

Causal effect of urban parks on children's happiness

Nayeon Kwon^a, Chanmin Kim^{1,a}

^aDepartment of Statistics, SungKyunKwan University

Abstract

Many existing studies have found significant correlations between green spaces, including urban parks, and children's happiness. Furthermore, it was implied that the area/proximity of the urban park would be effective in enhancing infancy happiness. However, inferring causal effects from observed data requires appropriate adjustment of confounding variables, and from this perspective, the causal relationship between the area of urban parks and children's happiness has not been well understood. The causal effect of urban parks on children's happiness was estimated in this study using data from the panel study on Korean children. As methods for adjusting confounding variables, regression adjustment using a regression method, weighting method, and matching method were used, and key concepts of each method were described before the analysis results. Confounders were chosen for the analysis using a directed acyclic graph. In contrast to previous research, the analysis found no significant causal relationship between the size of the city park and children's happiness.

Keywords: panel study on Korean children, causal inference, weighting, matching

1. 서론

도시공원은 도시지역에서 도시자연경관을 보호하고 시민의 건강·휴양 및 정서 생활을 향상하는 데에 이바지하기 위한 공간을 뜻한다 (국도교통부, 2021). 그 기능 및 주제에 따라 생활권공원(어린이공원, 근린공원, 소공원)과 주제공원(역사공원, 문화공원, 묘지공원, 체육공원, 수변공원과 특별시, 광역시 또는 도의 조례가 정하는 공원)으로 분류된다.

도시공원은 거주자에게 다양한 이점을 제공한다. Hartig 등 (2014)은 도시공원 등의 녹지가 대기질 개선, 신체활동 촉진, 스트레스 감소, 사회적 응집력이라는 네 가지 주된 경로를 통하여 건강에 기여한다고 밝혔다. 또한, Kim과 Kaplan (2004)에 의하면 녹지는 사회적 상호작용 및 소속감 촉진에 중요한 역할을 한다. 그리고 도시공원 등의 자연 편의시설은 물과 공기 정화, 방음, 도시 내 미기후 안정화의 역할을 할 뿐만 아니라, 거주자의 주관적 행복감을 증진하는 사회·정서적 서비스도 제공한다 (Chiesura, 2004). 이뿐만 아니라 도시공원에 대한 접근과 이용은 전 연령대의 신체활동, 앉아있는 시간의 감소와 연관성이 존재한다 (Epstein 등, 2006; Kaczynski와 Henderson, 2007; Kaczynski 등, 2008).

특히 도시공원을 비롯한 녹지는 유아의 정서, 인지 및 행동발달에 긍정적인 영향을 미친다. 유아는 만 3세부터 초등학교 취학 전까지의 어린이를 뜻한다 (교육부, 2022). Balseviciene 등 (2014)에 의하면 도시공원에 대한 접근성은 어머니의 학력 수준이 낮은 유아의 정신건강에 긍정적인 영향을 미친다. 또한 자연 기반의

This work is supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (NRF-2020R1F1A1A01048168, NRF-2022R1F1A1062904).

¹ Corresponding author: Department of Statistics, SungKyunKwan University, 25-2 Seonggyunwan-ro, Jongno-gu, Seoul 03063, Korea. E-mail: chanmin.kim@skku.edu

프로그램에 참여하는 것은 (유아를 포함한) 어린이의 자아 존중감과 행복감을 증진하며, 이는 부모의 사회·경제적 지위가 낮은 어린이에게 더 큰 영향을 미친다 (Readdick과 Schaller, 2005). 그리고 Barrera-Hernández 등 (2020)는 자연과 많이 접촉하는 어린이가 이타심, 공정심, 검소함 등의 긍정적인 생각을 가지고 있을 뿐만 아니라, 더 높은 수준의 행복감을 보인다는 것을 밝혔다. 인지발달 측면에서 거주지 주변에 뛰놀 수 있는 자연공간이 갖추어진 경우 유아는 더 높은 인지 기능 점수를 보였다 (Mårtensson 등, 2009). 또한 Faber Taylor와 Kuo (2011)는 녹지를 이용하는 것이 전 연령대 어린이의 ADHD 증상을 완화한다는 것을 밝혔다. 행동발달 측면에서 녹지에 대한 접근성이 있을 때 어린이는 더 높은 신체활동 수준을 보였다 (Boldemann 등, 2006; Cohen 등, 2006; Epstein 등, 2006; Roemmich 등, 2006).

지금까지 도시공원과 행복감 간의 연관성 연구는 주로 성인을 대상으로 행해졌다 (Houlden 등, 2018; Kim과 Jin, 2018). 그러나 한 사람의 생애 전반에 영향을 미치는 유아기 행복감에 주목할 필요성이 있다. 선행 연구에 의하면 행복은 유아기에 형성되며, 생애 초기의 행복 수준은 생애 후기까지 지속적인 영향을 미친다 (Gwon과 Sung 2010; Lee, 2016). 또한 유아기 행복 경험은 장기적 삶과 성취의 기본이 되는 정서적 안정감과 자존감, 건강한 신체와 정신, 대인관계, 사회적 유능성 등에 긍정적인 영향을 미친다 (Kim, 2011; Compton, 2005; Fredrickson, 2001). 그러므로 본 논문은 도시공원의 면적이 유아 행복감에 미치는 영향에 대한 인과효과를 분석하였다.

본 논문에서는 한국아동패널 7차자료의 유아 행복감 변수와 전국도시공원 정보표준자료 및 행정구역(시군구)별 주민등록 인구현황 자료를 활용하여 생성한 도시공원 면적 변수를 이용하여 인과효과를 추정하고자 한다. 관측 연구(observational study) 하에서 선택편향(selection bias)으로 인해 발생하는 문제를 최소화하기 위한 방법으로 성향 점수(propensity score) 방법 등이 제시되어 있다. 가중치 기법, 매칭 등의 성향 점수 방법을 이용한 교란 변수 영향의 통제 결과에 어떠한 차이가 있는지 데이터 분석을 통해 비교한다. 특히 본 연구에서는 한국아동패널자료 연구에서 적용되지 않았던 교란 변수의 영향 통제 방법으로 회귀 모형을 이용한 조정(regression adjustment), 가중치 기법(weighting methods), 그리고 매칭(matching methods)을 비교 분석하여 각 분석 방법의 특징과 차별점을 논의하는 데 목적을 두었다.

인과효과 추정의 관심 모수인 평균처치효과(average causal effect; ACE)는 ‘모집단 전체’에 대한 평균처치효과(average treatment effect; ATE)와 ‘처리군’에 대한 평균처치효과(average treatment effect for the treated; ATT)로 나눌 수 있다. 이 연구에서는 실증 자료 분석에 기반하여 교란 변수(confounder) 영향의 통제가 인과효과(ATE, ATT) 추정에서 어떤 효과가 있는지 알아보고자 한다. 정확한 인과효과 추정을 위해서 잠재적 반응 변수 체계(potential outcomes framework)를 사용하였고, 유향비순환그래프(directed acyclic graph; DAG)를 이용하여 교란 변수를 선택하였다. 2절에서는 인과효과추론에 필요한 가정과 성향 점수(propensity score), 그리고 인과효과 추정 방법을 설명한다. 3절에서는 분석에서 사용된 자료를 설명하고 변수들을 명확히 정의한다. 4절에서는 본 분석과 민감도 분석의 결과를 설명한다. 5절에서는 본 논문의 의의와 한계를 설명하고자 한다.

2. 방법론

2.1. 인과관계추론

처치 변수(treatment)의 지표는 Z_i 로 표기하며 ($Z_i \in \{0, 1\}$), $Y_i(0)$ 과 $Y_i(1)$ 은 i 번째 개체의 잠재적 반응 변수(potential outcome)이다. $Y_i(0)$ 은 i 번째 개체가 처치 변수 ($Z_i = 0$) 값을 받았을 때의 잠재적 반응 변수이고, $Y_i(1)$ 은 처치 변수 ($Z_i = 1$) 값을 받았을 때의 잠재적 반응 변수이다. 각 개체의 인과효과는 두 잠재적 반응 변수의 차이/비교를 통해 측정한다. 하지만, 각 개체에서 두 잠재적 반응 변수는 동시에 관찰될 수 없다 (the fundamental problem of causal inference) (Holland, 1986).

그러므로 관찰된 데이터로부터 인과효과를 추정하기 위해서는 몇 가지 가정들이 필요하다. 첫 번째, SUTVA (stable unit treatment value assumption) (Rubin, 1980) 가정이다. SUTVA는 비간섭(no interference)

과 일관성(consistency) 가정을 함의한다. 비간섭 가정은 각 개체의 잠재적 반응 변수가 다른 개체에 할당된 처치 변수 값에 영향을 받지 않는다는 가정이며 일관성 가정은 처치 변수 값 ($Z = 1, Z = 0$)이 명확한 의미를 갖고 있어야 한다는 가정이다. SUTVA 가정을 통해 관찰된 반응 변수 (Y_i)와 잠재적 반응 변수의 관계를 도출할 수 있다:

$$Y_i = Z_i Y_i(1) + (1 - Z_i) Y_i(0), \text{ for } i = 1, \dots, n. \quad (2.1)$$

두 번째는 무시성(ignorability) 가정이다 (Rosenbaum와 Rubin, 1983). 편향 없는 인과효과를 추정하기 위해서는 교란 변수(confounder)를 통제해야한다. 교란 변수란 처치 변수를 받기 전에 측정된 변수로써 처치 변수와 잠재적 반응 변수 모두에 영향을 미치는 변수이다. X_i 가 모든 교란 변수를 포함하고 있다고 가정했을 때, 개체들이 X_i 를 공유한다면 다음을 만족한다:

$$(Y_i(1), Y_i(0)) \perp Z_i \mid X_i. \quad (2.2)$$

가정 (2.1)과 (2.2) 하에 평균처치효과(average treatment effect; ATE)는 다음과 같이 정의된다. 평균처치효과 (ATE)는 전체 모집단에 대한 처치효과를 의미한다.

$$\begin{aligned} \Delta^{\text{ATE}} &= \mathbb{E}[Y_i(1) - Y_i(0)] \\ &= \mathbb{E}_x[\mathbb{E}[Y_i(1) \mid X_i]] - \mathbb{E}_x[\mathbb{E}[Y_i(0) \mid X_i]] \\ &= \mathbb{E}_x[\mathbb{E}[Y_i(1) \mid Z_i = 1, X_i]] - \mathbb{E}_x[\mathbb{E}[Y_i(0) \mid Z_i = 0, X_i]] \\ &= \mathbb{E}_x[\mathbb{E}[Y_i \mid Z_i = 1, X_i]] - \mathbb{E}_x[\mathbb{E}[Y_i \mid Z_i = 0, X_i]]. \end{aligned}$$

마지막으로 오버랩(overlap) 가정은 각 개체가 모든 처치 변수 값에 할당될 확률이 존재함을 뜻한다.

$$0 < P(Z_i = 1 \mid X_i) < 1, \text{ for all } i.$$

위의 세 가정은 임의화 실험(randomized experiments) 하에서 만족한다. 그러므로 임의화 실험은 인과관계 추론의 이상적인 환경이다. 하지만 윤리적·법적 문제로 인하여 처치 변수 값을 임의적으로 할당할 수 없는 경우, 피실험자가 특정 처치 변수 값을 선호하여 대조군에 속하기를 거부하는 경우, 그리고 선택편향이 존재하는 상황에서는 임의화 실험을 진행할 수 없다. 이 경우 관측 연구(observational study)를 진행한다. 그러나 관측 연구는 임의화 실험과 달리 각 개체가 처치 변수 값을 랜덤으로 할당받지 않는다. 그래서 관측 연구에서는 편향 없는 인과효과 추정을 위하여 교란 변수의 영향을 통제하는 것이 중요 과제이다.

2.2. 성향 점수

성향 점수(propensity score)는 공변량 X_i 를 조건부로 취했을 때 처치 변수를 받을 확률을 뜻한다 (Rosenbaum와 Rubin, 1983). X_i 는 다차원일 수 있으며, 성향 점수 $e(x)$ 는 일차원 값이다.

$$e(x) = P(Z_i = 1 \mid X_i = x).$$

다음의 성질로 인하여 성향 점수는 인과효과 추정 시 유용하게 사용될 수 있다. $e(x)$ 은 가장 성긴 균형 점수(the coarsest balancing score) $b(x)$ 이며, $e(x)$ 는 모든 $b(x)$ 의 함수 형태로 표현될 수 있다.

$$X_i \perp Z_i \mid b(x).$$

그러므로 다음과 같이 동일한 $e(x)$ 값을 가진 처리군과 대조군의 개체들은 같은 공변량 분포를 보인다.

$$X_i \perp Z_i \mid e(x).$$

만약 가정 (2.2)가 성립한다면, 다차원 X_i 대신 일차원 $e(x)$ 를 사용하여도 다음의 관계가 만족한다 (Rosenbaum 과 Rubin, 1983):

$$(Y_i(1), Y_i(0)) \perp Z_i | e(x).$$

임의화 실험 진행 시 $e(x)$ 는 실험 설계에 의해 정의되며 알려진 값이다. 반면, 관측 연구에서 $e(x)$ 는 알려지지 않은 값이기 때문에 데이터로부터의 추정이 필요하다 (Austin, 2011). 만약 X_i 가 저차원이라면 $e(x)$ 는 비모수 적으로 추정 가능하지만, X_i 가 고차원이라면 일반적으로 모수적 방법을 이용하여 추정한다. $e(x)$ 추정 시 가장 많이 사용되는 모형은 $e(X) = \{1 + \exp(-X^T\beta)\}^{-1}$ 와 같은 로지스틱 회귀 모형이다. 로지스틱 회귀 모형에서 추정량 β 는 최대가능도 방법(maximum likelihood estimation)을 이용하여 추정할 수 있다.

$$\sum_{i=1}^n \frac{Z_i - e(X_i)}{e(X_i)\{1 - e(X_i)\}} \frac{\partial}{\partial \beta} \{e(X_i)\} = \mathbf{0}.$$

하지만 이는 공변량의 균형이 아닌 가능도 함수를 최대로 만들어주는 모수의 추정량을 찾는 방법이다. 로지스틱 회귀 모형이 잘 설계되었을 경우 최대가능도 추정량은 처리군과 대조군의 공변량 분포 균형을 최대로 맞춰주는 추정량이다. 하지만 이 모형이 잘못 설계되었을 경우 추정량은 공변량의 균형을 잘 맞춰주는 추정량이 아닐 수 있다 (Wyss 등, 2014). 매칭 시 이러한 문제를 해결하기 위한 방안으로 공변량 균형 성향 점수 (covariate balancing propensity score; CBPS)가 고안되었다. 또한 가중치 기법에서 회귀 모형과 성향 점수 모형이 모두 잘못 설계되었을 경우 공변량 균형 성향 점수는 처치 효과의 추정량을 더 강건하게 한다 (Imai와 Ratkovic, 2014).

2.3. 인과효과 추정 방법

2.3.1. 회귀 모형을 이용한 조정(regression adjustment)

가장 널리 사용되는 인과효과 추정 방법은 회귀 모형을 이용한 조정(regression adjustment)이다. Y_i 는 연속형 변수임을 가정할 때, 다음의 선형 회귀 모형을 고려할 수 있다.

$$\mathbb{E}[Y_i | Z_i = z, X_i] = \beta_0 + \beta_z z + \beta_x^T X_i.$$

만약 가정 (2.2)를 만족한다면 다음의 식이 성립한다.

$$\begin{aligned} \mathbb{E}[Y_i | Z_i = 1, X_i] &= \mathbb{E}[Y_i(1) | X_i] = \beta_0 + \beta_z + \beta_x^T X_i, \\ \mathbb{E}[Y_i | Z_i = 0, X_i] &= \mathbb{E}[Y_i(0) | X_i] = \beta_0 + \beta_x^T X_i. \end{aligned}$$

그러므로 상수 인과효과를 가정하였을 때의 평균 처치 효과인 ATE는 다음과 같으며,

$$\mathbb{E}_x[\mathbb{E}[Y_i | Z_i = 1, X_i] - \mathbb{E}[Y_i | Z_i = 0, X_i]] = \mathbb{E}_x[\beta_z] = \beta_z.$$

ATE 추정량 $\hat{\Delta}_R^{\text{ATE}}$ 는 다음과 같이 표현할 수 있다.

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

여기서 $m_z(X_i; \hat{\alpha}_z) = \mathbb{E}[Y_i | Z_i = z, X_i]$ 이고, $\hat{\alpha}_z$ 는 각 회귀 모형의 모수이다.

회귀 모형을 이용한 조정을 통해 인과효과를 추정 시 m_1 과 m_0 을 정확히 설계하는 것이 중요하다. 회귀 모형을 이용하여 교란 변수의 영향을 통제할 경우, 만약 회귀 모형이 처치 변수와 교란 변수, 그리고 반응 변수와의 관계를 잘 반영하였다면, 추정량은 일치성을 만족하며 점근적으로는 비편향성을 만족한다 (Funk 등, 2011).

2.3.2. 가중치 기법(weighting methods)

가중치 기법(weighting methods)은 $\hat{e}(x)$ 를 이용하여 Δ^{ATE} 를 추정한다. 가중치 추정 시 반응 변수와 독립 변수의 모형화가 필요하지 않으며 (Rosenbaum, 1987), 한번 구한 가중치 집합은 다른 반응 변수에 대한 인과효과 추정 시에도 사용될 수 있다 (Little과 Rubin, 2019)는 점에서 실용적이다. 인과효과 추정 시 가중치 기법의 목적은 교란 변수에 의한 편향을 제거하여 안정된 Δ^{ATE} 를 추정하는 것이다. 매우 큰 가중치는 전반적인 분석 결과에 큰 영향을 미칠 수 있으며, 일정하지 않은 가중치는 분산을 크게 만들 수 있다 (Little과 Rubin, 2019).

역확률 가중치 기법(inverse probability weighting; IPW)

가장 많이 사용되는 가중치 기법은 역확률 가중치 기법(inverse probability weighting; IPW)이다 (Rosenbaum, 1987). 여기서 가중치는 $w_i = Z_i/e(X_i) + (1 - Z_i)/(1 - e(X_i))$ 로 정의된다. (2.1)에 의하여 Δ^{ATE} 는 다음과 같이 정의된다:

$$\Delta^{\text{ATE}} = \mathbb{E} \left[\frac{Z_i Y_i}{e(X_i)} - \frac{(1 - Z_i) Y_i}{1 - e(X_i)} \right].$$

그리고 (2.2) 하에 다음의 식이 성립한다.

$$\mathbb{E} \left[\frac{Z_i Y_i}{e(X_i)} \right] = \mathbb{E} \left\{ \frac{1}{e(X_i)} \mathbb{E}[Z_i Y_i(1) | X_i] \right\} = \mathbb{E} \left\{ \frac{\mathbb{E}[Z_i | X_i]}{e(X_i)} \mathbb{E}[Y_i(1) | X_i] \right\} = \mathbb{E}[Y_i(1)].$$

동일한 방식으로 $\mathbb{E}[(1 - Z)Y/(1 - e(x))] = \mathbb{E}[Y(0)]$ 이다. 이를 통하여 다음의 ATE 추정량 $\hat{\Delta}_{\text{IPW}}^{\text{ATE}}$ 을 도출할 수 있다.

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

여기서 $\hat{\Delta}_{\text{IPW}}^{\text{ATE}}$ 의 성능은 $\hat{e}(X_i)$ 에 의존한다. $0 < \hat{e}(X_i) < 1$ 이며 $\hat{e}(X_i)$ 의 값이 0에 가깝다면 처리군의 개체가 받는 가중치 $1/\hat{e}(X_i)$ 는 극단적으로 커진다. 또는 $\hat{e}(X_i)$ 의 값이 1에 가깝다면 대조군 개체의 가중치 $1/(1 - \hat{e}(X_i))$ 가 매우 커진다. 극단적인 가중치는 준모집단(pseudo-population) 내 소수의 개체들이 너무 큰 대표성을 가지게 하며, $\hat{\Delta}_{\text{IPW}}^{\text{ATE}}$ 의 분산을 커지게 한다 (Hernán 등, 2000).

안정화된 역확률 가중치 기법(stabilized inverse probability weighting; SIPW)

역확률 가중치 기법을 보완하기 위한 방법으로 안정화된 역확률 가중치 기법(stabilized inverse probability weighting; SIPW)이 제시되었다 (Hernán 등, 2000). 안정화된 역확률 가중치 기법의 추정량 $\hat{\Delta}_{\text{SIPW}}^{\text{ATE}}$ 는 $\hat{\Delta}_{\text{IPW}}^{\text{ATE}}$ 와 달리 경계가 있는 값이므로, 극단적인 가중치의 영향이 경감된다는 이점이 있다.

$$\mathbb{E} \left[\frac{Z_i}{e(X_i)} \right] = \mathbb{E} \left[\mathbb{E} \left[\frac{Z_i}{e(X_i)} \mid \mathbf{X}_i \right] \right] = \mathbb{E} \left[\frac{\mathbb{E}[Z_i | \mathbf{X}_i]}{e(X_i)} \right] = \mathbb{E} \left[\frac{e(X_i)}{e(X_i)} \right] = 1.$$

동일한 방식으로 $\mathbb{E}[(1 - Z_i)/(1 - e(X_i))] = 1$ 이다. 이를 이용하여 Δ^{ATE} 를 다시 표현할 수 있다.

$$\Delta^{\text{ATE}} = \frac{\mathbb{E}[Z_i Y_i / e(X_i)]}{\mathbb{E}[Z_i / e(X_i)]} - \frac{\mathbb{E}[(1 - Z_i) Y_i / (1 - e(X_i))]}{\mathbb{E}[(1 - Z_i) / (1 - e(X_i))]}.$$

그러므로 $\hat{\Delta}_{\text{SIPW}}^{\text{ATE}}$ 는 다음과 같이 정의된다.

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

수식의 첫 번째 부분은 가중치가 $(Z_i/\hat{e}(X_i))/(\sum_{k=0}^n Z_k/\hat{e}(X_k))$ 인 Y 의 가중 평균으로 볼 수 있으며, 두 번째 부분은 가중치가 $((1-Z_i)/(1-\hat{e}(X_i)))/(\sum_{k=1}^n (1-Z_k)/(1-\hat{e}(X_k)))$ 인 Y 의 가중 평균으로 볼 수 있다. 또한 모든 개체의 가중치 합은 1이며, 각 개체의 가중치는 0과 1을 포함한 그 사이의 값을 갖는다. 그러므로 수식의 첫 번째 부분에 대하여 다음이 성립한다.

$$\min_i Y_i \leq \left(\sum_{i=1}^n \frac{Z_i}{\hat{e}(X_i)} \right)^{-1} \sum_{i=1}^n \frac{Z_i Y_i}{\hat{e}(X_i)} \leq \max_i Y_i.$$

이와 유사한 방식으로, 수식의 두 번째 부분은 $[\min Y_i, \max Y_i]$ 의 경계값을 갖는다. 그러므로 $\hat{\Delta}_{SIPW}^{ATE}$ 는 다음과 같이 표현된다.

$$\min Y_i - \max Y_i \leq \hat{\Delta}_{SIPW}^{ATE} \leq \max Y_i - \min Y_i.$$

만약 성향 점수 모형인 $e(X_i)$ 가 정확히 설계되었다면, $\hat{\Delta}_{IPW}^{ATE}$ 와 $\hat{\Delta}_{SIPW}^{ATE}$ 는 모두 일치 추정량이다 (Lunceford와 Davidian, 2004).

이중 강건 추정기법(doubly-robust estimation; DR)

다음으로 Robins 등 (1994)이 제안한 이중 강건 추정기법(doubly-robust estimation; DR)은 회귀 모형과 성향 점수 모형 중 하나라도 정확하게 설계되었을 경우 비편향 추정량을 구할 수 있는 방법이다. 추정량 $\hat{\Delta}_R^{ATE}$ 와 $\hat{\Delta}_{(S)IPW}^{ATE}$ 는 다른 모형의 설계를 요구한다. 만약 회귀 모형은 정확히 설계되었지만 성향 점수 모형이 잘못 설계된 경우 $\hat{\Delta}_R^{ATE}$ 은 비편향 추정량이지만 $\hat{\Delta}_{(S)IPW}^{ATE}$ 에는 편향이 존재할 것이다. 이중 강건 추정기법은 두 경우를 모두 포함하므로 비편향 추정량을 얻을 가능성을 높일 수 있다 (Funk 등, 2011). 이때 추정량 $\hat{\Delta}_{DR}^{ATE}$ 은 다음과 같이 정의된다.

$$\hat{\Delta}_{DR}^{ATE} = n^{-1} \sum_{i=1}^n \frac{Z_i Y_i - (Z_i - \hat{e}(X_i)) m_1(X_i, \alpha_1)}{\hat{e}(X_i)} - n^{-1} \sum_{i=1}^n \frac{(1 - Z_i) Y_i + (Z_i - \hat{e}(X_i)) m_0(X_i, \alpha_0)}{1 - \hat{e}(X_i)}.$$

수식의 첫 번째 부분의 기댓값은 다음과 같이 표현될 수 있으며,

$$\mathbb{E} \left[\frac{Z_i Y_i}{e(X_i)} \right] - \mathbb{E} \left[\left\{ \frac{Z_i}{e(X_i)} - 1 \right\} m_1(X_i, \alpha_1) \right],$$

만약 성향 점수 모형이 정확히 설계되었다면 다음의 식을 만족한다.

$$\mathbb{E} \left[\left\{ \frac{Z_i}{e(X_i)} - 1 \right\} m_1(X_i, \alpha_1) \right] = 0.$$

만약 회귀 모형이 정확히 설계되었다면 다음의 식이 성립한다.

$$\mathbb{E} \left[\frac{Z_i}{e(X_i)} \{Y_i - m_1(X_i, \alpha_1)\} \right] = 0.$$

하지만 만약 회귀 모형과 성향 점수 모형이 모두 정확히 설계되지 않은 경우, $\hat{\Delta}_{DR}^{ATE}$ 는 일치 추정량이 아니다. 또한 $\hat{\Delta}_{DR}^{ATE}$ 가 $\hat{\Delta}_{(S)IPW}^{ATE}$ 보다 좋은 성능을 보인다는 근거는 없다 (Kang과 Schafer, 2007). Imai와 Ratkovic (2014)의 시뮬레이션 결과에 따르면, 두 모형이 모두 잘못 설계된 경우 공변량 균형 성향 점수(CBPS)를 기반으로 한 이중 강건 추정기법의 추정량은 로지스틱 회귀 모형을 통한 성향 점수를 기반으로 하였을 때보다 편향과 평균제곱오차(root mean square error; RMSE) 모두 상당히 감소한 결과를 보였다.

2.3.3. 매칭(Matching methods)

매칭(matching methods)은 관측 연구 하에서 처리군과 대조군 간 공변량 분포의 균형을 맞추주기 위한 방법이다. 매칭은 실험 설계와 결과 분석 단계로 나뉜다. 실험 설계는 관측 연구가 임의화 실험 환경을 모방할 수 있도록 매칭을 시행하는 단계이다. 이 단계에서는 거리 측정 방법을 정의한 후 두 집단 내 개체 간 거리를 계산한다. 그리고 이에 기반하여 매칭을 시행하고 매칭의 질을 평가한다. 결과 분석은 매칭된 표본을 이용하여 인과효과를 추정하는 단계이다 (Stuart, 2010).

매칭의 장점은 다음과 같다. 매칭은 회귀 모형을 이용한 조정과 상호보완적으로 사용할 수 있다. 두 방법을 함께 사용하였을 때 교란 변수로 인한 인과효과 추정의 편향을 가장 효과적으로 줄일 수 있다. 그리고 매칭은 처리군과 대조군 간의 공변량 분포가 충분히 겹치지 않는 부분에 주목한다. 두 그룹 간의 공변량 분포가 다른 경우 인과효과 추정 결과는 외삽법(extrapolation)에 크게 의존한다. 매칭은 연구자가 이를 인지함으로써 추정의 질을 파악할 수 있게 한다 (Stuart, 2010). 또한 매칭은 실험 설계와 결과 분석 단계가 분리되어 있으므로 연구의 투명성을 보장한다.

매칭 시 일반적으로 관찰된 처리군에 대한 평균처리효과(average treatment effect for the treated; ATT)를 추정한다 (Imbens, 2004).

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

가장 이상적인 매칭은 같은 공변량 X_i 를 가진 처리군과 대조군의 개체를 매칭하는 정확 매칭(exact matching)을 시행하는 것이다. 하지만 실제 데이터에 이를 활용하는 경우 많은 샘플을 버리게 된다. 특히 X_i 가 고차원이거나 연속형 변수인 경우 정확히 같은 공변량을 갖도록 두 개체를 매칭하는 것은 불가능하다 (Ho 등, 2007). 그리고 이로 인해 샘플 내의 많은 개체들이 버려진다면, 이는 유사한 공변량 분포를 가진 개체들을 매칭하여 더 많은 수의 개체들을 매칭에 사용하였을 때보다 더 큰 편향을 유발한다 (Rosenbaum과 Rubin, 1985). 그러므로 공변량의 분포를 최대한 유사하게 맞추주는 다음의 방법들을 고려할 수 있다.

첫 번째는 성향 점수 추정량 $\hat{e}(x)$ 를 이용한 성향 점수 매칭(propensity score matching; PSM)이다 (Rosenbaum과 Rubin, 1983). 그리고 $\hat{e}(x)$ 와 마할라노비스 거리(Mahalanobis distance)를 통합하여 사용하는 성향 점수 범위짜깁기 매칭(propensity score caliper matching; Caliper)을 이용할 수 있다 (Rubin과 Thomas, 2000). 다음으로 X_i 가 고차원일 때 정확 매칭을 활용할 수 있도록 고안한 성긴 정확 매칭(coarsened exact matching; CEM)을 고려할 수 있다 (Iacus 등, 2012). 마지막으로 연구자가 설정한 공변량 균형의 기준을 만족하는 매칭된 표본의 크기를 최대화하는 카디널리티(cardinality) 매칭이 있다 (Zubizarreta 등, 2014).

성향 점수 매칭(propensity score matching; PSM)

관찰 연구 하에서 $e(x)$ 는 알려지지 않은 값이므로, 성향 점수 매칭은 $\hat{e}(x)$ 를 이용한다. 처리군의 i 번째 개체와 대조군의 j 번째 개체의 거리는 다음과 같이 계산된다.

$$\delta_{i,j} = \left| \hat{e}(x_i) - \hat{e}(x_j) \right|.$$

다음의 함수 η 를 통해 처리군의 i 번째 개체와 대조군의 j 번째 개체의 매칭이 결정된다.

$$\min_{\eta} \sum_i \sum_{j=\eta(i)} \delta_{i,j}.$$

성향 점수 매칭은 공변량의 균형을 얻기까지 많은 반복을 필요로 하며 (Hainmueller, 2012), 이 과정은 공변량 균형 성향 점수(covariate balancing propensity score; CBPS)를 이용함으로써 개선될 수 있다. 또한 성향 점수 매칭은 거리 계산 시 X_i 가 아닌 $\hat{e}(x)$ 를 이용하므로, 매칭된 두 개체가 유사한 x_i 값을 가지는 것을 보장할 수

없다. 이를 보완하기 위하여 성향 점수 범위짜짓기 매칭을 이용할 수 있다.

성향 점수 범위짜짓기 매칭(propensity score caliper matching; Caliper)

성향 점수 범위짜짓기 매칭(propensity score caliper matching; Caliper)은 유사한 X_i 와 $\hat{e}(x)$ 를 가진 처리군과 대조군의 개체를 매칭하는 방법이다 (Rubin과 Thomas, 2000). 처리군의 i 번째 개체와 대조군의 j 번째 개체의 거리 $d(x_i, x_j)$ 는 다음과 같이 계산된다.

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

만약 $\delta_{i,j} \leq w$ 이면 순위 기반 마할라노비스 거리(rank-based Mahalanobis distance)를 이용하여 거리를 계산하고, $\delta_{i,j} > w$ 이면 거리를 ∞ 로 정의하여 매칭되지 못하게 한다. 일반적으로 초기 w 설정 시 성향 점수 표준편차의 20%를 선택한다. 그리고 공변량의 균형이 맞을 때까지 w 의 값을 조정한다. 이 과정에서 반응 변수는 사용되지 않으므로, 이 과정의 반복은 인과효과 추정에 영향을 미치지 않는다.

순위 기반 마할라노비스 거리는 다음과 같이 계산된다 (Rosenbaum, 2002).

$$(r(x_i) - r(x_j))^T (D \hat{\Sigma} D)^{-1} (r(x_i) - r(x_j)).$$

여기서 $r(x_i), r(x_j)$ 는 각각 처리군의 i 번째 개체와 대조군의 j 번째 개체의 공변량들을 순위로 변환한 값이다. 동물인 경우 순위의 평균을 사용한다. $D \hat{\Sigma} D$ 는 수정된 공분산 행렬을 나타내며, 이는 순위로 변환한 공분산 행렬 $\hat{\Sigma}$ 의 전후에 대각 행렬 D 를 곱한 값이다.

$$D = \begin{bmatrix} \frac{S_u}{S_{i,1}} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \frac{S_u}{S_{i,k}} \end{bmatrix}.$$

여기서 k 번째 공변량의 동물이 아닌 순위들의 표준 편차는 S_u 이며, 동물인 순위의 표준 편차는 $S_{i,l}$ 로 나타낸다.

데이터가 정규 분포를 따르지 않거나 데이터 내에 이상치가 존재하는 경우 마할라노비스 거리를 이용하여 거리 계산 시 표준 편차가 팽창하는 등의 이상 현상을 보인다. 그러므로 순위 기반 마할라노비스 거리를 이용한다.

성긴 정확 매칭(coarsened exact matching; CEM)

성긴 정확 매칭(coarsened exact matching; CEM)은 공변량을 성기계(coarsening) 한 뒤 정확 매칭을 시행한다. 성긴 정확 매칭의 알고리즘은 다음의 순서로 진행된다 (Blackwell 등, 2009).

Step 1: 공변량 X 의 복사본 X^* 를 생성한다.

Step 2: 연구자가 직접 정의한 절단점(cutpoint) 또는 자동적인 구간화(binining) 알고리즘(e.g., R 패키지 CEM)에 따라 X^* 를 성기계 한다.

Step 3: X^* 의 각 구간을 하나의 층(stratum)으로 간주하고, 각 층에 속해있는 개체를 확인한다.

Step 4: 각 층에서 처리군과 대조군의 개체가 최소 하나씩 포함되지 않았을 경우, 원본 데이터 X 에서 해당 층과 그 구간에 속해있는 개체를 삭제한다.

성긴 정확 매칭은 단조 불균형 경계(monotonic imbalance bounding; MIB)에 속하는 방법 중 하나로서, 연구자의 사전 선택을 통해서 공변량 분포의 최대 불균형을 통제할 수 있다는 장점이 있다. 이는 변수를 성기계 하는 과정을 통해 분포 내 최대 불균형(비선형성, 변수 간 상호작용, 적률, 분위수 등)을 통제한다. 또한 성긴 정확 매칭은 앞선 방법들과 달리 각 공변량의 최대 불균형을 독립적으로 개선시킬 수 있다. 한 공변량의 최대 불균형을 조정하는 것은 다른 공변량들의 최대 불균형에 영향을 미치지 않기 때문이다 (Iacus 등, 2012). 변수를 성기계 하는 과정에서 R 패키지 CEM을 이용할 수도 있지만, 이 매칭의 장점을 활용하기 위해서는 변수에 대한 사전 지식을 바탕으로 연구자가 직접 변수를 성기계 하는 것이 권장된다 (Iacus 등, 2012).

카디널리티(cardinality) 매칭

카디널리티(cardinality) 매칭은 성긴 정확 매칭처럼 매칭 시 공변량 원형을 이용한다. 카디널리티 매칭은 다음의 순서로 진행된다 (Visconti와 Zubizarreta, 2018).

Step 1 : 공변량 균형 제약(mean balance 등)을 설정한다.

Step 2 : 앞서 설정한 공변량 균형 제약을 만족하는 최대 크기의 매칭된 표본을 찾는다.

Step 3 : 공변량들의 거리를 최소화하는 처리군과 대조군 쌍을 찾기 위해 단계 2에서 얻은 표본을 재매칭한다.

카디널리티 매칭의 목표는 균형을 이룬 최대 크기의 1 : 1로 매칭된 표본을 찾는 것이며, 목적 함수는 다음과 같다.

$$\arg \max_m \sum_{t \in T} \sum_{c \in C} m_{tc}.$$

매칭 전 처리군과 대조군의 집합은 $\mathcal{T} = \{1, \dots, T\}$ 와 $\mathcal{C} = \{1, \dots, C\}$ 이며, $T \leq C$ 이다. $t \in \mathcal{T}$ 와 $c \in \mathcal{C}$ 는 두 집합에 대응하는 지표(index)를 나타낸다. 이진 결정 변수를 사용하여 처리군의 개체 t 가 대조군의 개체 c 와 매칭된다면 $m_{tc} = 1$, 매칭되지 않는다면 $m_{tc} = 0$ 이다.

다음의 (2.3)과 (2.4)는 매칭 시 제약조건을 나타내며, 쌍들을 형성할 때 적용되는 매칭 제약 조건은 다음과 같고,

$$\begin{aligned} \sum_{c \in C} m_{tc} &\leq 1, \quad t \in T, \\ \sum_{t \in T} m_{tc} &\leq 1, \quad c \in C. \end{aligned} \quad (2.3)$$

유사한 공변량 분포를 가지도록 1 : 1로 매칭된 표본에 가하는 공변량 균형의 제약 조건은 다음과 같다 (Visconti와 Zubizarreta, 2018).

$$\left| \sum_{t \in T} \sum_{c \in C} m_{tc} (f_k(x_{tp}) - f_k(x_{cp})) \right| \leq \varepsilon_p \sum_{t \in T} \sum_{c \in C} m_{tc}. \quad (2.4)$$

여기서 $\mathcal{P} = \{1, \dots, p\}$ 는 관찰된 공변량들의 지표이며, $x_t = \{x_{t1}, \dots, x_{tp}\}$ 와 $x_c = \{x_{c1}, \dots, x_{cp}\}$ 는 각각 처리군의 개체 t 와 대조군의 개체 c 에 대한 공변량들의 벡터이다. f_k 는 각 균형 조건 $k \in \mathcal{K}$ 에 대하여 관찰된 공변량들을 변형한 적절한 함수를 나타내며, $\varepsilon_p \geq 0$ 는 용인 수준(tolerance)이다. 이는 모두 연구자에 의해서 정의되는 값이다.

최적 매칭(optimal matching)

최적 매칭(optimal matching)은 처리군과 대조군 내의 개체를 매칭하는 순서와 관련된 매칭 방법이다. 각 그룹의 거리가 가장 가까운 쌍부터 차례대로 매칭하는 최근접 이웃 매칭(nearest neighbor matching)과 달리 최적 매칭은 매칭된 그룹의 거리 합을 최소화하는 방식으로 작동한다. 그러므로 한번 매칭이 이루어진 후에도 그 결정을 재고할 수 있다. 최적 매칭이 최근접 이웃 매칭보다 공변량 균형이 개선된 그룹을 형성하는 것은 아니지만, 개별 쌍의 거리를 줄인다는 측면에서 더 나은 성능을 보인다.

두 그룹 간 관찰된 공변량의 균형을 확인할 때 가장 널리 이용되는 방법은 표준화된 평균 차의 절댓값(absolute standardized mean difference; absolute SMD)이다.

$$\text{abs.SMD} = \frac{|\bar{X}_{Tk} - \bar{X}_{Ck}|}{\sqrt{(s_{Tk}^2 + s_{Ck}^2)/2}}.$$

공변량 X_k 에 대하여, \bar{X}_{Tk} , \bar{X}_{Ck} 는 각각 처리군과 대조군의 평균이며, s_{Tk}^2 , s_{Ck}^2 는 각 그룹의 표본 분산이다. 매칭 후의 표준화된 평균 차의 절댓값 계산 시에도 분모는 동일하며, 분자 부분만 $|\bar{X}_{Tk} - \bar{X}_{Ck}^{\text{after}}|$ 가 된다. 매칭을 통해 표준화된 평균 차의 절댓값을 작게 만들어주는 것이 목표이며, 합의된 기준은 없지만 absolute SMD ≤ 0.1 이 선호된다 (Rosenbaum와 Rubin, 1985).

또한 개별 공변량의 균형 확인 시 표본 크기에 대한 정보를 통합하는 가설 검정과 p -value를 이용할 수 있다. 표준화된 평균 차의 절댓값은 t -분포를 따르므로 (Park, 2021), $H_0 : \mu_{Tk} = \mu_{Ck}$ 를 통하여 k 번째 교란 변수의 평균이 동일한지에 대한 가설 검정을 시행할 수 있다. 만약 귀무 가설을 기각하지 못한다면, 처리군과 대조군의 개별 교란 변수의 평균이 동일하다고 할 수 있다.

두 그룹 간의 포괄적인 균형을 측정할 때 Iacus 등 (2011)이 제안한 L_1 통계량 $L_1(f, g)$ 을 이용할 수 있다.

$$L_1(f, g) = \frac{1}{2} \sum_{\ell_1 \dots \ell_k} |f_{\ell_1 \dots \ell_k} - g_{\ell_1 \dots \ell_k}|.$$

여기서 f 와 g 는 각각 처리군과 대조군의 개체들에 대한 상대빈도분포를 나타낸다. $f_{\ell_1 \dots \ell_k}$ 는 다차원 교차표(cross-tabulation)의 좌표가 $\ell_1 \dots \ell_k$ 인 셀에 속한 개체들의 상대빈도이며, $g_{\ell_1 \dots \ell_k}$ 의 경우에도 동일하다. $0 \leq L_1(f, g) \leq 1$ 이고, $L_1(f, g)$ 값이 0에 가까울 수록 두 그룹 간 포괄적인 균형이 잘 맞춰졌다고 할 수 있다. 완전히 포괄적인 균형을 이룰 때 $L_1(f, g) = 0$ 이며, 완전 불균형일 때 $L_1(f, g) = 1$ 이다. 매칭된 표본의 처리군과 대조군 상대빈도를 f^m 과 g^m 라고 할 때, 좋은 매칭은 L_1 통계량을 감소시킬 것이다. 즉, $L_1(f^m, g^m) \leq L_1(f, g)$ 이다. $L_1(f, g)$ 을 이용하기 위해서는 연속형 변수를 범주화하는 과정이 필요하다. 예를 들어, 가구 소득 변수의 경우 10분위로 분류된 범주형 변수로 변환 가능하다. 그리고 처리군과 대조군 각각 범주화된 변수들로 $X_1 \times \dots \times X_k$ 와 같이 교차표를 생성한다. 다음으로, 처리군 $f_{\ell_1 \dots \ell_k}$ 와 대조군 $g_{\ell_1 \dots \ell_k}$ 개체들에 대해 k -차원 상대 빈도를 기록한다. 전체적인 분포의 불균형 측정은 모든 셀 값들의 차이의 절댓값의 형태로 나타난다 (Blackwell 등, 2009).

매칭은 교란 변수의 영향을 통제하기 위하여 데이터 전처리 단계에서 사용되는 방법이다. 그러므로 결과 분석 단계에서 추정량을 구하는 방법을 필요로 한다. 매칭 후 처리군과 대조군 간의 평균 비교를 통해 인과 효과를 추정할 수도 있지만, 매칭 후 모수적 모형을 이용하면 인과 효과의 추정량을 더욱 개선할 수 있다 (Ho 등, 2007). 매칭 후 모수적 모형을 이용하여 인과효과 추정 시 장점은 다음과 같다. 연구자는 회귀 분석 등의 친숙한 모수적 분석 방법론들을 이용할 수 있다. 또한 매칭을 통해 처치 변수와 교란 변수 간의 관계를 끊어 내었으므로 이후 모수적 모형을 기반으로 인과효과 추정 시 모형의 선택과 설계에 덜 의존한다. 그리고 교란 변수 영향의 통제는 대부분 비모수적으로 행해졌으므로 원본 데이터를 이용하여 모수적 분석을 시행했을 때와 비교하여 인과효과 편향의 가능성이 크게 줄어든다 (Ho 등, 2007).

위에서 설명한 회귀 모형을 이용한 조정, 가중치 기법, 그리고 매칭 방법들을 이용하여 도시공원의 면적이 유아의 행복감에 미치는 영향에 대한 인과효과를 분석할 것이다.

Table 1: Summary statistics

Covariates		처리군 ($N_1 = 441$)	대조군 ($N_0 = 512$)
**유아 행복감	Mean (SD)	26.091 (2.563)	25.979 (2.498)
유아 문해 및 언어능력	Mean (SD)	53.971 (11.137)	53.246 (11.979)
*유아 자아존중감	Mean (SD)	59.968 (6.436)	60.562 (7.140)
*유아 사회적유능감 (주장성)	Mean (SD)	12.49 (2.135)	12.26 (2.284)
아버지 양육스트레스	Mean (SD)	26.011 (6.692)	26.041 (6.451)
*아버지 행복감	Mean (SD)	21.138 (4.202)	21.479 (3.910)
*아버지 애정적 부모간양육행동	Mean (SD)	20.556 (4.334)	21.053 (3.913)
아버지 최종 학력	고등학교 졸업	29%	25%
	전문대 졸업	24%	20%
	대학교 졸업	39%	45%
	대학원 졸업	8%	10%
*어머니 양육효능감	Mean (SD)	49.150 (3.618)	48.988 (4.084)
*어머니 애정적 부모간양육행동	Mean (SD)	22.12 (4.136)	22.27 (3.649)
*어머니 행복감	Mean (SD)	21.054 (4.241)	21.008 (4.186)
*어머니 통제적 양육행동	Mean (SD)	20.719 (2.877)	20.938 (3.048)
*어머니 통합적 양육행동	Mean (SD)	16.490 (3.299)	16.719 (3.237)
어머니 최종 학력	고등학교 졸업	28%	30%
	전문대 졸업	32%	25%
	대학교 졸업	35%	40%
	대학원 졸업	5%	5%
*교사 유아의 기관선호도	Mean (SD)	8.442 (1.065)	8.396 (1.128)
*월별 아파트 매매가격 평균	Mean (SD)	187029.5 (79591.08)	275061.3 (163358.60)
가구 월 평균 소득	Mean (SD)	447.141 (191.657)	437.559 (170.110)
수도권/비수도권	수도권	30%	51%
	비수도권	70%	49%

**outcome, *Additional predictors for the regression model

3. 자료 설명

3.1. 한국아동패널 자료

본 분석에서는 육아정책연구소에서 진행한 한국아동패널(PSKC)의 7차 (2014) 자료를 이용하였다. 이는 만 6세 유아를 대상으로 수집된 자료이다. 전체 자료의 관측값 수는 2,150개이며, 변수의 수는 1,425개이다. 본 자료의 변수는 크게 유아 변인, 어머니 변인, 아버지 변인, 그리고 교사 변인으로 분류된다. 총 1,425개의 변수 중 유아 행복감의 관련 변인을 연구한 Lee (2017)를 바탕으로 유아 행복감과 이에 영향을 미치는 변수들을 선정하였다. 선정된 변수는 총 15개이며, 결측값을 제외한 관측값의 수는 953개이다.

3.2. 추가 자료

치치 변수 생성 시 각 지방자치단체에서 제공하는 전국도시공원정보표준자료와 행정안전부 제공의 행정구역(시군구)별 주민등록 인구현황 자료 (2014)를 활용하였다. 전국도시공원정보표준데이터는 공원 명, 공원 구분, 공원면적, 소재지 주소 등을 포함한 전국 도시공원 정보 자료이며, 변수는 총 19개이고 관측값 수는 18,146이다. 그리고 행정구역별 주민등록 인구현황 자료는 주민등록이 되어 있는 자의 수, 남녀, 세대, 그리고 연령별 현황에 대한 자료이며, 총 관측치의 수는 279개이다. 회귀 모형을 이용한 조정(regression adjustment)을 통해

Table 2: Results from regression adjustment

	ATE	SE	95% Lower bound	95% Upper bound
Regression adjustment	0.096	0.156	-0.210	0.402

인과효과 추정 시 한국부동산원에서 제공하는 아파트 매매가격 자료 (2014)를 이용하여 추가적인 독립 변수를 생성하였다. 이는 매월 전국의 행정 구역별 아파트 가격 동향을 조사한 자료이다. 전체 관측치의 수는 227개이며, 17개의 관련 변수를 포함하고 있다. 행정 구역의 월별 자료가 비슷한 추세를 보이므로, 2014년도의 행정구역별 아파트 매매가격의 월평균 변수를 이용하였다.

3.3. 변수 정의

처치 변수의 할당 기준은 유아가 거주하는 행정구역 (시군구) 내 도시공원 면적이 충분한지의 여부이다. 각 행정구역의 도시공원 면적이 동일할지라도 인구밀도에 따라 도시공원에 대한 거주자의 접근성이 달라지므로, 비교 적도르써 1인당 도시공원 면적(m^2 /인)을 이용하였다. 표본의 1인당 도시공원 면적 분포는 web supplementary material (<https://github.com/lit777/UrbanPark>에서 다운로드 가능)의 Figure A.1에서 확인할 수 있다.

본 분석의 처치 변수 할당 기준은 중앙값($7.7m^2$)이다. 처리군 ($>7.7m^2$)과 대조군 ($\leq 7.7m^2$)에 속한 관측치는 각각 441개, 512개이다. 다른 환경에 노출된 처리군과 대조군 비교 시 은닉 편향(hidden biases)에 덜 민감한 결과를 보이는 경향이 있다 (Rosenbaum, 2004). 따라서, 그룹 간 차이를 증가시키는 처치 변수 할당 기준을 적용한 민감도 분석을 진행하였다. 처치 변수 할당 시 표본의 1인당 도시공원 면적을 4분위로 나눈 뒤 4분위에 해당하면 처리군, 1분위에 해당하면 대조군으로 분류하였다. 처리군과 대조군에 속한 관측치는 각각 233개, 255개이며, 2-3분위에 해당하는 관측치(465개)는 표본에서 제외하였다.

반응 변수는 유아의 행복감이며, 이는 유아의 주관적 행복감, 어머니가 응답한 유아의 행복감, 그리고 아버지가 응답한 유아의 행복감을 모두 합한 값이다 (Lee, 2017). 해당 변수는 한국아동패널 자료에서 추출되었다. 교란 변수를 선택하기 위해 선행 연구를 바탕으로 유향비순환그래프(directed acyclic graph; DAG)를 사용하였다. R 패키지 dagitty를 이용하여 시각화를 진행하였다 (Textor 등, 2016). 시각화된 유향비순환그래프와 이를 도출하는 방법은 web supplementary material에서 자세히 설명하였다. 유향비순환그래프를 통해서 선택된 교란 변수는 수도권/비수도권 변수, 가구 월 평균 소득, 아버지 최종 학력, 어머니 최종 학력, 유아 문해 및 언어 능력, 그리고 아버지 양육 스트레스 변수이다. Table 1은 본 분석의 처리군과 대조군을 기준으로 한 반응 변수와 교란 변수, 그리고 추정의 정확도(precision)를 높이기 위하여 회귀 모형에 추가한 독립 변수들의 요약 통계량을 나타낸다.

4. 분석 결과

4.1. 본 분석

4.1.1. 회귀 모형을 이용한 조정 분석 결과

회귀 모형을 이용한 조정을 통해 ATE 추정 시 상수 인과효과를 가정하였으며, 모형의 정확도를 높이기 위하여 교란 변수 외의 반응 변수에 영향을 미치는 변수들을 독립 변수로 사용하였다. Table 1을 통해 추가적으로 사용한 독립 변수들을 확인할 수 있다. Table 2는 회귀 모형을 이용한 조정의 분석 결과이다. ATE는 0.096이며, 표준 오차는 0.156이다. 그리고 95% 신뢰 구간은 $[-0.210, 0.402]$ 으로 나타났다. 점 추정 결과 유아의 거주지에 도시공원 면적이 충분하다면 유아의 행복감은 더 높다. 하지만 95% 신뢰 구간이 0을 포함하고 있으므로 유의미한 인과효과가 존재한다고 할 수 없다.

Table 3: Results from the weighting methods based on propensity score

	ATE	SE	95% Lower bound	95% Upper bound
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

Table 4: Results from the weighting methods based on CBPS

	ATE	SE	95% Lower bound	95% Upper bound
IPW	-1.008	0.622	-2.361	0.078
SIPW	0.057	0.168	-0.275	0.382
DR	0.090	0.158	-0.213	0.407

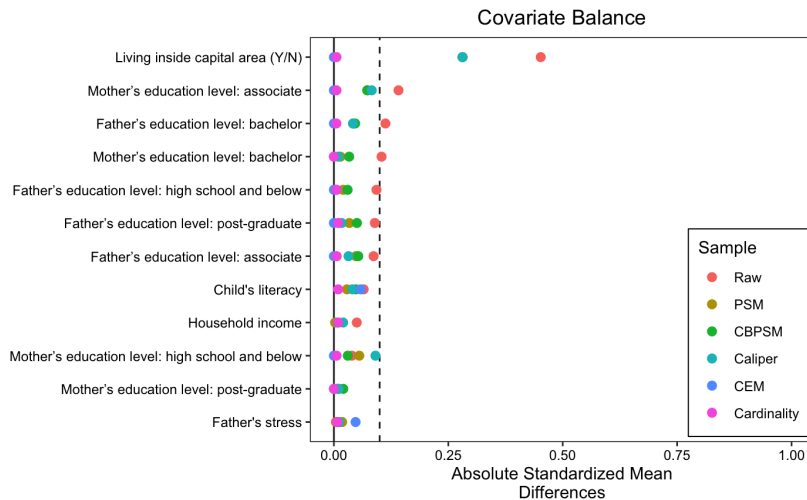


Figure 1: Absolute SMD before and after matching.

4.1.2. 가중치 기법 분석 결과

성향 점수 추정 시에는 로지스틱 회귀 분석을 이용하였다. Web supplementary material의 Figure C.1에서 처리군과 대조군 간의 성향 점수/CBPS의 히스토그램을 확인할 수 있다. 이를 통해 성향 점수와 CBPS 모두 전 구간에서 그룹 간 분포가 겹친다는 것을 알 수 있다.

Table 3은 성향 점수를 이용한 가중치 기법 분석 결과를 나타낸다. 역확률 가중치 기법(inverse probability weighting; IPW), 안정화된 역확률 가중치 기법(stabilized inverse probability weighting; SIPW), 그리고 이중강건 추정기법(doubly-robust estimation; DR)을 이용하였으며, 표준 오차와 신뢰 구간 추정 시 붓스트랩을 사용하였다. 세 방법을 통한 ATE 점추정 값은 각각 0.040, 0.047, 0.090으로 일관된 결과를 나타냈다. 하지만 세 방법의 95% 신뢰구간이 모두 0을 포함하므로, 도시공원의 면적과 유아 행복감 간의 유의미한 인과효과가 없다고 해석할 수 있다.

Table 4는 CBPS를 이용한 가중치 기법 분석 결과이다. 위와 동일한 분석 방법을 이용하였다. 역확률 가중치 기법(IPW)의 점 추정 결과, ATE 값이 -1.008로 다른 두 방법과 상반된 결과를 보였다. 또한, 상대적으로 큰 표준 오차(0.622) 값을 보였다. 이는 극단적인 가중치에 영향을 받았기 때문이다. 반면 안정화된 역확률 가중

Table 5: Covariate balance before matching

Group		처리군	대조군	Two sample <i>t</i> -test (<i>p</i> -value)
L_1		0.95		
n		441	512	
Covariates mean (standard deviation)	가구 월 평균 소득	447.14 (191.66)	437.56 (170.11)	0.8877
	유아 문해 및 언어능력	53.97 (11.14)	53.25 (11.98)	0.90
	아버지 양육 스트레스	26.01 (6.69)	26.04 (6.45)	0.91
	수도권/비수도권.0	0.70 (0.46)	0.49 (0.50)	0.94
	수도권/비수도권.1	0.30 (0.46)	0.51 (0.50)	0.94
	아버지 최종 학력.4	0.29 (0.45)	0.25 (0.43)	0.94
	아버지 최종 학력.5	0.24 (0.43)	0.20 (0.40)	0.93
	아버지 최종 학력.6	0.39 (0.49)	0.45 (0.50)	0.94
	아버지 최종 학력.7	0.08 (0.27)	0.10 (0.30)	0.90
	어머니 최종 학력.4	0.28 (0.45)	0.30 (0.46)	0.94
	어머니 최종 학력.5	0.32 (0.47)	0.25 (0.44)	0.94
	어머니 최종 학력.6	0.35 (0.48)	0.40 (0.49)	1.00
	어머니 최종 학력.7	0.05 (0.22)	0.05 (0.22)	1.00

4-7: 고등학교 졸업, 전문대 졸업, 대학교 졸업, 대학원 졸업 순
4-7: 고등학교 졸업, 전문대 졸업, 대학교 졸업, 대학원 졸업 순

치 기법(SIPW)과 이중 강건 추정기법(DR)의 결과는 ATE 값이 각각 0.057, 0.090으로 일관적인 값을 보였다. 또한 역확률 가중치 기법(IPW)과 비교했을 때 신뢰 구간의 길이가 상대적으로 짧았다. 하지만 성향 점수를 이용한 분석 결과와 동일하게 모든 방법의 95% 신뢰 구간이 0을 포함하므로 두 변수 간 유의미한 인과효과는 존재하지 않는다.

4.1.3. 매칭 분석 결과

매칭 방법 중 성향 점수 매칭(propensity score matching; PSM), 공변량 균형 성향 점수 매칭(covariate balancing propensity score; CBPSM), 성향 점수 범위짜깁기 매칭(propensity score caliper; Caliper), 성긴 정확 매칭(coarsened exact matching; CEM), 그리고 카디널리티(Cardinality) 매칭을 이용하였다. 분석 시 R 패키지 MatchIt를 활용하였다 (Stuart 등, 2011). Figure 1는 12개의 공변량에 대한 매칭 전과 후의 표준화된 평균 차의 절댓값(absolute SMD)을 나타낸다. 일반적으로 공변량 균형 평가 시 표준화된 평균 차의 절댓값이 0.1 이하이면 균형적이라고 판단한다. 이는 플롯의 점선으로 표시된다. 또한 Table 5와 Table 6은 매칭 전/후의 처리군과 대조군 간 공변량 균형을 나타낸다. 개별 공변량 균형은 표준화된 평균 차의 절댓값과 *t*-검정, 그리고 공변량 결합 분포 균형은 L_1 통계량을 통해 평가하였다.

매칭 전 처리군과 대조군 내 관측값 수는 각각 441개, 512개이며, Figure 1을 통해 매칭 전 대부분의 단일 공변량은 균형적임을 알 수 있다. 하지만 Table 5를 통해 확인한 L_1 통계량은 0.95로써, 두 그룹 간 결합 분포는 완전 불균형에 가까웠다. 성향 점수를 이용하여 1 : 1 매칭(PSM, CBPSM, Caliper)을 실시한 결과 441쌍이 매칭되었다. 매칭의 순서 결정 시에는 최적 매칭(optimal matching)과 비복원추출방법을 이용하였다. 세 매칭의 질에는 차이가 없었으며, 매칭 전 단일 분포가 매우 불균형했던 수도권 거주 여부 변수(living inside capital area (Y/N))의 경우 매칭 후에도 불균형함이 남아있었다. 성긴 정확 매칭(CEM) 이용 시 연속형 변수들을 구간화하였다. 가구 월평균소득 변수는 10분위, 유아의 문해 및 언어능력 변수와 아버지 양육스트레스 변수는 4분위로 구간화하였다. 그 결과 모든 공변량의 단일 분포가 균형을 이뤘지만, 가장 적은 수인 101쌍이 매칭되었다. 카디널리티(cardinality) 매칭의 경우 성향 점수를 이용한 매칭보다 적은 385쌍이 매칭되었지만,

Table 6: Covariate balance after matching

Method	PSM	CBPSM	Caliper	CEM	Cardinality	
Matched pairs	441	441	441	101	385	
L_1	0.95	0.95	0.94	0.82	0.96	
Covariates differences (<i>p</i> -value)	가구 월 평균 소득	-0.00 (0.97)	0.01 (0.89)	0.02 (0.76)	0.01 (0.95)	0.01 (0.89)
	유아 문해 및 언어능력	0.03 (0.68)	0.05 (0.49)	0.04 (0.55)	-0.06 (0.69)	0.01 (0.90)
	아버지 양육 스트레스	-0.02 (0.78)	0.01 (0.86)	-0.01 (0.86)	0.05 (0.71)	-0.01 (0.91)
	수도권/비수도권.0	0.28 (0.00)	0.28 (0.00)	0.28 (0.00)	0.00 (1.00)	0.01 (0.94)
	수도권/비수도권.1	-0.28 (0.00)	-0.28 (0.00)	-0.28 (0.00)	0.00 (1.00)	-0.01 (0.94)
	아버지 최종 학력.4	0.02 (0.77)	0.03 (0.65)	0.01 (0.94)	0.00 (1.00)	0.01 (0.94)
	아버지 최종 학력.5	0.05 (0.47)	0.05 (0.42)	0.03 (0.63)	0.00 (1.00)	0.01 (0.93)
	아버지 최종 학력.6	-0.04 (0.54)	-0.05 (0.49)	-0.04 (0.54)	0.00 (1.00)	-0.01 (0.94)
	아버지 최종 학력.7	-0.03 (0.63)	-0.05 (0.47)	0.02 (0.8)	0.00 (1.00)	-0.01 (0.90)
	어머니 최종 학력.4	-0.06 (0.42)	-0.03 (0.66)	-0.09 (0.19)	0.00 (1.00)	-0.01 (0.94)
	어머니 최종 학력.5	0.07 (0.27)	0.07 (0.27)	0.08 (0.21)	0.00 (1.00)	0.01 (0.94)
	어머니 최종 학력.6	-0.01 (0.83)	-0.03 (0.62)	0.01 (0.89)	0.00 (1.00)	0.00 (1.00)
	어머니 최종 학력.7	-0.01 (0.88)	-0.02 (0.77)	-0.01 (0.88)	0.00 (1.00)	0.00 (1.00)

Table 7: Results from the matching methods based on propensity score

	ATT	SE	95% Lower bound	95% Upper bound
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

Table 8: Results from regression adjustment

	ATE	SE	95% Lower bound	95% Upper bound
Regression adjustment	-0.069	0.25	-0.560	0.422

모든 공변량의 표준화된 평균 차의 절댓값이 0.1 이하였다. 또한 매칭 후 *t*-검정을 실시하였을 때 처리군과 대조군의 개별 교란 변수의 평균이 동일하다는 가설을 지지하였다. 공변량의 결합 분포를 기준으로 매칭의 질을 평가할 시, 성긴 정확 매칭(CEM)의 L_1 통계량은 0.82로 다른 매칭 방법들보다 개선된 결과를 보였다 (Table 6 참조).

Table 7은 매칭 후의 표본을 이용한 인과효과 추정 결과이다. 매칭 후에도 남아있는 그룹 간 공변량 균형을 잡아주기 위하여 회귀 분석을 실시하였다. 성긴 정확 매칭(CEM)을 제외한 모든 매칭 방법의 ATT 추정값은 양수로 나타났다. 하지만 구간 추정 결과, 모든 방법이 95% 신뢰 구간의 하한과 상한 사이에 0을 포함하고 있으므로, 유의미한 인과효과는 존재하지 않는다고 해석할 수 있다.

4.2. 민감도 분석

본 분석과 동일한 방법을 이용하여 민감도 분석을 진행하였다. 3.3절에서 설명하였듯이 표본의 처치 변수를 4분위로 나누어 4분위에 해당하면 처리군, 1분위에 해당하면 대조군으로 구분하였다. Table 8은 회귀 모형을 이용한 교란 변수 조정의 분석 결과이다. 점 추정 결과 음의 인과효과가 존재한다. 하지만 95% 신뢰 구간은

Table 9: Results from the weighting methods based on propensity score

	ATE	SE	95% Lower bound	95% Upper bound
IPW	-0.628	1.030	-2.532	1.507
SIPW	-0.107	0.303	-0.745	0.443
DR	-0.060	0.302	-0.700	0.484

Table 10: Results from the weighting methods based on CBPS

	ATE	SE	95% Lower bound	95% Upper bound
IPW	-1.592	0.848	-3.541	-0.217
SIPW	-0.097	0.289	-0.696	0.435
DR	-0.070	0.282	-0.671	0.432

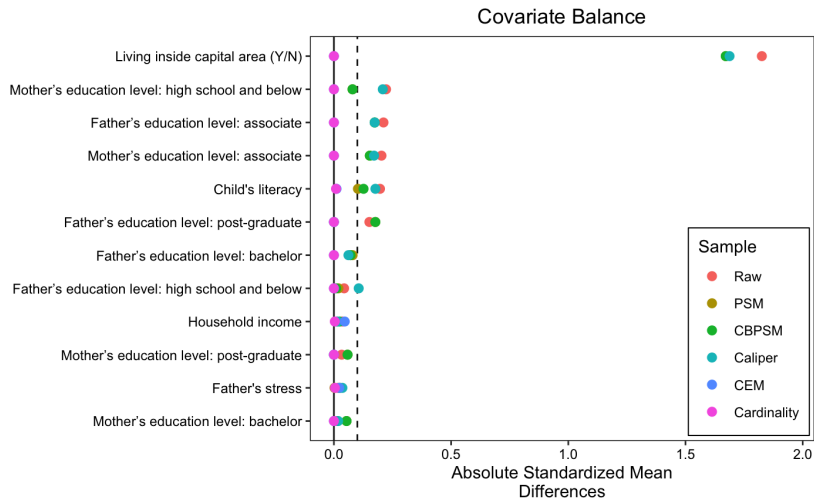


Figure 2: Absolute SMD before and after matching.

$[-0.560, 0.422]$ 으로, 신뢰 구간의 하한과 상한 사이에 0을 포함하고 있으므로 유의미한 인과효과는 존재하지 않는다. 이는 본 분석과 동일한 결과이다.

Table 9와 Table 10은 성향 점수/CBPS를 이용한 가중치 기법 분석 결과이다. 역확률 가중치 기법(IPW), 안정화된 역확률 가중치 기법(SIPW), 그리고 이중 강건 추정기법(DR)으로 추정된 ATE 값은 모두 음수로 나타났다. 두 분석 결과의 역확률 가중치 기법(IPW) 추정량의 표준 오차는 각각 1.030, 0.848으로 상대적으로 큰 값을 보였다. 이는 역확률 가중치 기법(IPW)으로 구한 추정량의 경우 극단적인 가중치에 영향을 받기 때문이다. 구간 추정 결과, 본 분석과 마찬가지로 95% 신뢰 구간의 상한과 하한 사이에 0이 포함되므로 유의미한 인과효과가 존재한다고 할 수 없다.

4.2.1. 매칭 분석 결과

Figure 2는 12개의 공변량에 대한 매칭 전과 후의 표준화된 평균 차의 절댓값을 나타낸다. 매칭 전/후 그들 간 공변량 균형에 대한 자세한 내용은 web supplementary material의 Table D.1와 Table D.2를 통해 확인할 수 있다. 본 분석과 동일한 공변량 균형 평가 기준을 적용하였다.

Table 11: Results from the matching methods based on propensity score

	ATT	SE	95% Lower bound	95% Upper bound
PSM	-0.051	0.276	-0.583	0.481
CBPSM	-0.054	0.278	-0.585	0.478
Caliper	-0.033	0.271	-0.563	0.497
CEM	-0.850	0.524	-1.930	0.230
Cardinality	-0.178	0.318	-0.773	0.417

민감도 분석의 처리군과 대조군 내 관측값 수는 각각 233개, 255개이다. Figure 2를 통해 매칭 전 매우 불균형한 단일 공변량은 수도권 거주 여부 변수(living inside capital area (Y/N))이며, 그 밖에 다섯 개의 단일 공변량이 다소 불균형적임을 알 수 있다. Web supplementary material의 Table D.1을 통해 확인할 수 있는 표본의 L_1 통계량은 0.95로써 두 그룹 간 공변량 결합 분포는 불균형했다. 성향 점수를 이용한 매칭(PSM, CBPSM, Caliper) 시 233개의 쌍들이 매칭되었다. 세 매칭의 질은 유사했으며, 매칭 후에도 여전히 몇몇 단일 공변량들은 불균형했다. 성긴 정확 매칭(CEM)은 상대적으로 적은 40개의 쌍들을 매칭하였지만, 성향 점수 매칭 방법들보다 단일 공변량의 균형을 개선하였다. 카디널리티(Cardinality) 매칭은 129개의 쌍을 매칭하였으며, 매칭 후 모든 단일 공변량이 균형을 이루었다. 공변량들의 결합 분포를 기준으로 공변량 균형을 평가하였을 때, 성긴 정확 매칭(CEM)의 L_1 통계량은 0.88로써, 이를 비교적 많이 개선하였다 (web supplementary material의 Table D.2 참조).

Table 11은 각각의 매칭 방법들을 시행하여 얻은 표본을 이용한 인과효과 추정 결과이다. 회귀 모형을 통해 ATT와 표준 오차, 그리고 95% 신뢰 구간을 추정하였다. 그 결과, 본 분석과 마찬가지로 유의미한 인과효과는 존재하지 않는다.

5. 토의

선행 연구에 의하면 도시공원을 비롯한 녹지는 유아의 정서에 긍정적인 영향을 미친다 (Balseviciene 등, 2014). 또한 자연과의 접촉과 유아의 행복감은 상관관계가 있다 (Barrera-Hernández 등, 2020). 이를 통해 도시공원과 유아 행복감 간의 인과관계를 유추해볼 수 있다. 하지만 도시공원과 유아 행복감 간의 인과관계는 지금까지 연구되지 않았다. 기존의 연구에서 잠재적인 교란 변수의 효과를 제거하지 않고 분석을 진행했기에, 이를 통해서 나타난 결과로 두 변수간의 인과관계를 유추할 수는 없다. 본 아동패널자료 연구에서는 인과효과 분석에 특화된 방법들(회귀 모형을 이용한 교란 변수 조정, 가중치 방법, 매칭 방법 등)을 이용하여 도시공원의 면적이 유아 행복감에 미치는 인과효과를 추정하였다. 모든 방법을 이용한 분석 결과, 도시공원의 면적이 유아 행복감에 미치는 인과효과는 통계적으로 유의미한 수준이 아니었다. 이는 기존의 연구에서 두 변수 간 양의 상관관계를 찾아낸 것과 다른 결론이다. 또한 유향비순환그래프(DAG)를 통해 선택한 교란 변수 중 처리군과 대조군 간 분포가 가장 불균형했던 교란 변수는 수도권 거주 여부였다. 그러므로 이 교란 변수의 효과 때문에 두 변수의 상관관계가 크게 나타났다는 것을 유추할 수 있다.

본 연구를 진행함에 있어서 측정되지 않은 다른 교란 변수는 없다는 가정(no unmeasured confounders)을 바탕으로 연구를 진행했다. 최대한 가정을 만족시키기 위해서 선행 연구에서 언급된 변수가 포함된 모든 수집 가능한 자료를 이용하였다. 하지만 자료 수집의 문제로 인하여 측정되지 않은 교란 변수를 분석에서 사용할 수 없었다는 한계점이 있다.

기존의 연구 (Balseviciene 등, 2014)에 따르면 도시공원의 면적이 아닌 도시공원에 대한 접근성(거주지에서 생활권 공원까지의 거리)이 유아 행복감에 영향을 미치는 주요 변인일 가능성이 있다. 그러므로 추후 공간 데이터를 이용하여 도시공원의 접근성이 유아 행복감에 미치는 인과 효과에 대한 연구를 진행할 수도 있을

것이다.

References

- Austin PC (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies, *Multivariate Behavioral Research*, **46**, 399–424.
- Balseviciene B, Sinkariova L, Grazuleviciene R, Andrusaityte S, Uzdanaviciute I, Dedele A, and Nieuwenhuijsen MJ (2014). Impact of residential greenness on preschool children’s emotional and behavioral problems, *International Journal of Environmental Research and Public Health*, **11**, 6757–6770.
- Barrera-Hernández LF, Sotelo-Castillo MA, Echeverría-Castro SB, and Tapia-Fonllem CO (2020). Connectedness to nature: Its impact on sustainable behaviors and happiness in children, *Frontiers in Psychology*, **11**, 276.
- Blackwell M, Iacus S, King G, and Porro G (2009). Cem: Coarsened exact matching in stata, *The Stata Journal*, **9**, 524–546.
- Boldemann C, Blennow M, Dal H, Mårtensson F, Raustorp A, Yuen K, and Wester U (2006). Impact of preschool environment upon children’s physical activity and sun exposure, *Preventive Medicine*, **42**, 301–308.
- Chiesura A (2004). The role of urban parks for the sustainable city, *Landscape and Urban Planning*, **68**, 129–138.
- Cohen DA, Ashwood JS, Scott MM *et al.* (2006). Public parks and physical activity among adolescent girls, *Pediatrics*, **118**, e1381–e1389.
- Compton WC (2005). *An Introduction to Positive Psychology*, Australia ; Belmont, CA, Thomson/Wadsworth.
- Epstein LH, Raja S, Gold SS, Paluch RA, Pak Y, and Roemmich JN (2006). Reducing sedentary behavior: The relationship between park area and the physical activity of youth, *Psychological Science*, **17**, 654–659.
- Faber Taylor A and Kuo FE (2011). Could exposure to everyday green spaces help treat ADHD? Evidence from children’s play settings, *Applied Psychology: Health and Well-Being*, **3**, 281–303.
- Fredrickson BL (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions, *American Psychologist*, **56**, 218.
- Funk MJ, Westreich D, Wiesen C, Stürmer T, Brookhart MA, and Davidian M (2011). Doubly robust estimation of causal effects, *American Journal of Epidemiology*, **173**, 761–767.
- Gwon GN and Sung MY (2010). The influences of young children’s happiness on behavior problems and interpersonal problem solving strategies, *Korean Journal of Human Ecology*, **19**, 257–270.
- Hainmueller J (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies, *Political Analysis*, **20**, 25–46.
- Hartig T, Mitchell R, De Vries S, and Frumkin H (2014). Nature and health. Annual review of public health, *Annual Review of Public Health*, **35**, 207–228.
- Hernán MÁ, Brumback B, and Robins JM (2000). Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men, *Epidemiology*, **11**, 561–570.
- Ho D, Imai K, King G, and Stuart E (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference, *Political Analysis*, **15**, 199–236.
- Holland Paul W (1986). Statistics and causal inference, *Journal of the American Statistical Association*, **81**, 945–960.
- Houlden V, Weich S, Porto de Albuquerque J, Jarvis S, and Rees K (2018). The relationship between greenspace and the mental wellbeing of adults: A systematic review, *PLOS ONE*, **13**, e0203000.

- Iacus SM, King G, and Porro G (2011). Multivariate matching methods that are monotonic imbalance bounding, *Journal of the American Statistical Association*, **106**, 345–361.
- Iacus SM, King G, and Porro G (2012). Causal inference without balance checking: Coarsened exact matching, *Political Analysis*, **20**, 1–24.
- Imai K and Ratkovic M (2014). Covariate balancing propensity score, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **76**, 243–263.
- Imbens GW (2004). Nonparametric estimation of average treatment effects under exogeneity: A review, *Review of Economics and Statistics*, **86**, 4–29.
- Kaczynski AT and Henderson KA (2007). Environmental correlates of physical activity: A review of evidence about parks and recreation, *Leisure Sciences*, **29**, 315–354.
- Kaczynski AT, Potwarka LR, and Saelens BE (2008). Association of park size, distance, and features with physical activity in neighborhood parks, *American Journal of Public Health*, **98**, 1451–1456.
- Kang JD and Schafer JL (2007). Demystifying double robustness: A comparison of alternative strategies for estimating a population mean from incomplete data, *Statistical Science*, **22**, 523–539.
- Kim D and Jin J (2018). Does happiness data say urban parks are worth it?, *Landscape and Urban Planning*, **178**, 1–11.
- Kim J and Kaplan R (2004). Physical and psychological factors in sense of community: New urbanist Kentlands and nearby Orchard village, *Environment and Behavior*, **36**, 313–340.
- Kim SB (2011). A study on the influence of ecological factors on children's happiness (Master's thesis), Chongshin University, Seoul.
- Korea Real Estate Board (2014). Available from: <https://kosis.kr/search/search.do?query=아파트매매>
- Lee GS (2017). Exploring variables influencing the happiness of young children, *Journal of Future Early Childhood Education*, **24**, 305–324.
- Lee KM (2016). Focus group interview for exploring constructive components and valid assessment method of happiness of young children, *Journal of Early Childhood Education*, **36**, 499–514.
- Little RJ and Rubin DB (2019). *Statistical Analysis with Missing Data*, John Wiley & Sons, New Jersey.
- Lunceford JK and Davidian M (2004). Stratification and weighting via the propensity score in estimation of causal treatment effects: A comparative study, *Statistics in Medicine*, **23**, 2937–2960.
- Mårtensson F, Boldemann C, Söderström M, Blennow M, Englund JE, and Grahn P (2009). Outdoor environmental assessment of attention promoting settings for preschool children, *Health & Place*, **15**, 1149–1157.
- Ministry of Education (2022). Available from: <https://www.law.go.kr/법령/유아교육법>
- Ministry of the Interior and Safety (2014). Available from: <https://jumin.mois.go.kr>
- Ministry of Land, Infrastructure and Transport (2021). Available from: <https://www.law.go.kr/법령/도시공원및 녹지등에관한법률>
- Ministry of Land, Infrastructure and Transport (2022). Available from: <https://www.data.go.kr/data/15012890/standard.do>
- OECD (2014). *Education at a Glance 2014: OECD Indicators*, OECD Publishing, Paris.
- Park SD (2021). A causal inference approach to the study of the effect of bereavement on mental health (Master's thesis), Sungkyunkwan University, Seoul.
- PSKC (2014). Available from: <https://panel.kicce.re.kr/pskc/index.do>
- Readdick CA and Schaller GR (2005). Summer camp and self-esteem of school-age inner-city children, *Perceptual and Motor Skills*, **101**, 121–130.

- Robins JM, Rotnitzky A, and Zhao LP (1994). Estimation of regression coefficients when some regressors are not always observed, *Journal of the American Statistical Association*, **89**, 846–866.
- Roemmich JN, Epstein LH, Raja S, Yin L, Robinson J, and Winiewicz D (2006). Association of access to parks and recreational facilities with the physical activity of young children, *Preventive Medicine*, **43**, 437–441.
- Rosenbaum PR and Rubin DB (1983). The central role of the propensity score in observational studies for causal effects, *Biometrika*, **70**, 41–55.
- Rosenbaum PR and Rubin DB (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score, *The American Statistician*, **39**, 33–38.
- Rosenbaum PR (1987). Model-based direct adjustment, *Journal of the American Statistical Association*, **82**, 387–394.
- Rosenbaum PR (2002). Overt bias in observational studies, *Observational Studies* (pp. 71–104), Springer, New York.
- Rosenbaum PR (2004). Design sensitivity in observational studies, *Biometrika*, **91**, 153–164.
- Rubin DB (1980). Randomization analysis of experimental data: The Fisher randomization test comment, *Journal of the American Statistical Association*, **75**, 591–593.
- Rubin DB and Thomas N (2000). Combining propensity score matching with additional adjustments for prognostic covariates, *Journal of the American Statistical Association*, **95**, 573–585.
- Stuart EA (2010). Matching methods for causal inference: A review and a look forward, *Statistical Science: A Review Journal of the Institute of Mathematical Statistics*, **25**, 1–21.
- Stuart EA, King G, Imai K, and Ho D (2011). MatchIt: Nonparametric preprocessing for parametric causal inference, *Journal of Statistical Software*, **42**, 1–28.
- Textor J, Van der Zander B, Gilthorpe MS, Liškiewicz M, and Ellison GT (2016). Robust causal inference using directed acyclic graphs: The R-package ‘dagitty’, *International Journal of Epidemiology*, **45**, 1887–1894.
- Visconti G and Zubizarreta JR (2018). Handling limited overlap in observational studies with cardinality matching, *Observational Studies*, **4**, 217–249.
- Wyss R, Ellis AR, Brookhart MA, Girman CJ, Jonsson Funk M, LoCasale R, and Stürmer T (2014). The role of prediction modeling in propensity score estimation: An evaluation of logistic regression, bCART, and the covariate-balancing propensity score, *American Journal of Epidemiology*, **180**, 645–655.
- Zubizarreta JR, Paredes RD, and Rosenbaum PR (2014). 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**, 204–231.

Received November 2, 2022; Revised November 11, 2022; Accepted November 13, 2022

도시공원 면적이 유아 행복감에 미치는 영향에 대한 인과관계 연구

권나연^a, 김찬민^{1,a}

^a성균관대학교 통계학과

요약

기존의 많은 연구에서 도시공원을 비롯한 녹지와 유아 행복감 간에 유의미한 상관관계를 도출했다. 또한 이를 통해 간접적으로 도시공원의 면적/근접성이 유아기의 행복감 증진에 효과가 있을 것이라 유추하였다. 하지만 관찰된 자료를 통한 인과효과 추정은 교란 변수의 적절한 조정을 필요로 하고, 이런 관점에서 도시공원의 면적과 유아 행복감의 인과관계는 명확히 밝혀지지 않았다고 할 수 있다. 본 연구에서는 한국아동패널 자료를 이용하여 도시공원의 면적이 유아 행복감에 미치는 영향에 대한 인과효과를 추정하였다. 교란 변수를 조정하기 위한 방법으로 회귀 모형을 이용한 조정(regression adjustment), 가중치 기법(weighting), 그리고 매칭(matching) 등을 이용하였고, 각 방법들의 중요 개념을 분석 결과에 앞서 기술하였다. 교란 변수의 선택에 있어서 유향 비순환 그래프(directed acyclic graph)를 사용하였다. 분석 결과, 기존의 상관관계를 이용한 결론과는 다르게 도시공원의 면적과 유아 행복감 간에는 유의미한 인과효과가 존재하지 않았다.

주요용어: 한국아동패널, 인과효과 추정, 가중치 기법, 매칭

본 연구는 한국연구재단의 지원에 의해 수행한 연구입니다 (NRF-2020R1F1A1A01048168, NRF-2022R1F1A1062904).

¹교신저자: (03063) 서울특별시 종로구 성균관로 25-2, 성균관대학교 통계학과. E-mail: chanmin.kim@skku.edu