \text{exchanging } i \text{ and } -i, \text{Lemma 15 (4)})
\end{aligned} \tag{117}
$$

□

LEMMA 16 (SQUARE OF SUM). *Let l be index of $\boldsymbol{Q}$ where, in the case of $l = i$ and $\mathbb{L} = \mathbb{I} = \{1, 2\}$, given j,*

$$Q_l = Q_{ij}, \quad \sum_{l\in\mathbb{L}} = \sum_{i\in\mathbb{I}},$$

in the case of $l = j$ and $\mathbb{L} = \mathbb{J} \subseteq \{1, 2, \ldots, n\}$,

$$Q_l = Q_{\cdot j}, \quad \sum_{l\in\mathbb{L}} = \sum_{j\in\mathbb{J}},$$

in the case of $l = ij$ and $\mathbb{L} = \mathbb{I} \times \mathbb{J}$,

$$Q_l = Q_{ij}, \quad \sum_{l\in\mathbb{L}} = \sum_{j\in\mathbb{J}}\sum_{i\in\mathbb{I}},$$

and, in the case of $l \in \mathbb{L} = \{0, 1\}$,

$$Q_l = f_Q^{(l)}(\boldsymbol{Q}), \quad \sum_{l\in\mathbb{L}} = \sum_{l\in\{0,1\}}.$$

(1)

$$\sum_{l(1)\in\mathbb{L}(1)}\sum_{l(2)\in\mathbb{L}(2)}\prod_{h=1}^{2} Q_{l(h)} = \prod_{h=1}^{2}\sum_{l(h)\in\mathbb{L}(h)} Q_{l(h)}$$

(2) For $h \in \{1,2\}$, suppose that, for some $f_X(\cdot), \boldsymbol{q}^{(h)}, \overline{f^2}_{EQ}$, and $\overline{f^2}_{DF}$, all of which are constant irrespective of $\boldsymbol{X}^{(h)} \in \{\boldsymbol{X}^T, \boldsymbol{X}^C\}$,

$$Q^{(h)}_{l(h)} = f_X(X^{(h)}_{l(h)}) q^{(h)}_{l(h)} \tag{118}$$

$$\mathbb{E}\Big\{ \prod_{h=1}^{2} f_X(X^{(h)}_{l(h)}) \Big\} = \begin{cases} \overline{f^2}_{EQ} & \text{if} \quad l(1) = l(2) \\ \overline{f^2}_{DF} & \text{if} \quad l(1) \neq l(2). \end{cases} \tag{119}$$

and for any $l(1) \in \mathbb{L}(1), l(2) \in \mathbb{L}(2)$,

$$\mathbb{X}_{def}\Big(\sum_{l'(1)\in\mathbb{L}(1)} \sum_{l'(2)\in\mathbb{L}(2)} \prod_{h=1}^{2} Q^{(h)}_{l'(h)} \Big) = \mathbb{X}_{def}\Big(\prod_{h=1}^{2} Q^{(h)}_{l(h)} \Big). \tag{120}$$

It follows

$$\mathbb{E}\Big(\sum_{l(1)\in\mathbb{L}(1)} \sum_{l(2)\in\mathbb{L}(2)} \prod_{h=1}^{2} Q_{l(h)} \Big) = (\overline{f^2}_{EQ} - \overline{f^2}_{DF}) \sum_{l\in\bigcap_{h=1}^{2}\mathbb{L}(h)} \prod_{h=1}^{2} q^{(h)}_{l} + \overline{f^2}_{DF} \sum_{l(1)\in\mathbb{L}(1)} \sum_{l(2)\in\mathbb{L}(2)} \prod_{h=1}^{2} q^{(h)}_{l(h)}$$

Proof. (1)

$$\begin{aligned} \sum_{l(1)\in\mathbb{L}(1)} \sum_{l(2)\in\mathbb{L}(2)} \prod_{h=1}^{2} Q_{l(h)} &= \sum_{l(1)\in\mathbb{L}(1)} \sum_{l(2)\in\mathbb{L}(2)} Q_{l(1)} Q_{l(2)} \\ &= \sum_{l(1)\in\mathbb{L}(1)} Q_{l(1)} \sum_{l(2)\in\mathbb{L}(2)} Q_{l(2)} \\ &= \Big(\sum_{l(1)\in\mathbb{L}(1)} Q_{l(1)} \Big) \Big(\sum_{l(2)\in\mathbb{L}(2)} Q_{l(2)} \Big) \\ &= \prod_{h=1}^{2} \sum_{l(h)\in\mathbb{L}(h)} Q_{l(h)} \end{aligned}$$

(2)

$$\begin{aligned} \mathbb{E}\Big(\prod_{h=1}^{2} Q^{(h)}_{l(h)} \Big) &= \mathbb{E}\Big\{ \prod_{h=1}^{2} f_X(X^{(h)}_{l(h)}) q^{(h)}_{l(h)} \Big\} \quad (\because \text{Equations 118}) \\ &= \mathbb{E}\Big\{ \prod_{h=1}^{2} f_X(X^{(h)}_{l(h)}) \Big\} \prod_{h=1}^{2} q^{(h)}_{l(h)} \quad (\because \text{Lemma 10 (2)}) \end{aligned} \tag{121}$$

Therefore,

$$
\begin{aligned}
&\mathbb{E}\Big(\sum_{l(1)\in\mathbb{L}(1)}\sum_{l(2)\in\mathbb{L}(2)}\prod_{h=1}^{2}Q_{l(h)}^{(h)}\Big)\\
&=\sum_{l(1)\in\mathbb{L}(1)}\sum_{l(2)\in\mathbb{L}(2)}\mathbb{E}\Big(\prod_{h=1}^{2}Q_{l(h)}^{(h)}\Big)\quad(\because \text{Lemma 10 (1), Equation 120})\\
&=\sum_{l\in\bigcap_{h=1}^{2}\mathbb{L}(h)}\mathbb{E}\Big(\prod_{h=1}^{2}Q_{l}^{(h)}\Big)+\sum_{l(1)\in\mathbb{L}(1)}\sum_{l(2)\in\mathbb{L}(2)\setminus\{l(1)\}}\mathbb{E}\Big(\prod_{h=1}^{2}Q_{l(h)}^{(h)}\Big)\\
&=\sum_{l\in\bigcap_{h=1}^{2}\mathbb{L}(h)}\overline{f^2}_{EQ}\prod_{h=1}^{2}q_{l}^{(h)}+\sum_{l(1)\in\mathbb{L}(1)}\sum_{l(2)\in\mathbb{L}(2)\setminus\{l(1)\}}\overline{f^2}_{DF}\prod_{h=1}^{2}q_{l(h)}^{(h)}\quad(\because \text{Equations 119 and 121})\\
&=\overline{f^2}_{EQ}\sum_{l\in\bigcap_{h=1}^{2}\mathbb{L}(h)}\prod_{h=1}^{2}q_{l}^{(h)}+\overline{f^2}_{DF}\Big(\sum_{l(1)\in\mathbb{L}(1)}\sum_{l(2)\in\mathbb{L}(2)}\prod_{h=1}^{2}q_{l(h)}^{(h)}-\sum_{l\in\bigcap_{h=1}^{2}\mathbb{L}(h)}\prod_{h=1}^{2}q_{l}^{(h)}\Big)\\
&=(\overline{f^2}_{EQ}-\overline{f^2}_{DF})\sum_{l\in\bigcap_{h=1}^{2}\mathbb{L}(h)}\prod_{h=1}^{2}q_{l}^{(h)}+\overline{f^2}_{DF}\sum_{l(1)\in\mathbb{L}(1)}\sum_{l(2)\in\mathbb{L}(2)}\prod_{h=1}^{2}q_{l(h)}^{(h)}
\end{aligned}
$$

□

LEMMA 17 (SQUARED MEAN). *Suppose that* $\boldsymbol{K}_G\in\mathbb{U}$ *satisfies Condition 1. For* $h\in\{1,2\}$, *let* $\boldsymbol{\omega}^{(h)}\in\mathbb{W},\boldsymbol{k}_G,\boldsymbol{k}_G^{(h)}\in\{\boldsymbol{k}_G^T,\boldsymbol{k}_G^C\},\boldsymbol{X}^{(h)}\in\{\boldsymbol{X}^T,\boldsymbol{X}^C\}$.

(1)

$$
\begin{aligned}
E\Big(\prod_{h=1}^{2}\boldsymbol{\omega}^{(h)}\Big|\boldsymbol{k}_G\boldsymbol{X}^A\Big)&=E\Big(\prod_{h=1}^{2}\boldsymbol{\omega}^{(h)}\Big|\boldsymbol{k}_G\Big)\\
&=V_{h=1}^{2}(\boldsymbol{\omega}^{(h)}|\boldsymbol{k}_G)
\end{aligned}
$$

(2) When $\boldsymbol{k}_G^{(h)}\neq\boldsymbol{0}$,

$$
\mathbb{E}\Big\{\prod_{h=1}^{2}E(\boldsymbol{\omega}^{(h)}|\boldsymbol{k}_G^{(h)}\boldsymbol{X}^{(h)})\Big\}=\{2\cdot I(\boldsymbol{X}^{(1)}=\boldsymbol{X}^{(2)})-1\}\frac{n_G^{(1)\cdot(2)}}{n_G^{(1)}n_G^{(2)}}V_{h=1}^{2}\Big(\boldsymbol{\omega}^{(h)}\Big|\prod_{h=1}^{2}\boldsymbol{k}_G^{(h)}\Big)
$$

In particular, when $\boldsymbol{k}_G\neq\boldsymbol{0}$,

$$
\mathbb{E}[\{E(\boldsymbol{\omega}|\boldsymbol{k}_G\boldsymbol{X}^A)\}^2]=\frac{1}{n_G}V^2(\boldsymbol{\omega}|\boldsymbol{k}_G)
$$

PROOF. (1) When $\boldsymbol{k}_G=\boldsymbol{0}$,

$$
\begin{aligned}
E\Big(\prod_{h=1}^{2}\boldsymbol{\omega}^{(h)}\Big|\boldsymbol{k}_G\boldsymbol{X}^A\Big)&=E\Big(\prod_{h=1}^{2}\boldsymbol{\omega}^{(h)}\Big|\boldsymbol{k}_G\Big)\quad(\because \boldsymbol{k}_G\boldsymbol{X}^A=\boldsymbol{k}_G=\boldsymbol{0})\\
&=0\quad(\because \text{Equation 5})\\
V_{h=1}^{2}(\boldsymbol{\omega}^{(h)}|\boldsymbol{k}_G)&=0\quad(\because \text{Equation 91})
\end{aligned}
$$

When $\boldsymbol{k}_G \neq \mathbf{0}$,

$$
\begin{aligned}
& E\Big(\prod_{h=1}^{2} \boldsymbol{\omega}^{(h)} \Big| \boldsymbol{k}_G \boldsymbol{X}^A\Big) \\
&= \frac{1}{n_G} \sum_j \sum_i k_{G,\cdot j} X_{ij}^A \prod_{h=1}^{2} \omega_{ij}^{(h)} \quad (\because \text{Equations 4, 9, and 36, Condition 1 }) \\
&= \frac{1}{n_G} \sum_j k_{G,\cdot j} \Big(X_{ij}^A \prod_{h=1}^{2} \omega_{ij}^{(h)} + X_{-ij}^A \prod_{h=1}^{2} \omega_{-ij}^{(h)} \Big) \\
&= \frac{1}{n_G} \sum_j k_{G,\cdot j} \Big\{ X_{ij}^A \prod_{h=1}^{2} \omega_{ij}^{(h)} + (1 - X_{ij}^A) \prod_{h=1}^{2} -\omega_{ij}^{(h)} \Big\} \quad (\because \text{Equations 2 and 21}) \\
&= \frac{1}{n_G} \sum_j k_{G,\cdot j} \prod_{h=1}^{2} \omega_{ij}^{(h)} \\
&= \frac{1}{n_G} \sum_j k_{G,\cdot j} \frac{1}{2} \Big(\prod_{h=1}^{2} \omega_{ij}^{(h)} + \prod_{h=1}^{2} -\omega_{ij}^{(h)} \Big) \\
&= \frac{1}{2n_G} \sum_j k_{G,\cdot j} \Big(\prod_{h=1}^{2} \omega_{ij}^{(h)} + \prod_{h=1}^{2} \omega_{-ij}^{(h)} \Big) \quad (\because \text{Equations 2 and 21}) \\
&= \frac{\sum_j \sum_i k_{G,\cdot j} \prod_{h=1}^{2} \omega_{ij}^{(h)}}{\sum_j \sum_i k_{G,\cdot j}} \\
&= E\Big(\prod_{h=1}^{2} \boldsymbol{\omega}^{(h)} \Big| \boldsymbol{k}_G\Big) \quad (\because \text{Equation 4})
\end{aligned}
$$

It also holds that

$$
\begin{aligned}
V_{h=1}^{2}(\boldsymbol{\omega}^{(h)} | \boldsymbol{k}_G) &= E\Big(\prod_{h=1}^{2} \boldsymbol{\omega}^{(h)} \Big| \boldsymbol{k}_G\Big) - \prod_{h=1}^{2} E(\boldsymbol{\omega}^{(h)} | \boldsymbol{k}_G) \quad (\because \text{Lemma 15 (7)}) \\
&= E\Big(\prod_{h=1}^{2} \boldsymbol{\omega}^{(h)} \Big| \boldsymbol{k}_G\Big) \quad (\because \text{Lemma 12 (1)})
\end{aligned}
$$

(2) Denote

$$
\overline{X^2}_{EQ}\{i(1), i(2)\} \equiv \mathbb{E}\Big(\prod_{h=1}^{2} X_{i(h)j}^{(h)}\Big) \tag{122}
$$

and, for any $j(1) \neq j(2)$,

$$
\begin{aligned}
\overline{X^2}_{DF} &\equiv \mathbb{E}\Big(\prod_{h=1}^{2} X_{i(h)j(h)}^{(h)}\Big) \\
&= \mathbb{E}\{X_{i(1)j(1)}^{(1)} \mathbb{E}(X_{i(2)j(2)}^{(2)} | X_{i(1)j(1)}^{(1)})\} \\
&= \frac{1}{2} \mathbb{E}(X_{i(1)j(1)}^{(1)}) \quad (\because j(1) \neq j(2), \text{ Lemmas 10 (2) and 11 (2)}) \\
&= \frac{1}{4} \quad (\because \text{Lemma 11 (1)})
\end{aligned} \tag{123}
$$

It holds that

$$
\begin{aligned}
&\mathbb{E}\Big\{\prod_{h=1}^{2} E(\boldsymbol{\omega}^{(h)}|\boldsymbol{k}_G^{(h)}\boldsymbol{X}^{(h)})\Big\}\\
&=\mathbb{E}\Big(\prod_{h=1}^{2}\frac{\sum_{j(h)}\sum_{i(h)} k_{G,i(h)j(h)}^{(h)} X_{i(h)j(h)}^{(h)}\omega_{i(h)j(h)}^{(h)}}{\sum_{j(h)}\sum_{i(h)} k_{G,i(h)j(h)}^{(h)} X_{i(h)j(h)}^{(h)}}\Big) \quad (\because \text{Equations 4 and 36, } \boldsymbol{k}_G^{(h)}\neq\boldsymbol{0})\\
&=\frac{1}{n_G^{(1)}n_G^{(2)}}\mathbb{E}\Big(\prod_{h=1}^{2}\sum_{j(h)}\sum_{i(h)} k_{G,i(h)j(h)}^{(h)} X_{i(h)j(h)}^{(h)}\omega_{i(h)j(h)}^{(h)}\Big) \quad (\because n_G^{(1)}n_G^{(2)} \text{ is constant irrespective of } \boldsymbol{X}^A)\\
&=\frac{1}{n_G^{(1)}n_G^{(2)}}\mathbb{E}\Big(\sum_{i(1)}\sum_{i(2)}\sum_{j(1)}\sum_{j(2)}\prod_{h=1}^{2} k_{G,\cdot j(h)}^{(h)} X_{i(h)j(h)}^{(h)}\omega_{i(h)j(h)}^{(h)}\Big) \quad (\because \text{Lemma 16 (1)},\\
&\quad \text{where } l(h)=i(h)j(h), \mathbb{L}(h)=\mathbb{I}\times\mathbb{J}, \mathbb{J}=\{1,2,\ldots,n\}, Q_{l(h)}^{(h)}=k_{G,\cdot j(h)}^{(h)}X_{i(h)j(h)}^{(h)}\omega_{i(h)j(h)}^{(h)})\\
&=\frac{1}{n_G^{(1)}n_G^{(2)}}\sum_{i(1)}\sum_{i(2)}\Big([\overline{X^2}_{EQ}\{i(1),i(2)\}-\overline{X^2}_{DF}]\sum_{j}\prod_{h=1}^{2} k_{G,\cdot j(h)}^{(h)}\omega_{i(h)j(h)}^{(h)}\\
&\quad +\overline{X^2}_{DF}\sum_{j(1)}\sum_{j(2)}\prod_{h=1}^{2} k_{G,\cdot j(h)}^{(h)}\omega_{i(h)j(h)}^{(h)}\Big)\\
&\quad (\because \text{Lemma 16 (2), where } l(h)=j(h), \mathbb{L}(h)=\{1,2,\ldots,n\}, f_X(X_{l(h)}^{(h)})=X_{i(h)j(h)}^{(h)},\\
&\quad q_{l(h)}^{(h)}=k_{G,\cdot j(h)}^{(h)}\omega_{i(h)j(h)}^{(h)}, \overline{f^2}_{EQ}=\overline{X^2}_{EQ}\{i(1),i(2)\}, \overline{f^2}_{DF}=\overline{X^2}_{DF})\\
&=\frac{1}{n_G^{(1)}n_G^{(2)}}\Big\{\sum_{j}\Big(\prod_{h=1}^{2} k_{G,\cdot j}^{(h)}\Big)\sum_{i(1)}\sum_{i(2)}\Big[\overline{X^2}_{EQ}\{i(1),i(2)\}-\overline{X^2}_{DF}\Big]\prod_{h=1}^{2}\omega_{i(h)j}^{(h)}\\
&\quad +\overline{X^2}_{DF}\sum_{j(1)}\sum_{j(2)}\sum_{i(1)}\sum_{i(2)}\prod_{h=1}^{2} k_{G,\cdot j(h)}^{(h)}\omega_{i(h)j(h)}^{(h)}\Big\} \quad (\because \text{Equation 123})\\
&=\frac{1}{n_G^{(1)}n_G^{(2)}}\Big\{\sum_{j}\Big(\prod_{h=1}^{2} k_{G,\cdot j}^{(h)}\Big)\sum_{i}\Big[\Big\{\overline{X^2}_{EQ}(i,i)-\overline{X^2}_{DF}\Big\}\prod_{h=1}^{2}\omega_{ij}^{(h)}\\
&\quad +\Big\{\overline{X^2}_{EQ}(i,-i)-\overline{X^2}_{DF}\Big\}\omega_{ij}^{(1)}\omega_{-ij}^{(2)}\Big]+\overline{X^2}_{DF}\sum_{j(1)}\sum_{j(2)}\prod_{h=1}^{2}\sum_{i(h)} k_{G,\cdot j(h)}^{(h)}\omega_{i(h)j(h)}^{(h)}\Big\}\\
&\quad (\because \text{Lemma 16 (1), where } l(h)=i(h), \mathbb{L}(h)=\{1,2\}, Q_{l(h)}^{(h)}=k_{G,\cdot j(h)}^{(h)}\omega_{i(h)j(h)}^{(h)})\\
&=\frac{1}{n_G^{(1)}n_G^{(2)}}\sum_{j}\Big(\prod_{h=1}^{2} k_{G,\cdot j}^{(h)}\Big)\sum_{i}\Big[\{\overline{X^2}_{EQ}(i,i)-\overline{X^2}_{DF}\}\prod_{h=1}^{2}\omega_{ij}^{(h)}-\{\overline{X^2}_{EQ}(i,-i)-\overline{X^2}_{DF}\}\prod_{h=1}^{2}\omega_{ij}^{(h)}\Big]\\
&\quad (\because \text{Lemma 5 (1), Equation 21})\\
&=\frac{1}{n_G^{(1)}n_G^{(2)}}\sum_{j}\Big(\prod_{h=1}^{2} k_{G,\cdot j}^{(h)}\Big)\sum_{i}\{\overline{X^2}_{EQ}(i,i)-\overline{X^2}_{EQ}(i,-i)\}\prod_{h=1}^{2}\omega_{ij}^{(h)}
\end{aligned}
\tag{124}
$$

When $\boldsymbol{X}^{(1)} = \boldsymbol{X}^{(2)} \equiv \boldsymbol{X}^A$, it follows that

$$\begin{aligned}\overline{X^2}_{EQ}\{i(1), i(2)\} &= \mathbb{E}\Big(\prod_{h=1}^{2} X^A_{i(h)j}\Big) \quad (\because \text{Equation 122}) \\ &= \begin{cases} \dfrac{1}{2} \quad \text{if} \quad i(1) = i(2) \quad (\because \text{Lemmas 1 (5) and 11 (1)}) \\ 0 \quad \text{if} \quad i(1) \neq i(2) \quad (\because \text{Lemma 2 (3)}) \end{cases}\end{aligned} \tag{125}$$

and, therefore,

$$\begin{aligned}\mathbb{E}\Big\{\prod_{h=1}^{2} E(\boldsymbol{\omega}^{(h)}|\boldsymbol{k}_G^{(h)}\boldsymbol{X}^{(h)})\Big\} &= \frac{1}{n_G^{(1)}n_G^{(2)}}\sum_j \Big(\prod_{h=1}^{2} k^{(h)}_{G,\cdot j}\Big)\sum_i \frac{1}{2}\prod_{h=1}^{2}\omega^{(h)}_{ij} \quad (\because \text{Equations 124 and 125}) \\ &= \frac{n_G^{(1)\cdot(2)}}{n_G^{(1)}n_G^{(2)}} E\Big(\prod_{h=1}^{2}\boldsymbol{\omega}^{(h)}\Big|\prod_{h=1}^{2}\boldsymbol{k}_G^{(h)}\Big) \\ &= \frac{n_G^{(1)\cdot(2)}}{n_G^{(1)}n_G^{(2)}} V^2_{h=1}\Big(\boldsymbol{\omega}^{(h)}\Big|\prod_{h=1}^{2}\boldsymbol{k}_G^{(h)}\Big) \quad (\because \text{Lemma 17 (1)})\end{aligned} \tag{126}$$

In addition, when $\boldsymbol{\omega}^{(1)} = \boldsymbol{\omega}^{(2)} \equiv \boldsymbol{\omega}, \boldsymbol{k}_G^{(1)} = \boldsymbol{k}_G^{(2)} \equiv \boldsymbol{k}_G$, it follows that

$$\begin{aligned}\mathbb{E}\Big\{\prod_{h=1}^{2} E(\boldsymbol{\omega}^{(h)}|\boldsymbol{k}_G^{(h)}\boldsymbol{X}^{(h)})\Big\} &= \frac{1}{n_G} V^2_{h=1}\Big(\boldsymbol{\omega}\Big|\prod_{h=1}^{2}\boldsymbol{k}_G\Big) \quad (\because \text{Equation 126, Lemma 14 (3)}) \\ &= \frac{1}{n_G} V^2(\boldsymbol{\omega}|\boldsymbol{k}_G) \quad (\because \boldsymbol{k}_G \in \mathbb{U}, \text{Lemma 1 (5), Equation 90})\end{aligned}$$

Similarly, when $\boldsymbol{X}^{(1)} = \boldsymbol{X}^A, \boldsymbol{X}^{(2)} = \boldsymbol{X}^{-A}$, it follows that

$$\begin{aligned}\overline{X^2}_{EQ}\{i(1), i(2)\} &= \mathbb{E}(X^A_{i(1)j}X^{-A}_{i(2)j}) \quad (\because \text{Equation 122}) \\ &= \begin{cases} 0 \quad \text{if } i(1) = i(2) \quad (\because \text{Lemma 2 (3)}) \\ \dfrac{1}{2} \quad \text{if } i(1) \neq i(2) \quad (\because \text{Lemmas 2 (4) and 11 (1)}) \end{cases}\end{aligned} \tag{127}$$

and, therefore,

$$\begin{aligned}&\mathbb{E}\Big\{\prod_{h=1}^{2} E(\boldsymbol{\omega}^{(h)}|\boldsymbol{k}_G^{(h)}\boldsymbol{X}^{(h)})\Big\} \\ &= \frac{1}{n_G^{(1)}n_G^{(2)}}\sum_j \Big(\prod_{h=1}^{2} k^{(h)}_{G,\cdot j}\Big)\sum_i \Big(-\frac{1}{2}\Big)\prod_{h=1}^{2}\omega^{(h)}_{ij} \quad (\because \text{Equations 124 and 127}) \\ &= -\frac{n_G^{(1)\cdot(2)}}{n_G^{(1)}n_G^{(2)}} V^2_{h=1}\Big(\boldsymbol{\omega}^{(h)}\Big|\prod_{h=1}^{2}\boldsymbol{k}_G^{(h)}\Big)\end{aligned}$$

□

Proposition 2 (Variance of ATE Estimators: FS). *(1) Under Assumption 1,*

$$\mathbb{V}^2(\hat{\tau}_F) = \frac{1}{n_F} V^2(\boldsymbol{\omega}^T + \boldsymbol{\omega}^C|\boldsymbol{k}_F).$$

(2) Under Assumption 3 and $n_P \geq 1$,

$$\mathbb{V}^2(\hat{\tau}_P) = \frac{1}{n_P} V^2(\boldsymbol{\omega}^T + \boldsymbol{\omega}^C|\boldsymbol{k}_P).$$

(3) Under Assumption 2 and $n_U^T, n_U^C \geq 1$,

$$\mathbb{V}^2(\hat{\tau}_U) = \frac{1}{n_U^T} V^2(\boldsymbol{\omega}^T | \boldsymbol{k}_U^T) + \frac{1}{n_U^C} V^2(\boldsymbol{\omega}^C | \boldsymbol{k}_U^C) + \frac{2n_P}{n_U^T n_U^C} V(\boldsymbol{\omega}^T, \boldsymbol{\omega}^C | \boldsymbol{k}_P).$$

PROOF. Suppose that $\boldsymbol{K}_G \in \mathbb{U}$ satisfies Condition 1 and $\mathbb{X}_{\text{def}}(\hat{\tau}_G) = \mathbb{X}_{\max}$. Thus and according to Equation 36, for any $\boldsymbol{X} \in \mathbb{X}_{\max}$, it holds that $N_G^A = n_G^A \geq 1$. According to Lemma 14 (4), it holds that $\boldsymbol{k}_G^A \neq \boldsymbol{0}$. It follows that

$$\begin{aligned}
&\mathbb{V}^2(\hat{\tau}_G) \\
&= \mathbb{E}[\{\hat{\tau}_G - \mathbb{E}(\hat{\tau}_G)\}^2] \quad (\because \text{Equation 94 where } f(\boldsymbol{X}^A) = \hat{\tau}_G) \\
&= \mathbb{E}([(\hat{\tau}_G - \bar{\tau}) - \{\mathbb{E}(\hat{\tau}_G) - \bar{\tau}\}]^2) \\
&= \mathbb{E}([\{E(\boldsymbol{\beta}^T + \boldsymbol{\omega}^T | \boldsymbol{K}_G \boldsymbol{X}^T) - E(\boldsymbol{\beta}^C + \boldsymbol{\omega}^C | \boldsymbol{K}_G \boldsymbol{X}^C)\} - \{E(\boldsymbol{\beta}^T | \boldsymbol{k}_G^T) - E(\boldsymbol{\beta}^C | \boldsymbol{k}_G^C)\}]^2) \\
&\quad (\because \text{Condition 1}, \mathbb{X}_{\text{def}}(\hat{\tau}_G) = \mathbb{X}_{\max}, N_G^A \geq 1 \text{ for any } \boldsymbol{X} \in \mathbb{X}_{\max}, \text{Equations 24 and 43}) \\
&= \mathbb{E}([\{E(\boldsymbol{\beta}^T + \boldsymbol{\omega}^T | \boldsymbol{k}_G^T \boldsymbol{X}^T) - E(\boldsymbol{\beta}^C + \boldsymbol{\omega}^C | \boldsymbol{k}_G^C \boldsymbol{X}^C)\} - \{E(\boldsymbol{\beta}^T | \boldsymbol{k}_G^T \boldsymbol{X}^T) - E(\boldsymbol{\beta}^C | \boldsymbol{k}_G^C \boldsymbol{X}^C)\}]^2) \\
&\quad (\because \text{Lemmas 8 and 12 (2)}) \\
&= \mathbb{E}[\{E(\boldsymbol{\omega}^T | \boldsymbol{k}_G^T \boldsymbol{X}^T) - E(\boldsymbol{\omega}^C | \boldsymbol{k}_G^C \boldsymbol{X}^C)\}^2] \quad (\because \text{Lemma 3 (1)}) \\
&= \mathbb{E}[\{E(\boldsymbol{\omega}^T | \boldsymbol{k}_G^T \boldsymbol{X}^T)\}^2 + \{E(\boldsymbol{\omega}^C | \boldsymbol{k}_G^C \boldsymbol{X}^C)\}^2 - 2E(\boldsymbol{\omega}^T | \boldsymbol{k}_G^T \boldsymbol{X}^T) E(\boldsymbol{\omega}^C | \boldsymbol{k}_G^C \boldsymbol{X}^C)] \\
&= \mathbb{E}[\{E(\boldsymbol{\omega}^T | \boldsymbol{k}_G^T \boldsymbol{X}^T)\}^2] + \mathbb{E}[\{E(\boldsymbol{\omega}^C | \boldsymbol{k}_G^C \boldsymbol{X}^C)\}^2 - 2\mathbb{E}\{E(\boldsymbol{\omega}^T | \boldsymbol{k}_G^T \boldsymbol{X}^T) E(\boldsymbol{\omega}^C | \boldsymbol{k}_G^C \boldsymbol{X}^C)\} \\
&\quad (\because \text{Lemma 10 (1)}) \\
&= \frac{1}{n_G^T} V^2(\boldsymbol{\omega}^T | \boldsymbol{k}_G^T) + \frac{1}{n_G^C} V^2(\boldsymbol{\omega}^C | \boldsymbol{k}_G^C) + \frac{2n_G^{TC}}{n_G^T n_G^C} V(\boldsymbol{\omega}^T, \boldsymbol{\omega}^C | \boldsymbol{k}_G^{TC}) \\
&\quad (\because \boldsymbol{k}_G^T, \boldsymbol{k}_G^C \neq \boldsymbol{0}, \text{Lemmas 14 (2) and 17 (2)}).
\end{aligned} \tag{128}$$

In particular, when $\boldsymbol{k}_G^T = \boldsymbol{k}_G^C \equiv \boldsymbol{k}_G$,

$$\begin{aligned}
\mathbb{V}^2(\hat{\tau}_G) &= \frac{1}{n_G} V^2(\boldsymbol{\omega}^T | \boldsymbol{k}_G) + \frac{1}{n_G} V^2(\boldsymbol{\omega}^C | \boldsymbol{k}_G) + \frac{2}{n_G} V(\boldsymbol{\omega}^T, \boldsymbol{\omega}^C | \boldsymbol{k}_G) \quad (\because \text{Equation 128, Lemma 14 (3)}) \\
&= \frac{1}{n_G} V^2(\boldsymbol{\omega}^T + \boldsymbol{\omega}^C | \boldsymbol{k}_G) \quad (\because \text{Lemma 15 (6)}).
\end{aligned} \tag{129}$$

(1) Under Assumption 1, it holds that $\mathbb{X}_{\text{def}}(\hat{\tau}_F) = \mathbb{X}_{\max}$. When $\boldsymbol{K}_G = \boldsymbol{K}_F \in \mathbb{U}$, according to Lemma 13 (1), it follows that Equation 129 is equivalent to the desired result.

(2) Under Assumption 3 and $N_P = n_P \geq 1$ ($\because$ Lemma 9 (1)), it holds that $\mathbb{X}_{\text{def}}(\hat{\tau}_P) = \mathbb{X}_{\max}$. When $\boldsymbol{K}_G = \boldsymbol{K}_P \in \mathbb{U}$, according to Lemma 13 (2), it follows that Equation 129 is equivalent to the desired result.

(3) Under Assumption 2 and $N_U^A = n_U^A \geq 1$ ($\because$ Lemma 9 (2)), it holds that $\mathbb{X}_{\text{def}}(\hat{\tau}_U) = \mathbb{X}_{\max}$. When $\boldsymbol{K}_G = \boldsymbol{K}_U \in \mathbb{U}$, according to Lemma 13 (3), it follows that Equation 128 is equivalent to the desired result.

□

In terms of my notation, Imbens & Rubin (2015, 227) and Imai (2008, 4861, Equation (8)) formalize $\mathbb{V}^2(\hat{\tau}_F)$ as

$$\frac{1}{(2n)^2} \sum_{j=1}^{2n/2} \{(y_{1j}^C + y_{1j}^T) - (y_{2j}^C + y_{2j}^T)\}^2$$

which leads to

$$\begin{aligned}
&\frac{1}{4n^2}\sum_j\{(\mu^C+\beta^C_{1j}+\omega^C_{1j})+(\mu^T+\beta^T_{1j}+\omega^T_{1j})\\
&-(\mu^C+\beta^C_{2j}+\omega^C_{2j})-(\mu^T+\beta^T_{2j}+\omega^T_{2j})\}^2 \quad (\because \text{Lemma 4 (1)})\\
=&\frac{1}{4n^2}\sum_j\{(\omega^C_{1j}+\omega^T_{1j})-(\omega^C_{2j}+\omega^T_{2j})\}^2 \quad (\because \text{Equation 18})\\
=&\frac{1}{4n^2}\sum_j\{2(\omega^C_{1j}+\omega^T_{1j})\}^2 \quad (\because \text{Equations 19,})\\
=&\frac{1}{n}\frac{\sum_j(\omega^C_{ij}+\omega^T_{ij})^2}{\sum_j 1}\\
=&\frac{1}{n}E\{(\boldsymbol{\omega}^C+\boldsymbol{\omega}^T)^2|\mathbf{1}\} \quad (\because \text{Equation 12 (2)})\\
&\text{where } \boldsymbol{\beta}\equiv(\boldsymbol{\omega}^C+\boldsymbol{\omega}^T)^2=\{-(\boldsymbol{\omega}^C_{-i}+\boldsymbol{\omega}^T_{-i})\}^2=(\boldsymbol{\omega}^C_{-i}+\boldsymbol{\omega}^T_{-i})^2=\boldsymbol{\beta}_{-i})\\
=&\frac{1}{n}V^2(\boldsymbol{\omega}^T+\boldsymbol{\omega}^C|\boldsymbol{k}_F) \quad (\because \text{Lemma 17 (1)})\\
=&\mathbb{V}^2(\hat{\tau}_F) \quad (\because \text{Proposition 2 (1)})
\end{aligned}$$

Thus, our representation and theirs of $\mathbb{V}^2(\hat{\tau}_F)$ are equivalent.

It follows:

$$\begin{aligned}
&\mathbb{V}^2(\hat{\tau}_U)-\mathbb{V}^2(\hat{\tau}_P)\\
=&\frac{1}{n^T_U}V^2(\boldsymbol{\omega}^T|\boldsymbol{k}^T_U)+\frac{1}{n^C_U}V^2(\boldsymbol{\omega}^C|\boldsymbol{k}^C_U)+\frac{2n_P}{n^T_U n^C_U}V(\boldsymbol{\omega}^T,\boldsymbol{\omega}^C|\boldsymbol{k}_P)\\
&-\Big\{\frac{1}{n_P}V^2(\boldsymbol{\omega}^T|\boldsymbol{k}_P)+\frac{1}{n_P}V^2(\boldsymbol{\omega}^C|\boldsymbol{k}_P)+\frac{2n_P}{n_P n_P}V(\boldsymbol{\omega}^T,\boldsymbol{\omega}^C|\boldsymbol{k}_P)\Big\}\\
&(\because \text{Proposition 2 (3), Equation 128})\\
=&\Big\{\frac{n_P V^2(\boldsymbol{\omega}^T|\boldsymbol{\kappa}^T_U)}{n^T_U V^2(\boldsymbol{\omega}^T|\boldsymbol{\kappa}_P)}-1\Big\}\frac{1}{n_P}V^2(\boldsymbol{\omega}^T|\boldsymbol{\kappa}_P)+\Big\{\frac{n_P V^2(\boldsymbol{\omega}^C|\boldsymbol{\kappa}^C_U)}{n^C_U V^2(\boldsymbol{\omega}^C|\boldsymbol{\kappa}_P)}-1\Big\}\frac{1}{n_P}V^2(\boldsymbol{\omega}^C|\boldsymbol{\kappa}_P)\\
&+2\Big(\frac{n_P^2}{n^T_U n^C_U}-1\Big)\frac{1}{n_P}V(\boldsymbol{\omega}^T,\boldsymbol{\omega}^C|\boldsymbol{\kappa}_P).
\end{aligned}$$

Note that

$$\begin{aligned}
V^2(\boldsymbol{\omega}^A|\boldsymbol{\kappa}^A_U)&=E\{(\boldsymbol{\omega}^A)^2|\boldsymbol{\kappa}^A_U\} \quad (\because \text{Lemma 17 (1)})\\
&=\frac{n_P}{n^A_U}E\{(\boldsymbol{\omega}^A)^2|\boldsymbol{\kappa}_P\}+\Big(1-\frac{n_P}{n^A_U}\Big)E\{(\boldsymbol{\omega}^A)^2|\boldsymbol{\kappa}^A_U-\boldsymbol{\kappa}_P\} \quad (\because \text{Lemma 3 (5)}, \boldsymbol{\kappa}^A_U\neq\mathbf{0})\\
&=\frac{n_P}{n^A_U}V^2(\boldsymbol{\omega}^A|\boldsymbol{\kappa}_P)+\Big(1-\frac{n_P}{n^A_U}\Big)V^2(\boldsymbol{\omega}^A|\boldsymbol{\kappa}^A_U-\boldsymbol{\kappa}_P) \quad (\because \text{Lemma 17 (1)})
\end{aligned}$$

Therefore,

$$V^2(\boldsymbol{\omega}^A|\boldsymbol{\kappa}^A_U)-V^2(\boldsymbol{\omega}^A|\boldsymbol{\kappa}_P)=\frac{n^A_U-n_P}{n^A_U}\{V^2(\boldsymbol{\omega}^A|\boldsymbol{\kappa}^A_U-\boldsymbol{\kappa}_P)-V^2(\boldsymbol{\omega}^A|\boldsymbol{\kappa}_P)\}. \tag{130}$$

In the case of $\frac{n^A_U-n_P}{n^A_U}>0$, if and only if $V^2(\boldsymbol{\omega}^A|\boldsymbol{\kappa}^A_U-\boldsymbol{\kappa}_P)>V^2(\boldsymbol{\omega}^A|\boldsymbol{\kappa}_P)$, it follows that $V^2(\boldsymbol{\omega}^A|\boldsymbol{\kappa}^A_U)>V^2(\boldsymbol{\omega}^A|\boldsymbol{\kappa}_P)$.

2.4. Variance Estimator

2.4.1. Derivation

Neyman Variance Estimator. First, I derive Neyman variance estimator in a different way from Imbens & Rubin (2015, 92). For $\boldsymbol{K}_G \in \mathbb{U}$, suppose $N_G^A \geq 2$. According to Lemma 4 (1),

$$\boldsymbol{\omega}^A = \boldsymbol{y}^A - \boldsymbol{\mu}^A - \boldsymbol{\beta}^A.$$

A natural estimator of μ^A is $E(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^A)$. When we observe $X_{ij}^A = 1, K_{G,ij} = 1$ and, thus, $y_{ij}^A = Y_{ij}$ and regard $\boldsymbol{\beta}^A$ as $\boldsymbol{0}$, we may estimate ω_{ij}^A by

$$\hat{\omega}_{ij}^A \equiv Y_{ij} - 0 - E(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^A).$$

Therefore, we may estimate $\boldsymbol{\omega}^A$ by

$$\hat{\boldsymbol{\omega}}^A \equiv \boldsymbol{K}_G\boldsymbol{X}^A\boldsymbol{D}(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^A). \tag{131}$$

A natural estimator of $V^2(\boldsymbol{\omega}^A|\boldsymbol{k}_G^A)$ is

$$\begin{aligned}
&\frac{N_G^A}{N_G^A - 1}V^2(\hat{\boldsymbol{\omega}}^A|\boldsymbol{K}_G\boldsymbol{X}^A)\\
&= \frac{N_G^A}{N_G^A - 1}V^2\{\boldsymbol{K}_G\boldsymbol{X}^A\boldsymbol{D}(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^A)|\boldsymbol{K}_G\boldsymbol{X}^A\} \quad (\because \text{Equation 131})\\
&= \frac{N_G^A}{N_G^A - 1}V^2(\boldsymbol{D}(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^A)|\boldsymbol{K}_G\boldsymbol{X}^A) \quad (\because \text{Lemma 15 (3)}, \boldsymbol{K}_G, \boldsymbol{X}^A \in \mathbb{U}, \text{Lemma 1 (5) and (6)})\\
&= \frac{N_G^A}{N_G^A - 1}V^2(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^A) \quad (\because \text{Lemma 15 (2)}).
\end{aligned} \tag{132}$$

By estimating $V^2(\boldsymbol{\omega}^A|\boldsymbol{k}_G^A)$ in Equation 128 by Equation 132, replacing n_G^A in Equation 128 by N_G^A, and dismissing the third term of Equation 128, in the case of $N_G^A \geq 2$, I derive Neyman variance estimator of $\mathbb{V}^2(\hat{\tau}_G)$ as

$$\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G) \equiv \sum_A \frac{1}{N_G^A - 1}V^2(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^A), \tag{133}$$

though we do not suppose Condition 1 and $\mathbb{X}_{\text{def}}(\hat{\tau}_G) = \mathbb{X}_{\max}$ (as we do in deriving Equation 128). Note Equations 12 and 16.

In terms of my notation, Imbens & Rubin (2015, 92, 228) regard a pairwise randomized experiment as a completely randomized experiment and apply $\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_F)$, which they define as

$$\sum_A \frac{1}{n}\frac{1}{n-1}\sum_j\sum_i X_{ij}^A\{Y_{ij} - E(\boldsymbol{Y}|\boldsymbol{X}^A)\}^2$$

and which leads to

$$\begin{aligned}
&\sum_A \frac{1}{n}\frac{1}{n-1}\frac{n}{\sum_j\sum_i X_{ij}^A}\sum_j\sum_i X_{ij}^A\{Y_{ij} - E(\boldsymbol{Y}|\boldsymbol{X}^A)\}^2\\
&= \sum_A \frac{1}{n-1}E[\{\boldsymbol{Y} - \boldsymbol{E}(\boldsymbol{Y}|\boldsymbol{X}^A)\}^2|\boldsymbol{X}^A] \quad (\because \text{Equation 4})\\
&= \sum_A \frac{1}{n-1}V^2(\boldsymbol{Y}|\boldsymbol{X}^A) \quad (\because \text{Equations 89 and 90})\\
&= \hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_F) \quad (\because \text{Equations 12 and 133})
\end{aligned}$$

Thus, our definition and theirs of $\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_F)$ are equivalent.

Adjusted Neyman Variance Estimator. Actually, thanks to pairwise randomization, we do not have to give up estimating the third term of Equation 128. Even if $X_{ij}^A = 1$, noting $\omega^{-A} = -\omega_{-i}^{-A}$ due to Equation 19, in the case of $K_{G,-ij} = 1$, we may estimate ω_{ij}^{-A} by $\hat{\omega}_{ij}^{-A} \equiv -\hat{\omega}_{-ij}^{-A} = -\{Y_{-ij} - E(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^{-A})\}$, which is available because $X_{-ij}^{-A} = 1$ and $y_{-ij}^{-A} = Y_{-ij}$ is observed. When $N_G^{TC} \geq 2$, it follows that $\boldsymbol{K}_G^{SQ}\boldsymbol{X}^T \neq \boldsymbol{0}$ (for, otherwise, due to Equation 96, $N_G^{TC} = 0$, a contradiction) and we may estimate $V(\boldsymbol{\omega}^T, \boldsymbol{\omega}^C|\boldsymbol{k}_G^{TC})$ by

$$\begin{aligned}
&\frac{N_G^{TC}}{N_G^{TC}-1}V(\hat{\boldsymbol{\omega}}^T, -\hat{\boldsymbol{\omega}}_{-i}^C|\boldsymbol{K}_G\boldsymbol{X}^T\boldsymbol{K}_{G,-i}\boldsymbol{X}_{-i}^C)\\
&= -\frac{N_G^{TC}}{N_G^{TC}-1}V\{\boldsymbol{K}_G\boldsymbol{X}^T\boldsymbol{D}(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^T), \boldsymbol{K}_{G,-i}\boldsymbol{X}_{-i}^C\boldsymbol{D}(\boldsymbol{Y}_{-i}|\boldsymbol{K}_{G,-i}\boldsymbol{X}_{-i}^C)|\boldsymbol{K}_G^{SQ}\boldsymbol{X}^T\}\\
&\quad (\because \text{Equations 95 and 131, Lemmas 2 (4) and 15 (4)})\\
&= -\frac{N_G^{TC}}{N_G^{TC}-1}V\{\boldsymbol{D}(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^T), \boldsymbol{D}(\boldsymbol{Y}_{-i}|\boldsymbol{K}_{G,-i}\boldsymbol{X}_{-i}^C)|\boldsymbol{K}_G^{SQ}\boldsymbol{X}^T)\}\\
&\quad (\because \boldsymbol{K}_G, \boldsymbol{X}^A \in \mathbb{U}, \text{Equation 95, Lemmas 1 (5), 2 (4), and 15 (3), where } \boldsymbol{Z}^{(1)} = \boldsymbol{K}_G\boldsymbol{X}^T,\\
&\quad \boldsymbol{Q}^{(1)} = \boldsymbol{D}(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^T), \boldsymbol{Z}^{(2)} = \boldsymbol{K}_{G,-i}\boldsymbol{X}_{-i}^C, \boldsymbol{Q}^{(2)} = \boldsymbol{D}(\boldsymbol{Y}_{-i}|\boldsymbol{K}_{G,-i}\boldsymbol{X}_{-i}^C), \boldsymbol{U} = \boldsymbol{K}_G^{SQ}\boldsymbol{X}^T)\\
&= -\frac{N_G^{TC}}{N_G^{TC}-1}E\Big[\boldsymbol{D}\Big\{\boldsymbol{D}(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^T)\Big|\boldsymbol{K}_G^{SQ}\boldsymbol{X}^T\Big\}\boldsymbol{D}\Big\{\boldsymbol{D}(\boldsymbol{Y}_{-i}|\boldsymbol{K}_{G,-i}\boldsymbol{X}_{-i}^C)\Big|\boldsymbol{K}_G^{SQ}\boldsymbol{X}^T\Big\}\Big|\boldsymbol{K}_G^{SQ}\boldsymbol{X}^T\Big]\\
&\quad (\because \text{Equation 89})\\
&= -\frac{N_G^{TC}}{N_G^{TC}-1}E\{\boldsymbol{D}(\boldsymbol{Y}|\boldsymbol{K}_G^{SQ}\boldsymbol{X}^T)\boldsymbol{D}(\boldsymbol{Y}_{-i}|\boldsymbol{K}_G^{SQ}\boldsymbol{X}^T)|\boldsymbol{K}_G^{SQ}\boldsymbol{X}^T\} \quad (\because \text{Lemma 15 (1)}, \boldsymbol{K}_G^{SQ}\boldsymbol{X}^T \neq \boldsymbol{0})\\
&= -\frac{N_G^{TC}}{N_G^{TC}-1}V(\boldsymbol{Y}, \boldsymbol{Y}_{-i}|\boldsymbol{K}_G^{SQ}\boldsymbol{X}^T) \quad (\because \text{Equation 90})\\
&= -\frac{N_G^{TC}}{N_G^{TC}-1}V(\boldsymbol{Y}, \boldsymbol{Y}_{-i}|\boldsymbol{K}_G^{SQ}\boldsymbol{X}^A) \quad (\because \text{Equation 95, Lemma 15 (8)}).
\end{aligned} \tag{134}$$

By estimating $V(\boldsymbol{\omega}^T, \boldsymbol{\omega}^C|\boldsymbol{k}_G^{TC})$ by Equation 134 and replacing n_G^{TC} by N_G^{TC} in the third term of Equation 128 and estimating the first and second terms of Equation 128 by $\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)$, in the case of $N_G^{TC} \geq 2$ (which implies $N_G^A \geq 2$ according to Lemma 14 (5)), I propose adjusted Neyman variance estimator of $\mathbb{V}^2(\hat{\tau}_G)$ as

$$\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G) \equiv \sum_A \frac{1}{N_G^A - 1}V^2(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^A) - \frac{2(N_G^{TC})^2}{N_G^T N_G^C(N_G^{TC}-1)}V(\boldsymbol{Y}, \boldsymbol{Y}_{-i}|\boldsymbol{K}_G^{SQ}\boldsymbol{X}^A), \tag{135}$$

though we do not suppose Condition 1 and $\mathbb{X}_{\text{def}}(\hat{\tau}_G) = \mathbb{X}_{\max}$. Note Equations 12, 16, 97 through 102.

Pairwise Variance Estimator. Suppose $\boldsymbol{K}_G = \boldsymbol{K}_{G,-i}$ and $N_G \geq 2$ (c.f. Lemma 14 (1)). Note

$$\begin{aligned}
\hat{\boldsymbol{\omega}}^T - \hat{\boldsymbol{\omega}}_{-i}^C &= \boldsymbol{K}_G\boldsymbol{X}^T\boldsymbol{D}(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^T) - \boldsymbol{K}_{G,-i}\boldsymbol{X}_{-i}^C\boldsymbol{D}(\boldsymbol{Y}_{-i}|\boldsymbol{K}_{G,-i}\boldsymbol{X}_{-i}^C) \quad (\because \text{Equation 131})\\
&= \boldsymbol{K}_G\boldsymbol{X}^T\{\boldsymbol{Y} - \boldsymbol{E}(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^T)\} - \boldsymbol{K}_G\boldsymbol{X}^T\{\boldsymbol{Y}_{-i} - \boldsymbol{E}(\boldsymbol{Y}_{-i}|\boldsymbol{K}_G\boldsymbol{X}^T)\}\\
&\quad (\because \text{Equation 88}, \boldsymbol{K}_G = \boldsymbol{K}_{G,-i}, \text{Lemma 2 (1)})\\
&= \boldsymbol{K}_G\boldsymbol{X}^T\{(\boldsymbol{Y} - \boldsymbol{Y}_{-i}) - \boldsymbol{E}(\boldsymbol{Y} - \boldsymbol{Y}_{-i}|\boldsymbol{K}_G\boldsymbol{X}^T)\} \quad (\because \text{Lemma 3 (1)})\\
&= \boldsymbol{K}_G\boldsymbol{X}^T\boldsymbol{D}(\boldsymbol{Y} - \boldsymbol{Y}_{-i}|\boldsymbol{K}_G\boldsymbol{X}^T) \quad (\because \text{Equation 88})
\end{aligned} \tag{136}$$

By estimating $\boldsymbol{\omega}^A$ by Equations 19 and 131, we may estimate $V^2(\boldsymbol{\omega}^T + \boldsymbol{\omega}^C|\boldsymbol{k}_G)$ by

$$\frac{N_G}{N_G - 1}V^2(\hat{\boldsymbol{\omega}}^T - \hat{\boldsymbol{\omega}}_{-i}^C|\boldsymbol{K}_G\boldsymbol{X}^T). \tag{137}$$

When $N_G \geq 2$, by substituting Equation 137 and replacing n_G by N_G in Equation 129, I define the pairwise variance estimator of $\mathbb{V}^2(\hat{\tau}_G)$ as

$$
\begin{aligned}
\hat{\mathbb{V}}^{\text{pair}}(\hat{\tau}_G) &\equiv \frac{1}{N_G}\Big\{\frac{N_G}{N_G-1}V^2(\hat{\boldsymbol{\omega}}^T - \hat{\boldsymbol{\omega}}^C_{-i}|\boldsymbol{K}_G\boldsymbol{X}^T)\Big\} \\
&= \frac{1}{N_G-1}V^2\{\boldsymbol{K}_G\boldsymbol{X}^T\boldsymbol{D}(\boldsymbol{Y}-\boldsymbol{Y}_{-i}|\boldsymbol{K}_G\boldsymbol{X}^T)|\boldsymbol{K}_G\boldsymbol{X}^T\} \quad (\because \text{Equation 136}) \\
&= \frac{1}{N_G-1}V^2\{\boldsymbol{D}(\boldsymbol{Y}-\boldsymbol{Y}_{-i}|\boldsymbol{K}_G\boldsymbol{X}^T)|\boldsymbol{K}_G\boldsymbol{X}^T\} \quad (\because \text{Lemma 15 (3)}) \\
&= \frac{1}{N_G-1}V^2(\boldsymbol{Y}-\boldsymbol{Y}_{-i}|\boldsymbol{K}_G\boldsymbol{X}^T) \quad (\because \text{Lemma 15 (2)}) \\
&= \frac{1}{N_G-1}V^2(\boldsymbol{Y}-\boldsymbol{Y}_{-i}|\boldsymbol{K}_G\boldsymbol{X}^A) \quad (\because \text{Lemma 15 (8)})
\end{aligned}
\tag{138}
$$

Note that we can define $\hat{\mathbb{V}}^{\text{pair}}(\hat{\tau}_F)$ ($\because$ $\boldsymbol{K}_F = \boldsymbol{K}_{F,-i}, N_F \geq 2$, Equation 12) and, in the case of $N_P \geq 2$, $\hat{\mathbb{V}}^{\text{pair}}(\hat{\tau}_P)$ ($\because$ Equation 25) but not $\hat{\mathbb{V}}^{\text{pair}}(\hat{\tau}_U)$ (because it is not always true that $\boldsymbol{K}_U = \boldsymbol{K}_{U,-i}$). Note that Equations 12 and 16.

In terms of my notation, Imbens & Rubin (2015, 227) and Imai (2008, 4861, Equation (9)) define $\hat{\mathbb{V}}^{\text{pair}}(\hat{\tau}_F)$ as

$$
\frac{4}{2n(2n-2)}\sum_j\{\hat{\tau}(j) - \hat{\tau}_F\}^2, \tag{139}
$$

where

$$
\hat{\tau}(j) \equiv \sum_i(2X^T_{ij}-1)Y_{ij}. \tag{140}
$$

Note

$$
\begin{aligned}
\hat{\tau}_F &= E(\boldsymbol{Y}|\boldsymbol{K}_F\boldsymbol{X}^T) - E(\boldsymbol{Y}|\boldsymbol{K}_F\boldsymbol{X}^C) \quad (\because \text{Equation 11 and 12}) \\
&= E(\boldsymbol{Y}|\boldsymbol{K}_F\boldsymbol{X}^T) - E(\boldsymbol{Y}|\boldsymbol{K}_{F,-i}\boldsymbol{X}^T_{-i}) \quad (\because \text{Lemma 2 (1), Equation 12}) \\
&= E(\boldsymbol{Y}|\boldsymbol{K}_F\boldsymbol{X}^T) - E(\boldsymbol{Y}_{-i}|\boldsymbol{K}_F\boldsymbol{X}^T) \quad (\because \text{exchanging } i \text{ and } -i) \\
&= E(\boldsymbol{Y}-\boldsymbol{Y}_{-i}|\boldsymbol{K}_F\boldsymbol{X}^T) \quad (\because \text{Lemma 3})
\end{aligned}
\tag{141}
$$

Equation 139 leads to

$$
\begin{aligned}
&\frac{1}{n(n-1)}\sum_j\Big\{\sum_i(2X^T_{ij}-1)Y_{ij} - E(\boldsymbol{Y}-\boldsymbol{Y}_{-i}|\boldsymbol{K}_F\boldsymbol{X}^T)\Big\}^2 \quad (\because \text{Equations 140 and 141}) \\
&= \frac{1}{n(n-1)}\sum_j\sum_i K_{F,ij}X^T_{ij}\{(2X^T_{ij}-1)Y_{ij} + (2X^T_{-ij}-1)Y_{-ij} - E(\boldsymbol{Y}-\boldsymbol{Y}_{-i}|\boldsymbol{K}_F\boldsymbol{X}^T)\}^2 \\
&= \frac{1}{n(n-1)}\sum_j\sum_i K_{F,ij}[\{2(X^T_{ij})^2 - X^T_{ij}\}Y_{ij} + (2X^T_{ij}X^T_{-ij} - X^T_{ij})Y_{-ij} - X^T_{ij}E(\boldsymbol{Y}-\boldsymbol{Y}_{-i}|\boldsymbol{K}_F\boldsymbol{X}^T)]^2 \\
&= \frac{1}{n(n-1)}\sum_j\sum_i K_{F,ij}\{X^T_{ij}Y_{ij} - X^T_{ij}Y_{-ij} - X^T_{ij}E(\boldsymbol{Y}-\boldsymbol{Y}_{-i}|\boldsymbol{K}_F\boldsymbol{X}^T)\}^2 \quad (\because \text{Lemma 1 (5) and 2 (3)}) \\
&= \frac{1}{n(n-1)}\frac{n}{\sum_j\sum_i K_{F,ij}X^T_{ij}}\sum_j\sum_i K_{F,ij}X^T_{ij}\{(Y_{ij}-Y_{-ij}) - E(\boldsymbol{Y}-\boldsymbol{Y}_{-i}|\boldsymbol{K}_F\boldsymbol{X}^T)\}^2 \\
&= \frac{1}{n-1}E[\{(\boldsymbol{Y}-\boldsymbol{Y}_{-i}) - \boldsymbol{E}(\boldsymbol{Y}-\boldsymbol{Y}_{-i}|\boldsymbol{K}_F\boldsymbol{X}^T)\}^2|\boldsymbol{K}_F\boldsymbol{X}^T] \quad (\because \text{Equations 3 (1) and 4}) \\
&= \frac{1}{n-1}E[\{\boldsymbol{D}(\boldsymbol{Y}-\boldsymbol{Y}_{-i}|\boldsymbol{K}_F\boldsymbol{X}^T)\}^2|\boldsymbol{K}_F\boldsymbol{X}^T] \quad (\because \text{Equation 88}) \\
&= \frac{1}{N_{F-1}}V^2(\boldsymbol{Y}-\boldsymbol{Y}_{-i}|\boldsymbol{K}_F\boldsymbol{X}^T) \quad (\because \text{Equation 90}) \\
&= \hat{\mathbb{V}}^{\text{pair}}(\hat{\tau}_F). \quad (\because \text{Equations 12 and 138})
\end{aligned}
$$

Thus, our definition and theirs of $\hat{\mathbb{V}}^{\text{pair}}(\hat{\tau}_F)$ are equivalent. Note that, in deriving the variance estimator, Imbens & Rubin (2015) "assume that the treatment effect is constant and additive, not only within pairs but also across pairs" (226), namely, according to Lemmas 4 (3) and 5 (1), $\boldsymbol{\omega}^T - \boldsymbol{\omega}^C = \mathbf{0}$ and, thus, $\boldsymbol{\beta}^T - \boldsymbol{\beta}^C = \mathbf{0}$, though we assume $\boldsymbol{\beta}^T = \boldsymbol{\beta}^C = \mathbf{0}$.

2.4.2. Properties

Lemma 18 (FS Expectation of FS Covariance). *Suppose that $\boldsymbol{K}_G \in \mathbb{U}$ satisfies Condition 1. For $h \in \{1,2\}$, let $\boldsymbol{\omega}, \boldsymbol{\omega}^{(h)} \in \mathbb{W}, \boldsymbol{\beta}, \boldsymbol{\beta}^{(h)} \in \mathbb{B}, \boldsymbol{k}_G \in \{\boldsymbol{k}_G^T, \boldsymbol{k}_G^C\}$.*

(1)

$$\mathbb{E}\{V_{h=1}^2(\boldsymbol{\beta}^{(h)}|\boldsymbol{k}_G\boldsymbol{X}^A)\} = V_{h=1}^2(\boldsymbol{\beta}^{(h)}|\boldsymbol{k}_G)$$

(2) When $\boldsymbol{k}_G \neq \mathbf{0}$,

$$\mathbb{E}\{V_{h=1}^2(\boldsymbol{\omega}^{(h)}|\boldsymbol{k}_G\boldsymbol{X}^A)\} = \frac{n_G - 1}{n_G} V_{h=1}^2(\boldsymbol{\omega}^{(h)}|\boldsymbol{k}_G)$$

(3)

$$\mathbb{E}\{V(\boldsymbol{\beta}, \boldsymbol{\omega}|\boldsymbol{k}_G\boldsymbol{X}^A)\} = 0$$

Proof. (1) Note

$$\begin{aligned} \boldsymbol{D}(\boldsymbol{\beta}^{(h)}|\boldsymbol{k}_G\boldsymbol{X}^A) &= \boldsymbol{\beta}^{(h)} - \boldsymbol{E}(\boldsymbol{\beta}^{(h)}|\boldsymbol{k}_G\boldsymbol{X}^A) \quad (\because \text{Equation 88}) \\ &= \boldsymbol{\beta}^{(h)} - \boldsymbol{E}(\boldsymbol{\beta}^{(h)}|\boldsymbol{k}_G) \quad (\because \text{Lemma 12 (2)}) \\ &= \boldsymbol{D}(\boldsymbol{\beta}^{(h)}|\boldsymbol{k}_G), \quad (\because \text{Equation 88}) \end{aligned} \tag{142}$$

Let

$$\boldsymbol{\beta}' = \prod_{h=1}^{2} \boldsymbol{D}(\boldsymbol{\beta}^{(h)}|\boldsymbol{k}_G) \tag{143}$$

It follows

$$\begin{aligned} \boldsymbol{\beta}'_{-i} &= \prod_{h=1}^{2} \boldsymbol{D}(\boldsymbol{\beta}_{-i}^{(h)}|\boldsymbol{k}_G) \quad (\because \text{Equation 143}) \\ &= \prod_{h=1}^{2} \boldsymbol{D}(\boldsymbol{\beta}^{(h)}|\boldsymbol{k}_G) \quad (\because \text{Equation 20}) \\ &= \boldsymbol{\beta}' \quad (\because \text{Equation 143}) \\ \therefore \boldsymbol{\beta}' &\in \mathbb{B} \end{aligned} \tag{144}$$

Thus,

$$\begin{aligned} &\mathbb{E}\{V_{h=1}^2(\boldsymbol{\beta}^{(h)}|\boldsymbol{k}_G\boldsymbol{X}^A)\} \\ &= \mathbb{E}\Big[E\Big\{\prod_{h=1}^{2} \boldsymbol{D}(\boldsymbol{\beta}^{(h)}|\boldsymbol{k}_G\boldsymbol{X}^A)\Big|\boldsymbol{k}_G\boldsymbol{X}^A\Big\}\Big] \quad (\because \text{Equation 89} \\ &= \mathbb{E}\Big[E\Big\{\prod_{h=1}^{2} \boldsymbol{D}(\boldsymbol{\beta}^{(h)}|\boldsymbol{k}_G)\Big|\boldsymbol{k}_G\boldsymbol{X}^A\Big\}\Big] \quad (\because \text{Equation 142}) \\ &= E\Big\{\prod_{h=1}^{2} \boldsymbol{D}(\boldsymbol{\beta}^{(h)}|\boldsymbol{k}_G)\Big|\boldsymbol{k}_G\Big\} \quad (\because \text{Lemma 12 (2), Equations 143 and 144}) \\ &= V_{h=1}^2(\boldsymbol{\beta}^{(h)}|\boldsymbol{k}_G) \quad (\because \text{Equation 89}) \end{aligned}$$

(2) Note

$$\begin{aligned} \boldsymbol{D}(\boldsymbol{\omega}|\boldsymbol{k}_G) &= \boldsymbol{\omega} - \boldsymbol{E}(\boldsymbol{\omega}|\boldsymbol{k}_G) \quad (\because \text{Equation 88}) \\ &= \boldsymbol{\omega} \quad (\because \text{Lemma 12 (1)}) \end{aligned} \tag{145}$$

Let

$$\boldsymbol{\beta}' = \prod_{h=1}^{2} \boldsymbol{\omega}^{(h)} \tag{146}$$

It follows

$$\begin{aligned} \boldsymbol{\beta}'_{-i} &= \prod_{h=1}^{2} \boldsymbol{\omega}^{(h)}_{-i} \quad (\because \text{Equation 146}) \\ &= \prod_{h=1}^{2} -\boldsymbol{\omega}^{(h)} \quad (\because \text{Equation 21}) \\ &= \prod_{h=1}^{2} \boldsymbol{\omega}^{(h)} \\ &= \boldsymbol{\beta}' \quad (\because \text{Equation 146}) \\ \therefore \boldsymbol{\beta}' &\in \mathbb{B} \end{aligned} \tag{147}$$

It holds

$$\begin{aligned} &\mathbb{E}\Big\{E\Big(\prod_{h=1}^{2} \boldsymbol{\omega}^{(h)}\Big|\boldsymbol{k}_G\boldsymbol{X}^A\Big)\Big\} \\ &= E\Big(\prod_{h=1}^{2} \boldsymbol{\omega}^{(h)}\Big|\boldsymbol{k}_G\Big) \quad (\because \text{Lemma 12 (2), Equations 146 and 147}) \\ &= E\Big\{\prod_{h=1}^{2} \boldsymbol{D}(\boldsymbol{\omega}^{(h)}|\boldsymbol{k}_G)\Big|\boldsymbol{k}_G\Big\}. \quad (\because \text{Equation 145}) \\ &= V_{h=1}^{2}(\boldsymbol{\omega}^{(h)}|\boldsymbol{k}_G) \quad (\because \text{Equation 90}) \end{aligned} \tag{148}$$

Thus,

$$\begin{aligned} &\mathbb{E}\{V_{h=1}^{2}(\boldsymbol{\omega}^{(h)}|\boldsymbol{k}_G\boldsymbol{X}^A)\} \\ &= \mathbb{E}\Big\{E\Big(\prod_{h=1}^{2} \boldsymbol{\omega}^{(h)}\Big|\boldsymbol{k}_G\boldsymbol{X}^A\Big) - \prod_{h=1}^{2} E(\boldsymbol{\omega}^{(h)}|\boldsymbol{k}_G\boldsymbol{X}^A)\Big\} \quad (\because \text{Lemma 15 (7)}) \\ &= V_{h=1}^{2}(\boldsymbol{\omega}^{(h)}|\boldsymbol{k}_G) - \frac{1}{n_G}V_{h=1}^{2}(\boldsymbol{\omega}^{(h)}|\boldsymbol{k}_G) \quad (\because \boldsymbol{k}_G \neq \boldsymbol{0}, \text{Equation 148, Lemmas 10, 14 (3) and 17 (2)}) \\ &= \frac{n_G - 1}{n_G}V_{h=1}^{2}(\boldsymbol{\omega}^{(h)}|\boldsymbol{k}_G) \end{aligned}$$

(3) Let

$$\boldsymbol{\omega}' = \boldsymbol{\beta}\boldsymbol{\omega}. \tag{149}$$

It follows

$$\begin{aligned} \boldsymbol{\omega}'_{-i} &= \boldsymbol{\beta}_{-i}\boldsymbol{\omega}_{-i} \quad (\because \text{Equation 149}) \\ &= \boldsymbol{\beta}(-\boldsymbol{\omega}) \quad (\because \text{Equations 20 and 21}) \\ &= -\boldsymbol{\omega}' \quad (\because \text{Equation 149}) \\ \therefore \boldsymbol{\omega}' &\in \mathbb{W} \end{aligned} \tag{150}$$

Thus,

$$
\begin{aligned}
&\mathbb{E}\{V(\boldsymbol{\beta},\boldsymbol{\omega}|\boldsymbol{k}_G\boldsymbol{X}^A)\}\\
&=\mathbb{E}\{E(\boldsymbol{\beta}\boldsymbol{\omega}|\boldsymbol{k}_G\boldsymbol{X}^A)-E(\boldsymbol{\beta}|\boldsymbol{k}_G\boldsymbol{X}^A)E(\boldsymbol{\omega}|\boldsymbol{k}_G\boldsymbol{X}^A)\}\quad(\because \text{Lemma 15 (7)})\\
&=\mathbb{E}\{E(\boldsymbol{\omega}'|\boldsymbol{k}_G\boldsymbol{X}^A)\}-E(\boldsymbol{\beta}|\boldsymbol{k}_G)\mathbb{E}\{E(\boldsymbol{\omega}|\boldsymbol{k}_G\boldsymbol{X}^A)\}\quad(\because \text{Lemma 10 and 12 (2), Equation 149})\\
&=0-E(\boldsymbol{\beta}|\boldsymbol{k}_G)\cdot 0\quad(\because \text{Lemma 12 (1), Equation 150})\\
&=0
\end{aligned}
\tag{151}
$$

□

Proposition 3 (Bias of the Neyman Variance Estimators: FS). *(1) Under Assumption 1,*

$$\mathbb{E}\{\hat{\mathbb{V}}^{Neyman}(\hat{\tau}_F)\}-\mathbb{V}^2(\hat{\tau}_F)=\frac{1}{n_F-1}\{V^2(\boldsymbol{\beta}^T|\boldsymbol{k}_F)+V^2(\boldsymbol{\beta}^C|\boldsymbol{k}_F)\}-\frac{2}{n_F}V(\boldsymbol{\omega}^T,\boldsymbol{\omega}^C|\boldsymbol{k}_F).$$

(2) Under Assumption 3 and $n_P\geq 2$,

$$\mathbb{E}\{\hat{\mathbb{V}}^{Neyman}(\hat{\tau}_P)\}-\mathbb{V}^2(\hat{\tau}_P)=\frac{1}{n_P-1}\{V^2(\boldsymbol{\beta}^T|\boldsymbol{k}_P)+V^2(\boldsymbol{\beta}^C|\boldsymbol{k}_P)\}-\frac{2}{n_P}V(\boldsymbol{\omega}^T,\boldsymbol{\omega}^C|\boldsymbol{k}_P).$$

(3) Under Assumption 2 and $n_U^T,n_U^C\geq 2$,

$$\mathbb{E}\{\hat{\mathbb{V}}^{Neyman}(\hat{\tau}_U)\}-\mathbb{V}^2(\hat{\tau}_U)=\frac{1}{n_U^T-1}V^2(\boldsymbol{\beta}^T|\boldsymbol{k}_U^T)+\frac{1}{n_U^C-1}V^2(\boldsymbol{\beta}^C|\boldsymbol{k}_U^C)-\frac{2n_P}{n_U^Tn_U^C}V(\boldsymbol{\omega}^T,\boldsymbol{\omega}^C|\boldsymbol{k}_P).$$

Proof. Note

$$
\begin{aligned}
\boldsymbol{D}(\boldsymbol{y}^A|\boldsymbol{Z})&=\boldsymbol{y}^A-\boldsymbol{E}(\boldsymbol{y}^A|\boldsymbol{Z})\quad(\because \text{Equation 88})\\
&=\boldsymbol{\mu}^A+\boldsymbol{\beta}^A+\boldsymbol{\omega}^A-\boldsymbol{E}(\boldsymbol{\mu}^A+\boldsymbol{\beta}^A+\boldsymbol{\omega}^A|\boldsymbol{Z})\quad(\because \text{Lemma 4 (1)})\\
&=(\boldsymbol{\mu}^A-\boldsymbol{\mu}^A)+\{\boldsymbol{\beta}^A-\boldsymbol{E}(\boldsymbol{\beta}^A|\boldsymbol{Z})\}+\{\boldsymbol{\omega}^A-\boldsymbol{E}(\boldsymbol{\omega}^A|\boldsymbol{Z})\}\quad(\because \text{Lemma 3})\\
&=\boldsymbol{D}(\boldsymbol{\beta}^A|\boldsymbol{Z})+\boldsymbol{D}(\boldsymbol{\omega}^A|\boldsymbol{Z})\quad(\because \text{Equation 88}).
\end{aligned}
\tag{152}
$$

Suppose that $\boldsymbol{K}_G\in\mathbb{U}$ satisfies Condition 1 and $\mathbb{X}_{\text{def}}\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)\}=\mathbb{X}_{\max}$. It follows that $n_G^A\geq 2$, $\boldsymbol{k}_G^A\neq\boldsymbol{0}$ (due to Lemma 14 (4)), and

$$
\begin{aligned}
&\mathbb{E}\{V^2(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^A)\}\\
&=\mathbb{E}\{V^2(\boldsymbol{Y}\boldsymbol{X}^A|\boldsymbol{K}_G\boldsymbol{X}^A)\}\quad(\because \boldsymbol{X}^A\in\mathbb{U},\text{Lemma 1 (5) and 15 (3)})\\
&=\mathbb{E}\{V^2(\boldsymbol{y}^A\boldsymbol{X}^A|\boldsymbol{k}_G^A\boldsymbol{X}^A)\}\quad(\because \text{Equation 22, Lemma 8})\\
&=\mathbb{E}\{V^2(\boldsymbol{y}^A|\boldsymbol{k}_G^A\boldsymbol{X}^A)\}\quad(\because \text{Lemma 15 (3)})\\
&=\mathbb{E}(E[\{\boldsymbol{D}(\boldsymbol{y}^A|\boldsymbol{k}_G^A\boldsymbol{X}^A)\}^2|\boldsymbol{k}_G^A\boldsymbol{X}^A])\quad(\because \text{Equation 90})\\
&=\mathbb{E}(E[\{\boldsymbol{D}(\boldsymbol{\beta}^A|\boldsymbol{k}_G^A\boldsymbol{X}^A)+\boldsymbol{D}(\boldsymbol{\omega}^A|\boldsymbol{k}_G^A\boldsymbol{X}^A)\}^2|\boldsymbol{k}_G^A\boldsymbol{X}^A])\quad(\because \text{Equation 152})\\
&=\mathbb{E}(E[\{\boldsymbol{D}(\boldsymbol{\beta}^A|\boldsymbol{k}_G^A\boldsymbol{X}^A)\}^2|\boldsymbol{k}_G^A\boldsymbol{X}^A])+\mathbb{E}(E[\{\boldsymbol{D}(\boldsymbol{\omega}^A|\boldsymbol{k}_G^A\boldsymbol{X}^A)\}^2|\boldsymbol{k}_G^A\boldsymbol{X}^A])\\
&\quad+2\mathbb{E}[E\{\boldsymbol{D}(\boldsymbol{\beta}^A|\boldsymbol{k}_G^A\boldsymbol{X}^A)\boldsymbol{D}(\boldsymbol{\omega}^A|\boldsymbol{k}_G^A\boldsymbol{X}^A)|\boldsymbol{k}_G^A\boldsymbol{X}^A\}]\quad(\because \text{Lemma 10})\\
&=\mathbb{E}\{V^2(\boldsymbol{\beta}^A|\boldsymbol{k}_G^A\boldsymbol{X}^A)\}+\mathbb{E}\{V^2(\boldsymbol{\omega}^A|\boldsymbol{k}_G^A\boldsymbol{X}^A)\}+2\mathbb{E}\{V(\boldsymbol{\beta}^A,\boldsymbol{\omega}^A|\boldsymbol{k}_G^A\boldsymbol{X}^A)\}\quad(\because \text{Equation 89}\\
&=V^2(\boldsymbol{\beta}^A|\boldsymbol{k}_G^A)+\frac{n_G^A-1}{n_G^A}V^2(\boldsymbol{\omega}^A|\boldsymbol{k}_G^A)+0\quad(\because \boldsymbol{k}_G^A\neq\boldsymbol{0},\text{Lemma 18})
\end{aligned}
\tag{153}
$$

Since $\mathbb{X}_{\max} = \mathbb{X}_{\text{def}}\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)\} \subseteq \mathbb{X}_{\text{def}}(\hat{\tau}_G) \subseteq \mathbb{X}_{\max}$, it holds that $\mathbb{X}_{\text{def}}(\hat{\tau}_G) = \mathbb{X}_{\max}$. Therefore,

$$\begin{aligned}
&\mathbb{E}\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)\} - \mathbb{V}^2(\hat{\tau}_G)\\
&= \mathbb{E}\Big\{\sum_A \frac{1}{N_G^A - 1} V^2(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^A)\Big\} - \Big\{\sum_A \frac{1}{n_G^A} V^2(\boldsymbol{\omega}^A|\boldsymbol{k}_G^A) + \frac{2n_G^{TC}}{n_G^T n_G^C} V(\boldsymbol{\omega}^T, \boldsymbol{\omega}^C|\boldsymbol{k}_G^{TC})\Big\}\\
&\quad (\because \text{Equations 128 and 133, } \mathbb{X}_{\text{def}}(\hat{\tau}_G) = \mathbb{X}_{\max})\\
&= \sum_A \Big[\frac{1}{n_G^A - 1}\Big\{V^2(\boldsymbol{\beta}^A|\boldsymbol{k}_G^A) + \frac{n_G^A - 1}{n_G^A} V^2(\boldsymbol{\omega}^A|\boldsymbol{k}_G^A)\Big\} - \frac{1}{n_G^A} V^2(\boldsymbol{\omega}^A|\boldsymbol{k}_G^A)\Big] - \frac{2n_G^{TC}}{n_G^T n_G^C} V(\boldsymbol{\omega}^T, \boldsymbol{\omega}^C|\boldsymbol{k}_G^{TC})\\
&\quad (\because \text{Equations 36 and 153, Lemma 10 })\\
&= \sum_A \frac{1}{n_G^A - 1} V^2(\boldsymbol{\beta}^A|\boldsymbol{k}_G^A) - \frac{2n_G^{TC}}{n_G^T n_G^C} V(\boldsymbol{\omega}^T, \boldsymbol{\omega}^C|\boldsymbol{k}_G^{TC})
\end{aligned} \tag{154}$$

(1) Under Assumption 1, it holds that $\mathbb{X}_{\text{def}}\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_F)\} = \mathbb{X}_{\max}$ ($\because$ Equation 12, $N_F^A = n \geq 2$). When $\boldsymbol{K}_G = \boldsymbol{K}_F \in \mathbb{U}$, according to Lemma 13 (1), it follows that Equation 154 is equivalent to the desired result.

(2) Under Assumption 3 and $N_P^A = N_P = n_P \geq 2$ ($\because$ Equation 16, Lemma 9 (1)), it holds that $\mathbb{X}_{\text{def}}\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_P)\} = \mathbb{X}_{\max}$. When $\boldsymbol{K}_G = \boldsymbol{K}_P \in \mathbb{U}$, according to Lemma 13 (2), it follows that Equation 154 is equivalent to the desired result.

(3) Under Assumption 2 and $N_U^A = n_U^A \geq 2$ ($\because$ Lemma 9 (2)), it holds that $\mathbb{X}_{\text{def}}\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_U)\} = \mathbb{X}_{\max}$. When $\boldsymbol{K}_G = \boldsymbol{K}_U \in \mathbb{U}$, according to Lemma 13 (3), it follows that Equation 154 is equivalent to the desired result.

□

Proposition 4 (Bias of the Adjusted Neyman Variance Estimators: FS). *(1) Under Assumption 1,*

$$\mathbb{E}\{\hat{\mathbb{V}}^{\textit{Adj-Neyman}}(\hat{\tau}_F)\} - \mathbb{V}^2(\hat{\tau}_F) = \frac{1}{n_F - 1} V^2(\boldsymbol{\beta}^T - \boldsymbol{\beta}^C|\boldsymbol{k}_F) \geq 0.$$

(2) Under Assumption 3 and $n_P \geq 2$,

$$\mathbb{E}\{\hat{\mathbb{V}}^{\textit{Adj-Neyman}}(\hat{\tau}_P)\} - \mathbb{V}^2(\hat{\tau}_P) = \frac{1}{n_P - 1} V^2(\boldsymbol{\beta}^T - \boldsymbol{\beta}^C|\boldsymbol{k}_P) \geq 0.$$

(3) Under Assumption 2 and $n_P \geq 2$,

$$\begin{aligned}
&\mathbb{E}\{\hat{\mathbb{V}}^{\textit{Adj-Neyman}}(\hat{\tau}_U)\} - \mathbb{V}^2(\hat{\tau}_U)\\
&= \frac{1}{n_U^T - 1} V^2(\boldsymbol{\beta}^T|\boldsymbol{k}_U^T) + \frac{1}{n_U^C - 1} V^2(\boldsymbol{\beta}^C|\boldsymbol{k}_U^C) - \frac{2n_P^2}{n_U^T n_U^C (n_P - 1)} V(\boldsymbol{\beta}^T, \boldsymbol{\beta}^C|\boldsymbol{k}_P).
\end{aligned}$$

Note that, unlike Proposition 3 (3), one of the conditions of Proposition 4 (3) is not $n_U^T, n_U^C \geq 2$ but $n_P \geq 2$.

Proof. Suppose that $\boldsymbol{K}_G \in \mathbb{U}$. Note

$$\begin{aligned}
\boldsymbol{K}_G^{SQ}\boldsymbol{X}^T &= \boldsymbol{K}_G\boldsymbol{K}_{G,-i}\boldsymbol{X}^T\boldsymbol{X}_{-i}^C \quad (\because \text{Equation 95, Lemma 2 (4)})\\
&= \boldsymbol{k}_G^T\boldsymbol{X}^T\boldsymbol{k}_{G,-i}^C\boldsymbol{X}_{-i}^C \quad (\because \text{Lemma 8})\\
&= \boldsymbol{k}_G^{TC}\boldsymbol{X}^T\boldsymbol{X}_{-i}^C \quad (\because \text{Equation 103})
\end{aligned} \tag{155}$$

Suppose that $\boldsymbol{K}_G$ satisfies Condition 1 and $\mathbb{X}_{\text{def}}\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G)\} = \mathbb{X}_{\max}$. It follows that

$n_G^{TC} \geq 2$ and

$$\begin{aligned}
&\mathbb{E}\{V(\boldsymbol{Y}, \boldsymbol{Y}_{-i}|\boldsymbol{K}_G^{SQ}\boldsymbol{X}^A)\} \\
&= \mathbb{E}\{V(\boldsymbol{Y}\boldsymbol{X}^A, \boldsymbol{Y}_{-i}\boldsymbol{X}_{-i}^{-A}|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A\boldsymbol{X}_{-i}^{-A})\} \\
&\quad (\because \boldsymbol{X}^A, \boldsymbol{X}^{-A} \in \mathbb{U}, \text{Lemmas 1 (5) and 15 (3), Equation 155}, \boldsymbol{X}^T\boldsymbol{X}_{-i}^C = \boldsymbol{X}^C\boldsymbol{X}_{-i}^T) \\
&= \mathbb{E}\{V(\boldsymbol{y}^T\boldsymbol{X}^A, \boldsymbol{y}_{-i}^C\boldsymbol{X}_{-i}^{-A}|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A\boldsymbol{X}_{-i}^{-A})\} \quad (\because \text{Equation 22}) \\
&= \mathbb{E}\{V(\boldsymbol{y}^T, \boldsymbol{y}_{-i}^C|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A\boldsymbol{X}_{-i}^{-A})\} \quad (\because \text{Lemma 15 (3)}) \\
&= \mathbb{E}[E\{\boldsymbol{D}(\boldsymbol{y}^T|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A)\boldsymbol{D}(\boldsymbol{y}_{-i}^C|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A)|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A\}] \quad (\because \text{Equation 89, Lemma 2 (1) and (4)}) \\
&= \mathbb{E}(E[\{\boldsymbol{D}(\boldsymbol{\beta}^T|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A) + \boldsymbol{D}(\boldsymbol{\omega}^T|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A)\}\{\boldsymbol{D}(\boldsymbol{\beta}_{-i}^C|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A) + \boldsymbol{D}(\boldsymbol{\omega}_{-i}^C|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A)\}|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A]) \\
&\qquad (\because \text{Equation 152}) \\
&= \mathbb{E}[E\{\boldsymbol{D}(\boldsymbol{\beta}^T|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A)\boldsymbol{D}(\boldsymbol{\beta}^C|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A)|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A\}] \\
&\quad + \mathbb{E}[E\{\boldsymbol{D}(\boldsymbol{\omega}^T|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A)\boldsymbol{D}(-\boldsymbol{\omega}^C|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A)\}|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A\}] \\
&\quad + \mathbb{E}[E\{\boldsymbol{D}(\boldsymbol{\beta}^T|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A)\boldsymbol{D}(-\boldsymbol{\omega}^C|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A)\}|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A\}] \\
&\quad + \mathbb{E}[E\{\boldsymbol{D}(\boldsymbol{\beta}^C|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A)\boldsymbol{D}(\boldsymbol{\omega}^T|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A)|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A\}] \\
&\quad (\because \text{Lemmas 3 (1) and 10 (1), Equations 18 and 19}) \\
&= \mathbb{E}\{V(\boldsymbol{\beta}^T, \boldsymbol{\beta}^C|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A)\} - \mathbb{E}\{V(\boldsymbol{\omega}^T, \boldsymbol{\omega}^C|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A)\} - \mathbb{E}\{V(\boldsymbol{\beta}^T, \boldsymbol{\omega}^C|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A)\} \\
&\quad + \mathbb{E}\{V(\boldsymbol{\beta}^C, \boldsymbol{\omega}^T|\boldsymbol{k}_G^{TC}\boldsymbol{X}^A)\} \quad (\because \text{Equation 89, Lemma 10 (2) and 15 (4)}) \\
&= V(\boldsymbol{\beta}^T, \boldsymbol{\beta}^C|\boldsymbol{k}_G^{TC}) - \frac{n_G^{TC}-1}{n_G^{TC}}V(\boldsymbol{\omega}^T, \boldsymbol{\omega}^C|\boldsymbol{k}_G^{TC}) - 0 + 0 \\
&\quad (\because n_G^{TC} \geq 2, \boldsymbol{k}_G^{TC} \neq \boldsymbol{0}, \text{Lemmas 14 (4) and 18, Equation 109})
\end{aligned} \tag{156}$$

Since $\mathbb{X}_{\max} = \mathbb{X}_{\text{def}}\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G)\} \subseteq \mathbb{X}_{\text{def}}\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)\} \subseteq \mathbb{X}_{\max}$, it holds that $\mathbb{X}_{\text{def}}\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)\} = \mathbb{X}_{\max}$. Therefore,

$$\begin{aligned}
&\mathbb{E}\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G)\} - \mathbb{V}(\hat{\tau}_G) \\
&= \mathbb{E}\Big\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G) - \frac{2n_G^{TC}}{n_G^T n_G^C}\frac{n_G^{TC}}{n_G^{TC}-1}V(\boldsymbol{Y}, \boldsymbol{Y}_{-i}|\boldsymbol{K}_G^{SQ}\boldsymbol{X}^A)\Big\} - \mathbb{V}(\hat{\tau}_G) \\
&\quad (\because \text{Equations 36, 133 and 135, Lemma14 (2)}, n_G^A \geq n_G^{TC} \geq 2) \\
&= \Big[\mathbb{E}\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)\} - \mathbb{V}(\hat{\tau}_G)\Big] - \mathbb{E}\Big\{\frac{2n_G^{TC}}{n_G^T n_G^C}\frac{n_G^{TC}}{n_G^{TC}-1}V(\boldsymbol{Y}, \boldsymbol{Y}_{-i}|\boldsymbol{K}_G^{SQ}\boldsymbol{X}^A)\Big\} \\
&\quad (\because \text{Lemma 10 (1)}, \mathbb{X}_{\text{def}}\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G)\} = \mathbb{X}_{\text{def}}\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)\} = \mathbb{X}_{\max}) \\
&= \sum_A \frac{1}{n_G^A - 1}V^2(\boldsymbol{\beta}^A|\boldsymbol{k}_G^A) - \frac{2n_G^{TC}}{n_G^T n_G^C}V(\boldsymbol{\omega}^T, \boldsymbol{\omega}^C|\boldsymbol{k}_G^{TC}) \\
&\quad - \frac{2n_G^{TC}}{n_G^T n_G^C}\frac{n_G^{TC}}{n_G^{TC}-1}\Big\{V(\boldsymbol{\beta}^T, \boldsymbol{\beta}^C|\boldsymbol{k}_G^{TC}) - \frac{n_G^{TC}-1}{n_G^{TC}}V(\boldsymbol{\omega}^T, \boldsymbol{\omega}^C|\boldsymbol{k}_G^{TC})\Big\} \\
&\quad (\because \mathbb{X}_{\text{def}}\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)\} = \mathbb{X}_{\max}, \text{Equations 154 and 156, Lemma 10 (2)}) \\
&= \sum_A \frac{1}{n_G^A - 1}V^2(\boldsymbol{\beta}^A|\boldsymbol{k}_G^A) - \frac{2n_G^{TC}}{n_G^T n_G^C}\frac{n_G^{TC}}{n_G^{TC}-1}V(\boldsymbol{\beta}^T, \boldsymbol{\beta}^C|\boldsymbol{k}_G^{TC})
\end{aligned} \tag{157}$$

When $\boldsymbol{k}_G^T = \boldsymbol{k}_G^C \equiv \boldsymbol{k}_G$,

$$\begin{aligned}\mathbb{E}\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G) - \mathbb{V}(\hat{\tau}_G)\} &= \sum_A \frac{1}{n_G - 1} V^2(\boldsymbol{\beta}^A|\boldsymbol{k}_G) - \frac{2}{n_G - 1} V(\boldsymbol{\beta}^T, \boldsymbol{\beta}^C|\boldsymbol{k}_G) \\ &\quad (\because \text{Equation 157, Lemma 14 (3)}) \\ &= \frac{1}{n_G - 1} V^2(\boldsymbol{\beta}^T - \boldsymbol{\beta}^C|\boldsymbol{k}_G) \quad (\because \text{Lemma 15 (6)})\end{aligned} \tag{158}$$

(1) Under Assumption 1, it holds that $\mathbb{X}_{\text{def}}\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_F)\} = \mathbb{X}_{\max}$ ($\because$ Equations 12 and 98, $N_F^{TC} = N_F = n \geq 2$). When $\boldsymbol{K}_G = \boldsymbol{K}_F \in \mathbb{U}$, according to Lemma 13 (1), it follows that Equation 158 is equivalent to the desired result.

(2) Under Assumption 3 and $N_P^{TC} = N_P = n_P \geq 2$ ($\because$ Equation 102, Lemma 9 (1)), it holds that $\mathbb{X}_{\text{def}}\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_P)\} = \mathbb{X}_{\max}$. When $\boldsymbol{K}_G = \boldsymbol{K}_P \in \mathbb{U}$, according to Lemma 13 (2), it follows that Equation 158 is equivalent to the desired result.

(3) Under Assumption 2 and $N_U^{TC} = N_P = n_P \geq 2$ ($\because$ Equation 100, Lemma 9 (2)), it holds that $\mathbb{X}_{\text{def}}\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_U)\} = \mathbb{X}_{\max}$. When $\boldsymbol{K}_G = \boldsymbol{K}_U \in \mathbb{U}$, according to Lemma 13 (3), it follows that Equation 157 is equivalent to the desired result.

□

PROPOSITION 5 (EQUIVALENCE BETWEEN VARIANCE ESTIMATORS). *(1)*

$$\hat{\mathbb{V}}^{\textit{Adj-Neyman}}(\hat{\tau}_F) = \hat{\mathbb{V}}^{\textit{pair}}(\hat{\tau}_F).$$

(2)

$$\hat{\mathbb{V}}^{\textit{Adj-Neyman}}(\hat{\tau}_P) = \hat{\mathbb{V}}^{\textit{pair}}(\hat{\tau}_P).$$

PROOF. When $\boldsymbol{K}_G = \boldsymbol{K}_{G,-i}$ and $N_G \geq 2$ (c.f. Lemma 14 (1)),

$$\begin{aligned}&\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G) \\ &= \frac{1}{N_G - 1} V^2(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^T) + \frac{1}{N_G - 1} V^2(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^C) - \frac{2}{N_G - 1} V(\boldsymbol{Y}, \boldsymbol{Y}_{-i}|\boldsymbol{K}_G\boldsymbol{X}^T) \\ &\quad (\because \text{Equation 135, Lemma 14 (1)}) \\ &= \frac{1}{N_G - 1}[E\{\boldsymbol{D}(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^T)^2|\boldsymbol{K}_G\boldsymbol{X}^T\} + E\{\boldsymbol{D}(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^C)^2|\boldsymbol{K}_G\boldsymbol{X}^C\} \\ &\quad - 2E\{\boldsymbol{D}(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^T)\boldsymbol{D}(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^C)|\boldsymbol{K}_G\boldsymbol{X}^T\}] \quad (\because \text{Equations 89 and 90}) \\ &= \frac{1}{N_G - 1}[E\{\boldsymbol{D}(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^T)^2|\boldsymbol{K}_G\boldsymbol{X}^T\} + E\{\boldsymbol{D}(\boldsymbol{Y}|\boldsymbol{K}_{G,-i}\boldsymbol{X}_{-i}^T)^2|\boldsymbol{K}_{G,-i}\boldsymbol{X}_{-i}^T\} \\ &\quad - 2E\{\boldsymbol{D}(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^T)\boldsymbol{D}(\boldsymbol{Y}|\boldsymbol{K}_{G,-i}\boldsymbol{X}_{-i}^T)|\boldsymbol{K}_G\boldsymbol{X}^T\}] \quad (\because \text{Lemma 2 (1)}, \boldsymbol{K}_G = \boldsymbol{K}_{G,-i}) \\ &= \frac{1}{N_G - 1}[E\{\boldsymbol{D}(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^T)^2|\boldsymbol{K}_G\boldsymbol{X}^T\} + E\{\boldsymbol{D}(\boldsymbol{Y}_{-i}|\boldsymbol{K}_G\boldsymbol{X}^T)^2|\boldsymbol{K}_G\boldsymbol{X}^T\} \\ &\quad - 2E\{\boldsymbol{D}(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^T)\boldsymbol{D}(\boldsymbol{Y}_{-i}|\boldsymbol{K}_G\boldsymbol{X}^T)|\boldsymbol{K}_G\boldsymbol{X}^T\}] \quad (\because \text{exchanging } i \text{ and } -i) \\ &= \frac{1}{N_G - 1} E\{\boldsymbol{D}(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^T)^2 + \boldsymbol{D}(\boldsymbol{Y}_{-i}|\boldsymbol{K}_G\boldsymbol{X}^T)^2 - 2\boldsymbol{D}(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^T)\boldsymbol{D}(\boldsymbol{Y}_{-i}|\boldsymbol{K}_G\boldsymbol{X}^T)|\boldsymbol{K}_G\boldsymbol{X}^T\} \\ &\quad (\because \text{Lemma 3 (1)}) \\ &= \frac{1}{N_G - 1} E[\{\boldsymbol{D}(\boldsymbol{Y}|\boldsymbol{K}_G\boldsymbol{X}^T) - \boldsymbol{D}(\boldsymbol{Y}_{-i}|\boldsymbol{K}_G\boldsymbol{X}^T)\}^2|\boldsymbol{K}_G\boldsymbol{X}^T] \\ &= \frac{1}{N_G - 1} E[\{\boldsymbol{D}(\boldsymbol{Y} - \boldsymbol{Y}_{-i}|\boldsymbol{K}_G\boldsymbol{X}^T)\}^2|\boldsymbol{K}_G\boldsymbol{X}^T] \quad (\because \text{Lemma 15 (5)}) \\ &= \frac{1}{N_G - 1} V^2(\boldsymbol{Y} - \boldsymbol{Y}_{-i}|\boldsymbol{K}_G\boldsymbol{X}^T) \quad (\because \text{Lemma 90}) \\ &= \hat{\mathbb{V}}^{\text{pair}}(\hat{\tau}_G) \quad (\because \text{Equation 138})\end{aligned}$$

By substituting $\boldsymbol{K}_G = \boldsymbol{K}_F$ or $\boldsymbol{K}_G = \boldsymbol{K}_P$ and noting Equations 10, 11, 12, 15, and 25, the desired results follow.

□

Brief comments follow. First, unlike Proposition 3 (3), one of the conditions of Proposition 4 (3) is not $n_U^T, n_U^C \geq 2$ but $n_P \geq 2$ because $N^{TC}(\boldsymbol{K}_U) = N_P = n_P$ should be not fewer than two. Second, Propositions 3 and 4 hold whether or not Assumption 4 or 5 holds.

Note that $\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_F) = \hat{\mathbb{V}}^{\text{pair}}(\hat{\tau}_F)$ and $\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_P) = \hat{\mathbb{V}}^{\text{pair}}(\hat{\tau}_P)$ have positive bias, though the directions of the other variance estimators' bias are unknown. Note that

$$\begin{aligned}
&\mathbb{E}\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_U)\} - \mathbb{V}^2(\hat{\tau}_U) \\
&= \sum_A \frac{1}{n_U^A - 1} V^2(\boldsymbol{\beta}^A | \boldsymbol{k}_U^A) - \frac{n_P^2}{n_U^T n_U^C (n_P - 1)} V(\boldsymbol{\beta}^T, \boldsymbol{\beta}^C | \boldsymbol{k}_P) \\
&= \sum_A \left\{ \frac{1}{n_U^A - 1} V^2(\boldsymbol{\beta}^A | \boldsymbol{k}_U^A) - \frac{2n_P^2}{n_U^T n_U^C (n_P - 1)} V^2(\boldsymbol{\beta}^A | \boldsymbol{k}_P) \right\} + \frac{2n_P^2}{n_U^T n_U^C (n_P - 1)} V^2(\boldsymbol{\beta}^A - \boldsymbol{\beta}^C | \boldsymbol{k}_P).
\end{aligned}$$

In a similar way to Eqution 130, we can easily show that $V^2(\boldsymbol{\beta}^A | \boldsymbol{k}_U^A)$ can be either larger or smaller than $V^2(\boldsymbol{\beta}^A | \boldsymbol{k}_P)$. Note also that $\frac{1}{n_U^A - 1}$ can be either larger or smaller than $\frac{n_P^2}{n_U^A n_U^{-A} (n_P - 1)}$. For instance, when $n_U^A = 3, n_U^{-A} = n_P = 2$, the former is smaller than the latter.

In terms of my notation, Imbens & Rubin (2015, 222, 227) formalize the bias of $\hat{\mathbb{V}}^{\text{pair}}(\hat{\tau}_F)$ as

$$\frac{4}{2n(2n-2)} \sum_j \{\tau(j) - \bar{\tau}\}^2 \tag{159}$$

where

$$\begin{aligned}
\tau(j) &\equiv \frac{1}{2} \sum_i (y_{ij}^T - y_{ij}^C) \\
&= \frac{1}{2} \sum_i \{(\mu_{ij}^T + \beta_{ij}^T + \omega_{ij}^T) - (\mu_{ij}^C + \beta_{ij}^C + \omega_{ij}^C)\} \quad (\because \text{Lemma 4 (1)}) \\
&= \frac{1}{2} \cdot 2(\bar{\tau} + \beta_{\cdot j}^T - \beta_{\cdot j}^C) \quad (\because \text{Lemmas 4 (2) and 5 (1), Equation 18}) \\
&= \bar{\tau} + \beta_{\cdot j}^T - \beta_{\cdot j}^C.
\end{aligned} \tag{160}$$

Equation 159 leads to

$$\begin{aligned}
&\frac{1}{n(n-1)} \sum_j (\bar{\tau} + \beta_{\cdot j}^T - \beta_{\cdot j}^C - \bar{\tau})^2 \quad (\because \text{Equation 160}) \\
&= \frac{1}{n-1} \frac{\sum_j k_{F,\cdot j} (\beta_{\cdot j}^T - \beta_{\cdot j}^C)^2}{\sum_j k_{F,\cdot j}} \quad (\because \text{Equation 12}) \\
&= \frac{1}{n-1} E\{(\boldsymbol{\beta}^T - \boldsymbol{\beta}^C)^2 | \boldsymbol{k}_F\} \quad (\because \text{Lemma 12 (2)}) \\
&= \frac{1}{n-1} E[\{\boldsymbol{\beta}^T - \boldsymbol{\beta}^C - \boldsymbol{E}(\boldsymbol{\beta}^T - \boldsymbol{\beta}^C | \boldsymbol{k}_F)\}^2 | \boldsymbol{k}_F] \quad (\because \text{Lemma 5 (2), Equation 12}) \\
&= \frac{1}{n-1} V(\boldsymbol{\beta}^T - \boldsymbol{\beta}^C | \boldsymbol{k}_F) \quad (\because \text{Equation 90}) \\
&= \mathbb{E}\{\hat{\mathbb{V}}^{\text{pair}}(\hat{\tau}_F)\} - \mathbb{V}^2(\hat{\tau}_F) \quad (\because \text{Propositions 5 and 4 (1), Equation 12})
\end{aligned} \tag{161}$$

In terms of my notation, Imai (2008, 4862, Proposition 1) formalizes the bias of $\hat{\mathbb{V}}^{\text{pair}}(\hat{\tau}_F)$ as

$$\frac{1}{(n-1)n} \sum_j \left\{ \tau(j) - \frac{\sum_j \tau(j)}{n} \right\}^2. \tag{162}$$

Equation 162 leads to

$$\begin{aligned}
&\frac{1}{(n-1)n}\sum_j \left(\bar{\tau}+\beta^T_{\cdot j}-\beta^C_{\cdot j}-\frac{\sum_j \bar{\tau}+\beta^T_{\cdot j}-\beta^C_{\cdot j}}{n}\right)^2 \quad (\because \text{Equation 160})\\
&=\frac{1}{(n-1)n}\sum_j (\bar{\tau}+\beta^T_{\cdot j}-\beta^C_{\cdot j}-\bar{\tau})^2 \quad (\because \text{Lemma 5 (2)})\\
&=\mathbb{E}\{\hat{\mathbb{V}}^{\text{pair}}(\hat{\tau}_F)\}-\mathbb{V}^2(\hat{\tau}_F) \quad (\because \text{Equation 161})
\end{aligned}$$

Note that efficiency analysis by Imai (2008, 4863-4866) compares $\hat{\mathbb{V}}^{\text{pair}}(\hat{\tau}_F)$ in pairwise randomized experiments and $\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_F)$ *in completely randomized experiments*. Thus, his interest lies in comparison between the two randomization experiment designs. By contrast, I compare $\hat{\mathbb{V}}^{\text{pair}}(\hat{\tau}_F)=\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_F)$ and $\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_F)$, *both* in pairwise randomized experiments. That is, I focus on comparison between the two variance estimators in pairwise randomized experiments.

Imai (2008, p. 4866) warns that "the variance estimator used when breaking the matches [i.e., $\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_F)$] will be biased," though neither Imai (2008) nor Imbens & Rubin (2015) nor Snedecor & Cochran (1980, pp. 99–102) derives the bias as Proposition 3 (1) does. Note that $\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_F)$ is also biased (Proposition 4 (1)).

3. SUPER-POPULATION

3.1. Setting

3.1.1. Definition

I introduce super-population from which the above n pairs are drawn. I define super-population variables and operators in the same way as, and denote them by adding superscript $*$ to, the corresponding variables and operators for the above n pairs of finite sample. All lemmas and equations in Section 2 hold in the super-population by adding superscript $*$ to the corresponding variables and operators unless they are not concerned with both $\boldsymbol{S}^*$ and $\mathbb{E}^*(\cdot)$, which I define shortly. Accordingly, there are $n^*(>n)$ pairs in the super-population, and unit $i^*\in\{1,2\}$ in pair $j^*\in\{1,2,\ldots,n^*\}$ is denoted by unit i^*j^*. Define

- $S^*_{i^*j^*}$: the sampling indicator of unit i^*j^*

$$S^*_{i^*j^*}=\begin{cases}1 \text{ if pair } j^* \text{ is sampled.}\\ 0 \text{ otherwise.}\end{cases}$$

Thus,

$$\begin{aligned}
S^*_{1j^*}=S^*_{2j^*}&\equiv S^*_{\cdot j^*}\\
\sum_{j^*} S^*_{\cdot j^*}&=n.
\end{aligned}$$

When unit i^*j^* is sampled ($S^*_{i^*j^*}=1$), the unit is denoted by ij in the sampled n pairs, where

$$\begin{aligned}
i&=I(i^*j^*|\boldsymbol{S}^*)\equiv i^*\\
j&=J(j^*|\boldsymbol{S}^*)\equiv\sum_{j^{*\prime}=1}^{j^*} S^*_{\cdot j^{*\prime}}.
\end{aligned}$$

For $i\in\{1,2\}$ and $j\in\{1,\ldots,n\}$, define

$$\begin{aligned}
I^*(ij|\boldsymbol{S}^*)&\equiv I^{-1}(i|\boldsymbol{S}^*)=i\\
J^*(j|\boldsymbol{S}^*)&\equiv J^{-1}(j|\boldsymbol{S}^*).
\end{aligned}$$

Below, for ease of presentation, I simply notate $I(i^*j^*)$, $J(j^*)$, $I^*(ij)$, and $J^*(j)$ without clarifying that they are conditioned on $\boldsymbol{S}^*$. The generic super-population vector is denoted by

$$\boldsymbol{Q}^* \equiv \{Q^*_{i^*j^*}\}_{11}^{2n^*}$$

Denote the generic sampled vector by

$$\boldsymbol{Q}^{(*)}(\boldsymbol{S}^*) \equiv \{Q^*_{I^*(ij)J^*(j)}\}_{11}^{2n} = \{Q_{ij}\}_{11}^{2n} = \boldsymbol{Q} \tag{163}$$

It follows

$$\begin{aligned} \boldsymbol{y}^A &= \boldsymbol{y}^{A(*)}(\boldsymbol{S}^*) \\ \boldsymbol{r}^A &= \boldsymbol{r}^{A(*)}(\boldsymbol{S}^*) \\ \boldsymbol{X} &= \boldsymbol{X}^{(*)}(\boldsymbol{S}^*) \\ \boldsymbol{K}_G &= \boldsymbol{K}_G^{(*)}(\boldsymbol{S}^*) \\ \boldsymbol{k}_G^A &= \boldsymbol{k}_G^{A(*)}(\boldsymbol{S}^*) \\ \boldsymbol{q} &= \boldsymbol{q}^{(*)}(\boldsymbol{S}^*). \end{aligned} \tag{164}$$

When $\boldsymbol{Z}^* \neq \boldsymbol{0}^*$, the super-population weighted mean operator is defined as

$$E^*(\boldsymbol{Q}^*|\boldsymbol{Z}^*) \equiv \frac{\sum_{j^*}\sum_{i^*} Z^*_{i^*j^*}Q^*_{i^*j^*}}{\sum_{j^*}\sum_{i^*} Z^*_{i^*j^*}}, \tag{165}$$

where $Z_{ij} \geq 0, \sum_{j^*} \equiv \sum_{j=1}^{n^*}$ and $\sum_{i^*} = \sum_{i^*=1}^{2}$, and, by abusing notation,

$$E^*(\boldsymbol{Q}^*|\boldsymbol{0}^*) \equiv 0. \tag{166}$$

Denote

$$\boldsymbol{E}^{(*)}(\boldsymbol{Q}^*|\boldsymbol{Z}^*) \equiv \{E^*(\boldsymbol{Q}^*|\boldsymbol{Z}^*)\}_{11}^{2n}.$$

In particular, the super-population estimand, the super-population average treatment effect, is defined as

$$\bar{\tau}^* \equiv E^*(\boldsymbol{\tau}^*|\boldsymbol{1}^*). \tag{167}$$

Lemma 19 (Sample and Super-population). *(1)*

$$\sum_j \sum_i Q_{ij}^{(*)}(\boldsymbol{S}^*) = \sum_{j^*}\sum_{i^*} S^*_{i^*j^*}Q^*_{i^*j^*}$$

(2) When $\sum_h \boldsymbol{Q}^{(h)*} = \boldsymbol{Q}^*$,

$$\sum_h \boldsymbol{Q}^{(h)(*)}(\boldsymbol{S}^*) = \boldsymbol{Q}^{(*)}(\boldsymbol{S}^*)$$

(3) When $\prod_h \boldsymbol{Q}^{(h)*} = \boldsymbol{Q}^*$,

$$\prod_h \boldsymbol{Q}^{(h)(*)}(\boldsymbol{S}^*) = \boldsymbol{Q}^{(*)}(\boldsymbol{S}^*)$$

(4)

$$E\{\boldsymbol{Q}^{(*)}(\boldsymbol{S}^*)|\boldsymbol{Z}^{(*)}(\boldsymbol{S}^*)\} = E^*(\boldsymbol{Q}^*|\boldsymbol{S}^*\boldsymbol{Z}^*)$$

(5)

$$\begin{aligned} \mu^A &= \mu^{A*} + E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{S}^*) \\ \boldsymbol{\beta}^A &= \boldsymbol{\beta}^{A(*)}(\boldsymbol{S}^*) - \boldsymbol{E}^{(*)}(\boldsymbol{\beta}^{A*}|\boldsymbol{S}^*) \\ \boldsymbol{\omega}^A &= \boldsymbol{\omega}^{A(*)}(\boldsymbol{S}^*) \\ \bar{\tau} &= \bar{\tau}^* + E^*(\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*}|\boldsymbol{S}^*) \end{aligned}$$

(6)

$$\sum_j \prod_i Q_{ij}^{(*)}(\boldsymbol{S}^*) = \sum_{j^*} S^*_{\cdot j^*} \prod_{i^*} Q^*_{i^*j^*}$$

PROOF. (1)

$$\begin{aligned}
\sum_{j^*}\sum_{i^*} S^*_{i^*j^*} Q^*_{i^*j^*} &= \sum_{j^*} S^*_{\cdot j^*} \sum_{i^*} Q^*_{i^*j^*} \\
&= \sum_{j^*:S^*_{\cdot j^*}=1} 1 \sum_{i^*} Q^*_{i^*j^*} + \sum_{j^*:S^*_{\cdot j^*}=0} 0 \sum_{i^*} Q^*_{i^*j^*} \\
&= \sum_{J(j^*):S^*_{\cdot J^{-1}\{J(j^*)\}}=1} \sum_{I(i^*j^*)} Q^*_{I^{-1}\{I(i^*j^*)J(j^*)\}J^{-1}\{J(j^*)\}} \\
&\quad (\because \text{changing the index from } j^* \text{ and } i^* \text{ to } J(j^*) \text{ and } I(i^*j^*) \text{, respectively}) \\
&= \sum_j \sum_i Q^*_{I^*(ij)J^*(j)} \quad (\because \text{substituting } J(j^*) \text{ with } j \text{ and } I(i^*j^*) \text{ with } i) \\
&= \sum_j \sum_i Q^{(*)}_{ij}(\boldsymbol{S}^*) \quad (\because \text{Equation 163})
\end{aligned}$$

(2)

$$\begin{aligned}
\sum_h \boldsymbol{Q}^{(h)(*)}(\boldsymbol{S}^*) &= \Big\{ \sum_h Q^{(h)*}_{I^*(ij)J^*(j)} \Big\}_{11}^{2n^*} \quad (\because \text{Equation 163}) \\
&= \{Q^*_{I^*(ij)J^*(j)}\}_{11}^{2n^*} \quad (\because \sum_h \boldsymbol{Q}^{(h)*} = \boldsymbol{Q}^*) \\
&= \boldsymbol{Q}^{(*)}(\boldsymbol{S}^*) \quad (\because \text{Equation 163})
\end{aligned}$$

(3)

$$\begin{aligned}
\prod_h \boldsymbol{Q}^{(h)(*)}(\boldsymbol{S}^*) &= \Big\{ \prod_h Q^{(h)*}_{I^*(ij)J^*(j)} \Big\}_{11}^{2n^*} \quad (\because \text{Equation 163}) \\
&= \{Q^*_{I^*(ij)J^*(j)}\}_{11}^{2n^*} \quad (\because \prod_h \boldsymbol{Q}^{(h)*} = \boldsymbol{Q}^*) \\
&= \boldsymbol{Q}^{(*)}(\boldsymbol{S}^*) \quad (\because \text{Equation 163})
\end{aligned}$$

(4) Let $\boldsymbol{Q}^{*\prime} = \boldsymbol{Z}^*\boldsymbol{Q}^*$. When $\boldsymbol{S}^*\boldsymbol{Z}^* \neq \boldsymbol{0}^*$, it follows

$$\begin{aligned}
E^*(\boldsymbol{Q}^*|\boldsymbol{S}^*\boldsymbol{Z}^*) &= \frac{\sum_{j^*}\sum_{i^*} S^*_{i^*j^*} Z^*_{i^*j^*} Q^*_{i^*j^*}}{\sum_{j^*}\sum_{i^*} S^*_{i^*j^*} Z^*_{i^*j^*}} \quad (\because \text{Equation 165}) \\
&= \frac{\sum_{j^*}\sum_{i^*} S^*_{i^*j^*} Q^{*\prime}_{i^*j^*}}{\sum_{j^*}\sum_{i^*} S^*_{i^*j^*} Z^*_{i^*j^*}} \quad (\because \boldsymbol{Q}^{*\prime} = \boldsymbol{Z}^*\boldsymbol{Q}^*) \\
&= \frac{\sum_j \sum_i Q^{(*)\prime}_{ij}(\boldsymbol{S}^*)}{\sum_j \sum_i Z^{(*)}_{ij}(\boldsymbol{S}^*)} \quad (\because \text{Lemma 19 (1)}) \\
&= \frac{\sum_j \sum_i Z^{(*)}_{ij}(\boldsymbol{S}^*) Q^{(*)}_{ij}(\boldsymbol{S}^*)}{\sum_j \sum_i Z^{(*)}_{ij}(\boldsymbol{S}^*)} \quad (\because \text{Lemma 19 (3)}) \\
&= E\{\boldsymbol{Q}^{(*)}(\boldsymbol{S}^*)|\boldsymbol{Z}^{(*)}(\boldsymbol{S}^*)\} \quad (\because \text{Equation 4})
\end{aligned}$$

When $\boldsymbol{S}^*\boldsymbol{Z}^* = \boldsymbol{0}^*$,

$$\begin{aligned}
\sum_j \sum_i Z_{ij}^{(*)}(\boldsymbol{S}^*) &= \sum_{j^*} \sum_{i^*} S_{i^*j^*}^* Z_{i^*j^*}^* \quad (\because \text{Lemma 19 (1)}) \\
&= 0 \quad (\because \boldsymbol{S}^*\boldsymbol{Z}^* = \boldsymbol{0}^*) \\
\therefore \boldsymbol{Z}^{(*)}(\boldsymbol{S}^*) &= \boldsymbol{0} \quad (\because Z_{i^*j^*}^* \geq 0) \\
\therefore E\{\boldsymbol{Q}^{(*)}(\boldsymbol{S}^*)|\boldsymbol{Z}^{(*)}(\boldsymbol{S}^*)\} &= 0 \quad (\because \text{Equation 5}) \\
E^*(\boldsymbol{Q}^*|\boldsymbol{S}^*\boldsymbol{Z}^*) &= E^*(\boldsymbol{Q}^*|\boldsymbol{0}^*) \quad (\because \boldsymbol{S}^*\boldsymbol{Z}^* = \boldsymbol{0}^*) \\
&= 0 \quad (\because \text{Equation 166}) \\
\therefore E^*(\boldsymbol{Q}^*|\boldsymbol{S}^*\boldsymbol{Z}^*) &= E\{\boldsymbol{Q}^{(*)}(\boldsymbol{S}^*)|\boldsymbol{Z}^{(*)}(\boldsymbol{S}^*)\}
\end{aligned}$$

(5) It holds

$$\begin{aligned}
\omega_{ij}^{A(*)}(\boldsymbol{S}^*) &= \omega_{I^*(ij)J^*(j)}^{A*} \quad (\because \text{Equation 163}) \\
&= y_{I^*(ij)J^*(j)}^{A*} - \frac{1}{2}\sum_i y_{I^*(ij)J^*(j)}^{A*} \quad (\because \text{Equation 19}) \\
&= y_{ij}^{A} - \frac{1}{2}\sum_i y_{ij}^{A} \quad (\because \text{Equation 164}) \\
&= \omega_{ij}^{A} \quad (\because \text{Equation 19})
\end{aligned} \tag{168}$$

Note

$$\begin{aligned}
E^*(\boldsymbol{\omega}^{A*}|\boldsymbol{S}^*) &= E^*(\boldsymbol{\omega}^{A*}|\boldsymbol{S}^*\boldsymbol{1}^*) \\
&= E\{\boldsymbol{\omega}^{A(*)}(\boldsymbol{S}^*)|\boldsymbol{1}^{(*)}(\boldsymbol{S}^*)\} \quad (\because \text{Lemma 19 (4)}) \\
&= E(\boldsymbol{\omega}^{A}|\boldsymbol{1}) \quad (\because \text{Equations 164 and 168}) \\
&= 0 \quad (\because \text{Lemma 12 (1)})
\end{aligned} \tag{169}$$

It follows

$$\begin{aligned}
\mu^A &= E(\boldsymbol{y}^A|\boldsymbol{1}) \quad (\because \text{Equation 17}) \\
&= E\{\boldsymbol{y}^{A(*)}(\boldsymbol{S}^*)|\boldsymbol{1}^{(*)}(\boldsymbol{S}^*)\} \quad (\because \text{Equation 164}) \\
&= E^*(\boldsymbol{y}^{A*}|\boldsymbol{S}^*\boldsymbol{1}^*) \quad (\because \text{Lemma 19 (4)}) \\
&= E^*(\boldsymbol{\mu}^{A*} + \boldsymbol{\beta}^{A*} + \boldsymbol{\omega}^{A*}|\boldsymbol{S}^*) \quad (\because \text{Lemma 4 (1)}) \\
&= \mu^{A*} + E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{S}^*) \quad (\because \text{Lemma 3 (1) and (2), Equation 169})
\end{aligned} \tag{170}$$

It also holds

$$\begin{aligned}
\beta_{\cdot j}^{A} &= \frac{1}{2}\sum_i y_{ij}^{A} - \mu^A \quad (\because \text{Equation 18}) \\
&= \frac{1}{2}\sum_i y_{I^*(ij)J^*(j)}^{A*} - \mu^A \quad (\because \text{Equation 164}) \\
&= \frac{1}{2}\Big(\sum_i \mu^{A*} + \beta_{\cdot J^*(j)}^{A*} + \omega_{I^*(ij)J^*(j)}^{A*}\Big) - \{\mu^{A*} + E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{S}^*)\} \quad (\because \text{Lemma 4 (1), Equation 170}) \\
&= \beta_{\cdot J^*(j)}^{A*} - E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{S}^*) \quad (\because \text{Lemma 5 (1)}) \\
&= \beta_{\cdot j}^{A(*)}(\boldsymbol{S}^*) - E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{S}^*) \quad (\because \text{Equation 163})
\end{aligned} \tag{171}$$

It follows

$$\begin{aligned}
\bar{\tau} &= \mu^T - \mu^C \quad (\because \text{Lemma 4 (2)}) \\
&= \{\mu^{T*} + E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{S}^*)\} - \{\mu^{C*} + E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{S}^*)\} \quad (\because \text{Equation 170}) \\
&= \bar{\tau}^* + E^*(\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*}|\boldsymbol{S}^*) \quad (\because \text{Lemma 3 (1) and 4 (2)})
\end{aligned}$$

(6)

$$
\begin{aligned}
&\sum_{j^*} S^*_{\cdot j^*} \prod_{i^*} Q^*_{i^*j^*} \\
&= \sum_{j^*:S^*_{\cdot j^*}=1} 1 \prod_{i^*} Q^*_{i^*j^*} + \sum_{j^*:S^*_{\cdot j^*}=0} 0 \prod_{i^*} Q^*_{i^*j^*} \\
&= \sum_{J(j^*):S^*_{\cdot J^{-1}\{J(j^*)\}}=1} \prod_{I(i^*j^*)} Q^*_{I^{-1}\{I(i^*j^*)J(j^*)\}J^{-1}\{J(j^*)\}} \\
&\quad (\because \text{changing the index from } j^* \text{ and } i^* \text{ to } J(j^*) \text{ and } I(i^*j^*), \text{respectively}) \\
&= \sum_j \prod_i Q^*_{I^*(ij)J^*(j)} \quad (\because \text{substituting } J(j^*) \text{ and } I(i^*j^*) \text{ with } j \text{ and } i, \text{respectively}) \\
&= \sum_j \prod_i Q^{(*)}_{ij}(\boldsymbol{S}^*) \quad (\because \text{Equation 163})
\end{aligned}
$$

□

LEMMA 20 (ESTIMATION ERROR: SP). *For any* $\boldsymbol{K}^*_G \in \mathbb{U}^*$, *when* $N^A_G \geq 1$,

$$\hat{\tau}_G - \bar{\tau}^* = E^*(\boldsymbol{\beta}^{T*} + \boldsymbol{\omega}^{T*}|\boldsymbol{S}^*\boldsymbol{K}^*_G\boldsymbol{X}^{T*}) - E^*(\boldsymbol{\beta}^{C*} + \boldsymbol{\omega}^{C*}|\boldsymbol{S}^*\boldsymbol{K}^*_G\boldsymbol{X}^{C*})$$

PROOF. Let

$$
\begin{aligned}
\boldsymbol{\delta}^A &\equiv \boldsymbol{\beta}^A + \boldsymbol{\omega}^A \\
\boldsymbol{\delta}^{A*} &\equiv \boldsymbol{\beta}^{A*} - \boldsymbol{E}^*(\boldsymbol{\beta}^{A*}|\boldsymbol{S}^*) + \boldsymbol{\omega}^{A*} \\
\boldsymbol{Z}^A &\equiv \boldsymbol{K}_G\boldsymbol{X}^A \\
\boldsymbol{Z}^{A*} &\equiv \boldsymbol{K}^*_G\boldsymbol{X}^{A*}.
\end{aligned} \tag{172}
$$

It follows

$$
\begin{aligned}
\boldsymbol{\delta}^{A(*)}(\boldsymbol{S}^*) &= \boldsymbol{\beta}^{A(*)}(\boldsymbol{S}^*) - \boldsymbol{E}^{(*)}(\boldsymbol{\beta}^{A*}|\boldsymbol{S}^*)(\boldsymbol{S}^*) + \boldsymbol{\omega}^{A(*)}(\boldsymbol{S}^*) \quad (\because \text{Equation 172, Lemma 19 (2)}) \\
&= \{\boldsymbol{\beta}^A + \boldsymbol{E}^{(*)}(\boldsymbol{\beta}^{A*}|\boldsymbol{S}^*)\} - \boldsymbol{E}^{(*)}(\boldsymbol{\beta}^{A*}|\boldsymbol{S}^*) + \boldsymbol{\omega}^A \quad (\because \text{Lemma 19 (5), Equation 164}) \\
&= \boldsymbol{\beta}^A + \boldsymbol{\omega}^A \\
&= \boldsymbol{\delta}^A \quad (\because \text{Equation 172})
\end{aligned} \tag{173}
$$

It also holds

$$
\begin{aligned}
\boldsymbol{Z}^{A(*)}(\boldsymbol{S}^*) &= \boldsymbol{K}^{(*)}_G(\boldsymbol{S}^*)\boldsymbol{X}^{A(*)}(\boldsymbol{S}^*) \quad (\because \text{Equation 172, Lemma 19 (3)}) \\
&= \boldsymbol{K}_G\boldsymbol{X}^A \quad (\because \text{Equation 164}) \\
&= \boldsymbol{Z}^A \quad (\because \text{Equation 172})
\end{aligned} \tag{174}
$$

Therefore,

$$
\begin{aligned}
&\hat{\tau}_G - \bar{\tau}^* \\
&= (\hat{\tau}_G - \bar{\tau}) - (\bar{\tau}^* - \bar{\tau}) \\
&= \{E(\boldsymbol{\beta}^T + \boldsymbol{\omega}^T|\boldsymbol{K}_G\boldsymbol{X}^T) - E(\boldsymbol{\beta}^C + \boldsymbol{\omega}^C|\boldsymbol{K}_G\boldsymbol{X}^C)\} + E^*(\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*}|\boldsymbol{S}^*) \\
&\quad (\because N_G^A \geq 1, \text{Equation 24, Lemma 19 (5)}) \\
&= E(\boldsymbol{\delta}^T|\boldsymbol{Z}^T) - E(\boldsymbol{\delta}^C|\boldsymbol{Z}^C) + E^*(\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*}|\boldsymbol{S}^*) \quad (\because \text{Equation 172}) \\
&= E\{\boldsymbol{\delta}^{T(*)}(\boldsymbol{S}^*)|\boldsymbol{Z}^{T(*)}(\boldsymbol{S}^*)\} - E\{\boldsymbol{\delta}^{C(*)}(\boldsymbol{S}^*)|\boldsymbol{Z}^{C(*)}(\boldsymbol{S}^*)\} + E^*(\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*}|\boldsymbol{S}^*) \\
&\quad (\because \text{Equations 173, and 174}) \\
&= E^*(\boldsymbol{\delta}^{T*}|\boldsymbol{S}^*\boldsymbol{Z}^{T*}) - E^*(\boldsymbol{\delta}^{C*}|\boldsymbol{S}^*\boldsymbol{Z}^{C*}) + E^*(\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*}|\boldsymbol{S}^*) \quad (\because \text{Lemma 19 (4)}) \\
&= E^*\{\boldsymbol{\beta}^{T*} - \boldsymbol{E}^*(\boldsymbol{\beta}^{T*}|\boldsymbol{S}^*) + \boldsymbol{\omega}^{T*}|\boldsymbol{S}^*\boldsymbol{K}_G^*\boldsymbol{X}^{T*}\} - E^*\{\boldsymbol{\beta}^{C*} - \boldsymbol{E}^*(\boldsymbol{\beta}^{C*}|\boldsymbol{S}^*) + \boldsymbol{\omega}^{C*}|\boldsymbol{S}^*\boldsymbol{K}_G^*\boldsymbol{X}^{C*}\} \\
&\quad + E^*(\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*}|\boldsymbol{S}^*) \quad (\because \text{Equation 172}) \\
&= E^*(\boldsymbol{\beta}^{T*} + \boldsymbol{\omega}^{T*}|\boldsymbol{S}^*\boldsymbol{K}_G^*\boldsymbol{X}^{T*}) - E^*(\boldsymbol{\beta}^{C*} + \boldsymbol{\omega}^{C*}|\boldsymbol{S}^*\boldsymbol{K}_G^*\boldsymbol{X}^{C*}) \quad (\because \text{Lemma 3})
\end{aligned}
$$

□

3.2. Bias

Sampling. Denote the maximum sampling space by

$$\mathbb{S}^*_{\max} \equiv \left\{\boldsymbol{s}^* \middle| s^* \in \mathbb{U}^*, s^*_{1j^*} = s^*_{2j^*}, \sum_{j^*} s^*_{\cdot j^*} = n\right\}.$$

Let

$$\{\boldsymbol{\Theta}^*\} \equiv \{(\boldsymbol{y}^{T*}, \boldsymbol{y}^{C*}, \boldsymbol{r}^{T*}, \boldsymbol{r}^{C*})|\boldsymbol{y}^{T*}, \boldsymbol{y}^{C*} \in \mathbb{Y}^*_{\max}, \boldsymbol{r}^{T*}, \boldsymbol{r}^{C*} \in \mathbb{U}^*\}$$

We assume random sampling, that is,

- Ignorability of sampling: for any $\Theta^* \in \{\boldsymbol{\Theta}^*\}, \boldsymbol{s}^* \in \mathbb{S}^*_{\max}$,

$$\Pr(\boldsymbol{S}^* = \boldsymbol{s}^*|\boldsymbol{S}^* \in \mathbb{S}^*_{\max}, \Theta^*) = \Pr(\boldsymbol{S}^* = \boldsymbol{s}^*|\boldsymbol{S}^* \in \mathbb{S}^*_{\max}). \tag{175}$$

- Isoprobability of sampling: for any $\boldsymbol{s}^{(1)*}, \boldsymbol{s}^{(2)*} \in \mathbb{S}^*_{\max}$,

$$\Pr(\boldsymbol{S}^* = \boldsymbol{s}^{(1)*}|\boldsymbol{S}^* \in \mathbb{S}^*_{\max}) = \Pr(\boldsymbol{S}^* = \boldsymbol{s}^{(2)*}|\boldsymbol{S}^* \in \mathbb{S}^*_{\max}). \tag{176}$$

Unlike X_{1j}'s, $S^*_{\cdot j}$'s are not independent of each other. Instead, Equation 176 leads to conditional independence (Equation 336). Note that, since $n^* > n$, it follows that $|\mathbb{S}^*_{\max}| \geq n + 1 \geq 2$ (see Equations 192 and 194).

Note that, for any $\boldsymbol{s}^* \in \mathbb{S}^*_{\max}$ and the generic sampling space, $\mathbb{S}^* \subseteq \mathbb{S}^*_{\max}$, when $\boldsymbol{s}^* \in \mathbb{S}^* \neq \emptyset$,

$$\Pr(\boldsymbol{S}^* = \boldsymbol{s}^*|\boldsymbol{S}^* \in \mathbb{S}^*) = \frac{\Pr(\boldsymbol{S}^* = \boldsymbol{s}^*|\boldsymbol{S}^* \in \mathbb{S}^*_{\max})}{\Pr(\boldsymbol{S}^* \in \mathbb{S}^*|\boldsymbol{S}^* \in \mathbb{S}^*_{\max})}, \tag{177}$$

and, when $\boldsymbol{s}^* \notin \mathbb{S}^*$,

$$\Pr(\boldsymbol{S}^* = \boldsymbol{s}^*|\boldsymbol{S}^* \in \mathbb{S}^*) = 0.$$

By abusing notation, when $\mathbb{S}^* = \emptyset$, I define

$$\Pr(\boldsymbol{S}^* = \boldsymbol{s}^*|\boldsymbol{S}^* \in \mathbb{S}^*) = 0. \tag{178}$$

Assumption of Potential Outcome and Response. For $\boldsymbol{K}^*_G \in \mathbb{U}^*$, like Condition 1, we define the following condition:

Condition 1* (Matched Attrition: SP).

$$\boldsymbol{k}_G^{A*} = \boldsymbol{k}_{G,-i}^{A*}.$$

Under Condition 1*, for any $\boldsymbol{s}^* \in \mathbb{S}^*_{\max}$, Condition 1 holds.

While the above assumptions of random sampling are mandatory, the following five assumptions of potential response (and outcome) are optional; we may invoke one or more of them in some lemmas and propositions below.

When $\boldsymbol{K}_G^* = \boldsymbol{K}_F^*$, Condition 1* always holds and $\mathbb{S}^*_{\text{def}}(\hat{\tau}_G) = \mathbb{S}_{\max}$ (which we will explain shortly) is equivalent to

Assumption 1* (No Attrition: SP).

$$\boldsymbol{r}^{T*} = \boldsymbol{r}^{C*} = \boldsymbol{1}^*.$$

When $\boldsymbol{K}_G^* = \boldsymbol{K}_U^*$, Condition 1* is equivalent to

Assumption 2* (Unitwise Matched Attrition: SP).

$$\boldsymbol{r}^{T*} = \boldsymbol{r}_{-i}^{T*}$$
$$\boldsymbol{r}^{C*} = \boldsymbol{r}_{-i}^{C*}.$$

When $\boldsymbol{K}_G^* = \boldsymbol{K}_P^*$, Condition 1* is equivalent to

Assumption 3* (Pairwise Matched Attrition: SP).

$$\boldsymbol{r}^{T*}\boldsymbol{r}_{-i}^{C*} = \boldsymbol{r}_{-i}^{T*}\boldsymbol{r}^{C*}.$$

Like Condition 2, we define the following condition as well:

Condition 2* (Ignorable Attrition: SP). *For any $y^* \in \mathbb{Y}^*$,*

$$P(\boldsymbol{k}_G^{A*} = 1 | \boldsymbol{y}^{A*} = y^*) = P(\boldsymbol{k}_G^{A*} = 1).$$

Unlike Condition 1*, under Condition 2*, it is not necessarily true that, for any $\boldsymbol{s}^* \in \mathbb{S}^*_{\max}$, Condition 2 holds.

When $\boldsymbol{K}_G^* = \boldsymbol{K}_F^*$, Condition 2* always holds.

When $\boldsymbol{K}_G^* = \boldsymbol{K}_U^*$, Condition 2* is equivalent to

Assumption 4* (Unitwise Ignorable Attrition: SP). *For any $y^{T*} \in \mathbb{Y}^{T*}$ and $y^{C*} \in \mathbb{Y}^{C*}$,*

$$P^*(\boldsymbol{r}^{T*} = 1 | \boldsymbol{y}^{T*} = y^{T*}) = P^*(\boldsymbol{r}^{T*} = 1)$$
$$P^*(\boldsymbol{r}^{C*} = 1 | \boldsymbol{y}^{C*} = y^{C*}) = P^*(\boldsymbol{r}^{C*} = 1)$$

When $\boldsymbol{K}_G^* = \boldsymbol{K}_P^*$, Condition 2* is equivalent to

Assumption 5* (Pairwise Ignorable Attrition: SP). *For any $y^{T*} \in \mathbb{Y}^{T*}$ and $y^{C*} \in \mathbb{Y}^{C*}$,*

$$\begin{aligned} & P^*(\boldsymbol{r}^{T*}\boldsymbol{r}_{-i}^{C*} = 1 | \boldsymbol{y}^{T*} = y^{T*}) = P^*(\boldsymbol{r}^{T*}\boldsymbol{r}_{-i}^{C*} = 1) \\ = \; & P^*(\boldsymbol{r}_{-i}^{T*}\boldsymbol{r}^{C*} = 1 | \boldsymbol{y}^{C*} = y^{C*}) = P^*(\boldsymbol{r}_{-i}^{T*}\boldsymbol{r}^{C*} = 1) \end{aligned}$$

On the one hand, under Assumption 1*, 2*, and 3*, for any sample $\boldsymbol{s}^* \in \mathbb{S}^*_{\max}$, Assumption 1, 2, and 3 hold, respectively. On the other hand, under Assumptions 4* or 5*, it is *not* guaranteed that, for any sample $\boldsymbol{s}^* \in \mathbb{S}^*_{\max}$, Assumption 4 and/or 5 hold. In addition, when the outcome is a continuous variable, it will be less likely that the ignorable attrition assumptions hold. Recall that the response is a binary variable. In these senses, one may say that the ignorable attrition assumptions are stronger than the matched attrition assumptions.

Number of Pairs. For a while, we suppose that $\boldsymbol{K}_G^* \in \mathbb{U}^*$ satisfies Condition 1*. Like Equation 31, we can define

$$n_G^{A*} \equiv \sum_{j^*} k_{G,\cdot j^*}^{A*}. \tag{179}$$

For $g \in \{0,1\}$, define

$$\boldsymbol{k}_G^{A*}(g) \equiv \begin{cases} \boldsymbol{k}_G^{A*} & \text{if } g = 1 \\ \mathbf{1} - \boldsymbol{k}_G^{A*} & \text{if } g = 0, \end{cases}$$

denote the set of pairs where $k_{G,\cdot j^*}^{A*}(g) = 1$ by

$$\mathbb{J}_G^{A*}(g) \equiv \{j^* | k_{G,\cdot j^*}^{A*}(g) = 1\},$$

the number of super-population pairs in $\mathbb{J}_G^{A*}(g)$ (namely, $|\mathbb{J}_G^{A*}(g)|$) by

$$\begin{aligned} n_G^{A*}(g) &\equiv \sum_{j^*} k_{G,\cdot j^*}^{A*}(g) \\ &= \begin{cases} n_G^{A*} & \text{if } g = 1 \\ n^* - n_G^{A*} & \text{if } g = 0, \end{cases} \quad (\because \text{Equation 179}) \end{aligned} \tag{180}$$

and the number of finite-sample pairs from $\mathbb{J}_G^{A*}(g)$ by

$$\begin{aligned} n_G^{A}(g) &\equiv \sum_{j^*} S_{\cdot j^*}^* k_{G,\cdot j^*}^{A*}(g) \\ &= \sum_{j} k_{G,\cdot j}^{A}(g) \quad (\because \text{Lemma 19 (1)}) \\ &= \begin{cases} n_G^{A} & \text{if } g = 1 \\ n - n_G^{A} & \text{if } g = 0, \end{cases} \quad (\because \text{Condition 1 holds, Equation 31}) \end{aligned}$$

which is equal to the number of finite-sample pairs *in* $\mathbb{J}_G^{A}(g)$ (namely, $|\mathbb{J}_G^{A}(g)|$).

Similarly, for $(g^T, g^C) \in \{0,1\}^2 \equiv \{(0,0),(0,1),(1,0),(1,1)\}$, define

$$\boldsymbol{k}_G^{*}(g^T, g^C) \equiv \prod_{A} \boldsymbol{k}_G^{A*}(g^A),$$

denote the set of pairs where $k_{G,\cdot j^*}^{A*}(g^T, g^C) = 1$ by

$$\mathbb{J}_G^{*}(g^T, g^C) \equiv \{j^* | k_{G,\cdot j^*}^{*}(g^T, g^C) = 1\},$$

the number of super-population pairs in $\mathbb{J}_G^{*}(g^T, g^C)$ (namely, $|\mathbb{J}_G^{*}(g^T, g^C)|$) by

$$n_G^{*}(g^T, g^C) \equiv \sum_{j^*} k_{G,\cdot j^*}^{*}(g^T, g^C), \tag{181}$$

and the number of finite-sample pairs from $\mathbb{J}_G^{*}(g^T, g^C)$ by

$$\begin{aligned} n_G(g^T, g^C) &\equiv \sum_{j^*} S_{\cdot j^*}^* k_{G,\cdot j^*}^{*}(g^T, g^C) \\ &= \sum_{j} k_{G,\cdot j}(g^T, g^C). \end{aligned} \tag{182}$$

I also denote

$$\begin{aligned} \boldsymbol{k}_G^{A*}(g^A, g^{-A}) &\equiv \boldsymbol{k}_G^{A*}(g^A) \boldsymbol{k}_G^{-A*}(g^{-A}) \\ &= \boldsymbol{k}_G^{*}(g^T, g^C), \end{aligned}$$

and $\mathbb{J}_G^{A*}(g^A, g^{-A})$, $n_G^{A*}(g^A, g^{-A})$, and $n_G^A(g^A, g^{-A})$ accordingly. Specifically, for instance,

$$\begin{aligned}
n_G^{A*}(1,1) &= \sum_{j^*} k_{G,\cdot j^*}^{T*} k_{G,\cdot j^*}^{C*} \\
&= n_G^{TC*} \quad (\because \text{Equation 109}) \\
n_G^{A*}(1,0) &= \sum_{j^*} k_{G,\cdot j^*}^{A*}(1 - k_{G,\cdot j^*}^{-A*}) \\
&= \sum_{j^*} (k_{G,\cdot j^*}^{A*} - k_{G,\cdot j^*}^{A*} k_{G,\cdot j^*}^{-A*}) \\
&= \sum_{j^*} k_{G,\cdot j^*}^{A*} - \sum_{j^*} k_{G,\cdot j^*}^{T*} k_{G,\cdot j^*}^{C*} \\
&= n_G^{A*} - n_G^{TC*} \quad (\because \text{Equations 109 and 179}) \\
n_G^{A*}(0,1) &= n_G^{-A*}(1,0) \\
&= n_G^{-A*} - n_G^{TC*} \\
n_G^{A*}(0,0) &= \sum_{j^*} (1 - k_{G,\cdot j^*}^{T*})(1 - k_{G,\cdot j^*}^{C*}) \\
&= \sum_{j^*} 1 - \sum_{j^*} k_{G,\cdot j^*}^{T*} - \sum_{j^*} k_{G,\cdot j^*}^{C*} + \sum_{j^*} k_{G,\cdot j^*}^{T*} k_{G,\cdot j^*}^{C*} \\
&= n^* - n_G^{T*} - n_G^{C*} + n_G^{TC*}
\end{aligned} \tag{183}$$

For any $g \in \{0,1\}$, thanks to $k_{G,\cdot j}^A(g) \in \{0,1\}$ and Equation 180, it follows that $0 \le n_G^A(g) \le n$. For any $(g^T, g^C) \in \{0,1\}^2$, thanks to $k_{G,\cdot j}(g^T, g^C) \in \{0,1\}$ and Equation 182, it follows that $0 \le n_G(g^T, g^C) \le n$. Therefore,

$$\max(0, n_G^T - n, n_G^C - n, n_G^T + n_G^C - n) \le n_G^{TC} \le \min(n, n_G^T, n_G^C, n_G^T + n_G^C)$$
$$\therefore \max(n_G^T + n_G^C - n, 0) \le n_G^{TC} \le \min(n_G^T, n_G^C).$$

Accordingly, define the maximum set of the triplet numbers of observed units as

$$\mathbb{N}_{\max} \equiv \{(n^T, n^C, n^{TC}) | n^T, n^C, n^{TC} \in \{0,1,\ldots,n\}, \max(n^T + n^C - n, 0) \le n^{TC} \le \min(n^T, n^C)\}, \tag{184}$$

and, for $\bar{\mathbf{n}} \in \mathbb{N}_{\max}$, where there are bar and no subscript G, and $g \in \{0,1\}$, define

$$\begin{aligned}
\bar{n}^A(g) &\equiv \begin{cases} \bar{n}^A & \text{if } g = 1 \\ n - \bar{n}^A & \text{if } g = 0 \end{cases} \\
\bar{n}(g^T, g^C) &\equiv \begin{cases} \bar{n}^{TC} & \text{if } (g^T, g^C) = (1,1) \\ \bar{n}^T - \bar{n}^{TC} & \text{if } (g^T, g^C) = (1,0) \\ \bar{n}^C - \bar{n}^{TC} & \text{if } (g^T, g^C) = (0,1) \\ n - \bar{n}^T - \bar{n}^C + \bar{n}^{TC} & \text{if } (g^T, g^C) = (0,0) \end{cases}
\end{aligned} \tag{185}$$

I summarize difference among these notations of the number of pairs:

- $\bar{n}^A(g)$ and $\bar{n}(g^T, g^C)$ are functions of $\bar{\mathbf{n}} \in \mathbb{N}_{\max}$ and constant irrespective of $\boldsymbol{K}_G^*$, $\boldsymbol{S}^*$, and $\boldsymbol{X}$.
- n_G^{A*} and $n_G^*(g^T, g^C)$ are functions of $\boldsymbol{K}_G^*$ and constant irrespective of $\boldsymbol{S}^*$ and $\boldsymbol{X}$.
- n_G^A and $n(g^T, g^C)$ are functions of $\boldsymbol{K}_G^*$ and $\boldsymbol{S}^*$ and constant irrespective of $\boldsymbol{X} = \boldsymbol{X}^{(*)}(\boldsymbol{s}^*)$ given $\boldsymbol{S}^* = \boldsymbol{s}^*$.

For $(g^{TC}, g^{-TC}) \in \{0,1\}^2$, define

$$\boldsymbol{k}_G^{TC*}(g^{TC}, g^{-TC}) \equiv \begin{cases} \boldsymbol{k}_G^{TC*} & \text{if } (g^{TC}, g^{-TC}) = (1,1) \\ \mathbf{1} - \boldsymbol{k}_G^{TC*} & \text{if } (g^{TC}, g^{-TC}) = (0,1) \\ \mathbf{0} & \text{if } (g^{TC}, g^{-TC}) = (1,0) \text{ or } (g^{TC}, g^{-TC}) = (0,0) \end{cases} \tag{186}$$

and define $n_G^{TC*}(g^{TC}, g^{-TC})$ and $n_G^{TC}(g^{TC}, g^{-TC})$ accordingly. Specifically, for instance,

$$\begin{aligned} n_G^{TC*}(g^{TC}, g^{-TC}) &= \sum_{j^*} \kappa_{G,\cdot j^*}^{TC*}(g^{TC}, g^{-TC}) \\ &= \begin{cases} n_G^{TC*} & \text{if } (g^{TC}, g^{-TC}) = (1,1) \\ n^* - n_G^{TC*} & \text{if } (g^{TC}, g^{-TC}) = (0,1) \\ 0 & \text{if } (g^{TC}, g^{-TC}) = (1,0) \text{ or } (g^{TC}, g^{-TC}) = (0,0) \end{cases} \end{aligned} \tag{187}$$

Denote the triplet numbers of finite-sample pairs as

$$\mathbf{n}_G \equiv \mathbf{n}(\boldsymbol{K}_G^*) \equiv (n_G^T, n_G^C, n_G^{TC}).$$

Specifically, according to Lemma 13 and Equation 164,

$$\mathbf{n}_F \equiv \mathbf{n}(\boldsymbol{K}_F^*) = (n, n, n), \tag{188}$$

under Assumption 3*

$$\mathbf{n}_P \equiv \mathbf{n}(\boldsymbol{K}_P^*) = (n_P, n_P, n_P), \tag{189}$$

and, under Assumption 2*

$$\mathbf{n}_U \equiv \mathbf{n}(\boldsymbol{K}_U^*) = (n_U^T, n_U^C, n_P). \tag{190}$$

For $\mathbb{S}^* \subseteq \mathbb{S}^*_{\max}$, let $\mathbb{S}^*(\cdot)$ denote the subset of $\mathbb{S}^*$ conditioned that the argument holds for any $\boldsymbol{X} \in \mathbb{X}_{\max}$. For instance, for $\bar{\mathbf{n}} \in \mathbb{N}_{\max}$,

$$\mathbb{S}^*(\mathbf{n}_G = \bar{\mathbf{n}}) \equiv \{\boldsymbol{s}^* | \boldsymbol{s}^* \in \mathbb{S}^*, \mathbf{n}_G = \bar{\mathbf{n}}\}. \tag{191}$$

For non-negative integers q^* and q, when $q^* \geq q$, denote

$${}_{q^*}C_q \equiv \frac{q^*!}{q!(q^*-q)!} \tag{192}$$

and, otherwise, by abusing notation, ${}_{q^*}C_q = 0$.

Lemma 21 (Isoprobability of Sampling). *(1) For* $\mathbb{S}^* \subseteq \mathbb{S}^*_{\max}, \mathbb{S}^* \neq \emptyset, \boldsymbol{s}^* \in \mathbb{S}^*$,

$$\Pr(\boldsymbol{S}^* = \boldsymbol{s}^* | \boldsymbol{S}^* \in \mathbb{S}^*) = \frac{1}{|\mathbb{S}^*|}.$$

(2) For $\boldsymbol{s}^* \in \mathbb{S}^*_{\max}$,

$$\Pr(\boldsymbol{S}^* = \boldsymbol{s}^* | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}) = \frac{n!(n^*-n)!}{n^*!}.$$

Proof. (1) Since $\boldsymbol{s}^* \in \mathbb{S}^*$, it follows $\mathbb{S}^* \neq \emptyset$. When $|\mathbb{S}^*| = 1$ and, thus, $\mathbb{S}^* = \{\boldsymbol{s}^*\}$, the desired result immediately follows. Suppose that $|\mathbb{S}^*| \geq 2$ and $\boldsymbol{s}^{(1)*}, \boldsymbol{s}^{(2)*} \in \mathbb{S}^*$. It follows

$$\begin{aligned} \Pr(\boldsymbol{S}^* = \boldsymbol{s}^{(1)*} | \boldsymbol{S}^* \in \mathbb{S}^*) &= \frac{\Pr(\boldsymbol{S}^* = \boldsymbol{s}^{(1)*} | \boldsymbol{S}^* \in \mathbb{S}^*_{\max})}{\Pr(\boldsymbol{S}^* \in \mathbb{S}^* | \boldsymbol{S}^* \in \mathbb{S}^*_{\max})} \quad (\because \text{Equation } 177, \boldsymbol{s}^{(1)*} \in \mathbb{S}^* \neq \emptyset) \\ &= \frac{\Pr(\boldsymbol{S}^* = \boldsymbol{s}^{(2)*} | \boldsymbol{S}^* \in \mathbb{S}^*_{\max})}{\Pr(\boldsymbol{S}^* \in \mathbb{S}^* | \boldsymbol{S}^* \in \mathbb{S}^*_{\max})} \quad (\because \text{Equation } 176, \boldsymbol{s}^{(1)*}, \boldsymbol{s}^{(2)*} \in \mathbb{S}^* \subseteq \mathbb{S}^*_{\max}) \\ &= \Pr(\boldsymbol{S}^* = \boldsymbol{s}^{(2)*} | \boldsymbol{S}^* \in \mathbb{S}^*) \quad (\because \text{Equation } 177, \boldsymbol{s}^{(2)*} \in \mathbb{S}^* \neq \emptyset) \end{aligned} \tag{193}$$

According to the axiom of probability and Equation 193,

$$1 = \sum_{\boldsymbol{s}^* \in \mathbb{S}^*} \Pr(\boldsymbol{S}^* = \boldsymbol{s}^* | \boldsymbol{S}^* \in \mathbb{S}^*) \quad (\because \text{the axiom of probability})$$
$$= |\mathbb{S}^*| \times \Pr(\boldsymbol{S}^* = \boldsymbol{s}^* | \boldsymbol{S}^* \in \mathbb{S}^*) \quad (\text{for any } \boldsymbol{s}^* \in \mathbb{S}^*, \because \text{Equation 193})$$
$$\therefore \Pr(\boldsymbol{S}^* = \boldsymbol{s}^* | \boldsymbol{S}^* \in \mathbb{S}^*) = \frac{1}{|\mathbb{S}^*|} \quad (\because |\mathbb{S}^*| \geq 2)$$

(2) Since we sample n pairs from $\mathbb{J}^* = \{1, 2, \ldots, n^*\}$, whose size is n^*, the number of sets of such n pairs and, thus, the number of values $\boldsymbol{s}^* \in \mathbb{S}^*_{\max}$ can take is combination of n sampled j^*'s such that $s^*_{\cdot j^*} = 1$, namely,

$$|\mathbb{S}^*_{\max}| = {}_{n^*}C_n. \tag{194}$$

Therefore,

$$\Pr(\boldsymbol{S}^* = \boldsymbol{s}^* | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}) = \frac{1}{|\mathbb{S}^*_{\max}|} \quad (\because \text{Lemma 21 (1)})$$
$$= \frac{n!(n^* - n)!}{n^*!} \quad (\because \text{Equations 192 and 194})$$

□

Lemma 22 (Conditional Probability of Sampling). *Suppose that $\boldsymbol{K}^*_G \in \mathbb{U}^*$ satisfies Condition 1*, $\bar{\mathbf{n}} \in \mathbb{N}_{\max}$, and $\mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}}) \neq \emptyset$.*

(1) Given $(g^T, g^C) \in \{0, 1\}^2$, for any $j^* \in \mathbb{J}^*_G(g^T, g^C)$,

$$\Pr\{S^*_{\cdot j^*} = 1 | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}})\} = \frac{\bar{n}(g^T, g^C)}{n^*_G(g^T, g^C)}$$

(2) Given $(g^{T(h)}, g^{C(h)}) \in \{0, 1\}^2$ for each $h \in \{1, 2\}$, for any $j^*(h) \in \mathbb{J}^*_G(g^{T(h)}, g^{C(h)})$ for each $h \in \{1, 2\}$, when $j^*(1) \neq j^*(2)$,

$$\Pr\left\{ \prod_{h=1}^2 S^*_{\cdot j^*(h)} = 1 \middle| \boldsymbol{S}^* \in \mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}})\right\}$$
$$= \begin{cases} \dfrac{\bar{n}(g^T, g^C)\{\bar{n}(g^T, g^C) - 1\}}{n^*_G(g^T, g^C)\{n^*_G(g^T, g^C) - 1\}} & \text{if } (g^{T(1)}, g^{C(1)}) = (g^{T(2)}, g^{C(2)}) \equiv (g^T, g^C) \\ \displaystyle\prod_{h=1}^2 \frac{\bar{n}(g^{T(h)}, g^{C(h)})}{n^*_G(g^{T(h)}, g^{C(h)})} & \text{if } (g^{T(1)}, g^{C(1)}) \neq (g^{T(2)}, g^{C(2)}) \end{cases}$$

Proof. Suppose $\mathbf{n}_G = \bar{\mathbf{n}}$. Since $\mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}}) \neq \emptyset$, it follows that $n^*_G(g^{T\prime}, g^{C\prime}) \geq \bar{n}(g^{T\prime}, g^{C\prime})$ for any $(g^{T\prime}, g^{C\prime}) \in \{0, 1\}^2$. For, otherwise, $n^*_G(g^{T\prime}, g^{C\prime}) < \bar{n}(g^{T\prime}, g^{C\prime})$ for some $(g^{T\prime}, g^{C\prime}) \in \{0, 1\}^2$ and we cannot sample $\bar{n}(g^{T\prime}, g^{C\prime})$ pairs from $\mathbb{J}^*_G(g^{T\prime}, g^{C\prime})$, which implies $\mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}}) = \emptyset$, a contradiction.

We sample $\bar{n}(g^{T\prime}, g^{C\prime})$ pairs from $\mathbb{J}^*_G(g^{T\prime}, g^{C\prime})$, whose size is $n^*_G(g^{T\prime}, g^{C\prime})$. The number of sets of such $\bar{n}(g^{T\prime}, g^{C\prime})$ pairs (including $\emptyset$ when $\bar{n}(g^{T\prime}, g^{C\prime}) = 0$) is ${}_{n^*_G(g^{T\prime}, g^{C\prime})}C_{\bar{n}(g^{T\prime}, g^{C\prime})}$ (note ${}_{q^*}C_0 = 1$). Thus, the number of values $\boldsymbol{s}^* \in \mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}})$ can take is

$$|\mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}})| = \prod_{(g^{T\prime}, g^{C\prime}) \in \{0,1\}^2} {}_{n^*_G(g^{T\prime}, g^{C\prime})}C_{\bar{n}(g^{T\prime}, g^{C\prime})} \tag{195}$$

because $\mathbb{J}^*_G(g^{T\prime}, g^{C\prime}) \cap \mathbb{J}^*_G(g^{T\prime\prime}, g^{C\prime\prime}) = \emptyset$ when $(g^{T\prime}, g^{C\prime}) \neq (g^{T\prime\prime}, g^{C\prime\prime})$.

(1) Since $j^* \in \mathbb{J}^*_G(g^T, g^C)$, $n^*_G(g^T, g^C) \geq 1$. When $\bar{n}(g^T, g^C) < 1$, we cannot sample any pair from $\mathbb{J}^*_G(g^T, g^C)$. Thus, both sides of the equation in Lemma 22 (1) are equal to each other (zero). Below, we suppose $\bar{n}(g^T, g^C) \geq 1$.

When $S^*_{\cdot j^*} = 1$, we sample pair j^* from $\mathbb{J}^*_G(g^T, g^C)$ and $\bar{n}(g^T, g^C) - 1$ pairs from $\mathbb{J}^*_G(g^T, g^C) \setminus \{j^*\}$, whose size is $n^*_G(g^T, g^C) - 1$. The number of sets of such $\bar{n}(g^T, g^C) - 1$ pairs is ${}_{\{n^*_G(g^T, g^C) - 1\}}C_{\bar{n}(g^T, g^C) - 1}$.

We also sample, for every $(g^{T\prime}, g^{C\prime}) \in \{0,1\}^2 \setminus \{(g^T, g^C)\}$, $\bar{n}(g^{T\prime}, g^{C\prime})$ pairs from $\mathbb{J}^*_G(g^{T\prime}, g^{C\prime})$. Thus, the number of values $\boldsymbol{s}^* \in \mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}}, S^*_{\cdot j^*} = 1)$ can take is

$$|\mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}}, S^*_{\cdot j^*} = 1)| = {}_{n^*_G(g^T,g^C)-1}C_{\bar{n}(g^T,g^C)-1} \prod_{(g^{T\prime},g^{C\prime})\neq(g^T,g^C)} {}_{n^*_G(g^{T\prime},g^{C\prime})}C_{\bar{n}(g^{T\prime},g^{C\prime})}. \quad (196)$$

Therefore,

$$\begin{aligned}
\Pr\{S^*_{\cdot j^*} = 1 | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}})\} &= \Pr\{\boldsymbol{S}^* \in \mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}}, S^*_{\cdot j^*} = 1) | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}})\} \\
&= \frac{\Pr(\boldsymbol{S}^* = \boldsymbol{s}^* | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}) \times |\mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}}, S^*_{\cdot j^*} = 1)|}{\Pr(\boldsymbol{S}^* = \boldsymbol{s}^* | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}) \times |\mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}})|} \\
&\quad (\because \mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}}, S^*_{\cdot j^*} = 1) \subset \mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}}), \text{Equation 176}) \\
&= \frac{{}_{n^*_G(g^T,g^C)-1}C_{\bar{n}(g^T,g^C)-1}}{{}_{n^*_G(g^T,g^C)}C_{\bar{n}(g^T,g^C)}} \quad (\because \text{Equations 195 and 196}) \\
&= \frac{\bar{n}(g^T, g^C)}{n^*_G(g^T, g^C)} \quad (\because \text{Equation 192})
\end{aligned}$$

(2) In the case of $(g^{T(1)}, g^{C(1)}) = (g^{T(2)}, g^{C(2)}) \equiv (g^T, g^C)$:

Since $j^*(h) \in \mathbb{J}^*_G(g^T, g^C)$ for every $h \in \{1,2\}$ and $j^*(1) \neq j^*(2)$, $n^*_G(g^T, g^C) \geq 2$. When $\bar{n}(g^T, g^C) < 2$, we cannot sample two pairs from $\mathbb{J}^*_G(g^T, g^C)$. Thus, both sides of the equation in Lemma 22 (2) are equal to each other (zero). Below, we suppose $\bar{n}(g^T, g^C) \geq 2$.

When $\prod_{h=1}^2 S^*_{\cdot j^*(h)} = 1$, we sample pairs $j^*(1)$ and $j^*(2)$ from $\mathbb{J}^*_G(g^T, g^C)$ and $\bar{n}(g^T, g^C) - 2$ pairs from $\mathbb{J}^*_G(g^T, g^C) \setminus \{j^*(1), j^*(2)\}$, whose size is $n^*_G(g^T, g^C) - 2$ ($\because j^*(1) \neq j^*(2)$). The number of sets of such $\bar{n}(g^T, g^C) - 2$ pairs is ${}_{\{n^*_G(g^T,g^C)-2\}}C_{\bar{n}(g^T,g^C)-2}$. We also sample, for every $(g^{T\prime}, g^{C\prime}) \in \{0,1\}^2 \setminus \{(g^T, g^C)\}$, $\bar{n}(g^{T\prime}, g^{C\prime})$ pairs from $\mathbb{J}^*_G(g^{T\prime}, g^{C\prime})$. Thus, the number of values $\boldsymbol{s}^* \in \mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}}, \prod_{h=1}^2 S^*_{\cdot j^*(h)} = 1)$ can take is

$$\left|\mathbb{S}^*_{\max}\left(\mathbf{n}_G = \bar{\mathbf{n}}, \prod_{h=1}^2 S^*_{\cdot j^*(h)} = 1\right)\right| = {}_{n^*_G(g^T,g^C)-2}C_{\bar{n}(g^T,g^C)-2} \prod_{(g^{T\prime},g^{C\prime})\neq(g^T,g^C)} {}_{n^*_G(g^{T\prime},g^{C\prime})}C_{\bar{n}(g^{T\prime},g^{C\prime})}. \quad (197)$$

Therefore,

$$\begin{aligned}
&\Pr\left\{\prod_{h=1}^2 S^*_{\cdot j^*(h)} = 1 \middle| \boldsymbol{S}^* \in \mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}})\right\} \\
&= \frac{|\mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}}, \prod_{h=1}^2 S^*_{\cdot j^*(h)} = 1)|}{|\mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}})|} \\
&\quad \left(\because \mathbb{S}^*_{\max}\left(\mathbf{n}_G = \bar{\mathbf{n}}, \prod_{h=1}^2 S^*_{\cdot j^*(h)} = 1\right) \in \mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}}), \text{Equation 176}\right) \\
&= \frac{{}_{n^*_G(g^T,g^C)-2}C_{\bar{n}(g^T,g^C)-2}}{{}_{n^*_G(g^T,g^C)}C_{\bar{n}(g^T,g^C)}} \quad (\because \text{Equations 195 and 197}) \\
&= \frac{\bar{n}(g^T, g^C)\{\bar{n}(g^T, g^C) - 1\}}{n^*_G(g^T, g^C)\{n^*_G(g^T, g^C) - 1\}} \quad (\because \text{Equation 192})
\end{aligned}$$

In the case of $(g^{T(1)}, g^{C(1)}) \neq (g^{T(2)}, g^{C(2)})$:

Since $j^*(h) \in \mathbb{J}^*_G(g^{T(h)}, g^{C(h)})$ for every $h \in \{1,2\}$ and $\bigcap_{h=1}^2 \mathbb{J}^*_G(g^{T(h)}, g^{C(h)}) = \emptyset$ ($\because (g^{T(1)}, g^{C(1)}) \neq (g^{T(2)}, g^{C(2)})$), it follows that $n^*_G(g^T, g^C) \geq 1$. When $\bar{n}(g^{T(h')}, g^{C(h')}) < 1$ for some $h' \in \{1,2\}$, we cannot sample any pair from $\mathbb{J}^*_G(g^{T(h')}, g^{C(h')})$. Thus, both sides of the equation in Lemma 22 (2) are equal to each other (zero). Below, we suppose $\bar{n}(g^{T(h)}, g^{C(h)}) \geq 1$ for every $h \in \{1,2\}$.

When $\prod_{h=1}^{2} S^*_{\cdot j^*(h)} = 1$, for every $h \in \{1,2\}$, we sample pair $j^*(h)$ from $\mathbb{J}^*_G(g^{T(h)}, g^{C(h)})$ and $\bar{n}(g^{T(h)}, g^{C(h)}) - 1$ pairs from $\mathbb{J}^*_G(g^{T(h)}, g^{C(h)}) \setminus \{j^*(h)\}$, whose size is $n^*_G(g^{T(h)}, g^{C(h)}) - 1$ ($\because \bigcap_{h=1}^{2} \mathbb{J}^*_G(g^{T(h)}, g^{C(h)}) = \emptyset$). The number of sets of such $\bar{n}(g^{T(h)}, g^{C(h)}) - 1$ pairs is ${}_{\{n^*_G(g^{T(h)},g^{C(h)})-1\}}C_{\{\bar{n}(g^{T(h)},g^{C(h)})-1\}}$. We also sample, for every $(g^{T\prime}, g^{C\prime}) \in \{0,1\}^2 \backslash \{(g^{T(h)}, g^{C(h)}), h \in \{1,2\}\}$, $\bar{n}(g^{T\prime}, g^{C\prime})$ pairs from $\mathbb{J}^*_G(g^{T\prime}, g^{C\prime})$. Thus, the number of values $\boldsymbol{s}^* \in \mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}}, \prod_{h=1}^{2} S^*_{\cdot j^*(h)} = 1)$ can take is

$$\begin{aligned}
&\Big|\mathbb{S}^*_{\max}\Big(\mathbf{n}_G = \bar{\mathbf{n}}, \prod_{h=1}^{2} S^*_{\cdot j^*(h)} = 1\Big)\Big| \\
&= \Big[\prod_{h=1}^{2} {}_{\{n^*_G(g^{T(h)},g^{C(h)})-1\}}C_{\{\bar{n}(g^{T(h)},g^{C(h)})-1\}}\Big] \times \Big[\prod_{(g^{T\prime},g^{C\prime}) \notin \{(g^{T(h)},g^{C(h)}),h\in\{1,2\}\}} {}_{n^*_G(g^{T\prime},g^{C\prime})}C_{\bar{n}(g^{T\prime},g^{C\prime})}\Big].
\end{aligned} \tag{198}$$

Therefore,

$$\begin{aligned}
&\Pr\Big\{\prod_{h=1}^{2} S^*_{\cdot j^*(h)} = 1 \Big| \boldsymbol{S}^* \in \mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}})\Big\} \\
&= \prod_{h=1}^{2} \frac{{}_{\{n^*_G(g^{T(h)},g^{C(h)})-1\}}C_{\{\bar{n}(g^{T(h)},g^{C(h)})-1\}}}{{}_{n^*_G(g^{T(h)},g^{C(h)})}C_{\bar{n}(g^{T(h)},g^{C(h)})}} \quad (\because \text{Equations 192, 195, and 198}) \\
&= \prod_{h=1}^{2} \frac{\bar{n}(g^{T(h)}, g^{C(h)})}{n^*_G(g^{T(h)}, g^{C(h)})} \quad (\because \text{Equation 176})
\end{aligned}$$

□

Expectation. Let

$$\begin{aligned}
\Theta^{(*)}(\boldsymbol{S}^*) &\equiv (\boldsymbol{y}^{T(*)}(\boldsymbol{S}^*), \boldsymbol{y}^{T(*)}(\boldsymbol{S}^*), \boldsymbol{r}^{T(*)}(\boldsymbol{S}^*), \boldsymbol{r}^{T(*)}(\boldsymbol{S}^*)) \\
&= (\boldsymbol{y}^T, \boldsymbol{y}^C, \boldsymbol{r}^T, \boldsymbol{r}^C) \\
&\equiv \Theta
\end{aligned}$$

For the generic function of the super-population treatment assignment vector and the super-population sampling indicator vector,

$$f^*_{X,S}(\boldsymbol{X}^*, \boldsymbol{S}^*|\Theta^*) \equiv f_X\{\boldsymbol{X}^{(*)}(\boldsymbol{S}^*)|\Theta^{(*)}(\boldsymbol{S}^*)\},$$

define the well behaved sampling space as

$$\mathbb{S}^*_{\text{def}}(f^*_{X,S}) \equiv \{\boldsymbol{s}^* | \boldsymbol{s}^* \in \mathbb{S}^*_{\max}, \mathbb{X}_{\text{def}}[f_X\{\boldsymbol{X}^{(*)}(\boldsymbol{s}^*)|\Theta^{(*)}(\boldsymbol{s}^*)\}] = \mathbb{X}_{\max}\}.$$

For $f^*_{X,S}(\boldsymbol{X}^*, \boldsymbol{S}^*)$ and $\mathbb{S}^* \subseteq \mathbb{S}^*_{\text{def}}(f^*_{X,S})$, the super-population expectation operator is defined as

$$\mathbb{E}^*\{f^*_{X,S}(\boldsymbol{X}^*, \boldsymbol{S}^*|\Theta^*)|\mathbb{S}^*\} \equiv \mathbb{E}^*_S[\mathbb{E}_X\{f_{X,S}(\boldsymbol{X}^*, \boldsymbol{S}^*|\Theta^*)\}|\mathbb{S}^*], \tag{199}$$

where

$$\begin{aligned}
\mathbb{E}_X\{f_X(\boldsymbol{X})\} &\equiv \mathbb{E}\{f_X(\boldsymbol{X})\} \\
\therefore \mathbb{E}_X\{f_{X,S}(\boldsymbol{X}^*, \boldsymbol{s}^*|\Theta^*)\} &= \mathbb{E}[f_X\{\boldsymbol{X}^{(*)}(\boldsymbol{s}^*)|\Theta^{(*)}(\boldsymbol{s}^*)\}] \\
&= \sum_{\boldsymbol{x}\in\mathbb{X}_{\max}} \Pr\{\boldsymbol{X}^{(*)}(\boldsymbol{s}^*) = \boldsymbol{x}\} f_X\{\boldsymbol{x}|\Theta^{(*)}(\boldsymbol{s}^*)\} \\
&\equiv f^*_S(\boldsymbol{s}^*|\Theta^*) \\
\mathbb{E}^*_S\{f^*_S(\boldsymbol{S}^*|\Theta^*)|\mathbb{S}^*\} &\equiv \sum_{\boldsymbol{s}^*\in\mathbb{S}^*} \Pr(\boldsymbol{S}^* = \boldsymbol{s}^*|\boldsymbol{S}^* \in \mathbb{S}^*) f^*_S(\boldsymbol{s}^*|\Theta^*).
\end{aligned} \tag{200}$$

Note

$$\mathbb{E}^*\{f^*_{X,S}(\boldsymbol{X}^*,\boldsymbol{S}^*|\Theta^*)|\emptyset\} = 0 \quad (\because \text{Equations 178, 199, and 200}) \tag{201}$$

Suppose that $\boldsymbol{K}^*_G \in \mathbb{U}^*$ satisfies Condition 1^*, $\mathbb{S}^* \subseteq \mathbb{S}^*_{\max}$, and $\bar{\mathbf{n}} \in \mathbb{N}_{\max}$. Denote expectation conditional on the triplet numbers of pairs as

$$\mathbb{E}^*_{S|N}\{f^*_S(\boldsymbol{S}^*)|\mathbb{S}^*(\mathbf{n}_G = \bar{\mathbf{n}})\} \equiv \mathbb{E}^*\{f^*_S(\boldsymbol{S}^*)|\mathbb{S}^*(\mathbf{n}_G = \bar{\mathbf{n}})\} \tag{202}$$

and expectation over the triplet numbers of pairs as

$$\mathbb{E}^*_N\{f^*_N(\mathbf{n}_G)|\mathbb{S}^*\} \equiv \mathbb{E}^*\{f^*_N(\mathbf{n}_G)|\mathbb{S}^*\} \tag{203}$$

Lemma 23 (Arithmetic of Expectation: SP). *(1) When*

$$\mathbb{S}^* \subseteq \Big\{\mathbb{S}^*_{def}\Big(\sum_h f^{(h)*}_{X,S}\Big) \cap \bigcap_h \mathbb{S}^*_{def}(f^{(h)*}_{X,S})\Big\},$$

It follows

$$\mathbb{E}^*\Big\{\sum_h f^{(h)*}_{X,S}(\boldsymbol{X}^*,\boldsymbol{S}^*)\Big|\mathbb{S}^*\Big\} = \sum_h \mathbb{E}^*\{f^{(h)*}_{X,S}(\boldsymbol{X}^*,\boldsymbol{S}^*)|\mathbb{S}^*\}$$

(2) Suppose that $\boldsymbol{K}^*_G \in \mathbb{U}^*$ *satisfies Condition* 1^* *and* $\mathbb{S}^* \subseteq \mathbb{S}^*_{\max}$. *It follows*

$$\begin{aligned}\mathbb{E}^*_S\{f^*_S(\boldsymbol{S}^*)|\mathbb{S}^*\} &= \mathbb{E}^*_N[\mathbb{E}^*_{S|N}\{f^*_S(\boldsymbol{S}^*)|\mathbb{S}^*(\mathbf{n}_G)\}|\mathbb{S}^*] \\ &= \mathbb{E}^*_N\{f^*_N(\mathbf{n}_G)|\mathbb{S}^*\} \\ &= \sum_{\bar{\mathbf{n}}\in\mathbb{N}_{\max}} \Pr(\mathbf{n}_G = \bar{\mathbf{n}}|\boldsymbol{S}^* \in \mathbb{S}^*) f^*_N(\bar{\mathbf{n}}),\end{aligned}$$

where

$$f^*_N(\bar{\mathbf{n}}) \equiv \mathbb{E}^*\{f^*_S(\boldsymbol{S}^*)|\mathbb{S}^*(\mathbf{n}_G = \bar{\mathbf{n}})\}.$$

(3) When $\mathbb{S}^* \neq \emptyset$,

$$\mathbb{E}(q|\mathbb{S}^*) = q.$$

Proof. (1)

$$\begin{aligned}&\mathbb{E}^*\Big\{\sum_h f^{(h)*}_{X,S}(\boldsymbol{X}^*,\boldsymbol{S}^*)\Big|\mathbb{S}^*\Big\} \\ &= \sum_{\boldsymbol{s}^*\in\mathbb{S}^*} \Pr(\boldsymbol{S}^* = \boldsymbol{s}^*|\boldsymbol{S}^* \in \mathbb{S}^*) \sum_{\boldsymbol{x}\in\mathbb{X}_{\max}} \Pr\{\boldsymbol{X}^{(*)}(\boldsymbol{s}^*) = \boldsymbol{x}|\boldsymbol{X}^{(*)}(\boldsymbol{s}^*) \in \mathbb{X}_{\max}\} \sum_h f^{(h)}_X\{\boldsymbol{x}(\boldsymbol{s}^*)\} \\ &\quad (\because \text{Equation 200, } \mathbb{S}^* \subseteq \mathbb{S}^*_{\text{def}}\Big(\sum_h f^{(h)*}_{X,S}\Big)) \\ &= \sum_h \sum_{\boldsymbol{s}^*\in\mathbb{S}^*} \Pr(\boldsymbol{S}^* = \boldsymbol{s}^*|\boldsymbol{S}^* \in \mathbb{S}^*) \sum_{\boldsymbol{x}\in\mathbb{X}_{\max}} \Pr\{\boldsymbol{X}^{(*)}(\boldsymbol{s}^*) = \boldsymbol{x}|\boldsymbol{X}^{(*)}(\boldsymbol{s}^*) \in \mathbb{X}_{\max}\} f^{(h)}_X\{\boldsymbol{x}(\boldsymbol{s}^*)\} \\ &\quad (\because \mathbb{S}^* \subseteq \bigcap_h \mathbb{S}^*_{\text{def}}(f^{(h)*}_{X,S})) \\ &= \sum_h \mathbb{E}^*\{f^{(h)*}_{X,S}(\boldsymbol{X}^*,\boldsymbol{S}^*)|\mathbb{S}^*\} \quad (\because \text{Equation 200})\end{aligned}$$

(2)

$$
\begin{aligned}
&\mathbb{E}_S^*\{f_S^*(\boldsymbol{S}^*)|\mathbb{S}^*\} \\
&= \sum_{\boldsymbol{s}^*\in\mathbb{S}^*} \Pr(\boldsymbol{S}^* = \boldsymbol{s}^*|\boldsymbol{S}^* \in \mathbb{S}^*) f_S^*(\boldsymbol{s}^*) \quad (\because \text{Equation 200}) \\
&= \sum_{\bar{\mathbf{n}}\in\mathbb{N}_{\max}} \sum_{\boldsymbol{s}^*\in\mathbb{S}^*(\mathbf{n}_G=\bar{\mathbf{n}})} \Pr\{\boldsymbol{S}^* = \boldsymbol{s}^*, \mathbf{n}_G = \bar{\mathbf{n}}|\boldsymbol{S}^* \in \mathbb{S}^*\} f_S^*(\boldsymbol{s}^*) \\
&= \sum_{\bar{\mathbf{n}}\in\mathbb{N}_{\max}} \Pr(\mathbf{n}_G = \bar{\mathbf{n}}|\boldsymbol{S}^* \in \mathbb{S}^*) \sum_{\boldsymbol{s}^*\in\mathbb{S}^*(\mathbf{n}_G=\bar{\mathbf{n}})} \Pr\{\boldsymbol{S}^* = \boldsymbol{s}^*|\boldsymbol{S}^* \in \mathbb{S}^*(\mathbf{n}_G = \bar{\mathbf{n}})\} f_S^*(\boldsymbol{s}^*) \\
&= \sum_{\bar{\mathbf{n}}\in\mathbb{N}_{\max}} \Pr(\mathbf{n}_G = \bar{\mathbf{n}}|\boldsymbol{S}^* \in \mathbb{S}^*) \mathbb{E}_{S|N}^*\{f_S^*(\boldsymbol{S}^*)|\mathbb{S}^*(\mathbf{n}_G = \bar{\mathbf{n}})\} \quad (\because \text{Equations 200 and 202}) \\
&= \sum_{\bar{\mathbf{n}}\in\mathbb{N}_{\max}} \Pr(\mathbf{n}_G = \bar{\mathbf{n}}|\boldsymbol{S}^* \in \mathbb{S}^*) f_N^*(\bar{\mathbf{n}}) \\
&= \sum_{\bar{\mathbf{n}}\in\mathbb{N}_{\max}} \sum_{\boldsymbol{s}^*\in\mathbb{S}^*} \Pr(\boldsymbol{S}^* = \boldsymbol{s}^*, \mathbf{n}_G = \bar{\mathbf{n}}|\boldsymbol{S}^* \in \mathbb{S}^*) f_N^*(\bar{\mathbf{n}}) \\
&= \sum_{\boldsymbol{s}^*\in\mathbb{S}^*} \sum_{\bar{\mathbf{n}}\in\mathbb{N}_{\max}} \Pr(\mathbf{n}_G = \bar{\mathbf{n}}|\boldsymbol{S}^* \in \mathbb{S}^*, \boldsymbol{S}^* = \boldsymbol{s}^*) \Pr(\boldsymbol{S}^* = \boldsymbol{s}^*|\boldsymbol{S}^* \in \mathbb{S}^*) f_N^*(\bar{\mathbf{n}}) \\
&= \sum_{\boldsymbol{s}^*\in\mathbb{S}^*} \sum_{\bar{\mathbf{n}}\in\mathbb{N}_{\max}} \{I(\mathbf{n}_G = \bar{\mathbf{n}}) \cdot 1 \cdot \Pr(\boldsymbol{S}^* = \boldsymbol{s}^*|\boldsymbol{S}^* \in \mathbb{S}^*) f_N^*(\mathbf{n}_G) \\
&\quad + I(\mathbf{n}_G \neq \bar{\mathbf{n}}) \cdot 0 \cdot \Pr(\boldsymbol{S}^* = \boldsymbol{s}^*|\boldsymbol{S}^* \in \mathbb{S}^*) f_N^*(\bar{\mathbf{n}})\} \\
&= \sum_{\boldsymbol{s}^*\in\mathbb{S}^*} \Pr(\boldsymbol{S}^* = \boldsymbol{s}^*|\boldsymbol{S}^* \in \mathbb{S}^*) f_N^*(\mathbf{n}_G) \\
&= \mathbb{E}_N^*\{f_N^*(\mathbf{n}_G)|\mathbb{S}^*\} \quad (\because \text{Equations 200 and 203}) \\
&= \mathbb{E}_N^*[\mathbb{E}_{S|N}^*\{f_S^*(\boldsymbol{S}^*)|\mathbb{S}^*(\mathbf{n}_G)\}|\mathbb{S}^*]
\end{aligned}
$$

(3)

$$
\begin{aligned}
\mathbb{E}^*(q|\mathbb{S}^*) &= \sum_{\boldsymbol{s}^*\in\mathbb{S}^*} \Pr\{\boldsymbol{S}^* = \boldsymbol{s}^*|\boldsymbol{S}^* \in \mathbb{S}^*\} q \quad (\because \text{Equation 200}) \\
&= q \sum_{\boldsymbol{s}^*\in\mathbb{S}^*} \Pr\{\boldsymbol{S}^* = \boldsymbol{s}^*|\boldsymbol{S}^* \in \mathbb{S}^*\} \\
&= q \cdot 1 \quad (\because \mathbb{S}^* \neq \emptyset, \text{the axiom of the probability}) \\
&= q.
\end{aligned}
$$

□

For $m \in \{1, 2, \ldots, n\}$, denote

$$
\begin{aligned}
\mathbb{N}^m &\equiv \{\mathbf{n}|\mathbf{n} \in \mathbb{N}_{\max}, n^T, n^C \geq m\} \\
\mathbb{N}^{Am} &\equiv \{\mathbf{n}|\mathbf{n} \in \mathbb{N}_{\max}, n^A \geq m\}.
\end{aligned} \tag{204}
$$

It holds $\mathbb{N}^{Am} \supseteq \mathbb{N}^m$ and, when $A(1) \neq A(2)$, it holds $\bigcap_{h=1}^2 \mathbb{N}^{A(h)m} = \mathbb{N}^m$.

For $\bar{\mathbf{n}} \in \mathbb{N}^{A1}$, define mean of $\boldsymbol{Q}^*$ weighted by $\boldsymbol{k}_G^{A*}$ and adjusted by $\bar{\mathbf{n}}$ as

$$
E_N^*(\boldsymbol{Q}^*|\boldsymbol{k}_G^{A*}, \bar{\mathbf{n}}) \equiv \sum_{g^{-A}=0}^{1} \frac{\bar{n}^A(1, g^{-A})}{\bar{n}^A} E^*\{\boldsymbol{Q}^*|\boldsymbol{k}_G^{A*}(1, g^{-A})\} \tag{205}
$$

and when $\bar{n}^A = 0$,

$$
E_N^*(\boldsymbol{Q}^*|\boldsymbol{k}_G^{A*}, \bar{\mathbf{n}}) \equiv 0. \tag{206}
$$

Lemma 24 (Expectation of Mean). *Suppose that* $\boldsymbol{K}_G^* \in \mathbb{U}^*$ *satisfies Condition* 1^*, $\boldsymbol{\beta}^* \in \mathbb{B}^*, \bar{\mathbf{n}} \in \mathbb{N} \subseteq \mathbb{N}_{\max}, \mathbb{S}_N^* \equiv \mathbb{S}_{\max}^*(\mathbf{n}_G = \bar{\mathbf{n}}), \mathbb{S}_N^* \neq \emptyset$, *and* $\mathbb{S}_G^* \equiv \mathbb{S}_{\max}^*(\mathbf{n}_G \in \mathbb{N})$.

(1) For any $(g^T, g^C) \in \{0,1\}^2$, *when* $\bar{n}(g^T, g^C) \geq 1$ *or* $n_G^*(g^T, g^C) = 0$,

$$\mathbb{E}_{S|N}^*[E^*\{\boldsymbol{\beta}^*|\boldsymbol{S}^*\boldsymbol{k}_G^*(g^T, g^C)\}|\mathbb{S}_N^*] = E^*\{\boldsymbol{\beta}^*|\boldsymbol{k}_G^*(g^T, g^C)\}.$$

(2)

$$\mathbb{E}_{S|N}^*\{E^*(\boldsymbol{\beta}^*|\boldsymbol{S}^*\boldsymbol{k}_G^{A*})|\mathbb{S}_N^*\} = E_N^*(\boldsymbol{\beta}^*|\boldsymbol{k}_G^{A*}, \bar{\mathbf{n}}).$$

(3)

$$\mathbb{E}^*\{E^*(\boldsymbol{\beta}^*|\boldsymbol{S}^*\boldsymbol{k}_G^{A*})|\mathbb{S}_G^*\} = \mathbb{E}^*\{E_N^*(\boldsymbol{\beta}^*|\boldsymbol{k}_G^{A*}, \mathbf{n}_G)|\mathbb{S}_G^*\},$$

Proof. (1) For any $\mathbb{S}^* \subseteq \mathbb{S}_{\max}^*$,

$$\begin{aligned}\mathbb{E}^*(S_{\cdot j^*}^*|\mathbb{S}^*) &= \sum_{\boldsymbol{s}^* \in \mathbb{S}^*} \Pr(\boldsymbol{S}^* = \boldsymbol{s}^*|\boldsymbol{S}^* \in \mathbb{S}^*) s_{\cdot j^*}^* \quad (\because \text{Equation 200}) \\ &= \sum_{s=0}^{1} \Pr(S_{\cdot j^*}^* = s|\boldsymbol{S}^* \in \mathbb{S}^*) s \\ &= \Pr(S_{\cdot j^*}^* = 1|\boldsymbol{S}^* \in \mathbb{S}^*).\end{aligned} \tag{207}$$

When $\mathbf{n}_G = \bar{\mathbf{n}}$,

$$\begin{aligned}n_G^*(g^T, g^C) &\geq n_G(g^T, g^C) \quad (\because \text{Equations 181 and 182}) \\ &= \bar{n}(g^T, g^C) \quad (\because \mathbf{n}_G = \bar{\mathbf{n}}) \\ &\geq 0 \quad (\because \bar{\mathbf{n}} \in \mathbb{N}_{\max})\end{aligned} \tag{208}$$

In addition, when $\bar{n}(g^T, g^C) \geq 1$,

$$\begin{aligned}n_G^*(g^T, g^C) &\geq 1 (\because \text{Equation 208}) \\ \therefore \boldsymbol{k}_{G,\cdot j^*}^*(g^T, g^C) &\neq \mathbf{0}^* \quad (\because \text{Lemma 14 (4)})\end{aligned} \tag{209}$$

It follows

$$\begin{aligned}&\mathbb{E}_{S|N}^*[E^*\{\boldsymbol{\beta}^*|\boldsymbol{S}^*\boldsymbol{k}_G^*(g^T, g^C)\}|\mathbb{S}_N^*] \\ &= \mathbb{E}_{S|N}^*\left\{\frac{\sum_{j^*} S_{j^*}^* k_{G,\cdot j^*}^*(g^T, g^C)\beta_{\cdot j^*}^*}{\sum_{j^*} S_{j^*}^* k_{G,\cdot j^*}^*(g^T, g^C)}\middle|\mathbb{S}_N^*\right\} \quad (\because \text{Equations 165 and 209}) \\ &= \frac{1}{\bar{n}(g^T, g^C)} \sum_{j^*} \mathbb{E}_{S|N}^*(S_{j^*}^*|\mathbb{S}_N^*) k_{G,\cdot j^*}^*(g^T, g^C)\beta_{\cdot j^*}^* \\ &\quad (\because \text{Equation 36, Lemmas 10, 19 (1), and 23 (1)}, \mathbf{n}_G = \bar{\mathbf{n}}, \bar{n}(g^T, g^C) \geq 1) \\ &= \frac{1}{\bar{n}(g^T, g^C)} \sum_{j^*} \Pr(S_{\cdot j^*}^* = 1|\boldsymbol{S}^* \in \mathbb{S}_N^*) k_{G,\cdot j^*}^*(g^T, g^C)\beta_{\cdot j^*}^* \quad (\because \text{Equation 207}) \\ &= \frac{1}{\bar{n}(g^T, g^C)} \sum_{j^* \in \mathbb{J}_G^*(g^T, g^C)} \Pr(S_{\cdot j^*}^* = 1|\boldsymbol{S}^* \in \mathbb{S}_N^*) k_{G,\cdot j^*}^*(g^T, g^C)\beta_{\cdot j^*}^* \\ &\quad (\because \text{when } j^* \notin \mathbb{J}_G^*(g^T, g^C),\ k_{G,\cdot j^*}^*(g^T, g^C) = 0) \\ &= \frac{1}{\bar{n}(g^T, g^C)} \sum_{j^* \in \mathbb{J}_G^*(g^T, g^C)} \frac{\bar{n}(g^T, g^C)}{n_G^*(g^T, g^C)} k_{G,\cdot j^*}^*(g^T, g^C)\beta_{\cdot j^*}^* \quad (\because \mathbb{S}_N^* \neq \emptyset, \text{Lemma 22 (1), Equation 209}) \\ &= \sum_{j^*} \frac{k_{G,\cdot j^*}^*(g^T, g^C)\beta_{\cdot j^*}^*}{n_G^*(g^T, g^C)} \quad (\because \text{when } j^* \notin \mathbb{J}_G^*(g^T, g^C),\ k_{G,\cdot j^*}^*(g^T, g^C) = 0) \\ &= E^*\{\boldsymbol{\beta}^*|\boldsymbol{k}_G^*(g^T, g^C)\} \quad (\because \text{Equations 165, 181, and 209})\end{aligned}$$

When $n^*_G(g^T, g^C) = 0$, it follows that

$$\boldsymbol{k}^*_G(g^T, g^C) = \mathbf{0}^* \quad (\because \text{Lemma 14 (4)})$$
$$\therefore E^*\{\boldsymbol{\beta}^* | \boldsymbol{k}^*_G(g^T, g^C)\} = 0 \quad (\because \text{Equation 166})$$

and

$$E^*\{\boldsymbol{\beta}^* | \boldsymbol{S}^* \boldsymbol{k}^*_G(g^T, g^C)\} = 0 \quad (\because \boldsymbol{S}^* \boldsymbol{k}^*_G(g^T, g^C) = \mathbf{0}^*, \text{Equation 166})$$
$$\therefore \mathbb{E}^*_{S|N}[E^*\{\boldsymbol{\beta}^* | \boldsymbol{S}^* \boldsymbol{k}^*_G(g^T, g^C)\} | \mathbb{S}^*_N] = 0$$

For reference, since $\mathbb{S}^*_N \neq \emptyset$, according to Equation 208, $n^*_G(g^T, g^C) = 0$ leads to $\bar{n}(g^T, g^C) = 0$.

(2) Suppose $\mathbf{n}_G = \bar{\mathbf{n}}$. Thus, it holds that $n^A_G = \bar{n}^A$ and $\boldsymbol{S}^* \in \mathbb{S}^*_N$.

When $\bar{n}^A = 0$, it follows that

$$E^*_N(\boldsymbol{\beta}^* | \boldsymbol{k}^{A*}_G, \bar{\mathbf{n}}) = 0 \quad (\because \text{Equation 206})$$

and

$$E^*(\boldsymbol{\beta}^* | \boldsymbol{S}^* \boldsymbol{k}^{A*}_G) = 0 \quad (\because \text{Lemma 14 (4)}, \boldsymbol{S}^* \boldsymbol{k}^{A*}_G = \mathbf{0}^*, \text{Equation 166})$$
$$\therefore \mathbb{E}^*_{S|N}\{E^*(\boldsymbol{\beta}^* | \boldsymbol{S}^* \boldsymbol{k}^{A*}_G) | \mathbb{S}^*_N\} = 0$$
$$= E^*_N(\boldsymbol{\beta}^* | \boldsymbol{k}^{A*}_G, \bar{\mathbf{n}}).$$

Below, suppose $\bar{n}^A \geq 1$. It follows

$$\begin{aligned}
&\mathbb{E}^*_{S|N}\{E^*(\boldsymbol{\beta}^* | \boldsymbol{S}^* \boldsymbol{k}^{A*}_G) | \mathbb{S}^*_N\} \\
&= \mathbb{E}^*_{S|N}\Big[\sum_{g^{-A}=0}^{1} \frac{n^A_G(1, g^{-A})}{n^A_G} E^*\{\boldsymbol{\beta}^* | \boldsymbol{S}^* \boldsymbol{k}^{A*}_G(1, g^{-A})\} \Big| \mathbb{S}^*_N\Big] \quad (\because \text{Lemma 3 (5)}, \\
&\quad \text{where } \boldsymbol{Z}^{(1)} = \boldsymbol{S}^* \boldsymbol{k}^{A*}_G \neq \mathbf{0}^*, \boldsymbol{Z}^{(2)} = \boldsymbol{k}^{-A*}_G, \text{Lemmas 14 (4) and 19 (1), Equations 9 and 182)} \\
&= \sum_{g^{-A}=0}^{1} \mathbb{E}^*_{S|N}\Big[\frac{n^A_G(1, g^{-A})}{n^A_G} E^*\{\boldsymbol{\beta}^* | \boldsymbol{S}^* \boldsymbol{k}^{A*}_G(1, g^{-A})\} \Big| \mathbb{S}^*_N\Big] \quad (\because \text{Lemma 23 (1)}) \\
&= \sum_{g^{-A}=0}^{1} \frac{\bar{n}^A(1, g^{-A})}{\bar{n}^A} \mathbb{E}^*_{S|N}[E^*\{\boldsymbol{\beta}^* | \boldsymbol{S}^* \boldsymbol{k}^{A*}_G(1, g^{-A})\} | \mathbb{S}^*_N] \quad (\because \text{Lemma 10 (2)}, \mathbf{n}_G = \bar{\mathbf{n}})
\end{aligned} \tag{210}$$

where

$$\mathbb{S}^*_{\text{def}}\Big[\sum_{g^{-A}=0}^{1} \frac{n^A_G(1, g^{-A})}{n^A_G} E^*\{\boldsymbol{\beta}^* | \boldsymbol{S}^* \boldsymbol{k}^{A*}_G(1, g^{-A})\}\Big] = \mathbb{S}^*_{\text{def}}\Big[\frac{n^A_G(1, g^{-A})}{n^A_G} E^*\{\boldsymbol{\beta}^* | \boldsymbol{S}^* \boldsymbol{k}^{A*}_G(1, g^{-A})\}\Big]$$
$$= \mathbb{S}^*_{\max}(n^A_G \geq 1)$$
$$\mathbb{S}^*_N \in \mathbb{S}^*_{\max}(n^A_G \geq 1).$$

When $\bar{n}^A(1, g^{-A}) \geq 1$ for any $g^{-A} \in \{0, 1\}$, Equation 210 is equal to

$$\sum_{g^{-A}=0}^{1} \frac{\bar{n}^A(1, g^{-A})}{\bar{n}^A} E^*\{\boldsymbol{\beta}^* | \boldsymbol{k}^*_G(1, g^{-A})\} \quad (\because \text{Lemma24 (1)})$$
$$= E^*_N(\boldsymbol{\beta}^* | \boldsymbol{k}^{A*}_G, \bar{\mathbf{n}}) \quad (\because \text{Equation 205})$$

When $\bar{n}^A(1, g^{-A}) \geq 1$ and $\bar{n}^A(1, 1 - g^{-A}) = 0$ for either $g^{-A} = 0$ or $g^{-A} = 1$, Equation 210 is equal to

$$\frac{\bar{n}^A(1, g^{-A})}{\bar{n}^A} E^*\{\boldsymbol{\beta}^* | \boldsymbol{k}^*_G(1, g^{-A})\} + \frac{0}{\bar{n}^A} \cdot 0$$
$$= E^*_N(\boldsymbol{\beta}^* | \boldsymbol{k}^{A*}_G, \bar{\mathbf{n}}).$$

It is not possible that $\bar{n}^A(1, g^{-A}) = 0$ for any $g^{-A} \in \{0, 1\}$. Suppose otherwise. It would hold

$$\begin{aligned}\bar{n}^A &= \sum_{g^{-A}=0}^{1} \bar{n}^A(1, g^{-A}) \quad (\because \text{Equation 185}) \\ &= 0,\end{aligned}$$

a contradiction. This completes the proof.

(3)

$$\begin{aligned}\mathbb{E}^*\{E^*(\boldsymbol{\beta}^*|\boldsymbol{S}^*\boldsymbol{k}_G^{A*})|\mathbb{S}_G^*\} &= \mathbb{E}_N^*[\mathbb{E}_{S|N}^*\{E^*(\boldsymbol{\beta}^*|\boldsymbol{S}^*\boldsymbol{k}_G^{A*})|\mathbb{S}_N^*\}|\mathbb{S}_G^*] \quad (\because \text{Lemma 23 (2)}) \\ &= \mathbb{E}^*\{E_N^*(\boldsymbol{\beta}^*|\boldsymbol{k}_G^{A*}, \mathbf{n}_G)|\mathbb{S}_G^*\} \quad (\because \mathbb{S}_N^* \neq \emptyset, \text{Lemma 24 (2), Equation 203})\end{aligned}$$

□

Suppose that $\boldsymbol{K}_G^* \in \mathbb{U}^*$ satisfies Condition 1*. For $m \in \{1, 2, \dots, n\}$, define m-th moment feasible sampling space as

$$\begin{aligned}\mathbb{S}_G^{m*} &\equiv \mathbb{S}^{m*}(\boldsymbol{K}_G^*) \equiv \mathbb{S}_{\max}^*(\mathbf{n}_G \in \mathbb{N}^m) \\ \mathbb{S}_G^{Am*} &\equiv \mathbb{S}^{Am*}(\boldsymbol{K}_G^*) \equiv \mathbb{S}_{\max}^*(\mathbf{n}_G \in \mathbb{N}^{Am}).\end{aligned} \tag{211}$$

Specifically,

$$\begin{aligned}\mathbb{S}_F^{m*} &\equiv \mathbb{S}^{m*}(\boldsymbol{K}_F^*) \\ &= \mathbb{S}_{\max}^*(\mathbf{n}_F \in \mathbb{N}^m) \quad (\because \text{Lemma 13 (1), Equations 188 and 211}) \\ &= \mathbb{S}_{\max}^*(n \geq m) \quad (\because \text{Equations 188 and 204}) \\ &= \mathbb{S}_{\max}^* \quad (\because m \in \{1, 2, \dots, n\})\end{aligned} \tag{212}$$

under Assumption 3*

$$\begin{aligned}\mathbb{S}_P^{m*} &\equiv \mathbb{S}^{m*}(\boldsymbol{K}_P^*) \\ &= \mathbb{S}_{\max}^*(\mathbf{n}_P \in \mathbb{N}^m) \quad (\because \text{Lemma 13 (2), Equations 189 and 211}) \\ &= \mathbb{S}_{\max}^*(n_P \geq m) \quad (\because \text{Equations 189 and 204})\end{aligned} \tag{213}$$

and under Assumption 2*

$$\begin{aligned}\mathbb{S}_U^{m*} &\equiv \mathbb{S}^{m*}(\boldsymbol{K}_U^*) \\ &= \mathbb{S}_{\max}^*(\mathbf{n}_U \in \mathbb{N}^m) \quad (\because \text{Lemma 13 (3), Equations 190 and 211}) \\ &= \mathbb{S}_{\max}^*(n_U^A \geq m) \quad (\because \text{Equations 190 and 204})\end{aligned} \tag{214}$$

Like Equation 48, define

$$\begin{aligned}\bar{\tau}_P^* &\equiv E^*(\boldsymbol{\tau}^*|\boldsymbol{k}_P^*) \\ &= \bar{\tau}^* + E^*(\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*}|\boldsymbol{k}_P^*).\end{aligned} \tag{215}$$

Proposition 1* (Bias of ATE Estimators: SP). *(1) Under Assumption* 1*, *it holds that* $\mathbb{S}_{def}^*(\hat{\tau}_F) = \mathbb{S}_{\max}^*$ *and*

$$\mathbb{E}^*(\hat{\tau}_F|\mathbb{S}_{\max}^*) - \bar{\tau}^* = 0.$$

(2) Under Assumption 3*, *it holds that* $\mathbb{S}_{def}^*(\hat{\tau}_P) = \mathbb{S}_P^{1*}$ *and, when* $\mathbb{S}_P^{1*} \neq \emptyset$,

$$\mathbb{E}^*(\hat{\tau}_P|\mathbb{S}_P^{1*}) - \bar{\tau}^* = E^*(\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*}|\boldsymbol{k}_P^*).$$

(3) Under Assumption 2*, *it holds that* $\mathbb{S}_{def}^*(\hat{\tau}_U) = \mathbb{S}_U^{1*}$ *and, when* $\mathbb{S}_U^{1*} \neq \emptyset$,

$$\begin{aligned}\mathbb{E}^*(\hat{\tau}_U|\mathbb{S}_U^{1*}) - \bar{\tau}^* = &\left\{\mathbb{E}^*\left(\frac{n_P}{n_U^T}\middle|\mathbb{S}_U^{1*}\right)E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{k}_P^*) + \mathbb{E}^*\left(\frac{n_U^T - n_P}{n_U^T}\middle|\mathbb{S}_U^{1*}\right)E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{k}_U^{T*} - \boldsymbol{k}_P^*)\right\} \\ &- \left\{\mathbb{E}^*\left(\frac{n_P}{n_U^C}\middle|\mathbb{S}_U^{1*}\right)E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{k}_P^*) + \mathbb{E}^*\left(\frac{n_U^C - n_P}{n_U^C}\middle|\mathbb{S}_U^{1*}\right)E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{k}_U^{C*} - \boldsymbol{k}_P^*)\right\}.\end{aligned}$$

(4) Under Assumptions 3^* *and* 5^**, it holds that* $\mathbb{S}^*_{def}(\hat{\tau}_P) = \mathbb{S}^{1*}_P$ *and, when* $\mathbb{S}^{1*}_P \neq \emptyset$,

$$\mathbb{E}^*(\hat{\tau}_P|\mathbb{S}^{1*}_P) - \bar{\tau}^* = 0.$$

(5) Under Assumptions 2^*, 4^*, *and* 5^**, it holds that* $\mathbb{S}^*_{def}(\hat{\tau}_U) = \mathbb{S}^{1*}_U$ *and, when* $\mathbb{S}^{1*}_U \neq \emptyset$,

$$\mathbb{E}^*(\hat{\tau}_U|\mathbb{S}^{1*}_U) - \bar{\tau}^* = 0.$$

PROOF. Suppose that $\boldsymbol{K}^*_G \in \mathbb{U}^*$ satisfies Condition 1^*, $\mathbb{S}^* \subseteq \mathbb{S}^*_{\text{def}}(\hat{\tau}_G)$, and $\mathbb{S}^* \neq \emptyset$. It follows

$$\begin{aligned}
&\mathbb{E}^*(\hat{\tau}_G|\mathbb{S}^*) - \bar{\tau}^* \\
&= \mathbb{E}^*(\hat{\tau}_G|\mathbb{S}^*) - \mathbb{E}^*(\bar{\tau}^*|\mathbb{S}^*) \quad (\because \mathbb{S}^* \neq \emptyset, \text{Lemma 23 (3)}) \\
&= \mathbb{E}^*(\hat{\tau}_G - \bar{\tau}^*|\mathbb{S}^*) \quad (\because \text{Lemma 23 (1)}) \\
&= \mathbb{E}^*_S\{\mathbb{E}_X(\hat{\tau}_G - \bar{\tau}^*)|\mathbb{S}^*\} \quad (\because \text{Equation 199}) \\
&= \mathbb{E}^*_S[\mathbb{E}_X\{E^*(\boldsymbol{\beta}^{T*} + \boldsymbol{\omega}^{T*}|\boldsymbol{S}^*\boldsymbol{K}^*_G\boldsymbol{X}^{T*}) - E^*(\boldsymbol{\beta}^{C*} + \boldsymbol{\omega}^{C*}|\boldsymbol{S}^*\boldsymbol{K}^*_G\boldsymbol{X}^{C*})\}|\mathbb{S}^*] \\
&\quad (\because \mathbb{S}^* \subseteq \mathbb{S}^*_{\text{def}}(\hat{\tau}_G), N^A_G \geq 1, \text{Lemma 20}) \\
&= \mathbb{E}^*_S\{E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{S}^*\boldsymbol{k}^{T*}_G) - E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{S}^*\boldsymbol{k}^{C*}_G)|\mathbb{S}^*\} \quad (\because \text{Lemmas 8 and 12}) \\
&= \mathbb{E}^*\{E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{S}^*\boldsymbol{k}^{T*}_G) - E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{S}^*\boldsymbol{k}^{C*}_G)|\mathbb{S}^*\} \quad (\because \text{Equation 199})
\end{aligned} \tag{216}$$

When, for $\mathbb{N} \subseteq \mathbb{N}^1$, $\mathbb{S}^* = \mathbb{S}^*_G \equiv \mathbb{S}^*_{\max}(\mathbf{n}_G \in \mathbb{N})$ (note that $\mathbb{S}^*_G \subseteq \mathbb{S}^*_{\text{def}}(\hat{\tau}_G)$), Equation 216 leads to

$$\begin{aligned}
&\mathbb{E}^*\{E^*_N(\boldsymbol{\beta}^{T*}|\boldsymbol{k}^{T*}_G, \mathbf{n}_G) - E^*_N(\boldsymbol{\beta}^{C*}|\boldsymbol{k}^{C*}_G, \mathbf{n}_G)|\mathbb{S}^*_G\} \quad (\because \text{Lemmas 23 (1) and 24 (3)}) \\
&= \left\{\mathbb{E}^*\left(\frac{n^{TC}_G}{n^T_G}\middle|\mathbb{S}^*_G\right)E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{k}^{TC*}_G) + \mathbb{E}^*\left(\frac{n^T_G - n^{TC}_G}{n^T_G}\middle|\mathbb{S}^*_G\right)E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{k}^{T*}_G - \boldsymbol{k}^{TC*}_G)\right\} \\
&\quad - \left\{\mathbb{E}^*\left(\frac{n^{TC}_G}{n^C_G}\middle|\mathbb{S}^*_G\right)E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{k}^{TC*}_G) + \mathbb{E}^*\left(\frac{n^C_G - n^{TC}_G}{n^C_G}\middle|\mathbb{S}^*_G\right)E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{k}^{C*}_G - \boldsymbol{k}^{TC*}_G)\right\} \\
&(\because \text{Lemma 23 (1) and, thanks to } \boldsymbol{S} \in \mathbb{S}^*_G, \text{ it follows } n^A_G \geq m \geq 1 \text{ and Equation 205, })
\end{aligned} \tag{217}$$

In particular, when $\boldsymbol{k}^{T*}_G = \boldsymbol{k}^{C*}_G \equiv \boldsymbol{k}^*_G$, Equation 217 leads to

$$\begin{aligned}
&\left\{\mathbb{E}^*\left(\frac{n_G}{n_G}\middle|\mathbb{S}^*_G\right)E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{k}^*_G) + \mathbb{E}^*\left(\frac{n_G - n_G}{n_G}\middle|\mathbb{S}^*_G\right)E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{k}^*_G - \boldsymbol{k}^*_G)\right\} \\
&\quad - \left\{\mathbb{E}^*\left(\frac{n_G}{n_G}\middle|\mathbb{S}^*_G\right)E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{k}^*_G) + \mathbb{E}^*\left(\frac{n_G - n_G}{n_G}\middle|\mathbb{S}^*_G\right)E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{k}^*_G - \boldsymbol{k}^*_G)\right\} \quad (\because \text{Lemma 14 (3)}) \\
&= E^*(\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*}|\boldsymbol{k}^*_G) \quad (\because \mathbb{S}^*_G = \mathbb{S}^* \neq \emptyset, \text{Lemmas 3 (1) and 23 (3)})
\end{aligned} \tag{218}$$

(1) Under Assumption 1^*, it holds that $\mathbb{S}^*_{\text{def}}(\hat{\tau}_F) = \mathbb{S}^*_{\max}$ ($\because$ Assumption 1). When $\boldsymbol{K}^*_G = \boldsymbol{K}^*_F \in \mathbb{U}^*$ and $\mathbb{N} = \mathbb{N}^n$ (note that $\mathbb{N}^n = \{(n, n, n)\} \subseteq \mathbb{N}^1$ and, obviously, $\mathbb{S}^*_{\max} \neq \emptyset$), it follows that $\boldsymbol{K}_G = \boldsymbol{K}_F$ ($\because$ Equation 164), $\mathbb{S}^* = \mathbb{S}^*_G = \mathbb{S}^{n*}_F = \mathbb{S}^*_{\max}$ ($\because$ Equations 211 and 212),

$$\begin{aligned}
\mathbb{E}^*(\hat{\tau}_F|\mathbb{S}^*_{\max}) - \bar{\tau}^* &= E^*(\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^C|\boldsymbol{k}^*_F) \quad (\because \text{Lemma 13 (1), Equations 216 through 218}) \\
&= 0 \quad (\because \text{Lemmas 3 (1) and 5 (2), Equation 12})
\end{aligned}$$

(2) Under Assumption 3^*, it holds that $\mathbb{S}^*_{\text{def}}(\hat{\tau}_P) = \mathbb{S}^{1*}_P$ ($\because$ Lemma 9 (1), $N_P = n_P \geq 1$, Equations 204 and 211). When $\boldsymbol{K}^*_G = \boldsymbol{K}^*_P \in \mathbb{U}^*$, $\mathbb{N} = \mathbb{N}^1$, and $\mathbb{S}^{1*}_P \neq \emptyset$, it follows that $\boldsymbol{K}_G = \boldsymbol{K}_P$ ($\because$ Equation 164) and, according to Lemma 13 (2), $\mathbb{S}^* = \mathbb{S}^*_G = \mathbb{S}^{1*}_P$ ($\because$ Equations 211 and 213) and Equations 216 through 218 are equivalent to the desired result.

(3) Under Assumption 2^*, it holds that $\mathbb{S}^*_{\text{def}}(\hat{\tau}_U) = \mathbb{S}^{1*}_U$ ($\because$ Lemma 9 (2), $N^A_U = n^A_U \geq 1$, Equations 204 and 211). When $\boldsymbol{K}^*_G = \boldsymbol{K}^*_U \in \mathbb{U}^*$, $\mathbb{N} = \mathbb{N}^1$, and $\mathbb{S}^{1*}_U \neq \emptyset$, it follows that

$\boldsymbol{K}_G = \boldsymbol{K}_U$ ($\because$ Equation 164) and, according to Lemma 13 (3), $\mathbb{S}^* = \mathbb{S}^*_G = \mathbb{S}^{1*}_U$ ($\because$ Equations 211 and 214) and Equations 216 and 217 are equivalent to the desired result, where $\mathbb{S}^* = \mathbb{S}^{1*}_U$.

(4) Substitute $\boldsymbol{K}^*_G = \boldsymbol{K}^*_P$. Under Assumption 3* and 5*, it follows that Conditions 1* and 2* (by definition) and Proposition 1* (2) hold. Thus, according to Equation 46,

$$E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^*_P) = 0. \tag{219}$$

When $\mathbb{S}^{1*}_P \neq \emptyset$,

$$\begin{aligned}\mathbb{E}^*(\hat{\tau}_P|\mathbb{S}^{1*}_P) - \bar{\tau}^* &= E^*(\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*}|\boldsymbol{k}^*_P) \quad (\because \text{Proposition } 1^* \text{ (2)})\\ &= 0 \quad (\because \text{Lemma 3 (1), Equation 219})\end{aligned}$$

(5) Substitute $\boldsymbol{K}^*_G = \boldsymbol{K}^*_U$. Under Assumption 2* and 4*, it follows that Conditions 1* and 2* hold (by definition). According to Equation 46,

$$E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_U) = 0. \tag{220}$$

Substitute $\boldsymbol{K}^*_G = \boldsymbol{K}^*_P$. Under Assumption 2* and 5*, it follows that Assumption 3* ($\because$ Lemma 9 (5)), Conditions 1* and 2* and, thus, Equation 219 hold.

Note that

$$\boldsymbol{k}^{A*}_U - \boldsymbol{k}^*_P = \boldsymbol{k}^{A*}_U(1 - \boldsymbol{k}^{-A*}_U) \quad (\because \text{Lemma 9 (2)}) \tag{221}$$

When $\boldsymbol{k}^{A*}_U - \boldsymbol{k}^*_P \neq \mathbf{0}$, it follows that $\boldsymbol{k}^{A*}_U \neq \mathbf{0}$ ($\because$ Lemma 221) and

$$\begin{aligned}&E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_U - \boldsymbol{k}^*_P)\\ &= E^*\{\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_U(1 - \boldsymbol{k}^{-A*}_U)\} \quad (\because \text{Equation 221})\\ &= \Big(\frac{n^{A*}_U - n^*_P}{n^{A*}_U}\Big)^{-1}\Big\{E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_U) - \frac{n^*_P}{n^{A*}_U}E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_U\boldsymbol{k}^{-A*}_U)\Big\}\\ &\quad(\because \text{Lemmas 3 (5) and 14 (4)}, \boldsymbol{k}^{A*}_U - \boldsymbol{k}^*_P, \boldsymbol{k}^{A*}_U \neq \mathbf{0},)\\ &= \frac{n^{A*}_U}{n^{A*}_U - n^*_P}\Big(0 - \frac{n^*_P}{n^{A*}_U} \times 0\Big) \quad (\because \text{Equations 219 and 220, Lemma 9 (2)})\\ &= 0.\end{aligned} \tag{222}$$

Even when $\boldsymbol{k}^{A*}_U - \boldsymbol{k}^*_P = \mathbf{0}$, thanks to Equation 166, both ends of Equation 222 are equal to each other. Therefore, when $\mathbb{S}^{1*}_U \neq \emptyset$, by applying Equations 219 and 222 to Proposition 1* (3) ($\because$ Assumption 2*), the desired result follows.

□

The remarks following Proposition 1 apply to Proposition 1* with necessary modification, though $\mathbb{E}(\hat{\tau}_G)$ and $\mathbb{E}^*(\hat{\tau}_G)$ are different in a few ways. First, even when attrition is ignorable in the *super-population* (Assumption 4* or 5*), it is not necessarily true that, in all *finite samples* (for all $\boldsymbol{s}^* \in \mathbb{S}^*_{\max}$), attrition is ignorable (Assumption 4 or 5) and, accordingly, even under Assumption 2* (or 3*), $\hat{\tau}_U$ (or $\hat{\tau}_P$) is unbiased for $\bar{\tau}$.

Second, unlike Proposition 1 (3), the first two terms of the right hand side of Proposition 1* (3) are not necessarily equal to

$$E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{k}^{T*}_U) = \frac{n^*_P}{n^{T*}_U}E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{k}^*_P) + \frac{n^{T*}_U - n^*_P}{n^{T*}_U}E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{k}^{T*}_U - \boldsymbol{k}^*_P) \tag{223}$$

because

$$\mathbb{E}^*\Big(\frac{n_P}{n^T_U}\Big|\mathbb{S}^{1*}_U\Big) \geq \frac{n^*_P}{n^{T*}_U}, \quad \mathbb{E}^*\Big(\frac{n^T_U - n_P}{n^T_U}\Big|\mathbb{S}^{1*}_U\Big) \leq \frac{n^{T*}_U - n^*_P}{n^{T*}_U}. \tag{224}$$

In the case of $\boldsymbol{S}^* \in \mathbb{S}^*_{\max}$, the probabilities experimenters sample a pair from $\mathbb{J}^*_P$ and $\mathbb{J}^{T*}_U \setminus \mathbb{J}^*_P$ are equal to each other. In the case of $\boldsymbol{S}^* \in \mathbb{S}^{1*}_U$, however, if they draw no pairs from $\mathbb{J}^{C*}_U \setminus \mathbb{J}^*_P$

(where $\mathbb{J}_U^{C*} \equiv \mathbb{J}^*(\boldsymbol{k}_U^{C*})$), they have to draw at least one pair from $\mathbb{J}_P^*$ so that $n_U^C \geq 1$. Thus, the probability that a pair in $\mathbb{J}_P^*$ is sampled is not smaller than the probability for a pair in $\mathbb{J}_U^{T*} \setminus \mathbb{J}_P^*$. Therefore, Equation 224 is derived. The difference between both sides of these inequalities does not decrease in n^*. The case of the last two terms of the right hand side of Proposition 1* (3) is similar.

Third, Equation 223 also explains why Assumption 5* is necessary for unbiasedness of $\hat{\tau}_U$ in Proposition 1* (5) even though Proposition 1 (5) do not premise Assumption 5. Assumption 2* and 4* lead to $E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{k}_U^{T*}) = E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{k}_U^{C*}) = 0$ but not to

$$E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{k}_P^*) = E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{k}_U^{T*} - \boldsymbol{k}_P^*) = E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{k}_P^*) = E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{k}_U^{C*} - \boldsymbol{k}_P^*) = 0, \tag{225}$$

which would hold if I add Assumption 5*. By substituting Equation 225 into Proposition 1* (3), I obtain Proposition 1* (5).

For reference, under Assumption 3*, it holds that $\mathbb{S}_{\text{def}}^*(\hat{\tau}_F) = \mathbb{S}_P^{n*}$. When $\mathbb{N} = \mathbb{N}^n$ and $\mathbb{S}_P^{n*} \neq \emptyset$, it follows that $\mathbb{S}_P^{1*} \neq \emptyset$ ($\because \mathbb{S}_P^{1*} \supseteq \mathbb{S}_P^{n*} \neq \emptyset$) and

$$\begin{aligned}
&\mathbb{E}^*(\hat{\tau}_F|\mathbb{S}_P^{n*}) - \bar{\tau}^* \\
&= \mathbb{E}^*\{E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{S}^*\boldsymbol{k}_F^*) - E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{S}^*\boldsymbol{k}_F^*)|\mathbb{S}_P^{n*}\} \\
&\quad (\because \text{Equation 216 where } \boldsymbol{K}_G^* = \boldsymbol{K}_F^* \in \mathbb{U}^*, \mathbb{S}^* = \mathbb{S}_P^{n*} = \mathbb{S}_{\text{def}}^*(\hat{\tau}_F) \neq \emptyset, \text{Lemma 13 (1)}) \\
&= \mathbb{E}^*\{E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{S}^*\boldsymbol{k}_P^*) - E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{S}^*\boldsymbol{k}_P^*)|\mathbb{S}_P^{n*}\} \quad (\because \boldsymbol{S}^* \in \mathbb{S}_P^{n*}) \\
&= E^*(\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*}|\boldsymbol{k}_P^*) \\
&\quad (\because \text{Equations 216 through 218 where } \boldsymbol{K}_G^* = \boldsymbol{K}_P^* \in \mathbb{U}^*, \mathbb{S}^* = \mathbb{S}_G^* = \mathbb{S}_P^{n*}, \text{Lemma 13 (2)}) \\
&= \mathbb{E}^*(\hat{\tau}_P|\mathbb{S}_P^{1*}) - \bar{\tau}^* \quad (\because \mathbb{S}_P^{1*} \neq \emptyset, \text{Proposition 1* (2)})
\end{aligned}$$

Note that

$$\begin{aligned}
\mathbb{E}^*(\hat{\tau}_P|\mathbb{S}_P^{1*}) - \bar{\tau}_P^* &= \{\mathbb{E}^*(\hat{\tau}_P|\mathbb{S}_P^{1*}) - \bar{\tau}^*\} - (\bar{\tau}_P^* - \bar{\tau}^*) \\
&= E^*(\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*}|\boldsymbol{k}_P^*) - E^*(\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*}|\boldsymbol{k}_P^*) \quad (\because \text{Equation 215}) \\
&= 0.
\end{aligned}$$

<u>Closed Form.</u> Before deriving the closed form of Proposition 1* (3) in limit, we have to prove some lemmas.

Suppose that $\boldsymbol{K}_G^* \in \mathbb{U}^*$ satisfies Condition 1*, In order to obtain simplified and essential expression in Lemmas 25 and 26 and Propositions 2* through 4* and, thus, closed forms of Propositions 1* (3) through 4* in limit, I also suppose that for every $(g^T, g^C) \in \{0,1\}^2$, either $\lim_{n^* \to \infty} n_G^*(g^T, g^C) = \infty$ or $\lim_{n^* \to \infty} n_G^*(g^T, g^C) < 2$, and

$$\lim_{n^* \to \infty} \frac{n_G^*(g^T, g^C)}{n^*} \equiv p_G^*(g^T, g^C) \tag{226}$$

exists. Accordingly,

$$\begin{aligned}
p_G^{A*}(g^A) &\equiv \lim_{n^* \to \infty} \frac{n_G^{A*}(g^A)}{n^*} \\
&= \lim_{n^* \to \infty} \frac{\sum_{g^{-A}=0}^{1} n_G^{A*}(g^A, g^{-A})}{n^*} \\
&= \sum_{g^{-A}=0}^{1} p_G^{A*}(g^A, g^{-A})
\end{aligned} \tag{227}$$

also exists. Denote

$$p_G^{A*} \equiv p_G^{A*}(1). \tag{228}$$

For $g \in \{0,1\}$, since $0 \leq n_G^{A*}(g) \leq n^*$ and $0 \leq n_G^*(g^T, g^C) \leq n^*$, it holds $0 \leq p_G^{A*}(g) \leq 1$ and $0 \leq p_G^*(g^T, g^C) \leq 1$.

Note that, for $\bar{\mathbf{n}} \in \mathbb{N}_{\max}$, $\mathbb{S}^* \subseteq \mathbb{S}^*_{\max}$, and $\mathbb{S}^* \neq \emptyset$,

$$\begin{aligned} \Pr(\mathbf{n}_G = \bar{\mathbf{n}} | \boldsymbol{S}^* \in \mathbb{S}^*) &= \sum_{\boldsymbol{s}^* \in \mathbb{S}^*(\mathbf{n}_G = \bar{\mathbf{n}})} \Pr(\boldsymbol{S}^* = \boldsymbol{s}^* | \boldsymbol{S}^* \in \mathbb{S}^*) \\ &= \frac{|\mathbb{S}^*(\mathbf{n}_G = \bar{\mathbf{n}})|}{|\mathbb{S}^*|} \quad (\because \text{Lemma 21 (1)}, |\mathbb{S}^*| \neq 0) \end{aligned} \tag{229}$$

Lemma 25 (Probability of Number of Pairs). *Suppose that $\boldsymbol{K}^*_G \in \mathbb{U}^*$ satisfies Condition 1^* and $\bar{\mathbf{n}} \in \mathbb{N}_{\max}$.*

(1)

$$\lim_{n^* \to \infty} \Pr(n^A_G = \bar{n}^A | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}) = n! \prod_{g=0}^{1} \frac{\{p^{A*}_G(g)\}^{\bar{n}^A(g)}}{\bar{n}^A(g)!}$$

(2)

$$\lim_{n^* \to \infty} \Pr(\mathbf{n}_G = \bar{\mathbf{n}} | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}) = n! \prod_{(g^T, g^C) \in \{0,1\}^2} \frac{\{p^*_G(g^T, g^C)\}^{\bar{n}(g^T, g^C)}}{\bar{n}(g^T, g^C)!}$$

Proof. (1) Suppose $n^A_G = \bar{n}^A$.

When $n^{A*}_G(g) \geq \bar{n}^A(g)$ for any $g \in \{0, 1\}$, we sample $\bar{n}^A(g)$ pairs from $\mathbb{J}^{A*}_G(g)$, whose size is $n^{A*}_G(g)$. Note that $\bigcap_{g=0}^{1} \mathbb{J}^{A*}_G(g) = \emptyset$. Thus, the number of values $\boldsymbol{s}^* \in \mathbb{S}^*_{\max}(n^A_G = \bar{n}^A)$ can take is

$$|\mathbb{S}^*_{\max}(n^A_G = \bar{n}^A)| = \prod_{g=0}^{1} {}_{n^{A*}_G(g)}C_{\bar{n}^A(g)}. \tag{230}$$

It follows

$$\begin{aligned} \Pr(n^A_G = \bar{n}^A | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}) &= \frac{|\mathbb{S}^*_{\max}(n^A_G = \bar{n}^A)|}{|\mathbb{S}^*_{\max}|} \quad (\because \text{Equation 229}, \mathbb{S}^*_{\max} \neq \emptyset) \\ &= \Big(\prod_{g=0}^{1} {}_{n^{A*}_G(g)}C_{\bar{n}^A(g)} \Big) \div {}_{n^*}C_n, \quad (\because \text{Equations 194 and 230}) \end{aligned} \tag{231}$$

and, therefore,

$$\begin{aligned} &\lim_{n^* \to \infty} \Pr(n^A_G = \bar{n}^A | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}) \\ &= \lim_{n^* \to \infty} \bigg(\Big[\prod_{g=0}^{1} \frac{n^{A*}_G(g)!}{\bar{n}^A(g)! \{n^{A*}_G(g) - \bar{n}^A(g)\}!} \Big] \div \frac{n^*!}{n!(n^* - n)!} \bigg) \quad (\because n^{A*}_G(g) \geq \bar{n}^A(g), \text{Equations 192}) \\ &= \lim_{n^* \to \infty} \bigg(\Big[\prod_{g=0}^{1} \frac{1}{\bar{n}^A(g)!} \prod_{h=1}^{\bar{n}^A(g)} \{n^{A*}_G(g) + 1 - h\} \Big] \div \Big\{ \frac{1}{n!} \prod_{h=1}^{n} (n^* + 1 - h) \Big\} \bigg) \\ &= n! \prod_{g=0}^{1} \frac{1}{\bar{n}^A(g)!} \prod_{h=1}^{\bar{n}^A(g)} \lim_{n^* \to \infty} \frac{n^{A*}_G(g) + 1 - h}{n^* + 1 - \{h + g \cdot \bar{n}^A(0)\}} \quad (\because \text{Equation 185}, \sum_{g=0}^{1} \bar{n}^A(g) = n) \\ &= n! \prod_{g=0}^{1} \frac{1}{\bar{n}^A(g)!} \prod_{h=1}^{\bar{n}^A(g)} p^{A*}_G(g) \quad (\because \text{Equation 227}) \\ &= n! \prod_{g=0}^{1} \frac{\{p^{A*}_G(g)\}^{\bar{n}^A(g)}}{\bar{n}^A(g)!}. \end{aligned}$$

When $n_G^{A*}(g) < \bar{n}^A(g)$ for some $g \in \{0,1\}$, we cannot sample $\bar{n}^A(g)$ pairs from $\mathbb{J}_G^{A*}(g)$. Thus,

$$\begin{aligned}\Pr(n_G^A = \bar{n}^A | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}) &= \Big(\prod_{g=0}^{1} {}_{n_G^{A*}(g)}C_{\bar{n}^A(g)}\Big) \div {}_{n^*}C_n \\ &= 0. \quad (\because {}_{n_G^{A*}(g)}C_{\bar{n}^A(g)} = 0 \text{ for some } g \in \{0,1\}).\end{aligned}$$

For any $g^{-A} \in \{0,1\}$, it follows $\lim_{n^*\to\infty} n_G^{A*}(g^A, g^{-A}) < 2$ where $g^A = g$. For, otherwise, $\lim_{n^*\to\infty} n_G^{A*}(g^A, g^{-A}) = \infty$ for some $g^{-A} \in \{0,1\}$ and

$$\begin{aligned}\lim_{n^*\to\infty} n_G^{A*}(g^A) &= \lim_{n^*\to\infty} \sum_{g^{-A}\in\{0,1\}} n_G^{A*}(g^A, g^{-A}) \quad (\because \text{Equations 180 and 181}) \\ &= \sum_{g^{-A}\in\{0,1\}} \lim_{n^*\to\infty} n_G^{A*}(g^A, g^{-A}) \\ &= \infty \quad (\because \lim_{n^*\to\infty} n_G^{A*}(g^A, g^{-A}) = \infty, \lim_{n^*\to\infty} n_G^{A*}(g^A, 1-g^{-A}) = 0 \text{ or } \infty)\end{aligned}$$

and, thus, $n_G^{A*}(g^A) > n \geq \bar{n}^A(g^A)$, a contradiction. Therefore, $p_G^{A*}(g^A) = 0$ ($\because$ Equation 227). It follows

$$n! \prod_{g^A=0}^{1} \frac{\{p_G^{A*}(g^A)\}^{\bar{n}^A(g^A)}}{\bar{n}^A(g^A)!} = 0.$$

Thus, the desired result follows.

(2) When $n_G^*(g^T, g^C) \geq \bar{n}(g^T, g^C)$ for every $(g^T, g^C) \in \{0,1\}^2$, it follows

$$\begin{aligned}\Pr(\mathbf{n}_G = \bar{\mathbf{n}} | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}) &= \frac{|\mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}})|}{|\mathbb{S}^*_{\max}|} \quad (\because \text{Equation } 229, \mathbb{S}^*_{\max} \neq \emptyset) \\ &= \Big(\prod_{(g^T,g^C)\in\{0,1\}^2} {}_{n_G^*(g^T,g^C)}C_{\bar{n}(g^T,g^C)}\Big) \div {}_{n^*}C_n, \quad (\because \text{Equations 194 and 195})\end{aligned} \tag{232}$$

and, therefore,

$$\begin{aligned}&\lim_{n^*\to\infty} \Pr(\mathbf{n}_G = \bar{\mathbf{n}} | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}) \\ &= \lim_{n^*\to\infty} \Big(\Big[\prod_{g^T=0}^{1}\prod_{g^C=0}^{1} \frac{n_G^*(g^T,g^C)!}{\bar{n}(g^T,g^C)!\{n_G^*(g^T,g^C) - \bar{n}(g^T,g^C)\}!}\Big] \div \frac{n^*!}{n!(n^*-n)!}\Big) \\ &\quad (\because n_G^*(g^T,g^C) \geq \bar{n}(g^T,g^C), \text{Equations 192 and 195, Lemma 21 (2)}) \\ &= \lim_{n^*\to\infty} \Big(\Big[\prod_{(g^T,g^C)\in\{0,1\}^2} \frac{1}{\bar{n}(g^T,g^C)!}\Big\{\prod_{h=1}^{\bar{n}(g^T,g^C)} n_G^*(g^T,g^C) + 1 - h\Big\}\Big] \div \Big\{\frac{1}{n!}\Big(\prod_{h=1}^{n} n^* + 1 - h\Big)\Big\}\Big) \\ &= n! \prod_{(g^T,g^C)\in\{0,1\}^2} \frac{1}{\bar{n}(g^T,g^C)!} \\ &\quad \times \prod_{h=1}^{\bar{n}(g^T,g^C)} \lim_{n^*\to\infty} \frac{n_G^*(g^T,g^C) + 1 - h}{n^* + 1 - [h + g^T n(0,0) + g^C\{n(0,0) + n(1,0)\} + g^T g^C\{n(0,1) - n(0,0)\}]} \\ &\quad (\because \text{Equation } 185, \sum_{g^T=0}^{1}\sum_{g^C=0}^{1} \bar{n}(g^T,g^C) = n) \\ &= n! \prod_{(g^T,g^C)\in\{0,1\}^2} \frac{\{p_G^*(g^T,g^C)\}^{\bar{n}(g^T,g^C)}}{\bar{n}(g^T,g^C)!} \quad (\because \text{Equation 226})\end{aligned}$$

When $n_G^*(g^T, g^C) < \bar{n}(g^T, g^C)$ for some $(g^T, g^C) \in \{0,1\}^2$, we cannot sample $\bar{n}(g^T, g^C)$ pairs from $\mathbb{J}_G^*(g^T, g^C)$. Thus,

$$\begin{aligned}\Pr(\mathbf{n}_G = \bar{\mathbf{n}} | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}) &= \Big(\prod_{(g^T,g^C)\in\{0,1\}^2} {}_{n_G^*(g^T,g^C)}C_{\bar{n}(g^T,g^C)} \Big) \div {}_{n^*}C_n \\ &= 0. \quad (\because {}_{n_G^*(g^T,g^C)}C_{\bar{n}(g^T,g^C)} = 0 \text{ for some } (g^T, g^C) \in \{0,1\}^2)\end{aligned}$$

It follows $\lim_{n^*\to\infty} n_G^*(g^T, g^C) < 2$. For, otherwise, $\lim_{n^*\to\infty} n_G^*(g^T, g^C) = \infty > n \geq \bar{n}(g^T, g^C)$, a contradiction. Therefore, $p_G^*(g^T, g^C) = 0$ ($\because$ Equation 226). It follows

$$n! \prod_{(g^T,g^C)\in\{0,1\}^2} \frac{\{p_G^*(g^T, g^C)\}^{\bar{n}(g^T,g^C)}}{\bar{n}(g^T, g^C)!} = 0.$$

Thus, the desired result follows.

□

Lemma 26 (Probability of Sampling Space). *Suppose that $\boldsymbol{K}_G^* \in \mathbb{U}^*$ satisfies Condition 1^*, $\mathbb{S}^* \subseteq \mathbb{S}^*_{\max}$, and $\mathbb{N} \subseteq \mathbb{N}_{\max}$,*

$$\Pr\{\boldsymbol{S}^* \in \mathbb{S}^*(\mathbf{n}_G \in \mathbb{N}) | \boldsymbol{S}^* \in \mathbb{S}^*\} = \sum_{\bar{\mathbf{n}}\in\mathbb{N}} \Pr(\mathbf{n}_G = \bar{\mathbf{n}} | \boldsymbol{S}^* \in \mathbb{S}^*).$$

In particular, when $\mathbb{S}^* = \mathbb{S}^*_{\max}$, $\mathbb{N} = \mathbb{N}^1$, it follows that $\mathbb{S}^*(\mathbf{n}_G \in \mathbb{N}) = \mathbb{S}_G^{1*}$ and, in limit,

$$\lim_{n^*\to\infty} \Pr(\boldsymbol{S}^* \in \mathbb{S}_G^{1*} | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}) = 1 - [(1 - p_G^{T*})^n + (1 - p_G^{C*})^n - \{1 - (p_G^{T*} + p_G^{C*} - p_G^{TC*})\}^n]$$

Proof.

$$\begin{aligned}\Pr\{\boldsymbol{S}^* \in \mathbb{S}^*(\mathbf{n}_G \in \mathbb{N}) | \boldsymbol{S}^* \in \mathbb{S}^*\} &= \sum_{s^*\in\mathbb{S}^*} \Pr(\boldsymbol{S}^* = s^*, \mathbf{n}_G \in \mathbb{N} | \boldsymbol{S}^* \in \mathbb{S}^*) \\ &= \sum_{\bar{\mathbf{n}}\in\mathbb{N}} \sum_{\boldsymbol{s}^*\in\mathbb{S}^*} I(\mathbf{n}_G = \bar{\mathbf{n}}) \Pr(\boldsymbol{S}^* = s^* | \boldsymbol{S}^* \in \mathbb{S}^*) \\ &= \sum_{\bar{\mathbf{n}}\in\mathbb{N}} \Pr(\mathbf{n}_G = \bar{\mathbf{n}} | \boldsymbol{S}^* \in \mathbb{S}^*).\end{aligned}$$

In particular, when $\mathbb{S}^* = \mathbb{S}^*_{\max}$, $\mathbb{N} = \mathbb{N}^1$, in limit,

$$\begin{aligned}&\lim_{n^*\to\infty} \Pr(\boldsymbol{S}^* \in \mathbb{S}_G^{1*} | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}) \\ &= \lim_{n^*\to\infty} \Pr(n_G^T \geq 1 \wedge n_G^C \geq 1 | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}) \\ &= 1 - \lim_{n^*\to\infty} \Pr(n_G^T = 0 \vee n_G^C = 0 | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}) \\ &= 1 - \lim_{n^*\to\infty} [\Pr(n_G^T = 0 | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}) + \Pr(n_G^C = 0 | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}) - \Pr\{\mathbf{n}_G = (0,0,0) | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}\}] \\ &= 1 - [\{p_G^{T*}(0)\}^n + \{p_G^{C*}(0)\}^n - \{p_G^*(0,0)\}^n] \\ &\quad (\because \text{Lemma 25, where } \bar{n}^A(1) = 0, \bar{n}^A(0) = n \text{ or } \bar{n}^A(1,1) = \bar{n}^A(1,0) = 0, \bar{n}(0,0) = n) \\ &= 1 - [(1 - p_G^{T*})^n + (1 - p_G^{C*})^n - \{1 - (p_G^{T*} + p_G^{C*} - p_G^{TC*})\}^n] \\ &\quad (\because \text{Equations 185, 226, 227, and 228})\end{aligned}$$

□

The closed forms of Proposition 1^* (3) in limit is obtained as follows. Note that, when

$\mathbb{S}_U^{1*} \neq \emptyset$, it follows that $\Pr(\boldsymbol{S}^* \in \mathbb{S}_U^{1*} | \boldsymbol{S}^* \in \mathbb{S}_{\max}^*) > 0$ (which implies $p_U^{A*} > 0$) and

$$\begin{aligned}
\lim_{n^* \to \infty} \Pr(\mathbf{n}_U = \bar{\mathbf{n}} | \boldsymbol{S}^* \in \mathbb{S}_U^{1*}) &= \lim_{n^* \to \infty} \Pr(\mathbf{n}_U = \bar{\mathbf{n}} | \boldsymbol{S}^* \in \mathbb{S}_{\max}^*) \div \lim_{n^* \to \infty} \Pr(\boldsymbol{S}^* \in \mathbb{S}_U^{1*} | \boldsymbol{S}^* \in \mathbb{S}_{\max}^*) \\
&= \Big[n! \prod_{(g^T, g^C) \in \{0,1\}^2} \frac{\{p_U^*(g^T, g^C)\}^{\bar{n}(g^T, g^C)}}{\bar{n}(g^T, g^C)!} \Big] \\
&\quad \times \Big(1 - [(1 - p_U^{T*})^n + (1 - p_U^{C*})^n - \{1 - (p_U^{T*} + p_U^{C*} - p_P^*)\}^n] \Big)^{-1}. \\
&\quad (\because \text{Lemmas 25 (2) and 26, } p_U^{TC*} = p_P^*)
\end{aligned} \tag{233}$$

Thus, under Assumption 2*,

$$\begin{aligned}
&\lim_{n^* \to \infty} \mathbb{E}^* \{ f_N^*(\mathbf{n}_U) | \mathbb{S}_U^{1*} \} \\
&= \lim_{n^* \to \infty} \sum_{\bar{\mathbf{n}} \in \mathbb{N}_{\max}} \Pr(\mathbf{n}_U = \bar{\mathbf{n}} | \boldsymbol{S}^* \in \mathbb{S}_U^{1*}) f_N^*(\bar{\mathbf{n}}) \quad (\because \text{Lemma 23 (2), Equation 203}) \\
&= \sum_{\bar{n}^T=1}^{n} \sum_{\bar{n}^C=1}^{n} \sum_{\bar{n}^{TC}=\max(\bar{n}^T+\bar{n}^C-n,0)}^{\min(\bar{n}^T, \bar{n}^C)} \lim_{n^* \to \infty} \Pr(\mathbf{n}_U = \bar{\mathbf{n}} | \boldsymbol{S}^* \in \mathbb{S}_U^{1*}) f_N^*(\bar{\mathbf{n}}) \\
&\quad (\because \boldsymbol{S}^* \in \mathbb{S}_U^{1*}, \mathbf{n}_U \in \mathbb{N}^1, \text{Equations 184 and 204}) \\
&= n! \Big(1 - [(1 - p_U^{T*})^n + (1 - p_U^{C*})^n - \{1 - (p_U^{T*} + p_U^{C*} - p_P^*)\}^n] \Big)^{-1} \\
&\quad \times \sum_{\bar{n}^T=1}^{n} \sum_{\bar{n}^C=1}^{n} \sum_{\bar{n}^{TC}=\max(\bar{n}^T+\bar{n}^C-n,0)}^{\min(\bar{n}^T, \bar{n}^C)} \Big[\prod_{(g^T, g^C) \in \{0,1\}^2} \frac{\{p_U^*(g^T, g^C)\}^{\bar{n}(g^T, g^C)}}{\bar{n}(g^T, g^C)!} \Big] f_N^*(\bar{\mathbf{n}}) \\
&\quad (\because \text{Equation 233})
\end{aligned} \tag{234}$$

Specifically,

$$\begin{aligned}
&\lim_{n^* \to \infty} \mathbb{E}^* (\hat{\tau}_U | \mathbb{S}_U^{1*}) - \bar{\tau}^* \\
&= \lim_{n^* \to \infty} \mathbb{E}^* \{ E_N^*(\boldsymbol{\beta}^{T*} | \boldsymbol{k}_U^{T*}, \mathbf{n}_U) - E_N^*(\boldsymbol{\beta}^{C*} | \boldsymbol{k}_U^{C*}, \mathbf{n}_U) | \mathbb{S}_U^{1*} \} \quad (\because \text{Proposition } 1^* \text{ (3)}) \\
&= n! \Big(1 - [(1 - p_U^{T*})^n + (1 - p_U^{C*})^n - \{1 - (p_U^{T*} + p_U^{C*} - p_P^*)\}^n] \Big)^{-1} \\
&\quad \times \sum_{\bar{n}^T=1}^{n} \sum_{\bar{n}^C=1}^{n} \sum_{\bar{n}^{TC}=\max(\bar{n}^T+\bar{n}^C-n,0)}^{\min(\bar{n}^T, \bar{n}^C)} \Big[\prod_{(g^T, g^C) \in \{0,1\}^2} \frac{\{p_U^*(g^T, g^C)\}^{\bar{n}(g^T, g^C)}}{\bar{n}(g^T, g^C)!} \Big] \\
&\quad \times \Big[\sum_{g^{C\prime}=0}^{1} \frac{\bar{n}^T(1, g^{C\prime})}{\bar{n}^T} E^* \{ \boldsymbol{\beta}^{T*} | \boldsymbol{k}_U^{T*}(1, g^{C\prime}) \} - \sum_{g^{T\prime}=0}^{1} \frac{\bar{n}^C(1, g^{T\prime})}{\bar{n}^C} E^* \{ \boldsymbol{\beta}^{C*} | \boldsymbol{k}_U^{C*}(1, g^{T\prime}) \} \Big] \\
&\quad (\because \text{Equations 205 and 234})
\end{aligned}$$

<u>Conditional Sampling Probability.</u> Make Assumption 2*. When $\boldsymbol{S}^* \in \mathbb{S}_{\max}^*$, $j^* \in \mathbb{J}_U^*(g^T, g^C)$, and $S_{\cdot j^*}^* = 1$, we sample pair j^* from $\mathbb{J}_U^*(g^T, g^C)$ and $n - 1$ pairs from $\mathbb{J}^* \setminus \{j^*\}$, whose size is $n^* - 1$, where $\mathbb{J}^* \equiv \{1, 2, \ldots, n^*\}$. The number of sets of such $n - 1$ pairs is ${}_{n^*-1}C_{n-1}$. Thus, the number of values $\boldsymbol{s}^* \in \mathbb{S}_{\max}^*(S_{\cdot j^*}^* = 1)$ can take is

$$|\mathbb{S}_{\max}^*(S_{\cdot j^*}^* = 1)| = {}_{n^*-1}C_{n-1}. \tag{235}$$

Therefore,

$$\begin{aligned}\Pr(S^*_{\cdot j^*} = 1 | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}) &= \frac{|\mathbb{S}^*_{\max}(S^*_{\cdot j^*} = 1)|}{|\mathbb{S}^*_{\max}|} \\ &= \frac{{}_{n^*-1}C_{n-1}}{{}_{n^*}C_n} \quad (\because \text{Equations 195 and 235}) \\ &= \frac{n}{n^*}, \quad (\because \text{Equation 192})\end{aligned} \tag{236}$$

which is constant for any $(g^T, g^C) \in \{0,1\}^2$. Note that $\mathbb{J}^*_P = \mathbb{J}^*_U(1,1)$ and $\mathbb{J}^{A*}_U \setminus \mathbb{J}^*_P = \mathbb{J}^{A*}_U(1,0)$. Thus, in the case of $\boldsymbol{S}^* \in \mathbb{S}^*_{\max}$, the probabilities we sample a pair from $\mathbb{J}^*_P$ and $\mathbb{J}^{A*}_U \setminus \mathbb{J}^*_P$ are equal to each other.

When $\boldsymbol{S}^* \in \mathbb{S}^{1*}_U \neq \emptyset$, $j^* \in \mathbb{J}^*_U(1,1)$, and $S^*_{\cdot j^*} = 1$, we sample pair j^* from $\mathbb{J}^*_U(1,1)$, which guarantees $\mathbf{n}_U \in \mathbb{N}^1$, and $n-1$ pairs from $\mathbb{J}^* \setminus \{j^*\}$. Thus,

$$\begin{aligned}\Pr(S^*_{\cdot j^*} = 1 | \boldsymbol{S}^* \in \mathbb{S}^{1*}_U) &= \frac{\Pr(S^*_{\cdot j^*} = 1, \boldsymbol{S}^* \in \mathbb{S}^{1*}_U | \boldsymbol{S}^* \in \mathbb{S}^*_{\max})}{\Pr(\boldsymbol{S}^* \in \mathbb{S}^{1*}_U | \boldsymbol{S}^* \in \mathbb{S}^*_{\max})} \\ &= \frac{|\mathbb{S}^*_{\max}(S^*_{\cdot j^*} = 1, \mathbf{n}_U \in \mathbb{N}^1)|}{|\mathbb{S}^*_{\max}(\mathbf{n}_U \in \mathbb{N}^1)|} \\ &= \frac{|\mathbb{S}^*_{\max}(S^*_{\cdot j^*} = 1)|}{|\mathbb{S}^*_{\max}(\mathbf{n}_U \in \mathbb{N}^1)|} \\ &= \frac{{}_{n^*-1}C_{n-1}}{|\mathbb{S}^*_{\max}(\mathbf{n}_U \in \mathbb{N}^1)|}. \quad (\because \text{Equation 235})\end{aligned} \tag{237}$$

Consider the case where $j^* \in \mathbb{J}^{A*}_U(1,0)$, $S^*_{\cdot j^*} = 1$, and $\mathbb{S}^{1*}_U \neq \emptyset$. Suppose $n^* - n^{-A*}_U \geq n$ for a moment. When $\boldsymbol{S}^* \notin \mathbb{S}^{1*}_U$, we sample pair j^* from $\mathbb{J}^{A*}_U(1,0)$, and $n-1$ pairs from $\mathbb{J}^{-A*}_U(0) \setminus \{j^*\}$, whose size is $n^* - n^{-A*}_U - 1$. The number of sets of such $n-1$ pairs is ${}_{n^*-n^{-A*}_U-1}C_{n-1}$. Thus, the number of values $\boldsymbol{s}^* \in \mathbb{S}^*_{\max}(S^*_{\cdot j^*} = 1, \mathbf{n}_U \notin \mathbb{N}^1)$ can take is

$$|\mathbb{S}^*_{\max}(S^*_{\cdot j^*} = 1, \mathbf{n}_U \notin \mathbb{N}^1)| = {}_{n^*-n^{-A*}_U-1}C_{n-1}. \tag{238}$$

Therefore

$$\begin{aligned}|\mathbb{S}^*_{\max}(S^*_{\cdot j^*} = 1, \mathbf{n}_U \in \mathbb{N}^1)| &= |\mathbb{S}^*_{\max}(S^*_{\cdot j^*} = 1)| - |\mathbb{S}^*_{\max}(S^*_{\cdot j^*} = 1, \mathbf{n}_U \notin \mathbb{N}^1)| \\ &= {}_{n^*-1}C_{n-1} - {}_{n^*-n^{-A*}_U-1}C_{n-1} \quad (\because \text{Equations 235 and 238})\end{aligned} \tag{239}$$

Thus,

$$\begin{aligned}\Pr(S^*_{\cdot j^*} = 1 | \boldsymbol{S}^* \in \mathbb{S}^{1*}_U) &= \frac{|\mathbb{S}^*_{\max}(S^*_{\cdot j^*} = 1, \mathbf{n}_U \in \mathbb{N}^1)|}{|\mathbb{S}^*_{\max}(\mathbf{n}_U \in \mathbb{N}^1)|} \\ &= \frac{{}_{n^*-1}C_{n-1} - {}_{n^*-n^{-A*}_U-1}C_{n-1}}{|\mathbb{S}^*_{\max}(\mathbf{n}_U \in \mathbb{N}^1)|} \quad (\because \text{Equation 239})\end{aligned} \tag{240}$$

Accordingly,

$$\begin{aligned}
&\frac{\Pr\{S^*_{\cdot j^*}=1|\boldsymbol{S}^*\in\mathbb{S}^{1*}_U, j^*\in\mathbb{J}^{A*}_U(1,0)\}}{\Pr\{S^*_{\cdot j^*}=1|\boldsymbol{S}^*\in\mathbb{S}^{1*}_U, j^*\in\mathbb{J}^{A*}_U(1,1)\}}\\
&=\frac{{}_{n^*-1}C_{n-1}-{}_{n^*-n_U^{-A*}-1}C_{n-1}}{|\mathbb{S}^*_{\max}(\mathbf{n}_U\in\mathbb{N}^1)|}\div\frac{{}_{n^*-1}C_{n-1}}{|\mathbb{S}^*_{\max}(\mathbf{n}_U\in\mathbb{N}^1)|}\quad(\because\text{Equations 237 and 240})\\
&=\frac{{}_{n^*-1}C_{n-1}-{}_{n^*-n_U^{-A*}-1}C_{n-1}}{{}_{n^*-1}C_{n-1}}\\
&=1-\frac{(n^*-n_U^{-A*}-1)!}{(n-1)!\{(n^*-n_U^{-A*}-1)-(n-1)\}!}\div\frac{(n^*-1)!}{(n-1)!\{(n^*-1)-(n-1)\}!}\\
&\quad(\because\text{Equation 192})\\
&=1-\prod_{h=1}^{n-1}\frac{n^*-n_U^{-A*}-h}{n^*-h}\\
&<1.\quad\left(\because\mathbb{S}^{1*}_U\neq\emptyset, n^{-A*}>0, n^*-n_U^{-A*}\geq n>n-1, 0<\frac{n^*-n_U^{-A*}-h}{n^*-h}<1\right)
\end{aligned}\tag{241}$$

In limit,

$$\begin{aligned}
\lim_{n^*\to\infty}\frac{\Pr\{S^*_{\cdot j^*}=1|\boldsymbol{S}^*\in\mathbb{S}^{1*}_U, j^*\in\mathbb{J}^{A*}_U(1,0)\}}{\Pr\{S^*_{\cdot j^*}=1|\boldsymbol{S}^*\in\mathbb{S}^{1*}_U, j^*\in\mathbb{J}^{A*}_U(1,1)\}}&=1-\prod_{h=1}^{n-1}\lim_{n^*\to\infty}\frac{n^*-n_U^{-A*}-h}{n^*-h}\quad(\because\text{Equation 241})\\
&=1-\prod_{h=1}^{n-1}\left(1-\lim_{n^*\to\infty}\frac{n_U^{-A*}}{n^*}\div\lim_{n^*\to\infty}\frac{n^*-h}{n^*}\right)\\
&=1-(1-p_U^{-A*})^{n-1}\\
&\leq 1,
\end{aligned}$$

where equality hold if and only if $p_U^{-A*}=1$.

When $n^*-n_U^{-A*}<n$, it follows that ${}_{n^*-n_U^{-A*}-1}C_{n-1}=0$ and, thus, Equations 238, 239, and 240 hold and

$$\begin{aligned}
\frac{\Pr\{S^*_{\cdot j^*}=1|\boldsymbol{S}^*\in\mathbb{S}^{1*}_U, j^*\in\mathbb{J}^{A*}_U(1,0)\}}{\Pr\{S^*_{\cdot j^*}=1|\boldsymbol{S}^*\in\mathbb{S}^{1*}_U, j^*\in\mathbb{J}^{A*}_U(1,1)\}}&=\frac{{}_{n^*-1}C_{n-1}-{}_{n^*-n_U^{-A*}-1}C_{n-1}}{{}_{n^*-1}C_{n-1}}\\
&=\frac{{}_{n^*-1}C_{n-1}}{{}_{n^*-1}C_{n-1}}\\
&=1.
\end{aligned}\tag{242}$$

Unlike Equation 236, according to Equations 241 and 242, in the case of $\boldsymbol{S}^*\in\mathbb{S}^{1*}_U$, the probability that a pair in $\mathbb{J}^*_P=\mathbb{J}^*_U(1,1)$ is sampled is not smaller than the probability for a pair in $\mathbb{J}^{A*}_U\backslash\mathbb{J}^*_P=\mathbb{J}^{A*}_U(1,0)$.

Equation 224 is derived as follows. Let $\bar{\mathbf{n}}\in\mathbb{N}_{\max}, \bar{n}^{-A}=0, \mathbb{S}^*_{\max}(n_U^A=\bar{n}^A)\neq\emptyset$ (that is,

$n_U^{A*}(g) \geq \bar{n}^A(g)$ for all $g \in \{0,1\}$). Denote

$$\begin{aligned}
p_N^{-A1} &\equiv \Pr\{n_U^{-A} \geq 1 | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}(n_U^A = \bar{n}^A)\} \\
&= 1 - \Pr\{n_U^{-A} = n_U^{TC} = \bar{n}^{-A} = 0 | \boldsymbol{S}^* \in \mathbb{S}^*_{\max}(n_U^A = \bar{n}^A)\} \\
&= 1 - \frac{\Pr(n_U^A = \bar{n}^A, n_U^{-A} = n_U^{TC} = \bar{n}^{-A} = 0 | \boldsymbol{S}^* \in \mathbb{S}^*_{\max})}{\Pr(n_U^A = \bar{n}^A | \boldsymbol{S}^* \in \mathbb{S}^*_{\max})} \\
&\quad (\because \text{Equation 177}, \mathbb{S}^*_{\max}(n_U^A = \bar{n}^A) \neq \emptyset) \\
&= 1 - \left\{ \left(\prod_{(g^A, g^{-A}) \in \{0,1\}^2} {}_{n_U^{A*}(g^A, g^{-A})}C_{\bar{n}^A(g^A, g^{-A})} \right) \div {}_{n^*}C_n \right\} \div \left\{ \left(\prod_{g=0}^{1} {}_{n_U^{A*}(g)}C_{\bar{n}^A(g)} \right) \div {}_{n^*}C_n \right\} \\
&\quad (\because \text{Equations 231 and 232}) \\
&= 1 - \left(\prod_{g^A=0}^{1} {}_{n_U^{A*}(g^A, 0)}C_{\bar{n}^A(g^A)} \right) \div \left(\prod_{g=0}^{1} {}_{n_U^{A*}(g)}C_{\bar{n}^A(g)} \right) \\
&\quad (\because \bar{n}^A(g^A, 1) = 0, \bar{n}^A(g^A, 0) = \bar{n}^A(g^A)) \\
&\leq 1.
\end{aligned} \tag{243}$$

Case 1: When $n_U^{A*}(g^A, 0) \geq \bar{n}^A(g^A)$ for all $g^A \in \{0,1\}$, Equations 192 and 243 lead to

$$\begin{aligned}
p_N^{-A1} &= 1 - \prod_{g^A=0}^{1} \frac{\{n_U^{A*}(g^A, 0)\}!}{\{\bar{n}^A(g^A)\}!\{n_U^{A*}(g^A, 0) - \bar{n}^A(g^A)\}!} \div \frac{\{n_U^{A*}(g^A)\}!}{\{\bar{n}^A(g^A)\}!\{n_U^{A*}(g^A) - \bar{n}^A(g^A)\}!} \\
&< 1. \quad (\because {}_{n_U^{A*}(g^A, 0)}C_{\bar{n}^A(g^A)} > 0)
\end{aligned} \tag{244}$$

Case 1.1: When $\bar{n}^A(g^A) \neq 0$ for all $g^A \in \{0,1\}$, Equation 244 leads to

$$p_N^{-A1} = 1 - \prod_{g^A=0}^{1} \prod_{h=0}^{\bar{n}^A(g^A)-1} \frac{n_U^{A*}(g^A, 0) - h}{n_U^{A*}(g^A) - h}. \tag{245}$$

Case 1.1.1: When $n_U^{A*}(g^A) - n_U^{A*}(g^A, 0) = n_U^{A*}(g^A, 1) = 0$ for all $g^A \in \{0,1\}$ (that is, $n_U^{-A*} = 0$, namely, $\boldsymbol{r}_U^{-A*} = \boldsymbol{0}^*$), it follows

$$\frac{n_U^{A*}(g^A, 0) - h}{n_U^{A*}(g^A) - h} = 1 \quad (\because n_U^{A*}(g^A) = n_U^{A*}(g^A, 0))$$

$$\therefore p_N^{-A1} = 0. \quad (\because \text{Equation 245})$$

Case 1.1.2: When $n_U^{A*}(g^A) - n_U^{A*}(g^A, 0) = n_U^{A*}(g^A, 1) > 0$ for some $g^A \in \{0,1\}$, it follows

$$\begin{aligned}
n_U^{A*}(g^A, 0) &\geq \bar{n}^A(g^A) \quad (\because \text{Case 1}) \\
&> \bar{n}^A(g^A) - 1 \\
0 &< \frac{n_U^{A*}(g^A, 0) - h}{n_U^{A*}(g^A) - h} \quad (\because n_U^{A*}(g^A, 0) > \bar{n}^A(g^A) - 1) \\
&< 1 \quad (\because n_U^{A*}(g^A) > n_U^{A*}(g^A, 0)) \\
\therefore 0 &< p_N^{-A1} \quad (\because \text{Equation 245}) \\
&< 1.
\end{aligned}$$

Case 1.2: When $\bar{n}^A(g^A) = n, \bar{n}^A(1 - g^A) = 0$ for some $g^A \in \{0,1\}$, Equation 244 leads to

$$p_N^{-A1} = 1 - \prod_{h=0}^{n-1} \frac{n_U^{A*}(g^A, 0) - h}{n_U^{A*}(g^A) - h}.$$

Case 1.2.1: When $n_U^{A*}(g^A) - n_U^{A*}(g^A, 0) = n_U^{A*}(g^A, 1) = 0$, it follows

$$p_N^{-A1} = 0.$$

Case 1.2.2: When $n_U^{A*}(g^A) - n_U^{A*}(g^A, 0) = n_U^{A*}(g^A, 1) > 0$, it follows

$$0 < p_N^{-A1} < 1.$$

Case 1.3: $\bar{n}^A(1) = \bar{n}^A(0) = 0$. This is impossible because $\bar{n}^A(1) + \bar{n}^A(0) = n$.

Case 2: When $n_U^{A*}(g^A, 0) < \bar{n}^A(g^A)$ for some $g^A \in \{0, 1\}$, it follows that ${}_{n_U^{A*}(g^A,0)}C_{\bar{n}^A(g^A)} = 0$ and Equation 243 leads to

$$\begin{aligned} p_N^{-A1} &= 1 - 0 \div \prod_{g=0}^{1} {}_{n_U^{A*}(g)}C_{\bar{n}^A(g)} \\ &= 1. \end{aligned}$$

It is easier to understand p_N^{-A1} in limit. When $p_U^{A*} > 0$,

$$\begin{aligned} \lim_{n^* \to \infty} p_N^{-A1} &= 1 - \frac{\lim_{n^* \to \infty} \Pr(n_U^A = \bar{n}^A, n_U^{-A} = n_U^{TC} = 0 | \boldsymbol{S}^* \in \mathbb{S}^*_{\max})}{\lim_{n^* \to \infty} \Pr(n_U^A = \bar{n}^A | \boldsymbol{S}^* \in \mathbb{S}^*_{\max})} \\ &\quad (\because \text{Equation 243, Lemma 25}, p_U^{A*} > 0, \bar{n}^A \geq 1) \\ &= 1 - n! \prod_{g^A=0}^{1} \frac{\{p_U^{A*}(g^A, 0)\}^{\bar{n}^A(g^A)}}{\bar{n}^A(g^A)!} \div n! \prod_{g^A=0}^{1} \frac{\{p_U^{A*}(g^A)\}^{\bar{n}^A(g^A)}}{\bar{n}^A(g^A)!}. \\ &\quad (\because \text{Lemma 25}, \bar{n}^A(g^A, 1) = 0, \bar{n}^A(g^A, 0) = \bar{n}^A(g^A)) \end{aligned}$$

.

When $0 < p_U^{A*} < 1$, it follows that

$$\begin{aligned} \lim_{n^* \to \infty} p_N^{-A1} &= 1 - \prod_{g^A=0}^{1} \left\{ \frac{p_U^*(g^A, 0)}{p_U^*(g^A)} \right\}^{\bar{n}^A(g^A)} \\ &\leq 1, \end{aligned}$$

where equality holds if and only if $p_U^{A*}(1, 0) = 0$ or $p_U^{A*}(0, 0) = 0, \bar{n}^A(0) > 0$.

When $p_U^{A*} = 1$, it follows that $\bar{n}^A(1) = n, \bar{n}^A(0) = 0, p_U^{A*}(1, 0) = p_U^{-A*}(0), p_U^{A*}(0, 0) = 0$, and

$$\begin{aligned} \lim_{n^* \to \infty} p_N^{-A1} &= 1 - \{p_U^{-A*}(0)\}^n \\ &\leq 1, \end{aligned}$$

where equality holds if and only if $p_U^{A*}(1, 0) = 0$ (when $p_U^{-A*}(0) = \sum_{g^A=0}^{1} p_U^{A*}(g^A, 0) = 0$ and, thus, $p_U^{-A*} = 1$).

Let $\bar{\mathbf{n}} \in \mathbb{N}_{\max}$ ($\bar{n}^{-A}$ is not necessarily equal to zero). Suppose that, for some $g^A \in \{0, 1\}$, it holds $n_U^A(g^A) = \bar{n}^A(g^A), n_U^A(g^A, 1) = \bar{n}^A(g^A, 1), n_U^{A*}(1 - g^A) \geq \bar{n}^A(1 - g^A)$, and $n_U^{A*}(g^A, g^{-A}) \geq \bar{n}^A(g^A, g^{-A})$ for every $g^{-A} \in \{0, 1\}$. We sample $\bar{n}^A(g^A, g^{-A})$ pairs from $\mathbb{J}_U^{A*}(g^A, g^{-A})$ for every $g^{-A} \in \{0, 1\}$ and $\bar{n}^A(1 - g^A)$ pairs from $\mathbb{J}_U^{A*}(1 - g^A)$. Thus, the number of values $\boldsymbol{s}^* \in \mathbb{S}^*_{\max}\{n_U^A(g^A) = \bar{n}^A(g^A), n_U^A(g^A, 1) = \bar{n}^A(g^A, 1)\}$ can take is

$$\begin{aligned} &|\mathbb{S}^*_{\max}\{n_U^A(g^A) = \bar{n}^A(g^A), n_U^A(g^A, 1) = \bar{n}^A(g^A, 1)\}| \\ &= {}_{n_U^{A*}(1-g^A)}C_{\bar{n}^A(1-g^A)} \prod_{g^{-A}=0}^{1} {}_{n_U^{A*}(g^A, g^{-A})}C_{\bar{n}^A(g^A, g^{-A})}. \end{aligned} \tag{246}$$

This equation holds even if $n_U^{A*}(1 - g^A) < \bar{n}^A(1 - g^A)$ or $n_U^{A*}(g^A, g^{-A}) < \bar{n}^A(g^A, g^{-A})$, when both sides of the equation are equal to zero.

It follows

$$\begin{aligned}
&\Pr[n_U^A(g^A,1)=\bar{n}^A(g^A,1)|\boldsymbol{S}^*\in\mathbb{S}^*_{\max}\{n_U^A(g^A)=\bar{n}^A(g^A)\}] \\
&=\frac{\Pr\{n_U^A(g^A)=\bar{n}^A(g^A),n_U^A(g^A,1)=\bar{n}^A(g^A,1)|\boldsymbol{S}^*\in\mathbb{S}^*_{\max}\}}{\Pr\{n_U^A(g^A)=\bar{n}^A(g^A)|\boldsymbol{S}^*\in\mathbb{S}^*_{\max}\}} \quad (\because \mathbb{S}^*_{\max}\{n_U^A(g^A)=\bar{n}^A(g^A)\}\neq\emptyset) \\
&=\frac{|\mathbb{S}^*_{\max}\{n_U^A(g^A)=\bar{n}^A(g^A),n_U^A(g^A,1)=\bar{n}^A(g^A,1)\}|}{|\mathbb{S}^*_{\max}\{n_U^A(g^A)=\bar{n}^A(g^A)\}|} \\
&=\left({}_{n_U^{A*}(1-g^A)}C_{\bar{n}^A(1-g^A)}\prod_{g^{-A}=0}^{1}{}_{n_U^{A*}(g^A,g^{-A})}C_{\bar{n}^A(g^A,g^{-A})}\right)\div\prod_{g^{A\prime}=0}^{1}{}_{n_U^{A*}(g^{A\prime})}C_{\bar{n}^A(g^{A\prime})} \\
&\quad(\because \text{Equations 230 and 246}) \\
&=\left(\prod_{g^{-A}=0}^{1}{}_{n_U^{A*}(g^A,g^{-A})}C_{\bar{n}^A(g^A,g^{-A})}\right)\div\prod_{g^A=0}^{1}{}_{n_U^{A*}(g^A)}C_{\bar{n}^A(g^A)}.
\end{aligned} \tag{247}$$

Suppose $\mathbb{S}_U^{-A1*}(n_U^A=\bar{n}^A)\neq\emptyset$ (namely, $p_N^{-A1}>0$, that is, for some $g^A\in\{0,1\}$, it holds $n_U^{A*}(g^A,0)<\bar{n}^A(g^A)\leq n_U^{A*}(g^A)$ or $n_U^{A*}(g^A)\geq\bar{n}^A(g^A)>0, n_U^{A*}(g^A,1)>0$). For integer $\bar{n}^{TC}\geq 1$, it follows that

$$\begin{aligned}
&\Pr\{n_P=\bar{n}^{TC}|\boldsymbol{S}^*\in\mathbb{S}_U^{-A1*}(n_U^A=\bar{n}^A)\} \\
&=\frac{\Pr(n_U^A=\bar{n}^A,n_P=\bar{n}^{TC}|\boldsymbol{S}^*\in\mathbb{S}^*_{\max})}{\Pr(n_U^A=\bar{n}^A,n_U^{-A}\geq 1|\boldsymbol{S}^*\in\mathbb{S}^*_{\max})} \quad (\because \text{Equation 177, 204, and 211,} \\
&\quad \mathbb{S}_U^{-A1*}(n_U^A=\bar{n}^A)=\mathbb{S}^*_{\max}(n_U^A=\bar{n}^A,n_U^{-A}\geq 1)\neq\emptyset, n_U^{-A}\geq n_P=\bar{n}^{TC}\geq 1) \\
&=\frac{\Pr(n_U^A=\bar{n}^A,n_P=\bar{n}^{TC}|\boldsymbol{S}^*\in\mathbb{S}^*_{\max})}{\Pr\{\mathbb{S}^*_{\max}(n_U^A=\bar{n}^A)|\boldsymbol{S}^*\in\mathbb{S}^*_{\max}\}}\div\frac{\Pr(n_U^A=\bar{n}^A,n_U^{-A}\geq 1|\boldsymbol{S}^*\in\mathbb{S}^*_{\max})}{\Pr\{\mathbb{S}^*_{\max}(n_U^A=\bar{n}^A)|\boldsymbol{S}^*\in\mathbb{S}^*_{\max}\}} \\
&\quad(\because \mathbb{S}^*_{\max}(n_U^A=\bar{n}^A)\supseteq\mathbb{S}_U^{-A1*}(n_U^A=\bar{n}^A)\neq\emptyset) \\
&=\Pr\{n_P=\bar{n}^{TC}|\boldsymbol{S}^*\in\mathbb{S}^*_{\max}(n_U^A=\bar{n}^A)\}\div\Pr\{n_U^{-A}\geq 1|\boldsymbol{S}^*\in\mathbb{S}^*_{\max}(n_U^A=\bar{n}^A)\} \quad (\because \text{Equation 177}) \\
&=\Pr\{n_P=\bar{n}^{TC}|\boldsymbol{S}^*\in\mathbb{S}^*_{\max}(n_U^A=\bar{n}^A)\}\div p_N^{-A1} \quad (\because \text{Equation 243})
\end{aligned} \tag{248}$$

Note that, since $\mathbb{S}_U^{-A1*}(n_U^A=\bar{n}^A)\neq\emptyset$, it follows $p_N^{-A1}>0$.

For positive integers q^* and q, when $q^*\geq q$, it follows

$$\begin{aligned}
{}_{q^*}C_q&=\frac{q^*!}{q!(q^*-q)!} \quad (\because \text{Equation 192}, q^*\geq q\geq 1>0) \\
&=\frac{q^*(q^*-1)!}{q(q-1)!\{(q^*-1)-(q-1)\}!} \\
&=\frac{q^*}{q}{}_{q^*-1}C_{q-1}. \quad (\because \text{Equation 192}, q^*-1\geq q-1\geq 0)
\end{aligned} \tag{249}$$

Equation 249 holds even when $q^*<q$ because ${}_{q^*}C_q=0, 0<q^*-1<q-1, {}_{q^*-1}C_{q-1}=0$.

Suppose $\bar{n}^A \geq 1$. When $n_P^* \geq 1$,

$$
\begin{aligned}
&\mathbb{E}^*_{S|N}\{n_P|\mathbb{S}^*_{\max}(n_U^A=\bar{n}^A)\}\\
&=\sum_{\bar{n}^{TC}=0}^{\bar{n}^A}\Pr\{n_P=\bar{n}^{TC}|\boldsymbol{S}^*\in\mathbb{S}^*_{\max}(n_U^A=\bar{n}^A)\}\bar{n}^{TC}\\
&=\sum_{\bar{n}^{TC}=1}^{\bar{n}^A}\Big\{\Big(\prod_{g^{-A}=0}^{1}{}_{n_U^{A*}(1,g^{-A})}C_{\bar{n}^A(1,g^{-A})}\Big)\div{}_{n_U^{A*}}C_{\bar{n}^A}\Big\}\bar{n}^{TC}\quad(\because\text{Equation 247},\mathbb{S}^*_{\max}(n_U^A=\bar{n}^A)\neq\emptyset)\\
&=\sum_{\bar{n}^{TC}=1}^{\bar{n}^A}\Big[\Big\{{}_{n_U^{A*}(1,0)}C_{\bar{n}^A(1,0)}\frac{n_U^{A*}(1,1)}{\bar{n}^A(1,1)}{}_{n_U^{A*}(1,1)-1}C_{\bar{n}^A(1,1)-1}\Big\}\div\frac{n_U^{A*}}{\bar{n}^A}{}_{n_U^{A*}-1}C_{\bar{n}^A-1}\Big]\bar{n}^A(1,1)\\
&\quad(\because\text{Equation 249},n_U^{A*}(1,1)=n_P^*\geq1,\bar{n}^A(1,1)=\bar{n}^{TC}\geq1,\mathbb{S}^*_{\max}(n_U^A=\bar{n}^A)\neq\emptyset,n_U^{A*}\geq\bar{n}^A\geq1)\\
&=\bar{n}^A\frac{n_P^*}{n_U^{A*}}\sum_{\bar{n}^{TC\prime}=0}^{\bar{n}^{A\prime}}\Big(\prod_{g^{-A}=0}^{1}{}_{n_U^{A*\prime}(1,g^{-A})}C_{\bar{n}^{A\prime}(1,g^{-A})}\Big)\div{}_{n_U^{A*\prime}}C_{\bar{n}^{A\prime}}\\
&=\bar{n}^A\frac{n_P^*}{n_U^{A*}}\sum_{\bar{n}^{TC\prime}=0}^{\bar{n}^{A\prime}}\Pr\{n_P=\bar{n}^{TC\prime}|\boldsymbol{S}^{*\prime}\in\mathbb{S}^{*\prime}_{\max}(n_U^A=\bar{n}^{A\prime})\}\quad(\because\text{Equation 247},\mathbb{S}^{*\prime}_{\max}(n_U^A=\bar{n}^{A\prime})\neq\emptyset)\\
&=\bar{n}^A\frac{n_P^*}{n_U^{A*}}.\quad(\because\text{the axiom of the probability})
\end{aligned}
\tag{250}
$$

where, in the last three lines, we suppose an alternative setup where $n_U^{A*\prime}(1,1)\equiv n_U^{A*}(1,1)-1\geq \bar{n}^{A\prime}(1,1)\equiv\bar{n}^A(1,1)-1\geq0$ and $n_U^{A*\prime}(1,0)\equiv n_U^{A*}(1,0)\geq\bar{n}^{A\prime}(1,0)\equiv\bar{n}^A(1,0)$ (and, thus, $n_U^{A*\prime}= n_U^{A*}-1\geq\bar{n}^{A\prime}=\bar{n}^A-1\geq0,\mathbb{S}^{*\prime}_{\max}(n_U^A=\bar{n}^{A\prime})\neq\emptyset$). Note that, since $\mathbb{S}^*_{\max}(n_U^A=\bar{n}^A)\neq\emptyset,\bar{n}^A\geq1$, it follows $n_U^{A*}\geq n_U^A=\bar{n}^A\geq1$.

When $n_P^*=0$,

$$
\begin{aligned}
n_P&=0\\
\therefore\mathbb{E}^*_{S|N}\{n_P|\mathbb{S}^*_{\max}(n_U^A=\bar{n}^A)\}&=0.
\end{aligned}
$$

Thus, both ends of Equation 250 are equal to each other (zero).

Therefore,

$$
\begin{aligned}
&\mathbb{E}^*_{S|N}\{n_P|\mathbb{S}_U^{-A1*}(n_U^A=\bar{n}^A)\}\\
&=\sum_{\bar{n}^{TC}=0}^{\bar{n}^A}\Pr\{n_P=\bar{n}^{TC}|\boldsymbol{S}^*\in\mathbb{S}_U^{-A1*}(n_U^A=\bar{n}^A)\}\bar{n}^{TC}\quad(\because\text{Equation 202, Lemma 23 (2)})\\
&=\sum_{\bar{n}^{TC}=1}^{\bar{n}^A}\frac{1}{p_N^{-A1}}\Pr\{n_P=\bar{n}^{TC}|\boldsymbol{S}^*\in\mathbb{S}^*_{\max}(n_U^A=\bar{n}^A)\}\bar{n}^{TC}+0\\
&\quad(\because\text{Equation 248},\mathbb{S}_U^{-A1*}(n_U^A=\bar{n}^A)\neq\emptyset,\bar{n}^{TC}\geq1)\\
&=\frac{1}{p_N^{-A1}}\sum_{\bar{n}^{TC}=0}^{\bar{n}^A}\Pr\{n_P=\bar{n}^{TC}|\boldsymbol{S}^*\in\mathbb{S}^*_{\max}(n_U^A=\bar{n}^A)\}\bar{n}^{TC}\\
&=\frac{1}{p_N^{-A1}}\bar{n}^A\frac{n_P^*}{n_U^{A*}}.\quad(\because\text{Equation 250})
\end{aligned}
\tag{251}
$$

It follows

$$\begin{aligned}\mathbb{E}^*_{S|N}\Big\{\frac{n_P}{n_U^A}\Big|\mathbb{S}_U^{-A1*}(n_U^A=\bar{n}^A)\Big\} &= \frac{1}{\bar{n}^A}\mathbb{E}^*_{S|N}\{n_P|\mathbb{S}_U^{-A1*}(n_U^A=\bar{n}^A)\} \quad (\because \text{Lemma 10 (2)}, \bar{n}^A\geq 1)\\ &= \frac{1}{\bar{n}^A}\frac{1}{p_N^{-A1}}\bar{n}^A\frac{n_P^*}{n_U^{A*}} \quad (\because \text{Equation 251})\\ &= \frac{1}{p_N^{-A1}}\frac{n_P^*}{n_U^{A*}}.\end{aligned} \tag{252}$$

When $\mathbb{S}_U^{1*}\neq\emptyset$,

$$\begin{aligned}\mathbb{E}^*\Big(\frac{n_P}{n_U^A}\Big|\mathbb{S}_U^{1*}\Big) &= \mathbb{E}^*_N\Big[\mathbb{E}^*_{S|N}\Big\{\frac{n_P}{n_U^A}\Big|\mathbb{S}_U^{-A1*}(n_U^A)\Big\}\Big|\mathbb{S}_U^{1*}\Big] \quad (\because \text{Lemma 23 (2)}, \mathbb{S}_U^{-A1*}(n_U^A)\subseteq\mathbb{S}_U^{1*})\\ &= \mathbb{E}^*\Big(\frac{1}{p_N^{-A1}}\frac{n_P^*}{n_U^{A*}}\Big|\mathbb{S}_U^{1*}\Big) \quad (\because \text{Equations 203 and 252}, \mathbb{S}_U^{-A1*}\neq\emptyset, p_N^{-A1}\neq 0)\\ &\geq \mathbb{E}^*\Big(\frac{n_P^*}{n_U^{A*}}\Big|\mathbb{S}_U^{1*}\Big) \quad (\because \text{Equation 243})\\ &= \frac{n_P^*}{n_U^{A*}}, \quad (\because \mathbb{S}_U^{1*}\neq\emptyset, \text{Lemma 23 (3)})\end{aligned} \tag{253}$$

where equality holds if and only if, for all $\boldsymbol{S}^*\in\mathbb{S}_U^{1*}$ (i.e., for all values n_U^A can take), it holds $p_N^{-A1}=1$, that is, for some $g^A\in\{0,1\}$, it holds $n_U^{A*}(g^A,0)<n_U^A(g^A)$ (Case 2 above). Since it holds $\min\{n_U^A(1)\}=\max\{1,n-n_U^{A*}(0)\}$, $\min\{n_U^A(0)\}=\max\{0,n-1-n_U^{A*}(1)\}$ and $n_U^{A*}(g^A,0)\geq 0$, the condition is equivalent to $n_U^{A*}(1,0)=0$ or $n_U^{A*}(1,0)<n-n_U^{A*}(0)$ or $n_U^{A*}(0,0)<n-1-n_U^{A*}(1)$, where the last two conditions are summarized as

$$\begin{aligned}n_U^{A*}(g^A,0)+n_U^{A*}(1-g^A) &= n_U^{A*}(g^A,0)+n_U^{A*}(g^A,1)+n_U^{A*}(1-g^A)-n_U^{A*}(g^A,1)\\ &= n^*-n_U^{A*}(g^A,1)\\ &< n-(1-g^A).\end{aligned}$$

In limit, when $p_U^{A*}>0$,

$$\begin{aligned}\lim_{n^*\to\infty}\mathbb{E}^*\Big(\frac{n_P}{n_U^A}\Big|\mathbb{S}_U^{1*}\Big) &= \mathbb{E}^*\Big(\lim_{n^*\to\infty}\frac{1}{p_N^{-A1}}\lim_{n^*\to\infty}\frac{n_P^*}{n_U^{A*}}\Big|\mathbb{S}_U^{1*}\Big) \quad (\because \text{Equation 253})\\ &\geq \lim_{n^*\to\infty}\frac{n_P^*}{n_U^{A*}} \quad (\because \text{Equations 243}, \mathbb{S}_U^{1*}\neq\emptyset, \text{Lemma 23 (3)})\\ &= \lim_{n^*\to\infty}\frac{n_P^*}{n^*}\div\lim_{n^*\to\infty}\frac{n_U^{A*}}{n^*} \quad (\because p_U^{A*}>0)\\ &= \frac{p_P^*}{p_U^{A*}},\end{aligned}$$

where equality holds if and only if, for all $\boldsymbol{S}^*\in\mathbb{S}_U^{1*}$, it holds $\lim_{n^*\to\infty}p_N^{-A1}=1$, that is, $p_U^{A*}(1,0)=0$ ($\because p_U^{A*}>0, \Pr(n_U^A=n|\mathbb{S}_U^{1*})>0$ and, when $p_U^{A*}<1, n_U^A=n$, it does not suffice that $p_U^{A*}(0,0)=0$), in which case it holds

$$\begin{aligned}\lim_{n^*\to\infty}\mathbb{E}^*\Big(\frac{n_P}{n_U^A}\Big|\mathbb{S}_U^{1*}\Big) &= \frac{p_P^*}{p_U^{A*}}\\ &= 1.\end{aligned}$$

Equation 253 is the first inequality in Equation 224, where $A=T$. By subtracting 1 from both sides and multiplying them by -1, we obtain the second inequality.

3.3. Variance

Suppose that $\boldsymbol{K}_G^* \in \mathbb{U}^*$ satisfies Condition 1*. Define expectation of variance between between $\boldsymbol{Q}^{(1)*}$ and $\boldsymbol{Q}^{(2)*}$ weighted by $\boldsymbol{k}_G^{A*}$ and adjusted by $\bar{\mathbf{n}} \in \mathbb{N}_{\max}$ as, when $\bar{n}^A \geq 1$,

$$EV_{N,h=1}^{2*}(\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A*}, \bar{\mathbf{n}}) \equiv \sum_{g^{-A}=0}^{1} \frac{\bar{n}^A(1, g^{-A})}{\bar{n}^A} V_{h=1}^{2*}\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A*}(1, g^{-A})\} \tag{254}$$

and, when $\bar{n}^A = 0$,

$$EV_{N,h=1}^{2*}(\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A*}, \bar{\mathbf{n}}) \equiv 0.$$

I also define, in the case of $\bar{n}^{TC} \geq 1$,

$$\begin{aligned}
&EV_{N,h=1}^{2*}\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{TC*}, \bar{\mathbf{n}}\} \\
&\equiv \sum_{g^{-TC}=0}^{1} \frac{\bar{n}^{TC}(1, g^{-TC})}{\bar{n}^{TC}} V_{h=1}^{2*}\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{TC*}(1, g^{-TC})\} \\
&= \frac{\bar{n}^{TC}(1,1)}{\bar{n}^{TC}} V_{h=1}^{2*}\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{TC*}(1,1)\} + \frac{\bar{n}^{TC}(1,0)}{\bar{n}^{TC}} V_{h=1}^{2*}\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{TC*}(1,0)\} \\
&= \frac{\bar{n}^{TC}}{\bar{n}^{TC}} V_{h=1}^{2*}(\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{TC*}) + \frac{0}{\bar{n}^{TC}} V_{h=1}^{2*}(\boldsymbol{Q}^{(h)*}|\mathbf{0}^*) \quad (\because \text{Equations 186 and 187}) \\
&= V_{h=1}^{2*}(\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{TC*}) \quad (\because \text{Equation 91})
\end{aligned} \tag{255}$$

and, in the case of $\bar{n}^{TC} < 1$,

$$EV_{N,h=1}^{2*}\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{TC*}, \bar{\mathbf{n}}\} \equiv 0.$$

Define, for non-negative integers q^* and q,

$$d(q^*, q) \equiv \begin{cases} \dfrac{q-1}{q^*-1} & \text{if } q^* \geq 2 \\ 0 & \text{if } q^* < 2 \end{cases} \tag{256}$$

and degree-of-freedom adjusted covariance as

$$\begin{aligned}
\widetilde{V_{h=1}^{2*}}\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A*}(1, g^{-A}), \bar{\mathbf{n}}\} &\equiv [1 - d\{n_G^{A*}(1, g^{-A}), \bar{n}^A(1, g^{-A})\}] V_{h=1}^{2*}\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A*}(1, g^{-A})\} \\
\widetilde{V_{h=1}^{2*}}\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{TC*}(1, g^{-TC}), \bar{\mathbf{n}}\} &\equiv [1 - d\{n_G^{TC*}(1, g^{-TC}), \bar{n}^{TC}(1, g^{-TC})\}] V_{h=1}^{2*}\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{TC*}(1, g^{-TC})\}
\end{aligned} \tag{257}$$

Define degree-of-freedom adjusted expectation of variance as, in the case of $\bar{n}^A \geq 1$,

$$\widetilde{EV_{N,h=1}^{2*}}(\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A*}, \bar{\mathbf{n}}) \equiv \sum_{g^{-A}=0}^{1} \frac{\bar{n}^A(1, g^{-A})}{\bar{n}^A} \widetilde{V_{h=1}^{2*}}\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A*}(1, g^{-A}), \bar{\mathbf{n}}\}, \tag{258}$$

in the case of $\bar{n}^A < 1$,

$$\widetilde{EV_{N,h=1}^{2*}}(\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A*}, \bar{\mathbf{n}}) \equiv 0,$$

in the case of $\bar{n}^{TC} \geq 1$,

$$\begin{aligned}
&\widetilde{EV_{N,h=1}^{2*}}\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{TC*}, \bar{\mathbf{n}}\} \\
&\equiv \sum_{g^{-TC}=0}^{1} \frac{\bar{n}^{TC}(1, g^{-TC})}{\bar{n}^{TC}} \widetilde{V_{h=1}^{2*}}\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{TC*}(1, g^{-TC}), \bar{\mathbf{n}}\} \\
&= \widetilde{V_{h=1}^{2*}}(\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{TC*}, \bar{\mathbf{n}}) \quad (\because \text{Equations 91, 186, and 187}) \\
&= \{1 - d(n_G^{TC*}, \bar{n}^{TC})\} V_{h=1}^{2*}(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}_G^{TC*}) \quad (\because \text{Equations 186, 187, and 257})
\end{aligned} \tag{259}$$

and, in the case of $\bar{n}^{TC} < 1$,

$$\widetilde{EV^{2*}_{N,h=1}}\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}^{TC*}_G, \bar{\mathbf{n}}\} \equiv 0.$$

Therefore,

$$\bar{n}^{TC}\widetilde{EV^{2*}_{N,h=1}}\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}^{TC*}_G, \bar{\mathbf{n}}\} = \bar{n}^{TC}\widetilde{V^{2*}_{h=1}}(\boldsymbol{Q}^{(h)*}|\boldsymbol{k}^{TC*}_G, \bar{\mathbf{n}}) \tag{260}$$

Note that $\mathbb{V}^{2*}_{h=1}\{f^{(h)*}_{X,S}(\cdot)|\mathbb{S}^*\}$ can be defined only when $\mathbb{S}^* \subseteq \bigcap^2_{h=1}\mathbb{S}^*_{\text{def}}(f^{(h)*}_{X,S})$. Note

$$\begin{aligned}\mathbb{V}^{2*}_{h=1}\{f^{(h)*}_{X,S}(\cdot)|\emptyset\} &= \mathbb{E}^*\Big[\prod^2_{h=1}\mathbb{D}^{(h)*}\{f^{(h)*}_{X,S}(\cdot)|\emptyset\}\Big|\emptyset\Big] \quad (\because \text{Equation 93})\\ &= \mathbb{E}^*\Big\{\prod^2_{h=1}f^{(h)*}_{X,S}(\cdot)\Big|\emptyset\Big\} \quad (\because \text{Equations 92 and 201})\\ &= 0 \quad (\because \text{Equation 201})\end{aligned} \tag{261}$$

Define

$$\begin{aligned}n^{A(1)A(2)}_G &\equiv \sum_j\prod^2_{h=1}k^{A(h)}_{G,\cdot j}\\ &= \begin{cases} n^{TC}_G & \text{when } A(1) \neq A(2)\\ n^A_G & \text{when } A(1) = A(2) \equiv A\end{cases}\end{aligned} \tag{262}$$

Accordingly, define

$$\bar{n}^{A(1)A(2)} \equiv \begin{cases} \bar{n}^{TC} & \text{when } A(1) \neq A(2)\\ \bar{n}^A & \text{when } A(1) = A(2) \equiv A\end{cases}$$

Denote

$$\Delta E^*(\boldsymbol{Q}^*|\boldsymbol{k}^{A*}_G) \equiv E^*\{\boldsymbol{Q}^*|\boldsymbol{k}^{A*}_G(1,1)\} - E^*\{\boldsymbol{Q}^*|\boldsymbol{k}^{A*}_G(1,0)\}. \tag{263}$$

LEMMA 27 (SP COVARIANCE OF FS MEAN). *Let $A(h) \in \{T, C\}$ for $h \in \{1,2\}$. Suppose that $\boldsymbol{K}^*_G \in \mathbb{U}^*$ satisfies Condition 1^*, $\mathbb{N} \subseteq \bigcap^2_{h=1}\mathbb{N}^{A(h)1}$, $\mathbb{S}^*_G \equiv \mathbb{S}^*_{\max}(\mathbf{n}_G \in \mathbb{N})$, $\boldsymbol{\beta}^{(h)*} \in \mathbb{B}^*$, and $\boldsymbol{\omega}^{(h)*} \in \mathbb{W}^*$ for $h \in \{1,2\}$.*

(1)

$$\begin{aligned}&\mathbb{V}^{2*}_{h=1}\{E^*(\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}^{A(h)*}_G)|\mathbb{S}^*_G\}\\ &= \mathbb{E}^*\Big\{\frac{n^{A(1)A(2)}_G}{\prod^2_{h=1}n^{A(h)}_G}\widetilde{EV^{2*}_{N,h=1}}\Big(\boldsymbol{\beta}^{(h)*}\Big|\prod^2_{h=1}\boldsymbol{k}^{A(h)*}_G, \mathbf{n}_G\Big)\Big|\mathbb{S}^*_G\Big\} + \mathbb{V}^{2*}_{h=1}\Big(\frac{n^{TC}_G}{n^{A(h)}_G}\Big|\mathbb{S}^*_G\Big)\prod^2_{h=1}\Delta E^*(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A(h)*}_G)\end{aligned}$$

(2)

$$\begin{aligned}&\mathbb{V}^{2*}_{h=1}\{E(\boldsymbol{\omega}^{(h)}|\boldsymbol{k}^{A(h)}_G\boldsymbol{X}^{A(h)})|\mathbb{S}^*_G\}\\ &= [2\cdot I\{A(1) = A(2)\} - 1]\mathbb{E}^*\Big\{\frac{n^{A(1)A(2)}_G}{\prod^2_{h=1}n^{A(h)}_G}EV^{2*}_{N,h=1}\Big(\boldsymbol{\omega}^{(h)*}\Big|\prod^2_{h=1}\boldsymbol{k}^{A(h)*}_G, \mathbf{n}_G\Big)\Big|\mathbb{S}^*_G\Big\}\end{aligned}$$

PROOF. Given $g^{-A(h)} \in \{0,1\}$ for all $h \in \{1,2\}$, denote

$$\mathbb{J}^{(h)*} \equiv \mathbb{J}^{A(h)*}_G(1, g^{-A(h)}) \tag{264}$$

Note $\mathbb{J}^{(1)*} = \mathbb{J}^{(2)*}$ if and only if $A(1) = A(2)$ or $g^{-A(1)} = g^{-A(2)} = 1$ because, in the case of $A(1) = A(2) = A$, obviously,

$$\boldsymbol{k}^{A(1)*}_G(1, g^{-A(1)}) = \boldsymbol{k}^{A(2)*}_G(1, g^{-A(2)}) = \boldsymbol{k}^{A*}_G(1, g^{-A}),$$

in the case of $g^{-A(1)} = g^{-A(2)} = 1$,

$$\boldsymbol{k}_G^{A(1)*}(1, g^{-A(1)}) = \boldsymbol{k}_G^{A(2)*}(1, g^{-A(2)}) = \boldsymbol{k}_G^{*}(1,1) \equiv \boldsymbol{k}_G^{A*}(1, g^{-A}),$$

and, otherwise,

$$\begin{aligned}
\prod_{h=1}^{2} \boldsymbol{k}_G^{A(h)*}(1, g^{-A(h)}) &= \boldsymbol{k}_G^{T*}(1, g^C)\boldsymbol{k}_G^{C*}(1, g^T) \quad (\because A(1) \neq A(2)) \\
&= \{\boldsymbol{k}_G^{T*}(\boldsymbol{k}_G^{C*})^{g^C}(\mathbf{1} - \boldsymbol{k}_G^{C*})^{1-g^C}\}\{\boldsymbol{k}_G^{C*}(\boldsymbol{k}_G^{T*})^{g^T}(\mathbf{1} - \boldsymbol{k}_G^{T*})^{1-g^T}\} \\
&= \boldsymbol{k}_G^{T*}(\mathbf{1} - \boldsymbol{k}_G^{T*})^{1-g^T}\boldsymbol{k}_G^{C*}(\mathbf{1} - \boldsymbol{k}_G^{C*})^{1-g^C} \quad (\because \boldsymbol{k}_G^{A*}(\boldsymbol{k}_G^{A*})^{g^A} = \boldsymbol{k}_G^{A*}) \\
&= \mathbf{0} \quad (\because g^T = 0 \text{ or } g^C = 0) \\
\therefore \bigcap_{h=1}^{2} \mathbb{J}^{(h)*} &= \emptyset.
\end{aligned}$$

When $\mathbb{J}^{(1)*} = \mathbb{J}^{(2)*}$, for $h \in \{1, 2\}$, we define

$$\begin{aligned}
n_G^{A*}(1, g^{-A}) &\equiv n_G^{A(1)*}(1, g^{-A(1)}) = n_G^{A(2)*}(1, g^{-A(2)}) \\
\bar{n}^{A}(1, g^{-A}) &\equiv \bar{n}^{A(1)}(1, g^{-A(1)}) = \bar{n}^{A(2)}(1, g^{-A(2)}).
\end{aligned}$$

For any $\mathbb{S}^* \subseteq \mathbb{S}^*_{\max}$,

$$\begin{aligned}
\mathbb{E}^*\Big(\prod_{h=1}^{2} S^*_{\cdot j^*(h)}\Big|\mathbb{S}^*\Big) &= \sum_{\boldsymbol{s}^* \in \mathbb{S}^*} \Pr(\boldsymbol{S}^* = \boldsymbol{s}^*\Big|\boldsymbol{S}^* \in \mathbb{S}^*)\prod_{h=1}^{2} s^*_{\cdot j^*(h)} \quad (\because \text{Equation 200}) \\
&= \sum_{s^{(1)}=0}^{1}\sum_{s^{(2)}=0}^{1} \Pr(S^*_{\cdot j^*(1)} = s^{(1)}, S^*_{\cdot j^*(2)} = s^{(2)}|\boldsymbol{S}^* \in \mathbb{S}^*)\prod_{h=1}^{2} s^{(h)} \qquad (265) \\
&= \Pr\Big(\prod_{h=1}^{2} S^*_{\cdot j^*(h)} = 1\Big|\boldsymbol{S}^* \in \mathbb{S}^*\Big)
\end{aligned}$$

Given $\bar{\mathbf{n}} \in \mathbb{N}$, denote $\mathbb{S}^*_N \equiv \mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}})$. We define

$$\overline{S^2}_{EQ} \equiv \begin{cases} \dfrac{\bar{n}^A(1, g^{-A})}{n_G^{A*}(1, g^{-A})} & \text{if } \mathbb{J}^{(1)*} = \mathbb{J}^{(2)*}, \mathbb{S}^*_N \neq \emptyset, n_G^{A*}(1, g^{-A}) \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\overline{S^2}_{DF} \equiv \begin{cases} \dfrac{\bar{n}^A(1, g^{-A})\{\bar{n}^A(1, g^{-A}) - 1\}}{n_G^{A*}(1, g^{-A})\{n_G^{A*}(1, g^{-A}) - 1\}} & \text{if } \mathbb{J}^{(1)*} = \mathbb{J}^{(2)*}, \mathbb{S}^*_N \neq \emptyset, n_G^{A*}(1, g^{-A}) \geq 2 \\ \displaystyle\prod_{h=1}^{2} \frac{\bar{n}^{A(h)}(1, g^{-A(h)})}{n_G^{A(h)*}(1, g^{-A(h)})} & \text{if } \mathbb{J}^{(1)*} \neq \mathbb{J}^{(2)*}, \mathbb{S}^*_N \neq \emptyset, n_G^{A(h)*}(1, g^{-A(h)}) \geq 1 \\ 0 & \text{otherwise} \end{cases} \qquad (266)$$

When $j^*(h) \in \mathbb{J}^{(h)*}$ for $h \in \{1, 2\}$, according to Equations 265 and 266 and Lemma 22, it holds

$$\mathbb{E}^*_{S|N}\Big\{\prod_{h=1}^{2} S^*_{\cdot j^*(h)}\Big|\mathbb{S}^*_N\Big\} = \begin{cases} \overline{S^2}_{EQ} & \text{if } j^*(1) = j^*(2) \\ \overline{S^2}_{DF} & \text{if } j^*(1) \neq j^*(2). \end{cases} \qquad (267)$$

Note that, when $\bar{n}^{A(h)}(1, g^{-A(h)}) \geq 1$ for any $h \in \{1, 2\}$,

$$
\begin{aligned}
& \mathbb{E}^*_{S|N}\Big[\prod_{h=1}^{2} \bar{n}^{A(h)}(1, g^{-A(h)}) E^*\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{S}^* \boldsymbol{k}_G^{A(h)*}(1, g^{-A(h)})\}\Big| \mathbb{S}^*_N\Big] \\
&= \mathbb{E}^*_{S|N}\Big\{\prod_{h=1}^{2} \bar{n}^{A(h)}(1, g^{-A(h)}) \frac{\sum_{j^*(h)} S^*_{\cdot j^*(h)} k^{A(h)*}_{G,\cdot j^*(h)}(1, g^{-A(h)}) \beta^{(h)*}_{\cdot j^*(h)}}{\bar{n}^{A(h)}(1, g^{-A(h)})}\Big| \mathbb{S}^*_N\Big\} \\
& (\because \text{Lemma 12 (2), Equation 182}, \mathbf{n}_G = \bar{\mathbf{n}}) \\
&= \mathbb{E}^*_{S|N}\Big\{\sum_{j^*(1)} \sum_{j^*(2)} \prod_{h=1}^{2} S^*_{\cdot j^*(h)} k^{A(h)*}_{G,\cdot j^*(h)}(1, g^{-A(h)}) \beta^{(h)*}_{\cdot j^*(h)}\Big| \mathbb{S}^*_N\Big\} \quad (\because \text{Lemma 16 (1)} \\
&= \mathbb{E}^*_{S|N}\Big\{\sum_{g^{A(1)\prime}=0}^{1} \sum_{g^{-A(1)\prime}=0}^{1} \sum_{g^{A(2)\prime}=0}^{1} \sum_{g^{-A(2)\prime}=0}^{1} \sum_{j^*(1) \in \mathbb{J}_G^{A(1)*}(g^{A(1)\prime}, g^{-A(1)\prime})} \sum_{j^*(2) \in \mathbb{J}_G^{A(2)*}(g^{A(2)\prime}, g^{-A(2)\prime})} \\
& \prod_{h=1}^{2} S^*_{\cdot j^*(h)} k^{A(h)*}_{G,\cdot j^*(h)}(1, g^{-A(h)}) \beta^{(h)*}_{\cdot j^*(h)}\Big| \mathbb{S}^*_N\Big\} \\
&= \mathbb{E}^*_{S|N}\Big\{\sum_{j^*(1) \in \mathbb{J}^{(1)*}} \sum_{j^*(2) \in \mathbb{J}^{(2)*}} \prod_{h=1}^{2} S^*_{\cdot j^*(h)} k^{A(h)*}_{G,\cdot j^*(h)}(1, g^{-A(h)}) \beta^{(h)*}_{\cdot j^*(h)}\Big| \mathbb{S}^*_N\Big\} \\
& (\because \prod_{h=1}^{2} k^{A(h)*}_{G,\cdot j^*(h)}(1, g^{-A(h)}) = 0 \text{ for } j^*(h) \notin \mathbb{J}^{(h)*}) \\
&= \{\overline{S^2}_{EQ} - \overline{S^2}_{DF}\} \sum_{j^* \in \bigcap_{h=1}^{2} \mathbb{J}^{(h)*}} \prod_{h=1}^{2} k^{A(h)*}_{G,\cdot j^*}(1, g^{-A(h)}) \beta^{(h)*}_{\cdot j^*} \\
& \quad + \overline{S^2}_{DF} \prod_{h=1}^{2} \sum_{j^*(h) \in \mathbb{J}^{(h)*}} k^{A(h)*}_{G,\cdot j^*(h)}(1, g^{-A(h)}) \beta^{(h)*}_{\cdot j^*(h)} \quad (\because \text{Lemma 16 (2), Equation 267} \\
&= \{\overline{S^2}_{EQ} - \overline{S^2}_{DF}\} \sum_{j^*} \prod_{h=1}^{2} k^{A(h)*}_{G,\cdot j^*}(1, g^{-A(h)}) \beta^{(h)*}_{\cdot j^*} + \overline{S^2}_{DF} \prod_{h=1}^{2} \sum_{j^*(h)} k^{A(h)*}_{G,\cdot j^*(h)}(1, g^{-A(h)}) \beta^{(h)*}_{\cdot j^*(h)} \\
& (\because k^{A(h)*}_{G,\cdot j^*(h)}(1, g^{-A(h)}) = 0 \text{ for } j^*(h) \notin \mathbb{J}^{(h)*})
\end{aligned}
\tag{268}
$$

where, in applying Lemma 16, we substitute

$$
\begin{aligned}
l(h) &= j^*(h) \\
\mathbb{L}(h) &= \{1, 2, \ldots, n^*\} \text{ or } \mathbb{J}^{(h)*} \\
q^{(h)}_{l(h)} &= k^{A(h)*}_{G,\cdot j^*(h)}(1, g^{-A(h)}) \beta^{(h)*}_{\cdot j^*(h)} \\
f\{X^{A(h)}_{l(h)}\} &= S^*_{\cdot j^*(h)} \\
\mathbb{E}(\cdot) &= \mathbb{E}^*_{S|N}(\cdot | \mathbb{S}^*_N) \\
\overline{f^2}_{EQ} &= \overline{S^2}_{EQ} \\
\overline{f^2}_{DF} &= \overline{S^2}_{DF} \\
\mathbb{X}_{\text{def}}\{f_X(\boldsymbol{Q})\} &= \mathbb{S}_{\text{def}}\{f^*_{X,S}(\boldsymbol{Q}^*)\}
\end{aligned}
$$

and Equation 120 holds because, for any $j^*(h) \in \mathbb{J}^{(h)*}$,

$$
\begin{aligned}
&\mathbb{S}_{\text{def}}\Big(\sum_{j^{*\prime}(1)\in\mathbb{J}^{(1)*}} \sum_{j^{*\prime}(2)\in\mathbb{J}^{(2)*}} \prod_{h=1}^{2} S^*_{\cdot j^{*\prime}(h)} k^{A(h)*}_{G,\cdot j^{*\prime}(h)}(1, g^{-A(h)}) \beta^{(h)*}_{\cdot j^{*\prime}(h)} \Big) \\
&= \mathbb{S}_{\text{def}}\Big(\prod_{h=1}^{2} S^*_{\cdot j^*(h)} k^{A(h)*}_{G,\cdot j^*(h)}(1, g^{-A(h)}) \beta^{(h)*}_{\cdot j^*(h)} \Big) \\
&= \mathbb{S}^*_{\max} \\
&\supset \mathbb{S}^*_N.
\end{aligned}
$$

Note that both ends of Equation 268 are equal to each other (zero) even when $\bar{n}^{A(h)}(1, g^{-A(h)}) < 1$ for some $h \in \{1, 2\}$ because $\overline{S^2}_{EQ} = \overline{S^2}_{DF} = 0$.

For a moment, we suppose $\mathbb{S}^*_N \neq \emptyset$. When $\mathbb{J}^{(1)*} \neq \mathbb{J}^{(2)*}, n^{A(h)*}_G(1, g^{-A(h)}) \geq 1$ for $h \in \{1, 2\}$, Equation 268 is equal to

$$
\begin{aligned}
&\overline{S^2}_{DF} \prod_{h=1}^{2} \sum_{j^*(h)} k^{A(h)*}_{G,\cdot j^*(h)}(1, g^{-A(h)}) \beta^{(h)*}_{\cdot j^*(h)} \quad (\because \prod_{h=1}^{2} k^{A(h)*}_{G,\cdot j^*}(1, g^{-A(h)}) = 0) \\
&= \Big\{ \prod_{h=1}^{2} \frac{\bar{n}^{A(h)}(1, g^{-A(h)})}{n^{A(h)*}_G(1, g^{-A(h)})} \Big\} \prod_{h=1}^{2} n^{A(h)*}_G(1, g^{-A(h)}) E^*\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{k}^{A(h)*}_G(1, g^{-A(h)})\} \\
&(\because \text{Equations 165 and 266}) \\
&= \prod_{h=1}^{2} \bar{n}^{A(h)}(1, g^{-A(h)}) E^*\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{k}^{A(h)*}_G(1, g^{-A(h)})\}
\end{aligned} \tag{269}
$$

Note that both ends of the equation are equal to each other (zero) even when $n^{A(h)*}_G(1, g^{-A(h)}) < 1$ for some $h \in \{1, 2\}$ because

$$
\begin{aligned}
\overline{S^2}_{DF} &= 0 \quad (\because \text{Equation 266}) \\
\boldsymbol{k}^{A(h)*}_G(1, g^{-A(h)}) &= \mathbf{0}^* \quad (\because \text{Lemma 14 (4)}) \\
E^*\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{k}^{A(h)*}_G(1, g^{-A(h)})\} &= 0 \quad (\because \text{Equation 166}) \\
\bar{n}^{A(h)}(1, g^{-A(h)}) &= 0 \quad (\because \text{ Equation 208})
\end{aligned} \tag{270}
$$

When $\mathbb{J}^{(1)*} = \mathbb{J}^{(2)*}$, Equation 268 is equal to

$$\begin{aligned}
&(\overline{S^2}_{EQ} - \overline{S^2}_{DF}) \sum_{j^*} k^{A*}_{G,\cdot j^*}(1, g^{-A}) \prod_{h=1}^{2} \beta^{(h)*}_{\cdot j^*} + \overline{S^2}_{DF} \prod_{h=1}^{2} \sum_{j^*(h)} k^{A*}_{G,\cdot j^*(h)}(1, g^{-A}) \beta^{(h)*}_{\cdot j^*(h)} \\
&(\because \text{Lemma 14 (3)}) \\
&= (\overline{S^2}_{EQ} - \overline{S^2}_{DF}) n^{A*}_{G}(1, g^{-A}) E^*\Big\{ \prod_{h=1}^{2} \boldsymbol{\beta}^{(h)*} \Big| \boldsymbol{k}^{A*}_{G}(1, g^{-A}) \Big\} \\
&\quad + \overline{S^2}_{DF} \prod_{h=1}^{2} n^{A*}_{G}(1, g^{-A}) E^*\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{k}^{A*}_{G}(1, g^{-A})\} \quad (\because \text{Lemma 12 (2), Equation 181}) \\
&= (\overline{S^2}_{EQ} - \overline{S^2}_{DF}) n^{A*}_{G}(1, g^{-A}) \Big[V^{2*}_{h=1}\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{k}^{A*}_{G}(1, g^{-A})\} + \prod_{h=1}^{2} E^*\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{k}^{A*}_{G}(1, g^{-A})\} \Big] \\
&\quad + \overline{S^2}_{DF} \{n^{A*}_{G}(1, g^{-A})\}^2 \prod_{h=1}^{2} E^*\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{k}^{A*}_{G}(1, g^{-A})\} \quad (\because \text{Lemma 15 (7)}) \\
&= n^{A*}_{G}(1, g^{-A}) \Big((\overline{S^2}_{EQ} - \overline{S^2}_{DF}) V^{2*}_{h=1}\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{k}^{A*}_{G}(1, g^{-A})\} \\
&\quad + [\overline{S^2}_{EQ} + \overline{S^2}_{DF}\{n^{A*}_{G}(1, g^{-A}) - 1\}] \prod_{h=1}^{2} E^*\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{k}^{A*}_{G}(1, g^{-A})\} \Big)
\end{aligned} \tag{271}$$

When $n^{A*}_{G}(1, g^{-A}) \geq 2$, Equation 271 is equal to

$$\begin{aligned}
&n^{A*}_{G}(1, g^{-A}) \Big(\Big\{ \frac{\bar{n}^{A}(1, g^{-A})}{n^{A*}_{G}(1, g^{-A})} - \frac{\bar{n}^{A}(1, g^{-A})}{n^{A*}_{G}(1, g^{-A})} \frac{\bar{n}^{A}(1, g^{-A}) - 1}{n^{A*}_{G}(1, g^{-A}) - 1} \Big\} V^{2*}_{h=1}\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{k}^{A*}_{G}(1, g^{-A})\} \\
&\quad + \Big[\frac{\bar{n}^{A}(1, g^{-A})}{n^{A*}_{G}(1, g^{-A})} + \frac{\bar{n}^{A}(1, g^{-A})}{n^{A*}_{G}(1, g^{-A})} \frac{\bar{n}^{A}(1, g^{-A}) - 1}{n^{A*}_{G}(1, g^{-A}) - 1} \{n^{A*}_{G}(1, g^{-A}) - 1\} \Big] \\
&\quad \times \prod_{h=1}^{2} E^*\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{k}^{A*}_{G}(1, g^{-A})\} \Big) \quad (\because \text{Equation 266}) \\
&= \bar{n}^{A}(1, g^{-A}) \widetilde{V^{2*}_{h=1}}\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{k}^{A*}_{G}(1, g^{-A})\} + \prod_{h=1}^{2} \bar{n}^{A}(1, g^{-A}) E^*\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{k}^{A*}_{G}(1, g^{-A})\} \\
&(\because \text{Equation 257})
\end{aligned} \tag{272}$$

When $n^{A*}_{G}(1, g^{-A}) = 1$, it follows that $\bar{n}^{A}(1, g^{-A}) \in \{0, 1\}$. ($\because \mathbb{S}^*_N \neq \emptyset$, Equation 208.) There is only one such pair $j^{*(\text{one})}$ that $k^{A*}_{G}(1, g^{-A})_{\cdot j^{*(\text{one})}} = 1$. That is, for $j^* \neq j^{*(\text{one})}$, $k^{A*}_{G}(1, g^{-A})_{\cdot j^*} = 0$. It follows

$$\begin{aligned}
E^*\{\boldsymbol{Q}^{(h)*} | \boldsymbol{k}^{A*}_{G}(1, g^{-A})\} &= Q^{(h)*}_{\cdot j^{*(\text{one})}} \quad (\because \text{Equation 165}) \\
\therefore V^{2*}_{h=1}\{\boldsymbol{Q}^{(h)*} | \boldsymbol{k}^{A*}_{G}(1, g^{-A})\} &= \frac{(Q^{(h)*}_{\cdot j^{*(\text{one})}} - Q^{(h)*}_{\cdot j^{*(\text{one})}})^2}{1} \quad (\because \text{Equation 89}) \\
&= 0
\end{aligned} \tag{273}$$

Therefore, Equation 271 is equal to

$$
\begin{aligned}
&\bar{n}^A(1, g^{-A}) V_{h=1}^{2*}\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{k}_G^{A*}(1, g^{-A})\} + \bar{n}^A(1, g^{-A}) \prod_{h=1}^{2} E^*\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{k}_G^{A*}(1, g^{-A})\} \\
&(\because \text{Equation 266}) \\
&= \bar{n}^A(1, g^{-A}) \widetilde{V_{h=1}^{2*}}\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{k}_G^{A*}(1, g^{-A})\} + \prod_{h=1}^{2} \bar{n}^A(1, g^{-A}) E^*\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{k}_G^{A*}(1, g^{-A})\} \\
&(\because \text{Equation 257, } n_G^{A*}(1, g^{-A}) = 1, \bar{n}^A(1, g^{-A}) \in \{0, 1\}, \{\bar{n}^A(1, g^{-A})\}^2 = \bar{n}^A(1, g^{-A}))
\end{aligned}
\tag{274}
$$

When $n_G^{A*}(1, g^{-A}) = 0$, thanks to Equation 270, Equation 271 is equal to

$$
\bar{n}^A(1, g^{-A}) \widetilde{V_{h=1}^{2*}}\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{k}_G^{A*}(1, g^{-A})\} + \prod_{h=1}^{2} \bar{n}^A(1, g^{-A}) E^*\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{k}_G^{A*}(1, g^{-A})\} = 0. \tag{275}
$$

For reference, note also that

$$
\begin{aligned}
\overline{S^2}_{EQ} &= 0 \quad (\because \text{Equation 266}) \\
V_{h=1}^{2*}\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{k}_G^{A*}(1, g^{-A})\} &= 0 \quad (\because \text{Equations 89 and 166}) \\
\widetilde{V_{h=1}^{2*}}\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{k}_G^{A*}(1, g^{-A}), \bar{\mathbf{n}}\} &= 0 \quad (\because \text{Equation 257})
\end{aligned}
\tag{276}
$$

According to Equations 268, 269, 271, 272, 274, and 275, it follows that, when $\mathbb{S}_N^* \neq \emptyset$,

$$
\begin{aligned}
&\mathbb{E}_{S|N}^*\Big[\prod_{h=1}^{2} \bar{n}^{A(h)}(1, g^{-A(h)}) E^*\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{S}^* \boldsymbol{k}_G^{A(h)*}(1, g^{-A(h)})\} \Big| \mathbb{S}_N^*\Big] \\
&= \prod_{h=1}^{2} \bar{n}^{A(h)}(1, g^{-A(h)}) E^*\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{k}_G^{A(h)*}(1, g^{-A(h)})\} \\
&\quad + I(\mathbb{J}^{(1)*} = \mathbb{J}^{(2)*}) \bar{n}^A(1, g^{-A}) \widetilde{V_{h=1}^{2*}}\{\boldsymbol{\beta}^{(h)*} | \boldsymbol{k}_G^{A*}(1, g^{-A}), \bar{\mathbf{n}}\}.
\end{aligned}
\tag{277}
$$

Denote

$$
\begin{aligned}
\mathbb{D}_{S|N}^*\{f_S(\boldsymbol{S}^*) | \mathbb{S}_N^*\} &\equiv f_S(\boldsymbol{S}^*) - \mathbb{E}_{S|N}^*\{f_S(\boldsymbol{S}^*) | \mathbb{S}_N^*\} \\
\mathbb{V}_{S|N,h=1}^{2*}\{f_S^{(h)}(\boldsymbol{S}^*) | \mathbb{S}_N^*\} &\equiv \mathbb{E}_{S|N}^*\Big[\prod_{h=1}^{2} \mathbb{D}_{S|N}^*\{f_S^{(h)}(\boldsymbol{S}^*) | \mathbb{S}_N^*\}\Big].
\end{aligned}
\tag{278}
$$

When $\bar{n}^{A(h)}(1, g^{-A(h)}) \geq 1$ for all $h \in \{1, 2\}$, it follows that

$$\begin{aligned}
&\Big\{\prod_{h=1}^{2} \bar{n}^{A(h)}(1, g^{-A(h)})\Big\} \mathbb{V}^{2*}_{S|N,h=1}[E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}_G^{A(h)*}(1, g^{-A(h)})\}|\mathbb{S}^*_N] \\
&= \mathbb{E}^*_{S|N}\Big[\prod_{h=1}^{2} \bar{n}^{A(h)}(1, g^{-A(h)}) E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}_G^{A(h)*}(1, g^{-A(h)})\}\Big|\mathbb{S}^*_N\Big] \\
&\quad - \prod_{h=1}^{2} \bar{n}^{A(h)}(1, g^{-A(h)}) \mathbb{E}^*_{S|N}[E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}_G^{A(h)*}(1, g^{-A(h)})\}|\mathbb{S}^*_N] \quad (\because \text{Lemma 15 (7)}) \\
&= \prod_{h=1}^{2} \bar{n}^{A(h)}(1, g^{-A(h)}) E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}_G^{A(h)*}(1, g^{-A(h)})\} \\
&\quad + I(\mathbb{J}^{(1)*} = \mathbb{J}^{(2)*}) \bar{n}^{A}(1, g^{-A}) \widetilde{V^{2*}_{h=1}}\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}_G^{A*}(1, g^{-A}), \bar{\mathbf{n}}\} \\
&\quad - \prod_{h=1}^{2} \bar{n}^{A(h)}(1, g^{-A(h)}) E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}_G^{A(h)*}(1, g^{-A(h)})\} \\
&\quad (\because \text{Equation 277, Lemma 24 (1)}, \mathbb{S}^*_N \neq \emptyset, \bar{n}^{A(h)}(1, g^{-A(h)}) \geq 1) \\
&= I(\mathbb{J}^{(1)*} = \mathbb{J}^{(2)*}) \bar{n}^{A}(1, g^{-A}) \widetilde{V^{2*}_{h=1}}\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}_G^{A*}(1, g^{-A}), \bar{\mathbf{n}}\}.
\end{aligned} \tag{279}$$

and

$$\begin{aligned}
&\mathbb{D}^*_{S|N}\{E^*(\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}_G^{A(h)*})|\mathbb{S}^*_N\} \\
&= E^*(\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}_G^{A(h)*}) - \mathbb{E}^*_{S|N}\{E^*(\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}_G^{A(h)*})|\mathbb{S}^*_N\} \quad (\because \text{Equation 278}) \\
&= E^*(\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}_G^{A(h)*}) - E^*_N(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}_G^{A(h)*}, \bar{\mathbf{n}}) \quad (\because \text{Lemma 24 (2)}, \mathbb{S}^*_N \neq \emptyset) \\
&= \sum_{g^{-A(h)}=0}^{1} \frac{\bar{n}^{A(h)}(1, g^{-A(h)})}{\bar{n}^{A(h)}} [E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}_G^{A(h)*}(1, g^{-A(h)})\} - E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}_G^{A(h)*}(1, g^{-A(h)})\}] \\
&\quad (\because \text{Lemma 3 (5), where } \boldsymbol{Z}^{(1)} = \boldsymbol{S}^*\boldsymbol{k}^{A(h)*}, \boldsymbol{Z}^{(2)} = \boldsymbol{k}^{-A(h)*}, \text{and Equation 205, } \mathbb{N} \subseteq \mathbb{N}^{A(h)1}) \\
&= \sum_{g^{-A(h)}=0}^{1} \frac{\bar{n}^{A(h)}(1, g^{-A(h)})}{\bar{n}^{A(h)}} \mathbb{D}^*_{S|N}[E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}_G^{A(h)*}(1, g^{-A(h)})\}|\mathbb{S}^*_N] \\
&\quad (\because \text{Lemma 24 (1)}, \bar{n}^{A(h)}(1, g^{-A(h)}) \geq 1, \text{Equation 278})
\end{aligned} \tag{280}$$

Even when $n^{A(h)}(1, g^{-A(h)}) < 1$ for some $h \in \{1, 2\}$, both ends of Equation 279 (or 280) are equal to each other (zero).

When $A(1) \neq A(2)$,

$$
\begin{aligned}
&\mathbb{V}^{2*}_{S|N,h=1}\{E^*(\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}_G^{A(h)*})|\mathbb{S}^*_N\} \\
&= \mathbb{E}^*_{S|N}\Big[\prod_{h=1}^{2}\mathbb{D}^*_{S|N}\{E^*(\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}_G^{A(h)*})|\mathbb{S}^*_N\}\Big|\mathbb{S}^*_N\Big] \quad (\because \text{Equation 93}) \\
&= \mathbb{E}^*_{S|N}\Big(\prod_{h=1}^{2}\sum_{g^{-A(h)}=0}^{1}\frac{\bar{n}^{A(h)}(1,g^{-A(h)})}{\bar{n}^{A(h)}}\mathbb{D}^*_{S|N}[E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}_G^{A(h)*}(1,g^{-A(h)})\}|\mathbb{S}^*_N]\Big|\mathbb{S}^*_N\Big) \\
&(\because \text{Equation 280}) \\
&= \mathbb{E}^*_{S|N}\Big(\sum_{g^{-A(1)}=0}^{1}\sum_{g^{-A(2)}=0}^{1}\prod_{h=1}^{2}\frac{\bar{n}^{A(h)}(1,g^{-A(h)})}{\bar{n}^{A(h)}}\mathbb{D}^*_{S|N}[E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}_G^{A(h)*}(1,g^{-A(h)})\}|\mathbb{S}^*_N]\Big|\mathbb{S}^*_N\Big) \\
&(\because A(1)\neq A(2), \text{Lemma 16 (1), where } l(h)=g^{-A(h)}, \mathbb{L}(h)=\{0,1\} \\
&Q_{l(h)} = \frac{\bar{n}^{A(h)}(1,g^{-A(h)})}{\bar{n}^{A(h)}}\mathbb{D}^*_{S|N}[E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}_G^{A(h)*}(1,g^{-A(h)})\}|\mathbb{S}^*_N]) \\
&= \frac{1}{\prod_{h=1}^{2}\bar{n}^{A(h)}}\sum_{g^{-A(1)}=0}^{1}\sum_{g^{-A(2)}=0}^{1}\Big\{\prod_{h=1}^{2}\bar{n}^{A(h)}(1,g^{-A(h)})\Big\} \\
&\quad \mathbb{E}^*_{S|N}\Big\{\prod_{h=1}^{2}\mathbb{D}^*_{S|N}[E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}_G^{A(h)*}(1,g^{-A(h)})\}|\mathbb{S}^*_N]\Big|\mathbb{S}^*_N\Big\} \quad (\because \text{Lemmas 10 (2) and 23 (1)}) \\
&= \frac{1}{\prod_{h=1}^{2}\bar{n}^{A(h)}}\sum_{g^{-A(1)}=0}^{1}\sum_{g^{-A(2)}=0}^{1}\Big\{\prod_{h=1}^{2}\bar{n}^{A(h)}(1,g^{-A(h)})\Big\} \\
&\quad \mathbb{V}^{2*}_{S|N,h=1}[E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}_G^{A(h)*}(1,g^{-A(h)})\}|\mathbb{S}^*_N](\because \text{Equation 93}) \\
&= \frac{1}{\prod_{h=1}^{2}\bar{n}^{A(h)}}\sum_{g^{-A(1)}=0}^{1}\sum_{g^{-A(2)}=0}^{1}I(\mathbb{J}^{(1)*}=\mathbb{J}^{(2)*})\bar{n}^{A}(1,g^{-A})\widetilde{V^{2*}_{h=1}}\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}_G^{A*}(1,g^{-A}),\bar{\mathbf{n}}\} \\
&(\because \text{Equations 264 and 279}) \\
&= \frac{\bar{n}^{A(1)A(2)}}{\prod_{h=1}^{2}\bar{n}^{A(h)}}\widetilde{V^{2*}_{h=1}}\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}_G^{*}(1,1),\bar{\mathbf{n}}\} \\
&(\because \mathbb{J}^{(1)*}=\mathbb{J}^{(2)*} \text{ only when } g^{-A(1)}=g^{-A(2)}=1 \because A(1)\neq A(2), \bar{n}^{A}(1,1)=\bar{n}^{A(1)A(2)}) \\
&= \frac{\bar{n}^{A(1)A(2)}}{\prod_{h=1}^{2}\bar{n}^{A(h)}}\widetilde{EV^{2*}_{N,h=1}}\Big(\boldsymbol{\beta}^{(h)*}\Big|\prod_{h=1}^{2}\boldsymbol{k}_G^{A(h)*},\bar{\mathbf{n}}\Big) \quad (\because A(1)\neq A(2), \text{Lemma 14 (2), Equation 260})
\end{aligned}
\tag{281}
$$

Similarly, when $A(1) = A(2) = A$, it follows that $\mathbb{J}^{(1)*} = \mathbb{J}^{(2)*}$ and

$$
\begin{aligned}
&\mathbb{V}^{2*}_{S|N,h=1}\{E^*(\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}_G^{A(h)*})|\mathbb{S}^*_N\} \\
&= \frac{1}{(\bar{n}^A)^2}\sum_{g^{-A}=0}^{1}\{\bar{n}^A(1,g^{-A})\}^2\mathbb{V}^{2*}_{S|N,h=1}[E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}_G^{A*}(1,g^{-A})\}|\mathbb{S}^*_N] \\
&= \frac{1}{\bar{n}^A}\sum_{g^{-A}=0}^{1}\frac{\bar{n}^A(1,g^{-A})}{\bar{n}^A}\widetilde{V^{2*}_{h=1}}\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}_G^{A*}(1,g^{-A}),\bar{\mathbf{n}}\} \qquad (282)\\
&= \frac{1}{\bar{n}^A}\widetilde{EV^{2*}_{N,h=1}}(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}_G^{A*},\bar{\mathbf{n}}) \quad (\because \text{Equation 258}) \\
&= \frac{\bar{n}^{A(1)A(2)}}{\prod_{h=1}^{2}\bar{n}^{A(h)}}\widetilde{EV^{2*}_{N,h=1}}\left(\boldsymbol{\beta}^{(h)*}\middle|\prod_{h=1}^{2}\boldsymbol{k}_G^{A(h)*},\bar{\mathbf{n}}\right)
\end{aligned}
$$

where the last line follows because

$$
\begin{aligned}
\frac{\bar{n}^{A(1)A(2)}}{\prod_{h=1}^{2}\bar{n}^{A(h)}} &= \frac{\bar{n}^A}{(\bar{n}^A)^2} \quad (\because A(1) = A(2) = A) \\
&= \frac{1}{\bar{n}^A} \\
\prod_{h=1}^{2}\boldsymbol{k}_G^{A(h)*} &= (\boldsymbol{k}_G^{A*})^2 \quad (\because A(1) = A(2) = A) \\
&= \boldsymbol{k}_G^{A*} \quad (\because \boldsymbol{k}_G^{A*} \in \mathbb{U}^*, \text{Lemma 1 (5)})
\end{aligned}
$$

When $\mathbb{S}^*_N = \emptyset$, according to Equation 261, it follows

$$
\mathbb{V}^{2*}_{S|N,h=1}\{E^*(\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}_G^{A(h)*})|\mathbb{S}^*_N\} = 0. \qquad (283)
$$

Let $\mathbb{S}'^*_N \equiv \mathbb{S}^*(\mathbf{n}_G = \bar{\mathbf{n}})$. Note that, in general,

$$\begin{aligned}
&\mathbb{V}^{2*}_{h=1}\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}^*\} \\
&= \mathbb{E}^*\Big(\prod_{h=1}^{2}[f_S^{(h)}(\boldsymbol{S}^*) - \mathbb{E}^*\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}^*\}]\Big|\mathbb{S}^*\Big) \quad (\because \text{Equations 92 and 93}) \\
&= \mathbb{E}^*\Big\{\prod_{h=1}^{2}\Big([f_S^{(h)}(\boldsymbol{S}^*) - \mathbb{E}^*_{S|N}\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}'^*_N\}] + [\mathbb{E}^*_{S|N}\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}'^*_N\} - \mathbb{E}^*\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}^*\}]\Big)\Big|\mathbb{S}^*\Big\} \\
&= \mathbb{E}^*\Big\{\prod_{h=1}^{2}\Big(\mathbb{D}^*_{S|N}\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}'^*_N\} + \mathbb{D}^*[\mathbb{E}^*_{S|N}\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}'^*_N\}|\mathbb{S}^*]\Big)\Big|\mathbb{S}^*\Big\} \\
&\quad (\because \text{Equations 203 and 278, Lemma 23 (2)}) \\
&= \mathbb{E}^*\Big(\prod_{h=1}^{2}\mathbb{D}^*_{S|N}\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}'^*_N\}\Big|\mathbb{S}^*\Big) + \mathbb{E}^*\Big(\prod_{h=1}^{2}\mathbb{D}^*[\mathbb{E}^*_{S|N}\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}'^*_N\}|\mathbb{S}^*]\Big|\mathbb{S}^*\Big) \\
&\quad + 2\mathbb{E}^*\Big(\mathbb{D}^*_{S|N}\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}'^*_N\}\mathbb{D}^*[\mathbb{E}^*_{S|N}\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}'^*_N\}|\mathbb{S}^*]\Big|\mathbb{S}^*\Big) \quad (\because \text{Lemma 23 (1)}) \\
&= \mathbb{E}^*_N\Big\{\mathbb{E}^*_{S|N}\Big(\prod_{h=1}^{2}\mathbb{D}^*_{S|N}\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}'^*_N\}\Big|\mathbb{S}'^*_N\Big)\Big|\mathbb{S}^*\Big\} + \mathbb{V}^{2*}_{h=1}[\mathbb{E}^*_{S|N}\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}'^*_N\}|\mathbb{S}^*] \\
&\quad + 2\mathbb{E}^*_N\Big\{\mathbb{E}^*_{S|N}\Big(\mathbb{D}^*_{S|N}\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}'^*_N\}\mathbb{D}^*[\mathbb{E}^*_{S|N}\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}'^*_N\}|\mathbb{S}^*]\Big|\mathbb{S}'^*_N\Big)\Big|\mathbb{S}^*\Big\} \\
&\quad (\because \text{Lemma 23 (2), Equations 92 and 93}) \\
&= \mathbb{E}^*[\mathbb{V}^{2*}_{S|N,h=1}\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}'^*_N\}|\mathbb{S}^*] + \mathbb{V}^{2*}_{h=1}[\mathbb{E}^*_{S|N}\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}'^*_N\}|\mathbb{S}^*] \\
&\quad + 2\mathbb{E}^*\Big(\mathbb{E}^*_{S|N}[\mathbb{D}^*_{S|N}\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}'^*_N\}|\mathbb{S}'^*_N]\mathbb{D}^*[\mathbb{E}^*_{S|N}\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}'^*_N\}|\mathbb{S}^*]\Big|\mathbb{S}^*\Big) \\
&\quad (\because \text{Equations 92, 93, and 203, Lemma 10 (2)}) \\
&= \mathbb{E}^*[\mathbb{V}^{2*}_{S|N,h=1}\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}'^*_N\}|\mathbb{S}^*] + \mathbb{V}^{2*}_{h=1}[\mathbb{E}^*_{S|N}\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}'^*_N\}|\mathbb{S}^*] \\
&\quad (\because \mathbb{E}^*_{S|N}[\mathbb{D}^*_{S|N}\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}'^*_N\}|\mathbb{S}'^*_N] = \mathbb{E}^*_{S|N}\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}'^*_N\} - \mathbb{E}^*_{S|N}\{f_S^{(h)}(\boldsymbol{S}^*)|\mathbb{S}'^*_N\} = 0)
\end{aligned} \tag{284}$$

Note

$$\begin{aligned}
\mathbb{D}^*\Big\{\frac{n_G^A(1,0)}{n_G^A}\Big|\mathbb{S}^*\Big\} &= \mathbb{D}^*\Big\{1 - \frac{n_G^A(1,1)}{n_G^A}\Big|\mathbb{S}^*\Big\}) \quad (\because \text{Equations 180 and 182}) \\
&= -\mathbb{D}^*\Big\{\frac{n_G^A(1,1)}{n_G^A}\Big|\mathbb{S}^*\Big\} \quad (\because \text{Lemma 15 (4)})
\end{aligned} \tag{285}$$

When $\mathbb{S}^* \subseteq \bigcap_{h=1}^{2} \mathbb{S}_G^{A(h)1*}$,

$$\begin{aligned}
&\mathbb{V}_{h=1}^{2*}\{E_N^*(\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A(h)*},\mathbf{n}_G)|\mathbb{S}^*\}\\
&=\mathbb{E}^*\Big(\prod_{h=1}^{2}\mathbb{D}^*\Big[\sum_{g^{-A(h)}=0}^{1}\frac{n_G^{A(h)}(1,g^{-A(h)})}{n_G^{A(h)}}E^*\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A(h)*}(1,g^{-A(h)})\}\Big|\mathbb{S}^*\Big]\Big|\mathbb{S}^*\Big)\\
&\quad(\because \mathbb{S}^*\subseteq\bigcap_{h=1}^{2}\mathbb{S}_G^{A(h)1*}, \text{Equations 93 and 205})\\
&=\mathbb{E}^*\Big[\prod_{h=1}^{2}\sum_{g^{-A(h)}=0}^{1}\mathbb{D}^*\Big\{\frac{n_G^{A(h)}(1,g^{-A(h)})}{n_G^{A(h)}}\Big|\mathbb{S}^*\Big\}E^*\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A(h)*}(1,g^{-A(h)})\}\Big|\mathbb{S}^*\Big]\\
&\quad(\because \text{Lemma 15 (5), Equation 115})\\
&=\mathbb{E}^*\Big(\prod_{h=1}^{2}\mathbb{D}^*\Big\{\frac{n_G^{A(h)}(1,1)}{n_G^{A(h)}}\Big|\mathbb{S}^*\Big\}[E^*\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A(h)*}(1,1)\}-E^*\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A(h)*}(1,0)\}]\Big|\mathbb{S}^*\Big)\\
&\quad(\because \text{Equation 285})\\
&=\mathbb{E}^*\Big\{\prod_{h=1}^{2}\mathbb{D}^*\Big(\frac{n_G^{TC}}{n_G^{A(h)}}\Big|\mathbb{S}^*\Big)\Big|\mathbb{S}^*\Big\}\prod_{h=1}^{2}\Delta E^*(\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A(h)*})\\
&\quad(\because \text{Equation 263, Lemma 10 (2)})\\
&=\mathbb{V}_{h=1}^{2*}\Big(\frac{n_G^{TC}}{n_G^{A(h)}}\Big|\mathbb{S}^*\Big)\prod_{h=1}^{2}\Delta E^*(\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A(h)*})\quad(\because \text{Equation 93})
\end{aligned}\tag{286}$$

To conclude,

$$\begin{aligned}
&\mathbb{V}_{h=1}^{2*}\{E^*(\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}_G^{A(h)*})|\mathbb{S}_G^*\}\\
&=\mathbb{E}^*[\mathbb{V}_{S|N,h=1}^{2*}\{E^*(\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}_G^{A(h)*})|\mathbb{S}_N^*\}|\mathbb{S}_G^*]+\mathbb{V}_{h=1}^{2*}[\mathbb{E}_{S|N}^*\{E^*(\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}_G^{A(h)*})|\mathbb{S}_N^*\}|\mathbb{S}_G^*]\\
&\quad(\because \text{Equation 284 where } \mathbb{S}^*=\mathbb{S}_G^*, \mathbb{S}_N'^*=\mathbb{S}_G^*(\mathbf{n}_G=\bar{\mathbf{n}})=\mathbb{S}_N^*)\\
&=\mathbb{E}^*\Big\{\frac{n_G^{A(1)A(2)}}{\prod_{h=1}^{2}n_G^{A(h)}}\widetilde{EV_{N,h=1}^{2*}}\Big(\boldsymbol{\beta}^{(h)*}\Big|\prod_{h=1}^{2}\boldsymbol{k}_G^{A(h)*},\mathbf{n}_G\Big)\Big|\mathbb{S}_G^*\Big\}+\mathbb{V}_{h=1}^{2*}\Big(\frac{n_G^{TC}}{n_G^{A(h)}}\Big|\mathbb{S}^*\Big)\prod_{h=1}^{2}\Delta E^*(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}_G^{A(h)*})\\
&\quad(\because \text{Equations 281, 282, 283, and 286, where } \mathbb{S}^*=\mathbb{S}_G^*\subseteq\bigcap_A\mathbb{S}_G^{A1*}(\because \mathbb{N}\subseteq\bigcap_{h=1}^{2}\mathbb{N}^{A(h)1}),\\
&\quad\text{Lemma 24 (2)})
\end{aligned}$$

(2) Note that

$$\begin{aligned}
V_{h=1}^{2}\Big(\boldsymbol{\omega}^{(h)}\Big|\prod_{h=1}^{2}\boldsymbol{k}_G^{A(h)}\Big)&=E\Big(\prod_{h=1}^{2}\boldsymbol{\omega}^{(h)}\Big|\prod_{h=1}^{2}\boldsymbol{k}_G^{A(h)}\Big)\quad(\because \text{Lemma 17 (1)})\\
&=E\Big(\prod_{h=1}^{2}\boldsymbol{\omega}^{(h)(*)}(\boldsymbol{S}^*)\Big|\prod_{h=1}^{2}\boldsymbol{k}_G^{A(h)(*)}(\boldsymbol{S}^*)\Big)\quad(\because \text{Lemma 19 (5), Equation 164})\\
&=E^*\Big(\prod_{h=1}^{2}\boldsymbol{\omega}^{(h)*}\Big|\boldsymbol{S}^*\prod_{h=1}^{2}\boldsymbol{k}_G^{A(h)*}\Big)\quad(\because \text{Lemma 19 (4)})
\end{aligned}\tag{287}$$

Given $\bar{\mathbf{n}}\in\mathbb{N}$, let $\mathbb{S}_N^*=\mathbb{S}_{\max}^*(\mathbf{n}_G=\bar{\mathbf{n}})$. Suppose that $\mathbb{S}_N^*\neq\emptyset$. When $A(1)=A(2)\equiv A$, it follows

$$
\begin{aligned}
&\mathbb{E}^*_{S|N}\Big\{V^2_{h=1}\Big(\boldsymbol{\omega}^{(h)}\Big|\prod_{h=1}^{2}\boldsymbol{k}_G^{A(h)})\Big|\mathbb{S}^*_N\Big\}\\
&=\mathbb{E}^*_{S|N}\Big\{E^*\Big(\prod_{h=1}^{2}\boldsymbol{\omega}^{(h)*}\Big|\boldsymbol{S}^*\boldsymbol{k}_G^{A*}\Big)\Big|\mathbb{S}^*_N\Big\} \quad (\because \text{Equation 287, Lemma 14 (3)})\\
&=E^*_N\Big(\prod_{h=1}^{2}\boldsymbol{\omega}^{(h)*}\Big|\boldsymbol{k}_G^{A*},\bar{\mathbf{n}}\Big) \quad (\because \mathbb{S}^*_N\neq\emptyset,\text{ Lemma 24 (2), where we substitute }\prod_{h=1}^{2}\boldsymbol{\omega}^{(h)*}\in\mathbb{B}\text{ with }\boldsymbol{\beta}^*)\\
&=\sum_{g^{-A}=0}^{1}\frac{\bar{n}^A(1,g^{-A})}{\bar{n}^A}E^*\Big\{\prod_{h=1}^{2}\boldsymbol{\omega}^{(h)*}\Big|\boldsymbol{k}_G^{A*}(1,g^{-A})\Big\} \quad (\because \text{Equation 205, }\bar{\mathbf{n}}\in\mathbb{N}\subseteq\mathbb{N}^{A1}))\\
&=\sum_{g^{-A}=0}^{1}\frac{\bar{n}^A(1,g^{-A})}{\bar{n}^A}V^{2*}_{h=1}\{\boldsymbol{\omega}^{(h)*}|\boldsymbol{k}_G^{A*}(1,g^{-A})\} \quad (\because \text{Lemma 17 (1)})\\
&=EV^{2*}_{N,h=1}(\boldsymbol{\omega}^{(h)*}|\boldsymbol{k}_G^{A*},\bar{\mathbf{n}}) \quad (\because \text{Equation 254})\\
&=EV^{2*}_{N,h=1}\Big(\boldsymbol{\omega}^{(h)*}\Big|\prod_{h=1}^{2}\boldsymbol{k}_G^{A(h)*},\bar{\mathbf{n}}\Big) \quad (\because \text{Lemma 14 (3)})
\end{aligned}
\tag{288}
$$

When $A(1)\neq A(2)$, due to Lemma 14 (2), it follows $\prod_{h=1}^{2}\boldsymbol{k}_G^{A(h)*}=\boldsymbol{k}_G^{TC*}=\boldsymbol{k}_G^*(1,1)$ and

$$
\begin{aligned}
&\mathbb{E}^*_{S|N}\Big\{V^2_{h=1}\Big(\boldsymbol{\omega}^{(h)}\Big|\prod_{h=1}^{2}\boldsymbol{k}_G^{A(h)})\Big|\mathbb{S}^*_N\Big\}\\
&=\mathbb{E}^*_{S|N}\Big[E^*\Big\{\prod_{h=1}^{2}\boldsymbol{\omega}^{(h)*}\Big|\boldsymbol{S}^*\boldsymbol{k}_G^*(1,1)\Big\}\Big|\mathbb{S}^*_N\Big] \quad (\because \text{Equation 287, }\prod_{h=1}^{2}\boldsymbol{k}_G^{A(h)*}=\boldsymbol{k}_G^*(1,1))\\
&=E^*\Big\{\prod_{h=1}^{2}\boldsymbol{\omega}^{(h)*}\Big|\boldsymbol{k}_G^*(1,1)\Big\} \quad (\because \mathbb{S}^*_N\neq\emptyset,\text{ Lemma 24 (1),where we substitute }\prod_{h=1}^{2}\boldsymbol{\omega}^{(h)*}\in\mathbb{B}\text{ with }\boldsymbol{\beta}^*)\\
&=V^{2*}_{h=1}\{\boldsymbol{\omega}^{(h)*}|\boldsymbol{k}_G^*(1,1)\} \quad (\because \text{Lemma 17 (1)})\\
&=EV^{2*}_{N,h=1}\Big(\boldsymbol{\omega}^{(h)*}\Big|\prod_{h=1}^{2}\boldsymbol{k}_G^{A(h)*},\bar{\mathbf{n}}\Big) \quad (\because \text{Equation 255, }\prod_{h=1}^{2}\boldsymbol{k}_G^{A(h)*}=\boldsymbol{k}_G^{TC*}=\boldsymbol{k}_G^*(1,1))
\end{aligned}
\tag{289}
$$

Thus,

$$\begin{aligned}
&\mathbb{V}^{2*}_{h=1}\{E(\boldsymbol{\omega}^{(h)}|\boldsymbol{k}_G^{A(h)}\boldsymbol{X}^{A(h)})|\mathbb{S}^*_G\}\\
&=\mathbb{E}^*\Big(\prod_{h=1}^{2}\Big[E(\boldsymbol{\omega}^{(h)}|\boldsymbol{k}_G^{A(h)}\boldsymbol{X}^{A(h)})-\mathbb{E}^*\{E(\boldsymbol{\omega}^{(h)}|\boldsymbol{k}_G^{A(h)}\boldsymbol{X}^{A(h)})|\mathbb{S}^*_G\}\Big]\Big|\mathbb{S}^*_G\Big)\quad(\because \text{Equation 93})\\
&=\mathbb{E}\Big\{\prod_{h=1}^{2}E(\boldsymbol{\omega}^{(h)}|\boldsymbol{k}_G^{A(h)}\boldsymbol{X}^{A(h)})\Big|\mathbb{S}^*_G\Big\}\quad(\because \text{Lemma 12 (1)})\\
&=\mathbb{E}^*_S\Big[\mathbb{E}_X\Big\{\prod_{h=1}^{2}E(\boldsymbol{\omega}^{(h)}|\boldsymbol{k}_G^{A(h)}\boldsymbol{X}^{A(h)})\Big\}\Big|\mathbb{S}^*_G\Big]\quad(\because \text{Equation 199})\\
&=\mathbb{E}^*_S\Big\{[2\cdot I\{A(1)=A(2)\}-1]\frac{n_G^{A(1)A(2)}}{\prod_{h=1}^{2}n_G^{A(h)}}V^2_{h=1}\Big(\boldsymbol{\omega}^{(h)}\Big|\prod_{h=1}^{2}\boldsymbol{k}_G^{A(h)*}\Big)\Big|\mathbb{S}^*_G\Big\}\\
&\quad(\because \text{Lemma 17 (2)}, \mathbf{n}_G\in\mathbb{N}\subseteq\mathbb{N}^{A(h)1})\\
&=[2\cdot I\{A(1)=A(2)\}-1]\mathbb{E}^*_N\Big[\frac{n_G^{A(1)A(2)}}{\prod_{h=1}^{2}n_G^{A(h)}}\mathbb{E}^*_{S|N}\Big\{V^2_{h=1}\Big(\boldsymbol{\omega}^{(h)}\Big|\prod_{h=1}^{2}\boldsymbol{k}_G^{A(h)}\Big)\Big|\mathbb{S}^*_N\Big\}\Big|\mathbb{S}^*_G\Big]\\
&\quad(\because \text{Lemmas 10 (2) and 23 (2)})\\
&=[2\cdot I\{A(1)=A(2)\}-1]\mathbb{E}^*_N\Big[\frac{n_G^{A(1)A(2)}}{\prod_{h=1}^{2}n_G^{A(h)}}\Big\{I(\mathbb{S}^*_N\neq\emptyset)EV^{2*}_{N,h=1}\Big(\boldsymbol{\omega}^{(h)*}\Big|\prod_{h=1}^{2}\boldsymbol{k}_G^{A(h)*},\mathbf{n}_G\Big)\\
&\quad+I(\mathbb{S}^*_N=\emptyset)\cdot 0\Big\}\Big|\mathbb{S}^*_G\Big]\quad(\because \text{Equations 201, 288, and 289})\\
&=[2\cdot I\{A(1)=A(2)\}-1]\sum_{\bar{\mathbf{n}}\in\mathbb{N}_{\max}}\Pr(\mathbf{n}_G=\bar{\mathbf{n}}|\boldsymbol{S}^*\in\mathbb{S}^*_G)\frac{\bar{n}^{A(1)A(2)}}{\prod_{h=1}^{2}\bar{n}^{A(h)}}I(\mathbb{S}^*_N\neq\emptyset)\\
&\quad\times EV^{2*}_{N,h=1}\Big(\boldsymbol{\omega}^{(h)*}\Big|\prod_{h=1}^{2}\boldsymbol{k}_G^{A(h)*},\bar{\mathbf{n}}\Big)\quad(\because \text{Lemma 23 (2)})\\
&=[2\cdot I\{A(1)=A(2)\}-1]\sum_{\bar{\mathbf{n}}\in\mathbb{N}_{\max}}\Pr(\mathbf{n}_G=\bar{\mathbf{n}}|\boldsymbol{S}^*\in\mathbb{S}^*_G)\frac{\bar{n}^{A(1)A(2)}}{\prod_{h=1}^{2}\bar{n}^{A(h)}}\\
&\quad\times EV^{2*}_{N,h=1}\Big(\boldsymbol{\omega}^{(h)*}\Big|\prod_{h=1}^{2}\boldsymbol{k}_G^{A(h)*},\bar{\mathbf{n}}\Big)\quad(\because \text{when } \mathbb{S}^*_N=\emptyset, \text{ it follows } \Pr(\mathbf{n}_G=\bar{\mathbf{n}}|\boldsymbol{S}^*\in\mathbb{S}^*_G)=0)\\
&=[2\cdot I\{A(1)=A(2)\}-1]\mathbb{E}^*\Big\{\frac{n_G^{A(1)A(2)}}{\prod_{h=1}^{2}n_G^{A(h)}}EV^{2*}_{N,h=1}\Big(\boldsymbol{\omega}^{(h)*}\Big|\prod_{h=1}^{2}\boldsymbol{k}_G^{A(h)*},\mathbf{n}_G\Big)\Big|\mathbb{S}^*_G\Big\}\\
&\quad(\because \text{Lemma 23 (2)})
\end{aligned}$$

□

Proposition 2^* (Variance of ATE Estimators: SP). *(1) Under Assumption* 1^**, it holds that* $\mathbb{S}^*_{def}(\hat{\tau}_F)=\mathbb{S}^*_{\max}$ *and*

$$\lim_{n^*\to\infty}\mathbb{V}^{2*}(\hat{\tau}_F|\mathbb{S}^*_{\max})=\frac{1}{n_F}\{V^{2*}(\boldsymbol{\beta}^{T*}-\boldsymbol{\beta}^{C*}|\boldsymbol{k}_F^*)+V^{2*}(\boldsymbol{\omega}^{T*}+\boldsymbol{\omega}^{C*}|\boldsymbol{k}_F^*)\}.$$

(2) Under Assumption 3^**, it holds that* $\mathbb{S}^*_{def}(\hat{\tau}_P)=\mathbb{S}^{1*}_P$ *and, when* $\mathbb{S}^{1*}_P\neq\emptyset$*,*

$$\lim_{n^*\to\infty}\mathbb{V}^{2*}(\hat{\tau}_P|\mathbb{S}^{1*}_P)=\mathbb{E}^*\Big(\frac{1}{n_P}\Big|\mathbb{S}^{1*}_P\Big)\{V^{2*}(\boldsymbol{\beta}^{T*}-\boldsymbol{\beta}^{C*}|\boldsymbol{k}_P^*)+V^{2*}(\boldsymbol{\omega}^{T*}+\boldsymbol{\omega}^{C*}|\boldsymbol{k}_P^*)\}$$

(3) Under Assumption 2, it holds that* $\mathbb{S}^*_{def}(\hat{\tau}_U) = \mathbb{S}^{1*}_U$ *and, when* $\mathbb{S}^{1*}_U \neq \emptyset$,

$$
\begin{aligned}
\lim_{n^*\to\infty} \mathbb{V}^{2*}(\hat{\tau}_U|\mathbb{S}^{1*}_U) = &\mathbb{E}^*\Big\{\frac{n_P}{(n^T_U)^2}\Big|\mathbb{S}^{1*}_U\Big\}\{V^{2*}(\boldsymbol{\beta}^{T*}|\boldsymbol{k}^*_P) + V^{2*}(\boldsymbol{\omega}^{T*}|\boldsymbol{k}^*_P)\}\\
&+ \mathbb{E}^*\Big\{\frac{n^T_U - n_P}{(n^T_U)^2}\Big|\mathbb{S}^{1*}_U\Big\}\{V^{2*}(\boldsymbol{\beta}^{T*}|\boldsymbol{k}^{T*}_U - \boldsymbol{k}^*_P) + V^{2*}(\boldsymbol{\omega}^{T*}|\boldsymbol{k}^{T*}_U - \boldsymbol{k}^*_P)\}\\
&+ \mathbb{E}^*\Big\{\frac{n_P}{(n^C_U)^2}\Big|\mathbb{S}^{1*}_U\Big\}\{V^{2*}(\boldsymbol{\beta}^{C*}|\boldsymbol{k}^*_P) + V^{2*}(\boldsymbol{\omega}^{C*}|\boldsymbol{k}^*_P)\}\\
&+ \mathbb{E}^*\Big\{\frac{n^C_U - n_P}{(n^C_U)^2}\Big|\mathbb{S}^{1*}_U\Big\}\{V^{2*}(\boldsymbol{\beta}^{C*}|\boldsymbol{k}^{C*}_U - \boldsymbol{k}^*_P) + V^{2*}(\boldsymbol{\omega}^{C*}|\boldsymbol{k}^{C*}_U - \boldsymbol{k}^*_P)\}\\
&- 2\mathbb{E}^*\Big\{\frac{n_P}{n^T_U n^C_U}\Big|\mathbb{S}^{1*}_U\Big\}\{V^*(\boldsymbol{\beta}^{T*}, \boldsymbol{\beta}^{C*}|\boldsymbol{k}^*_P) - V^*(\boldsymbol{\omega}^{T*}, \boldsymbol{\omega}^{C*}|\boldsymbol{k}^*_P)\}\\
&+ \mathbb{V}^{2*}\Big(\frac{n_P}{n^T_U}\Big|\mathbb{S}^{1*}_U\Big)\{\Delta E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{k}^{T*}_U)\}^2 + \mathbb{V}^{2*}\Big(\frac{n_P}{n^C_U}\Big|\mathbb{S}^{1*}_U\Big)\{\Delta E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{k}^{C*}_U)\}^2\\
&- 2\mathbb{V}^*\Big(\frac{n_P}{n^T_U}, \frac{n_P}{n^C_U}\Big|\mathbb{S}^{1*}_U\Big)\Delta E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{k}^{T*}_U)\Delta E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{k}^{C*}_U).
\end{aligned}
$$

In fact, I do not have to condition Propositions 2* (2) and (3) on $\mathbb{S}^{1*}_P \neq \emptyset$ and $\mathbb{S}^{1*}_U \neq \emptyset$, respectively. Similar notes apply to Propositions 3* and 4* as well.

In Proposition 2* (3), the first five lines of the right hand side correspond to expectation of conditional variance of $\hat{\tau}_U$ ($\mathbb{E}^*\{\mathbb{V}^{2*}(\hat{\tau}_U|\mathbb{S}^{1*}_U, \mathbf{n}_U)|\mathbb{S}^{1*}_U\}$), while the last two lines represent variance of conditional expectation of $\hat{\tau}_U$ ($\mathbb{V}^{2*}\{\mathbb{E}^*(\hat{\tau}_U|\mathbb{S}^{1*}_U, \mathbf{n}_U)|\mathbb{S}^{1*}_U\}$), where both conditions depend on $\mathbf{n}_U \equiv (n^T_U, n^C_U, n_P)$. The first and second lines refer to expectation of the conditional variance of the average outcome for the treated group ($\mathbb{E}^*[\mathbb{V}^{2*}\{E(\boldsymbol{Y}|\boldsymbol{K}_U\boldsymbol{X})|\mathbb{S}^{1*}_U, \mathbf{n}_U\}|\mathbb{S}^{1*}_U]$), which is reduced to a weighted average of the variances of the treated potential outcomes for the pairs in $\mathbb{J}^*_P$ ($V^{2*}(\boldsymbol{y}^{T*}|\boldsymbol{k}^*_P)$, the first line) and in $\mathbb{J}^{T*}_U \setminus \mathbb{J}^*_P$ ($V^{2*}(\boldsymbol{y}^{T*}|\boldsymbol{k}^{T*}_U - \boldsymbol{k}^*_P)$, the second line) divided by n^T_U, where weights are n_P/n^T_U and $(n^T_U - n_P)/n^T_U$, respectively. The third and fourth lines indicate the equivalent for the control group ($\mathbb{E}^*(\mathbb{V}^{2*}[E\{\boldsymbol{Y}|\boldsymbol{K}_U(\mathbf{1}-\boldsymbol{X})\}|\mathbb{S}^{1*}_U, \mathbf{n}_U]|\mathbb{S}^{1*}_U)$). The fifth line deals with the covariance between both groups ($-2\mathbb{E}^*(\mathbb{V}^*[E(\boldsymbol{Y}|\boldsymbol{K}_U\boldsymbol{X}), E\{\boldsymbol{Y}|\boldsymbol{K}_U(\mathbf{1}-\boldsymbol{X})\}|\mathbb{S}^{1*}_U, \mathbf{n}_U]|\mathbb{S}^{1*}_U)$). The first and second terms in the sixth line express variances of the conditional expectation of the average outcome for the treated group ($\mathbb{V}^{2*}[\mathbb{E}^*\{E(\boldsymbol{Y}|\boldsymbol{K}_U\boldsymbol{X})|\mathbb{S}^{1*}_U, \mathbf{n}_U\}|\mathbb{S}^{1*}_U]$) and for the control group ($\mathbb{V}^{2*}(\mathbb{E}^*[E\{\boldsymbol{Y}|\boldsymbol{K}_U(\mathbf{1}-\boldsymbol{X})\}|\mathbb{S}^{1*}_U, \mathbf{n}_U]|\mathbb{S}^{1*}_U)$), respectively. The seventh line stands for the covariance between both groups ($-2\mathbb{V}^*(\mathbb{E}^*\{E(\boldsymbol{Y}|\boldsymbol{K}_U\boldsymbol{X})|\mathbb{S}^{1*}_U, \mathbf{n}_U\}\mathbb{E}^*[E\{\boldsymbol{Y}|\boldsymbol{K}_U(\mathbf{1}-\boldsymbol{X})\}|\mathbb{S}^{1*}_U, \mathbf{n}_U]|\mathbb{S}^{1*}_U)$).

The closed forms are obtained by using Lemmas 23 (2), 25, 26 and 28 (I will show Lemmad 28 after the proof of this proposition) and Equations 205 and 254. If we do not take the limit, Equations 297 and 306 show variance of the general case, $\mathbb{V}(\hat{\tau}_G|\mathbb{S}^*_G)$.

Under Assumption 2*, 4* and 5*, it follows that Proposition 2* (3) holds and

$$
\begin{aligned}
&\Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_U)\\
&= E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^*_P) - E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_U - \boldsymbol{k}^*_P) \quad (\because \text{Equation 263, Assumption 2}^*\text{, Lemma 9 (2)})\\
&= 0 - 0 \quad (\because \text{Assumption 2}^*\text{, 4}^* \text{ and 5}^*\text{, Equations 219 and 222})\\
&= 0.
\end{aligned}
\tag{290}
$$

Applying Equation 290 to Proposition 2* (3), we obtain

$$\begin{aligned}\lim_{n^*\to\infty}\mathbb{V}^{2*}(\hat{\tau}_U|\mathbb{S}_U^{1*}) = {}& \mathbb{E}^*\Big\{\frac{n_P}{(n_U^T)^2}\Big|\mathbb{S}_U^{1*}\Big\}\{V^{2*}(\boldsymbol{\beta}^{T*}|\boldsymbol{k}_P^*)+V^{2*}(\boldsymbol{\omega}^{T*}|\boldsymbol{k}_P^*)\}\\ &+\mathbb{E}^*\Big\{\frac{n_U^T-n_P}{(n_U^T)^2}\Big|\mathbb{S}_U^{1*}\Big\}\{V^{2*}(\boldsymbol{\beta}^{T*}|\boldsymbol{k}_U^{T*}-\boldsymbol{k}_P^*)+V^{2*}(\boldsymbol{\omega}^{T*}|\boldsymbol{k}_U^{T*}-\boldsymbol{k}_P^*)\}\\ &+\mathbb{E}^*\Big\{\frac{n_P}{(n_U^C)^2}\Big|\mathbb{S}_U^{1*}\Big\}\{V^{2*}(\boldsymbol{\beta}^{C*}|\boldsymbol{k}_P^*)+V^{2*}(\boldsymbol{\omega}^{C*}|\boldsymbol{k}_P^*)\}\\ &+\mathbb{E}^*\Big\{\frac{n_U^C-n_P}{(n_U^C)^2}\Big|\mathbb{S}_U^{1*}\Big\}\{V^{2*}(\boldsymbol{\beta}^{C*}|\boldsymbol{k}_U^{C*}-\boldsymbol{k}_P^*)+V^{2*}(\boldsymbol{\omega}^{C*}|\boldsymbol{k}_U^{C*}-\boldsymbol{k}_P^*)\}\\ &-2\mathbb{E}^*\Big\{\frac{n_P}{n_U^Tn_U^C}\Big|\mathbb{S}_U^{1*}\Big\}\{V^*(\boldsymbol{\beta}^{T*},\boldsymbol{\beta}^{C*}|\boldsymbol{k}_P^*)-V^*(\boldsymbol{\omega}^{T*},\boldsymbol{\omega}^{C*}|\boldsymbol{k}_P^*)\}.\end{aligned}$$

PROOF. Suppose that $\boldsymbol{K}_G^*\in\mathbb{U}^*$ satisfies Condition 1*, $\mathbb{N}\subseteq\mathbb{N}^1$, $\mathbb{S}_G^*=\mathbb{S}_{\max}^*(\mathbf{n}_G\in\mathbb{N})$,

$$\begin{aligned}\Delta\overline{\boldsymbol{\omega}}&\equiv E(\boldsymbol{\omega}^T|\boldsymbol{k}_G^T\boldsymbol{X}^T)-E(\boldsymbol{\omega}^C|\boldsymbol{k}_G^C\boldsymbol{X}^C)\\ &=\mathbb{D}^*\{E(\boldsymbol{\omega}^T|\boldsymbol{k}_G^T\boldsymbol{X}^T)|\mathbb{S}_G^*\}-\mathbb{D}^*\{E(\boldsymbol{\omega}^C|\boldsymbol{k}_G^C\boldsymbol{X}^C)|\mathbb{S}_G^*\}\quad(\because \text{Equation 92, Lemma 12 (1)})\\ &=\mathbb{D}^*\{E(\boldsymbol{\omega}^T|\boldsymbol{k}_G^T\boldsymbol{X}^T)-E(\boldsymbol{\omega}^C|\boldsymbol{k}_G^C\boldsymbol{X}^C)|\mathbb{S}_G^*\}\quad(\because \text{Lemma 15 (5)})\end{aligned}\tag{291}$$

and

$$\begin{aligned}\Delta\overline{\boldsymbol{\beta}}&\equiv[E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{S}^*\boldsymbol{k}_G^{T*})-\mathbb{E}^*\{E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{S}^*\boldsymbol{k}_G^{T*})|\mathbb{S}_G^*\}]\\ &\quad-[E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{S}\boldsymbol{k}_G^{C*})-\mathbb{E}^*\{E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{S}^*\boldsymbol{k}_G^{C*})|\mathbb{S}_G^*\}]\\ &=\mathbb{D}^*\{E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{S}^*\boldsymbol{k}_G^{T*})|\mathbb{S}_G^*\}-\mathbb{D}^*\{E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{S}^*\boldsymbol{k}_G^{C*})|\mathbb{S}_G^*\}\quad(\because \text{Equatuion 92})\\ &=\mathbb{D}^*\{E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{S}^*\boldsymbol{k}_G^{T*})-E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{S}^*\boldsymbol{k}_G^{C*})|\mathbb{S}_G^*\}.\quad(\because \text{Lemma 15 (5)})\end{aligned}\tag{292}$$

It follows

$$\begin{aligned}&\mathbb{E}^*\{(\Delta\overline{\boldsymbol{\omega}})^2|\mathbb{S}_G^*\}\\ &=\mathbb{V}^{2*}\{E(\boldsymbol{\omega}^T|\boldsymbol{k}_G^T\boldsymbol{X}^T)-E(\boldsymbol{\omega}^C|\boldsymbol{k}_G^C\boldsymbol{X}^C)|\mathbb{S}_G^*\}\quad(\because \text{Equations 93 and 291})\\ &=\sum_A\mathbb{V}^{2*}\{E(\boldsymbol{\omega}^A|\boldsymbol{k}_G^A\boldsymbol{X}^A)|\mathbb{S}_G^*\}-2\mathbb{V}^*\{E(\boldsymbol{\omega}^T|\boldsymbol{k}_G^T\boldsymbol{X}^T),E(\boldsymbol{\omega}^{C*}|\boldsymbol{k}_G^C\boldsymbol{X}^C)|\mathbb{S}_G^*\}\quad(\because \text{Lemma 15 (6)})\\ &=\mathbb{E}^*\Big\{\sum_A\frac{1}{n_G^A}EV_N^{2*}(\boldsymbol{\omega}^{A*}|\boldsymbol{k}_G^{A*},\mathbf{n}_G)+\frac{2n_G^{TC}}{n_G^Tn_G^C}V^*(\boldsymbol{\omega}^{T*},\boldsymbol{\omega}^{C*}|\boldsymbol{k}_G^{TC*})\Big|\mathbb{S}_G^*\Big\}\\ &(\because \text{Lemmas 23 (1) and 27 (2)}, \mathbb{N}\subseteq\mathbb{N}^1)\end{aligned}\tag{293}$$

and

$$\begin{aligned}
&\mathbb{E}^*\{(\Delta\overline{\boldsymbol{\beta}})^2|\mathbb{S}^*_G\}\\
&= \mathbb{V}^{2*}\{E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{S}^*\boldsymbol{k}^{T*}_G) - E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{S}^*\boldsymbol{k}^{C*}_G)|\mathbb{S}^*_G\} \quad (\because \text{Equations 93 and 292})\\
&= \sum_A \mathbb{V}^{2*}\{E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{S}^*\boldsymbol{k}^{A*}_G)|\mathbb{S}^*_G\} - 2\mathbb{V}^*\{E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{S}^*\boldsymbol{k}^{T*}_G), E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{S}^*\boldsymbol{k}^{C*}_G)|\mathbb{S}^*_G\} \quad (\because \text{Lemma 15 (6)})\\
&= \sum_A \Big[\mathbb{E}^*\Big\{\frac{1}{n^A_G}\widetilde{EV^{2*}_N}(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G, \mathbf{n}_G)\Big|\mathbb{S}^*_G\Big\} + \mathbb{V}^{2*}\Big(\frac{n^{TC}_G}{n^A_G}\Big|\mathbb{S}^*_G\Big)\{\Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G)\}^2\Big]\\
&\quad - 2\Big[\mathbb{E}^*\Big\{\frac{n^{TC}_G}{n^T_G n^C_G}\widetilde{EV^{2*}_N}(\boldsymbol{\beta}^{T*}, \boldsymbol{\beta}^{C*}|\boldsymbol{k}^{TC*}_G, \mathbf{n}_G)\Big|\mathbb{S}^*_G\Big\} + \mathbb{V}^*\Big(\frac{n^{TC}_G}{n^T_G}, \frac{n^{TC}_G}{n^C_G}\Big|\mathbb{S}^*_G\Big)\prod_A \Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G)\Big]\\
&\quad (\because \text{Lemma 27 (1)}, \mathbb{N} \subseteq \mathbb{N}^1)\\
&= \mathbb{E}^*\Big\{\sum_A \frac{1}{n^A_G}\widetilde{EV^{2*}_N}(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G, \mathbf{n}_G) - \frac{2n^{TC}_G}{n^T_G n^C_G}\widetilde{V^*}(\boldsymbol{\beta}^{T*}, \boldsymbol{\beta}^{C*}|\boldsymbol{k}^{TC*}_G, \mathbf{n}_G)\Big|\mathbb{S}^*_G\Big\}\\
&\quad + \sum_A \mathbb{V}^{2*}\Big(\frac{n^{TC}_G}{n^A_G}\Big|\mathbb{S}^*_G\Big)\{\Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G)\}^2 - 2\mathbb{V}^*\Big(\frac{n^{TC}_G}{n^T_G}, \frac{n^{TC}_G}{n^C_G}\Big|\mathbb{S}^*_G\Big)\prod_A \Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G)\\
&\quad (\because \text{Lemma 23 (1), Equation 260})
\end{aligned} \tag{294}$$

and

$$\begin{aligned}
\mathbb{E}^*(\Delta\overline{\boldsymbol{\varpi}}\Delta\overline{\boldsymbol{\beta}}|\mathbb{S}^*_G) &= \mathbb{E}^*_S\{\mathbb{E}_X(\Delta\overline{\boldsymbol{\varpi}}\Delta\overline{\boldsymbol{\beta}})|\mathbb{S}^*_G\} \quad (\because \text{Equation 199})\\
&= \mathbb{E}^*_S\{\mathbb{E}_X(\Delta\overline{\boldsymbol{\varpi}})\Delta\overline{\boldsymbol{\beta}}|\mathbb{S}^*_G\} \quad (\because \text{Lemmas 10 (2)})\\
&= \mathbb{E}^*_S(0 \cdot \Delta\overline{\boldsymbol{\beta}}|\mathbb{S}^*_G) \quad (\because \text{Lemma 12 (1)})\\
&= 0.
\end{aligned} \tag{295}$$

Note that, for any $\boldsymbol{\beta} \in \mathbb{B}$,

$$\begin{aligned}
E(\boldsymbol{\beta}|\boldsymbol{k}^A_G) &= E\{\boldsymbol{\beta}^{(*)}(\boldsymbol{S}^*) - \boldsymbol{E}^{(*)}(\boldsymbol{\beta}^*|\boldsymbol{S}^*)|\boldsymbol{k}^{A(*)}_G(\boldsymbol{S}^*)\} \quad (\because \text{Lemma 19 (5), Equation 164})\\
&= E^*(\boldsymbol{\beta}^*|\boldsymbol{S}^*\boldsymbol{k}^{A*}_G) - E^*(\boldsymbol{\beta}^*|\boldsymbol{S}^*) \quad (\because \text{Lemmas 3 (1) and (3) and 19 (4)})
\end{aligned} \tag{296}$$

Suppose that $\mathbb{S}^*_G \subseteq \mathbb{S}^*_{\text{def}}(\hat{\tau}_G)$. It follows

$$
\begin{aligned}
&\mathbb{V}^{2*}(\hat{\tau}_G|\mathbb{S}^*_G)\\
&= \mathbb{E}^*[\{\hat{\tau}_G - \mathbb{E}^*(\hat{\tau}_G|\mathbb{S}^*_G)\}^2|\mathbb{S}^*_G] \quad (\because \text{Equation 94})\\
&= \mathbb{E}^*\Big\{\Big(\hat{\tau}_G - [\bar{\tau}^* + \mathbb{E}^*\{E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{S}^*\boldsymbol{k}^{T*}_G)|\mathbb{S}^*_G\} - \mathbb{E}^*\{E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{S}^*\boldsymbol{k}^{C*}_G)|\mathbb{S}^*_G\}]\Big)^2\Big|\mathbb{S}^*_G\Big\}\\
&(\because \text{Equation 216}, \mathbb{S}^*_G \subseteq \mathbb{S}^*_{\text{def}}(\hat{\tau}_G))\\
&= \mathbb{E}^*\Big\{\Big(\hat{\tau}_G - \{\bar{\tau} + E(\boldsymbol{\beta}^T|\boldsymbol{k}^T_G) - E(\boldsymbol{\beta}^C|\boldsymbol{k}^C_G)\}\\
&- [\bar{\tau}^* + \mathbb{E}^*\{E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{S}^*\boldsymbol{k}^{T*}_G)|\mathbb{S}^*_G\} - \mathbb{E}^*\{E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{S}^*\boldsymbol{k}^{C*}_G)|\mathbb{S}^*_G\}\\
&- \{\bar{\tau} + E(\boldsymbol{\beta}^T|\boldsymbol{k}^T_G) - E(\boldsymbol{\beta}^C|\boldsymbol{k}^C_G)\}]\Big)^2\Big|\mathbb{S}^*_G\Big\}\\
&(\because \text{adding and subtracting } \bar{\tau} + E(\boldsymbol{\beta}^T|\boldsymbol{k}^T_G) - E(\boldsymbol{\beta}^C|\boldsymbol{k}^C_G))\\
&= \mathbb{E}^*\Big\{\Big[\{E(\boldsymbol{\omega}^T + \boldsymbol{\beta}^T|\boldsymbol{k}^T_G\boldsymbol{X}^T) - E(\boldsymbol{\omega}^C + \boldsymbol{\beta}^C|\boldsymbol{k}^C_G\boldsymbol{X}^C)\} - \{E(\boldsymbol{\beta}^T|\boldsymbol{k}^T_G) - E(\boldsymbol{\beta}^C|\boldsymbol{k}^C_G)\}\\
&- \Big(\bar{\tau}^* + \mathbb{E}^*\{E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{S}^*\boldsymbol{k}^{T*}_G)|\mathbb{S}^*_G\} - \mathbb{E}^*\{E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{S}^*\boldsymbol{k}^{C*}_G)|\mathbb{S}^*_G\}\\
&- [\{\bar{\tau}^* + E^*(\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*}|\boldsymbol{S}^*)\}\\
&+ \{E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{S}^*\boldsymbol{k}^{T*}_G) - E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{S}^*)\} - \{E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{S}^*\boldsymbol{k}^{C*}_G) - E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{S}^*)\}]\Big)\Big]^2\Big|\mathbb{S}^*_G\Big\}\\
&(\because \text{Equations 24 and 296, Lemma 19 (5)})\\
&= \mathbb{E}^*\Big\{\Big[\{E(\boldsymbol{\omega}^T|\boldsymbol{k}^T_G\boldsymbol{X}^T) - E(\boldsymbol{\omega}^C|\boldsymbol{k}^C_G\boldsymbol{X}^C)\} + \Big([E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{S}^*\boldsymbol{k}^{T*}_G) - \mathbb{E}^*\{E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{S}^*\boldsymbol{k}^{T*}_G)|\mathbb{S}^*_G\}]\\
&- [E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{S}\boldsymbol{k}^{C*}_G) - \mathbb{E}^*\{E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{S}^*\boldsymbol{k}^{C*}_G)|\mathbb{S}^*_G\}]\Big)\Big]^2\Big|\mathbb{S}^*_G\Big\} \quad (\because \text{Lemma 12 (2)})\\
&= \mathbb{E}^*\{(\Delta\overline{\boldsymbol{\omega}} + \Delta\overline{\boldsymbol{\beta}})^2|\mathbb{S}^*_G\} \quad (\because \text{Equations 291 and 292})\\
&= \mathbb{E}^*\{(\Delta\overline{\boldsymbol{\omega}})^2|\mathbb{S}^*_G\} + \mathbb{E}^*\{(\Delta\overline{\boldsymbol{\beta}})^2|\mathbb{S}^*_G\} + 2\mathbb{E}^*(\Delta\overline{\boldsymbol{\omega}} \cdot \Delta\overline{\boldsymbol{\beta}}|\mathbb{S}^*_G) \quad (\because \text{Lemmas 23 (1)})\\
&= \mathbb{E}^*\Big[\sum_A \frac{1}{n^A_G}\{EV^{2*}_N(\boldsymbol{\omega}^{A*}|\boldsymbol{k}^{A*}_G, \mathbf{n}_G) + \widetilde{EV^{2*}_N}(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G, \mathbf{n}_G)\}\\
&+ \frac{2n^{TC}_G}{n^T_G n^C_G}\{V^*(\boldsymbol{\omega}^{T*}, \boldsymbol{\omega}^{C*}|\boldsymbol{k}^{TC*}_G) - \widetilde{V^*}(\boldsymbol{\beta}^{T*}, \boldsymbol{\beta}^{C*}|\boldsymbol{k}^{TC*}_G, \mathbf{n}_G)\}\Big|\mathbb{S}^*_G\Big]\\
&+ \sum_A \mathbb{V}^{2*}\Big(\frac{n^{TC}_G}{n^A_G}\Big|\mathbb{S}^*_G\Big)\{\Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G)\}^2 - 2\mathbb{V}^*\Big(\frac{n^{TC}_G}{n^T_G}, \frac{n^{TC}_G}{n^C_G}\Big|\mathbb{S}^*_G\Big)\prod_A \Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G)\\
&(\because \text{Equations 293, 294, and 295})
\end{aligned} \tag{297}
$$

Before we take the limit, we need some preparation. For non-negative integer $\bar{n}$, when $\lim_{n^* \to \infty} n^{A*}_G(1, g^{-A}) = \infty$,

$$
\begin{aligned}
\lim_{n^* \to \infty} d\{n^{A*}_G(1, g^{-A}), \bar{n}\} &= \lim_{n^* \to \infty} \frac{\bar{n} - 1}{n^{A*}_G(1, g^{-A}) - 1} \quad (\because \text{Equation 256})\\
&= 0 \quad (\because \lim_{n^* \to \infty} n^{A*}_G(1, g^{-A}) = \infty)
\end{aligned} \tag{298}
$$

and, when $\lim_{n^* \to \infty} n^{A*}_G(1, g^{-A}) < 2$,

$$
\lim_{n^* \to \infty} d\{n^{A*}_G(1, g^{-A}), \bar{n}\} = 0 \quad (\because \text{Equation 256}) \tag{299}
$$

Accordingly, for $\bar{\mathbf{n}} \in \mathbb{N}$,

$$\begin{aligned}
&\lim_{n^*\to\infty} \widetilde{V_{h=1}^{2*}}\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A*}(1,g^{-A}),\bar{\mathbf{n}}\} \\
&= [1-\lim_{n^*\to\infty} d\{n_G^{A*}(1,g^{-A}),\bar{n}^A(1,g^{-A})\}]V_{h=1}^{2*}\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A*}(1,g^{-A})\} \quad (\because \text{Equation 257}) \\
&= V_{h=1}^{2*}\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A*}(1,g^{-A})\} \quad (\because \text{Equations 298 and 299})
\end{aligned} \tag{300}$$

and, thus,

$$\begin{aligned}
&\lim_{n^*\to\infty} \widetilde{EV_{N,h=1}^{2*}}(\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A*},\bar{\mathbf{n}}) \\
&= \sum_{g^{-A}=0}^{1} \frac{\bar{n}^A(1,g^{-A})}{\bar{n}^A} \lim_{n^*\to\infty} \widetilde{V_{h=1}^{2*}}\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A*}(1,g^{-A}),\bar{\mathbf{n}}\} \quad (\because \text{Equation 258}, \mathbb{N}\subseteq\mathbb{N}^1) \\
&= \sum_{g^{-A}=0}^{1} \frac{\bar{n}^A(1,g^{-A})}{\bar{n}^A} V_{h=1}^{2*}\{\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A*}(1,g^{-A})\} \quad (\because \text{Equation 300}) \\
&= EV_{N,h=1}^{2*}(\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A*},\bar{\mathbf{n}}) \quad (\because \text{Equation 254})
\end{aligned} \tag{301}$$

When $\mathbb{S}^* \subseteq \mathbb{S}_G^{A1*}$,

$$\begin{aligned}
&\mathbb{E}^*\Big\{f(\mathbf{n}_G)EV_{N,h=1}^{2*}(\boldsymbol{Q}^*|\boldsymbol{k}_G^{A*},\mathbf{n}_G)\Big|\mathbb{S}^*\Big\} \\
&= \mathbb{E}^*\Big[\sum_{g^{-A}=0}^{1} f(\mathbf{n}_G)\frac{n_G^A(1,g^{-A})}{n_G^A}V_{h=1}^{2*}\{\boldsymbol{Q}^*|\boldsymbol{k}_G^{A*}(1,g^{-A})\}\Big|\mathbb{S}^*\Big] \quad (\because \mathbb{S}^*\subseteq\mathbb{S}_G^{A1*}, \text{Equation 254}) \\
&= \sum_{g^{-A}=0}^{1} \mathbb{E}^*\Big\{f(\mathbf{n}_G)\frac{n_G^A(1,g^{-A})}{n_G^A}\Big|\mathbb{S}^*\Big\}V_{h=1}^{2*}\{\boldsymbol{Q}^*|\boldsymbol{k}_G^{A*}(1,g^{-A})\} \quad (\because \text{Lemmas 10 (2) and 23 (1)})
\end{aligned} \tag{302}$$

To sum,

$$\begin{aligned}
&\lim_{n^*\to\infty} \mathbb{V}^{2*}(\hat{\tau}_G|\mathbb{S}_G^*) \\
&= \mathbb{E}^*\Big[\sum_A \frac{1}{n_G^A}\{EV_N^{2*}(\boldsymbol{\omega}^{A*}|\boldsymbol{k}_G^{A*},\mathbf{n}_G)+EV_N^{2*}(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*},\mathbf{n}_G)\} \\
&\quad + \frac{2n_G^{TC}}{n_G^T n_G^C}\{V^*(\boldsymbol{\omega}^{T*},\boldsymbol{\omega}^{C*}|\boldsymbol{k}_G^{TC*})-V^*(\boldsymbol{\beta}^{T*},\boldsymbol{\beta}^{C*}|\boldsymbol{k}_G^{TC*})\}\Big|\mathbb{S}_G^*\Big] \\
&\quad + \sum_A \mathbb{V}^{2*}\Big(\frac{n_G^{TC}}{n_G^A}\Big|\mathbb{S}_G^*\Big)\{\Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*})\}^2 - 2\mathbb{V}^*\Big(\frac{n_G^{TC}}{n_G^T},\frac{n_G^{TC}}{n_G^C}\Big|\mathbb{S}_G^*\Big)\prod_A \Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*}) \\
&\quad (\because \text{Equations 297, 300, and 301}) \\
&= \sum_A \sum_{g^{-A}=0}^{1} \mathbb{E}^*\Big\{\frac{n_G^A(1,g^{-A})}{(n_G^A)^2}\Big|\mathbb{S}_G^*\Big\}[V^{2*}\{\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*}(1,g^{-A})\}+V^{2*}\{\boldsymbol{\omega}^{A*}|\boldsymbol{k}_G^{A*}(1,g^{-A})\}] \\
&\quad - 2\mathbb{E}^*\Big\{\frac{n_G^{TC}}{n_G^T n_G^C}\Big|\mathbb{S}_G^*\Big\}\{V^*(\boldsymbol{\beta}^{T*},\boldsymbol{\beta}^{C*}|\boldsymbol{k}_G^{TC*})-V^*(\boldsymbol{\omega}^{T*},\boldsymbol{\omega}^{C*}|\boldsymbol{k}_G^{TC*})\} \\
&\quad + \sum_A \mathbb{V}^{2*}\Big(\frac{n_G^{TC}}{n_G^A}\Big|\mathbb{S}_G^*\Big)\{\Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*})\}^2 - 2\mathbb{V}^*\Big(\frac{n_G^{TC}}{n_G^T},\frac{n_G^{TC}}{n_G^C}\Big|\mathbb{S}_G^*\Big)\prod_A \Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*}) \\
&\quad (\because \text{Lemmas 10 (2) and 23 (1), Equation 302, where } \mathbb{S}^*=\mathbb{S}_G^*\subseteq(\mathbb{S}_G^{T1*}\cap\mathbb{S}_G^{C1*})(\because \mathbb{N}\subseteq\mathbb{N}^1)\ )
\end{aligned} \tag{303}$$

In particular, when $\boldsymbol{k}_G^{T*} = \boldsymbol{k}_G^{C*} \equiv \boldsymbol{k}_G^*$, it follows that, according to Lemma 14 (3),

$$\begin{aligned} \boldsymbol{k}_G^{A*}(1,1) &= \boldsymbol{k}_G^* \\ \boldsymbol{k}_G^{A*}(1,0) &= \mathbf{0}^* \\ n_G^{A*}(1,1) &= n_G^* \\ n_G^{A}(1,1) &= n_G^{A} = n_G \\ n_G^{A}(1,0) &= 0, \end{aligned} \tag{304}$$

and, thus,

$$\begin{aligned} EV_{N,h=1}^{2*}(\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A*}, \mathbf{n}_G) &= \frac{n_G^A}{n_G^A} V_{h=1}^{2*}(\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A*}) + \frac{0}{n_G^A} V_{h=1}^{2*}(\boldsymbol{Q}^{(h)*}|\mathbf{0}^*) \quad (\because \text{Equation 254}) \\ &= V_{h=1}^{2*}(\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^*) \\ \widetilde{EV_{N,h=1}^{2*}}(\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A*}, \mathbf{n}_G) &= \frac{n_G^A}{n_G^A} \widetilde{V_{h=1}^{2*}}(\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^{A*}, \mathbf{n}_G) + \frac{0}{n_G^A} \widetilde{V_{h=1}^{2*}}(\boldsymbol{Q}^{(h)*}|\mathbf{0}^*, \mathbf{n}_G) \quad (\because \text{Equation 258}) \\ &= \{1 - d(n_G^*, n_G)\} V_{h=1}^{2*}(\boldsymbol{Q}^{(h)*}|\boldsymbol{k}_G^*) \quad (\because \text{Equation 257}) \\ \Delta E^*(\boldsymbol{Q}^*|\boldsymbol{k}_G^{A*}) &= E^*(\boldsymbol{Q}^*|\boldsymbol{k}_G^{A*}) - E^*(\boldsymbol{Q}^*|\mathbf{0}^*) \quad (\because \text{Equation 263}) \\ &= E^*(\boldsymbol{Q}^*|\boldsymbol{k}_G^{A*}) \quad (\because \text{Equation 166}) \end{aligned} \tag{305}$$

and, therefore,

$$\begin{aligned} &\mathbb{V}^{2*}(\hat{\tau}_G|\mathbb{S}_G^*) \\ &= \mathbb{E}^*\Big(\sum_A \frac{1}{n_G} [V^{2*}(\boldsymbol{\omega}^{A*}|\boldsymbol{k}_G^*) + \{1 - d(n_G^*, n_G)\} V^{2*}(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^*, \mathbf{n}_G)] \\ &\quad + \frac{2}{n_G} [V^*(\boldsymbol{\omega}^{T*}, \boldsymbol{\omega}^{C*}|\boldsymbol{k}_G^*) - \{1 - d(n_G^*, n_G)\} V^*(\boldsymbol{\beta}^{T*}, \boldsymbol{\beta}^{C*}|\boldsymbol{k}_G^*)] \Big| \mathbb{S}_G^* \Big) \\ &\quad + \sum_A \mathbb{V}^{2*}\Big(\frac{n_G}{n_G}\Big|\mathbb{S}_G^*\Big) \{E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^*)\}^2 - 2\mathbb{V}^*\Big(\frac{n_G}{n_G}, \frac{n_G}{n_G}\Big|\mathbb{S}_G^*\Big) \prod_A E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^*) \\ &\quad (\because \text{Equation 297, 305, Lemma 14 (3)}) \\ &= \mathbb{E}^*\Big(\frac{1}{n_G}\Big|\mathbb{S}_G^*\Big) V^{2*}(\boldsymbol{\omega}^{T*} + \boldsymbol{\omega}^{C*}|\boldsymbol{k}_G^*) + \mathbb{E}^*\Big[\frac{1}{n_G}\{1 - d(n_G^*, n_G)\}\Big|\mathbb{S}_G^*\Big] V^{2*}(\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*}|\boldsymbol{k}_G^*) \\ &\quad \Big(\because \text{Lemmas 10 (2) and 15 (6)}, \mathbb{V}^{2*}\Big(\frac{n_G}{n_G}\Big|\mathbb{S}_G^*\Big) = \mathbb{V}^*\Big(\frac{n_G}{n_G}, \frac{n_G}{n_G}\Big|\mathbb{S}_G^*\Big) = \mathbb{V}^{2*}(1|\mathbb{S}_G^*) = 0\Big) \end{aligned} \tag{306}$$

In limit,

$$\begin{aligned} \lim_{n^* \to \infty} d(n_G^*, n_G) &= \lim_{n^* \to \infty} d\{n_G^{A*}(1,1), n_G\} \quad (\because \text{Equation 304}) \\ &= 0 \quad (\because \text{Equations 298 and 299, where } g^{-A} = 1, \bar{n} = n_G) \end{aligned} \tag{307}$$

and, thus,

$$\begin{aligned} \lim_{n^* \to \infty} \mathbb{V}^{2*}(\hat{\tau}_G|\mathbb{S}_G^*) &= \mathbb{E}^*\Big(\frac{1}{n_G}\Big|\mathbb{S}_G^*\Big) V^{2*}(\boldsymbol{\omega}^{T*} + \boldsymbol{\omega}^{C*}|\boldsymbol{k}_G^*) + \mathbb{E}^*\Big\{\frac{1}{n_G}(1 - 0)\Big|\mathbb{S}_G^*\Big\} V^{2*}(\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*}|\boldsymbol{k}_G^*) \\ &\quad (\because \text{Equations 306 and 307}) \\ &= \mathbb{E}^*\Big(\frac{1}{n_G}\Big|\mathbb{S}_G^*\Big) \{V^{2*}(\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*}|\boldsymbol{k}_G^*) + V^{2*}(\boldsymbol{\omega}^{T*} + \boldsymbol{\omega}^{C*}|\boldsymbol{k}_G^*)\}. \end{aligned} \tag{308}$$

(1) Under Assumption 1*, it holds that $\mathbb{S}_{\text{def}}^*(\hat{\tau}_F) = \mathbb{S}_{\max}^*$ ($\because$ Assumption 1). When $\boldsymbol{K}_G^* = \boldsymbol{K}_F^* \in \mathbb{U}^*$ and $\mathbb{N} = \mathbb{N}^n$, it follows that $\boldsymbol{K}_G = \boldsymbol{K}_F$ ($\because$ Equation 164) and, according to Lemma 13

(1), $\mathbb{S}^*_G = \mathbb{S}^{n*}_F = \mathbb{S}^*_{\max}$ ($\because$ Equations 211 and 212), and Equation 308 is equivalent to the desired result, where

$$\begin{aligned}\mathbb{E}^*\Big(\frac{1}{n_G}\Big|\mathbb{S}^*_G\Big) &= \mathbb{E}^*\Big(\frac{1}{n_F}\Big|\mathbb{S}^*_G\Big) \quad (\because \text{Lemma 13 (1)})\\ &= \frac{1}{n_F}. \quad (\because \text{Equation 35})\end{aligned} \tag{309}$$

(2) Under Assumption 3*, it holds that $\mathbb{S}^*_{\text{def}}(\hat{\tau}_P) = \mathbb{S}^{1*}_P$ ($\because$ Lemma 9 (1), $N_P = n_P \geq 1$, Equations 204 and 211). When $\boldsymbol{K}^*_G = \boldsymbol{K}^*_P \in \mathbb{U}^*$ and $\mathbb{N} = \mathbb{N}^1$, it follows that $\boldsymbol{K}_G = \boldsymbol{K}_P$ ($\because$ Equation 164) and, according to Lemma 13 (2), $\mathbb{S}^*_G = \mathbb{S}^{1*}_P$ ($\because$ Equations 211 and 213) and Equation 308 is equivalent to the desired result.

(3) Under Assumption 2*, it holds that $\mathbb{S}^*_{\text{def}}(\hat{\tau}_U) = \mathbb{S}^{1*}_U$ ($\because$ Lemma 9 (2), $N^A_U = n^A_U \geq 1$, Equations 204 and 211). When $\boldsymbol{K}^*_G = \boldsymbol{K}^*_U \in \mathbb{U}^*$ and $\mathbb{N} = \mathbb{N}^1$, it follows that $\boldsymbol{K}_G = \boldsymbol{K}_U$ ($\because$ Equation 164) and, according to Lemma 13 (3), $\mathbb{S}^*_G = \mathbb{S}^{1*}_U$ ($\because$ Equations 211 and 214) and Equation 303 is equivalent to the desired result.

□

Under Assumption 2*, 4* and 5*, the sixth and seventh lines in Proposition 2* (3) are equal to zero because Equation 225 leads to

$$\Delta E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{k}^{T*}_U) = \Delta E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{k}^{C*}_U) = 0, \tag{310}$$

or because the conditional expectation of $\hat{\tau}_U$ is constant ($\mathbb{E}^*(\hat{\tau}_U|\mathbb{S}^{1*}_U, \mathbf{n}_U) = \bar{\tau}^*$) and thus its variance is zero. Note that, under Assumption 3*, the conditional expectation of $\hat{\tau}_P$ is also constant ($\mathbb{E}^*(\hat{\tau}_P|\mathbb{S}^{1*}_U, n_P) = \bar{\tau}^*_P$), while $n_F = n$ is always constant. This is a reason why Propositions 2* (1) and (2) are simpler than (3). In addition, Propositions 2* (1) and (2) hold whether or not Assumption 4* and/or 5* hold.

For an analogous reason in the case of Propositions 2, $\mathbb{V}^{2*}(\hat{\tau}_P|\mathbb{S}^{1*}_P)$ can be smaller than $\mathbb{V}^{2*}(\hat{\tau}_U|\mathbb{S}^{1*}_U)$ (or $\mathbb{V}^{2*}(\hat{\tau}_U|\mathbb{S}^{1*}_P)$, where $\mathbb{S}^{1*}_P \subseteq \mathbb{S}^{1*}_U$).

Alternative proof of Proposition 2* (3) is as follows. According to Equations 284,

$$\mathbb{V}^{2*}(\hat{\tau}_U|\mathbb{S}^{1*}_U) = \mathbb{E}^*\{\mathbb{V}^{2*}(\hat{\tau}_U|\mathbb{S}^{1*}_U, \mathbf{n}_U)|\mathbb{S}^{1*}_U\} + \mathbb{V}^{2*}\{\mathbb{E}^*(\hat{\tau}_U|\mathbb{S}^{1*}_U, \mathbf{n}_U)|\mathbb{S}^{1*}_U\}.$$

It follows

$$\begin{aligned}&\mathbb{E}^*\{\mathbb{V}^{2*}(\hat{\tau}_U|\mathbb{S}^{1*}_U, \mathbf{n}_U)|\mathbb{S}^{1*}_U\}\\ &= \mathbb{E}^*[\mathbb{V}^{2*}\{E(\boldsymbol{Y}|\boldsymbol{K}_U\boldsymbol{X}^T) - E(\boldsymbol{Y}|\boldsymbol{K}_U\boldsymbol{X}^C)|\mathbb{S}^{1*}_U, \mathbf{n}_U\}|\mathbb{S}^{1*}_U] \quad (\because \text{Equations 10 and 13})\\ &= \mathbb{E}^*[\mathbb{V}^{2*}\{E(\boldsymbol{Y}|\boldsymbol{K}_U\boldsymbol{X}^T)|\mathbb{S}^{1*}_U, \mathbf{n}_U\}|\mathbb{S}^{1*}_U] + \mathbb{E}^*[\mathbb{V}^{2*}\{E(\boldsymbol{Y}|\boldsymbol{K}_U\boldsymbol{X}^C)|\mathbb{S}^{1*}_U, \mathbf{n}_U\}|\mathbb{S}^{1*}_U]\\ &\quad - 2\mathbb{E}^*[\mathbb{V}^*\{E(\boldsymbol{Y}|\boldsymbol{K}_U\boldsymbol{X}^T), E(\boldsymbol{Y}|\boldsymbol{K}_U\boldsymbol{X}^C)|\mathbb{S}^{1*}_U, \mathbf{n}_U\}|\mathbb{S}^{1*}_U] \quad (\because \text{Lemmas 15 (6) and 23 (1)})\end{aligned}$$

Let $\mathbf{n}_U \in \mathbb{N}^1, \mathbb{S}_N^* \equiv \mathbb{S}_{\max}^*(\mathbf{n}_U)$. Note that $\boldsymbol{S}^* \in \mathbb{S}_U^{1*}, \mathbf{n}_U$ if and only if $\boldsymbol{S}^* \in \mathbb{S}_N^*$ and

$$
\begin{aligned}
&\mathbb{V}_{h=1}^{2*}\{E(\boldsymbol{Y}|\boldsymbol{K}_U\boldsymbol{X}^{A(h)})|\mathbb{S}_U^{1*}, \mathbf{n}_U\} \\
&= \mathbb{V}_{h=1}^{2*}\{E(\boldsymbol{y}^{A(h)}|\boldsymbol{k}_U^{A(h)}\boldsymbol{X}^{A(h)})|\mathbb{S}_N^*\} \quad (\because \text{Equation 23}) \\
&= \mathbb{V}_{h=1}^{2*}\{E(\boldsymbol{\mu}^{A(h)} + \boldsymbol{\beta}^{A(h)} + \boldsymbol{\omega}^{A(h)}|\boldsymbol{k}_U^{A(h)}\boldsymbol{X}^{A(h)}) - \mu^{A(h)}|\mathbb{S}_N^*\} \quad (\because \text{Lemmas 4 (1) and 15 (4)}) \\
&= \mathbb{V}_{h=1}^{2*}\{E(\boldsymbol{\beta}^{A(h)}|\boldsymbol{k}_U^{A(h)}\boldsymbol{X}^{A(h)}) + E(\boldsymbol{\omega}^{A(h)}|\boldsymbol{k}_U^{A(h)}\boldsymbol{X}^{A(h)})|\mathbb{S}_N^*\} \quad (\because \text{Lemmas 3 (1) and (3)}) \\
&= \mathbb{V}_{h=1}^{2*}\{E(\boldsymbol{\beta}^{A(h)}|\boldsymbol{k}_U^{A(h)})|\mathbb{S}_N^*\} + \mathbb{V}_{h=1}^{2*}\{E(\boldsymbol{\omega}^{A(h)}|\boldsymbol{k}_U^{A(h)}\boldsymbol{X}^{A(h)})|\mathbb{S}_N^*\} \\
&\quad + \sum_h^2 \mathbb{V}^*\{E(\boldsymbol{\beta}^{A(h)}|\boldsymbol{k}_U^{A(h)}), E(\boldsymbol{\omega}^{A(-h)}|\boldsymbol{k}_U^{A(-h)}\boldsymbol{X}^{A(-h)})|\mathbb{S}_N^*\} \quad (\because \text{Lemmas 12 (2) and 15 (6)}) \\
&= \frac{n_U^{A(1)A(2)}}{\prod_{h=1}^2 n_U^{A(h)}} \widetilde{EV_N^{2*}}\Big(\boldsymbol{\beta}^{A(h)*}\Big| \prod_{h=1}^2 \boldsymbol{k}_U^{A(h)*}, \mathbf{n}_U\Big) \\
&+ [2 \cdot I\{A(1) = A(2)\} - 1] \frac{n_U^{A(1)A(2)}}{\prod_{h=1}^2 n_U^{A(h)}} EV_N^{2*}\Big(\boldsymbol{\omega}^{A(h)*}\Big| \prod_{h=1}^2 \boldsymbol{k}_U^{A(h)*}, \mathbf{n}_U\Big) \\
&\quad (\because \text{Lemma 19 (4) and 27 (2), Equation 282})
\end{aligned}
\tag{311}
$$

where the last line follows because

$$
\begin{aligned}
&\mathbb{V}^*\{E(\boldsymbol{\beta}^{A(h)}|\boldsymbol{k}_U^{A(h)}), E(\boldsymbol{\omega}^{A(-h)}|\boldsymbol{k}_U^{A(-h)}\boldsymbol{X}^{A(-h)})|\mathbb{S}_N^*\} \\
&= \mathbb{E}^*\Big([E(\boldsymbol{\beta}^{A(h)}|\boldsymbol{k}_U^{A(h)}) - \mathbb{E}^*\{E(\boldsymbol{\beta}^{A(h)}|\boldsymbol{k}_U^{A(h)})|\mathbb{S}_N^*\}] \\
&\quad \times [E(\boldsymbol{\omega}^{A(-h)}|\boldsymbol{k}_U^{A(-h)}\boldsymbol{X}^{A(-h)}) - \mathbb{E}^*\{E(\boldsymbol{\omega}^{A(-h)}|\boldsymbol{k}_U^{A(-h)}\boldsymbol{X}^{A(-h)})|\mathbb{S}_N^*\}]\Big|\mathbb{S}_N^*\Big) \quad (\because \text{Equation 93}) \\
&= \mathbb{E}_S^*\Big\{\mathbb{E}_X\Big([E(\boldsymbol{\beta}^{A(h)}|\boldsymbol{k}_U^{A(h)}) - \mathbb{E}^*\{E(\boldsymbol{\beta}^{A(h)}|\boldsymbol{k}_U^{A(h)})|\mathbb{S}_N^*\}]E(\boldsymbol{\omega}^{A(-h)}|\boldsymbol{k}_U^{A(-h)}\boldsymbol{X}^{A(-h)})\Big)\Big|\mathbb{S}_N^*\Big\} \\
&\quad (\because \text{Equation 199, Lemma 12 (1)}) \\
&= \mathbb{E}_S^*\Big([E(\boldsymbol{\beta}^{A(h)}|\boldsymbol{k}_U^{A(h)}) - \mathbb{E}^*\{E(\boldsymbol{\beta}^{A(h)}|\boldsymbol{k}_U^{A(h)})|\mathbb{S}_N^*\}]\mathbb{E}_X\{E(\boldsymbol{\omega}^{A(-h)}|\boldsymbol{k}_U^{A(-h)}\boldsymbol{X}^{A(-h)})\}\Big|\mathbb{S}_N^*\Big) \\
&\quad (\because \text{Lemmas 10 (2)}) \\
&= \mathbb{E}_S^*[E(\boldsymbol{\beta}^{A(h)}|\boldsymbol{k}_U^{A(h)}) - \mathbb{E}^*\{E(\boldsymbol{\beta}^{A(h)}|\boldsymbol{k}_U^{A(h)})|\mathbb{S}_N^*\} \cdot 0|\mathbb{S}_N^*] \quad (\because \text{Lemma 12 (1)}) \\
&= 0.
\end{aligned}
$$

Note that, when $A(1) = A(2) \equiv A$ and $A(1) \neq A(2)$, it follows $\prod_{h=1}^2 \boldsymbol{k}_U^{A(h)*} = \boldsymbol{k}_U^{A*} \equiv \boldsymbol{k}_G^{A*}$ and

$\prod_{h=1}^{2} \boldsymbol{k}_U^{A(h)*} = \boldsymbol{k}_P^* \equiv \boldsymbol{k}_G^{A*}$, respectively. Therefore,

$$\begin{aligned}
&\lim_{n^*\to\infty} \mathbb{E}^*[\mathbb{V}_{h=1}^{2*}\{E(\boldsymbol{Y}|\boldsymbol{K}_U\boldsymbol{X}^{A(h)})|\mathbb{S}_U^{1*}, \mathbf{n}_U\}|\mathbb{S}_U^{1*}] \\
&= \mathbb{E}^*\Big[\frac{n_U^{A(1)A(2)}}{\prod_{h=1}^{2} n_U^{A(h)}}\Big\{\lim_{n^*\to\infty} \widetilde{EV_N^{2*}}\Big(\boldsymbol{\beta}^{A(h)*}\Big|\prod_{h=1}^{2}\boldsymbol{k}_U^{A(h)*}, \mathbf{n}_U\Big) \\
&\quad + [2\cdot I\{A(1)=A(2)\}-1]\lim_{n^*\to\infty} EV_N^{2*}\Big(\boldsymbol{\omega}^{A(h)*}\Big|\prod_{h=1}^{2}\boldsymbol{k}_U^{A(h)*}, \mathbf{n}_U\Big)\Big\}\Big|\mathbb{S}_U^{1*}\Big] \quad (\because \text{Equation 311}) \\
&= \mathbb{E}^*\Big[\frac{n_U^{A(1)A(2)}}{\prod_{h=1}^{2} n_U^{A(h)}}\Big\{EV_N^{2*}(\boldsymbol{\beta}^{A(h)*}|\boldsymbol{k}_G^{A*}, \mathbf{n}_U) + [2\cdot I\{A(1)=A(2)\}-1]EV_N^{2*}(\boldsymbol{\omega}^{A(h)*}|\boldsymbol{k}_G^{A*}, \mathbf{n}_U)\Big\}\Big|\mathbb{S}_U^{1*}\Big] \\
&\quad (\because \text{Equation 301}) \\
&= \mathbb{E}^*\Big\{\frac{n_U^{A(1)A(2)}}{\prod_{h=1}^{2} n_U^{A(h)}}\sum_{g^{-A}=0}^{1}\frac{n_U^A(1,g^{-A})}{n_U^A}\Big(V^{2*}\{\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*}(1,g^{-A})\} \\
&+ [2\cdot I\{A(1)=A(2)\}-1]V^{2*}\{\boldsymbol{\omega}^{A*}|\boldsymbol{k}_G^{A*}(1,g^{-A})\}\Big)\Big|\mathbb{S}_U^{1*}\Big\} \quad (\because \text{Equation 254}) \\
&= \sum_{g^{-A}=0}^{1}\mathbb{E}^*\Big(\frac{n_U^{A(1)A(2)}}{\prod_{h=1}^{2} n_U^{A(h)}}\frac{n_U^A(1,g^{-A})}{n_U^A}\Big|\mathbb{S}_U^{1*}\Big)[V^{2*}\{\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*}(1,g^{-A})\} \\
&\quad + [2\cdot I\{A(1)=A(2)\}-1]V^{2*}\{\boldsymbol{\omega}^{A*}|\boldsymbol{k}_G^{A*}(1,g^{-A})\}] \quad (\because \text{Lemmas 10 (2) and 23 (1)})
\end{aligned}$$

Specifically, when $A(1) = A(2) = T$, the first and second lines of the right hand side of Proposition 2* (3) is equal to

$$\begin{aligned}
&\lim_{n^*\to\infty} \mathbb{E}^*[\mathbb{V}^{2*}\{E(\boldsymbol{Y}|\boldsymbol{K}_U\boldsymbol{X}^T)|\mathbb{S}_U^{1*}, \mathbf{n}_U\}|\mathbb{S}_U^{1*}] \\
&= \mathbb{E}^*\Big\{\frac{n_P}{(n_U^T)^2}\Big|\mathbb{S}_U^{1*}\Big\}\{V^{2*}(\boldsymbol{\beta}^{T*}|\boldsymbol{k}_P^*) + V^{2*}(\boldsymbol{\omega}^{T*}|\boldsymbol{k}_P^*)\} \\
&\quad + \mathbb{E}^*\Big\{\frac{n_U^T - n_P}{(n_U^T)^2}\Big|\mathbb{S}_U^{1*}\Big\}\{V^{2*}(\boldsymbol{\beta}^{T*}|\boldsymbol{k}_U^{T*} - \boldsymbol{k}_P^*) + V^{2*}(\boldsymbol{\omega}^{T*}|\boldsymbol{k}_U^{T*} - \boldsymbol{k}_P^*)\},
\end{aligned}$$

when $A(1) = A(2) = C$, the third and fourth lines correspond to

$$\begin{aligned}
&\lim_{n^*\to\infty} \mathbb{E}^*[\mathbb{V}^{2*}\{E(\boldsymbol{Y}|\boldsymbol{K}_U\boldsymbol{X}^C)|\mathbb{S}_U^{1*}, \mathbf{n}_U\}|\mathbb{S}_U^{1*}] \\
&\quad + \mathbb{E}^*\Big\{\frac{n_P}{(n_U^C)^2}\Big|\mathbb{S}_U^{1*}\Big\}\{V^{2*}(\boldsymbol{\beta}^{C*}|\boldsymbol{k}_P^*) + V^{2*}(\boldsymbol{\omega}^{C*}|\boldsymbol{k}_P^*)\} \\
&\quad + \mathbb{E}^*\Big\{\frac{n_U^C - n_P}{(n_U^C)^2}\Big|\mathbb{S}_U^{1*}\Big\}\{V^{2*}(\boldsymbol{\beta}^{C*}|\boldsymbol{k}_U^{C*} - \boldsymbol{k}_P^*) + V^{2*}(\boldsymbol{\omega}^{C*}|\boldsymbol{k}_U^{C*} - \boldsymbol{k}_P^*)\},
\end{aligned}$$

and, when $A(1) \neq A(2)$, the fifth line deals with

$$\begin{aligned}
&\lim_{n^*\to\infty} \mathbb{E}^*[\mathbb{V}^*\{E(\boldsymbol{Y}|\boldsymbol{K}_U\boldsymbol{X}^T), E(\boldsymbol{Y}|\boldsymbol{K}_U\boldsymbol{X}^C)|\mathbb{S}_U^{1*}, \mathbf{n}_U\}|\mathbb{S}_U^{1*}] \\
&= \mathbb{E}^*\Big\{\frac{n_P}{n_U^T n_U^C}\Big|\mathbb{S}_U^{1*}\Big\}\{V^*(\boldsymbol{\beta}^{T*}, \boldsymbol{\beta}^{C*}|\boldsymbol{k}_P^*) - V^*(\boldsymbol{\omega}^{T*}, \boldsymbol{\omega}^{C*}|\boldsymbol{k}_P^*)\}
\end{aligned}$$

Finally, the sixth and seventh lines represent

$$\begin{aligned}
&\mathbb{V}^{2*}\{\mathbb{E}^*(\hat{\tau}_U|\mathbb{S}_U^{1*},\mathbf{n}_U)|\mathbb{S}_U^{1*}\}\\
&=\mathbb{V}^{2*}[\mathbb{E}^*\{E(\boldsymbol{Y}|\boldsymbol{K}_U\boldsymbol{X}^T)-E(\boldsymbol{Y}|\boldsymbol{K}_U\boldsymbol{X}^C)|\mathbb{S}_U^{1*},\mathbf{n}_U\}|\mathbb{S}_U^{1*}] \quad (\because \text{Equations 10 and 13})\\
&=\mathbb{V}^{2*}[\mathbb{E}^*\{E(\boldsymbol{\mu}^T+\boldsymbol{\beta}^T+\boldsymbol{\omega}^T|\boldsymbol{k}_U^T\boldsymbol{X}^T)-E(\boldsymbol{\mu}^C+\boldsymbol{\beta}^C+\boldsymbol{\omega}^C|\boldsymbol{k}_U^C\boldsymbol{X}^C)|\mathbb{S}_U^{1*},\mathbf{n}_U\}-\bar{\tau}|\mathbb{S}_U^{1*}]\\
&\quad(\because \text{Equation 23, Lemmas 4 (1), 8, and 15 (4)})\\
&=\mathbb{V}^{2*}[\mathbb{E}^*\{E^*(\boldsymbol{\beta}^{T*}|S^*\boldsymbol{k}_U^{T*})-E^*(\boldsymbol{\beta}^{C*}|S^*\boldsymbol{k}_U^{C*})|\mathbb{S}_U^{1*},\mathbf{n}_U\}|\mathbb{S}_U^{1*}]\\
&\quad(\because \text{Lemmas 3 (1) and (3), 10 (2), 12, 23 (1), and 19 (4)})\\
&=\mathbb{V}^{2*}\{E_N^*(\boldsymbol{\beta}^{T*}|\boldsymbol{k}_U^{T*},\mathbf{n}_U)-E_N^*(\boldsymbol{\beta}^{C*}|\boldsymbol{k}_U^{C*},\mathbf{n}_U)|\mathbb{S}_U^{1*}\} \quad (\because \text{Lemmas 23 (1) and 24 (2)})\\
&=\sum_A \mathbb{V}^{2*}\{E_N^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_U^{A*},\mathbf{n}_U)|\mathbb{S}_U^{1*}\}-2\mathbb{V}^*\{E_N^*(\boldsymbol{\beta}^{T*}|\boldsymbol{k}_U^{T*},\mathbf{n}_U),E_N^*(\boldsymbol{\beta}^{C*}|\boldsymbol{k}_U^{C*},\mathbf{n}_U)|\mathbb{S}_U^{1*}\}\\
&\quad(\because \text{Lemma 15 (6)})\\
&=\sum_A \mathbb{V}^{2*}\Big(\frac{n_G^{TC}}{n_G^A}\Big|\mathbb{S}_U^{1*}\Big)\{\Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*})\}^2-2\mathbb{V}^*\Big(\frac{n_G^{TC}}{n_G^T n_G^C}\Big|\mathbb{S}_U^{1*}\Big)\prod_A \Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*}) \quad (\because \text{Equation 286})
\end{aligned}$$

LEMMA 28 (DECOMPOSITION OF EXPECTATION). *Suppose that $\boldsymbol{K}_G^* \in \mathbb{U}^*$ satisfies Condition 1^*, $\mathbb{S}^* \subseteq \mathbb{S}_{def}^*(f_{X,S}^*)$, and, for any $\bar{\mathbf{n}}$such that $\mathbb{S}^*(\mathbf{n}_G=\bar{\mathbf{n}}) \neq \emptyset$,*

$$\mathbb{E}_{S|N}^*[\mathbb{E}_X\{f_{X,S}^*(\boldsymbol{X}^*,\boldsymbol{S}^*)\}|\mathbb{S}^*(\mathbf{n}_G=\bar{\mathbf{n}})]=\sum_h q^{(h)}f_N^{(h)}(\bar{\mathbf{n}})$$

and

$$\mathbb{S}^* \subseteq \bigcap_h \mathbb{S}_{def}^*\{f_N^{(h)}(\mathbf{n}_G)\}.$$

It follows

$$\mathbb{E}^*\{f_{X,S}^*(\boldsymbol{X}^*,\boldsymbol{S}^*)|\mathbb{S}^*\}=\sum_h q^{(h)}\mathbb{E}^*\{f_N^{(h)}(\mathbf{n}_G)|\mathbb{S}^*\}.$$

PROOF.

$$\begin{aligned}
&\mathbb{E}^*\{f_{X,S}^*(\boldsymbol{X}^*,\boldsymbol{S}^*)|\mathbb{S}^*\}\\
&=\mathbb{E}_S^*[\mathbb{E}_X\{f_{X,S}^*(\boldsymbol{X}^*,\boldsymbol{S}^*)|\mathbb{S}^*\}] \quad (\because \text{Equation 199})\\
&=\mathbb{E}_N^*\Big(\mathbb{E}_{S|N}^*[\mathbb{E}_X\{f_{X,S}^*(\boldsymbol{X}^*,\boldsymbol{S}^*)\}|\mathbb{S}^*(\mathbf{n}_G)]\Big|\mathbb{S}^*\Big) \quad (\because \text{Lemma 23 (2)})\\
&=\mathbb{E}_N^*\Big\{\sum_h q^{(h)}f_N^{(h)}(\mathbf{n}_G)\Big|\mathbb{S}^*\Big\}\\
&=\sum_h q^{(h)}\mathbb{E}^*\{f_N^{(h)}(\mathbf{n}_G)|\mathbb{S}^*\} \quad (\because \text{Lemmas 10 (2) and 23 (1), Equation 203})
\end{aligned}$$

□

3.4. Variance Estimator

Suppose that $\boldsymbol{K}_G^* \in \mathbb{U}^*$ satisfies Condition 1^*. For $\boldsymbol{\beta}^{(h)*} \in \mathbb{B}^*$, $h \in \{1,2\}$, define variance of expectation between $\boldsymbol{\beta}^{(1)*}$ and $\boldsymbol{\beta}^{(2)*}$ weighted by $\boldsymbol{k}_G^{A*}$ and adjusted by $\bar{\mathbf{n}} \in \mathbb{N}^{A1}$ as

$$VE_{N,h=1}^{2*}(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}_G^{A*},\bar{\mathbf{n}}) \equiv \sum_{g^{-A}=0}^{1}\frac{\bar{n}^A(1,g^{-A})}{\bar{n}^A}\prod_{h=1}^{2}[E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}_G^{A*}(1,g^{-A})\}-E_N^*(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}_G^{A*},\bar{\mathbf{n}})]. \tag{312}$$

This quantity can be represented by two other ways:

$$\begin{aligned}
&VE^{2*}_{N,h=1}(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G, \bar{\mathbf{n}}) \\
&= \sum_{g^{-A}=0}^{1} \frac{\bar{n}^A(1,g^{-A})}{\bar{n}^A} \prod_{h=1}^{2} E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G(1,g^{-A})\} + \sum_{g^{-A}=0}^{1} \frac{\bar{n}^A(1,g^{-A})}{\bar{n}^A} \prod_{h=1}^{2} E^*_N(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G, \bar{\mathbf{n}}) \\
&\quad - \sum_{h=1}^{2} \sum_{g^{-A}=0}^{1} \frac{\bar{n}^A(1,g^{-A})}{\bar{n}^A} E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G(1,g^{-A})\} E^*_N(\boldsymbol{\beta}^{(-h)*}|\boldsymbol{k}^{A*}_G, \bar{\mathbf{n}}) \quad (\because \text{Equation 312}) \\
&= \sum_{g^{-A}=0}^{1} \frac{\bar{n}^A(1,g^{-A})}{\bar{n}^A} \prod_{h=1}^{2} E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G(1,g^{-A})\} + \prod_{h=1}^{2} E^*_N(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G, \bar{\mathbf{n}}) \\
&\quad - 2\prod_{h=1}^{2} E^*_N(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G, \bar{\mathbf{n}}) \quad (\because \text{Equation 205}) \\
&= \sum_{g^{-A}=0}^{1} \frac{\bar{n}^A(1,g^{-A})}{\bar{n}^A} \prod_{h=1}^{2} E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G(1,g^{-A})\} - \prod_{h=1}^{2} E^*_N(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G, \bar{\mathbf{n}})
\end{aligned} \tag{313}$$

and

$$\begin{aligned}
&VE^{2*}_{N,h=1}(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G, \bar{\mathbf{n}}) \\
&= \sum_{g^{-A}=0}^{1} \frac{\bar{n}^A(1,g^{-A})}{\bar{n}^A} \prod_{h=1}^{2} \frac{\bar{n}^A(1,1-g^{-A})}{\bar{n}^A} [E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G(1,g^{-A})\} - E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G(1,1-g^{-A})\}] \\
&(\because \text{Equations 185 and 312}, 1 - \frac{\bar{n}^A(1,g^{-A})}{\bar{n}^A} = \frac{\bar{n}^A(1,1-g^{-A})}{\bar{n}^A}) \\
&= \Big\{ \prod_{g^{-A}=0}^{1} \frac{\bar{n}^A(1,g^{-A})}{\bar{n}^A} \Big\} \sum_{g^{-A\prime}=0}^{1} \frac{\bar{n}^A(1,1-g^{-A\prime})}{\bar{n}^A} \prod_{h=1}^{2} [E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G(1,1)\} - E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G(1,0)\}] \\
&= \Big\{ \prod_{g^{-A}=0}^{1} \frac{\bar{n}^A(1,g^{-A})}{\bar{n}^A} \Big\} \prod_{h=1}^{2} \Delta E^*(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G) \quad (\because \text{Equation 263})
\end{aligned} \tag{314}$$

Define

$$\mathbb{N}^{TCm} \equiv \{\mathbf{n}|\mathbf{n} \in \mathbb{N}_{\max}, \bar{n}^{TC} \geq m\}. \tag{315}$$

For $\bar{\mathbf{n}} \in \mathbb{N}^{TC1}$, I define

$$\begin{aligned}
&VE^{2*}_{N,h=1}(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{TC*}_G, \bar{\mathbf{n}}) \\
&\equiv \sum_{g=0}^{1} \frac{\bar{n}^{TC}(1,g)}{\bar{n}^{TC}} \prod_{h=1}^{2} [E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{TC*}_G(1,g)\} - E^*_N(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{TC*}_G, \bar{\mathbf{n}})] \\
&= \frac{\bar{n}^{TC}(1,1)}{\bar{n}^{TC}} \prod_{h=1}^{2} [E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{TC*}_G(1,1)\} - E^*(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{TC*}_G)] \quad (\because \text{Equations 186 and 187}) \\
&= 0 \quad (\because \text{Equations 186 and 187})
\end{aligned} \tag{316}$$

LEMMA 29 (SP MEAN OF FS COVARIANCE). *Suppose that* $\boldsymbol{K}^*_G \in \mathbb{U}^*$ *satisfies Condition* 1^*. $\mathbb{N} \subseteq \mathbb{N}_{\max}$, $f_N(\bar{\mathbf{n}})$ *is a function of* $\bar{\mathbf{n}} \in \mathbb{N}$, $\boldsymbol{\beta}^{(h)*} \in \mathbb{B}^*$, $\boldsymbol{\omega}^{(h)*} \in \mathbb{W}^*$ *for* $h \in \{1,2\}$, $\mathbb{S}^*_G \equiv \mathbb{S}^*_{\max}(\mathbf{n}_G \in \mathbb{N}) \subseteq \mathbb{S}^*_{def}\{f_N(\mathbf{n}_G)\}$. *It follows*

(1) When $\mathbb{N} \subseteq \mathbb{N}^{A1}$,

$$\begin{aligned}&\mathbb{E}^*\{f_N(\mathbf{n}_G)V^{2*}_{h=1}(\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}^{A*}_G)|\mathbb{S}^*_G\}\\ &= \mathbb{E}^*\Big[f_N(\mathbf{n}_G)\Big\{EV^{2*}_{N,h=1}(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G,\mathbf{n}_G) - \frac{1}{n^A_G}\widetilde{EV^{2*}_{N,h=1}}(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G,\mathbf{n}_G)\\ &\quad + VE^{2*}_{N,h=1}(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G,\mathbf{n}_G)\Big\}\Big|\mathbb{S}^*_G\Big]\end{aligned}$$

(2) When $\mathbb{N} \subseteq \mathbb{N}^{TC1}$,

$$\begin{aligned}&\mathbb{E}^*\{f_N(\mathbf{n}_G)V^{2*}_{h=1}(\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}^{TC*}_G)|\mathbb{S}^*_G\}\\ &= \mathbb{E}^*\Big(f_N(\mathbf{n}_G)\Big[1 - \frac{1}{n^{TC}_G}\{1 - d(n^{TC*}_G, n^{TC}_G)\}\Big]\Big|\mathbb{S}^*_G\Big)V^{2*}_{h=1}(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{TC*}_G)\end{aligned}$$

(3)

$$\mathbb{E}^*\{f_N(\mathbf{n}_G)V^{2*}_{h=1}(\boldsymbol{\omega}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}^{A*}_G)|\mathbb{S}^*_G\} = \mathbb{E}^*\{f_N(\mathbf{n}_G)EV^{2*}_{N,h=1}(\boldsymbol{\omega}^{(h)*}|\boldsymbol{k}^{A*}_G,\mathbf{n}_G)|\mathbb{S}^*_G\}$$

(4)

$$\mathbb{E}^*\{f_N(\mathbf{n}_G)V^{2*}_{h=1}(\boldsymbol{\omega}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}^{TC*}_G)|\mathbb{S}^*_G\} = \mathbb{E}^*\{f_N(\mathbf{n}_G)|\mathbb{S}^*_G\}V^{2*}_{h=1}(\boldsymbol{\omega}^{(h)*}|\boldsymbol{k}^{TC*}_G)$$

Proof. For $\bar{\mathbf{n}} \in \mathbb{N}$, let $\mathbb{S}^*_N \equiv \mathbb{S}^*_{\max}(\mathbf{n}_G = \bar{\mathbf{n}})$. Suppose $\mathbb{S}^*_N \neq \emptyset$.
(1) Note

$$\begin{aligned}&\mathbb{E}^*_{S|N}\{V^{2*}_{h=1}(\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}^{A*}_G)|\mathbb{S}^*_N\}\\ &= \mathbb{E}^*_{S|N}\Big\{E^*\Big(\prod_{h=1}^{2}\boldsymbol{\beta}^{(h)*}\Big|\boldsymbol{S}^*\boldsymbol{k}^{A*}_G\Big) - \prod_{h=1}^{2}E^*(\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}^{A*}_G)\Big|\mathbb{S}^*_N\Big\} \quad (\because \text{Lemma 15 (7)})\\ &= E^*_N\Big(\prod_{h=1}^{2}\boldsymbol{\beta}^{(h)*}\Big|\boldsymbol{k}^{A*}_G,\bar{\mathbf{n}}\Big) - \mathbb{E}^*_{S|N}\Big\{\prod_{h=1}^{2}E^*(\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}^{A*}_G)\Big|\mathbb{S}^*_N\Big\}\\ &\quad (\because \text{Lemma 23 (1) and 24 (2), where we substitute } \prod_{h=1}^{2}\boldsymbol{\beta}^{(h)*} \in \mathbb{B} \text{ with } \boldsymbol{\beta}^*, \mathbb{S}^*_N \neq \emptyset)\\ &= \sum_{g^{-A}=0}^{1}\frac{\bar{n}^A(1,g^{-A})}{\bar{n}^A}E^*\Big\{\prod_{h=1}^{2}\boldsymbol{\beta}^{(h)*}\Big|\boldsymbol{k}^{A*}_G(1,g^{-A})\Big\} - \Big[\mathbb{V}^2_{S|N,h=1}\{E^*(\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}^{A*}_G)|\mathbb{S}^*_N\}\\ &\quad + \prod_{h=1}^{2}\mathbb{E}^*_{S|N}\{E^*(\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}^{A*}_G)|\mathbb{S}^*_N\}\Big] \quad (\because \text{Equation 205, Lemma 15 (7)}, \mathbb{N} \subseteq \mathbb{N}^{A1})\\ &= \sum_{g^{-A}=0}^{1}\frac{\bar{n}^A(1,g^{-A})}{\bar{n}^A}\Big[V^{2*}_{h=1}\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G(1,g^{-A})\} + \prod_{h=1}^{2}E^*\{\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G(1,g^{-A})\}\Big]\\ &\quad - \frac{1}{\bar{n}^A}\widetilde{EV^{2*}_{N,h=1}}(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G,\bar{\mathbf{n}}) - \prod_{h=1}^{2}E^*_N(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G,\bar{\mathbf{n}})\\ &\quad (\because \text{Lemmas 14 (3), 15 (7) and 24 (2), Equation 282})\\ &= EV^{2*}_{N,h=1}(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G,\bar{\mathbf{n}}) - \frac{1}{\bar{n}^A}\widetilde{EV^{2*}_{N,h=1}}(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G,\bar{\mathbf{n}}) + VE^{2*}_{N,h=1}(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{A*}_G,\bar{\mathbf{n}})\\ &\quad (\because \text{Equations 254 and 313})\end{aligned}$$

Note also that $\mathbb{S}^*_{\text{def}}\{f_N(\mathbf{n}_G)V^{2*}_{h=1}(\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}^{A*}_G)\} = \mathbb{S}^*_{\text{def}}\{f_N(\mathbf{n}_G)\} \supseteq \mathbb{S}^*_G$. Therefore, by applying Lemma 23 (2) and substituting $\bar{\mathbf{n}} = \mathbf{n}_G$, the desired result follows.

(2) We can substitute $A = TC$ in the proof of Lemma 29 (1) to obtain

$$\begin{aligned}
&\mathbb{E}^*\{f_N(\mathbf{n}_G)V^{2*}_{h=1}(\boldsymbol{\beta}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}^{TC*}_G)|\mathbb{S}^*_G\} \\
&= \mathbb{E}^*\Big[f_N(\mathbf{n}_G)\Big\{EV^{2*}_{N,h=1}(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{TC*}_G, \mathbf{n}_G) - \frac{1}{n^{TC}_G}\widetilde{EV^{2*}_{N,h=1}}(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{TC*}_G, \mathbf{n}_G) \\
&\quad + VE^{2*}_{N,h=1}(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{TC*}_G, \mathbf{n}_G)\Big\}\Big|\mathbb{S}^*_G\Big] \\
&= \mathbb{E}^*\Big(f_N(\mathbf{n}_G)\Big[1 - \frac{1}{n^{TC}_G}\{1 - d(n^{TC*}_G, n^{TC}_G)\}\Big]\Big|\mathbb{S}^*_G\Big)V^{2*}_{h=1}(\boldsymbol{\beta}^{(h)*}|\boldsymbol{k}^{TC*}_G) \\
&\quad (\because \mathbb{N} \subseteq \mathbb{N}^{TC1}, \text{Equations 186, 187, 255, 259, and 316, Lemma 10 (2)})
\end{aligned}$$

(3) When $\bar{n}^A \geq 1$,

$$\begin{aligned}
&\mathbb{E}^*_{S|N}\{V^{2*}_{h=1}(\boldsymbol{\omega}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}^{A*}_G)|\mathbb{S}^*_N\} \\
&= \mathbb{E}^*_{S|N}\Big\{E^*\Big(\prod_{h=1}^{2}\boldsymbol{\omega}^{(h)*}\Big|\boldsymbol{S}^*\boldsymbol{k}^{A*}_G\Big) - \prod_{h=1}^{2}E^*(\boldsymbol{\omega}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}^{A*}_G)\Big|\mathbb{S}^*_N\Big\} \quad (\because \text{Lemma 15 (7)}) \\
&= E^*_N\Big(\prod_{h=1}^{2}\boldsymbol{\omega}^{(h)*}\Big|\boldsymbol{k}^{A*}_G, \bar{\mathbf{n}}\Big) \quad (\because \text{Equations 146 and 147, Lemmas 12 (1), 23 (1), and 24 (2)}, \bar{n}^A \geq 1) \\
&= \sum_{g^{-A}=0}^{1}\frac{\bar{n}^A(1, g^{-A})}{\bar{n}^A}E^*\Big\{\prod_{h=1}^{2}\boldsymbol{\omega}^{(h)*}\Big|\boldsymbol{k}^{A*}_G(1, g^{-A})\Big\} \quad (\because \text{Equation 205}, \mathbb{N} \subseteq \mathbb{N}^{A1}) \\
&= \sum_{g^{-A}=0}^{1}\frac{\bar{n}^A(1, g^{-A})}{\bar{n}^A}V^{2*}_{h=1}\{\boldsymbol{\omega}^{(h)*}|\boldsymbol{k}^{A*}_G(1, g^{-A})\} \quad (\because \text{Lemma 17 (2)}) \\
&= EV^{2*}_{N,h=1}(\boldsymbol{\omega}^{(h)*}|\boldsymbol{k}^{A*}_G, \bar{\mathbf{n}}) \quad (\because \text{Equation 254})
\end{aligned}$$

Even when $\bar{n}^A < 1$, both ends of the above equation are equal to each other (zero). Note that $\mathbb{S}^*_{\text{def}}\{f_N(\mathbf{n}_G)V^{2*}_{h=1}(\boldsymbol{\omega}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}^{A*}_G)\} = \mathbb{S}^*_{\text{def}}\{f_N(\mathbf{n}_G)\} \supseteq \mathbb{S}^*_G$. Therefore, by applying Lemma 23 (2) and substituting $\bar{\mathbf{n}} = \mathbf{n}_G$, the desired result follows.

(4) We can substitute $A = TC$ in the proof of Lemma 29 (3) to obtain

$$\begin{aligned}
&\mathbb{E}^*\{f_N(\mathbf{n}_G)V^{2*}_{h=1}(\boldsymbol{\omega}^{(h)*}|\boldsymbol{S}^*\boldsymbol{k}^{TC*}_G)|\mathbb{S}^*_G\} \\
&= \mathbb{E}^*\Big\{f_N(\mathbf{n}_G)EV^{2*}_{N,h=1}(\boldsymbol{\omega}^{(h)*}|\boldsymbol{k}^{TC*}_G, \mathbf{n}_G)\Big|\mathbb{S}^*_G\Big\} \\
&= \mathbb{E}^*\{f_N(\mathbf{n}_G)|\mathbb{S}^*_G\}V^{2*}_{h=1}(\boldsymbol{\omega}^{(h)*}|\boldsymbol{k}^{TC*}_G) \quad (\because \text{Equation 255, Lemma 10 (2)})
\end{aligned}$$

□

Proposition 3* (Bias of the Neyman Variance Estimators: SP). *(1) Under Assumption 1*, it holds that* $\mathbb{S}^*_{def}\{\hat{\mathbb{V}}^{Neyman}(\hat{\tau}_F)\} = \mathbb{S}^*_{\max}$ *and*

$$\lim_{n^*\to\infty}[\mathbb{E}^*\{\hat{\mathbb{V}}^{Neyman}(\hat{\tau}_F)|\mathbb{S}^*_{\max}\} - \mathbb{V}^{2*}(\hat{\tau}_F|\mathbb{S}^*_{\max})] = \frac{2}{n_F}\{V^{2*}(\boldsymbol{\beta}^{T*}, \boldsymbol{\beta}^{C*}|\boldsymbol{k}^*_F) - V^{2*}(\boldsymbol{\omega}^{T*}, \boldsymbol{\omega}^{C*}|\boldsymbol{k}^*_F)\}.$$

(2) Under Assumption 3, it holds that* $\mathbb{S}^*_{def}\{\hat{\mathbb{V}}^{Neyman}(\hat{\tau}_P)\} = \mathbb{S}^{2*}_P$ *and, when* $\mathbb{S}^{2*}_P \neq \emptyset$,

$$\begin{aligned}
&\lim_{n^*\to\infty}[\mathbb{E}^*\{\hat{\mathbb{V}}^{Neyman}(\hat{\tau}_P)|\mathbb{S}^{2*}_P\} - \mathbb{V}^{2*}(\hat{\tau}_P|\mathbb{S}^{2*}_P)] \\
&= 2\mathbb{E}^*\Big(\frac{1}{n_P}\Big|\mathbb{S}^{2*}_P\Big)\{V^{2*}(\boldsymbol{\beta}^{T*}, \boldsymbol{\beta}^{C*}|\boldsymbol{k}^*_P) - V^{2*}(\boldsymbol{\omega}^{T*}, \boldsymbol{\omega}^{C*}|\boldsymbol{k}^*_P)\}.
\end{aligned}$$

(3) Under Assumption 2^*, *it holds that* $\mathbb{S}^*_{def}\{\hat{\mathbb{V}}^{Neyman}(\hat{\tau}_U)\} = \mathbb{S}^{2*}_U$ *and, when* $\mathbb{S}^{2*}_U \neq \emptyset$,

$$\begin{aligned}
&\lim_{n^*\to\infty}[\mathbb{E}^*\{\hat{\mathbb{V}}^{Neyman}(\hat{\tau}_U)|\mathbb{S}^{2*}_U\} - \mathbb{V}^{2*}(\hat{\tau}_U|\mathbb{S}^{2*}_U)] \\
&= 2\mathbb{E}^*\Big(\frac{n_P}{n^T_U n^C_U}\Big|\mathbb{S}^{2*}_U\Big)\{V^*(\boldsymbol{\beta}^{T*},\boldsymbol{\beta}^{C*}|\boldsymbol{k}^*_P) - V^*(\boldsymbol{\omega}^{T*},\boldsymbol{\omega}^{C*}|\boldsymbol{k}^*_P)\} \\
&\quad + \Big[\mathbb{E}^*\Big\{\frac{n_P(n^T_U - n_P)}{(n^T_U-1)(n^T_U)^2}\Big|\mathbb{S}^{2*}_U\Big\} - \mathbb{V}^{2*}\Big(\frac{n_P}{n^T_U}\Big|\mathbb{S}^{2*}_U\Big)\Big]\{\Delta E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{k}^{T*}_U)\}^2 \\
&\quad + \Big[\mathbb{E}^*\Big\{\frac{n_P(n^C_U - n_P)}{(n^C_U-1)(n^C_U)^2}\Big|\mathbb{S}^{2*}_U\Big\} - \mathbb{V}^{2*}\Big(\frac{n_P}{n^C_U}\Big|\mathbb{S}^{2*}_U\Big)\Big]\{\Delta E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{k}^{C*}_U)\}^2 \\
&\quad + 2\mathbb{V}^*\Big(\frac{n_P}{n^T_U},\frac{n_P}{n^C_U}\Big|\mathbb{S}^{2*}_U\Big)\Delta E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{k}^{T*}_U)\Delta E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{k}^{C*}_U).
\end{aligned}$$

The closed forms are obtained by using Lemmas 23 (2), 25, and 28 and Equation 205. If we do not take the limit, Equations 318 and 321 show bias of the general case, $\mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)|\mathbb{S}^*_G\} - \mathbb{V}^{2*}(\hat{\tau}_G|\mathbb{S}^*_G)$.

Under Assumption 2^*, 4^*, and 5^*, it follows that Proposition 3^* (3) and Equation 290 hold. Applying Equation 290 to Proposition 3^* (3), we obtaion

$$\begin{aligned}
&\lim_{n^*\to\infty}[\mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_U)|\mathbb{S}^{2*}_U\} - \mathbb{V}^{2*}(\hat{\tau}_U|\mathbb{S}^{2*}_U)] \\
&= 2\mathbb{E}^*\Big(\frac{n_P}{n^T_U n^C_U}\Big|\mathbb{S}^{2*}_U\Big)\{V^*(\boldsymbol{\beta}^{T*},\boldsymbol{\beta}^{C*}|\boldsymbol{k}^*_P) - V^*(\boldsymbol{\omega}^{T*},\boldsymbol{\omega}^{C*}|\boldsymbol{k}^*_P)\}.
\end{aligned}$$

Proof. Suppose that $\boldsymbol{K}^*_G \in \mathbb{U}^*$ satisfies Condition 1^*, $\mathbb{N} \subseteq \mathbb{N}^2$, and $\mathbb{S}^*_G \equiv \mathbb{S}^*_{\max}(\mathbf{n}_G \in \mathbb{N}) \subseteq \mathbb{S}^*_{\text{def}}\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)\}$. It follows that

$$\begin{aligned}
&\mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)|\mathbb{S}^*_G\} \\
&= \mathbb{E}^*_S[\mathbb{E}_X\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)\}|\mathbb{S}^*_G] \quad (\because \text{Equation 199}) \\
&= \mathbb{E}^*_S[\mathbb{E}_X\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G) - \mathbb{V}_X(\hat{\tau}_G)\} + \mathbb{V}_X(\hat{\tau}_G)|\mathbb{S}^*_G] \\
&= \mathbb{E}^*_S\Big\{\sum_A \frac{1}{n^A_G - 1}V^2(\boldsymbol{\beta}^A|\boldsymbol{k}^A_G) - \frac{2n^{TC}_G}{n^T_G n^C_G}V(\boldsymbol{\omega}^T,\boldsymbol{\omega}^C|\boldsymbol{k}^{TC}_G) \\
&\quad + \sum_A \frac{1}{n^A_G}V^2(\boldsymbol{\omega}^A|\boldsymbol{k}^A_G) + \frac{2n^{TC}_G}{n^T_G n^C_G}V(\boldsymbol{\omega}^T,\boldsymbol{\omega}^C|\boldsymbol{k}^{TC}_G)\Big|\mathbb{S}^*_G\Big\} \\
&\quad (\because \text{Equations 128 and 154}, \mathbb{S}^*_G \subseteq \mathbb{S}^*_{\text{def}}\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)\}, \mathbb{N} \subseteq \mathbb{N}^2) \\
&= \mathbb{E}^*\Big[\sum_A \Big\{\frac{1}{n^A_G - 1}V^2(\boldsymbol{\beta}^A|\boldsymbol{k}^A_G) + \frac{1}{n^A_G}V^2(\boldsymbol{\omega}^A|\boldsymbol{k}^A_G)\Big\}\Big|\mathbb{S}^*_G\Big] \quad (\because \text{Equation 199}) \\
&= \mathbb{E}^*\Big(\sum_A \Big[\frac{1}{n^A_G - 1}\Big\{EV^{2*}_N(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G,\mathbf{n}_G) - \frac{1}{n^A_G}\widetilde{EV^{2*}_N}(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G,\mathbf{n}_G) + VE^{2*}_N(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G,\mathbf{n}_G)\Big\} \\
&\quad + \frac{1}{n^A_G}EV^{2*}_N(\boldsymbol{\omega}^{A*}|\boldsymbol{k}^{A*}_G,\mathbf{n}_G)\Big]\Big|\mathbb{S}^*_G\Big) \quad (\because \text{Lemmas 23 (1) and 29 (1) and (3)}, \mathbb{N} \subseteq \mathbb{N}^2 \subseteq \mathbb{N}^1)
\end{aligned} \tag{317}$$

Therefore, the bias of Neyman variance estimator is equal to

$$
\begin{aligned}
&\mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)|\mathbb{S}^*_G\} - \mathbb{V}^{2*}(\hat{\tau}_G|\mathbb{S}^*_G) \\
&= \mathbb{E}^*\Big(\sum_A \Big[\frac{1}{n^A_G - 1}\Big\{EV^{2*}_N(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G, \mathbf{n}_G) - \frac{1}{n^A_G}\widetilde{EV^{2*}_N}(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G, \mathbf{n}_G) + VE^{2*}_N(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G, \mathbf{n}_G)\Big\} \\
&\quad + \frac{1}{n^A_G} EV^{2*}_N(\boldsymbol{\omega}^{A*}|\boldsymbol{k}^{A*}_G, \mathbf{n}_G)\Big]\Big|\mathbb{S}^*_G\Big) \\
&\quad - \mathbb{E}^*\Big[\sum_A \frac{1}{n^A_G}\{EV^{2*}_N(\boldsymbol{\omega}^{A*}|\boldsymbol{k}^{A*}_G, \mathbf{n}_G) + \widetilde{EV^{2*}_N}(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G, \mathbf{n}_G)\} \\
&\quad + \frac{2n^{TC}_G}{n^T_G n^C_G}\{V^*(\boldsymbol{\omega}^{T*}, \boldsymbol{\omega}^{C*}|\boldsymbol{k}^{TC*}_G) - \widetilde{V^*}(\boldsymbol{\beta}^{T*}, \boldsymbol{\beta}^{C*}|\boldsymbol{k}^{TC*}_G, \mathbf{n}_G)\}\Big|\mathbb{S}^*_G\Big] \\
&\quad - \sum_A \mathbb{V}^{2*}\Big(\frac{n^{TC}_G}{n^A_G}\Big|\mathbb{S}^*_G\Big)\{\Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G)\}^2 + 2\mathbb{V}^*\Big(\frac{n^{TC}_G}{n^T_G}, \frac{n^{TC}_G}{n^C_G}\Big|\mathbb{S}^*_G\Big)\prod_A \Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G) \\
&\quad (\because \text{Equations 297 and 317, } \mathbb{S}^*_G \subseteq \mathbb{S}^*_{\text{def}}\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)\} \subseteq \mathbb{S}^*_{\text{def}}(\hat{\tau}_G)) \\
&= 2\mathbb{E}^*\Big\{\frac{n^{TC}_G}{n^T_G n^C_G}\widetilde{V^*}(\boldsymbol{\beta}^{T*}, \boldsymbol{\beta}^{C*}|\boldsymbol{k}^{TC*}_G, \mathbf{n}_G)\Big|\mathbb{S}^*_G\Big\} - 2\mathbb{E}^*\Big(\frac{n^{TC}_G}{n^T_G n^C_G}\Big|\mathbb{S}^*_G\Big)V^*(\boldsymbol{\omega}^{T*}, \boldsymbol{\omega}^{C*}|\boldsymbol{k}^{TC*}_G) \\
&\quad + \mathbb{E}^*\Big[\sum_A \frac{1}{n^A_G - 1}\{VE^{2*}_N(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G, \mathbf{n}_G) + EV^{2*}_N(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G, \mathbf{n}_G) - \widetilde{EV^{2*}_N}(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G, \mathbf{n}_G)\}\Big|\mathbb{S}^*_G\Big] \\
&\quad - \sum_A \mathbb{V}^{2*}\Big(\frac{n^{TC}_G}{n^A_G}\Big|\mathbb{S}^*_G\Big)\{\Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G)\}^2 + 2\mathbb{V}^*\Big(\frac{n^{TC}_G}{n^T_G}, \frac{n^{TC}_G}{n^C_G}\Big|\mathbb{S}^*_G\Big)\prod_A \Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G) \\
&= 2\mathbb{E}^*\Big[\frac{n^{TC}_G}{n^T_G n^C_G}\{1 - d(n^{TC*}_G, n^{TC}_G)\}\Big|\mathbb{S}^*_G\Big]V^*(\boldsymbol{\beta}^{T*}, \boldsymbol{\beta}^{C*}|\boldsymbol{k}^{TC*}_G) - 2\mathbb{E}^*\Big(\frac{n^{TC}_G}{n^T_G n^C_G}\Big|\mathbb{S}^*_G\Big)V^*(\boldsymbol{\omega}^{T*}, \boldsymbol{\omega}^{C*}|\boldsymbol{k}^{TC*}_G) \\
&\quad + \mathbb{E}^*\Big(\sum_A \frac{1}{n^A_G - 1}\Big[\Big\{\prod_{g^{-A}=0}^{1} \frac{n^A_G(1, g^{-A})}{n^A_G}\Big\}\{\Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G)\}^2 \\
&\quad + \sum_{g^{-A}=0}^{1} \frac{n^A_G(1, g^{-A})}{n^A_G} d\{n^{A*}_G(1, g^{-A}), n^A_G(1, g^{-A})\}V^{2*}_{h=1}\{\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G(1, g^{-A})\}\Big]\Big|\mathbb{S}^*_G\Big) \\
&\quad - \sum_A \mathbb{V}^{2*}\Big(\frac{n^{TC}_G}{n^A_G}\Big|\mathbb{S}^*_G\Big)\{\Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G)\}^2 + 2\mathbb{V}^*\Big(\frac{n^{TC}_G}{n^T_G}, \frac{n^{TC}_G}{n^C_G}\Big|\mathbb{S}^*_G\Big)\prod_A \Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G) \\
&\quad (\because \text{Equations 254, 258, 260, and 314}) \\
&= 2\mathbb{E}^*\Big[\frac{n^{TC}_G}{n^T_G n^C_G}\{1 - d(n^{TC*}_G, n^{TC}_G)\}\Big|\mathbb{S}^*_G\Big]V^*(\boldsymbol{\beta}^{T*}, \boldsymbol{\beta}^{C*}|\boldsymbol{k}^{TC*}_G) - 2\mathbb{E}^*\Big(\frac{n^{TC}_G}{n^T_G n^C_G}\Big|\mathbb{S}^*_G\Big)V^*(\boldsymbol{\omega}^{T*}, \boldsymbol{\omega}^{C*}|\boldsymbol{k}^{TC*}_G) \\
&\quad + \sum_A \mathbb{E}^*\Big\{\frac{1}{n^A_G - 1}\prod_{g^{-A}=0}^{1} \frac{n^A_G(1, g^{-A})}{n^A_G}\Big|\mathbb{S}^*_G\Big\}\{\Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G)\}^2 \\
&\quad + \sum_A \sum_{g^{-A}=0}^{1} \mathbb{E}^*\Big[\frac{n^A_G(1, g^{-A})}{(n^A_G - 1)n^A_G} d\{n^{A*}_G(1, g^{-A}), n^A_G(1, g^{-A})\}\Big|\mathbb{S}^*_G\Big]V^{2*}_{h=1}\{\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G(1, g^{-A})\} \\
&\quad - \sum_A \mathbb{V}^{2*}\Big(\frac{n^{TC}_G}{n^A_G}\Big|\mathbb{S}^*_G\Big)\{\Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G)\}^2 + 2\mathbb{V}^*\Big(\frac{n^{TC}_G}{n^T_G}, \frac{n^{TC}_G}{n^C_G}\Big|\mathbb{S}^*_G\Big)\prod_A \Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^{A*}_G)
\end{aligned}
\tag{318}
$$

In limit, it follows

$$\begin{aligned}
&\lim_{n^*\to\infty}[\mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)|\mathbb{S}_G^*\} - \mathbb{V}^{2*}(\hat{\tau}_G|\mathbb{S}_G^*)] \\
&= 2\mathbb{E}^*\Big\{\frac{n_G^{TC}}{n_G^T n_G^C}(1-0)\Big|\mathbb{S}_G^*\Big\}V^*(\boldsymbol{\beta}^{T*},\boldsymbol{\beta}^{C*}|\boldsymbol{k}_G^{TC*}) - 2\mathbb{E}^*\Big(\frac{n_G^{TC}}{n_G^T n_G^C}\Big|\mathbb{S}_G^*\Big)V^*(\boldsymbol{\omega}^{T*},\boldsymbol{\omega}^{C*}|\boldsymbol{k}_G^{TC*}) \\
&\quad + \sum_A \mathbb{E}^*\Big\{\frac{1}{n_G^A-1}\prod_{g^{-A}=0}^{1}\frac{n_G^A(1,g^{-A})}{n_G^A}\Big|\mathbb{S}_G^*\Big\}\{\Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*})\}^2 \\
&\quad + \sum_A\sum_{g^{-A}=0}^{1}\mathbb{E}^*\Big\{\frac{n_G^A(1,g^{-A})}{(n_G^A-1)n_G^A}\cdot 0\Big|\mathbb{S}_G^*\Big\}V_{h=1}^{2*}\{\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*}(1,g^{-A})\} \\
&\quad - \sum_A \mathbb{V}^{2*}\Big(\frac{n_G^{TC}}{n_G^A}\Big|\mathbb{S}_G^*\Big)\{\Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*})\}^2 + 2\mathbb{V}^*\Big(\frac{n_G^{TC}}{n_G^T},\frac{n_G^{TC}}{n_G^C}\Big|\mathbb{S}_G^*\Big)\prod_A \Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*}) \qquad (319)\\
&\quad (\because \text{Equations 298 and 299}) \\
&= 2\mathbb{E}^*\Big(\frac{n_G^{TC}}{n_G^T n_G^C}\Big|\mathbb{S}_G^*\Big)\{V^*(\boldsymbol{\beta}^{T*},\boldsymbol{\beta}^{C*}|\boldsymbol{k}_G^{TC*}) - V^*(\boldsymbol{\omega}^{T*},\boldsymbol{\omega}^{C*}|\boldsymbol{k}_G^{TC*})\} \\
&\quad + \sum_A\Big[\mathbb{E}^*\Big\{\frac{n_G^{TC}(n_G^A-n_G^{TC})}{(n_G^A-1)(n_G^A)^2}\Big|\mathbb{S}_G^*\Big\} - \mathbb{V}^{2*}\Big(\frac{n_G^{TC}}{n_G^A}\Big|\mathbb{S}_G^*\Big)\Big]\{\Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*})\}^2 \\
&\quad + 2\mathbb{V}^*\Big(\frac{n_G^{TC}}{n_G^T},\frac{n_G^{TC}}{n_G^C}\Big|\mathbb{S}_G^*\Big)\prod_A \Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*}) \quad (\because \text{Lemma 10 (2)})
\end{aligned}$$

When $\boldsymbol{k}_G^{T*} = \boldsymbol{k}_G^{C*} \equiv \boldsymbol{k}_G^*$, it follows

$$\mathbb{V}^{2*}\Big(\frac{n_G}{n_G}\Big|\mathbb{S}_G^*\Big) = \mathbb{V}^*\Big(\frac{n_G}{n_G},\frac{n_G}{n_G}\Big|\mathbb{S}_G^*\Big) = \mathbb{V}^{2*}(1|\mathbb{S}_G^*) = 0 \qquad (320)$$

and Equation 318 leads to

$$\begin{aligned}
&\mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)|\mathbb{S}_G^*\} - \mathbb{V}^{2*}(\hat{\tau}_G|\mathbb{S}_G^*) \\
&= 2\mathbb{E}^*\Big[\frac{n_G}{n_G n_G}\{1-d(n_G^*,n_G)\}\Big|\mathbb{S}_G^*\Big]V^*(\boldsymbol{\beta}^{T*},\boldsymbol{\beta}^{C*}|\boldsymbol{k}_G^*) - 2\mathbb{E}^*\Big(\frac{n_G}{n_G n_G}\Big|\mathbb{S}_G^*\Big)V^*(\boldsymbol{\omega}^{T*},\boldsymbol{\omega}^{C*}|\boldsymbol{k}_G^*) \\
&\quad + \sum_A \mathbb{E}^*\Big(\frac{1}{n_G-1}\prod_{g=0}^{1}\frac{n_G(1,g)}{n_G}\Big|\mathbb{S}_G^*\Big)\{\Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^*)\}^2 \\
&\quad + \sum_A\sum_{g=0}^{1}\mathbb{E}^*\Big[\frac{n_G(1,g)}{(n_G-1)n_G}d\{n_G^*(1,g),n_G(1,g)\}\Big|\mathbb{S}_G^*\Big]V_{h=1}^{2*}\{\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^*(1,g)\} \\
&\quad - \sum_A \mathbb{V}^{2*}\Big(\frac{n_G}{n_G}\Big|\mathbb{S}_G^*\Big)\{\Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^*)\}^2 + 2\mathbb{V}^*\Big(\frac{n_G}{n_G^T},\frac{n_G}{n_G^C}\Big|\mathbb{S}_G^*\Big)\prod_A \Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^*) \\
&\quad (\because \text{Equations 257 and 305, Lemma 14 (3)}) \\
&= 2\mathbb{E}^*\Big\{\frac{1-d(n_G^*,n_G)}{n_G}\Big|\mathbb{S}_G^*\Big\}V^*(\boldsymbol{\beta}^{T*},\boldsymbol{\beta}^{C*}|\boldsymbol{k}_G^*) - 2\mathbb{E}^*\Big(\frac{1}{n_G}\Big|\mathbb{S}_G^*\Big)V^*(\boldsymbol{\omega}^{T*},\boldsymbol{\omega}^{C*}|\boldsymbol{k}_G^*) \\
&\quad + \mathbb{E}^*\Big\{\frac{d(n_G^*,n_G)}{n_G-1}\Big|\mathbb{S}_G^*\Big\}\sum_A V_{h=1}^{2*}(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^*) \\
&\quad (\because \text{Equations 304 and 320})
\end{aligned} \qquad (321)$$

In limit, it follows

$$\begin{aligned}
&\lim_{n^*\to\infty}[\mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)|\mathbb{S}^*_G\} - \mathbb{V}^{2*}(\hat{\tau}_G|\mathbb{S}^*_G)] \\
&= 2\mathbb{E}^*\Big(\frac{1-0}{n_G}\Big|\mathbb{S}^*_G\Big)V^*(\boldsymbol{\beta}^{T*}, \boldsymbol{\beta}^{C*}|\boldsymbol{k}^*_G) - 2\mathbb{E}^*\Big(\frac{1}{n_G}\Big|\mathbb{S}^*_G\Big)V^*(\boldsymbol{\omega}^{T*}, \boldsymbol{\omega}^{C*}|\boldsymbol{k}^*_G) \\
&\quad + \mathbb{E}^*\Big(\frac{0}{n_G-1}\Big|\mathbb{S}^*_G\Big)\sum_A V^{2*}(\boldsymbol{\beta}^{A*}|\boldsymbol{k}^*_G) \quad (\because \text{Equations 307 and 321}) \\
&= 2\mathbb{E}^*\Big(\frac{1}{n_G}\Big|\mathbb{S}^*_G\Big)\{V^*(\boldsymbol{\beta}^{T*}, \boldsymbol{\beta}^{C*}|\boldsymbol{k}^*_G) - V^*(\boldsymbol{\omega}^{T*}, \boldsymbol{\omega}^{C*}|\boldsymbol{k}^*_G)\}.
\end{aligned} \tag{322}$$

(1) Under Assumption 1^*, it holds that $\mathbb{S}^*_{\text{def}}\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_F)\} = \mathbb{S}^*_{\max}$ ($\because$ Equation 12, $N^A_F = n \geq 2$). When $\boldsymbol{K}^*_G = \boldsymbol{K}^*_F \in \mathbb{U}^*$ and $\mathbb{N} = \mathbb{N}^n$, it follows that $\boldsymbol{K}_G = \boldsymbol{K}_F$ ($\because$ Equation 164) and, according to Lemma 13 (1), $\mathbb{S}^*_G = \mathbb{S}^{n*}_F = \mathbb{S}^*_{\max}$ ($\because$ Equations 211 and 212), and Equations 309 and 322 lead to the desired result.

(2) Under Assumption 3^*, it holds that $\mathbb{S}^*_{\text{def}}\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_P)\} = \mathbb{S}^{2*}_P$ ($\because$ Equation 16, Lemma 9 (1), $N^A_P = N_P = n_P \geq 2$, Equations 204 and 211). When $\boldsymbol{K}^*_G = \boldsymbol{K}^*_P \in \mathbb{U}^*$ and $\mathbb{N} = \mathbb{N}^2$, it follows that $\boldsymbol{K}_G = \boldsymbol{K}_P$ ($\because$ Equation 164) and, according to Lemma 13 (2), $\mathbb{S}^*_G = \mathbb{S}^{2*}_P$ ($\because$ Equations 211 and 213) and Equation 322 is equivalent to the desired result.

(3) Under Assumption 2^*, it holds that $\mathbb{S}^*_{\text{def}}\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_U)\} = \mathbb{S}^{2*}_U$ ($\because$ Lemma 9 (2), $N^A_U = n^A_U \geq 2$, Equations 204 and 211). When $\boldsymbol{K}^*_G = \boldsymbol{K}^*_U \in \mathbb{U}^*$ and $\mathbb{N} = \mathbb{N}^2$, it follows that $\boldsymbol{K}_G = \boldsymbol{K}_U$ ($\because$ Equation 164) and, according to Lemma 13 (3), $\mathbb{S}^*_G = \mathbb{S}^{2*}_U$ ($\because$ Equations 211 and 214) and Equation 319 is equivalent to the desired result.

□

Here are short notes. First, the bias directions of the Neyman variance estimators for all of these three ATE estimators are unknown. When the pair matching is effective in the sense that $\boldsymbol{y}^{T*}$ and $\boldsymbol{y}^{C*}$ are sufficiently close to $\boldsymbol{y}^{T*}_{-i}$ and $\boldsymbol{y}^{C*}_{-i}$, respectively, and the treatment effects are not so heterogeneous across pairs, $\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_F)$ and $\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_P)$ are upwardly biased; otherwise, they are downwardly biased. The bias direction of $\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_U)$ in these cases is unknown. Second, under Assumption 2^*, 4^* and 5^*, the last three lines in Proposition 3^* (3) are equal to zero thanks to Equation 310. Propositions 3^* (1) and (2) hold whether or not Assumption 4^* and/or 5^* hold(s). Third, If only for $\hat{\tau}_F$, even Imai (2008) and Imbens & Rubin (2015) do not derive Proposition 3^* (1).

Proposition 4^* (Bias of the Adjusted Neyman Variance Estimators: SP). *(1) Under Assumption* 1^*, *it holds that* $\mathbb{S}^*_{def}\{\hat{\mathbb{V}}^{\textit{Adj-Neyman}}(\hat{\tau}_F)\} = \mathbb{S}^*_{\max}$ *and*

$$\lim_{n^*\to\infty}[\mathbb{E}^*\{\hat{\mathbb{V}}^{\textit{Adj-Neyman}}(\hat{\tau}_F)|\mathbb{S}^*_{\max}\} - \mathbb{V}^{2*}(\hat{\tau}_F|\mathbb{S}^*_{\max})] = 0.$$

(2) Under Assumption 3^*, *it holds that* $\mathbb{S}^*_{def}\{\hat{\mathbb{V}}^{\textit{Adj-Neyman}}(\hat{\tau}_P)\} = \mathbb{S}^{2*}_P$ *and, when* $\mathbb{S}^{2*}_P \neq \emptyset$,

$$\lim_{n^*\to\infty}[\mathbb{E}^*\{\hat{\mathbb{V}}^{\textit{Adj-Neyman}}(\hat{\tau}_P)|\mathbb{S}^{2*}_P\} - \mathbb{V}^{2*}(\hat{\tau}_P|\mathbb{S}^{2*}_P)] = 0.$$

(3) Under Assumption 2^*, *it holds that* $\mathbb{S}^*_{def}\{\hat{\mathbb{V}}^{\textit{Adj-Neyman}}(\hat{\tau}_U)\} = \mathbb{S}^{2*}_P$ *and, when* $\mathbb{S}^{2*}_P \neq \emptyset$,

$$\begin{aligned}
&\lim_{n^*\to\infty}[\mathbb{E}^*\{\hat{\mathbb{V}}^{\textit{Adj-Neyman}}(\hat{\tau}_U)|\mathbb{S}^{2*}_P\} - \mathbb{V}^{2*}(\hat{\tau}_U|\mathbb{S}^{2*}_P)] \\
&= \Big[\mathbb{E}^*\Big\{\frac{n_P(n^T_U - n_P)}{(n^T_U-1)(n^T_U)^2}\Big|\mathbb{S}^{2*}_P\Big\} - \mathbb{V}^{2*}\Big(\frac{n_P}{n^T_U}\Big|\mathbb{S}^{2*}_P\Big)\Big]\{\Delta E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{k}^{T*}_U)\}^2 \\
&\quad + \Big[\mathbb{E}^*\Big\{\frac{n_P(n^C_U - n_P)}{(n^C_U-1)(n^C_U)^2}\Big|\mathbb{S}^{2*}_P\Big\} - \mathbb{V}^{2*}\Big(\frac{n_P}{n^C_U}\Big|\mathbb{S}^{2*}_P\Big)\Big]\{\Delta E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{k}^{C*}_U)\}^2 \\
&\quad + 2\mathbb{V}^*\Big(\frac{n_P}{n^T_U}, \frac{n_P}{n^C_U}\Big|\mathbb{S}^{2*}_P\Big)\Delta E^*(\boldsymbol{\beta}^{T*}|\boldsymbol{k}^{T*}_U)\Delta E^*(\boldsymbol{\beta}^{C*}|\boldsymbol{k}^{C*}_U).
\end{aligned}$$

The closed forms are obtained by using Lemmas 23 (2), 25, and 28 and Equation 205. If we do not take the limit, Equations 324 and 326 show bias of the general case, $\mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G)|\mathbb{S}^*_G\} - \mathbb{V}^{2*}(\hat{\tau}_G|\mathbb{S}^*_G)$.

Under Assumption 2*, 4*, and 5*, it follows that Proposition 4* (3) and Equation 290 hold. Applying Equation 290 to Proposition 4* (3), we obtain

$$\lim_{n^*\to\infty}[\mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_U)|\mathbb{S}^{2*}_P\} - \mathbb{V}^{2*}(\hat{\tau}_U|\mathbb{S}^{2*}_P)] = 0.$$

PROOF. Suppose that $\boldsymbol{K}^*_G \in \mathbb{U}^*$ satisfies Condition 1*, $\mathbb{N} \subseteq \mathbb{N}^{TC2}$, $\mathbb{S}^*_G = \mathbb{S}^*_{\text{max}}(\mathbf{n}_G \in \mathbb{N}) \subseteq \mathbb{S}^*_{\text{def}}\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G)\}$. Since $\mathbb{S}^*_{\text{def}}\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G)\} \subseteq \mathbb{S}^*_{\text{def}}\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)\}$, it follows $\mathbb{S}^*_G \subseteq \mathbb{S}^*_{\text{def}}\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)\}$. In addition, since $\mathbb{N}^{TC2} \subseteq \mathbb{N}^2$, it follows $\mathbb{N} \subseteq \mathbb{N}^2$. Thus,

$$\begin{aligned}
&\mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G)|\mathbb{S}^*_G\}\\
&= \mathbb{E}^*_S[\mathbb{E}_X\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G)\}|\mathbb{S}^*_G] \quad (\because \text{Equation 199})\\
&= \mathbb{E}^*_S[\mathbb{E}_X\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G) - \mathbb{V}(\hat{\tau}_G)\} + \mathbb{V}(\hat{\tau}_G)|\mathbb{S}^*_G]\\
&= \mathbb{E}^*_S\Big\{\sum_A \frac{1}{n^A_G - 1}V^2(\boldsymbol{\beta}^A|\boldsymbol{k}^A_G) - \frac{2n^{TC}_G}{n^T_G n^C_G}\frac{n^{TC}_G}{n^{TC}_G - 1}V(\boldsymbol{\beta}^T,\boldsymbol{\beta}^C|\boldsymbol{k}^{TC}_G)\\
&\quad + \sum_A \frac{1}{n^A_G}V^2(\boldsymbol{\omega}^A|\boldsymbol{k}^A_G) + \frac{2n^{TC}_G}{n^T_G n^C_G}V(\boldsymbol{\omega}^T,\boldsymbol{\omega}^C|\boldsymbol{k}^{TC}_G)\Big|\mathbb{S}^*_G\Big\}\\
&(\because \text{Equations 128 and 157}, \mathbb{S}^*_G \subseteq \mathbb{S}^*_{\text{def}}\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G)\}, \mathbb{N} \subseteq \mathbb{N}^{TC2} \subseteq \mathbb{N}^2)\\
&= \mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)|\mathbb{S}^*_G\} - \mathbb{E}^*\Big[\frac{2n^{TC}_G}{n^T_G n^C_G}\Big\{\frac{n^{TC}_G}{n^{TC}_G - 1}V(\boldsymbol{\beta}^T,\boldsymbol{\beta}^C|\boldsymbol{k}^{TC}_G) - V(\boldsymbol{\omega}^T,\boldsymbol{\omega}^C|\boldsymbol{k}^{TC}_G)\Big\}\Big|\mathbb{S}^*_G\Big]\\
&(\because \text{Lemmas 23 (1), Equations 199 and 317, } \mathbb{S}^*_G \subseteq \mathbb{S}^*_{\text{def}}\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)\}\ )\\
&= \mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)|\mathbb{S}^*_G\}\\
&\quad - 2\mathbb{E}^*\Big(\frac{n^{TC}_G}{n^T_G n^C_G}\frac{n^{TC}_G}{n^{TC}_G - 1}\Big[1 - \frac{1}{n^{TC}_G}\{1 - d(n^{TC*}_G, n^{TC}_G)\}\Big]\Big|\mathbb{S}^*_G\Big)V^*(\boldsymbol{\beta}^{T*},\boldsymbol{\beta}^{C*}|\boldsymbol{k}^{TC*}_G)\\
&\quad + 2\mathbb{E}^*\Big(\frac{n^{TC}_G}{n^T_G n^C_G}\Big|\mathbb{S}^*_G\Big)V^*(\boldsymbol{\omega}^{T*},\boldsymbol{\omega}^{C*}|\boldsymbol{k}^{TC*}_G)\\
&(\because \text{Lemmas 10 (2), 23 (1), 29 (2) and (4)}, \mathbb{N} \subseteq \mathbb{N}^{TC2} \subseteq \mathbb{N}^{TC1}).
\end{aligned} \tag{323}$$

Therefore, the bias of Adjusted Neyman variance estimator is equal to

$$
\begin{aligned}
&\mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G)|\mathbb{S}_G^*\} - \mathbb{V}^*(\hat{\tau}_G|\mathbb{S}_G^*)\\
&= \mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)|\mathbb{S}_G^*\} - \mathbb{V}^{2*}(\hat{\tau}_G|\mathbb{S}_G^*)\\
&\quad - 2\mathbb{E}^*\Big(\frac{n_G^{TC}}{n_G^T n_G^C}\frac{n_G^{TC}}{n_G^{TC}-1}\Big[1-\frac{1}{n_G^{TC}}\{1-d(n_G^{TC*}, n_G^{TC})\}\Big]\Big|\mathbb{S}_G^*\Big)V^*(\boldsymbol{\beta}^{T*}, \boldsymbol{\beta}^{C*}|\boldsymbol{k}_G^{TC*})\\
&\quad + 2\mathbb{E}^*\Big(\frac{n_G^{TC}}{n_G^T n_G^C}\Big|\mathbb{S}_G^*\Big)V^*(\boldsymbol{\omega}^{T*}, \boldsymbol{\omega}^{C*}|\boldsymbol{k}_G^{TC*}) \quad (\because \text{Equation 323})\\
&= 2\mathbb{E}^*\Big[\frac{n_G^{TC}}{n_G^T n_G^C}\{1-d(n_G^{TC*}, n_G^{TC})\}\Big|\mathbb{S}_G^*\Big]V^*(\boldsymbol{\beta}^{T*}, \boldsymbol{\beta}^{C*}|\boldsymbol{k}_G^{TC*}) - 2\mathbb{E}^*\Big(\frac{n_G^{TC}}{n_G^T n_G^C}\Big|\mathbb{S}_G^*\Big)V^*(\boldsymbol{\omega}^{T*}, \boldsymbol{\omega}^{C*}|\boldsymbol{k}_G^{TC*})\\
&\quad + \sum_A \mathbb{E}^*\Big\{\frac{1}{n_G^A-1}\prod_{g^{-A}=0}^{1}\frac{n_G^A(1,g^{-A})}{n_G^A}\Big|\mathbb{S}_G^*\Big\}\{\Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*})\}^2\\
&\quad + \sum_A\sum_{g^{-A}=0}^{1}\mathbb{E}^*\Big[\frac{n_G^A(1,g^{-A})}{(n_G^A-1)n_G^A}d\{n_G^{A*}(1,g^{-A}), n_G^A(1,g^{-A})\}\Big|\mathbb{S}_G^*\Big]V_{h=1}^{2*}\{\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*}(1,g^{-A})\}\\
&\quad - \sum_A \mathbb{V}^{2*}\Big(\frac{n_G^{TC}}{n_G^A}\Big|\mathbb{S}_G^*\Big)\{\Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*})\}^2 + 2\mathbb{V}^*\Big(\frac{n_G^{TC}}{n_G^T}, \frac{n_G^{TC}}{n_G^C}\Big|\mathbb{S}_G^*\Big)\prod_A \Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*})\\
&\quad - 2\mathbb{E}^*\Big(\frac{n_G^{TC}}{n_G^T n_G^C}\frac{n_G^{TC}}{n_G^{TC}-1}\Big[1-\frac{1}{n_G^{TC}}\{1-d(n_G^{TC*}, n_G^{TC})\}\Big]\Big|\mathbb{S}_G^*\Big)V^*(\boldsymbol{\beta}^{T*}, \boldsymbol{\beta}^{C*}|\boldsymbol{k}_G^{TC*})\\
&\quad + 2\mathbb{E}^*\Big(\frac{n_G^{TC}}{n_G^T n_G^C}\Big|\mathbb{S}_G^*\Big)V^*(\boldsymbol{\omega}^{T*}, \boldsymbol{\omega}^{C*}|\boldsymbol{k}_G^{TC*}) \quad (\because \text{Equation 318})\\
&= \sum_A\Big[\mathbb{E}^*\Big\{\frac{1}{n_G^A-1}\prod_{g^{-A}=0}^{1}\frac{n_G^A(1,g^{-A})}{n_G^A}\Big|\mathbb{S}_G^*\Big\} - \mathbb{V}^{2*}\Big(\frac{n_G^{TC}}{n_G^A}\Big|\mathbb{S}_G^*\Big)\Big]\{\Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*})\}^2\\
&\quad + 2\mathbb{V}^*\Big(\frac{n_G^{TC}}{n_G^T}, \frac{n_G^{TC}}{n_G^C}\Big|\mathbb{S}_G^*\Big)\prod_A \Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*})\\
&\quad + \sum_A\sum_{g^{-A}=0}^{1}\mathbb{E}^*\Big[\frac{n_G^A(1,g^{-A})}{(n_G^A-1)n_G^A}d\{n_G^{A*}(1,g^{-A}), n_G^A(1,g^{-A})\}\Big|\mathbb{S}_G^*\Big]V_{h=1}^{2*}\{\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*}(1,g^{-A})\}\\
&\quad - 2\mathbb{E}^*\Big\{\frac{n_G^{TC}}{n_G^T n_G^C}\frac{n_G^{TC}}{n_G^{TC}-1}d(n_G^{TC*}, n_G^{TC})\Big|\mathbb{S}_G^*\Big\}V^*(\boldsymbol{\beta}^{T*}, \boldsymbol{\beta}^{C*}|\boldsymbol{k}_G^{TC*})
\end{aligned} \tag{324}
$$

In limit, it follows

$$\begin{aligned}
&\lim_{n^*\to\infty}[\mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G)|\mathbb{S}_G^*\} - \mathbb{V}^*(\hat{\tau}_G|\mathbb{S}_G^*)] \\
&= \sum_A \left[\mathbb{E}^*\left\{\frac{1}{n_G^A - 1}\prod_{g^{-A}=0}^{1}\frac{n_G^A(1,g^{-A})}{n_G^A}\middle|\mathbb{S}_G^*\right\} - \mathbb{V}^{2*}\left(\frac{n_G^{TC}}{n_G^A}\middle|\mathbb{S}_G^*\right)\right]\{\Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*})\}^2 \\
&\quad + 2\mathbb{V}^*\left(\frac{n_G^{TC}}{n_G^T}, \frac{n_G^{TC}}{n_G^C}\middle|\mathbb{S}_G^*\right)\prod_A \Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*}) \\
&\quad + \sum_A \sum_{g^{-A}=0}^{1} \mathbb{E}^*\left\{\frac{n_G^A(1,g^{-A})}{(n_G^A-1)n_G^A}\cdot 0\middle|\mathbb{S}_G^*\right\}V_{h=1}^{2*}\{\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*}(1,g^{-A})\} \\
&\quad - 2\mathbb{E}^*\left\{\frac{n_G^{TC}}{n_G^T n_G^C}\frac{n_G^{TC}}{n_G^{TC}-1}\cdot 0\middle|\mathbb{S}_G^*\right\}\cdot V^*(\boldsymbol{\beta}^{T*},\boldsymbol{\beta}^{C*}|\boldsymbol{k}_G^{TC*}) \quad (\because \text{Equations 183, 298, 299, and 324}) \\
&= \sum_A \left[\mathbb{E}^*\left\{\frac{1}{n_G^A - 1}\prod_{g^{-A}=0}^{1}\frac{n_G^A(1,g^{-A})}{n_G^A}\middle|\mathbb{S}_G^*\right\} - \mathbb{V}^{2*}\left(\frac{n_G^{TC}}{n_G^A}\middle|\mathbb{S}_G^*\right)\right]\{\Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*})\}^2 \\
&\quad + 2\mathbb{V}^*\left(\frac{n_G^{TC}}{n_G^T}, \frac{n_G^{TC}}{n_G^C}\middle|\mathbb{S}_G^*\right)\prod_A \Delta E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^{A*})
\end{aligned} \tag{325}$$

When $\boldsymbol{k}_G^{T*} = \boldsymbol{k}_G^{C*} \equiv \boldsymbol{k}_G^*$, Equation 324 leads to

$$\begin{aligned}
&\mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G)|\mathbb{S}_G^*\} - \mathbb{V}^*(\hat{\tau}_G|\mathbb{S}_G^*) \\
&= \sum_A \left[\mathbb{E}^*\left\{\frac{1}{n_G - 1}\prod_{g^{-A}=0}^{1}\frac{n_G(1,g^{-A})}{n_G}\middle|\mathbb{S}_G^*\right\} - \mathbb{V}^{2*}\left(\frac{n_G}{n_G}\middle|\mathbb{S}_G^*\right)\right]\{E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^*)\}^2 \\
&\quad + 2\mathbb{V}^*\left(\frac{n_G}{n_G}, \frac{n_G}{n_G}\middle|\mathbb{S}_G^*\right)\prod_A E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^*) \\
&\quad + \sum_A \sum_{g^{-A}=0}^{1} \mathbb{E}^*\left[\frac{n_G(1,g^{-A})}{(n_G-1)n_G}d\{n_G^*(1,g^{-A}), n_G(1,g^{-A})\}\middle|\mathbb{S}_G^*\right]V_{h=1}^{2*}\{\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^*(1,g^{-A})\} \\
&\quad - 2\mathbb{E}^*\left\{\frac{n_G}{n_G n_G}\frac{n_G}{n_G-1}d(n_G^*, n_G)\middle|\mathbb{S}_G^*\right\}V^*(\boldsymbol{\beta}^{T*},\boldsymbol{\beta}^{C*}|\boldsymbol{k}_G^*) \quad (\because \text{Lemma 14 (3), Equation 305}) \\
&= \sum_A \left[\mathbb{E}^*\left\{\frac{1}{n_G-1}\frac{n_G}{n_G}\frac{0}{n_G}\middle|\mathbb{S}_G^*\right\} - 0\right]\{E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^*)\}^2 + 2\cdot 0\cdot\prod_A E^*(\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^*) \\
&\quad + \sum_A \left(\mathbb{E}^*\left[\frac{n_G}{(n_G-1)n_G}d(n_G^*, n_G)\middle|\mathbb{S}_G^*\right]V_{h=1}^{2*}\{\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^*\}\right. \\
&\quad \left. + \mathbb{E}^*\left[\frac{n_G}{(n_G-1)n_G}d(n_G^*, n_G)\middle|\mathbb{S}_G^*\right]V_{h=1}^{2*}\{\boldsymbol{\beta}^{A*}|\boldsymbol{k}_G^*\}\right) \\
&\quad - 2\mathbb{E}^*\left\{\frac{n_G}{n_G n_G}\frac{n_G}{n_G-1}d(n_G^*, n_G)\middle|\mathbb{S}_G^*\right\}V^*(\boldsymbol{\beta}^{T*},\boldsymbol{\beta}^{C*}|\boldsymbol{k}_G^*) \quad (\because \text{Equations 304 and 320}) \\
&= \mathbb{E}^*\left\{\frac{d(n_G^*, n_G)}{n_G-1}\middle|\mathbb{S}_G^*\right\}V^{2*}(\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*}|\boldsymbol{k}_G^*) \quad (\because \text{Lemma 15 (6)}) \\
&= \begin{cases} \dfrac{1}{n_G^*-1}V^{2*}(\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*}|\boldsymbol{k}_G^*) & \text{if } n_G^* \geq 2 \because \text{Equation 256} \\ 0 & \text{if } n_G^* < 2 \because \text{Equation 256} \end{cases}
\end{aligned} \tag{326}$$

In limit, it follows

$$\lim_{n^*\to\infty}[\mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G)|\mathbb{S}_G^*\} - \mathbb{V}^*(\hat{\tau}_G|\mathbb{S}_G^*)] = 0. \tag{327}$$

(1) Under Assumption 1^*, it holds that $\mathbb{S}^*_{\text{def}}\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_F)\} = \mathbb{S}^*_{\max}$ ($\because$ Equation 12 and 98, $N_F^{TC} = N_F = n \geq 2$). When $\boldsymbol{K}^*_G = \boldsymbol{K}^*_F \in \mathbb{U}^*$ and $\mathbb{N} = \mathbb{N}^n$ (note $\mathbb{N}^n \subseteq \mathbb{N}^{TC2}$), it follows that $\boldsymbol{K}_G = \boldsymbol{K}_F$ ($\because$ Equation 164),

$$\begin{aligned}\mathbb{S}^*_G &= \mathbb{S}^*_{\max}(\mathbf{n}_G \in \mathbb{N}) \quad (\because \text{by definition})\\ &= \mathbb{S}^*_{\max}(\mathbf{n}_F \in \mathbb{N}^{TC2}) \quad (\because \text{Lemma 13 (1), Equations 188}, \mathbb{N} = \mathbb{N}^{TC2})\\ &= \mathbb{S}^*_{\max}(n \geq 2) \quad (\because \text{Equations 188 and 315})\\ &= \mathbb{S}^*_{\max} \quad (\because n \geq 2)\end{aligned}$$

and, according to Lemma 13 (1), Equations 309 and 327 lead to the desired result.

(2) Under Assumption 3^*, it holds that $\mathbb{S}^*_{\text{def}}\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_P)\} = \mathbb{S}^{2*}_P$ ($\because$ Equation 102, Lemma 9 (1), $N_P^{TC} = N_P = n_P \geq 2$, Equations 204 and 211). When $\boldsymbol{K}^*_G = \boldsymbol{K}^*_P \in \mathbb{U}^*$ and $\mathbb{N} = \mathbb{N}^{TC2}$, it follows that $\boldsymbol{K}_G = \boldsymbol{K}_P$ ($\because$ Equation 164),

$$\begin{aligned}\mathbb{S}^*_G &= \mathbb{S}^*_{\max}(\mathbf{n}_G \in \mathbb{N}) \quad (\because \text{by definition})\\ &= \mathbb{S}^*_{\max}(\mathbf{n}_P \in \mathbb{N}^{TC2}) \quad (\because \text{Lemma 13 (2), Equations 189}, \mathbb{N} = \mathbb{N}^{TC2})\\ &= \mathbb{S}^*_{\max}(n_P \geq 2) \quad (\because \text{Equations 189 and 315})\\ &= \mathbb{S}^{2*}_P \quad (\because \text{Equation 213})\end{aligned}$$

and, according to Lemma 13 (2), Equation 327 is equivalent to the desired result.

(3) Under Assumption 2^*, it holds that $\mathbb{S}^*_{\text{def}}\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_U)\} = \mathbb{S}^{2*}_P$ (not $\mathbb{S}^{2*}_U$, $\because$ Equation 100, Lemma 9 (2), $N_U^{TC} = N_P = n_P \geq 2$, Equations 204 and 211). When $\boldsymbol{K}^*_G = \boldsymbol{K}^*_U \in \mathbb{U}^*$ and $\mathbb{N} = \mathbb{N}^{TC2}$, it follows that $\boldsymbol{K}_G = \boldsymbol{K}_U$ ($\because$ Equation 164),

$$\begin{aligned}\mathbb{S}^*_G &= \mathbb{S}^*_{\max}(\mathbf{n}_G \in \mathbb{N}) \quad (\because \text{by definition})\\ &= \mathbb{S}^*_{\max}(\mathbf{n}_U \in \mathbb{N}^{TC2}) \quad (\because \text{Lemma 13 (3), Equations 190}, \mathbb{N} = \mathbb{N}^{TC2})\\ &= \mathbb{S}^*_{\max}(n_P \geq 2) \quad (\because \text{Equations 190 and 315})\\ &= \mathbb{S}^{2*}_P \quad (\because \text{Equation 213})\end{aligned}$$

(note that, except for knife-edge situation, $\mathbb{S}^*_G \neq \mathbb{S}^{2*}_U$), and, according to Lemma 13 (3), Equation 325 is equivalent to the desired result.

□

Proposition 4^* (2) is surprising. Its alternative and more intuitive proof is as follows. Recall that Proposition 4^* (1) has been already established (Imai 2008, 4862-4863, Equations (10) and (11)). Suppose the following alternative setup for every value of $n' \in \{2, 3, \ldots, n\}$. We draw $n' = n_P$ pairs of finite sample $\mathbb{J}'_{\max} \equiv \mathbb{J}_P$ from the super-population $\mathbb{J}'^*_{\max} \equiv \mathbb{J}^*_P$ which is composed of $n'^* \equiv n^*_P$ pairs. We also draw $n^- = n - n_P$ pairs of finite sample $\mathbb{J}^-_{\max} \equiv \mathbb{J}_{\max} \setminus \mathbb{J}_P$ from the super-population $\mathbb{J}^{-*}_{\max} \equiv \mathbb{J}^*_{\max} \setminus \mathbb{J}^*_P$ which is composed of $n^{-*} \equiv n^* - n^*_P$ pairs, where $\mathbb{J}_{\max} \equiv \{1, 2, \ldots, n\}$ and $\mathbb{J}^*_{\max} \equiv \{1, 2, \ldots, n^*\}$. If we combine the two sets of pairs, we obtain finite sample as we have dealt with it so far ($\mathbb{J}'_{\max} \cup \mathbb{J}^-_{\max} = \mathbb{J}_{\max}$).

Let $j' \in \mathbb{J}'_{\max}$. Denote the h-th smallest j' by $j'(h)$. Let

$$\boldsymbol{Q}' \equiv \{Q_{1j'(1)}, Q_{2j'(1)}, Q_{1j'(2)}, Q_{2j'(2)}, \ldots, Q_{1j'(n')}, Q_{2j'(n')}\}.$$

Note

$$\begin{aligned}\sum_j \sum_i K_{P,ij} Q_{ij} &= \sum_{j' \in \mathbb{J}'_{\max}} \sum_i 1 \cdot Q_{ij'} + \sum_{j^- \in \mathbb{J}^-_{\max}} \sum_i 0 \cdot Q_{ij^-}\\ &= \sum_{j' \in \mathbb{J}'_{\max}} \sum_i K_{F,ij'} Q_{ij'}.\end{aligned} \tag{328}$$

Therefore,

$$\begin{aligned} N'_F &= \sum_{j' \in \mathbb{J}'_{\max}} \sum_i K_{F,ij'} X^A_{ij'} \quad (\because \text{Equations 9 and 12}) \\ &= \sum_j \sum_i K_{P,ij} X^A_{ij'} \quad (\because \text{Equation 328}) \\ &= N_P \quad (\because \text{Equations 9 and 16}) \end{aligned} \tag{329}$$

and

$$\begin{aligned} E'(\boldsymbol{Q}'|\boldsymbol{K}'_F\boldsymbol{Z}') &= \frac{\sum_{j' \in \mathbb{J}'_{\max}} \sum_i K'_{F,ij'} Z'_{ij'} Q'_{ij'}}{\sum_{j' \in \mathbb{J}'_{\max}} \sum_i K'_{F,ij'} Z'_{ij'}} \quad (\because \text{Equation 4}) \\ &= \frac{\sum_j \sum_i K_{P,ij} Z_{ij} Q_{ij}}{\sum_j \sum_i K_{P,ij} Z_{ij}} \quad (\because \text{Equation 328}) \\ &= E(\boldsymbol{Q}|\boldsymbol{K}_P\boldsymbol{Z}) \quad (\because \text{Equation 4}) \end{aligned} \tag{330}$$

and

$$\begin{aligned} \hat{\tau}'_F &= E'(\boldsymbol{Y}'|\boldsymbol{K}'_F\boldsymbol{X}^{T\prime}) - E'(\boldsymbol{Y}'|\boldsymbol{K}'_F\boldsymbol{X}^{C\prime}) \quad (\because \text{Equations 10 and 12} \\ &= E(\boldsymbol{Y}|\boldsymbol{K}_P\boldsymbol{X}^T) - E(\boldsymbol{Y}|\boldsymbol{K}_P\boldsymbol{X}^C) \quad (\because \text{Equation 330}) \\ &= \hat{\tau}_P \quad (\because \text{Equations 10 and 16} \end{aligned}$$

and

$$\begin{aligned} &V'(\boldsymbol{Q}^{(1)\prime}, \boldsymbol{Q}^{(2)\prime}|\boldsymbol{K}'_F\boldsymbol{Z}') \\ &= E'[\{\boldsymbol{Q}^{(1)\prime} - \boldsymbol{E}'(\boldsymbol{Q}^{(1)\prime}|\boldsymbol{K}'_F\boldsymbol{Z}')\}\{\boldsymbol{Q}^{(2)\prime} - \boldsymbol{E}'(\boldsymbol{Q}^{(2)\prime}|\boldsymbol{K}'_F\boldsymbol{Z}')\}|\boldsymbol{K}'_F\boldsymbol{Z}'] \quad (\because \text{Equations 88 and 89}) \\ &= E'[\{\boldsymbol{Q}^{(1)\prime} - \boldsymbol{E}(\boldsymbol{Q}^{(1)}|\boldsymbol{K}_P\boldsymbol{Z})\}\{\boldsymbol{Q}^{(2)\prime} - \boldsymbol{E}(\boldsymbol{Q}^{(2)}|\boldsymbol{K}_P\boldsymbol{Z})\}|\boldsymbol{K}'_F\boldsymbol{Z}'] \quad (\because \text{Equation 330}) \\ &= E'(\boldsymbol{Q}'|\boldsymbol{K}'_F\boldsymbol{Z}') \\ &= E(\boldsymbol{Q}|\boldsymbol{K}_P\boldsymbol{Z}) \quad (\because \text{Equation 330}) \\ &= E'[\{\boldsymbol{Q}^{(1)} - \boldsymbol{E}(\boldsymbol{Q}^{(1)}|\boldsymbol{K}_P\boldsymbol{Z})\}\{\boldsymbol{Q}^{(2)} - \boldsymbol{E}(\boldsymbol{Q}^{(2)}|\boldsymbol{K}_P\boldsymbol{Z})\}|\boldsymbol{K}'_F\boldsymbol{Z}'] \\ &= V(\boldsymbol{Q}^{(1)}, \boldsymbol{Q}^{(2)}|\boldsymbol{K}_P\boldsymbol{Z}) \quad (\because \text{Equations 88 and 89}) \end{aligned} \tag{331}$$

where $\boldsymbol{Q}' \equiv \{\boldsymbol{Q}^{(1)\prime} - \boldsymbol{E}(\boldsymbol{Q}^{(1)}|\boldsymbol{K}_P\boldsymbol{Z})\}\{\boldsymbol{Q}^{(2)\prime} - \boldsymbol{E}(\boldsymbol{Q}^{(2)}|\boldsymbol{K}_P\boldsymbol{Z})\}$, and

$$\begin{aligned} &\hat{\mathbb{V}}^{\text{Adj-Neyman}\prime}(\hat{\tau}'_F) \\ &= \sum_A \frac{1}{N'_F - 1} V^{2\prime}(\boldsymbol{Y}'|\boldsymbol{K}'_F\boldsymbol{X}^{A\prime}) - \frac{2}{N'_F - 1} V'(\boldsymbol{Y}', \boldsymbol{Y}'_{-i}|\boldsymbol{K}'_F\boldsymbol{X}^{A\prime}) \quad (\because \text{Equation 135}) \\ &= \sum_A \frac{1}{N_P - 1} V^2(\boldsymbol{Y}|\boldsymbol{K}_P\boldsymbol{X}^A) - \frac{2}{N_P - 1} V(\boldsymbol{Y}, \boldsymbol{Y}_{-i}|\boldsymbol{K}_P\boldsymbol{X}^A) \quad (\because \text{Equations 329 and 331}) \\ &= \hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_P) \quad (\because \text{Equation 135}) \end{aligned} \tag{332}$$

Let $j'^* \in \mathbb{J}'^*_{\max}$. Denote the h-th smallest j'^* by $j'^*(h)$. Let

$$\boldsymbol{s}'^* \equiv (s^*_{\cdot j'^*(1)}, s^*_{\cdot j'^*(2)}, \ldots s^*_{\cdot j'^*(n'^*)})$$
$$\mathbb{S}'^*_{\max} \equiv \left\{ \boldsymbol{s}'^* \middle| s^*_{\cdot j'^*} = \{0, 1\}, \sum_{j'^* \in \mathbb{J}'^*_{\max}} s^*_{\cdot j'^*} = n' \right\}.$$

Let $j^{-*} \in \mathbb{J}^{-*}_{\max}$. Denote the h-th smallest j^{-*} by $j^{-*}(h)$. Let

$$\boldsymbol{s}^{-*} \equiv (s^*_{\cdot j^{-*}(1)}, s^*_{\cdot j^{-*}(2)}, \ldots s^*_{\cdot j^{-*}(n^{-*})})$$
$$\mathbb{S}^{-*}_{\max} \equiv \left\{ \boldsymbol{s}^{-*} \middle| s^*_{\cdot j^{-*}} = \{0, 1\}, \sum_{j^{-*} \in \mathbb{J}^{-*}_{\max}} s^*_{\cdot j^{-*}} = n^- \right\}.$$

For $\mathbb{S}^* \subseteq \mathbb{S}^*_{\max}$, let $\mathbb{S}^*_N \equiv \mathbb{S}^*(n_P = n')$. Define

$$\mathbb{S}'^* \equiv \{\boldsymbol{s}'^* | \boldsymbol{s}'^* \in \mathbb{S}'^*_{\max}, \boldsymbol{s}^* \in \mathbb{S}^*_N\}.$$
$$\mathbb{S}^{-*} \equiv \{\boldsymbol{s}^{-*} | \boldsymbol{s}^{-*} \in \mathbb{S}^{-*}_{\max}, \boldsymbol{s}^* \in \mathbb{S}^*_N\}.$$

For a while, suppose $n^- > 0$. Note

$$|\mathbb{S}^*_N| = |\mathbb{S}'^*| \cdot |\mathbb{S}^{-*}|. \tag{333}$$

It follows

$$\begin{aligned}\Pr(\boldsymbol{S}'^* = \boldsymbol{s}'^* | \boldsymbol{S}'^* \in \mathbb{S}'^*) &= \sum_{\boldsymbol{s}^{-*} \in \mathbb{S}^{-*}} \Pr(\boldsymbol{S}'^* = \boldsymbol{s}'^*, \boldsymbol{S}^{-*} = \boldsymbol{s}^{-*} | \boldsymbol{S}^* \in \mathbb{S}^*_N) \\ &= \frac{|\mathbb{S}^{-*}|}{|\mathbb{S}^*_N|} \quad (\because \text{Lemma 21 (1)}) \\ &= \frac{1}{|\mathbb{S}'^*|} \quad (\because \text{Equation 333})\end{aligned} \tag{334}$$

Similarly,

$$\Pr(\boldsymbol{S}^{-*} = \boldsymbol{s}^{-*} | \boldsymbol{S}^{-*} \in \mathbb{S}^{-*}) = \frac{1}{|\mathbb{S}^{-*}|}. \tag{335}$$

It follows

$$\begin{aligned}&\Pr(\boldsymbol{S}^* = \boldsymbol{s}^* | \boldsymbol{S}^* \in \mathbb{S}^*_N) \\ &= \frac{1}{|\mathbb{S}^*_N|} \quad (\because \text{Lemma 21 (1)}) \\ &= \frac{1}{|\mathbb{S}'^*|} \cdot \frac{1}{|\mathbb{S}^{-*}|} \quad (\because \text{Equation 333}) \\ &= \Pr(\boldsymbol{S}'^* = \boldsymbol{s}'^* | \boldsymbol{S}'^* \in \mathbb{S}'^*) \Pr(\boldsymbol{S}^{-*} = \boldsymbol{s}^{-*} | \boldsymbol{S}^{-*} \in \mathbb{S}^{-*}) \quad (\because \text{Equations 334 and 335})\end{aligned} \tag{336}$$

Suppose that $f'^*_S(\boldsymbol{s}'^*) = f^*_S(\boldsymbol{s}^*)$. It follows that

$$\begin{aligned}\mathbb{E}^*\{f^*_S(\boldsymbol{s}^*)|\mathbb{S}^*_N\} &= \sum_{\boldsymbol{s}^* \in \mathbb{S}^*_N} \Pr(\boldsymbol{S}^* = \boldsymbol{s}^* | \boldsymbol{S}^* \in \mathbb{S}^*_N) f^*_S(\boldsymbol{s}^*) \\ &= \sum_{\boldsymbol{s}^{-*} \in \mathbb{S}^{-*}} \sum_{\boldsymbol{s}'^* \in \mathbb{S}'^*} \Pr(\boldsymbol{S}^{-*} = \boldsymbol{s}^{-*} | \boldsymbol{S}^{-*} \in \mathbb{S}^{-*}) \Pr(\boldsymbol{S}'^* = \boldsymbol{s}'^* | \boldsymbol{S}'^* \in \mathbb{S}'^*) f'^*_S(\boldsymbol{s}'^*) \\ &\quad (\because \text{Equation 336}) \\ &= \Big\{ \sum_{\boldsymbol{s}^{-*} \in \mathbb{S}^{-*}} \Pr(\boldsymbol{S}^{-*} = \boldsymbol{s}^{-*} | \boldsymbol{S}^{-*} \in \mathbb{S}^{-*}) \Big\} \Big\{ \sum_{\boldsymbol{s}'^* \in \mathbb{S}'^*} \Pr(\boldsymbol{S}'^* = \boldsymbol{s}'^* | \boldsymbol{S}'^* \in \mathbb{S}'^*) f'^*_S(\boldsymbol{s}'^*) \Big\} \\ &= \mathbb{E}^*\{f'^*_S(\boldsymbol{s}'^*)|\mathbb{S}'^*\} \quad (\because \text{Axiom of probability, Equation 199})\end{aligned} \tag{337}$$

and

$$\mathbb{V}^{2*}\{f^*_S(\boldsymbol{s}^*)|\mathbb{S}^*_N\} = \mathbb{V}^{2*}\{f'^*_S(\boldsymbol{s}'^*)|\mathbb{S}'^*\} \quad (\because \text{Equation 337}) \tag{338}$$

Even when $n^- = 0$, both ends of Equations 337 and 338 are equal to each other, respectively. When $\mathbb{S}^* = \mathbb{S}^*_{\max}$, it follows that

$$\begin{aligned}\mathbb{S}'^* &= \mathbb{S}'^*_{\max} \\ \mathbb{S}^{-*} &= \mathbb{S}^{-*}_{\max}.\end{aligned} \tag{339}$$

Accordingly,

$$\begin{aligned}
&\mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_P)|\mathbb{S}_P^{2*}\} - \mathbb{V}^{2*}(\hat{\tau}_P|\mathbb{S}_P^{2*}) \\
&= \sum_{n'=2}^{n} \Pr(n_P = n'|\mathbb{S}_P^{2*})[\mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_P)|\mathbb{S}_P^{2*}, n_P = n'\} - \mathbb{V}^{2*}(\hat{\tau}_P|\mathbb{S}_P^{2*}, n_P = n')] \\
&= \sum_{n'=2}^{n} \Pr(n_P = n'|\mathbb{S}_P^{2*})[\mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_P)|\mathbb{S}_{\max}^*(n_P = n')\} - \mathbb{V}^{2*}(\hat{\tau}_P|\mathbb{S}_{\max}^*(n_P = n'))] \\
&= \sum_{n'=2}^{n} \Pr(n_P = n'|\mathbb{S}_P^{2*})[\mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Adj-Neyman}\prime}(\hat{\tau}'_F)|\mathbb{S}'^*_{\max}\} - \mathbb{V}^{2*}(\hat{\tau}'_F|\mathbb{S}'^*_{\max})] \\
&(\because \text{Equations 332, 337, 338, and 339})
\end{aligned}$$

Since $n_P \geq 2$, it follows that $n_P^* \geq n_P \geq 2$ and, thus, $\lim_{n^* \to \infty} n'^* = \lim_{n^* \to \infty} n_P^* = \infty$. Therefore,

$$\begin{aligned}
&\lim_{n^* \to \infty}[\mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_P)|\mathbb{S}_P^{2*}\} - \mathbb{V}^{2*}(\hat{\tau}_P|\mathbb{S}_P^{2*})] \\
&= \sum_{n'=2}^{n} \lim_{n^* \to \infty} \Pr(n_P = n'|\mathbb{S}_P^{2*}) \lim_{n'^* \to \infty}[\mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Adj-Neyman}\prime}(\hat{\tau}'_F)|\mathbb{S}'^*_{\max}\} - \mathbb{V}^{2*}(\hat{\tau}'_F|\mathbb{S}'^*_{\max})] \\
&= \sum_{n'=2}^{n} \lim_{n^* \to \infty} \Pr(n_P = n'|\mathbb{S}_P^{2*}) \cdot 0 \quad (\because \text{Proposition } 4^* \text{ (1)}) \\
&= 0.
\end{aligned}$$

Imai (2008, 4862-4863, Equations (10) and (11)) has already shown (1), but neither (2) nor (3), of Proposition 4*. In terms of my notation, Imai (2008, 4863, Equation (11)) formalizes $\lim_{n^* \to \infty} \mathbb{E}^*\{\hat{\mathbb{V}}^{\text{pair}}(\hat{\tau}_F)|\mathbb{S}_{\max}^*\}$ as

$$\frac{1}{n} V^{2*}(\boldsymbol{y}^{T*} - \boldsymbol{y}_{-i}^{C*}|\mathbf{1}^*).$$

Note that

$$\begin{aligned}
\boldsymbol{y}^{T*} - \boldsymbol{y}_{-i}^{C*} &= (\boldsymbol{\mu}^{T*} + \boldsymbol{\beta}^{T*} + \boldsymbol{\omega}^{T*}) - (\boldsymbol{\mu}_{-i}^{C*} + \boldsymbol{\beta}_{-i}^{C*} + \boldsymbol{\omega}_{-i}^{C*}) \quad (\because \text{Lemma 4 (1)}) \\
&= \boldsymbol{\tau}^* + (\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*}) + (\boldsymbol{\omega}^{T*} + \boldsymbol{\omega}^{C*}) \quad (\because \text{Equations 6, 17, 18, and 19}) \\
\therefore E^*(\boldsymbol{y}^{T*} - \boldsymbol{y}_{-i}^{C*}|\mathbf{1}^*) &= E^*(\boldsymbol{\tau}^*|\mathbf{1}^*) + E^*(\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*}|\mathbf{1}^*) + E^*(\boldsymbol{\omega}^{T*} + \boldsymbol{\omega}^{C*}|\mathbf{1}^*) \quad (\because \text{Lemmas 3 (1)}) \\
&= \bar{\tau}^* \quad (\because \text{Lemmas 3 (3) and 5})
\end{aligned} \tag{340}$$

Denote $\boldsymbol{\delta}^* \equiv (\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*})(\boldsymbol{\omega}^{T*} + \boldsymbol{\omega}^{C*})$. It follows that

$$\begin{aligned}
\boldsymbol{\delta}_{-i}^* &= (\boldsymbol{\beta}_{-i}^{T*} - \boldsymbol{\beta}_{-i}^{C*})(\boldsymbol{\omega}_{-i}^{T*} + \boldsymbol{\omega}_{-i}^{C*}) \\
&= (\boldsymbol{\beta}^{T*} - \boldsymbol{\beta}^{C*})(-\boldsymbol{\omega}^{T*} - \boldsymbol{\omega}^{C*}) \quad (\because \boldsymbol{\beta}^{A*} \in \mathbb{B}^*, \boldsymbol{\omega}^{A*} \in \mathbb{W}^*) \\
&= -\boldsymbol{\delta}^* \\
\therefore \boldsymbol{\delta}^* &\in \mathbb{W}^*
\end{aligned} \tag{341}$$

Therefore,

$$
\begin{aligned}
&\frac{1}{n}V^{2*}(\boldsymbol{y}^{T*}-\boldsymbol{y}_{-i}^{C*}|\mathbf{1}^*)\\
&=\frac{1}{n}E^*[\{\boldsymbol{y}^{T*}-\boldsymbol{y}_{-i}^{C*}-\boldsymbol{E}^*(\boldsymbol{y}^{T*}-\boldsymbol{y}_{-i}^{C*}|\mathbf{1}^*)\}^2|\mathbf{1}^*] \quad (\because \text{Equations 88, 89, and 90})\\
&=\frac{1}{n}E^*[\{(\boldsymbol{\beta}^{T*}-\boldsymbol{\beta}^{C*})+(\boldsymbol{\omega}^{T*}+\boldsymbol{\omega}^{C*})\}^2|\mathbf{1}^*] \quad (\because \text{Equation 340})\\
&=\frac{1}{n}[E^*\{(\boldsymbol{\beta}^{T*}-\boldsymbol{\beta}^{C*})^2|\mathbf{1}^*\}+E^*\{(\boldsymbol{\omega}^{T*}+\boldsymbol{\omega}^{C*})^2|\mathbf{1}^*\}+2E^*(\boldsymbol{\delta}^*|\mathbf{1}^*)]\\
&\quad(\because \text{Lemma 3 (1) and (2)})\\
&=\frac{1}{n}\Big(E^*[\{\boldsymbol{\beta}^{T*}-\boldsymbol{\beta}^{C*}-E^*(\boldsymbol{\beta}^{T*}-\boldsymbol{\beta}^{C*}|\mathbf{1}^*)\}^2|\mathbf{1}^*]+V^{2*}(\boldsymbol{\omega}^{T*}+\boldsymbol{\omega}^{C*}|\mathbf{1}^*)\Big)\\
&\quad(\because \text{Equation 341, Lemmas 5 and 17 (1)})\\
&=\frac{1}{n_F}\{V^{2*}(\boldsymbol{\beta}^{T*}-\boldsymbol{\beta}^{C*}|\boldsymbol{k}_F^*)+V^{2*}(\boldsymbol{\omega}^{T*}+\boldsymbol{\omega}^{C*}|\boldsymbol{k}_F^*)\} \quad (\because \text{Equations 35, 88, 89, and 90})\\
&=\lim_{n^*\to\infty}\mathbb{V}^{2*}(\hat{\tau}_F|\mathbb{S}^*_{\max}) \quad (\because \text{Proposition } 2^* \text{ (1)})
\end{aligned}
\tag{342}
$$

Thus, we confirm that Proposition 4^* (1):

$$
\begin{aligned}
&\lim_{n^*\to\infty}[\mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_F)|\mathbb{S}^*_{\max}\}-\mathbb{V}^{2*}(\hat{\tau}_F|\mathbb{S}^*_{\max})]\\
&=\frac{1}{n}V^{2*}(\boldsymbol{y}^{T*}-\boldsymbol{y}_{-i}^{C*}|\mathbf{1}^*)-\lim_{n^*\to\infty}\mathbb{V}^{2*}(\hat{\tau}_F|\mathbb{S}^*_{\max}) \quad (\because \text{Imai (2008, 4863, Equation (11)))}\\
&=0 \quad (\because \text{Equation 342})
\end{aligned}
$$

Here is another look. In fact, $\hat{\omega}_{ij}^T$ estimates not only ω_{ij}^T but also β_{ij}^T (Equation (**??**)). $\mathbb{V}^2(\hat{\tau}_P)$ depends on $\boldsymbol{\omega}^T$ but not $\boldsymbol{\beta}^T$, though $\mathbb{V}^{2*}(\hat{\tau}_P|\mathbb{S}_P^{2*})$ takes into account both $\boldsymbol{\omega}^T$ and $\boldsymbol{\beta}^T$. (The case for $\hat{\omega}_{ij}^C$ is similar.) This difference corresponds to the contrast where $\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_P)$ is biased for $\mathbb{V}^2(\hat{\tau}_P)$ (where the bias size is related to $\boldsymbol{\beta}^T$, Proposition 4 (2)) but unbiased for $\mathbb{V}^{2*}(\hat{\tau}_P|\mathbb{S}_P^{2*})$ (Proposition 4^* (2)). The above argument holds for $\hat{\tau}_F$ as well.

In Proposition 4^* (3), note that $\mathbb{S}^*_{\text{def}}\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_U)\}$ is not $\mathbb{S}_U^{2*}$ but $\mathbb{S}_P^{2*}$ because neither n_U^T nor n_U^C but $N^{TC}(\boldsymbol{K}_U)=N_P=n_P$ should be not fewer than two. The bias direction of $\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_U)$ is unknown. Under Assumption 2^*, 4^* and 5^*, and $\mathbb{S}_P^{2*}\neq\emptyset$, thanks to Equation 310, it follows that $\lim_{n^*\to\infty}[\mathbb{E}^*\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_U)|\mathbb{S}_P^{2*}\}-\mathbb{V}^{2*}(\hat{\tau}_U|\mathbb{S}_P^{2*})]=0$.

4. APPLICATION

4.1. Setting

In order to give readers the sense of how (well) $\hat{\tau}_P$, $\hat{\tau}_U$, $\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)$, and $\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G)$ work, this section reanalyzes the data of Angrist & Lavy (2009, Table A1, Copyright American Economic Association; reproduced with permission of the American Economic Review). The original authors matched $n=20$ pairs of Israeli schools based on the percentages of students who earned the high school matriculation certificate, or "Bagrut," in 1999 (hereafter, Bagrut rate). In December of 2000, they randomly assigned treatment of the Achievement Award program ($\boldsymbol{X}$) to one school of every pair. In the program, every student who received a Bagrut in 2001 was eligible for a payment. Orientation for principals and students took place in January 2001. Tests for Bagrut were taken in June 2001. Student is the unit of analysis in the original article, but in this manuscript the unit of analysis is school, the original unit of randomization. Analysts do not have to consider an interference issue among students in the same school which is always a concern for a cluster randomized design (Imai et al. 2009, pp. 32–34, 40).

Among several outcomes the original authors examine, this manuscript focuses on four annual Bagrut rates from 1999 to 2002. The outcome of main interest is the Bagrut rate in the

Table 1. Estimation of the Average Treatment Effects on Bagrut Rate by PDE ($\hat{\tau}_P$) and UDE ($\hat{\tau}_U$).

Outcome ($\boldsymbol{Y}^{(t)}$)	Year (t)		$\hat{\tau}_P$	$\hat{\tau}_U$
Matched-on	1999	$\hat{\tau}_G$	-0.28	-0.40
		$\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)\}^{1/2}$	(2.33)	(2.27)
		$\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G)\}^{1/2}$	[0.17]	[0.17]
Pretreatment	2000	$\hat{\tau}_G$	-1.58	-2.07
		$\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)\}^{1/2}$	(6.37)	(6.26)
		$\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G)\}^{1/2}$	[6.75]	[6.63]
Posttreatment	2001	$\hat{\tau}_G$	7.78	7.02
		$\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)\}^{1/2}$	(6.29)	(6.17)
		$\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G)\}^{1/2}$	[6.69]	[6.55]
Placebo	2002	$\hat{\tau}_G$	-5.50	-4.65
		$\{\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)\}^{1/2}$	(6.30)	(6.12)
		$\{\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G)\}^{1/2}$	[5.30]	[5.20]
N_U^T(2002 in parenthesis)				20 (20)
N_P, N_U^C(2002 in parenthesis)			19 (18)	19 (18)

posttreatment year, 2001 ($\boldsymbol{Y}^{(2001)}$), and the original authors theoretically expect that the ATE is positive; $\bar{\tau}^{(2001)} > 0$. By design, the Bagrut rates in 1999 ($\boldsymbol{Y}^{(1999)}$) and 2000 ($\boldsymbol{Y}^{(2000)}$) should be pretreatments. Therefore, *a priori*, scholars should be confident that $\boldsymbol{\tau}^{(1999)} = \boldsymbol{\tau}^{(2000)} = \mathbf{0}$ (sharp null) and, therefore, $\bar{\tau}^{(1999)} = \bar{\tau}^{(2000)} = 0$ (null ATE). Angrist & Lavy (2009) argue that the Bagrut rate in 2002 ($\boldsymbol{Y}^{(2002)}$) "can be seen as providing a sort of placebo control" because "[a]lthough seniors in the 2002 cohort were offered small payment ... as eleventh graders in 2001, no further incentives were offered to this cohort since the program was canceled before they began their senior year" (1395). Thus, the original authors believe $\boldsymbol{\tau}^{(2002)} = \mathbf{0}$ and $\bar{\tau}^{(2002)} = 0$, though, admittedly, it is arguable.

Two schools have missing values. Without loss of generality, for every pair j, I label the treated unit $i = 1$ and the controlled unit $i = 2$. Since the control school in pair 6 had closed before treatment assignment was announced, the original authors do not report its Bagrut rates for all four years ($R_{2,6}^{(t)} = 0$ for $t = 1999, \ldots, 2002$). In addition, the control school in pair 17 has a missing value only in 2002 ($R_{2,17}^{(2002)} = 0$). Since the Bagrut rates of this school were bad in 2000 ($Y_{2,17}^{(2000)} = 0.071$, the sixth worst of 39 schools, where the average Bagrut rate was 0.238) and 2001 ($Y_{2,17}^{(2001)} = 0.000$), it is reasonable to infer that $Y_{2,17}^{(2002)}$ would be low if it were observed. If the control school in pair 6 also closed because of its poor performance, attrition is likely to be non-ignorable. To sum, for $t = 1999, 2000, 2001$, $N_P^{(t)} = N_U^{C(t)} = 19$, while $N_P^{(2002)} = N_U^{C(2002)} = 18$. For all years, $N_U^{T(t)} = 20$. Since the original authors utilize the treated schools of pairs 6 and 17 (whose controlled schools have missing values) in student-level analysis, they employ $\hat{\tau}_U$ in effect.

4.2. Results

Table 1 displays the results. (The replication materials for this section can be found at `https://doi.org/10.7910/DVN/O9WE06`.) Four panels correspond to four annual Bagrut rates ($\boldsymbol{Y}^{(t)}$) from $t = 1999$ (top) to $t = 2002$ (bottom). In every panel, the point estimates of the ATE ($\hat{\tau}_G$, first row) and their standard errors based on $\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G)$(second row) and $\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G)$(third row) are presented. At the bottom of the table, I report N_P, N_U^T, and N_U^C. The first and second columns indicate $\hat{\tau}_P$ and $\hat{\tau}_U$, respectively.

At a first look, it seems that $\hat{\tau}_P$ and $\hat{\tau}_U$ make little difference; their point estimates and standard errors are close to each other; and both ATE estimators lead to the same conclusion that one cannot reject the null hypothesis $\bar{\tau}^{(t)} = 0$ from $t = 2000$ to $t = 2002$ at the 5% significance level.

There is, however, one important exception for $t = 1999$. Figure 1 illustrates the point estimates of the ATE ($\hat{\tau}_G^{(1999)}$, indicated by points) and their 95% confidence intervals (indicated by lines) based on $\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G^{(1999)})$ and a t-distribution with 19 degrees of freedom. I also demonstrate the case of a t-distribution with 20 degrees of freedom (Figure 2) and the case of a normal distribution (Figure 3). The upper and lower parts respond to $\hat{\tau}_P^{(1999)}$ and $\hat{\tau}_U^{(1999)}$, respectively. The dotted vertical line indicates the null hypothesis ($\bar{\tau}^{(1999)} = 0$). Accordingly, $\hat{\tau}_U^{(1999)}$ rejects the null hypothesis at the 5% significance level, though $\hat{\tau}_P^{(1999)}$ does not. Recall that, since $\boldsymbol{Y}^{(1999)}$ is a pretreatment, scholars should be sure of $\bar{\tau}^{(1999)} = 0$. Thus, in this particular case, inference based on $\hat{\tau}_U^{(1999)}$ is misleading, while that of $\hat{\tau}_P^{(1999)}$ is plausible. This might be because $\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_U^{(1999)}) < \mathbb{V}^2(\hat{\tau}_U^{(1999)})$, that is, $\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_U^{(1999)})$ is too optimistic, as Propositions 4 (3) and 4* (3) allow. I do not argue that this finding confirms that $\hat{\tau}_P$ always leads to more valid conclusion than $\hat{\tau}_U$; I simply call attention to the fact that both ATE estimators can make difference.

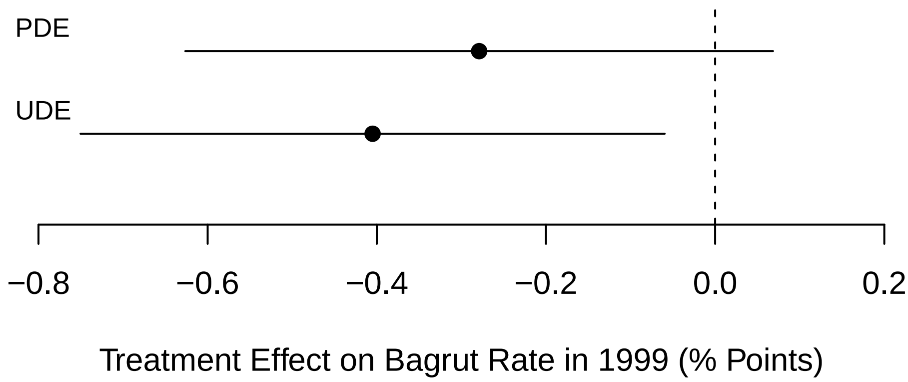

Fig. 1. Inference of the Average Treatment Effect on Bagrut Rate in 1999. Points indicate the point estimates of the ATE ($\hat{\tau}_G^{(1999)}$). Lines indicate their 95% confidence intervals based on $\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G^{(1999)})$. The upper and lower parts respond to $\hat{\tau}_P^{(1999)}$ and $\hat{\tau}_U^{(1999)}$, respectively. The dotted vertical line indicates the null hypothesis.

It is also worth mentioning one more important point for $t = 1999$: $\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G^{(1999)})$ is much smaller than $\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G^{(1999)})$. Since $\boldsymbol{Y}^{(1999)}$ is the matched-on variable to match pairs, it follows that $\boldsymbol{Y}^{(1999)} \approx \boldsymbol{Y}_{-i}^{(1999)}$, which is confirmed by $\hat{\tau}_P^{(1999)} \approx 0$. Given $\boldsymbol{\tau}^{(1999)} = \mathbf{0}$ (which implies that the treatment effect is perfectly homogeneous), I would infer that $\boldsymbol{y}^{T(1999)} = \boldsymbol{y}^{C(1999)}, \boldsymbol{\beta}^{T(1999)} = \boldsymbol{\beta}^{C(1999)}, \boldsymbol{y}^{T(1999)} \approx \boldsymbol{y}_{-i}^{T(1999)}, \boldsymbol{y}^{C(1999)} \approx \boldsymbol{y}_{-i}^{C(1999)}, \boldsymbol{\omega}^{T(1999)} = \boldsymbol{\omega}^{C(1999)} \approx \mathbf{0}$. Therefore, my conjecture is that $\mathbb{V}(\hat{\tau}_G^{(1999)})$ is negligible (Proposition 2), $\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G^{(1999)})$ is larger than $\mathbb{V}(\hat{\tau}_G^{(1999)})$ (Proposition 3), and $\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G^{(1999)})$ is as small as $\mathbb{V}(\hat{\tau}_G^{(1999)})$ (Proposition 4).

Some minor remarks are in order. First, $\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_P^{(t)}) > \hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_U^{(t)})$ and $\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_P^{(t)}) > \hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_U^{(t)})$ for all four years (t). Second, $\hat{\tau}_P^{(t)}$ is closer to $\bar{\tau}^{(t)} = 0$ than $\hat{\tau}_U^{(t)}$ for $t = 1999$ and 2000, though the opposite holds for $t = 2002$. Note that, however, the case of $\bar{\tau}^{(2002)} = 0$ is weaker than those of $\bar{\tau}^{(1999)} = 0$ or $\bar{\tau}^{(2000)} = 0$. Recall that the original authors expect $\bar{\tau}^{(2001)} > 0$. Third, $\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_G^{(t)})$ is larger than $\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_G^{(t)})$ for $t = 2000$ and 2001, though the opposite holds for $t = 2002$. Recall that $\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_P)$ is conservative in the finite sample

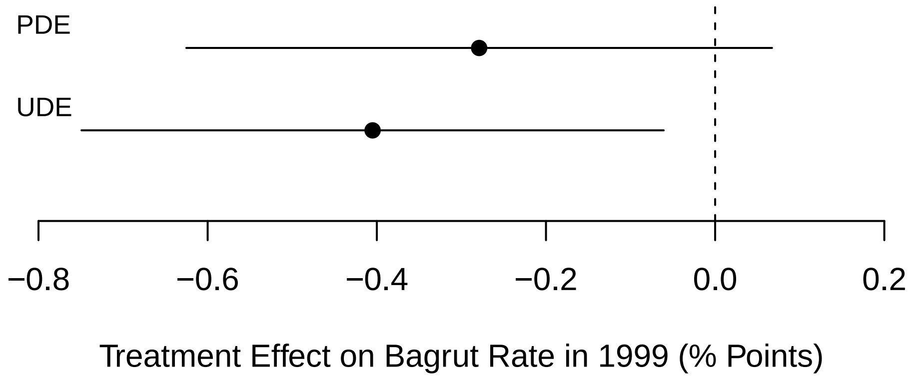

Fig. 2. Inference of the Average Treatment Effect on Bagrut Rate in 1999 based on a t-distribution with 20 degrees of freedom.

and unbiased in the super-population (Propositions 4 (2) and 4* (2)), while the bias direction of $\hat{\mathbb{V}}^{\text{Neyman}}(\hat{\tau}_P)$ is unknown (Propositions 3 (2) and 3* (2)). Note that, however, all of these points are concerned with only this application; I cannot generalize these minor remarks to other cases.

Finally, I emphasize that I do not aim to challenge any findings of the original analysis by Angrist & Lavy (2009) because their unit of observation is a student and mine is a school. I repeat that the goal of my reanalysis is to demonstrate how the two ATE estimators and the two variance estimators work in a real application, not to look for new substantive findings about the effect of the Achievement Award program on the Bagrut rate.

5. COMPARISON WITH IMAI AND JIANG (2018)

Imai & Jiang (2018) is very relevant to my study. In fact, Imai & Jiang (2018, 2908) cite an earlier version of my present manuscript, saying "[t]he analytical results and methodology presented in this paper *complement* the recent work by Fukumoto who examines the bias due to missing outcomes under the matched-pairs design" (emphasis added). I completely agree with them. Below, I compare the main manuscript with Imai & Jiang (2018).

Both studies share an estimand and an estimator. What I call "a kind of local average treatment effect (LATE) of 'always-reporting pairs' " (τ_P, p. 3) corresponds to "the average treatment effect for always-observed pairs (ATOP)," the key concept of Imai & Jiang (2018, 2909). Accordingly, the PDE ($\hat{\tau}_P$) is the same as the "naive difference-in-means estimator, applied to all the pairs without missing outcomes" ($\hat{\tau}_{\text{ATOP}}$) in Imai & Jiang (2018, 2909).

When Imai & Jiang (2018, 2910–2911) consider no-assumption bounds, they make neither my Assumption 2 nor my Assumption 3. In their sensitivity analysis, Imai & Jiang (2018, 2911) introduce their Assumption 2:

$$P\{R_{2j}(t) = r \mid R_{1j}(t) = r\} = P\{R_{1j}(t) = r \mid R_{2j}(t) = r\} \geq \gamma$$

for $r = 0, 1$ and $t = 0, 1$. When $\gamma = 1$, this is reduced to my Assumption 2. Thus, their Assumption 2 is more relaxed than my Assumption 2. All told, Imai & Jiang (2018) derive the bounds of the ATOP in their Theorems 1 and 3 under assumptions that are more relaxed than those of my paper.

On the other hand, if (but not only if) $\gamma = 1$, the bound is reduced to a point, the ATOP ($= \tau_P$) is identified, and not only my Assumption 2 but also my Assumptions 3 and 3* hold.

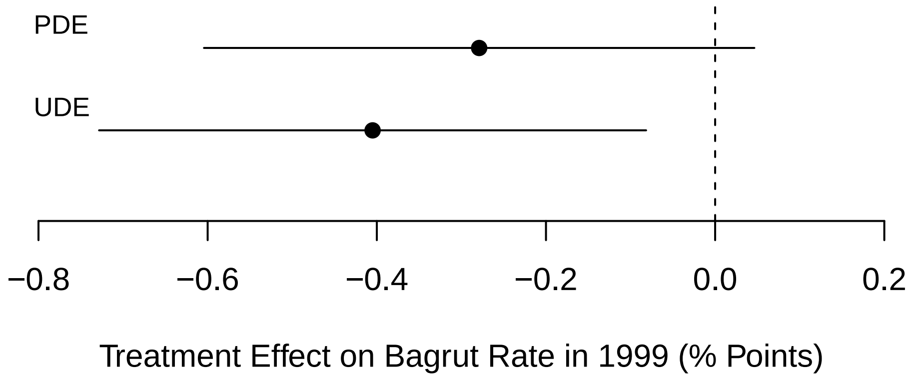

Fig. 3. Inference of the Average Treatment Effect on Bagrut Rate in 1999 based on a normal distribution.

My contribution is to show that under my Assumption 3 or 3*, we can derive the closed form of the variance of $\hat{\tau}_P = \hat{\tau}_{\mathsf{ATOP}}$ (my Proposition 2 (2) or 2* (2)) and the Adjusted Neyman variance estimator ($\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\cdot)$) of $\hat{\tau}_P = \hat{\tau}_{\mathsf{ATOP}}$ is only upwardly biased (and thus conservative) in the finite sample (my Proposition 4 (2)) or unbiased in a super-population (my Proposition 4* (2)). (In this sense, attention to the difference between a finite sample and a super-population is essential.) For instance, Table 1 of Imai & Jiang (2018, 2915) reports not only the "[n]aive estimates using available units" (which I denote by $\hat{\tau}_{\mathsf{ATOU}}$ and is equal to the UDE ($\hat{\tau}_U$) of my paper, the top panel) and $\hat{\tau}_{\mathsf{ATOP}}$ (= $\hat{\tau}_P$, the second panel) but also their corresponding confidence intervals. My paper provides some useful guidance about how to interpret the table. On the one hand, as long as readers suspect that my Assumption 3 or 3* is not violated, my Propositions 1 (2) and 4 (2) (or my Propositions 1* (2) and 4* (2)) will make them believe that the ATOP probably falls in the confidence interval of $\hat{\tau}_{\mathsf{ATOP}}$. On the other hand, even though readers allow my Assumption 2 or 2*, which is stronger than Assumption 3 or 3*, my Proposition 1 (3) or 1* (3) does not guarantee that $\hat{\tau}_{\mathsf{ATOU}}$ is unbiased for the average treatment effect for always-observed units (ATOU, $E\{Y_{ij}(1) - Y_{ij}(0) \mid R_{ij}(1) = R_{ij}(0) = 1\}$), and my Proposition 3 (3) or 3* (3) suggests that the confidence interval of $\hat{\tau}_{\mathsf{ATOU}}$ can be too narrow. Accordingly, readers should not be confident that the ATOU belongs to the confidence interval of $\hat{\tau}_{\mathsf{ATOU}}$.

Furthermore, I argue that we do not have to be as pessimistic as Imai & Jiang (2018). They warn "in practice, the treatment *often* affects the missingness pattern [$R_{ij}(1) \neq R_{ij}(0)$] and matching is imperfect [$R_{1j}(t) \neq R_{2j}(t)$]" (p. 2910, emphasis added and square brackets inserted). Here I hasten to add that this is *not always* the case. For instance, in the cases of blind tests, subliminal stimuli, and administrative records, it is likely that $R_{ij}(1) = R_{ij}(0)$ (p. 3); since unit i of pair j does not recognize treatment status, whether the unit responds or not (R_{ij}) will not depend on whether treatment is assigned or not. Or, if matched-on variables (e.g., (part of) DNA) completely explain the missingness pattern (e.g., in the cases of (monozygotic) twins and littermates of the same sex), $R_{1j}(t) = R_{2j}(t)$ should hold (p. 2). I also emphasize that my Assumption 3 is more relaxed than the conditions under which Imai & Jiang (2018, 2909–2910) argue $\hat{\tau}_{\mathsf{ATOP}}$ "is unbiased for ATOP." Their condition is either "$R_{ij}(1) = R_{ij}(0)$ for all i and j" or "$R_{1j}(t) = R_{2j}(t)$ for each $t = 0, 1$ and $j = 1, 2, \ldots, J$." Indeed, my Assumption 3 is, *for each pair* j, either $R_{ij}(1) = R_{ij}(0)$ for all i or $R_{1j}(t) = R_{2j}(t)$ for each $t = 0, 1$. Note also that even under my Assumption 3, $\hat{\tau}_{\mathsf{ATOP}}$ (= $\hat{\tau}_P$) is unbiased for ATOP (= τ_P) according to Section 2.2.

At least, even if we make my Assumption 3 for mathematical reasons, the closed form of the variance of $\hat{\tau}_P$ and the closed form of the bias of $\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_P)$ make it clear what is mainly responsible for the inefficiency and the bias. That is, the variance of $\hat{\tau}_P$ increases in the variance of within-pair deviation $(\omega_{ij}(1) + \omega_{ij}(0))$, which represents the quality of pair matching (my Proposition 2 (2)), while the bias of $\hat{\mathbb{V}}^{\text{Adj-Neyman}}(\hat{\tau}_P)$ increases in the variance of between-pair deviation $(\beta_{ij}(1) - \beta_{ij}(0))$, which indicates the between-pair heterogeneity of treatment effects (my Proposition 4 (2)).

To conclude, I emphasize that Imai & Jiang (2018) and my paper complement each other on the same topic: non-ignorable attrition in pairwise randomized experiments. Imai & Jiang (2018) are interested in bounds and its sensitivity to assumptions, while my main focus is on comparison between the PDE and UDE.

6. CONCLUDING REMARKS

I emphasize that both the PDE and the adjusted Neyman variance estimator have the favorable properties outlined above because they take advantage of pairwise randomization's design. It is clear that basically, one can translate propositions about the full sample estimator into those of the PDE by replacing $\boldsymbol{k}_F, \boldsymbol{k}_F^*$ and n_F by $\boldsymbol{k}_P, \boldsymbol{k}_P^*$ and n_P, respectively. This is because, under the assumption of pairwise matched attrition, the PDE regards the always-reporting pairs as the full sample. This is the case neither with the UDE nor with the Neyman variance estimator; they "break the match," or break the design of pairwise randomization. Furthermore, the framework of this study can also be extended to observational data where scholars apply a matching method, one outcome variable (e.g., event occurrence) has no missing values, and another outcome variable (e.g., time-to-event) has some. In this case, too, scholars should use the PDE rather than the UDE. The Neyman variance estimator can have either positive or negative bias for both ATE estimators (Propositions 3 and 3*) and thus is not recommended.

Non-ignorable attrition in pairwise randomized experiments has attracted less attention than it should and requires thorough consideration. This work aims to move towards a solution for this problem, though many questions remain unanswered. For instance, researchers can consider a pair as a stratum which is composed of two units, while, in many studies, a stratum is composed of more than two units (stratified randomized experiments, Imbens & Rubin 2015, ch. 9). In this case, inverse probability weighting (e.g. Little & Rubin 2002, pp. 46–47) is available, where the weighting class is a stratum and the weight variable (Z_{ij}) is not a dummy but a non-negative real number which can be larger than one. The question is whether or under what conditions the PDE is better than inverse probability weighting. Another example is non-compliance; units sometimes do not comply with treatment assignment even though their outcomes are observed. A typical solution is instrumental variable estimation (Angrist et al. 1996). This study sketches unbiasedness of both the PDE and the instrumental variable estimator for compliers' LATE, though it remains unclear which estimator is more efficient. Finally, pairwise randomization is often applied to cluster randomized experiments, and researchers may conduct not only cluster-level but also individual-level analyses (Donner & Klar 2000; Hayes & Moulton 2009; Imai et al. 2009). My conjecture is that the implication of the propositions of this manuscript still holds even at the individual level, even though specific equations should differ. Studying these topics is a future agenda for causal inference research.

References

Abadie, A. & Imbens, G. (2008), 'Estimation of the Conditional Variance in Paired Experiments', *Annales d'Economie et de Statistique* **91/92**, 175–187.

Angrist, J. D., Imbens, G. W. & Rubin, D. B. (1996), 'Identification of Causal Effects Using Instrumental Variables', *Journal of the American Statistical Association* **91**(434), 444–455.

Angrist, J. & Lavy, V. (2009), 'The Effects of High Stakes High School Achievement Awards: Evidence From a Randomized Trial', *American Economic Review* **99**(4), 1384–1414.

Barnard, J., Frangakis, C. E., Hill, J. L. & Rubin, D. B. (2003), 'Principal Stratification Approach to Broken Randomized Experiments', *Journal of the American Statistical Association* **98**(462), 299–323.

Box, G. E. P., Hunter, J. S. & Hunter, W. G. (2005), *Statistics for Experimenters: Design, Innovation, and Discovery*, 2nd edn, John Wiley & Sons, Hoboken.

Donner, A. & Klar, N. (2000), *Design and Analysis of Cluster Randomization Trials in Health Research*, Arnold, London.

Duflo, E., Glennerster, R. & Kremer, M. (2008), Using Randomization in Development Economics Research: A Toolkit, *in* 'Handbook of Development Economics', Vol. 4, Noth Holland, Amsterdam, pp. 3895–3962.

Dunning, T. (2011), 'Does Blocking Reduce Attrition Bias?', *Experimental Political Scientist: Newsletter of the APSA Experimental Section* **2**(1), 12–16.

Enos, R. D. (2014), 'Causal Effect of Intergroup Contact on Exclusionary Attitudes', *Proceedings of the National Academy of Sciences* **111**(10), 3699–3704.

Family Heart Study Group (1994), 'Randomised Controlled Trial Evaluating Cardiovascular Screening and Intervention in General Practice: Principal Results of British Family Heart Study', *British Medical Journal* **308**(6924), 313–320.

Frangakis, C. E. & Rubin, D. B. (1999), 'Addressing Complications of Intention-to-Treat Analysis in the Combined Presence of All-or-None Treatment-Noncompliance and Subsequent Missing Outcomes', *Biometrika* **86**(2), 365–379.

Gerber, A. S. & Green, D. P. (2012), *Field Experiments: Design, Analysis, and Interpretation*, W. W. Norton and Company, New York.

Glennerster, R. & Takavarasha, K. (2013), *Running Randomized Evaluations*, Princeton University Press, Princeton.

Hayes, R. J. & Moulton, L. H. (2009), *Cluster Randomised Trials*, Chapman & Hall/CRC, Boca Raton.

Hayward, A. C., Harling, R., Wetten, S., Johnson, A. M., Munro, S., Smedley, J., Murad, S. & Watson, J. M. (2006), 'Effectiveness of an Influenza Vaccine Programme for Care Home Staff to Prevent Death, Morbidity, and Health Service Use Among Residents: Cluster Randomised Controlled Trial', *British Medical Journal* **333**(7581), 1241–1244.

Imai, K. (2008), 'Variance Identification and Efficiency Analysis in Randomized Experiments Under the Matched-Pair Design', *Statistics in Medicine* **27**(24), 4857–4873.

Imai, K. & Jiang, Z. (2018), 'A sensitivity analysis for missing outcomes due to truncation by death under the matched-pairs design', *Statistics in Medicine* **37**(20), 2907–2922.

Imai, K., King, G. & Nall, C. (2009), 'The Essential Role of Pair Matching in Cluster-Randomized Experiments, With Application to the Mexican Universal Health Insurance Evaluation', *Statistical Science* **24**(1), 29–53.

Imbens, G. W. & Rubin, D. B. (2015), *Causal Inference for Statistics, Social, and Biomedical Sciences*, Cambridge University Press.

King, G., Gakidou, E., Ravishankar, N., Moore, R. T., Lakin, J., Vargas, M., Téllez-Rojo, M. M., Ávila, J. E. H., Ávila, M. H. & Llamas, H. H. (2007), 'A "Politically Robust" Experimental Design for Public Policy Evaluation, With Application to the Mexican Universal Health Insurance Program', *Journal of Policy Analysis and Management* **26**, 479–506.

Little, R. J. a. & Rubin, D. B. (2002), *Statistical Analysis with Missing Data*, 2nd edn, John Wiley & Sons, Hoboken.

Loeb, M., Carusone, S. C., Goeree, R., Walter, S. D., Brazil, K., Krueger, P., Simor, A., Moss, L. & Marrie, T. (2006), 'Effect of a Clinical Pathway to Reduce Hospitalizations in Nursing Home Residents With Pneumonia: A Randomized Controlled Trial', *Journal of the American Medical Association* **295**(21), 2503–2510.

Panagopoulos, C. & Green, D. P. (2008), 'Field Experiments Testing the Impact of Radio Advertisements on Electoral Competition', *American Journal of Political Science* **52**(1), 156–168.

Snedecor, G. W. & Cochran, W. G. (1980), *Statistical Methods*, 7th edn, The Iowa State University Press, Ames, Iowa.