Causal Effect of Urban Parks on Children's Happiness

Nayeon Kwon, Chanmin Kim
Department of Statistics, Sungkyunkwan University

Motivation

Previous Study
- Residential greenness influences children's emotional and behavioral problems. [1]
- Connectedness to nature has an impact on children's happiness. [2]

Purpose
- Few studies investigated the causal effect of nature, especially urban parks, on children's happiness.
- The purpose of this study is to examine the causal relationships between park area per capita and children's happiness.

Design

Data
- The 7th year (2014) survey from the Panel Study on Korean Children
- Unit: children aged 6

Outcome(Y): Child's Happiness
- The Degree of happiness at the time of the 7th survey

Treatment(Z): Park Area per Capita
- **Main analysis** Z=1: $>$ median(7.7m²), Z=0: $\leq$ median(7.7m²)
- **Sensitivity analysis** Z=1: $>$ Q4(16.7m²), Z=0: $\leq$ Q1(5.6m²)
- The number of units in each group

	Main	Sensitivity
Group	**# of units**	
Treated (Z=1)	441	233
Control (Z=0)	512	255
Total	953	488

Potential Outcomes Framework

Notation
- $Z_i \in \{0, 1\}$: treatment for unit i
- $Y_i(1)$: potential outcome if unit i receives a treatment
- $Y_i(0)$: potential outcome if unit i doesn't receive a treatment

Stable Unit Treatment Value Assumption (SUTVA) [6]

$$Y_i = Z_i Y_i(1) + (1 - Z_i) Y_i(0) \tag{1}$$

Ignorability Assumption [5]

$$(Y_i(1), Y_i(0)) \perp Z_i | X_i \tag{2}$$

Average Treatment Effect (ATE)

- $\Delta^{ATE} = E[Y_i(1) - Y_i(0)|X_i] = E[Y_i|X_i, Z_i = 1] - E[Y_i|X_i, Z_i = 0]$ identified by (1) and (2)

Average Treatment effect on the Treated (ATT)

- $\Delta^{ATT} = E[Y_i(1) - Y_i(0)|X_i, Z_i = 1]$

Propensity Score Methods

Propensity Score [5]: $e(x) = P(Z_i = 1|X_i)$ which is the probability of receiving treatment given the observed covariates
Properties of Propensity Score

$$X_i \perp Z_i \mid e(x), \quad (Y_i(1), Y_i(0)) \perp Z_i \mid e(x)$$

Covariate Balancing Propensity Score (CBPS)
- No guarantee of $X_i \perp Z_i \mid \hat{e}(X_i)$ in a given sample due to misspecified the propensity score model
- CBPS is robust to mild misspecification of the propensity score model.

Regression Adjustment

Regression Adjustment

$$\hat{\Delta}_R^{ATE} = \frac{1}{N}\sum_{i=1}^{N} m_1(X_i; \hat{\alpha}_1) - \frac{1}{N}\sum_{i=1}^{N} m_0(X_i; \hat{\alpha}_0),$$

where $m_z(\mathbf{X_i}; \hat{\alpha}_z) = \mathbb{E}[Y_i \mid Z_i = z, \mathbf{X_i}]$ with regression parameters $\hat{\alpha}_z$.

Weighting Methods

Inverse probability weighting (IPW) [4]

$$\hat{\Delta}_{IPW}^{ATE} = n^{-1}\sum_{i=1}^{n} \frac{Z_i Y_i}{\hat{e}(X_i)} - n^{-1}\sum_{i=1}^{n} \frac{(1 - Z_i) Y_i}{1 - \hat{e}(X_i)}.$$

Stabilized inverse probability weighting (SIPW)

$$\hat{\Delta}_{SIPW}^{ATE} = \left(\sum_{i=1}^{n} \frac{Z_i}{\hat{e}(X_i)}\right)^{-1} \sum_{i=1}^{n} \frac{Z_i Y_i}{\hat{e}(X_i)} - \left(\sum_{i=1}^{n} \frac{1 - Z_i}{1 - \hat{e}(X_i)}\right)^{-1} \sum_{i=1}^{n} \frac{(1 - Z_i) Y_i}{1 - \hat{e}(X_i)}.$$

Doubly-robust estimation (DR)

$$\hat{\Delta}_{DR}^{ATE} = \sum_{i=1}^{n} \frac{Z_i Y_i - (Z_i - \hat{e}(X_i)) m_1(X_i, \hat{\alpha}_1)}{n\hat{e}(X_i)} - n^{-1}\sum_{i=1}^{n} \frac{(1 - Z_i) Y_i + (Z_i - \hat{e}(X_i)) m_0(X_i, \hat{\alpha}_0)}{n(1 - \hat{e}(X_i))}.$$

Matching Methods

Propensity Score Matching (PSM) [5]
(1) Obtain $\delta_{i,j}$ which is the distance between treated unit i and control unit j; (2) Find the function η that indicates who is matched to whom by solving

$$\min_{\eta} \sum_{i} \sum_{j=\eta(i)} \delta_{i,j}, \quad \text{where } \delta_{i,j} = \left|\hat{e}(x_i) - \hat{e}(x_j)\right|$$

Covariate Balancing Propensity Score Matching (CBPSM)
- It applies the same process of matching with PSM but using CBPS.

Propensity Score Caliper Matching (Caliper)
- Define the distance between unit i and j as

$$d(\mathbf{x_i}, \mathbf{x_j}) = \begin{cases} Rank - Mahalanobis(\mathbf{x_i}, \mathbf{x_j}) & \text{if } \delta_{i,j} \leq w, \\ \infty & \text{if } \delta_{i,j} > w. \end{cases}$$

- A caliper of 20% of the standard deviation of the propensity score is a common choice.

Coarsened Exact Matching (CEM) [3]
- Step 1 Begin with the covariates $\mathbf{X}$ and make a copy, $\mathbf{X}^*$.
- Step 2 Coarsen $\mathbf{X}^*$ as stated by user-defined cutpoints.
- Step 3 Create one stratum per unique observation of $\mathbf{X}^*$, and put each observation in a stratum.
- Step 4 Assign these strata to the original data, $\mathbf{X}$, and drop any observation whose stratum does not contain at least one treated and one control unit.

Cardinality Matching [7]
- Step 1 Set the covariate balance requirements (e.g., mean balance).
- Step 2 Find the largest matched sample that satisfies the covariate balance requirements.
- Step 3 Rematch the matched sample to minimize covariate distances between matched units.

Confounders

Confounders
- Household income
- Child's literacy
- Father's stress
- Living inside capital area (Y/N)
- Father's education level
  - high school and below (4)
  - associate (5)
  - bachelor (6)
  - post-graduate (7)
- Mother's education level
  - high school and below (4)
  - associate (5)
  - bachelor (6)
  - post-graduate (7)

Covariate Balance Plots

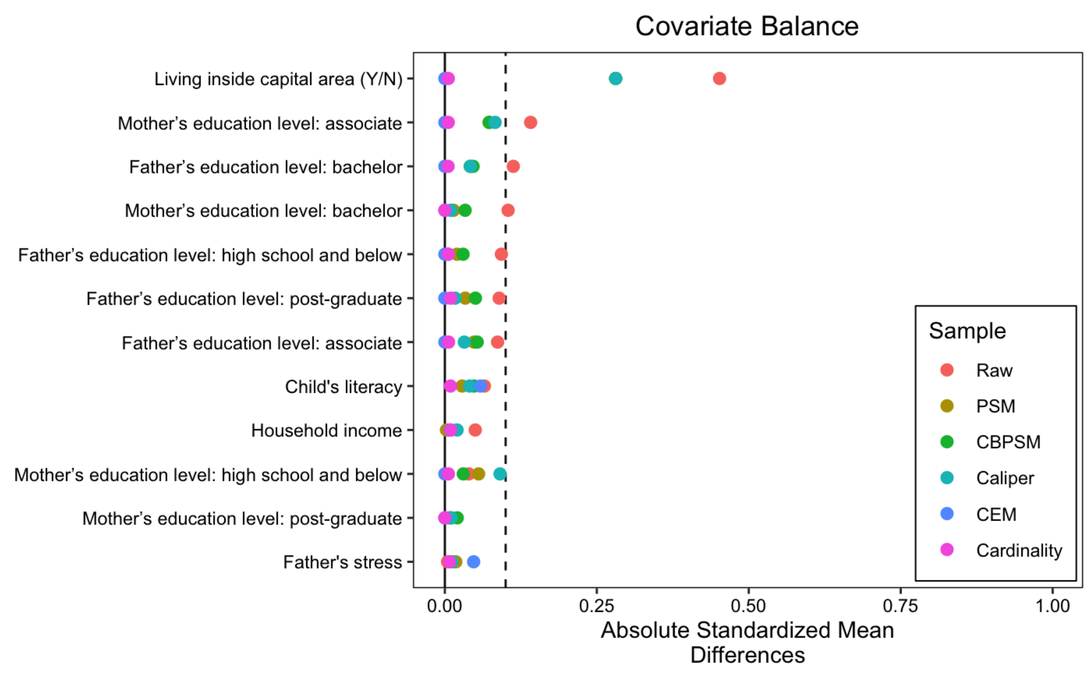

Fig. 1: Absolute SMD before and after matching

Results of The Main Analysis

Methods		ATE/ATT[1]	SD	95% LB	95% UB
Regression Adjustment[2]		0.096	0.156	-0.210	0.402
Weighting	IPW	0.040	0.227	-0.394	0.497
	SIPW	0.047	0.168	-0.287	0.372
	DR	0.090	0.159	-0.215	0.408
Matching	PSM	0.155	0.142	-0.185	0.494
	CBPSM	0.158	0.167	-0.181	0.498
	Caliper	0.103	0.170	-0.233	0.440
	CEM	-0.059	0.327	-0.716	0.597
	Cardinality	0.044	0.175	-0.312	0.400

[1] The estimand for matching
[2] Not only confounders but additional predictors are used (e.g., housing prices).

Discussion

Conclusion
- To satisfy ignorability (unconfoundedness) assumption, we used all possible data referred to in previous studies.
- No causal effect of urban parks on children's happiness

Limitation
- There might be unobserved, important confounders.

Future Study
- Using spatial data, we could study the causal effect of accessibility to urban parks on children's happiness.

References

[1] Balseviciene, B., Sinkariova, L., Grazuleviciene, R., Andrusaityte, S., Uzdanaviciute, I., Dedele, A., and Nieuwenhuijsen, M. J. Impact of residential greenness on preschool children's emotional and behavioral problems. *International journal of environmental research and public health 11*, 7 (2014), 6757–6770.

[2] Barrera-Hernández, L. F., Sotelo-Castillo, M. A., Echeverría-Castro, S. B., and Tapia-Fonllem, C. O. Connectedness to nature: Its impact on sustainable behaviors and happiness in children. *Frontiers in psychology 276* (2020).

[3] Blackwell, M., Iacus, S., King, G., and Porro, G. cem: Coarsened exact matching in stata. *The Stata Journal 9*, 4 (2009), 524–546.

[4] Rosenbaum, P. R. Model-based direct adjustment. *Journal of the American statistical Association 82*, 398 (1987), 387–394.

[5] Rosenbaum, P. R., and Rubin, D. B. The central role of the propensity score in observational studies for causal effects. *Biometrika 70*, 1 (1983), 41–55.

[6] Rubin, D. B. Randomization analysis of experimental data: The fisher randomization test comment. *Journal of the American statistical association 75*, 371 (1980), 591–593.

[7] Zubizarreta, J. R., Paredes, R. D., and Rosenbaum, P. R. Matching for balance, pairing for heterogeneity in an observational study of the effectiveness of for-profit and not-for-profit high schools in chile. *The Annals of Applied Statistics 8*, 1 (2014), 204–231.