SSM - Population Health 6 (2018) 9–15

Contents lists available at ScienceDirect

SSM - Population Health

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

Article

Walkable neighborhoods and obesity: Evaluating effects with a propensity score approach

Lori Kowaleski-Jones*, Cathleen Zick, Ken R. Smith, Barbara Brown, Heidi Hanson, Jessie Fan

University of Utah, USA

ARTICLE INFO

Keywords:
Obesity
Neighborhood walkability
Propensity score matching
Non-random selection bias
Built environment

ABSTRACT

Background: Research investigating the connection between neighborhood walkability and obesity often overlooks the issue of nonrandom residential selection.

Methods: We use propensity score methods to adjust for the nonrandom selection into residential neighborhoods in this cross-sectional, observational study. The sample includes 103,912 women residing in Salt Lake County, Utah age 20 or older. We measured percentage living in neighborhoods with more walkability, area level measures of neighborhood characteristics, and obesity (body mass index (BMI) > 30).

Results: Our findings confirm previous work that observes an association between living in more walkable neighborhoods and lower obesity. After adjusting for nonrandom selection, the odds of being obese when living in a less walkable neighborhood increase. Specifically, the odds ratio for being obese without the propensity score correction is 1.12. After adjusting for nonrandom selection, the odds ratio for being obese is 1.19, an increase of six percent.

Conclusion: Results demonstrate that residential selection bias inherent in cross-sectional analysis slightly attenuates the true association between neighborhood walkability and obesity. Results lend support to the growing body of research suggesting that more walkable neighborhoods have residents with a lower prevalence of obesity. Absent propensity score controls, the causal relationship between environment and obesity would be underestimated by 6%. Our analysis suggests there is an association between neighborhood walkability and obesity.

1. Introduction

A growing body of research provides evidence of the relationship between poor walkability of the residential environment and greater obesity prevalence (Mackenbach et al., 2014). However, many studies rely on correlational evidence, without considering that individuals may select their location in ways that relate to obesity prevalence. Active individuals may prefer to live in walkable environments and failing to account for this selection may lead to biased results. A number of design and statistical approaches are available to adjust for self-selection into neighborhoods, which we review briefly. We employ a propensity score approach to matching within a large sample frame of mothers (n= 105,770) drawn from birth certificates. We then test whether walkable neighborhoods relate to obesity prevalence, after applying propensity score adjustments.

1.1. Walkability

A wide variety of measures of walkability have been employed in past research, with specific measures often chosen to represent one or more of three defining features of walkability known as the “3Ds:” population density, a diversity of destinations, and pedestrian friendly design (Cervero & Kockelman, 1997). Population density may encourage walking because automotive traffic may be inconvenient and because there is a critical mass of residents needed to supply walking destinations, such as stores and mass transit stops. Diversity of destinations means that there are nearby jobs, stores, services, transit stops and other desired destinations within walking or biking distance. Pedestrian friendly designs include features that make active travel pleasant and safe, such as tree-shaded streets, well-connected streets with sidewalks, or pleasant aesthetic features. A recent review of 92 articles on this topic confirmed that land use diversity is consistently related to obesity, with other walkability factors sometimes related to obesity (Mackenbach et al., 2014).

In the current study, we employ measures of walkability that have proved useful in past research and are commonly available. We include population density, which has been found to relate to both more walking and lower obesity risk (Bodea, Garrow, Meyer, & Ross, 2009;

* Corresponding author.

https://doi.org/10.1016/j.ssmph.2017.11.005
Received 22 November 2016; Received in revised form 12 November 2017; Accepted 13 November 2017

Glazier et al., 2014; Hu et al., 2014; Rodríguez, Evenson, Diez Roux, & Brines, 2009; Rundle et al., 2007), although sometimes only for men (Smith et al., 2008; Wen and Kowaleski-Jones, 2012) or not at all (Gebel, Bauman, Sugiyama, & Owen, 2011; Ross, Crouse, Tremblay, Khan, Tremblay & Berthelot, 2007). We also include proximity to the central business district and the proportion of residents who actively commute by walking, biking, or taking public transit to work. These variables indicate whether residents are living in neighborhoods with sufficiently diverse land uses that they can walk, bike, or take transit to destinations and have been found to relate to BMI or walking in past research (Brown et al., 2013; Brown et al., 2014; Smith et al., 2016; Zick et al., 2013). A variable measuring neighborhood age captures pedestrian-friendly design, given that older neighborhoods often have amenities like tree-shaded sidewalks and narrow streets (Southworth & Ben-Joseph, 1995). In addition, older neighborhoods are related to more walking and lower BMI in past research (Berrigan & Troiano, 2002; Smith et al., 2008; Zick et al., 2013).

An important limitation in research that relates neighborhood design to obesity prevalence is that many of the studies on this topic are cross-sectional and thus they are typically unable to draw strong conclusions about causality. People may elect to live in neighborhoods that reinforce their physical activity and dietary predispositions or select neighborhoods based on reasons that correlate with physical activity and dietary patterns. If true, then conclusions drawn from cross-sectional analyses have the potential to misrepresent the underlying causal relationship between neighborhood environments and body weight.

1.2. Addressing residential selection: traditional approaches

Both cross-sectional and longitudinal designs have been used to address nonresidential selection. Among cross-sectional designs, asking residents about why they moved to their neighborhood and then adjusting for their responses is a common technique. Results of these adjustments often show that the strength of the relationship between the built environment and physical activity/BMI declines in magnitude, yet retains some statistical significance (Frank, Saelens, Powell, & Chapman, 2007; Handy, 2006,; Pinjari, Bhat, & Hensher, 2009; Smith et al., 2011), providing evidence that both residential selection and causal effects of neighborhoods are present. However, answering survey questions may also generate new sources of bias as responses are prone to error because of memory issues, dissonance reduction, and/or social desirability.

Other cross-sectional studies make use of a two-stage least squares or instrumental variables approaches in estimating the relationship between neighborhood choice and BMI (Terwee et al., 2010; Zick et al., 2013). Zick et al. found that the relationship between neighborhood walkability and obesity strengthened after instrumental variable adjustments and She et al. found that transit served areas also related to lower obesity risks, after adjustment (She, King, and Jacobson, 2017). Yet, these approaches are limited by the functional form that is chosen, the inability to control for unobservable characteristics related to the residential location decision, and by the reality that such methods may hide the fact that some in the “treated” sample have no matched controls in the non-treated sample.

Although longitudinal data provide an opportunity to use time ordering to quantify causal effects (Boone-Heinonen, Gordon-Larsen, Guilkey, Jacobs, & Popkin, 2011; Hirsch et al., 2014; Michael, Nagel, Gold, & Hillier, 2014; Smith et al., 2016), such data are relatively rare. Other studies have exploited the use of natural experiments to investigate the linkages between the built environment and body weight (Ludwig et al., 2011; Mayne, Auchincloss, & Michael, 2015), however this method is also difficult to implement because it requires foreknowledge of the environmental change. Given that each approach has strengths and weaknesses, greater confidence will be provided when researchers use a variety of approaches and reach similar conclusions. We propose that a propensity score matching approach is also useful in complementing these traditional approaches to accounting for nonrandom selection.

1.3. Residential selection threat: propensity score approaches

Rosenbaum and Rubin (1983, 1984) propose the use of the propensity score method to address the issue of nonrandom selection bias. Propensity scores adjust for the bias that is caused by the self-selection into a walkable neighborhood by creating matches between members of the treatment and control groups rather than through the random assignment used in true experiments.

Past research has employed propensity scores in studies that examine relationships between neighborhood qualities and walking (Cao, 2010; Mujahid et al., 2008), BMI or obesity (Leal, Bean, Thomas, & Chaix, 2011; Mujahid et al., 2008), and hypertension (Chiu et al., 2016). Three of these studies found no substantial differences in results between full and propensity score matched samples (Chiu et al., 2016) another found a reduction of effects for some but not all walkability features (Boer, Zheng, Overton, Ridgeway, & Cohen, 2007). Substantial reductions were found for self-reported attitudes toward walking and self-reported walking, but the author noted that, due to cross-sectional data, it is not clear whether attitudes actually influenced residential selection (Cao, 2010).

Building on the propensity score matching approaches used in physical activity and obesity research, this study tests whether the relationship between neighborhood characteristics and BMI remains significant after adopting propensity score controls. Our research extends prior work in this domain by utilizing a data source that provides expansive coverage of neighborhood residents and the ability to match based on numerous characteristics. Specifically, we assess the overall relationship between neighborhood design and obesity prevalence to test our two competing hypotheses: 1) the relationship is no longer significant after propensity score matching, suggesting that residential selection drives the association; 2) A significant relationship remains after adjusting for propensity score matching, suggesting a causal effect due to neighborhood design. In answering these questions, we explicitly adjust for residential self-selection within the context of a cross-sectional, non-experimental study design.

2. Methods

2.1. Data

The propensity score approach requires measures for obesity, neighborhood walkability, and measures of multiple variables that might relate to residents' choices to live in their neighborhood, and a set of matching procedures unique to propensity score methods, as described below. Recent research has noted that the choice of definitions for neighborhood likely matter when investigating the link between neighborhood characteristics and obesity (Fan et al., 2014; Hattori, An, & Sturm, 2013; Xu, Wen, & Wang, 2014; Yamada et al., 2012; Zhang et al., 2014). We elect to use the Census block group as our definition of neighborhood. The Census block group is a relatively small area (i.e., typically about 1500 residents, ranging from 300 to 3000) (U.S. Census Bureau, 2000) that generally approximates a local neighborhood. The data utilized in this study come from 550 of the 567 census block groups in Salt Lake County, UT, and are drawn from two different sources. (We eliminate 17 block groups that are at the periphery of the county with very few residents.).

The first data source used is the Utah Population Database (UPDB). UPDB is one of the world's richest sources of linked population-based information. One of its elements is a complete set of Utah birth certificates from 1930–2015. The birth certificates contain health and sociodemographic information for the mother, the father, and the child. Importantly for the purposes of our analyses, these data contain self-reported pre-*pregnancy* measures of the mother's height and weight (starting in 1989) used to construct her BMI for a large defined population. In addition, the birth certificates provide residential address information allowing us to locate a woman in a specific neighborhood at the time of the child's birth. Finally, UPDB and the birth certificates provide information on key socio-demographic measures including the mother's age, education, race/ethnicity, marital status, and her siblings' BMI.

This project has been approved by the University of Utah Institutional Review Board and the Utah Resource for Genetic and Epidemiologic Research. As part of this process, the UPDB staff retained identifying address information, linked birth certificate data to census-block groups based on Universal Transverse Mercator (UTM) coordinates (similar to longitude and latitude), and then provided the researchers with a data set without individual addresses.

Our UPDB sample was initially 108,920 women who had a birth during 1995–2005. Some women had multiple births during this period. In those cases, we chose the observation during this period that was closest to 2000, given the rich Census data that year. We further restricted the sample by eliminating women under age 20 (3021) and those with a pre-pregnancy BMI below 18.5 or above 49.9 (1987). We omitted young mothers (< age 20) because they were more likely to be unmarried and living with their family of origin, making residential selection a non-issue for them (Smith et al., 2011). With these restrictions in place, the sample size is 103,912.

UPDB data were linked to the 2000 U.S. Census. The Census contains numerous variables that capture neighborhood characteristics measured at the Census block group level. The use of the 2000 Census data allows us to describe the salient features of the neighborhood during the 10-year window of our birth certificate data.

2.2. Neighborhood walkability

Central to the propensity score approach is the identification of more walkable and less walkable neighborhoods. For the current analyses, we adopt the use of a walkability factor score that captures several domains of walkability that reflect evidence of active transportation (i.e., walking, biking, or taking transit) and the age and proximity to the central business district (CBD) of the houses in the neighborhood (Zick et al., 2013). Specifically, the factor score equals 0.30 proximity to the CBD, 0.17 population density, 0.24 proportion walking to work, 0.21 proportion biking to work, 0.24 proportion taking public transit to work, and 0.26 median age of the housing. The neighborhood walkability factor scores were derived from a confirmatory factor analysis where we assume only one factor and the items yielded a standardized Cronbach's alpha of 0.78 indicating acceptable levels of internal consistency (Nunnally, 1978). The factor scores distinguish high walkable neighborhoods from low walkable neighborhoods. As reported by Zick et al. (2013) the average proximity to the CBD is 14.17 miles (23 for the lowest and 4.66 for the highest factor score quartile). In comparison to neighborhoods that have high factor scores, neighborhoods with low factor scores are farther away from the CBD (23.00 vs. 4.66 miles), have lower population density (3703 vs 7473 residents per square mile), have a smaller proportion of residents who use public transit (0.02 vs 0.08) or walk (0.01 vs. 0.06) or bike (0.00 vs. 0.08) to and from work, and have a younger housing stock (12.79 vs. 46.12 years).

We follow past research in defining the CBD (Christakis & Fowler, 2007) and measure each individual's proximity as the street network distance between the centroid of each block group to the closest street intersection in the CBD measured in miles. The remaining variables contained in the factor score are based on U.S. Census 2000. Neighborhoods in the top quartile of walkability based on the factor score are categorized as 'more walkable neighborhoods' and those in the bottom three quartiles are considered 'less walkable neighborhoods'. Our choice to designate the top quartile of the walkability score as more walkable neighborhoods is motivated by other literature that employs similar designations to evaluate the linkages between walkability and health outcomes (Christian et al., 2011; Frank et al., 2007; Li, Harmer, Cardinal, & Vongjaturapat, 2009; Müller-Riemenschneider et al., 2013). Further, we conducted sensitivity analyses where we used varying designations (top 15 percent and top 30 percent of our walkability score) and found similar results (available by request to authors).

2.3. Obesity

From the birth certificate data contained in UPDB, height and pre-pregnancy weight information for the mothers are converted to BMI ([weight in kg]/[height in m]^2). Pre-pregnancy weight is self-reported in a clinical setting prior to the birth of the child. Obesity is coded 1 for BMI $\geq$ 30 and 0 otherwise, following the definitions from the Center from Disease Control. We choose obesity status over BMI as it represents a health concern at a threshold level as compared to a continuous measure.

2.4. Other individual characteristics

Other individual-level information that is available in the birth certificates includes marital status, mother's education, race/ethnicity, and age. All of these characteristics are included in the propensity score estimation.

2.5. Propensity score procedures

The propensity score methodology aims to balance the treatment group (i.e., residents of more walkable neighborhoods) with the control group (i.e., residents of less walkable neighborhoods) with regard to their covariates. The propensity score adjusts for the bias that may be caused by individuals self-selecting into certain neighborhoods by creating matches in terms of demographics between members of the treatment and control groups rather than through randomization used in true experiments. As such, the propensity score approach addresses concerns about functional form and the need to use only those observations in the common support region, defined below.

Three conditions must hold when using the propensity score approach. First, once we control for observable covariates, potential outcomes must be independent of the treatment selection. This is known as the conditional independence assumption (CIA). In our case, this means that residential choice should be random once we control for the covariates.

The second condition is the common support assumption. The common area of support or the degree of overlap in the propensity score between treated and untreated subjects. In this paper, this means that the estimated probabilities of living in a more walkable neighborhood for members of the treatment group must overlap with the estimated probabilities of living in a more walkable neighborhood for members of the control group and the probabilities have to be positive, irrespective of the covariate values (Caliendo & Kopeinig, 2008; Imbens, 2004; Smith & Todd, 2005). We define the common support region by dropping treatment observations whose propensity scores are higher than the maximum or less than the minimum propensity of the controls. This insures that individuals who live in a more walkable neighborhood have a counterpart living in a less walkable neighborhood but who nonetheless has the same estimated probability of living in the more walkable neighborhood.

The final condition is the stable unit treatment value assumption (SUTVA). SUTVA requires that the outcome for any individual depends on her/his residential choice only and not on the residential choice of any other individual in the sample. This leads us to restrict our analyses to women who are represented in the data with one observation. If a woman were allowed to enter the data set more than once because of a subsequent birth, this would violate the SUTVA assumption.

The propensity score approach relies first on estimating a logit equation where the dependent variable measures whether the

Table 1
Off and on support numbers by treated and untreated group: Results of Propensity Score Matching (Data from Utah Population Data Base).

	Off support	On support	Total
Untreated	58,462	17,386	75,848
Treated	8355	17,095	25,450
Total	66,817	34,481	101,298

respondent is living in a more walkable community. The independent variables include factors that are hypothesized to affect the decision to live in a walkable neighborhood and/or BMI. This constraint guides our choice of sociodemographic variables in the propensity score estimation and is the rationale behind not including other sociodemographic variables that are included in the birth certificate data. From the logit estimates, the predicted probabilities of living in a more walkable neighborhood are generated for all individuals. These predicted probabilities become the features on which treatment respondents are matched to controls.

Once the common support area is identified from the estimated probabilities, members of the treatment group are matched to members of the control group (See Table 1). Here the treatment group is defined as those individuals who live in the most walkable neighborhoods, as measured by the top quartile of walkability. There are varieties of ways in which propensity scores have been used in the literature with a range of positions on the use of matching techniques (King and Nielsen, 2016). However, matching methods have been widely used with a variety of research applications (Gibson-Davis & Foster, 2006). In this paper, we follow the example provided by Cao and colleagues (Cao, Xu, & Fan, 2010) and use a caliper method without replacement for matching but recognize that there is emerging scholarship that also advocates for the use of propensity score adjustment. All propensity score analyses are conducted using psmatch2 in STATA (Leuven & Sianesi, 2015). Non-replacement methods translate into a situation where an observation in the control group can only be used once. We also use the common option in STATA which drops treatment observations whose propensity scores are outside of the range of the propensity score of control observations. We set our caliper width to .01, which means that the non-treatment observations whose propensity scores are within .01 of the propensity score of the treatment observation are selected. It is important to note that use of these methods within the psmatch technique in STATA may result in the situation where we cannot find matches for some of the treatment observations and thus we discard some treatment observations in this process. (In supplementary analyses available upon request, other matching methods were used and they produced similar results.) As a measure of balance, we reported the standardized mean difference and follow prior empirical examples (Cao et al., 2010; Oakes & Johnson, 2006) where a standard difference less than 10 percent is considered an acceptable difference between the groups (See Table 4).

3. Results

Descriptive information for the variables used in both the estimation of the propensity scores as well as for the key treatment of interest, walkable neighborhoods, and outcome of obese versus non-obese are presented in Table 2. AT an average of 17, approximately 13% of the sample are classified as obese in these data. These figures are comparable to BRFSS statistics showing that regardless of gender, 24% of Utah adults are obese in 2000. These figures are also comparable to the 17.7% of women in Utah aged 18–34 who are obese ("Complete Health Indicator Report of Obesity Among Adults," 2016).

Table 3 shows the parameter estimates of the logistic regression that is used to predict the probability that an individual lives in a walkable neighborhood. Statistically significant covariates generally have the

Table 2
Descriptive statistics for 103,912 women residing in Salt Lake County, Utah age 20 or older (Data from Utah Population Data Base).

Variable	Percent/mean	S.D.
Percent of sample obese	13	0.13
Walkability top quartile (percent)	25	0.43
Race/Ethnicity of mother (percent)		
Black	.09	.09
Hispanic	17	.37
Pacific Islander	4	.19
Other race	1	.12
Total number of live births	2.09	1.36
Maternal age	27.32	5.80
Mother is married (percent)	77	.42
Educational level of mother (percent)		
High school education	32	.47
Some college	27	.44
College grad	23	.42
Mother is employed (percent)	70	.46
Birth year of observed pregnancy	1999	2.54

Data for total number of live births, maternal age and birth year of observed pregnancy are presented as means. The remaining descriptive characteristics are percentages.

Table 3
Logistic regression parameter estimates used to generate the propensity scores (dependent variable is top quartile of walkability measure; Data from Utah Population Data Base).

	Odds ratio	Std. Err.	P-value
Race/Ethnicity of mother (reference category is white)			
Black	2.99	.39	.00
Hispanic	1.95	.184	.00
American Indian	2.87	.31	.00
Pacific Islander	1.42	.16	.00
Other race	1.78	.29	.00
Birth order	.85	.02	.00
Maternal age	1.01	.01	.22
Mother is married	.77	.04	.00
Educational level of mother (reference category is less than high school)			
High school education	.63	.03	.00
Some college	.74	.07	.00
College grad	1.17	.14	.19
Mother is employed	.90	.02	.00
Birth year of observed pregnancy	.96	.01	.00

expected effects. After the matching is completed, we assess whether the variables included in the prediction equation are balanced between the matched groups. Descriptive information on the covariates used in the walkable neighborhood logit equations are presented in Table 4. Before matching, residents in high and lower walkable neighborhoods differ by a number of factors such as Hispanic origin, birth order, marital status and education. After matching, the standard differences of all variables is reduced to the acceptable level of 10% or less.

Table 5 displays the estimates of the average treatment effects for the treated (ATT) using the propensity scores. Approximately 12% of the women are obese when they live in walkable neighborhoods, which are defined as neighborhoods in the top quartile for walkability. In contrast, 13% of the women are obese when they live in less walkable neighborhoods (the lower three quartiles for walkability). After adjusting for the propensity score matching, the percentage of women who are obese living in walkable neighborhoods continues to be lower than the percentage of women who are obese living in less walkable neighborhoods (12% vs. 14%), which is consistent with prior research (Boone-Heinonen, Guilkey, Evenson, & Gordon-Larsen, 2010; Zick et al., 2013). Based on the estimates presented in Table 5, the odds ratio

Table 4

More walkable and less walkable neighborhood means for covariates used to estimate propensity scores: Unmatched and matched for a sample of 103,912 women residing in Salt Lake County, Utah age 20 or older.

Variable	Sample	Treated mean	Control mean	Std. Difference
Race/Ethnicity of mother				
Black	Unmatched	0.017	0.007	9.1
	Matched	0.014	0.016	-1.6
Hispanic	Unmatched	0.245	0.142	26.4
	Matched	0.213	0.229	-4.1
American Indian	Unmatched	0.018	0.008	9.1
	Matched	0.015	0.016	-1.2
Pacific Islander	Unmatched	0.042	0.038	2.5
	Matched	0.043	0.045	-1.2
Other Race	Unmatched	0.019	0.012	5.8
	Matched	0.017	0.017	-0.1
Birth Order	Unmatched	1.915	2.134	-16.6
	Matched	1.956	1.966	-0.8
Maternal Age	Unmatched	26.955	27.388	-7.5
	Matched	27.180	27.335	-2.7
Mother is married	Unmatched	0.715	0.790	-17.4
	Matched	0.732	0.719	3.2
Educational level of mother				
High school graduate	Unmatched	0.262	0.347	-18.6
	Matched	0.286	0.261	5.4
Some college	Unmatched	0.225	0.283	-13.3
	Matched	0.232	0.236	-1
College graduate	Unmatched	0.264	0.213	12.1
	Matched	0.265	0.273	-1.7
Mother is employed	Unmatched	0.684	0.709	-5.3
	Matched	0.703	0.677	5.7

for being obese without the propensity score correction is 1.12. After adjusting for nonrandom selection, the odds ratio for being obese is 1.19 which translates into odds increasing about 6% (1.19–1.12/1.12 = .06). These odds ratios are calculated using the inverse of a cross product ratio derived from the estimates of treated and controls from Table 5. In Table 5, note that the difference in proportion of individuals who are obese before and after adjustments for propensity scores have confidence intervals that do not overlap. These findings indicate that without correcting for nonrandom residential selection, the association between neighborhood walkability and obesity is understated by six percent.

We also perform some sensitivity analyses to test the robustness of our main result. First, we estimated a reduced-form regression equation where in the first iteration, obesity is regressed on the variable denoting the top quartile of walkability for the entire sample. Then, we restrict the sample to just those individuals who are in the support region created by propensity score matching techniques. Although the p value is smaller for the second set of results, the overall pattern of results remain significant across the two specifications. Second, we perform some additional analyses based on the principles of coarsened exact matching (Blackwell et al., 2009). We examined the variables used to produce our propensity scores and subsequent matching in Stata. Of these variables, the only variables that would benefit from additional coarsening would be maternal age and birth order. We set the following cut point for birth order to be first born versus all others. For maternal age, we set the categories as mothers under the age of 25 (and the reference category for the analysis), mothers who are aged 25 to 34 and mothers who are over the age of 35. We re-estimated the propensity score using these coarsened variables and then performed the matching procedures. We find that the value for the ATT on the treated is .121 (.120 is what we found prior to coarsening) and for the controls it is .127 (.138 before additional coarsening) which is comparable to our original estimates. These additional sensitivity tests bolster our confidence in our main findings.

4. Discussion

There is growing evidence of the potential importance of the built environment, especially land use diversity, for diet and physical activity (Mackenbach et al., 2014). Yet, the proportion of the association between the built environment and obesity that is causal is unknown as is the amount of the observed relationship that is a function of residential self-selection. In the absence of longitudinal data, it is difficult to disentangle the two effects.

We applied propensity score matching techniques to cross-sectional data from Salt Lake County to explore the potential for self-selection bias in the relationship between one summary measure of neighborhood walkability, and obesity. Our findings confirm previous work which observes an association between walkability and lower body weight among adults (Boone-Heinonen et al., 2010; Zick et al., 2013). More importantly, if our findings replicate in other studies, the results would suggest that the residential selection bias inherent in traditional cross-sectional analyses is likely modest, as this is a pattern noted in other research using nationally representative data (Hart et al., 2015). We find an underestimate of the relationship to be roughly 6 percent. This implies that the physical environment of neighborhoods may indeed affect obesity. As such, these results lend support to the growing body of research that concludes that neighborhoods that are more walkable are also less obesogenic.

Our results suggest that failing to account for residential selection may result in an underestimation of associations between neighborhood walkability and obesity. This is a finding that does run counter to longitudinal (Smith et al., 2011) and some cross-sectional studies (Ewing, Brownson, & Berrigan, 2006,; Lee, Ewing, & Sesso, 2009), which largely do not addresses potential selection effects but yet conclude that neighborhood effects are likely overstated in studies that do not correct for residential selection bias. However, there are two studies that have found that relationships environment and health behaviors increase after adjusting for selection (Boone-Heinonen et al., 2010; Zick et al., 2013), suggesting the importance of addressing the issue of nonrandom selection bias in estimating the association between neighborhood walkability and obesity. In particular, Zick and colleagues suggest that the mechanism behind this underestimation is complex but may involve competing dimensions of a desirable residential neighborhood. Following from their reasoning, it could be that walkability measures such as the ones captured by our walkability index are at odds with other desirable features such as school quality and housing costs and thus influence the residential choices of mothers included in this study.

Our study has several strengths. This investigation greatly benefits from the inclusion of a population-based sample of women, a design

Table 5

Neighborhood walkability residents and percent obese: unmatched and average treatment effect on the treated (ATT). Data for a sample of 103,912 women in Salt Lake County, Utah age 20 or older.

Variable	Sample	Treated	Controls	Difference	S.E.	95% C.I. for diff		t
	Unmatched	.121	.134	-.0105	.002	-.0107	-.0103	-4.34
	ATT	.119	.139	-.0201	.003	-.0204	-.0198	-6.03

Treatment in this table represents living in neighborhood where the walkability is in top quartile. The outcome of interest is percent obese.

feature not available to other investigations of implications of built environment on obesity. At the same time, it is important to note that our data come from one county and as such, caution should be used in generalizing to other locations that may have very different distributions of built environment. In addition, the use of the census block group as a definition of residential neighborhoods may over or under-approximate individuals' true neighborhoods though this remains an open question. We use birth certificate data that only include women who have had children so these findings are not generalizable to a population of childless women. Finally, we were able to employ a limited set of individual obesity risk factors. Future work should attempt to examine the role of residential selection in the estimated relationship between the built environment and obesity prevalence using samples from other locales and a more extensive list of individual level covariates.

A clear policy implication emerges from these analyses. Public health officials looking for levers to adjust Americans' risk of obesity should take note that our results suggest that walkable neighborhoods are associated with lower obesity prevalence. Moreover, this association is robust to tests of non-random residential selection. When designing new residential communities or renovating old communities, planners might be well served to include amenities that are often found in walkable neighborhoods. These communities are likely to be settings for lifestyles that promote healthy diets and physical activity.

Conflict of interest

The authors do not have any conflict of interest or any competing financial interests in relation to the work described. We acknowledge research support from grant number (R21 DK080406-01) National Institute for Health, The National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK).

Ethical statement

This research was reviewed and approved by the University of Utah Institutional Review Board. The authors do not have any conflict of interest or any competing financial interests in relation to the work described. We acknowledge research support from grant number (R21 DK080406-01) National Institute for Health, The National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) (R21 DK080406-01).

References

Berrigan, D., & Troiano, R. P. (2002). The association between urban form and physical activity in US adults. *American Journal of Preventive Medicine, 23*(2), 74–79.
Blackwell, M., Iacus, S. M., King, G., & Porro, G. (2009). cem: Coarsened exact matching in Stata. *The Stata Journal, 9*(4), 524–546.
Bodea, T. D., Garrow, L. A., Meyer, M. D., & Ross, C. L. (2009). Socio-demographic and built environment influences on the odds of being overweight or obese: The Atlanta experience. *Transportation Research Part A: Policy and Practice, 43*(4), 430–444.
Boer, R., Zheng, Y., Overton, A., Ridgeway, G. K., & Cohen, D. A. (2007). Neighborhood Design and Walking Trips in Ten U.S. Metropolitan Areas. *American Journal of Preventive Medicine, 32*(4), 298–304.
Boone-Heinonen, J., Gordon-Larsen, P., Guilkey, D. K., Jacobs, D. R., & Popkin, B. M. (2011). Environment and physical activity dynamics: The role of residential self-selection. *Psychology of Sport and Exercise, 12*, 54–60.
Boone-Heinonen, J., Guilkey, D. K., Evenson, K. R., & Gordon-Larsen, P. (2010). Residential self-selection bias in the estimation of built environment effects on physical activity between adolescence and young adulthood. *International Journal of Behavioral Nutrition and Physical Activity, 7*. https://doi.org/10.1186/1479-5868-7-70.
Brown, B. B., Smith, K. R., Hanson, H., Fan, J. X., Kowaleski-Jones, L., & Zick, C. D. (2013). Neighborhood design for walking and biking: Physical activity and body mass index. *American Journal of Preventive Medicine*.
Brown, S. C., Lombard, J., Toro, M., Huang, S., Perrino, T., Perez-Gomez, G., & Szapocznik, J. (2014). Walking and proximity to the urban growth boundary and central business district. *American Journal of Preventive Medicine, 47*(4), 481–486. https://doi.org/10.1016/j.amepre.2014.05.008.
Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys, 22*(1), 31–72. https://doi.org/10.1111/j.1467-6419.2007.00527.x.
Cao, X. (2010). Exploring causal effects of neighborhood type on walking behavior using stratification on the propensity score. *Environment and Planning A, 42*(2), 487–504.
Cao, X., Xu, Z., & Fan, Y. (2010). Exploring the connections among residential location, self-selection, and driving: Propensity score matching with multiple treatments. *Transportation Research Part A: Policy and Practice, 44*(10), 797–805.
Census Bureau U.S. (2000). Salt Lake County, Utah, Census 2000 Demographic Profile Highlights, Summary File 3. Retrieved (21 January 2008), from ⟨http://factfinder.census.gov⟩.
Cervero, R., & Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D-Transport and Environment, 2*(3), 199–219.
Chiu, M., Rezai, M.-R., Maclagan, L. C., Austin, P. C., Shah, B. R., Redelmeier, D. A., & Tu, J. V. (2016). Moving to a highly walkable neighborhood and incidence of hypertension: A propensity-score matched cohort study. *Environmental Health Perspectives (Online), 124*(6), 754.
Christakis, N. A., & Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England Journal of Medicine, 357*(4), 370–379 (doi: NEJMsa066082 (pii)(10.1056/NEJMsa066082).
Christian, H. E., Bull, F. C., Middleton, N. J., Knuiman, M. W., Divitini, M. L., Hooper, P., & Giles-Corti, B. (2011). How important is the land use mix measure in understanding walking behaviour? Results from the RESIDE study. *International Journal of Behavioral Nutrition and Physical Activity, 8*(1), 55.
Complete Health Indicator Report of Obesity Among Adults (2016). . *Public Health Indicator Based Information System (IBIS)*. Retrieved June 12, 2017, from ⟨https://ibis.health.utah.gov/indicator/complete_profile/Obe.html⟩.
Ewing, R., Brownson, R. C., & Berrigan, D. (2006). Relationship between urban sprawl and weight of United States youth. *American Journal of Preventive Medicine, 31*(6), 464–474.
Fan, J. X., Hanson, H. A., Zick, C. D., Brown, B. B., Kowaleski-Jones, L., & Smith, K. R. (2014). Geographic scale matters in detecting the relationship between neighbourhood food environments and obesity risk: An analysis of driver license records in Salt Lake County, Utah. *BMJ open, 4*(8), e005458.
Frank, L. D., Saelens, B. E., Powell, K. E., & Chapman, J. E. (2007). Stepping towards causation: Do built environments or neighborhood and travel preferences explain physical activity, driving, and obesity? *Social Science Medicine, 65*(9), 1898–1914.
Gebel, K., Bauman, A. E., Sugiyama, T., & Owen, N. (2011). Mismatch between perceived and objectively assessed neighborhood walkability attributes: Prospective relationships with walking and weight gain. *Health and Place, 17*(2), 519–524. https://doi.org/10.1016/j.healthplace.2010.12.008.
Gibson-Davis, C., & Foster, E. M. (2006). A cautionary tale: Using propensity scores to estimate the effect of food stamps on food insecurity. *Social Science Review,* 93–126.
Glazier, R. H., Creatore, M. I., Weyman, J. T., Fazli, G., Matheson, F. I., Gozdyra, P., & Booth, G. L. (2014). Density, destinations or both? A comparison of measures of walkability in relation to transportation behaviors, obesity and diabetes in Toronto, Canada. *PLoS One, 9*(1), https://doi.org/10.1371/journal.pone.0085295.
Handy (2006). Self-Selection in the Relationship between the Built Environment and Walking. *Journal of The American Planning Association, 72*, 55–74.
Hart, J. P., Arcaya, J. E., Feskanich, D., Laden, F., & Subramanian, S. V. (2015). Neighborhood self-selection: The role of pre-move health factors on the built and socioeconomic environment. *International journal of environmental research and public health, 12*(10), 12489–12504.
Hattori, A., An, R., & Sturm, R. (2013). Neighborhood Food Outlets, Diet, and Obesity Among California Adults, 2007 and 2009. *Preventing Chronic Disease, 10*.
Hirsch, J. A., Moore, K. A., Barrientos-Gutierrez, T., Brines, S. J., Zagorski, M. A., Rodriguez, D. A., & Diez Roux, A. V. (2014). Built environment change and change in BMI and waist circumference: Multi-ethnic Study of Atherosclerosis. *Obesity*. https://doi.org/10.1002/oby.20873.
Hu, H. H., Cho, J., Huang, G., Wen, F., Choi, S., Shih, M., & Lightstone, A. S. (2014). Neighborhood environment and health behavior in Los Angeles area. *Transport Policy, 33*, 40–47. https://doi.org/10.1016/j.tranpol.2014.02.004.
Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics, 86*(1), 4–29.
King, G., & Nielsen, R. (2016). Why propensity scores should not be used for matching. *Copy at http://j. mp/1sexgVw Download Citation BibTex Tagged XML Download Paper, 378*.
Leal, C., Bean, K., Thomas, F., & Chaix, B. (2011). Are associations between neighborhood socioeconomic characteristics and body mass index or waist circumference based on model extrapolations? *Epidemiology, 22*(5), 694–703.
Lee, I. M., Ewing, R., & Sesso, H. D. (2009). The built environment and physical activity levels: The Harvard Alumni Health Study. *American Journal of Preventive Medicine, 37*(4), 293–298.
Leuven, E., & Sianesi, B. (2015). PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. *Statistical Software Components*.
Li, F., Harmer, P., Cardinal, B. J., & Vongjaturapat, N. (2009). Built environment and changes in blood pressure in middle aged and older adults. *Preventive Medicine, 48*(3), 237–241.
Ludwig, J., Sanbonmatsu, L., Gennetian, L., Adam, E., Duncan, G. J., Katz, L. F., & McDade, T. W. (2011). Neighborhoods, obesity, and diabetes – A randomized social experiment. *New England Journal of Medicine, 365*(16), 1509–1519. https://doi.org/10.1056/NEJMsa1103216.
Mackenbach, J. D., Rutter, H., Compernolle, S., Glonti, K., Oppert, J. M., Charreire, H., & Lakerveld, J. (2014). Obesogenic environments: A systematic review of the association between the physical environment and adult weight status, the SPOTLIGHT project. *Bmc Public Health, 14*(1), https://doi.org/10.1186/1471-2458-14-233.
Mayne, S. L., Auchincloss, A. H., & Michael, Y. L. (2015). Impact of policy and built

environment changes on obesity-related outcomes: A systematic review of naturally occurring experiments. *Obesity Reviews, 16*(5), 362–375. https://doi.org/10.1111/obr.12269.

Michael, Y. L., Nagel, C. L., Gold, R., & Hillier, T. A. (2014). Does change in the neighborhood environment prevent obesity in older women? *Social Science and Medicine, 102*, 129–137. https://doi.org/10.1016/j.socscimed.2013.11.047.

Mujahid, M. S., Roux, A. V. D., Shen, M. W., Gowda, D., Sanchez, B., Shea, S., & Jackson, S. A. (2008). Relation between neighborhood environments and obesity in the multi-ethnic study of atherosclerosis. *American Journal of Epidemiology, 167*(11), 1349–1357.

Müller-Riemenschneider, F., Pereira, G., Villanueva, K., Christian, H., Knuiman, M., Giles-Corti, B., & Bull, F. C. (2013). Neighborhood walkability and cardiometabolic risk factors in australian adults: An observational study. *Bmc Public Health, 13*(1), 755.

Nunnally, J. (1978). *Psychometric theory* (2nd ed.). New York: McGraw Hill.

Oakes, J. M., & Johnson, P. J. (2006). Propensity score matching for social epidemiology. *Methods in social epidemiology, 1*, 370–393.

Pinjari, A. R., Bhat, C. R., & Hensher, D. A. (2009). Residential self-selection effects in an activity time-use behavior model. *Transportation Research Part B: Methodological, 43*(7), 729–748.

Rodríguez, D. A., Evenson, K. R., Diez Roux, A. V., & Brines, S. J. (2009). Land Use, Residential Density, and Walking. The Multi-Ethnic Study of Atherosclerosis. *American Journal of Preventive Medicine, 37*(5), 397–404.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41–55. https://doi.org/10.1093/biomet/70.1.41.

Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association, 79*(387), 516–524. https://doi.org/10.1080/01621459.1984.10478078.

Ross, N. A., Crouse, D., Tremblay, S., Khan, S., Tremblay, M., & Berthelot, J.-M. (2007). Body Mass Index in urban Canada: Neighborhood and metropolitan area effects. *American Journal of Public Health, 97*(3), 500–508.

Rundle, A., Roux, A. V. D., Freeman, L. M., Miller, D., Neckerman, K. M., & Weiss, C. C. (2007). The urban built environment and obesity in New York City: A multilevel analysis. *American Journal of Health Promotion, 21*(4), 326–334.

She, Z., King, D. M., & Jacobson, S. H. (2017). Analyzing the impact of public transit usage on obesity. *Preventive Medicine, 99*, 264–268. https://doi.org/10.1016/j.ypmed.2017.03.010.

Smith, J. A., & Todd, P. E. (2005). Does matching overcome LaLonde's critique of nonexperimental estimators? (1-2 SPEC. ISS.) *Journal of Econometrics, 125*, 305–353. https://doi.org/10.1016/j.jeconom.2004.04.011.

Smith, K. R., Brown, B. B., Yamada, I., Kowaleski-Jones, L., Zick, C. D., & Fan, J. X. (2008). Walkability and Body Mass Index: Density, design, and new diversity measures. *American Journal of Preventive Medicine, 35*(3), 237–244.

Smith, K. R., Hanson, H. A., Brown, B. B., Zick, C. D., Kowaleski-Jones, L., & Fan, J. X. (2016). Movers and stayers: How residential selection contributes to the association between female body mass index and neighborhood characteristics. *International Journal of Obesity, 40*(9), 1384–1391. https://doi.org/10.1038/ijo.2016.78.

Smith, K. R., Zick, C. D., Kowaleski-Jones, L., Brown, B. B., Fan, J. X., & Yamada, I. (2011). Effects of neighborhood walkability on healthy weight: Assessing selection and causal influences. *Social Science Research, 40*(5), 1445–1455. https://doi.org/10.1016/j.ssresearch.2011.04.009.

Southworth, M., & Ben-Joseph, E. (1995). Street standards and the shaping of suburbia. *Journal of The American Planning Association, 61*, 65–81.

Terwee, C. B., Mokkink, L. B., van Poppel, M. N. M., Chinapaw, M. J. M., van Mechelen, W., & de Vet, H. C. W. (2010). Qualitative attributes and measurement properties of physical activity questionnaires: a checklist. *Sports Medicine, 40*(7), 525–537 (510.2165/11531370-000000000-000000000).

Wen, M., & Kowaleski-Jones, L. (2012). The built environment and risk of obesity in the United States: Racial–ethnic disparities. *Health Place, 18*(6), 1314–1322. https://doi.org/10.1016/j.healthplace.2012.09.002.

Xu, Y., Wen, M., & Wang, F. (2014). Multilevel built environment features and individual odds of overweight and obesity in Utah. *Applied Geography*(60), https://doi.org/10.1016/j.apgeog.2014.10.006.

Yamada, I., Brown, B. B., Smith, K. R., Zick, C. D., Kowaleski-Jones, L., & Fan, J. X. (2012). Mixed land use and obesity: An empirical comparison of alternative land use measures and geographic scales. *Professional Geographer, 64*(2), 157–177. https://doi.org/10.1080/00330124.2011.583592.

Zhang, X., Holt, J. B., Lu, H., Onufrak, S., Yang, J., French, S. P., & Sui, D. Z. (2014). Neighborhood commuting environment and obesity in the United States: An urban–rural stratified multilevel analysis. *Preventive Medicine, 59*(0), 31–36. https://doi.org/10.1016/j.ypmed.2013.11.004.

Zick, C. D., Hanson, H. A., Fan, J. X., Smith, K. R., Kowaleski-Jones, L., Brown, B. B., & Yamada, I. (2013). Re-visiting the relationship between neighbourhood environment and BMI: An instrumental variables approach to correcting for residential selection bias. *International Journal of Behaviour Nutrition Physical Activity, 10*, 27. https://doi.org/10.1186/1479-5868-10-27.