# 1 Basic difference-in-differences example

Consider a simple difference-in-differences design with $m$ clusters and $n = 2$ time periods. Suppose that the first $m_0$ clusters remain untreated in the second time period and the remaining $m_1 = m - m_0$ clusters are treated in the second time period. The basic difference-in-differences model for this design is then

$$y_{it} = \alpha_i + \beta_t + \delta T_{it} + e_{it}, \tag{1}$$

where $T_{i1} = 1$ for $i = m_0 + 1, ..., m$, $T_{it} = 0$ otherwise, and $\delta$ is the average treatment effect.

Estimating $\delta$ by OLS is exactly equivalent to taking first differences and then calculating the mean difference between treated and untreated clusters. Let $d_i = y_{i1} - y_{i0}$ for $i = 1, ..., m$, $\bar{d}_0 = \sum_{i=1}^{m_0} d_i / m_0$, and $\bar{d}_1 = \sum_{i=m_0+1}^{m} d_i / m_1$. Then $\hat{\delta} = \bar{d}_1 - \bar{d}_0$. In this simplified representation of the model, it is clear that the null hypothesis $\delta = 0$ may be tested using a simple two-sample t-test on the difference scores, while allowing for unequal variances. The sampling variance of $\hat{\delta}$ can be estimated from the difference scores as

$$V_\Delta = \frac{1}{m_0(m_0-1)} \sum_{i=1}^{m_0} \left(d_i - \bar{d}_0\right)^2 + \frac{1}{m_1(m_1-1)} \sum_{i=m_0+1}^{m} \left(d_i - \bar{d}_1\right)^2.$$

Under a "working homoskedasticity" model, the degrees of freedom corresponding to $V_\Delta$ are

$$\nu_\Delta = \frac{m^2(m_0-1)(m_1-1)}{m_0^2(m_0-1) + m_1^2(m_1-1)}$$

(**?**).

We shall now consider the variance estimator and degrees of freedom generated by the CR2 correction as applied to the full difference-in-differences model (1), while estimating $\delta$ after absorbing the cluster- and period-specific effects. We use the "working independence" model for deriving the CR2 adjustment matrices and degrees of freedom. Following the notation of the main paper, this design has

$$\mathbf{R}_i = \left[\begin{array}{c} 0 \\ T_{i1} \end{array}\right] \qquad \mathbf{S}_i = \left[\begin{array}{c} 0 \\ 1 \end{array}\right] \qquad \mathbf{T}_i = \left[\begin{array}{c} 1 \\ 1 \end{array}\right] \left[\begin{array}{cccc} I(i=1) & I(i=2) & \cdots & I(i=m) \end{array}\right]$$

After absorption, $\ddot{\mathbf{R}}_i = \left(T_{i1} - m_1/m\right)/2 \left[\begin{array}{c} -1 \\ 1 \end{array}\right]'$, $\mathbf{M}_{\ddot{\mathbf{R}}} = 2m/(m_0 m_1)$, and

$$\mathbf{e}_i = \frac{d_i - \bar{d}_0}{2} \left[\begin{array}{c} 1 \\ 1 \end{array}\right] \quad \text{for} \quad i = 1, ..., m_0, \qquad \mathbf{e}_i = \frac{d_i - \bar{d}_1}{2} \left[\begin{array}{c} 1 \\ 1 \end{array}\right], \quad \text{for} \quad i = m_0 + 1, ..., m.$$

If the CR2 adjustment matrices are calculated based on the absorbed model only, then

$$\mathbf{A}_i = \left(\mathbf{I}_i - \ddot{\mathbf{R}}_i \mathbf{M}_{\ddot{\mathbf{R}}} \ddot{\mathbf{R}}_i'\right)^{+1/2} = \left[\begin{array}{cc} 1 + a_i & -a_i \\ -a_i & 1 + a_i \end{array}\right],$$

where

$$\begin{aligned} a_i &= \frac{1}{2}\left(\sqrt{\frac{m_0 m}{m_0 m - m_1}} - 1\right) \qquad & i = 1, ..., m_0 \\ a_i &= \frac{1}{2}\left(\sqrt{\frac{m_1 m}{m_1 m - m_0}} - 1\right) \qquad & i = m_0 + 1, ..., m. \end{aligned}$$

Using these adjustment matrices yields the variance estimator

$$V_{\ddot{\mathbf{R}}} = \frac{1}{m_0(m_0 - m_1/m)} \sum_{i=1}^{m_0} \left(d_i - \bar{d}_0\right)^2 + \frac{1}{m_1(m_1 - m_0/m)} \sum_{i=m_0+1}^{m} \left(d_i - \bar{d}_1\right)^2,$$

which will be slightly smaller than $V_\Delta$, with Satterthwaite degrees of freedom

$$\nu_{\ddot{\mathbf{R}}} = \frac{\left(\dfrac{m_0 - 1}{m_0(m_0 - m_1/m)} + \dfrac{m_1 - 1}{m_1(m_1 - m_0/m)}\right)^2}{\dfrac{1}{m_0(m_0 - m_1/m)} + \dfrac{1}{m_1(m_1 - m_0/m)}},$$

which will be slightly larger than $\nu_\Delta$.

Now consider calculating the adjustment matrices using the full design matrix, as recommended in the paper. Theorem 2 implies that the adjustment matrices can be calculated from $\ddot{\mathbf{U}}$, ignoring the cluster-specific effects. We then have

$$\mathbf{A}_i = \left(\mathbf{I}_i - \ddot{\mathbf{U}}_i \mathbf{M}_{\ddot{\mathbf{U}}} \ddot{\mathbf{U}}_i'\right)^{+1/2} = \begin{bmatrix} 1 + b_i & -b_i \\ -b_i & 1 + b_i \end{bmatrix},$$

where

$$\begin{aligned} b_i &= \frac{1}{2}\left(\sqrt{\frac{m_0}{m_0 - 1}} - 1\right) \qquad & i = 1, ..., m_0 \\ b_i &= \frac{1}{2}\left(\sqrt{\frac{m_1}{m_1 - 1}} - 1\right) \qquad & i = m_0 + 1, ..., m. \end{aligned}$$

It can be verified that using these adjustment matrices yields a variance estimator that is exactly equivalent to $V_\Delta$, with degrees of freedom equal to $\nu_\Delta$.