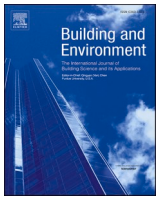
Building and Environment 207 (2022) 108401

Contents lists available at ScienceDirect

Building and Environment

journal homepage: www.elsevier.com/locate/buildenv



Rethinking the link between the availability of neighborhood PA facilities and PA behavior: A comparison between private and public housing

Ruoyu Wang [a], George Grekousis [b,c], Yi Lu [d,e,*]

[a] Institute of Geography, School of GeoSciences, University of Edinburgh, Edinburgh, UK
[b] School of Geography and Planning, Department of Urban and Regional Planning, Sun Yat-sen University, Xingang Xi Road, Guangzhou, 510275, China
[c] Guangdong Key Laboratory for Urbanization and Geo-simulation, Sun Yat-sen University, Xingang Xi Road, Guangzhou, 510275, China
[d] Department of Architecture and Civil Engineering, City University of Hong Kong, Hong Kong SAR, Hong Kong
[e] City University of Hong Kong Shenzhen Research Institute, Shenzhen, China

ARTICLE INFO

Keywords:
Built environment
Physical activity facilities
Physical activity behavior
Residential self-selection
Housing type

ABSTRACT

To examine whether residential self-selection bias influence the associations between neighborhood physical activity (PA) facilities and respondents' PA behavior. This study uses a natural experimental research design. Three waves of China Family Panel Studies data were used. The outcome variables were the weekly duration and frequency of PA. The availability of neighborhood PA facilities was measured using the presence of PA facilities and the number of PA facilities within residential neighborhoods. Multilevel regression models were used to examine the relationships between the availability of neighborhood PA facilities and respondents' PA behavior, adjusted for covariates. Associations were stratified by two groups (those living in private housing vs. public housing). Residents of private housing can choose their residential location, whereas those in public housing have little freedom to do so. Therefore, comparing these two groups can help us determine whether residential self-selection bias exists. The results show that both the presence and number of neighborhood PA facilities are positively associated with the frequency and duration of PA among residents of private housing. However, the evidence does not support an association between the availability of neighborhood PA facilities and PA behavior among residents of public housing. Residential self-selection bias may have a pronounced effect on the findings of studies of the association between neighborhood PA facilities and individuals' PA behavior. Therefore, it is important to control for residential self-selection bias when examining built environment–PA associations.

1. Introduction

The World Health Organization has recommended that adults should take physical activity (PA) of moderate intensity (i.e., cycling) for at least 150 min a week, or of vigorous intensity (i.e., running) for 75 min a week [1]. PA has an array of health benefits, such as promoting mental health [2] and lowering the risk of physical diseases such as obesity [3], cardiovascular disease [4], diabetes [5] and even some cancers [6]. However, the global prevalence of insufficient PA was 27.5% in 2016 [7], and a report in China showed that only 14.7% of adults took PA regularly in 2014 [8]. Therefore, encouraging PA has become a public health priority in China. For example, the State Council of China has declared the aims of increasing the proportion of people who regularly take PA to 37% and 40% by 2022 and 2030, respectively [9].

Awareness is mounting that the built environment influences people's PA behavior [10–21]. Built environment factors such as greenspace, land use mix, street connectivity, PA facilities and transit facilities all have proven associations with PA behavior [10,11] Among the various built environment factors, the availability of PA facilities such as football fields and gyms is directly related to PA behavior for two reasons [13,21] First, the availability of PA facilities in neighborhoods increases their accessibility to people and encourages PA [13]. Second, according to the theory of "social norms," individuals' attitudes and behavior are influenced by those of their friends or neighbors [22]. Therefore, residents of neighborhoods with a higher availability of PA facilities are more likely to be in contact with physically active friends and neighbors and to understand the health benefits of PA, which will facilitate their own willingness to partake in PA [13,21].

However, previous cross-sectional studies of environment–behavior associations face an important threat to their internal validity and

* Corresponding author. Department of Architecture and Civil Engineering, City University of Hong Kong, Hong Kong SAR, Hong Kong.
E-mail addresses: R.Wang-54@sms.ed.ac.uk (R. Wang), graikousis@mail.sysu.edu.cn (G. Grekousis), yilu24@cityu.edu.hk (Y. Lu).

https://doi.org/10.1016/j.buildenv.2021.108401
Received 4 June 2021; Received in revised form 28 August 2021; Accepted 25 September 2021
Available online 29 September 2021

causality due to residential self-selection bias [23–31]. Residential self-selection bias refers to the fact that personal attitudes and preferences regarding PA may affect both people's residential locations and their PA behavior [31]. As for the association between PA facilities and PA behaviors, residential self-selection bias refers to the situation that residents intentionally choose to live in areas with more or less PA facilities due to personal preference [24,26–28]. For example, a person who values the health benefits of PA is more likely to choose a neighborhood with more PA facilities and to have higher PA levels. Under such a circumstance, the observed PA facilities–PA behavior association would be explained by personal preferences rather than a true causal effect of PA facilities on PA behaviors.

Three approaches to mitigate residential self-selection bias have been proposed in the literature [31–40]. The first approach is to collect residents' personal preferences through questionnaires [32]. This can, in principle, control for the confounding effect of personal preferences regarding PA. However, in practice, people tend to overstate their preference for PA, which makes it difficult to accurately assess this confounder [31]. For example, Lin et al. [32] collected travel attitudes and behaviors at the same time and examined how travel attitudes may affect built environment-travel behaviors associations.

The second approach is to track the environmental changes for same people (e.g., the creation of new parks, or people relocate to different neighborhoods) and to estimate their effects on the changes in individuals' PA behavior using longitudinal data [31]. For instance, Boone-Heinonen et al. [31] used a longitudinal study in USA which was collected in 1994 and 2001 to examine how changes of neighborhood environment have influence on PA. Xie et al. [33] used two waves of longitudinal data collected in 2016 and 2019 respectively, to explore the creation of 102-km-long greenway on residents' PA levels in Wuhan, China. Nevertheless, if people still choose residential locations based on their personal preferences when relocating, longitudinal data will not be able to reduce residential self-selection bias.

Third, the natural experimental research design has been advocated to establish rigorous evidence on the association between neighborhood environments and human behavior by improving internal validity [34]. Natural experiment assigns participants to residential neighborhoods with different built environment exposure, which disentangles self-selection bias and allows researchers to study the effect of different exposure contexts on the changes of various outcomes. However, unlike randomized controlled trials (RCTs), the participant assignment and built environment interventions are naturally occurred rather than fully controlled by researchers in natural experiments. For example, Ludwig et al. [34] used data from Moving to Opportunity (MTO) to study the neighborhood effect on wellbeing. MTO project randomly assign participants living in high-poverty public housing projects into two groups, one group received housing vouchers that can be used in low-poverty areas, and the other group did not receive vouchers. Such research design allows researcher to infer the causal evidence that families who moved from high-poverty areas to low-poverty areas had lower rates of obesity and depression, higher level of happiness [34]. However, such method is expensive, labour intensive and may cause some potential ethical and political issues [34].

China's public housing system provides a rare opportunity to conduct such a large-scale quasi-natural experiment to address residential self-selection bias [32,41]. "Public housing" in China refers to housing built and owned by the government, which is usually rented to socioeconomically disadvantaged populations at an affordable rate [32,42,43]. Residents are often assigned to residential locations randomly, so they have little or no freedom to choose their location. Therefore, they can serve as a non-self-selecting group for investigating built environment–PA associations. Previous studies have used this method in the Chinese context to explore the association between neighborhood built environments and travel behavior [32,41]. For example, Zang et al. [41] found that some built environment characteristics are only associated with travel behavior for people in private housing, whereas such associations are not observed for people in public housing, indicating the existence of residential self-selection bias. In 2010, about 20% of the households accommodated in public housing projects in urban China who may not have freedom to select their residential locations, while the remaining 80% lived in private housing, who have the freedom to select their residential locations [32]. If the majority of the research samples live in private housing in China, then the study may be subjected to residential self-selection bias [32]. However, few studies have paid attention to such a context in China, and not addressing residential self-selection bias for a national representative survey data may prevent researchers from inferring causality between built environment and human behaviors in Chinese context.

Besides the vulnerability to residential self-section bias, studies examining PA facilities–PA behavior associations are often based on a single city, which limits the generalizability of the findings to other cities. This study aims to address residential self-selection bias in PA facilities–PA behavior associations using a natural experimental study design. This study extends previous research in several respects. First, the study is among the first to investigate the residential self-selection issue in the association between availability of neighborhood PA facilities and PA behavior in a densely populated Chinese context. Second, it uses national longitudinal data from a large representative sample, which greatly enhances the generalizability of our findings.

2. Methods

2.1. Study data

This study uses a dataset from three waves of the China Family Panel Studies (CFPS), specifically those conducted in 2014 (wave 2), 2016 (wave 3) and 2018 (wave 4) [https://www.isss.pku.edu.cn/cfps/index.htm]. The CFPS was conducted by the Institute of Social Science Survey (ISSS) of Peking University, China. Ethical approval for the study was granted by the Ethics Review Committee of Peking University, and all participants provided signed informed consent at the time of participation. The CFPS was a nationally representative survey that focused on the family welfare of Chinese residents, with questions relating to family relationships, health outcomes and demographic characteristics. Respondents were selected using a probability-proportional-to-size sampling technique with implicit stratification, multiple stages, multiple levels and population proportionality. The cross-sectional response rate of the survey in 2014 was 72.8%, and the tracking rate was 83.8%. We excluded data from 2010 (wave 0) and 2012 (wave 1) because our key variable (PA behavior) has only been included in the CFPS since 2014. After leaving out any observations with missing information for any individual- or neighborhood-level variables, the sample comprised 69,988 person-year records (33,837 individuals).

2.2. Outcomes

We used two measures of PA behavior in this study: the frequency of PA and duration of PA. Respondents were asked: "How many times have you taken PA in the past week?" and "For how many hours have you taken PA in the past week?"

2.3. Availability of neighborhood PA facilities

The availability of neighborhood PA facilities was measured via the number of PA facilities within a neighborhood as provided by the neighborhood section of the CFPS questionnaire. The two PA facilities-related questions were "Are there any PA facilities in the neighborhood?" and "How many PA facilities are there in the neighborhood?" Due to the CFPS's privacy policy, we could not access the respondents' home or neighborhood addresses to objectively assess the availability of PA facilities. First, we treated the number of PA facilities within a neighborhood as a binary variable to measure the presence of PA

facilities (1: number of PA facilities within the neighborhood >0; 0: number of PA facilities within the neighborhood = 0). Then, we treated the number of PA facilities within the neighborhood as a continuous variable to measure the availability of neighborhood PA facilities in terms of the number of facilities.

2.4. Using housing type to disentangle the impact of residential self-selection bias

People in different housing types (private housing vs. public housing) have different degrees of freedom to choose where to live [32,41]. In this study, people in private housing were assumed to have greater freedom of residence choice and were therefore classified into the residential self-selection group to account for the issue of residential self-selection. However, those in government-assigned public housing have limited freedom of residence choice and were therefore classified into the non-self-selection group due to the presumed lack of residential self-selection. This natural experimental method allowed us to distinguish self-selecting and non-self-selecting groups based on housing type.

Housing type was assessed by a single question: "Who has the property rights of the house/flat where you live now?" The response options for this question were as follows: (1) Family members have full property rights; (2) Family members have partial property rights; (3) Commercial housing is rented on the market; (4) A relative or friend has full property rights; (5) Others (i.e. unknown); (6) Public housing, also known as *gongfang* or *danwei*, wherein state-owned enterprises or public institutions, respectively, have full property rights; (7) Low-rent housing, also known as *lianjiazulinfang*, a kind of social security housing provided by the local government to provide a minimum living standard to urban families with housing difficulties, distributed in the form of rent subsidies, rent relief and housing apportionment; and (8) Publicly rented housing, also known as *gonggongzulinfang*, which refers to housing owned by a governmental or public agency that is rented to people with housing difficulties at below-market or affordable prices. Following previous studies [32,41], the first five housing types were treated as private housing (self-selection group), while the last three were treated as public housing (non-self-selection group).

2.5. Covariates

We adjusted for a series of confounding covariates, including gender, age, marital status, educational attainment, employment status, *hukou* status (registered permanent residence vs. registered temporary residence), medical insurance participation and annual household income. We also included neighborhood population density (persons/km^2) as a proxy for the level of urbanicity, following previous studies [44]. The summary statistics for all variables are shown in Table 1.

2.6. Statistical analysis

To assess the link between the availability of neighborhood PA facilities and individuals' PA behavior, we fitted multilevel linear regressions [45]. A multilevel approach was necessary due to the hierarchical structure of the data, as measurements at each wave were nested within individuals, and individuals were nested within neighborhoods. The full estimation model, including all predictors, was as follows:

$$PA_{tij} = \beta_0 + \beta_1 PAFacilities_j + \beta_2 Covariates_{tij} + \beta_3 Covariates_{ij} + \varepsilon_{tij} + \mu_{ij} + \phi_j$$

where t represents time, i represents individuals and j represents neighborhoods. β_0 is the random intercepts. PAFacilities j represents a vector of neighborhood-level variables. Covariates tij represents a vector of time-varying covariates such as age. Covariates ij represents a vector of time-invariant covariates such as gender. μ_{ij}, and $\phi_j$represent the random errors within individuals, between individuals and between neighborhoods, respectively.

Table 1
Characteristics of study participants in the baseline.

Variables	Proportion/Mean (SD)			
	Total	Private housing	Public housing	p-value
Dependent variables				
Weekly frequency of PA (times)	1.88 (2.95)	1.87 (2.94)	2.17 (3.05)	0.01[a]
Weekly duration of PA (hours)	2.86 (6.89)	2.84 (6.88)	3.61 (7.25)	0.01[a]
Independent variables				
Number of PA facilities within the neighborhood (numbers)	1.00 (1.53)	1.00 (1.51)	1.26 (2.06)	<0.01[a]
Controlled variables				
Gender (%)				
Male	48.89	48.91	48.07	0.68[b]
Female	51.11	51.09	51.93	
Age (years)	46.83 (16.66)	46.86 (16.65)	45.58 (17.12)	0.06[a]
Marital status (%)				
Single, divorced or widowed	14.22	14.11	18.65	<0.01[b]
Married	85.78	85.89	81.35	
Educational attainment (%)				
Primary school or below	50.91	51.39	32.8	<0.01[b]
High school	42.13	41.73	57.39	
College or above	6.96	6.88	9.81	
Employment status (%)				
Employed	65.68	65.93	56.11	<0.01[b]
Unemployed	34.32	34.07	43.89	
Hukou status (%)				
Local *hukou*	96.07	96.54	77.97	<0.01[b]
Non-local *hukou*	3.93	3.46	22.03	
Medical insurance (%)				
Having medical insurance	91.23	91.41	84.24	<0.01[b]
No medical insurance	8.77	8.59	15.76	
Annual household income (Chinese yuan)	11937.48 (13753.24)	11808.46 (13685.19)	16973.69 (15375.33)	<0.01[a]
Housing type (%)				
Private housing	97.50			
Public housing	2.50			
Neighborhood population density (persons/km^2)	4553.79 (14762.59)	4501.45 (14686.91)	6596.66 (17353.14)	<0.01[a]

SD = standard deviation; a = student's t-tests; b = chi-square tests.

The calculated variance inflation factors (all <3) suggested no severe multicollinearity among the dependent variables. The intra-class correlation coefficient (ICC) for the null model predicting frequency of PA is 0.08 at neighborhood level and 0.21 at individual level respectively; the ICC for the null model predicting duration of PA is 0.04 at neighborhood level and 0.09 at individual level respectively. This means that living within the same neighborhood accounted for 8% of total variation in respondents' frequency of physical activity, and for 4% of total variation in duration of PA. Such results confirmed the necessity of multilevel models.

First, we regressed the frequency and duration of PA separately on the presence of PA facilities (present vs. not) for all respondents (Models 1 and 2). Second, we regressed the frequency and duration of PA on the number of PA facilities separately for all respondents (Models 7 and 8). Third, we regressed the frequency and duration of PA on the presence of PA facilities for respondents living in private housing (Models 3 and 4). Fourth, we regressed the frequency and duration of PA on the number of PA facilities for respondents living in private housing (Models 9 and 10).

Fifth, we regressed the frequency and duration of PA on the presence of PA facilities for respondents living in public housing (Models 5 and 6). Last, we regressed the frequency and duration of PA on the number of PA facilities for respondents living in public housing (Models 11 and 12).

The results of the two groups (private housing vs. public housing) were compared to generate insights into the effect of self-selection. If residential self-selection bias exists, we would expect heterogeneity between the PA facilities–PA behavior associations of the two groups. However, if residential self-selection bias does not exist, we would expect similar results for the two groups.

3. Results

Table 1 summarizes the characteristics of the baseline population. The respondents' average weekly frequency of PA was 1.88 times, while their average weekly duration of PA was 2.86 h. The average number of PA facilities within the neighborhood was 1.00. In terms of socioeconomic and demographic characteristics, only 2.50% of the total respondents lived in public housing. The mean age of all respondents was 46.83 years, and 48.89% were male. Of the respondents, 85.78% were married, 42.13% had a high school degree and 6.96% had at least a college degree. Nearly all respondents had a local *hukou* (96.07%). Furthermore, 65.68% were employed, and a large proportion had medical insurance (91.23%). The average household income was 11,937.48 Chinese yuan, while the average neighborhood population density was 4553.79 persons/km^2.

Table 2 presents the results of a multilevel linear regression analysis to predict the frequency and duration of PA (where the independent variable is the presence of PA facilities). Model 1 and Model 2 show the associations between the presence of PA facilities and the frequency and duration, respectively, of PA for all samples. Compared with respondents living in neighborhoods without PA facilities, those who lived in neighborhoods with PA facilities had a higher frequency of PA (Coef. = 0.147, standard error [SE] = 0.067). However, the evidence does not support an association between the presence of PA facilities and the duration of PA (Coef. = 0.046, SE = 0.026).

Model 3 and Model 4 show the associations between the presence of PA facilities and the frequency and duration, respectively, of PA for respondents living in private housing. Compared with the respondents living in neighborhoods without PA facilities, those who lived in neighborhoods with PA facilities had a higher frequency of PA (Coef. = 0.147, SE = 0.067) and longer duration of PA (Coef. = 0.045, SE = 0.022). Model 5 and Model 6 show the associations between the presence of PA facilities and the frequency and duration, respectively, of PA for respondents living in public housing. However, the evidence does not support an association between the presence of PA facilities and the frequency of PA (Coef. = 0.060, SE = 0.201) or duration of PA (Coef. = 0.041, SE = 0.066).

Table 3 presents the results of a multilevel linear regression analysis to predict the frequency and duration of PA from the number of PA facilities. Model 7 and Model 8 show the associations between the availability of PA facilities and the frequency and duration, respectively, of PA for all samples. The number of PA facilities within the neighborhood is positively associated with the respondents' frequency of PA (Coef. = 0.047, SE = 0.021) but not with their duration of PA (Coef. = 0.015, SE = 0.008). Model 9 and Model 10 show the associations between the availability of PA facilities and the frequency and duration, respectively, of PA for respondents living in private housing. The number of PA facilities within the neighborhood is positively associated with both the respondents' frequency of PA (Coef. = 0.051, SE = 0.021) and duration of PA (Coef. = 0.016, SE = 0.008). Model 11 and Model 12 show the associations between the availability of PA facilities and the frequency

Table 2
Multilevel linear regression analysis to predict the frequency and duration of PA (independent variable: presence of PA facilities).

	All samples		Private housing		Public housing	
	Model 1 (DV = Frequency of PA)	Model 2 (DV = Duration of PA)	Model 3 (DV = Frequency of PA)	Model 4 (DV = Duration of PA)	Model 5 (DV = Frequency of PA)	Model 6 (DV = Duration of PA)
	Coef. (SE)	Coef. (SE)	Coef. (SE)	Coef. (SE)	Coef. (SE)	Coef. (SE)
Fixed part						
Independent variables						
With PA facilities (ref: without PA facilities)	0.147** (0.067)	0.046* (0.026)	0.147** (0.067)	0.045** (0.022)	0.060 (0.201)	0.041 (0.066)
Controlled variables						
Male (ref: female)	0.314*** (0.026)	0.115*** (0.009)	0.304*** (0.026)	0.111*** (0.009)	0.693*** (0.162)	0.281*** (0.054)
Age	0.024*** (0.001)	0.006*** (0.000)	0.023*** (0.001)	0.005*** (0.000)	0.039*** (0.006)	0.011*** (0.002)
Married (ref: single, divorced or widowed)	−0.378*** (0.033)	−0.142*** (0.011)	−0.375*** (0.034)	−0.142*** (0.011)	−0.531*** (0.195)	−0.124* (0.065)
Education (ref: primary school or below)						
High school	0.380*** (0.029)	0.140*** (0.010)	0.382*** (0.030)	0.142*** (0.010)	0.447** (0.195)	0.132** (0.065)
College or above	0.453*** (0.053)	0.199*** (0.018)	0.447*** (0.054)	0.196*** (0.018)	0.676** (0.283)	0.353*** (0.095)
Employed (ref: unemployed)	−0.713*** (0.027)	−0.197*** (0.009)	−0.705*** (0.027)	−0.193*** (0.009)	−1.016*** (0.173)	−0.343*** (0.058)
Local *hukou* (ref: non-local *hukou*)	0.155* (0.084)	0.054* (0.029)	0.091 (0.089)	0.024 (0.030)	0.629** (0.265)	0.284*** (0.088)
Medical insurance (ref: no medical insurance)	−0.210*** (0.039)	−0.055*** (0.013)	−0.217*** (0.040)	−0.057*** (0.014)	0.068 (0.221)	0.015 (0.074)
Household income (ref: no medical insurance)	0.164*** (0.01)	0.053*** (0.003)	0.162*** (0.010)	0.052*** (0.003)	0.206*** (0.062)	0.061*** (0.021)
Public housing (ref: private housing)	0.038 (0.082)	−0.004 (0.028)				
Neighborhood population density	−0.001 (0.004)	−0.002 (0.002)	−0.002 (0.005)	−0.002 (0.002)	0.033 (0.030)	0.007 (0.010)
Constant	0.095 (0.138)	0.178*** (0.048)	0.198 (0.142)	0.216*** (0.049)	−1.685** (0.714)	−0.447* (0.238)
Random part						
Var (Neighborhoods)	0.501***	0.501***	0.510***	0.075***	0.626***	0.061***
Var (Between individuals)	1.793***	1.793***	1.784***	0.199***	2.248***	0.254***
Var (Within individuals)	6.944***	6.943***	6.941***	0.805***	6.269***	0.701***
Log likelihood	−174680.4	−174680.33	−170649.87	−96857.341	−4024.487	−2263.277
AIC	349392.8	349392.7	341329.7	193744.7	8078.975	4556.556

DV = dependent variable; Coeff. = coefficient; *SE* = standard error; AIC = Akaike information criterion. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Table 3
Multilevel linear regression analysis to predict the frequency and duration of PA (independent variable: number of PA facilities).

	All samples		Private housing		Public housing	
	Model 7 (DV = Frequency of PA)	Model 8 (DV = Duration of PA)	Model 9 (DV = Frequency of PA)	Model 10 (DV = Duration of PA)	Model 11 (DV = Frequency of PA)	Model 12 (DV = Duration of PA)
	Coef. (SE)	Coef. (SE)	Coef. (SE)	Coef. (SE)	Coef. (SE)	Coef. (SE)
Fixed part						
Independent variables						
Number of PA facilities within the neighborhood	0.047** (0.021)	0.015* (0.008)	0.051** (0.021)	0.016** (0.008)	−0.019 (0.055)	0.001 (0.018)
Controlled variables						
Male (ref: female)	0.313*** (0.026)	0.115*** (0.009)	0.304*** (0.026)	0.111*** (0.009)	0.694*** (0.162)	0.281*** (0.054)
Age	0.024*** (0.001)	0.006*** (0.000)	0.023*** (0.001)	0.005*** (0.000)	0.039*** (0.006)	0.011*** (0.002)
Married (ref: single, divorced or widowed)	−0.378*** (0.033)	−0.142*** (0.011)	−0.375*** (0.034)	−0.142*** (0.011)	−0.533*** (0.195)	−0.125* (0.065)
Education (ref: primary school or below)						
High school	0.380*** (0.029)	0.140*** (0.010)	0.381*** (0.030)	0.142*** (0.010)	0.453** (0.195)	0.133** (0.065)
College and above	0.451*** (0.053)	0.199*** (0.018)	0.445*** (0.054)	0.195*** (0.018)	0.682*** (0.283)	0.353*** (0.095)
Employed (ref: unemployed)	−0.712*** (0.027)	−0.197*** (0.009)	−0.704*** (0.027)	−0.193*** (0.009)	−1.021*** (0.174)	−0.343*** (0.058)
Local *hukou* (ref: non-local *hukou*)	0.156* (0.084)	0.054* (0.029)	0.092 (0.089)	0.025 (0.030)	0.617** (0.265)	0.281*** (0.088)
Medical insurance (ref: no medical insurance)	−0.210*** (0.039)	−0.055*** (0.013)	−0.216*** (0.040)	−0.057*** (0.014)	0.069 (0.221)	0.015 (0.074)
Household income (ref: no medical insurance)	0.164*** (0.010)	0.053*** (0.003)	0.162*** (0.010)	0.052*** (0.003)	0.205*** (0.062)	0.060*** (0.021)
Public housing (ref: private housing)	0.037 (0.082)	−0.005 (0.028)				
Neighborhood population density	−0.001 (0.004)	−0.002 (0.002)	−0.002 (0.005)	−0.002 (0.002)	0.033 (0.030)	0.007 (0.010)
Constant	0.123 (0.135)	0.186*** (0.046)	0.222 (0.139)	0.223*** (0.048)	−1.621** (0.705)	−0.419* (0.235)
Random part						
Var (Neighborhoods)	0.074***	0.074***	0.509***	0.075***	0.633***	0.063***
Var (Between individuals)	0.200***	0.200***	1.785***	0.199***	2.244***	0.254***
Var (Within individuals)	0.804***	0.804***	6.941***	0.805***	6.269***	0.701***
Log likelihood	−99129.613	−99129.401	−170649.42	−96856.934	−4024.471	−2263.471
AIC	198291.2	198290.8	341328.8	193743.9	8078.944	4556.942

DV = dependent variable; Coeff. = coefficient; SE = standard error; AIC = Akaike information criterion. *p < 0.10, **p < 0.05, ***p < 0.01.

and duration, respectively, of PA for respondents living in public housing. The number of PA facilities is not associated with either the frequency of PA (Coef. = −0.019, SE = 0.055) or the duration of PA (Coef. = 0.001, SE = 0.018). As for sensitivity analysis, we kept neighborhoods with both public housing and private housing and reran the models. This can help us compare findings of the public housing and private housing from the same neighborhoods (Appendix: Tables S1 and S2). The results indicate that despite some differences in magnitude, the PA facilities-PA associations remained unchanged.

4. Discussion

This study extends previous research on the association between the availability of neighborhood PA facilities and PA behavior in three respects. First, the study is among the first to investigate the residential self-selection issue in a densely populated Chinese context. Second, it uses a natural experimental design with national-level longitudinal data to provide rigorous evidence on the association between the provision of neighborhood PA facilities and PA behavior. Third, it focuses on both the frequency and duration of PA to extend our understanding of the relationship between PA facilities and PA behavior.

Our results suggest that the availability of neighborhood PA facilities is associated with the frequency and duration of PA for people in private housing, whereas such associations are not observed for people in public housing. This implies that results measured elsewhere for residents of private housing may be influenced by residential self-selection bias. There are several potential causes of such bias. First, attitudes toward PA and health among private housing residents may affect their selection of residential location [31]. For example, people who enjoy taking PA may choose to live in a neighborhood with more PA facilities to meet their needs. In contrast, public housing residents who prefer to partake in PA for health reasons cannot act on this preference by choosing their place of residence, and may instead be assigned to a neighborhood with few PA facilities. Second, people's health-related knowledge may also affect both their residential locations and PA behavior [46]. Private housing residents may be better informed about the health benefits of PA and may choose to partake in it whether or not they find it enjoyable. Therefore, they are more likely to choose a neighborhood with a higher availability of PA facilities. Third, the social benefits of PA for private housing residents may also influence the disparity between residents of private and public housing in the observed built environment–PA behavior associations [47]. In China, people often regard group PA as an effective way to bond with neighbors and build social contacts [48,49]. Therefore, private housing residents may take frequent PA for the purpose of socializing and hence choose neighborhoods with a higher availability of PA facilities. In contrast, public housing residents have little freedom to choose their residential location even if group PA would be a way for them to build social contacts. Therefore, the built environment–PA behavior associations are more pronounced among private housing residents than among public housing residents.

It is important to acknowledge that this study does not provide conclusive proof of the existence of residential self-selection bias. First, residents of public housing usually have labor-intensive jobs in fields such as construction, restaurant service or sales [50], which require them to engage in a large amount of PA during the workday. Therefore, they may prefer not to take extra PA within their neighborhood after work, regardless of the availability of facilities. Second, residents of public housing have less time for recreation due to their typically long working hours and commuting times [51,52]. Under such circumstances, even those living in neighborhoods with a higher availability of PA facilities would have little free time during which to take PA. Third, this segment of the population also tends to have lower levels of

health-related knowledge and may not realize the health benefits of PA [53], making them more likely to engage in sedentary behavior (e.g. using smartphones or watching television) in their spare time, regardless of the existence of neighborhood PA facilities. Last, the size of the sample of residents of public housing was smaller than that of residents of private housing, which may have influenced the meaningfulness of the significance tests for the public housing samples [54]. Last, although most PA facilities are regarded as public infrastructures and provided by the government, there maintenance may still rely on local community [55]. Therefore, the quality of PA facilities for public housing residents may vary significantly across different neighborhoods. Furthermore, private housing in China is usually in the form of gated communities, while public housing is not [56]. Previous studies indicated that the effect of PA facilities in gated communities is more significant than non-gated communities, since the facilities in non-gated communities may be shared by a wider population, which may lead to be worn or crowded and hence discourage their usage [57].

The following limitations of this study should be noted. First, this research is subject to the modifiable areal unit problem (MAUP); that is, the definition of the boundaries of neighborhoods may have a confounding influence on environment–health associations [58]. The data were collected at the administrative neighborhood unit level, and the results might not hold true if we used fine-grained data on personal residential locations and local built environment characteristics. Second, the dependent variable was self-reported by the respondents, which may have caused self-report bias. Further research is needed to collect objective and accurate PA data. Third, detailed data on PA, such as intensity and type, were not included which prevents us from further identifying more precise environment-PA associations [59,60,61]. Fourth, a person's PA behavior may be affected not only by the PA facilities in their residential neighborhood but also by those near their workplace [13,21]. This may lead to an uncertain geographic context problem [62] and neighborhood effect averaging problem [63,64] regarding the effects of PA facilities on PA behavior. Fifth, our dependent variable only included the number of PA facilities within the neighborhood. Other aspects of PA facilities, such as quality, accessibility, type and size, are also important. Sixth, information regarding the respondents' usage of PA facilities was not included, which made it difficult for us to infer causation. Last, the respondents' attitudes toward and preference regarding PA behavior were not included, which prevents us from further investigating the mechanism by which residential self-selection influences PA behaviors.

4.1. Policy implications

The findings of this study have several policy implications. First, our results suggest that the association between the availability of neighborhood PA facilities and PA behavior may be influenced by residential self-selection. Therefore, some potential factors, such as the residents' attitudes toward and preferences regarding PA facilities and behavior, should be taken into account when planning for the provision of neighborhood PA facilities. Furthermore, the insignificant association between the availability of neighborhood PA facilities and PA behavior among residents of public housing may be due to the disadvantages facing this population group in terms of long working hours, labor-intensive jobs and limited health-related knowledge. Therefore, targeted policy interventions in such group should not be limited to increasing the provision of PA facilities. Other strategies, such as promoting awareness of the health benefits of PA and adjusting the opening hours of PA facilities to fit the residents' preferences and available leisure time, should also be pursued.

5. Conclusion

This study investigated the existence of residential self-selection bias in the association between the availability of neighborhood PA facilities and PA behavior by comparing residents of private and public housing. Residents of private housing have greater freedom of residential choice, whereas those in public housing are allocated to their residential locations randomly. Therefore, comparing these two groups can help us identify whether residential self-selection bias exists. The results showed that the availability of neighborhood PA facilities was positively associated with the frequency and duration of PA for residents of private housing. However, no association was observed for residents of public housing. This indicates that residential self-selection may cause bias when investigating the association between neighborhood PA facilities and individuals' PA behavior. Therefore, it is important to control for residential self-selection when examining the impact of PA facilities.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank the Institute of Social Science Survey of Peking University for providing CFPS data. We would like to thank the anonymous reviewer for their helpful and positive comments which have improved the quality of the paper. The work described in this paper was fully supported by the grants from National Natural Science Foundation of China (Project No. 51778552) and the Research Grants Council of the Hong Kong SAR (Project No. CityU11207520).

Abbreviations

PA physical activity

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.buildenv.2021.108401.

References

[1] World Health Organization, Global Recommendations on Physical Activity for Health, 2010. https://www.who.int/publications/i/item/9789241599979.
[2] G. Mammen, G. Faulkner, Physical activity and the prevention of depression: a systematic review of prospective studies, Am. J. Prev. Med. 45 (5) (2013) 649–657.
[3] H. Du, D. Bennett, L. Li, G. Whitlock, Y. Guo, R. Collins, J. Chen, Z. Bian, L. Hong, S. Feng, Physical activity and sedentary leisure time and their associations with BMI, waist circumference, and percentage body fat in 0.5 million adults: the China Kadoorie Biobank study, Am. J. Clin. Nutr. 97 (3) (2013) 487–496.
[4] P. Poirier, T.D. Giles, G.A. Bray, Y. Hong, J.S. Stern, F.X. Pi-Sunyer, R.H. Eckel, Obesity and cardiovascular disease: pathophysiology, evaluation, and effect of weight loss: an update of the 1997 American heart association scientific statement on obesity and heart disease from the obesity committee of the Council on nutrition, physical activity, and metabolism, Circulation 113 (6) (2006) 898–918.
[5] R.J. Sigal, G.P. Kenny, D.H. Wasserman, C. Castaneda-Sceppa, Physical activity/exercise and type 2 diabetes, Diabetes Care 27 (10) (2004) 2518–2539.
[6] H. Vainio, R. Kaaks, F. Bianchini, Weight control and physical activity in cancer prevention: international evaluation of the evidence, Eur. J. Canc. Prev. : the Official Journal of the European Cancer Prevention Organisation (ECP) 11 (2002) S94–S100.
[7] R. Guthold, G.A. Stevens, L.M. Riley, F.C. Bull, Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1· 9 million participants, The Lancet Global Health 6 (10) (2018) e1077–e1086.
[8] General Administration of Sport of China, Bulletin of National Fitness Activities, 2014, 2015, http://www.sport.gov.cn/n16/n1077/n1422/7300210.html.
[9] State Council, State Council's Opinions on Implementing Healthy China Action, 2019. http://www.gov.cn/zhengce/content/2019-07/15/content_5409492.htm.
[10] G.R. McCormack, A. Shiell, In search of causality: a systematic review of the relationship between the built environment and physical activity among adults, Int. J. Behav. Nutr. Phys. Activ. 8 (1) (2011) 1–11.
[11] A. Kajosaari, T.E. Laatikainen, Adults' leisure-time physical activity and the neighborhood built environment: a contextual perspective, Int. J. Health Geogr. 19 (1) (2020) 1–13.

[12] G.S. Lovasi, A.V. Moudon, A.L. Pearson, P.M. Hurvitz, E.B. Larson, D.S. Siscovick, E.M. Berke, T. Lumley, B.M. Psaty, Using built environment characteristics to predict walking for exercise, Int. J. Health Geogr. 7 (1) (2008) 1–13.
[13] Y. Liu, X. Wang, S. Zhou, W. Wu, The association between spatial access to physical activity facilities within home and workplace neighborhoods and time spent on physical activities: evidence from Guangzhou, China, Int. J. Health Geogr. 19 (1) (2020) 1–15.
[14] N. Karusisi, F. Thomas, J. Méline, B. Chaix, Spatial accessibility to specific sport facilities and corresponding sport practice: the RECORD Study, Int. J. Behav. Nutr. Phys. Activ. 10 (1) (2013) 1–10.
[15] X. Zhu, A. Yoshikawa, L. Qiu, Z. Lu, C. Lee, M. Ory, Healthy workplaces, active employees: a systematic literature review on impacts of workplace environments on employees' physical activity and sedentary behavior, Build. Environ. 168 (2020) 106455.
[16] B.Y.-M. Wong, S.-Y. Ho, W.-S. Lo, E. Cerin, K.-K. Mak, T.-H. Lam, Longitudinal relations of perceived availability of neighborhood sport facilities with physical activity in adolescents: an analysis of potential moderators, J. Phys. Activ. Health 11 (3) (2014) 581–587.
[17] J. Han, H.-J. Kang, G.H. Kwon, A measurement for evaluating the environmental quality of advanced healthcare facilities: intelligent healthscape quality for medical staff, Build. Environ. 144 (2018) 532–541.
[18] J.K. Vanos, A.J. Herdt, M.R. Lochbaum, Effects of physical activity and shade on the heat balance and thermal perceptions of children in a playground microclimate, Build. Environ. 126 (2017) 119–131.
[19] I. Markevych, M.P. Smith, S. Jochner, M. Standl, I. Brüske, A. von Berg, C.-P. Bauer, K. Fuks, S. Koletzko, D. Berdel, Neighbourhood and physical activity in German adolescents: GINIplus and LISAplus, Environ. Res. 147 (2016) 284–293.
[20] T.K. Kaufman, A. Rundle, K.M. Neckerman, D.M. Sheehan, G.S. Lovasi, J.A. Hirsch, Neighborhood recreation facilities and facility membership are jointly associated with objectively measured physical activity, J. Urban Health 96 (4) (2019) 570–582.
[21] L. Macdonald, Associations between spatial access to physical activity facilities and frequency of physical activity; how do home and workplace neighbourhoods in West Central Scotland compare? Int. J. Health Geogr. 18 (1) (2019) 1–10.
[22] K. Ball, R.W. Jeffery, G. Abbott, S.A. McNaughton, D. Crawford, Is healthy behavior contagious: associations of social norms with physical activity and healthy eating, Int. J. Behav. Nutr. Phys. Activ. 7 (1) (2010) 1–9.
[23] X. Cao, P.L. Mokhtarian, S.L. Handy, Examining the impacts of residential self-selection on travel behaviour: a focus on empirical findings, Transport Rev. 29 (3) (2009) 359–395.
[24] J.D. Mackenbach, M.G.M. de Pinho, E. Faber, N. den Braver, R. de Groot, H. Charreire, J.-M. Oppert, H. Bardos, H. Rutter, S. Compernolle, Exploring the cross-sectional association between outdoor recreational facilities and leisure-time physical activity: the role of usage and residential self-selection, Int. J. Behav. Nutr. Phys. Activ. 15 (1) (2018) 1–11.
[25] A.R. Pinjari, C.R. Bhat, D.A. Hensher, Residential self-selection effects in an activity time-use behavior model, Transp. Res. Part B Methodol. 43 (7) (2009) 729–748.
[26] D. Van Dyck, G. Cardon, B. Deforche, N. Owen, I. De Bourdeaudhuij, Relationships between neighborhood walkability and adults' physical activity: how important is residential self-selection? Health Place 17 (4) (2011) 1011–1014.
[27] Y. Lu, L. Chen, Y. Yang, Z. Gou, The association of built environment and physical activity in older adults: using a citywide public housing scheme to reduce residential self-selection bias, Int. J. Environ. Res. Publ. Health 15 (9) (2018) 1973.
[28] J. Baar, M. Romppel, U. Igel, E. Brähler, G. Grande, The independent relations of both residential self-selection and the environment to physical activity, Int. J. Environ. Res. Publ. Health 25 (3) (2015) 288–298.
[29] N.A. Howell, S. Farber, M.J. Widener, J. Allen, G.L. Booth, Association between residential self-selection and non-residential built environment exposures, Health Place 54 (2018) 149–154.
[30] G.R. McCormack, L. McLaren, G. Salvo, A. Blackstaffe, Changes in objectively-determined walkability and physical activity in adults: a quasi-longitudinal residential relocation study, Int. J. Environ. Res. Publ. Health 14 (5) (2017) 551.
[31] J. Boone-Heinonen, D.K. Guilkey, K.R. Evenson, P. Gordon-Larsen, Residential self-selection bias in the estimation of built environment effects on physical activity between adolescence and young adulthood, Int. J. Behav. Nutr. Phys. Activ. 7 (1) (2010) 1–11.
[32] T. Lin, D. Wang, X. Guan, The built environment, travel attitude, and travel behavior: residential self-selection or residential determination? J. Transport Geogr. 65 (2017) 111–122.
[33] B. Xie, Y. Lu, L. Wu, Z. An, Dose-response effect of a large-scale greenway intervention on physical activities: the first natural experimental study in China, Health Place 67 (2021) 102502.
[34] J. Ludwig, G.J. Duncan, L.A. Gennetian, L.F. Katz, R.C. Kessler, J.R. Kling, L. Sanbonmatsu, Neighborhood effects on the long-term well-being of low-income adults, Science 337 (6101) (2012) 1505–1510.
[35] G. Sun, N.M. Oreskovic, H. Lin, How do changes to the built environment influence walking behaviors? A longitudinal study within a university campus in Hong Kong, Int. J. Health Geogr. 13 (1) (2014) 1–10.
[36] A. Drewnowski, D. Arterburn, J. Zane, A. Aggarwal, S. Gupta, P. Hurvitz, A. Moudon, J. Bobb, A. Cook, P. Lozano, The Moving to Health (M2H) approach to natural experiment research: a paradigm shift for studies on built environment and health, SSM-Population Health 7 (2019) 100345.
[37] M. Kärmeniemi, T. Lankila, T. Ikäheimo, H. Koivumaa-Honkanen, R. Korpelainen, The built environment as a determinant of physical activity: a systematic review of longitudinal studies and natural experiments, Ann. Behav. Med. 52 (3) (2018) 239–251.
[38] J.S. Benton, J. Anderson, R.F. Hunter, D.P. French, The effect of changing the built environment on physical activity: a quantitative review of the risk of bias in natural experiments, Int. J. Behav. Nutr. Phys. Activ. 13 (1) (2016) 1–18.
[39] F. MacMillan, E.S. George, X. Feng, D. Merom, A. Bennie, A. Cook, T. Sanders, G. Dwyer, B. Pang, J.M. Guagliano, Do natural experiments of changes in neighborhood built environment impact physical activity and diet? A systematic review, Int. J. Environ. Res. Publ. Health 15 (2) (2018) 217.
[40] T.V. Calise, S.C. Dumith, W. DeJong, H.W. Kohl, The effect of a neighborhood built environment on physical activity behaviors, J. Phys. Activ. Health 9 (8) (2012) 1089–1097.
[41] P. Zang, Y. Lu, J. Ma, B. Xie, R. Wang, Y. Liu, Disentangling residential self-selection from impacts of built environment characteristics on travel behaviors for older adults, Soc. Sci. Med. 238 (2019) 112515.
[42] D. Wang, T. Lin, Residential self-selection, built environment, and travel behavior in the Chinese context, Journal of Transport and Land Use 7 (3) (2014) 5–14.
[43] D. Wang, Y. Chai, F. Li, Built environment diversities and activity–travel behaviour variations in Beijing, China, J. Transport Geogr. 19 (6) (2011) 1173–1186.
[44] Y. Liu, R. Wang, Y. Xiao, B. Huang, H. Chen, Z. Li, Exploring the linkage between greenness exposure and depression among Chinese people: mediating roles of physical activity, stress and social cohesion and moderating role of urbanicity, Health Place 58 (2019) 102168.
[45] S.W. Raudenbush, A.S. Bryk, Hierarchical Linear Models: Applications and Data Analysis Methods (Vol. 1), Sage2002.
[46] J. Boone-Heinonen, P. Gordon-Larsen, D.K. Guilkey, D.R. Jacobs Jr., B.M. Popkin, Environment and physical activity dynamics: the role of residential self-selection, Psychol. Sport Exerc. 12 (1) (2011) 54–60.
[47] A.E. Bauman, R.S. Reis, J.F. Sallis, J.C. Wells, R.J. Loos, B.W. Martin, L.P.A.S. W. Group, Correlates of physical activity: why are some people physically active and others not? The Lancet Global Health 380 (9838) (2012) 258–271.
[48] E.M. Bertera, Physical activity and social network contacts in community dwelling older adults, Activities, Adaptation & Aging 27 (3–4) (2003) 113–127.
[49] H. Litwin, Physical activity, social network type, and depressive symptoms in late life: an analysis of data from the National Social Life, Health and Aging Project, Aging Ment. Health 16 (5) (2012) 608–616.
[50] W. Zeng, P. Rees, L. Xiang, Do residents of Affordable Housing Communities in China suffer from relative accessibility deprivation? A case study of Nanjing, Cities 90 (2019) 141–156.
[51] Y. Zhao, Y. Chai, Residents' activity-travel behavior variation by communities in Beijing, China, Chin. Geogr. Sci. 23 (4) (2013) 492–505.
[52] S. McLafferty, V. Preston, Who has long commutes to low-wage jobs? Gender, race, and access to work in the New York region, Urban Geogr. 40 (9) (2019) 1270–1290.
[53] Y. Yu, A. He, S. Zheng, J. Jiang, J. Liang, B. Shrestha, P. Wang, How does health literacy affect the utilization of basic public health services in Chinese migrants? Health Promotion International, 2021.
[54] R.M. Royall, The effect of sample size on the meaning of significance tests, Am. Statistician 40 (4) (1986) 313–315.
[55] J. Shen, J. Cheng, W. Huang, F. Zeng, An exploration of spatial and social inequalities of urban sports facilities in Nanning City, China, Sustainability 12 (11) (2020) 4353.
[56] J. Chen, L. Yao, H. Wang, Development of public housing in post-reform China, China World Econ. 25 (4) (2017) 60–77.
[57] Y. Gul, Z. Sultan, M. Moeinaddini, G.A. Jokhio, The effects of physical activity facilities on vigorous physical activity in gated and non-gated neighborhoods, Land Use Pol. 77 (2018) 155–162.
[58] A.S. Fotheringham, D.W. Wong, The modifiable areal unit problem in multivariate statistical analysis, Environ. Plann. 23 (7) (1991) 1025–1044.
[59] B. Huang, T. Xiao, G. Grekousis, H. Zhao, J. He, G. Dong, Y. Liu, Greenness-air pollution-physical activity-hypertension association among middle-aged and older adults: evidence from urban and rural China, Environ. Res. 195 (2021) 110836.
[60] R. Wang, Y. Liu, Y. Lu, Y. Yuan, J. Zhang, P. Liu, Y. Yao, The linkage between the perception of neighbourhood and physical activity in Guangzhou, China: using street view imagery with deep learning techniques, Int. J. Health Geogr. 18 (1) (2019) 1–11.
[61] R. Wang, Z. Feng, J. Pearce, Y. Liu, G. Dong, Are greenspace quantity and quality associated with mental health through different mechanisms in Guangzhou, China: a comparison study using street view data, Environ. Pollut. (2021) 117976.
[62] M.-P. Kwan, The uncertain geographic context problem, Ann. Assoc. Am. Geogr. 102 (5) (2012) 958–968.
[63] M.-P. Kwan, The neighborhood effect averaging problem (NEAP): an elusive confounder of the neighborhood effect, Int. J. Environ. Res. Publ. Health 15 (9) (2018) 1841.
[64] R. Wang, Z. Feng, J. Pearce, S. Zhou, L. Zhang, Y. Liu, Dynamic greenspace exposure and residents' mental health in Guangzhou, China: from over-head to eye-level perspective, from quantity to quality, Landsc. Urban Plann. (2021) 104230.