

Online Appendix for: Difference-in-Differences with Geocoded Microdata

Kyle Butts*

October 12, 2021

1 – Cross-Sectional Data

In the case of cross sectional data, we have individuals i that appear in the data in period $t(i) \in \{0, 1\}$. However, since we no longer are able to observe units in both periods, we are not able to take first differences. Our model therefore will have a λ term for both periods. Therefore, λ includes the average of μ_i , covariates, and period shocks at a given distance.

$$Y_i = \tau(\text{Dist}_i)\mathbf{1}_{t(i)=1} + \lambda_{t(i)}(\text{Dist}_i) + \nu_{it}. \quad (1)$$

The parallel trends assumption must be modified now in the case of cross sections:

Assumption 1 (Local Parallel Trends (RC)). For a distance \bar{d} , we say that ‘local parallel trends’ hold if for all positive $d, d' \leq \bar{d}$, then $\lambda_1(d) - \lambda_0(d) = \lambda_1(d') - \lambda_0(d')$.

Assumption 2 (Average Parallel Trends (RC)). For a pair of distances d_t and d_c , we say that ‘average parallel trends’ hold if $\mathbb{E}[\lambda_1(d) \mid 0 \leq d \leq d_t] - \mathbb{E}[\lambda_0(d) \mid 0 \leq d \leq d_t] = \mathbb{E}[\lambda_1(d) \mid d_t < d \leq d_c] - \mathbb{E}[\lambda_0(d) \mid d_t < d \leq d_c]$.

The parallel trends assumption is a bit more complicated now and is theoretically more strict. **Local Parallel Trends (RC)** still require changes in outcomes over time for a given unit i must be constant across distance (or on average in the case of **Average Parallel Trends (RC)**). However since the composition of units can change over time, this also requires that the average of individual fixed effects must be constant across time. This is well understood in the hedonic pricing literature that the composition of homes being sold can not change over time for identification (e.g. [Linden and Rockoff \(2008\)](#)).

For completeness, I rewrite the other necessary assumption

*University of Colorado, Boulder. Email: kyle.butts@colorado.edu.

Assumption 3 (Correct d_t). A distance d_t satisfies this assumption if (i) for all $d \leq d_t$, $\tau(d) > 0$ and for all $d > d_t$, $\tau(d) = 0$ and (ii) $F(d_c) - F(d_t) > 0$.

The ring estimate in the case of cross-sections is given by:

$$\Delta Y_i = \beta_0 + \beta_1 \mathbf{1}_{i \in \mathcal{D}_c} \mathbf{1}_{t(i)=1} + \beta_2 \mathbf{1}_{i \in \mathcal{D}_t} \mathbf{1}_{t(i)=0} + \beta_4 \mathbf{1}_{i \in \mathcal{D}_t} \mathbf{1}_{t(i)=1} + u_{it}. \quad (2)$$

Proposition 1 (Decomposition of Ring Estimate (RC)). Given that units follow model (1),

(i) The estimate of β_4 in (2) has the following expectation:

$$\begin{aligned} \mathbb{E} [\hat{\beta}_4] &= \mathbb{E} [\Delta Y_{it} \mid \mathcal{D}_t] - \mathbb{E} [\Delta Y_{it} \mid \mathcal{D}_c] \\ &= \underbrace{\mathbb{E} [\tau(\text{Dist}) \mid \mathcal{D}_t] - \mathbb{E} [\tau(\text{Dist}) \mid \mathcal{D}_c]}_{\text{Difference in Treatment Effect}} \\ &\quad + (\mathbb{E} [\lambda_1(\text{Dist}) \mid \mathcal{D}_t, t(i) = 1] - \mathbb{E} [\lambda_0(\text{Dist}) \mid \mathcal{D}_t, t(i) = 0]) \\ &\quad - (\mathbb{E} [\lambda_1(\text{Dist}) \mid \mathcal{D}_c, t(i) = 1] - \mathbb{E} [\lambda_0(\text{Dist}) \mid \mathcal{D}_c, t(i) = 0]) \end{aligned}$$

(ii) If d_c satisfies **Local Parallel Trends (RC)** or, more weakly, if d_t and d_c satisfy **Average Parallel Trends (RC)**, then

$$\mathbb{E} [\hat{\beta}_4] = \underbrace{\mathbb{E} [\tau(\text{Dist}) \mid \mathcal{D}_t] - \mathbb{E} [\tau(\text{Dist}) \mid \mathcal{D}_c]}_{\text{Difference in Treatment Effect}}.$$

(iii) If d_c satisfies **Local Parallel Trends (RC)** and d_t satisfies **Assumption 3**, then

$$\mathbb{E} [\hat{\beta}_4] = \bar{\tau}.$$

Proof. With some algebraic manipulation, we can rewrite our difference-in-differences estimator as

$$\begin{aligned} &(\mathbb{E} [Y_i \mid \mathcal{D}_t, t(i) = 1] - \mathbb{E} [Y_i \mid \mathcal{D}_t, t(i) = 0]) - (\mathbb{E} [Y_i \mid \mathcal{D}_c, t(i) = 1] - \mathbb{E} [Y_i \mid \mathcal{D}_c, t(i) = 0]) \\ &= (\mathbb{E} [\tau(\text{Dist}_i) + \lambda_1(\text{Dist}_i) + \nu_{i1} \mid \mathcal{D}_t, t(i) = 1] - \mathbb{E} [\lambda_0(\text{Dist}_i) + \nu_{i0} \mid \mathcal{D}_t, t(i) = 0]) \\ &\quad - (\mathbb{E} [\tau(\text{Dist}_i) + \lambda_1(\text{Dist}_i) + \nu_{i1} \mid \mathcal{D}_c, t(i) = 1] - \mathbb{E} [\lambda_0(\text{Dist}_i) + \nu_{i0} \mid \mathcal{D}_c, t(i) = 0]) \\ &= \mathbb{E} [\tau(\text{Dist}_i) \mid \mathcal{D}_t, t(i) = 1] - \mathbb{E} [\tau(\text{Dist}_i) \mid \mathcal{D}_c, t(i) = 0] + \\ &\quad (\mathbb{E} [\lambda_1(\text{Dist}_i) \mid \mathcal{D}_t, t(i) = 1] - \mathbb{E} [\lambda_0(\text{Dist}_i) \mid \mathcal{D}_t, t(i) = 0]) \\ &\quad - (\mathbb{E} [\lambda_1(\text{Dist}_i) \mid \mathcal{D}_c, t(i) = 1] - \mathbb{E} [\lambda_0(\text{Dist}_i) \mid \mathcal{D}_c, t(i) = 0]), \end{aligned}$$

where the terms consisting of ν_{it} cancel out as they are uncorrelated with distance. Propositions (ii) and (iii) follow the same arguments as in the panel case. \square

Part (i) of this theorem shows that under no parallel trends assumption, the ‘Difference in Trends’ term becomes the change in λ_t for the treated ring minus the change in for the control ring. As discussed above, this change in lambdas can be do to period specific shocks or changes in the composition of units observed in each period.

1.1. Nonparametric Estimation

Since we can no longer perform a single nonparametric regression on first differences in the context of cross-sections, our nonparametric estimator must be adjusted. The modified procedure will fit a nonparametric estimate of $\mathbb{E}[Y_i | \text{Dist}_i, t]$ separately for $t = 0$ and $t = 1$ with a restriction that the bin intervals $\{\mathcal{D}_1, \dots, \mathcal{D}_L\}$ must be the same in both samples.¹ Then, for each distance bin we calculate an estimate of $\bar{Y}_{j,t}$ which corresponds to the sample average of observations in period t in bin \mathcal{D}_j .

Then estimates of τ_j can be formed as

$$\hat{\tau}_j = [\bar{Y}_{j,1} - \bar{Y}_{j,0}] - [\bar{Y}_{L,1} - \bar{Y}_{L,0}] ,$$

where, as before, the change in trends in the last ring serve as an estimate for the counterfactual trend. Under **Local Parallel Trends (RC)**, estimates of $\hat{\tau}_{au_j}$ are consistent for $\mathbb{E}[\tau(\text{Dist}_i) | i \in \mathcal{D}_j]$ and the treatment effect curve converges uniformly to the treatment effect curve $\tau(d)$.

Standard errors are formed similarly as before, but is the difference of four means so they can be formed as $\sqrt{\sigma_{j,1}^2 + \sigma_{j,0}^2 + \sigma_{L,1}^2 + \sigma_{L,0}^2}$. These individual estimates and standard errors can be produced by the Stata/R package `binsreg`.

1. The number of intervals are decided based on a different IMSE condition described in Cattaneo et al. (2019) and the quantiles are calculated using the distribution of distances in both periods.

References

- Cattaneo, Matias D., Richard K. Crump, Max H. Farrell, and Yingjie Feng. 2019. "On Binscatter." *arXiv:1902.09608 [econ, stat]*, <http://arxiv.org/abs/1902.09608>, arXiv: 1902.09608.
- Linden, Leigh, and Jonah E Rockoff. 2008. "Estimates of the Impact of Crime Risk on Property Values from Megan's Laws." *American Economic Review* 98 (3): 1103–1127. [10.1257/aer.98.3.1103](https://doi.org/10.1257/aer.98.3.1103).