Appendix

This appendix contains the proofs of the theorems presented in the paper entitled "Comparing Correction Methods to Reduce Misclassification Bias". Recall that we have assumed a population of size N in which a fraction $\alpha \coloneqq N_{1+}/N$ belongs to the class of interest, referred to as the class labelled as 1. We assume that a binary classification algorithm has been trained that correctly classifies a data point that belongs to class $i \in \{0, 1\}$ with probability $p_{ii} > 0.5$, independently across all data points. In addition, we assume that a test set of size $n \ll N$ is available and that it can be considered a simple random sample from the population. The classification probabilities p_{00} and p_{11} are estimated on that test set as described in Section 2. Finally, we assume that the classify-and-count estimator $\hat{\alpha}^*$ is distributed independently of $\hat{p}_{00}$ and $\hat{p}_{11}$, which is reasonable (at least as an approximation) when $n \ll N$.

It may be noted that the estimated probabilities $\hat{p}_{11}$ and $\hat{p}_{00}$ defined in Section 2 cannot be computed if $n_{1+} = 0$ or $n_{0+} = 0$. Similarly, the calibration probabilities c_{11} and c_{00} cannot be estimated if $n_{+1} = 0$ or $n_{+0} = 0$. We assume here that these events occur with negligible probability. This will be true when n is sufficiently large so that $n\alpha \gg 1$ and $n(1-\alpha) \gg 1$.

Preliminaries

Many of the proofs presented in this appendix rely on the following two mathematical results. First, we will use univariate and bivariate Taylor series to approximate the expectation of non-linear functions of random variables. That is, to estimate $E[f(X)]$ and $E[g(X,Y)]$ for sufficiently differentiable functions f and g, we will insert the Taylor series for f and g at $x_0 = E[X]$ and $y_0 = E[Y]$ up to terms of order 2 and utilize the linearity of the expectation. Second, we will use the following conditional variance decomposition for the variance of a random variable X:

$$V(X) = E[V(X \mid Y)] + V(E[X \mid Y]). \tag{19}$$

The conditional variance decomposition follows from the tower property of conditional expectations [10]. Before we prove the theorems presented in the paper, we begin by proving the following lemma.

Lemma 1. *The variance of the estimator $\hat{p}_{11}$ for p_{11} estimated on the test set is given by*

$$V(\hat{p}_{11}) = \frac{p_{11}(1-p_{11})}{n\alpha}\left[1 + \frac{1-\alpha}{n\alpha}\right] + O\left(\frac{1}{n^3}\right). \tag{20}$$

Similarly, the variance of $\hat{p}_{00}$ is given by

$$V(\hat{p}_{00}) = \frac{p_{00}(1-p_{00})}{n(1-\alpha)}\left[1 + \frac{\alpha}{n(1-\alpha)}\right] + O\left(\frac{1}{n^3}\right). \tag{21}$$

Moreover, $\hat{p}_{11}$ and $\hat{p}_{00}$ are uncorrelated: $C(\hat{p}_{11}, \hat{p}_{00}) = 0$.

Proof (of Lemma 1). We approximate the variance of $\hat{p}_{00}$ using the conditional variance decomposition and a second-order Taylor series, as follows:

$$\begin{aligned}
V(\hat{p}_{00}) &= V\left(\frac{n_{00}}{n_{0+}}\right) \\
&= E_{n_{0+}}\left[V\left(\frac{n_{00}}{n_{0+}} \mid n_{0+}\right)\right] + V_{n_{0+}}\left[E\left(\frac{n_{00}}{n_{0+}} \mid n_{0+}\right)\right] \\
&= E_{n_{0+}}\left[\frac{1}{n_{0+}^2} V(n_{00} \mid n_{0+})\right] + V_{n_{0+}}\left[\frac{1}{n_{0+}} E(n_{00} \mid n_{0+})\right] \\
&= E_{n_{0+}}\left[\frac{n_{0+} p_{00}(1-p_{00})}{n_{0+}^2}\right] + V_{n_{0+}}\left[\frac{n_{0+} p_{00}}{n_{0+}}\right] \\
&= E_{n_{0+}}\left[\frac{1}{n_{0+}}\right] p_{00}(1-p_{00}) \\
&= \left[\frac{1}{E[n_{0+}]} + \frac{1}{2}\frac{2}{E[n_{0+}]^3} \times V[n_{0+}]\right] p_{00}(1-p_{00}) + O\left(\frac{1}{n^3}\right) \\
&= \frac{p_{00}(1-p_{00})}{E[n_{0+}]}\left[1 + \frac{V[n_{0+}]}{E[n_{0+}]^2}\right] + O\left(\frac{1}{n^3}\right) \\
&= \frac{p_{00}(1-p_{00})}{n(1-\alpha)}\left[1 + \frac{\alpha}{n(1-\alpha)}\right] + O\left(\frac{1}{n^3}\right).
\end{aligned}$$

The variance of $\hat{p}_{11}$ is approximated in the exact same way.

Finally, to evaluate $C(\hat{p}_{11}, \hat{p}_{00})$ we use the analogue of (19) for covariances:

$$\begin{aligned}
C(\hat{p}_{11}, \hat{p}_{00}) &= C\left(\frac{n_{11}}{n_{1+}}, \frac{n_{00}}{n_{0+}}\right) \\
&= E_{n_{1+},n_{0+}}\left[C\left(\frac{n_{11}}{n_{1+}}, \frac{n_{00}}{n_{0+}} \mid n_{1+}, n_{0+}\right)\right] \\
&\qquad + C_{n_{1+},n_{0+}}\left[E\left(\frac{n_{11}}{n_{1+}} \mid n_{1+}, n_{0+}\right), E\left(\frac{n_{00}}{n_{0+}} \mid n_{1+}, n_{0+}\right)\right] \\
&= E_{n_{1+},n_{0+}}\left[\frac{1}{n_{1+} n_{0+}} C(n_{11}, n_{00} \mid n_{1+}, n_{0+})\right] \\
&\qquad + C_{n_{1+},n_{0+}}\left[\frac{1}{n_{1+}} E(n_{11} \mid n_{1+}), \frac{1}{n_{0+}} E(n_{00} \mid n_{0+})\right].
\end{aligned}$$

The second term is zero as before. The first term also vanishes because, conditional on the row totals n_{1+} and n_{0+}, the counts n_{11} and n_{00} follow independent binomial distributions, so $C(n_{11}, n_{00} \mid n_{1+}, n_{0+}) = 0$.

Note: in the remainder of this appendix, we will not add explicit subscripts to expectations and variances when their meaning is unambiguous.

Subtracted-bias estimator

We will now prove the bias and variance approximations for the subtracted-bias estimator $\hat{\alpha}_b$ that was defined in Equation (9).

Proof (of Theorem 1). The bias of $\hat{\alpha}_b$ is given by

$$\begin{aligned}
B(\hat{\alpha}_b) &= E\left[\hat{\alpha}^\star - \hat{B}[\hat{\alpha}^\star]\right] - \alpha \\
&= E[\hat{\alpha}^\star - \alpha] - E\left[\hat{B}[\hat{\alpha}^\star]\right] \\
&= B[\hat{\alpha}^\star] - E\left[\hat{B}[\hat{\alpha}^\star]\right] \\
&= [\alpha(p_{00} + p_{11} - 2) + (1 - p_{00})] - E\left[\hat{\alpha}^\star(\hat{p}_{00} + \hat{p}_{11} - 2) + (1 - \hat{p}_{00})\right].
\end{aligned}$$

Because $\hat{\alpha}^*$ and $(\hat{p}_{00} + \hat{p}_{11} - 2)$ are assumed to be independent, the expectation of their product equals the product of their expectations:

$$\begin{aligned}
B(\hat{\alpha}_b) &= \alpha(p_{00} + p_{11} - 2) + (1 - p_{00}) - E[\hat{\alpha}^\star](p_{00} + p_{11} - 2) - (1 - p_{00}) \\
&= (\alpha - E[\hat{\alpha}^\star])(p_{00} + p_{11} - 2) \\
&= B[\hat{\alpha}^\star](2 - p_{00} - p_{11}) \\
&= (1 - p_{00})(2 - p_{00} - p_{11}) - \alpha(p_{00} + p_{11} - 2)^2.
\end{aligned}$$

This proves the formula for the bias of $\hat{\alpha}_b$ as estimator for α. To approximate the variance of $\hat{\alpha}_b$, we apply the conditional variance decomposition (19) conditional on $\hat{\alpha}^*$ and look at the two resulting terms separately. First, consider the expectation of the conditional variance:

$$\begin{aligned}
E\left[V(\hat{\alpha}_b \mid \hat{\alpha}^*)\right] &= E\left[V(\hat{\alpha}^*(3 - \hat{p}_{00} - \hat{p}_{11}) - (1 - \hat{p}_{00}) \mid \hat{\alpha}^*)\right] \\
&= E\big[V(\hat{\alpha}^*(3 - \hat{p}_{00} - \hat{p}_{11}) \mid \hat{\alpha}^*) + V(1 - \hat{p}_{00} \mid \hat{\alpha}^*) \\
&\quad - 2C(\hat{\alpha}^*(3 - \hat{p}_{00} - \hat{p}_{11}), 1 - \hat{p}_{00} \mid \hat{\alpha}^*)\big] \\
&= E\big[(\hat{\alpha}^*)^2 V(3 - \hat{p}_{00} - \hat{p}_{11} \mid \hat{\alpha}^*) + V(1 - \hat{p}_{00} \mid \hat{\alpha}^*) \\
&\quad - 2\hat{\alpha}^* C(3 - \hat{p}_{00} - \hat{p}_{11}, 1 - \hat{p}_{00} \mid \hat{\alpha}^*)\big] \\
&= E\big[(\hat{\alpha}^*)^2 \left[V(\hat{p}_{00}) + V(\hat{p}_{11})\right] + V(\hat{p}_{00}) - 2\hat{\alpha}^* V(\hat{p}_{00})\big] \\
&= E\left[(\hat{\alpha}^*)^2\right] \left[V(\hat{p}_{00}) + V(\hat{p}_{11})\right] + V(\hat{p}_{00}) - 2E\left[\hat{\alpha}^*\right] V(\hat{p}_{00}).
\end{aligned}$$

In the penultimate line, we used that $C(\hat{p}_{11}, \hat{p}_{00}) = 0$. The second moment $E\left[(\hat{\alpha}^*)^2\right]$ can be written as $E\left[\hat{\alpha}^*\right]^2 + V(\hat{\alpha}^*)$. Because $V(\hat{\alpha}^*)$ is of order $1/N$, it can be neglected compared to $E\left[\hat{\alpha}^*\right]^2$, which is of order 1. In particular, we find that the expectation of the conditional variance equals:

$$\begin{aligned}
E\left[V(\hat{\alpha}_b \mid \hat{\alpha}^*)\right] &= E\left[(\hat{\alpha}^*)\right]^2 \left[V(\hat{p}_{00}) + V(\hat{p}_{11})\right] + V(\hat{p}_{00}) - 2E\left[\hat{\alpha}^*\right] V(\hat{p}_{00}) + O\left(\frac{1}{N}\right) \\
&= V(\hat{p}_{00}) \left[E\left[\hat{\alpha}^*\right] - 1\right]^2 + V(\hat{p}_{11}) E\left[\hat{\alpha}^*\right]^2 + O\left(\frac{1}{N}\right).
\end{aligned}$$

Next, the variance of the conditional expectation can be seen to be equal the following:

$$\begin{aligned}
V\left[E(\hat{\alpha}_b \mid \hat{\alpha}^*)\right] &= V\left[E(\hat{\alpha}^*(3 - \hat{p}_{00} - \hat{p}_{11}) - (1 - \hat{p}_{00}) \mid \hat{\alpha}^*)\right] \\
&= V\left[\hat{\alpha}^* E(3 - \hat{p}_{00} - \hat{p}_{11} \mid \hat{\alpha}^*) - E(1 - \hat{p}_{00} \mid \hat{\alpha}^*)\right] \\
&= V(\hat{\alpha}^*)(3 - p_{00} - p_{11})^2.
\end{aligned}$$

Because $V(\hat{\alpha}^*)$ is of order $1/N$, it can be neglected in the final formula. Furthermore, the variances of $\hat{p}_{00}$ and $\hat{p}_{11}$ can be written out using the result from Lemma 1:

$$\begin{aligned} V(\hat{\alpha}_b) = & \frac{\left[\alpha(p_{00}+p_{11}-1)-p_{00}\right]^2 p_{00}(1-p_{00})}{n(1-\alpha)}\left[1+\frac{\alpha}{n(1-\alpha)}\right] \\ & + \frac{\left[\alpha(p_{00}+p_{11}-1)+(1-p_{00})\right]^2 p_{11}(1-p_{11})}{n\alpha}\left[1+\frac{1-\alpha}{n\alpha}\right] \\ & + O\left(\max\left[\frac{1}{n^3},\frac{1}{N}\right]\right). \end{aligned}$$

This concludes the proof of Theorem 1.

Misclassification estimator

We will now prove the bias and variance approximations for the misclassification estimator $\hat{\alpha}_p$ as defined in Equation (12).

Proof (of Theorem 2). Under the assumption that $\hat{\alpha}^*$ is distributed independently of $(\hat{p}_{00}, \hat{p}_{11})$, it holds that

$$\begin{aligned} E(\hat{\alpha}_p) &= E\left(\frac{\hat{p}_{00}-1}{\hat{p}_{00}+\hat{p}_{11}-1}\right) + E\left[E\left(\frac{\hat{\alpha}^*}{\hat{p}_{00}+\hat{p}_{11}-1}\;\middle|\;\hat{\alpha}^*\right)\right] \\ &= E\left(\frac{\hat{p}_{00}-1}{\hat{p}_{00}+\hat{p}_{11}-1}\right) + E(\hat{\alpha}^*)E\left(\frac{1}{\hat{p}_{00}+\hat{p}_{11}-1}\right). \end{aligned} \tag{22}$$

$E(\hat{\alpha}^*)$ is known from (4). To evaluate the other two expectations, we use a second-order Taylor series approximation. The first- and second-order partial derivatives of $f(x,y) = 1/(x+y-1)$ and $g(x,y) = (x-1)/(x+y-1) = 1-[y/(x+y-1)]$ are given by:

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = \frac{-1}{(x+y-1)^2}, \tag{23}$$

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 f}{\partial y^2} = \frac{2}{(x+y-1)^3},$$

$$\frac{\partial g}{\partial x} = \frac{y}{(x+y-1)^2}, \tag{24}$$

$$\frac{\partial g}{\partial y} = \frac{-(x-1)}{(x+y-1)^2}, \tag{25}$$

$$\frac{\partial^2 g}{\partial x^2} = \frac{-2y}{(x+y-1)^3},$$

$$\frac{\partial^2 g}{\partial y^2} = \frac{2(x-1)}{(x+y-1)^3}.$$

Now also using that $C(\hat{p}_{11}, \hat{p}_{00}) = 0$, we obtain for the first expectation:

$$\begin{aligned}E\left(\frac{1}{\hat{p}_{00}+\hat{p}_{11}-1}\right) &= \frac{1}{p_{00}+p_{11}-1} + \frac{V(\hat{p}_{00})+V(\hat{p}_{11})}{(p_{00}+p_{11}-1)^3} + O(n^{-2}) \\ &= \frac{1}{p_{00}+p_{11}-1}\left[1+\frac{\frac{p_{00}(1-p_{00})}{n(1-\alpha)}+\frac{p_{11}(1-p_{11})}{n\alpha}}{(p_{00}+p_{11}-1)^2}\right] + O(n^{-2}).\end{aligned} \tag{26}$$

Here, we have included only the first term of the approximations to $V(\hat{p}_{00})$ and $V(\hat{p}_{11})$ from Lemma 1, since this suffices to approximate the bias up to terms of order $O(1/n)$. Similarly, for the second expectation we obtain:

$$\begin{aligned}E\left(\frac{\hat{p}_{00}-1}{\hat{p}_{00}+\hat{p}_{11}-1}\right) &= \frac{p_{00}-1}{p_{00}+p_{11}-1} + \frac{(p_{00}-1)V(\hat{p}_{11})-p_{11}V(\hat{p}_{00})}{(p_{00}+p_{11}-1)^3} + O(n^{-2}) \\ &= \frac{p_{00}-1}{p_{00}+p_{11}-1}\left[1+p_{11}\frac{\frac{1-p_{11}}{n\alpha}+\frac{p_{00}}{n(1-\alpha)}}{(p_{00}+p_{11}-1)^2}\right] + O(n^{-2}).\end{aligned} \tag{27}$$

Using (22), (4), (26), and (27), we conclude that:

$$\begin{aligned}E(\hat{\alpha}_p) = &\frac{\alpha(p_{00}+p_{11}-1)-(p_{00}-1)}{p_{00}+p_{11}-1}\left[1+\frac{\frac{p_{00}(1-p_{00})}{n(1-\alpha)}+\frac{p_{11}(1-p_{11})}{n\alpha}}{(p_{00}+p_{11}-1)^2}\right] \\ &+ \frac{p_{00}-1}{p_{00}+p_{11}-1}\left[1+p_{11}\frac{\frac{1-p_{11}}{n\alpha}+\frac{p_{00}}{n(1-\alpha)}}{(p_{00}+p_{11}-1)^2}\right] + O\left(\frac{1}{n^2}\right).\end{aligned}$$

From this, it follows that an approximation to the bias of $\hat{\alpha}_p$ that is correct up to terms of order $O(1/n)$ is given by:

$$\begin{aligned}B(\hat{\alpha}_p) = &\frac{\alpha(p_{00}+p_{11}-1)-(p_{00}-1)}{n(p_{00}+p_{11}-1)^3}\left[\frac{p_{00}(1-p_{00})}{1-\alpha}+\frac{p_{11}(1-p_{11})}{\alpha}\right] \\ &+ \frac{(p_{00}-1)p_{11}}{n(p_{00}+p_{11}-1)^3}\left[\frac{1-p_{11}}{\alpha}+\frac{p_{00}}{1-\alpha}\right] + O\left(\frac{1}{n^2}\right).\end{aligned}$$

By expanding the products in this expression and combining similar terms, the expression can be simplified to:

$$B(\hat{\alpha}_p) = \frac{p_{11}(1-p_{11})-p_{00}(1-p_{00})}{n(p_{00}+p_{11}-1)^2} + O\left(\frac{1}{n^2}\right).$$

Finally, using the identity $p_{11}(1-p_{11})-p_{00}(1-p_{00}) = (p_{00}+p_{11}-1)(p_{00}-p_{11})$, we obtain the required result for $B(\hat{\alpha}_p)$.

To approximate the variance of $\hat{\alpha}_p$, we apply the conditional variance decomposition conditional on $\hat{\alpha}^*$ and look at the two resulting terms separately. First, consider the variance of the conditional expectation:

$$
\begin{aligned}
V\left[E(\hat{\alpha}_p \mid \hat{\alpha}^*)\right] &= V\left[E\left(\hat{\alpha}^* \frac{1}{\hat{p}_{00}+\hat{p}_{11}-1} + \frac{\hat{p}_{00}-1}{\hat{p}_{00}+\hat{p}_{11}-1} \mid \hat{\alpha}^*\right)\right] \\
&= V\left[\hat{\alpha}^* \frac{1}{p_{00}+p_{11}-1}\right] \\
&= \frac{1}{(p_{00}+p_{11}-1)^2} V\left[\hat{\alpha}^*\right] = O\left(\frac{1}{N}\right), \qquad (28)
\end{aligned}
$$

where in the last line we used (6). Note: the factor $1/(p_{00}+p_{11}-1)^2$ can become arbitrarily large in the limit $p_{00}+p_{11} \to 1$. It will be seen below that this same factor also occurs in the lower-order terms of $V(\hat{\alpha}_p)$; hence, the relative contribution of (28) remains negligible even in the limit $p_{00}+p_{11} \to 1$.

Next, we compute the expectation of the conditional variance.

$$
\begin{aligned}
E\left[V(\hat{\alpha}_p \mid \hat{\alpha}^*)\right] &= E\left[V\left(\hat{\alpha}^* \frac{1}{\hat{p}_{00}+\hat{p}_{11}-1} + \frac{\hat{p}_{00}-1}{\hat{p}_{00}+\hat{p}_{11}-1} \mid \hat{\alpha}^\star\right)\right] \\
&= E\Bigg[V\left(\hat{\alpha}^* \frac{1}{\hat{p}_{00}+\hat{p}_{11}-1} \mid \alpha^\star\right) + V\left(\frac{\hat{p}_{00}-1}{\hat{p}_{00}+\hat{p}_{11}-1} \mid \hat{\alpha}^\star\right) \\
&\quad + 2C\left(\hat{\alpha}^* \frac{1}{\hat{p}_{00}+\hat{p}_{11}-1}, \frac{\hat{p}_{00}-1}{\hat{p}_{00}+\hat{p}_{11}-1} \mid \hat{\alpha}^\star\right)\Bigg] \\
&= E\left[(\hat{\alpha}^*)^2\right] V\left[\frac{1}{\hat{p}_{00}+\hat{p}_{11}-1}\right] + V\left[\frac{\hat{p}_{00}-1}{\hat{p}_{00}+\hat{p}_{11}-1}\right] \\
&\quad + 2E\left[\hat{\alpha}^\star\right] C\left[\frac{1}{\hat{p}_{00}+\hat{p}_{11}-1}, \frac{\hat{p}_{00}-1}{\hat{p}_{00}+\hat{p}_{11}-1}\right] \\
&= E\left[\hat{\alpha}^\star\right]^2 \left[1 + O\left(\frac{1}{N}\right)\right] V\left[\frac{1}{\hat{p}_{00}+\hat{p}_{11}-1}\right] + V\left[\frac{\hat{p}_{00}-1}{\hat{p}_{00}+\hat{p}_{11}-1}\right] \\
&\quad + 2E\left[\hat{\alpha}^\star\right] C\left[\frac{1}{\hat{p}_{00}+\hat{p}_{11}-1}, \frac{\hat{p}_{00}-1}{\hat{p}_{00}+\hat{p}_{11}-1}\right]. \qquad (29)
\end{aligned}
$$

To approximate the variance and covariance terms, we use a first-order Taylor series. Using the partial derivatives in (23), (24) and (25), we obtain:

$$
\begin{aligned}
V\left[\frac{1}{\hat{p}_{00}+\hat{p}_{11}-1}\right] &= \frac{V(\hat{p}_{00}) + V(\hat{p}_{11})}{(p_{00}+p_{11}-1)^4} + O(n^{-2}) \\
V\left[\frac{\hat{p}_{00}-1}{\hat{p}_{00}+\hat{p}_{11}-1}\right] &= \frac{V(\hat{p}_{00})(p_{11})^2}{(p_{00}+p_{11}-1)^4} + \frac{V(\hat{p}_{11})(1-p_{00})^2}{(p_{00}+p_{11}-1)^4} + O(n^{-2}) \\
C\left[\frac{1}{\hat{p}_{00}+\hat{p}_{11}-1}, \frac{\hat{p}_{00}-1}{\hat{p}_{00}+\hat{p}_{11}-1}\right] &= \frac{V(\hat{p}_{00})(-p_{11})}{(p_{00}+p_{11}-1)^4} + \frac{V(\hat{p}_{11})(p_{00}-1)}{(p_{00}+p_{11}-1)^4} + O(n^{-2}).
\end{aligned}
$$

Substituting these terms into Formula (29) and accounting for Formula (28) yields:

$$
\begin{aligned}
V(\hat{\alpha}_p) = & \frac{V(\hat{p}_{00})\left[E\left[\hat{\alpha}^\star\right]^2 - 2p_{11}E\left[\hat{\alpha}^\star\right] + p_{11}^2\right]}{(p_{00}+p_{11}-1)^4} \\
& + \frac{V(\hat{p}_{11})\left[E\left[\hat{\alpha}^\star\right]^2 - 2(1-p_{00})E\left[\hat{\alpha}^\star\right] + (1-p_{00})^2\right]}{(p_{00}+p_{11}-1)^4} + O\left(\max\left[\frac{1}{n^2},\frac{1}{N}\right]\right) \\
= & \frac{V(\hat{p}_{00})\left[E\left[\hat{\alpha}^\star\right] - p_{11}\right]^2}{(p_{00}+p_{11}-1)^4} + \frac{V(\hat{p}_{11})\left[E\left[\hat{\alpha}^\star\right] - (1-p_{00})\right]^2}{(p_{00}+p_{11}-1)^4} + O\left(\max\left[\frac{1}{n^2},\frac{1}{N}\right]\right) \\
= & \frac{V(\hat{p}_{00})(1-\alpha)^2}{(p_{00}+p_{11}-1)^2} + \frac{V(\hat{p}_{11})\alpha^2}{(p_{00}+p_{11}-1)^2} + O\left(\max\left[\frac{1}{n^2},\frac{1}{N}\right]\right).
\end{aligned}
$$

Finally, inserting the expressions for $V(\hat{p}_{00})$ and $V(\hat{p}_{11})$ from Lemma 1 yields:

$$
\begin{aligned}
V(\hat{\alpha}_p) = & \frac{\frac{p_{00}(1-p_{00})}{n(1-\alpha)}\left[1+\frac{\alpha}{n(1-\alpha)}\right](1-\alpha)^2}{(p_{00}+p_{11}-1)^2} + \frac{\frac{p_{11}(1-p_{11})}{n\alpha}\left[1+\frac{1-\alpha}{n\alpha}\right]\alpha^2}{(p_{00}+p_{11}-1)^2} \\
& + O\left(\max\left[\frac{1}{n^2},\frac{1}{N}\right]\right),
\end{aligned}
$$

from which expression (14) follows. This concludes the proof of Theorem 2.

Calibration estimator

We will now prove the bias and variance approximations for the calibration estimator $\hat{\alpha}_c$ that was defined in Equation (15).

Proof (of Theorem 3). To compute the expected value of $\hat{\alpha}_c$, we first compute its expectation conditional on the 4-vector $\boldsymbol{N} = (N_{00}, N_{01}, N_{10}, N_{11})$:

$$
\begin{aligned}
E(\hat{\alpha}_c \mid \boldsymbol{N}) &= E\left[\hat{\alpha}^* \frac{n_{11}}{n_{+1}} + (1-\hat{\alpha}^*)\frac{n_{10}}{n_{+0}} \mid \boldsymbol{N}\right] \\
&= \hat{\alpha}^* E\left[\frac{n_{11}}{n_{+1}} \mid \boldsymbol{N}\right] + (1-\hat{\alpha}^*)E\left[\frac{n_{10}}{n_{+0}} \mid \boldsymbol{N}\right] \\
&= \hat{\alpha}^* E\left[E\left(\frac{n_{11}}{n_{+1}} \mid \boldsymbol{N}, n_{+1}\right) \mid \boldsymbol{N}\right] \\
&\quad + (1-\hat{\alpha}^*)E\left[E\left(\frac{n_{10}}{n_{+0}} \mid \boldsymbol{N}, n_{+0}\right) \mid \boldsymbol{N}\right] \\
&= \frac{N_{+1}}{N}E\left[\frac{1}{n_{+1}}n_{+1}\frac{N_{11}}{N_{+1}} \mid \boldsymbol{N}\right] + \frac{N_{+0}}{N}E\left[\frac{1}{n_{+0}}n_{+0}\frac{N_{10}}{N_{+0}} \mid \boldsymbol{N}\right] \\
&= \frac{N_{11}}{N} + \frac{N_{10}}{N} \\
&= \frac{N_{1+}}{N} = \alpha. \qquad (30)
\end{aligned}
$$

By the tower property of conditional expectations, it follows that $E[\hat{\alpha}_c] = E\left[E(\hat{\alpha}_c \mid \boldsymbol{N})\right] = \alpha$. This proves that $\hat{\alpha}_c$ is an unbiased estimator for α.

To compute the variance of $\hat{\alpha}_c$, we use the conditional variance decomposition, again conditioning on the 4-vector $\boldsymbol{N}$. We remark that N_{0+} and N_{1+} are deterministic values, but that N_{+0} and N_{+1} are random variables. As shown above in Equation (30), the conditional expectation is deterministic, hence it has no variance: $V(E[\hat{\alpha}_c \mid \boldsymbol{N}]) = 0$. The conditional variance decomposition then simplifies to the following:

$$V(\hat{\alpha}_c) = E\left[V(\hat{\alpha}_c \mid \boldsymbol{N})\right]. \tag{31}$$

The conditional variance $V(\hat{\alpha}_c \mid \boldsymbol{N})$ can be written as follows:

$$\begin{aligned} V[\hat{\alpha}_c \mid \boldsymbol{N}] &= V\left[\hat{\alpha}^* \frac{n_{11}}{n_{+1}} + (1-\hat{\alpha}^*)\frac{n_{10}}{n_{+0}} \mid \boldsymbol{N}\right] \\ &= (\hat{\alpha}^*)^2 V\left[\frac{n_{11}}{n_{+1}} \mid \boldsymbol{N}\right] + (1-\hat{\alpha}^*)^2 V\left[\frac{n_{10}}{n_{+0}} \mid \boldsymbol{N}\right] \\ &\quad + 2\hat{\alpha}^*(1-\hat{\alpha}^*) C\left[\frac{n_{11}}{n_{+1}}, \frac{n_{10}}{n_{+0}} \mid \boldsymbol{N}\right]. \end{aligned} \tag{32}$$

We will consider these terms separately. First, the variance of n_{11}/n_{+1} can be computed by applying an additional conditional variance decomposition:

$$V\left[\frac{n_{11}}{n_{+1}} \mid \boldsymbol{N}\right] = V\left[E\left(\frac{n_{11}}{n_{+1}} \mid \boldsymbol{N}, n_{+1}\right) \mid \boldsymbol{N}\right] + E\left[V\left(\frac{n_{11}}{n_{+1}} \mid \boldsymbol{N}, n_{+1}\right) \mid \boldsymbol{N}\right].$$

The first term is zero, which can be shown as follows:

$$\begin{aligned} V\left[E\left(\frac{n_{11}}{n_{+1}} \mid \boldsymbol{N}, n_{+1}\right)\right] &= V\left[\frac{1}{n_{+1}} E(n_{11} \mid \boldsymbol{N}, n_{+1}) \mid \boldsymbol{N}\right] \\ &= V\left[\frac{1}{n_{+1}} n_{+1} \frac{N_{11}}{N_{+1}} \mid \boldsymbol{N}\right] \\ &= V\left[\frac{N_{11}}{N_{+1}} \mid \boldsymbol{N}\right] = 0. \end{aligned}$$

For the second term, we find under the assumption that $n \ll N$:

$$\begin{aligned} E\left[V\left(\frac{n_{11}}{n_{+1}} \mid \boldsymbol{N}, n_{+1}\right) \mid \boldsymbol{N}\right] &= E\left[\frac{1}{n_{+1}^2} V(n_{11} \mid \boldsymbol{N}, n_{+1}) \mid \boldsymbol{N}\right] \\ &= E\left[\frac{1}{n_{+1}^2} n_{+1} \frac{N_{11}}{N_{+1}} (1 - \frac{N_{11}}{N_{+1}}) \mid \boldsymbol{N}\right] \\ &= E\left[\frac{1}{n_{+1}} \mid \boldsymbol{N}\right] \frac{N_{11} N_{01}}{N_{+1}^2}. \end{aligned}$$

The expectation of $\frac{1}{n_{+1}}$ can be approximated with a second-order Taylor series:

$$\begin{aligned}
V\left[\frac{n_{11}}{n_{+1}} \mid \boldsymbol{N}\right] &= \left[\frac{1}{E[n_{+1} \mid \boldsymbol{N}]} + \frac{1}{2}\frac{2}{E[n_{+1} \mid \boldsymbol{N}]^3} V\left[n_{+1} \mid \boldsymbol{N}\right]\right] \frac{N_{11}N_{01}}{N_{+1}^2} + O(n^{-3}) \\
&= \frac{1}{E[n_{+1} \mid \boldsymbol{N}]}\left[1 + \frac{V\left[n_{+1} \mid \boldsymbol{N}\right]}{E[n_{+1} \mid \boldsymbol{N}]^2}\right] \frac{N_{11}N_{01}}{N_{+1}^2} + O(n^{-3}) \\
&= \frac{1}{n\hat{\alpha}^*}\left[1 + \frac{1-\hat{\alpha}^*}{n\hat{\alpha}^*}\right] \frac{N_{11}N_{01}}{N_{+1}^2} + O(n^{-3}).
\end{aligned} \tag{33}$$

The variance of n_{10}/n_{+0} can be approximated in the same way, which yields the following expression:

$$V\left[\frac{n_{10}}{n_{+0}} \mid \boldsymbol{N}\right] = \frac{1}{n(1-\hat{\alpha}^*)}\left[1 + \frac{\hat{\alpha}^*}{n(1-\hat{\alpha}^*)}\right] \frac{N_{00}N_{10}}{N_{+0}^2} + O(n^{-3}). \tag{34}$$

Finally, it can be shown that the covariance in the final term is equal to zero:

$$\begin{aligned}
C\left[\frac{n_{11}}{n_{+1}}, \frac{n_{10}}{n_{+0}} \mid \boldsymbol{N}\right] &= E\left[C\left(\frac{n_{11}}{n_{+1}}, \frac{n_{10}}{n_{+0}} \mid \boldsymbol{N}, n_{+0}, n_{+1}\right) \mid \boldsymbol{N}\right] \\
&\quad + C\left[E\left(\frac{n_{11}}{n_{+1}} \mid \boldsymbol{N}, n_{+0}, n_{+1}\right), E\left(\frac{n_{10}}{n_{+0}} \mid \boldsymbol{N}, n_{+0}, n_{+1}\right) \mid \boldsymbol{N}\right] \\
&= E\left[\frac{1}{n_{+0}n_{+1}} C\left(n_{11}, n_{10} \mid \boldsymbol{N}, n_{+0}, n_{+1}\right) \mid \boldsymbol{N}\right] \\
&\quad + C\left[\frac{1}{n_{+1}} E\left(n_{11} \mid \boldsymbol{N}, n_{+0}, n_{+1}\right), \frac{1}{n_{+0}} E\left(n_{10} \mid \boldsymbol{N}, n_{+0}, n_{+1}\right) \mid \boldsymbol{N}\right] \\
&= 0 + C\left[\frac{1}{n_{+1}} n_{+1} \frac{N_{11}}{N_{+1}}, \frac{1}{n_{+0}} n_{+0} \frac{N_{10}}{N_{+0}} \mid \boldsymbol{N}\right] = 0.
\end{aligned} \tag{35}$$

Combining Formulas (33), (34) and (35) with (32) gives:

$$\begin{aligned}
V[\hat{\alpha}_c \mid \boldsymbol{N}] &= \frac{N_{+1}^2}{N^2}\frac{1}{n\hat{\alpha}^*}\left[1 + \frac{1-\hat{\alpha}^*}{n\hat{\alpha}^*}\right] \frac{N_{11}N_{01}}{N_{+1}^2} \\
&\quad + \frac{N_{+0}^2}{N^2}\frac{1}{n(1-\hat{\alpha}^*)}\left[1 + \frac{\hat{\alpha}^*}{n(1-\hat{\alpha}^*)}\right] \frac{N_{00}N_{10}}{N_{+0}^2} + O(n^{-3}) \\
&= \frac{1}{n\hat{\alpha}^*}\left[1 + \frac{1-\hat{\alpha}^*}{n\hat{\alpha}^*}\right] \frac{N_{11}N_{01}}{N^2} \\
&\quad + \frac{1}{n(1-\hat{\alpha}^*)}\left[1 + \frac{\hat{\alpha}^*}{n(1-\hat{\alpha}^*)}\right] \frac{N_{00}N_{10}}{N^2} + O(n^{-3}).
\end{aligned}$$

Recall from Formula (31) that $\mathrm{V}\left[\hat{\alpha}_c\right] = \mathrm{E}\left[V[\hat{\alpha}_c \mid \boldsymbol{N}]\right] = \mathrm{E}\left[\mathrm{E}\left[V[\hat{\alpha}_c \mid \boldsymbol{N}] \mid N_{+1}\right]\right]$. Hence,

$$\begin{aligned}
V[\hat{\alpha}_c] = E\Bigg[&\frac{1}{n\hat{\alpha}^*}\left(1 + \frac{1-\hat{\alpha}^*}{n\hat{\alpha}^*}\right) E\left(\frac{N_{11}N_{01}}{N^2} \mid N_{+1}\right) \\
&+ \frac{1}{n(1-\hat{\alpha}^*)}\left(1 + \frac{\hat{\alpha}^*}{n(1-\hat{\alpha}^*)}\right) E\left(\frac{N_{00}N_{10}}{N^2} \mid N_{+1}\right)\Bigg] + O(n^{-3}).
\end{aligned} \tag{36}$$

To evaluate the expectations in this expression, we observe that, conditional on the column total N_{+1}, N_{11} is distributed as $Bin(N_{+1}, c_{11})$, where c_{11} is a calibration probability as defined in Section 2.5. Hence,

$$\begin{aligned} E\left[N_{11} \mid N_{+1}\right] &= N_{+1}c_{11} = \frac{N_{+1}\alpha p_{11}}{(1-\alpha)(1-p_{00})+\alpha p_{11}} \\ V\left[N_{11} \mid N_{+1}\right] &= N_{+1}c_{11}(1-c_{11}). \end{aligned} \tag{37}$$

Similarly, since $N = N_{+1} + N_{+0}$ is fixed,

$$\begin{aligned} E\left[N_{00} \mid N_{+1}\right] &= N_{+0}c_{00} = \frac{N_{+0}(1-\alpha)p_{00}}{(1-\alpha)p_{00}+\alpha(1-p_{11})} \\ V\left[N_{00} \mid N_{+1}\right] &= N_{+0}c_{00}(1-c_{00}). \end{aligned} \tag{38}$$

Using these results, we obtain:

$$\begin{aligned} E\left[\frac{N_{11}N_{01}}{N^2} \mid N_{+1}\right] &= \frac{1}{N^2}E\left[N_{11}N_{01} \mid N_{+1}\right] \\ &= \frac{1}{N^2}E\left[N_{11}(N_{+1}-N_{11}) \mid N_{+1}\right] \\ &= \frac{1}{N^2}\left[N_{+1}E\left[N_{11} \mid N_{+1}\right] - E\left[N_{11}^2 \mid N_{+1}\right]\right] \\ &= \frac{1}{N^2}\left[N_{+1}E\left[N_{11} \mid N_{+1}\right] - V\left[N_{11} \mid N_{+1}\right] - E\left[N_{11} \mid N_{+1}\right]^2\right] \\ &= \frac{1}{N^2}\left[N_{+1}^2c_{11} - N_{+1}c_{11}(1-c_{11}) - N_{+1}^2c_{11}^2\right] \\ &= \frac{N_{+1}^2}{N^2}c_{11}(1-c_{11}) + O\left(\frac{1}{N}\right), \end{aligned} \tag{39}$$

and similarly

$$E\left[\frac{N_{00}N_{10}}{N^2} \mid N_{+1}\right] = \frac{N_{+0}^2}{N^2}c_{00}(1-c_{00}) + O\left(\frac{1}{N}\right). \tag{40}$$

Substituting expressions (39) and (40) into (36) and noting that $N_{+1}^2/N^2 = (\hat{\alpha}^*)^2$ and $N_{+0}^2/N^2 = (1-\hat{\alpha}^*)^2$, we obtain:

$$\begin{aligned} V[\hat{\alpha}_c] = E\Bigg[&\frac{\hat{\alpha}^*}{n}\left(1+\frac{1-\hat{\alpha}^*}{n\hat{\alpha}^*}\right)c_{11}(1-c_{11}) \\ &+\frac{1-\hat{\alpha}^*}{n}\left(1+\frac{\hat{\alpha}^*}{n(1-\hat{\alpha}^*)}\right)c_{00}(1-c_{00})\Bigg] + O\left(\max\left[\frac{1}{n^3}, \frac{1}{Nn}\right]\right) \\ = &\left[\frac{E(\hat{\alpha}^*)}{n} + \frac{1-E(\hat{\alpha}^*)}{n^2}\right]c_{11}(1-c_{11}) \\ &+\left[\frac{1-E(\hat{\alpha}^*)}{n} + \frac{E(\hat{\alpha}^*)}{n^2}\right]c_{00}(1-c_{00}) + O\left(\max\left[\frac{1}{n^3}, \frac{1}{Nn}\right]\right). \end{aligned}$$

Finally, substituting the expressions for $E(\hat{\alpha}^*)$ from (4) and the expressions for c_{11} and c_{00} from (37) and (38), the desired expression (17) is obtained. This concludes the proof of Theorem 3.

Comparing mean squared errors

To conclude, we present the proof of Theorem 4, which essentially shows that the mean squared error (up to and including terms of order $1/n$) of the calibration estimator is lower than that of the misclassification estimator.

Proof (of Theorem 4). Recall that the bias of $\hat{\alpha}_p$ as an estimator for α is given by

$$B\left[\hat{\alpha}_p\right] = \frac{p_{00} - p_{11}}{n(p_{00} + p_{11} - 1)} + O\left(\frac{1}{n^2}\right).$$

Hence, $(B\left[\hat{\alpha}_p\right])^2 = O(1/n^2)$ is not relevant for $\widetilde{MSE}[\hat{\alpha}_p]$. It follows that $\widetilde{MSE}[\hat{\alpha}_p]$ is equal to the variance of $\hat{\alpha}_p$ up to order $1/n$. From (14) we obtain:

$$\widetilde{MSE}[\hat{\alpha}_p] = \frac{1}{n}\left[\frac{(1-\alpha)p_{00}(1-p_{00}) + \alpha p_{11}(1-p_{11})}{(p_{00}+p_{11}-1)^2}\right]. \tag{41}$$

Recall that $\hat{\alpha}_c$ is an unbiased estimator for α, i.e., $B[\hat{\alpha}_c] = 0$. Also recall the notation $\beta = (1-\alpha)(1-p_{00}) + \alpha p_{11}$. It follows from (17) that the variance, and hence the MSE, of $\hat{\alpha}_c$ up to terms of order $1/n$ can be written as:

$$\begin{aligned}
\widetilde{MSE}[\hat{\alpha}_c] &= \frac{1}{n}\left[\beta\frac{\alpha p_{11}}{\beta}\left(1 - \frac{\alpha p_{11}}{\beta}\right) + (1-\beta)\frac{(1-\alpha)p_{00}}{1-\beta}\left(1 - \frac{(1-\alpha)p_{00}}{1-\beta}\right)\right] \\
&= \frac{\alpha(1-\alpha)}{n}\left[\frac{(1-p_{00})p_{11}}{\beta} + \frac{p_{00}(1-p_{11})}{1-\beta}\right].
\end{aligned} \tag{42}$$

To prove Expression (18), first note that

$$\frac{(1-p_{00})p_{11}}{\beta} + \frac{p_{00}(1-p_{11})}{1-\beta} = \frac{(1-p_{00})p_{11} + \beta(p_{00}-p_{11})}{\beta(1-\beta)}. \tag{43}$$

The numerator of this equation can be rewritten as follows:

$$\begin{aligned}
&(1-p_{00})p_{11} + \beta(p_{00}-p_{11}) \\
&= (1-p_{00})p_{11} + (1-\alpha)p_{00}(1-p_{00}) + \alpha p_{00}p_{11} - (1-\alpha)(1-p_{00})p_{11} - \alpha p_{11}^2 \\
&= (1-\alpha)p_{00}(1-p_{00}) + \alpha p_{00}p_{11} + \alpha(1-p_{00})p_{11} - \alpha p_{11}^2 \\
&= (1-\alpha)p_{00}(1-p_{00}) + \alpha p_{11}(1-p_{11}).
\end{aligned}$$

Note that the obtained expression is equal to the numerator of Expression (41). Write $T = (1-\alpha)p_{00}(1-p_{00}) + \alpha p_{11}(1-p_{11})$ for that expression. It follows that

$$\begin{aligned}
&\widetilde{MSE}[\hat{\alpha}_p] - \widetilde{MSE}[\hat{\alpha}_c] \\
&= \frac{T}{n(p_{00}+p_{11}-1)^2} - \frac{T\alpha(1-\alpha)}{n\beta(1-\beta)} \\
&= \frac{T}{n(p_{00}+p_{11}-1)^2\beta(1-\beta)}\Big[\beta(1-\beta) - \alpha(1-\alpha)(p_{00}+p_{11}-1)^2\Big].
\end{aligned}$$

Writing out the second factor in the last expression gives the following:

$$
\begin{aligned}
&\beta(1-\beta) - \alpha(1-\alpha)(p_{00}+p_{11}-1)^2 \\
&= (1-\alpha)^2 p_{00}(1-p_{00}) + \alpha(1-\alpha)\Big((1-p_{00})(1-p_{11}) + p_{00}p_{11}\Big) + \alpha^2 p_{11}(1-p_{11}) \\
&\qquad - \alpha(1-\alpha)(p_{00}+p_{11}-1)^2 \\
&= (1-\alpha)^2 p_{00}(1-p_{00}) + \alpha(1-\alpha)\Big(p_{00}(1-p_{00}) + p_{11}(1-p_{11})\Big) + \alpha^2 p_{11}(1-p_{11}) \\
&= (1-\alpha)p_{00}(1-p_{00}) + \alpha p_{11}(1-p_{11}) \\
&= T.
\end{aligned}
$$

This concludes the proof of Theorem 4.