Multi-cutoff RD designs with observations located at each cutoff: solutions for practitioners

Margherita Fort† Andrea Ichino‡ Enrico Rettore¶ Giulio Zanella§

- ONLINE APPENDIX

January 14, 2022

Contents

List of Figures

List of Tables

†University of Bologna, IZA, and CESifo; margherita.fort@unibo.it.
‡European University Institute, University of Bologna, CEPR, CESifo, and IZA; andrea.ichino@eui.eu.
§University of Adelaide, University of Bologna, and IZA; giulio.zanella@adelaide.edu.au.
¶University of Trento, FBK-IRVAPP, and IZA; enrico.rettore@unitn.it.

Appendix to Section 4
The problematic multi-cutoff design

Heterogeneity of sites and sample size

Here we show what happens to the across-site variability of the site weights as the total sample size decreases if we impose the condition that the number of sites is fixed.

Following the notation in the main text, let J be the number of sites, N_j the number of applicants in site j, $\overline{N}$ the average number of applicants per site and $N = J\overline{N}$ the total number of applicants. Suppose we impose the condition that in each site there must be at least μ units. Then, the number of units in site j is:

$$N_j = \mu + (N - J\mu)p_j, \tag{A–1}$$

where p_j is the fraction of units assigned to site j out of those left after assigning μ units to each site.

The across-site variance of N_j is:

$$\mathrm{var}[N_j] = J^2(\overline{N} - \mu)^2\mathrm{var}[p_j], \tag{A–2}$$

while the across-site variance of the site weights is:

$$\mathrm{var}[\frac{N_j}{N}] = (1 - \frac{\mu}{\overline{N}})^2\mathrm{var}[p_j]. \tag{A–3}$$

As $\overline{N}$ decreases to μ the across-site variance of site weights decreases to zero.

Appendix to Section 5
Fixed effects: the safest option

A sufficient condition for the efficiency of SFE

Here we derive the condition (43) in Section 5.2. Under Allocation rule 1 the within-site deviation of the treatment status D_{ij} is:

$$SS_w = \sum_{j=1}^{J} n_j D_{.j}(1 - D_{.j}), \tag{A–4}$$

n_j the number of units in a suitable neighbourhood of the cutoff, $D_{.j}$ the proportion of exposed units among the n_j ones. Due to sampling variability:

$$D_{.j} = 0.5 + e_{.j}, \tag{A–5}$$

$e_{.j}$ the sampling error. As a result:

$$SS_w = 0.25n - \sum_{j=1}^{J} n_j e_{.j}^2, \tag{A–6}$$

where $n = \sum_{j=1}^{J} n_j$. The between-site deviation of D_{ij} is:

$$SS_b = \sum_{j=1}^{J} n_j (D_{.j} - D_{..})^2, \tag{A–7}$$

where $D_{..} = 0.5 + e_{..}$ is the overall proportion of treated units. With large n, the sampling error of $D_{..}$, $e_{..}$, is negligible. Then:

$$SS_b = \sum_{j=1}^{J} n_j e_{.j}^2. \tag{A–8}$$

As a result, the intraclass correlation of the treatment status is:

$$\frac{SS_b}{SS_b + SS_w} = \frac{\sum_{j=1}^{J} n_j e_{.j}^2}{0.25n} = \frac{\sum_{j=1}^{J} n_j Var[e_{.j}]}{0.25n}, \tag{A–9}$$

the last step holding when J is large. In sites with at least one treated and one untreated unit:

$$Var[e_{.j}] = \frac{0.25}{n_j}(1 - \frac{2}{n_j}), \tag{A–10}$$

implying that:

$$\frac{SS_b}{SS_b + SS_w} = \frac{0.25J - 0.5\sum_{j=1}^{J}\frac{1}{n_j}}{0.25n} = \frac{1}{\bar{n}}(1 - 2\frac{\sum_{j=1}^{J}\frac{1}{n_j}}{J}). \tag{A–11}$$

SFE is more precise than NP as long as the intraclass correlation of the composite error term is larger than the RHS of (A-11). Recalling that:

$$\frac{\sum_{j=1}^{J}\frac{1}{n_j}}{J} \geq \frac{1}{\bar{n}}, \tag{A–12}$$

a more straightforward sufficient condition is:

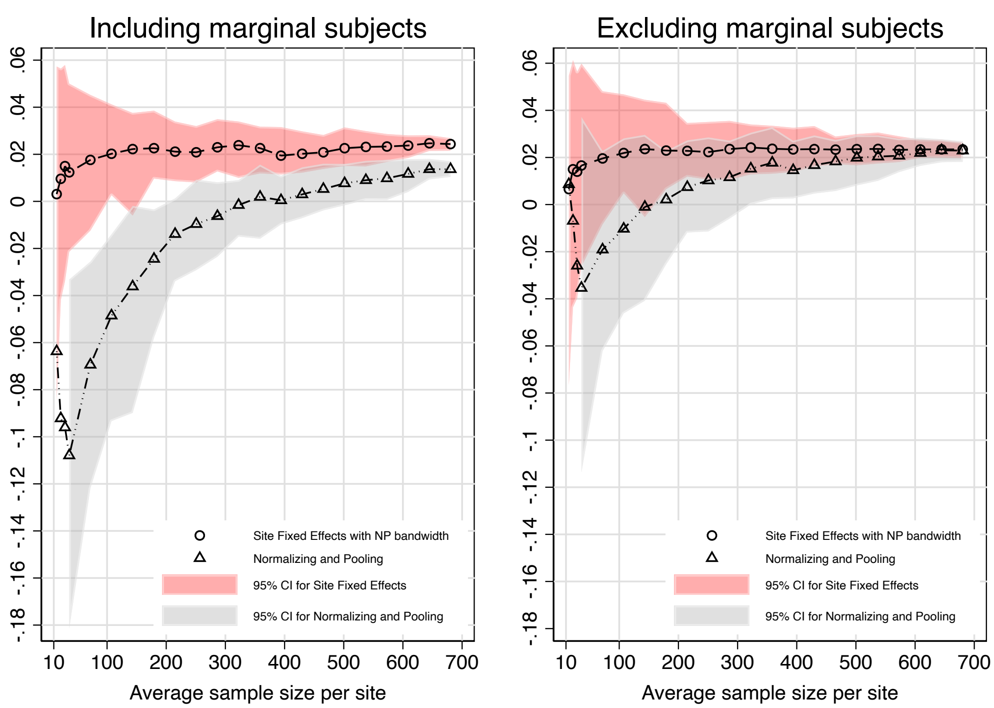
$$\frac{Var[u]}{Var[u] + Var[\epsilon]} > \frac{1}{\bar{n}}(1 - 2\frac{1}{\bar{n}}). \tag{A–13}$$

The largest value of the RHS is at $\bar{n} = 4$, as large as 0.125.

Appendix to Section 8
An empirical illustration

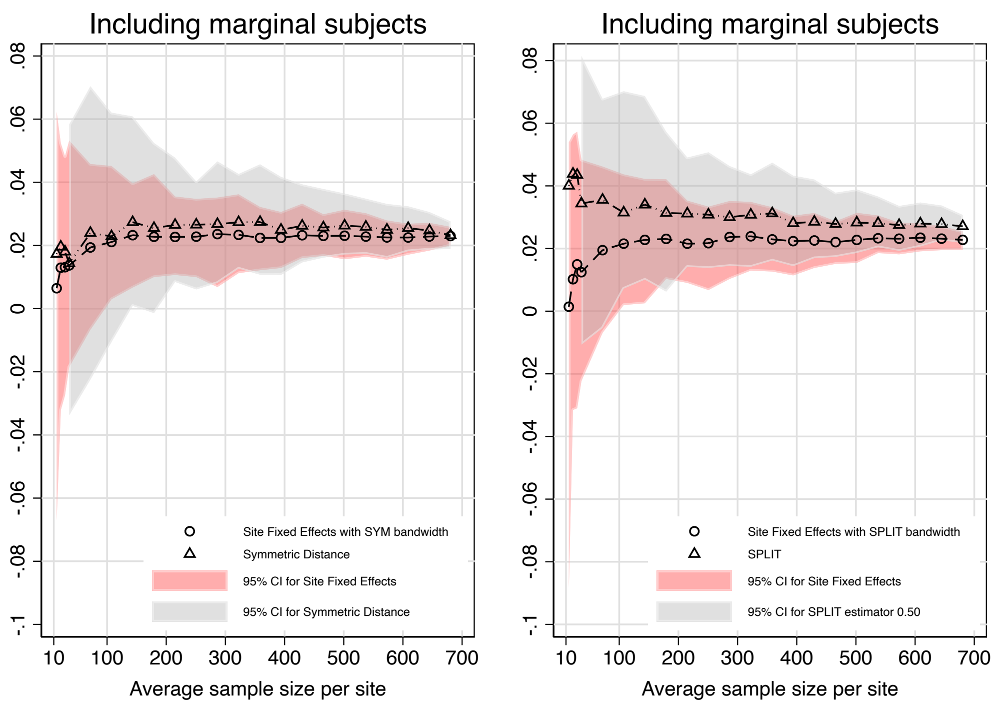
We report below figures that compare the estimators in the reduced samples, in which the 95% confidence intervals have been obtained with the conventional asymptotic approximation. The analogous figures of the main text display instead 95% empirical confidence intervals, computed as the corresponding appropriate percentiles of the distribution of 100 bootstrapped estimates with replacement. The two approaches give similar results.

Figure A–1: Normalizing and Pooling (NP) and corresponding Fixed Effects (FE) estimates



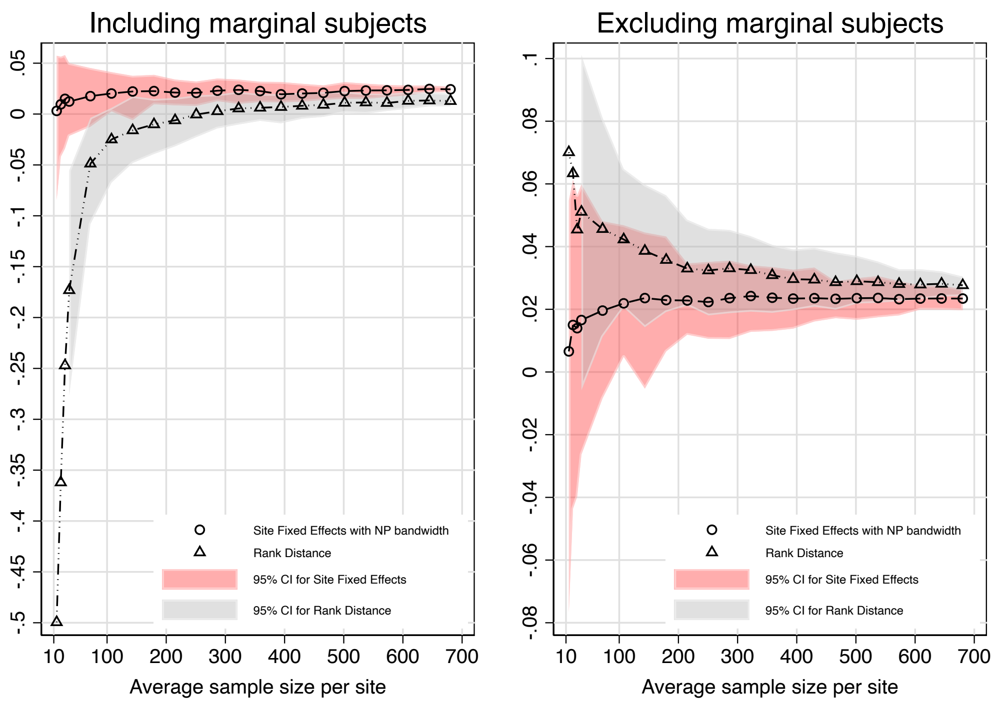
Notes: This figure displays the NP estimates (triangles) and the corresponding FE estimates (circles) in the reduced samples. All estimates are non-parametric, sharp RD. In the left panel marginal subjects (i.e., subjects located exactly at the cutoff), are included, while in the right panel they are dropped. The NP estimators are the sample analog of the Normalizing and Pooling estimands described in Section 3 of the main text, depending on whether the marginal subjects are included or not. The FE estimators are the sample analog of the Fixed Effect estimand described in Section 5. The reduced samples have been obtained from the Full Sample constructed with the original data of Pop-Eleches and Urquiola (2013), with the procedure described in Section 8. The RD estimates have been obtained with Local Linear Regressions using a triangular kernel and the optimal bandwidth from Calonico et al. (2014). The optimal bandwith for the FE estimator is the same as the optimal bandwith for the corresponding NP estimator. For each estimator, the shaded areas describe the 95% empirical confidence intervals of the estimates, obtained with the conventional asymptotic approximation.

Figure A–2: Estimates based on redefining the cutoff and corresponding Fixed Effects (FE) estimates



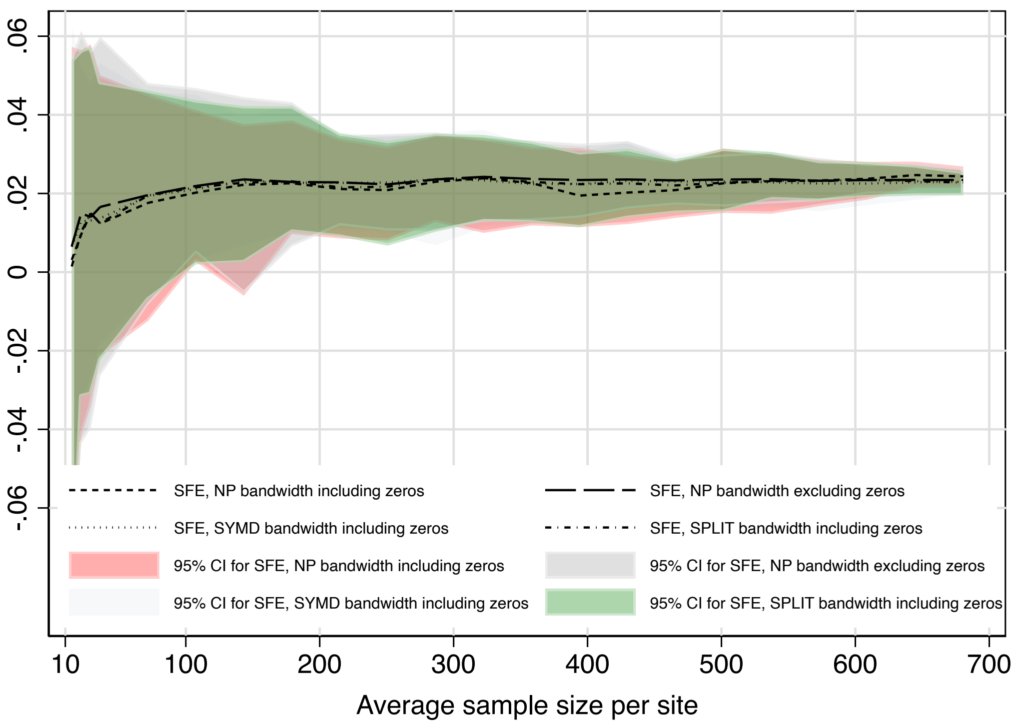
Notes: The left panel of this figure displays the SYM estimates (triangles) and the corresponding FE estimates (circles) in the reduced samples. The right panel displays instead the SPLIT estimates (triangles) and the corresponding FE estimates (circles) in the reduced samples. All estimates are non-parametric, sharp RD. The SYM and the SPLIT estimators are the sample analogues of the estimands described in Section 6. The FE estimators are the sample analog of the Fixed Effect estimand described in Section 5. In both panels marginal subjects (i.e., subjects located exactly at the cutoff), are included. The reduced samples have been obtained from the Full Sample constructed with the original data of Pop-Eleches and Urquiola (2013), with the procedure described in Section 8. The RD estimates have been obtained with Local Linear Regressions using a triangular kernel and the optimal bandwidth from Calonico et al. (2014). The optimal bandwith for the FE estimator is the same as the optimal bandwith for the corresponding SYM or SPLIT estimator. For each estimator, the shaded areas describe the 95% empirical confidence intervals of the estimates, obtained with the conventional asymptotic approximation.

Figure A–3: Rank distance (RK) and corresponding Fixed Effects (FE) estimates



Notes: This figure displays the RK estimates (triangles) and the FE estimates (circles) of the corresponding NP estimator (see Figure A–1) in the reduced samples. In the left panel marginal subjects (i.e., subjects located exactly at the cutoff), are included, while in the right panel they are dropped. All estimates are non-parametric, sharp RD. The RK estimators are the sample analog of the Rank Distance estimands described in Section 7. The FE estimators are the sample analog of the Fixed Effect estimands described in Section 5. The reduced samples have been obtained from the Full Sample constructed with the original data of Pop-Eleches and Urquiola (2013), with the procedure described in Section 8. The RD estimates have been obtained with Local Linear Regressions using a triangular kernel and the optimal bandwidth from Calonico et al. (2014). The optimal bandwith for the FE estimator is the same as the optimal bandwith for the corresponding NP estimators. For each estimator, the shaded areas describe the 95% empirical confidence intervals of the estimates, obtained with the conventional asymptotic approximation.

Figure A–4: Comparison of the Fixed Effects (FE) estimates



Notes: This figure displays the four non-parametric FE estimates in the reduced samples corresponding to the two NP, the SYM and the SPLIT estimators displayed in previous figures A–1–A–3. Marginal subjects (i.e., subjects located exactly at the cutoff), are dropped only for the NP estimator that excludes them. All the FE estimators are the sample analogs of the Fixed Effect estimands described in Section 5. They differ because of the optimal bandwiths that are set equal to those of the corresponding NP, SYM or SPLIT estimators. The reduced samples have been obtained from the Full Sample constructed with the original data of Pop-Eleches and Urquiola (2013), with the procedure described in Section 8. The RD estimates have been obtained with Local Linear Regressions using a triangular kernel and the optimal bandwidth from Calonico et al. (2014). For each estimator, the shaded areas describe the 95% empirical confidence intervals of the estimates, obtained with the conventional asymptotic approximation.

References

Calonico, S., Cattaneo, M.D., Titiunik, R., 2014. Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. Econometrica 82, 2295–2326.

Pop-Eleches, Urquiola, M., 2013. Going to a Better School: Effects and Behavioural Responses. American Economic Review 103, 1289–1324.