

If we consider the score with the constant factor  $\frac{1}{p}$  we have

$$\begin{aligned}\psi(W; \theta, \eta) &:= \frac{D(Y - g(0, X))}{p} - \frac{m(X)(1 - D)(Y - g(0, X))}{p(1 - m(x))} - \frac{D}{p}\theta \\ &= \psi_a(W; \eta)\theta + \psi_b(W; \eta)\end{aligned}$$

with  $\eta = (g, m, p)$  and where the components of the linear score are

$$\begin{aligned}\psi_a(W; \eta) &= -\frac{D}{p}, \\ \psi_b(W; \eta) &= \frac{D(Y - g(0, X))}{p} - \frac{m(X)(1 - D)(Y - g(0, X))}{p(1 - m(X))}.\end{aligned}$$

If the score is linear in the parameter, we get for the asymptotic variance

$$\sigma^2 = J_0^{-2} \mathbb{E}(\psi^2(W; \theta_0, \eta_0)),$$

with  $J_0 = \mathbb{E}(\psi_a(W; \eta_0))$ . For the ATTE score we get

$$J_0 = \mathbb{E}(\psi_a(W; \eta_0)) = \mathbb{E}\left(-\frac{D}{p_0}\right) = -\frac{\mathbb{E}(D)}{p_0} = -\frac{p_0}{p_0} = -1.$$

Therefore, the variance simplifies to  $\sigma^2 = \mathbb{E}(\psi^2(W; \theta_0, \eta_0))$ . This is also mentioned in Chernozhukov et al. (2018) on page C35.

If we instead consider the score without the constant factor  $\frac{1}{p}$  we get

$$\begin{aligned}\tilde{\psi}(W; \theta, \eta) &:= D(Y - g(0, X)) - \frac{m(X)(1 - D)(Y - g(0, X))}{(1 - m(x))} - D\theta \\ &= \tilde{\psi}_a(W; \eta)\theta + \tilde{\psi}_b(W; \eta)\end{aligned}$$

with  $\eta = (g, m)$  and where the components of the linear score are

$$\begin{aligned}\psi_a(W; \eta) &= -D, \\ \psi_b(W; \eta) &= D(Y - g(0, X)) - \frac{m(X)(1 - D)(Y - g(0, X))}{(1 - m(X))}.\end{aligned}$$

The two score functions are related via  $\psi(W; \theta, \eta) = \frac{1}{p}\tilde{\psi}(W; \theta, \eta)$ . We further get

$$\tilde{J}_0 = \mathbb{E}(\tilde{\psi}_a(W; \eta_0)) = \mathbb{E}(-D) = -p_0.$$

For the variance we therefore get

$$\tilde{\sigma}^2 = \tilde{J}_0^{-2} \mathbb{E}(\tilde{\psi}^2(W; \theta_0, \eta_0)) = \frac{\mathbb{E}(p_0^2 \psi^2(W; \theta_0, \eta_0))}{p_0^2} = \mathbb{E}(\psi^2(W; \theta_0, \eta_0)) = \sigma^2.$$

$\implies$  So it basically does not matter whether one adds the constant factor  $\frac{1}{p}$ , but with the constant the formula of the variance simplifies to  $\sigma^2 = \mathbb{E}(\psi^2(W; \theta_0, \eta_0))$ .

## References

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W. and Robins, J. (2018), Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21: C1-C68.