6-2017

Analyzing the impact of public transit usage on obesity

Zhaowei SHE
Singapore Management University, zwshe@smu.edu.sg

Douglas M. KING

Jacobson Sheldon SHELDON

Preventive Medicine 99 (2017) 264–268

Contents lists available at ScienceDirect

Preventive Medicine

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

Analyzing the impact of public transit usage on obesity

Zhaowei She [a], Douglas M. King [b,*], Sheldon H. Jacobson [c]

[a] H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, United States
[b] Department of Industrial and Enterprise Systems Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, United States
[c] Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL, United States

ARTICLE INFO

Article history:
Received 15 August 2016
Received in revised form 27 February 2017
Accepted 12 March 2017
Available online 16 March 2017

Keywords:
Obesity
Transportation
Environment design

ABSTRACT

The objective of this paper is to estimate the impact of county-level public transit usage on obesity prevalence in the United States and assess the potential for public transit usage as an intervention for obesity. This study adopts an *instrumental regression* approach to implicitly control for potential *selection bias* due to possible differences in commuting preferences among obese and non-obese populations. United States health data from the 2009 Behavioral Risk Factor Surveillance System and transportation data from the 2009 National Household Travel Survey are aggregated and matched at the county level. County-level public transit accessibility and vehicle ownership rates are chosen as *instrumental variables* to implicitly control for unobservable commuting preferences. The results of this instrumental regression analysis suggest that a one percent increase in county population usage of public transit is associated with a 0.221 percent decrease in county population obesity prevalence at the $\alpha = 0.01$ statistical significance level, when commuting preferences, amount of non-travel physical activity, education level, health resource, and distribution of income are fixed. Hence, this study provides empirical support for the effectiveness of encouraging public transit usage as an intervention strategy for obesity.

1. Introduction

Recent studies show that people's commuting choices are associated with their obesity status; more driving is positively associated with obesity prevalence, while higher public transit usage is negatively associated with obesity prevalence (Edwards, 2008; Besser and Dannenberg, 2005; Behzad et al., 2013; Jacobson et al., 2011). As such, the Centers for Disease Control and Prevention (CDC) encourages public transit usage as a possible obesity intervention strategy (Centers for Disease Control and Prevention, 2009; Centers for Disease Control and Prevention, 2011). This strategy's effectiveness can be affected by confounding factors; if the obese population has significantly lower preference for public transportation, a potential increase in public transit usage may not translate into lower obesity prevalence, since this increase is less likely to come from the obese population. Therefore, to justify obesity interventions based on encouraging public transit usage, it is important to understand whether the negative association between public transit usage and obesity prevalence is independent of confounding factors.

Two commonly-discussed confounding factors in this association are selection bias and potential substitution effects between travel-related and non-travel physical activity. Selection bias refers to the possibility that unobservable differences in people's commuting preferences can affect the estimated association between public transit usage and obesity prevalence. For example, Eid et al. (2008) found that people who are obese tend to prefer living in more sprawling neighborhoods, while Plantinga and Bernell (2007) noted that public transportation is less viable in these neighborhoods. In this case, obesity could be a cause of lower public transit usage, rather than a result of lower public transit usage; a simple statistical model associating obesity and public transit usage would only estimate how less likely an obese individual commutes via public transit, instead of the impact of public transit usage on obesity. Another source of confounding is the possible substitution effect between travel-related and non-travel physical activity, such that increasing travel-related physical activity may reduce non-travel physical activity (Saunders et al., 2013). For example, when returning home from a bus ride, one may be either too tired or not have sufficient time for additional physical exercises. In this case, an overweight individual may prefer driving to taking public transit even if they wish to lose weight. As such, the impact of public transit usage on obesity is inconclusive if the negative association between public transit usage and obesity is due to confounding effects.

To address possible self-selection estimation bias, this study proposes an *instrumental regression*, or two-stage least squares (2SLS) regression approach to estimate the impact of public transit usage on obesity prevalence at the county population level. In the estimation,

* Corresponding author at: 117 Transportation Building, MC-238, 104 S. Mathews Ave., Urbana, IL 61801, United States.
E-mail address: dmking@illinois.edu (D.M. King).

http://dx.doi.org/10.1016/j.ypmed.2017.03.010
0091-7435/

amount of non-travel physical activity, health resource, and distribution of income are explicitly controlled through data from multiple sources. Unobserved commuting preferences are implicitly controlled though two traffic-related instrumental variables: public transit accessibility and vehicle ownership rates. Hence, this approach focuses on people forced to use public transit due to traffic constraints. Therefore, variations in public transit usage due to commuting preferences have been statistically ruled out, and hence, should not bias the estimation. As such, this study addresses the limitations of earlier studies (Frank et al., 2007; Tiemann and Miller, 2013) and provides further evidence of the negative impact of public transit usage on obesity. By separating the impact of public transit usage on obesity from potential confounding effects, this study provides further support for the public health efforts to reduce obesity prevalence through encouraging public transit usage.

2. Methods

2.1. Data sets and data pre-processing

This study gathers and matches county-level aggregated health and transportation data from multiple sources. Health related variables are calculated from the 2009 Behavioral Risk Factor Surveillance System (BRFSS) (Centers for Disease Control and Prevention, 2015). Surveys of BRFSS have been conducted annually since 1984 by the Centers for Disease Control and Prevention (CDC) and other federal agencies through a nationwide random sample (one per household) of adults (18 + years) in the United States. Health data capture obesity status, and its associated risk factors, with health variables defined as:

- *OBESE*: Percentage of county population with Body Mass Index (BMI) at least 30 kg/m^2; (Ogden et al., 2014)
- *LTPA*: Percentage of county population engaging in leisure time physical activity (e.g., running, calisthenics, golf, gardening, walking);
- *Employed%*: Binary variable, with 1 indicating data point collected from respondents who were employed/self-employed in 2009; 0 otherwise;
- *Education:* Percentage of county population with education above the high school level (at least one year of college education)
- *Healthcare*: Percentage of county population with health care coverage (e.g., health insurance, prepaid plans, or Medicare).

The 2009 National Household Travel Survey (NHTS) database provides variables related to transportation patterns (U.S. Department of Transportation, Federal Highway Administration, 2009). The National Household Travel Survey is conducted to examine travel behavior at the individual and household level in the United States, and is publicly accessible through a database published by Federal Highway Administration of the U.S. Department of Transportation. This study utilizes a special research version with more detailed geographic information; to remain consistent with the age limits of the BRFSS, all individuals with age below 18 years are excluded. Transportation data describe transportation patterns and transit mode choice, with transportation variables defined as:

- *Transit%*: Percentage of the county population using public transit at least eleven times per month (i.e., two or more days a week);
- *Transit_Important%*: Percentage of the county population ranking accessibility/availability of public transit as their most important transportation issue, compared to other issues like highway congestion, lack of walkways or sidewalks, price of travel, aggressive/distracted drivers and safety concerns;
- *AverageVehicle*: Average number of vehicles per household at county level;
- *Rail*: Binary variable, with 1 indicating data point collected from respondents residing in a metropolitan area with subway/rail; 0 otherwise;
- *Employed%*: Same as for Health data.

This study also includes data to control for social-economic factors and spatial correlations in the associations between obesity and public transit usage. To control for income, this study includes *Income* (county level median household income) and *Poverty* (percentage of county population that lives below the poverty threshold) (United States Census Bureau, 2015). The variable *Income* is obtained from the U.S. Census Bureau (United States Census Bureau, Small Area Estimates Branch, 2009), as median income statistics for each county cannot be computed from either BRFSS or NHTS, which only provides a range, instead of the exact number, of each interviewee's income level. The U.S. Census Bureau derived this *Income* estimate through combining the decennial census and the direct estimates from the American Community Survey. The variable *Poverty* is computed as the average of estimates from BRFSS and NHTS. The *Poverty* estimate is updated by the U.S. Census Bureau each year using the change in the average annual Consumer Price Index for All Urban Consumers. To control for possible spatial correlations between county observations, this study includes fixed effects for a vector $\overrightarrow{State}$, which describes the state in which each county is located; hence, possible confounding effects due to spatial closeness can be addressed.

These data sets are aggregated and matched based on two identifier types: At the county aggregate level (given the large sample size in each county), and according to their employment status (*Employed%*) (to control for the difference in occupational physical activities and leisure time physical activities). Each county-level statistic is a weighted average of at least 30 individual observations from the raw datasets, with 318 counties from 44 U.S. states represented. Table 1 summarizes the descriptive statics of relevant variables.

2.2. Statistical analysis

This study uses 2SLS regression to address the potential influence of self-selection bias. Self-selection bias cannot be controlled explicitly through an ordinary least squares model, because subjective motives (e.g., personal preferences) are often not evaluated in nationwide surveys. The advantage of 2SLS regression is its ability to control for potential confounding variables without direct estimations of these variables (Wooldridge, 2012). Conceptually, one can understand 2SLS regressions as “causal path analysis” (Angrist and Krueger, 2001). From Fig. 1, personal preference (PP) for a sedentary lifestyle can simultaneously influence transit mode choice (PT) and obesity (OB), and cannot be explicitly controlled with the available data. To address this confounding effect, a vector of instrumental variables (IV) would be needed, with variations in IV only associated with variations in OB through PT. For example, in a study on the causal effect of obesity on wages, Cawley (2004) uses maternal body weight as an instrumental variable to estimate the causal impact of females' body weight on their wages. Here Cawley assumes that maternal body weight can only associate with females' wages through body weight inheritance. By regressing mother's BMI on daughter's BMI, he obtained a predicted value of daughter's BMI in the first stage of the 2SLS model. In the second stage, he regresses wage outcomes against this predicted BMI and other control variables to obtain unbiased estimates of the impact of body weight on wage outcomes. In this case, he implicitly controlled for risk factors in obesity due to low wages, for example unhealthy food, because maternal body weight can only change females' inherited body weight and has no impact on other risk factors in obesity due to low wages. A similar 2SLS approach is adopted in this study.

In the first stage regression of our 2SLS model, *Transit_Important%* and *AverageVehicle* are the instrumental variables chosen to characterize a county's traffic constraints. Regardless of their commuting preferences, people living in a county with high *Transit_Important%* are more

Table 1
Descriptive statistics calculated from a sample of the adult population from 318 counties across 44 U.S. states using 2009 BRFSS and NHTS.

Variable	Units	Employed			Unemployed		
		Average	Max	Min	Average	Max	Min
OBESE	%	26.78	51.22	7.57	27.64	53.86	6.1
LTPA	%	80.59	97.49	58.35	71.99	93.32	31.37
Education	%	72.04	95.52	39.84	54.99	88.72	22.44
Healthcare	%	87.84	100	49.76	82.91	99.37	37.55
Transit%	%	6.28	72.25	0	4.70	51.62	0
Transit_Important%	%	7.71	42.22	0	8.18	35.82	0
AverageVehicle	# vehicles	2.35	3.40	0.43	1.89	3.23	0.33
Income	$	53,457.59	102,325	30,360	50,874.02	102,325	27,421
Poverty	%	10.15	40.01	1.13	23.91	56.78	3.31

likely to have a high dependency on public transportations. Similarly, regardless of their commuting preferences, people living in a county with high *AverageVehicle* are less likely to have high dependency on public transportations. Bagley and Mokhtarian (2002) provide a logical explanation of this association between traffic constraints and public transit usage: people living in areas with heavy traffic (e.g., large metropolitan areas) or areas with efficient public transit infrastructure (e.g., college towns) have higher dependency on public transportation, while people living in areas with light traffic (e.g. small rural areas) or areas with less developed public transit infrastructure have lower dependency on public transportation. This type of public transit dependency is unlikely to directly associate with individual commuting preferences or obesity status, and hence, serves as an ideal candidate for instrumental variables. In other words, these two instrumental variables can only associate with obesity through public transit usage. Moreover, while an individual's commuting preferences can influence their choice of residential community, this influence is likely more influential over their choice of community within a county, rather than their choice between different counties; hence, the county-level analysis suppresses any influence of these preferences. Therefore, variations in public transit usage due to commuting preferences have been statistically ruled out in this study, and hence, cannot bias the estimation.

Several control variables are also included in our 2SLS model to account for other covariates: *LTPA* is a proxy for non-travel related physical activity, *Employed%* controls for occupational activity, *Education* captures countywide environmental and public health awareness, *Healthcare* is a surrogate for medical resources, *Income* controls for overall county income level, *Poverty* approximates variation of county income, and *Rail* and $\overrightarrow{State}$ capture spatial fixed effects, where $\overrightarrow{State}$ is a 43×1 binary vector.

In the first stage of the 2SLS regression,

$$\begin{aligned} Transit\% = {} & \beta_0' + \beta_1' \; Transit\_Important\% + \beta_2' \; AverageVehicle \\ & + \beta_3' \; LTPA + \beta_4' \; LTPA \times Employed\% + \beta_5' \; Education \\ & + \beta_6' \; Healthcare + \beta_7' \; Income + \beta_8' \; Poverty + \beta_9' \; Rail \\ & + \overrightarrow{\beta_{10}'}^T \; \overrightarrow{State} + \epsilon', \end{aligned} \tag{1}$$

provides a prediction of *Transit%* from instrumental variables. The predicted value $\widehat{Transit}\%$ replaces the two instrumental variables in the second stage regression,

$$\begin{aligned} OBESE = {} & \beta_0 + \beta_{Transit\%} \; \widehat{Transit}\% + \beta_{LTPA} \; LTPA + \beta_{LTPA \times Employed\%} \; LTPA \\ & \times Employed\% + \beta_{Education} \; Education + \beta_{Healthcare} \; Healthcare \\ & + \beta_{Income} \; Income + \beta_{Poverty} \; Poverty + \beta_{Rail} \; Rail + \overrightarrow{\beta_S}^T \; \overrightarrow{State} \\ & + \epsilon, \end{aligned} \tag{2}$$

while the control variables vector from the first stage remains unchanged in the second stage. Here, the model parameters in the second stage regression are interpreted as:

- $\beta_{Transit\%}$: Percentage change in county population obesity prevalence associated with a 1% increase of frequent public transit riders in the county population;
- β_{LTPA}: Percentage change in county population obesity prevalence associated with a 1% increase of people that engage in leisure time physical activity in the unemployed county population;
- $\beta_{LTPA} + \beta_{LTPA \times Employed\%}$: Percentage change in county population obesity prevalence associated with a 1% of people that engage in leisure time physical activity in the employed county population;
- $\beta_{Education}$: Percentage change in county population obesity prevalence associated with a 1% increase of percentage of the county's population with at least one year of college education.
- $\beta_{Healthcare}$: Percentage change in county population obesity prevalence associated with a 1% increase of health care coverage in the county population.
- β_{Income}: Percentage change in county population obesity prevalence associated with a \$1 increase in county median annual household income;
- $\beta_{Poverty}$: Percentage change in county population obesity prevalence associated with a 1% increase in the county poverty rate.

All parameters from the first stage regression (1), as well as the intercept (β_0), rail fixed effect (β_{Rail}), and state-specific fixed effects ($\overrightarrow{\beta_S}^T$) from the second stage regression (2) are included to facilitate estimation, with no specific meanings attached. Analysis was conducted in 2015 using the *R* programming language.

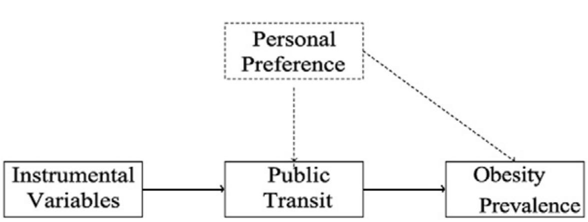

Fig. 1. Direct arrows in the graph stands for causal directions. The dashed circle and arrows indicate that these effects cannot be explicitly controlled with the available data.

3. Results

The 2SLS model results from Eq. (2) are reported in Table 2. The results show two findings. First, public transit usage is negatively associated with obesity after controlling for selection bias. Specifically, a 1% increase of frequent public transit riders in the county population is estimated to decrease county population obesity prevalence by 0.221% ($\alpha = 0.01$ significance level). Second, the impact of public transit usage on obesity prevalence is comparable to the impact of physical activities on obesity prevalence among the employed group; the estimated value of $\beta_{LTPA} + \beta_{LTPA \times Employed\%}$ quantifies this effect to be 0.186% ($\alpha = 0.01$ statistical significance level). Hence, among a county's employed subpopulation, a 1% increase in the number of people engaging in leisure time

Table 2
Estimates represent the percentage change in United States county population obesity prevalence associated with one unit increase in each factor, based on 2009 health data and transportation data. The unit in every factor in this table is percentage, except *Income*, whose unit is dollars.

Factor	Parameter	Estimate	Standard error	*p* value
(Intercept)	β_0	57.2	8.95	<0.001
Public transit usage	$\beta_{Transit\%}$	−0.221	0.0732	0.002
LTPA	β_{LTPA}	−0.254	0.0498	<0.001
LTPA among employed	$\beta_{LTPA \times Employed\%}$	0.0682	0.0119	<0.001
Education	$\beta_{Education}$	−0.157	0.0405	<0.001
Healthcare	$\beta_{Healthcare}$	0.143	0.0482	0.003
Income	β_{Income}	-1.08×10^{-4}	3.97×10^{-5}	0.007
Poverty	$\beta_{Poverty}$	0.123	0.0597	0.04

physical activity is estimated to decrease the population's obesity prevalence by 0.186%. Among the unemployed subpopulation, a 1% increase in the number of people engaging in leisure time physical activity is estimated to decrease the population's obesity prevalence by 0.254%.

4. Discussion

By combining health and transportation data, this paper provides a better understanding of the impact of public transit usage on obesity by quantifying the impact of additional physical activity involved in public transit usage on obesity, and addressing the limitations in previous research on the impact of public transit usage on obesity.

First, this study further confirms the negative associations between public transit usage and obesity reported in the literature (Edwards, 2008; Besser and Dannenberg, 2005; Frank et al., 2007; Tiemann and Miller, 2013; Flint et al., 2014). For the employed subpopulation, increasing public transit usage may be an equally effective strategy in losing weight in comparison to increasing leisure time physical activity. The overall negative association between leisure time physical activity and obesity is consistent with Tiemann and Miller (2013) and Flint et al. (2014). The effect difference in employed and unemployed groups is also supported by previous research, indicating that work-related physical activity can substitute the impact of leisure-time physical activity on obesity (Coleman and Dave, 2014). Moreover, though occupation activities differ in the amount of physical activities involved, this study did not find statistical differences in the impact of public transit usage on obesity among different types of occupations. Though some county residents may not be able to engage in public transit (e.g., workers whose jobs require them to either transport equipment or who otherwise need to commute by private car), thereby restricting increases in countywide *Transit%*, the estimation results support the hypothesis that increasing public transit usage alone is an effective strategy in reducing obesity prevalence.

Second, while it is believed that negative association between obesity and public transit usage is due to the extra physical activity associated with taking public transit, it is not clear from the data whether there exist substitution effects between travel-related and non-travel physical activity (Saunders et al., 2013). Some recent studies have answered this question through questionnaire surveys (Sahlqvist et al., 2013), natural experiment design (Miller et al., 2015) and analysis of data from mobile sensors (Saelens et al., 2014), and show that the proposed substitution effects are not supported by their data. However, all previous studies are only based on regional data, and thus may not have nationwide implications. This study conducts a nationwide study on this effect, by explicitly controlling for the impacts of recreational and occupational physical activities on obesity with variables *LTPA* and *LTPA* × *Employed%* with the 2SLS model in Eqs. (1) and (2). Even with levels of other non-travel physical activities explicitly controlled, the estimated values of $\beta_{Transit\%}$, β_{LTPA} and $\beta_{LTPA} + \beta_{LTPA \times Employed\%}$ are still jointly statistically significant at the $\alpha = 0.01$ level with negative signs. Therefore, the result in Table 2 shows that extra physical activity involved in public transit usage does not offset other non-travel physical activities in terms of reducing obesity. This result is also confirmed by the multicollinearity test, where VIF scores only need to fall below ten to justify the linear independence assumption between regressors (Allen, 1997). In this study, the VIF scores of $\beta_{Transit\%}$, β_{LTPA} and $\beta_{LTPA \times Employed\%}$ are all below four, suggesting little evidence of linear dependency between these variables. Therefore, this study quantifies the impact of the extra physical activity involved in public transit usage on obesity, and finds no substitution effect between travel-related and non-travel physical activity.

Third, this study addresses limitations in previous research on the impact of public transit usage on obesity. Tiemann and Miller (2013) employs a 2SLS regression model, but with population density and race distribution as instrumental variables, which has been criticized in the literature (Plantinga and Bernell, 2007; Eid et al., 2008). Eid et al. (2008) suggests that individual BMI exhibits no statistically significant change when a person moves between dense and less dense areas, and hence, built environment features such as population density and race distribution have no causal impact on obesity. One possible explanation is that while high density areas have strong public transit dependency, the association between density and public transit dependency area can be very weak in low density areas. Since most counties in the United States are not very population dense (e.g., two thirds of the counties in this study have population density < 1000 people/mi^2), population density is not an appropriate proxy of public transit dependency in a nationwide study. Therefore, Tiemann and Miller (2013) did not fully address the self-selection issue. Frank et al. (2007) provides a more convincing answer to the self-selection problem, explicitly controlling for potential self-selection bias through questionnaires that assess residents' commuting preferences. However, their data are only collected in Atlanta, GA, and may not be representative enough to draw nationwide conclusions. Our results are consistent with this literature and, by design, allow conclusions to be drawn regarding the association between obesity and public transit at the national level.

Finally, note that not all estimated coefficients in Eq. (2) are independent of confounding effects. Particularly, the 0.143 estimate of $\beta_{Healthcare}$ does not necessarily mean that a 1% increase of health care coverage in the county population will increase the county population obesity prevalence by 0.143%. On one hand, some research indeed states that people with health care coverage tend to have a moral hazard problem of living an unhealthy lifestyle and thus are more likely to be obese (Bhattacharya and Sood, 2007). Conversely, this estimate can also mean that obese population has higher demands for medical resources, and hence, are more likely to have health care coverage. Such confounding effects, which do not directly affect the association between public transit usage and obesity prevalence, fall out of the scope of this study.

5. Conclusions

This study establishes a statistically significant negative association between public transit usage and obesity prevalence, and shows that common confounding factors such as selection bias do not affect the result. This result suggests that increasing public transit usage alone can effectively decrease population obesity prevalence, providing further empirical support for encouraging public transit usage as an intervention strategy for obesity based on community designs.

The analyses presented in this study are limited due to many relevant variables being absent from the available survey data, and a more comprehensive dataset can possibly provide a clearer picture of the relation between transit mode choice and obesity prevalence. Nevertheless, the estimated association between public transit usage and obesity should be robust in the current model. To account for possible omitted variables, geographical (i.e. $\overrightarrow{State}$) and metropolitan status (i.e. *Rail*) binary variables are added as fixed effects; the association between public transit usage and obesity is statistically significant ($\alpha = 0.01$ level) even with these fixed effects added.

The results reported are limited by the cross sectional data used. Though several potential confounding effects are controlled by instrumental and controlled variables, it remains to be seen whether public transit usage's impact on obesity is indeed causal, and the 2SLS model estimation may still be biased. Moreover, it is not possible to directly test the unbiasedness hypothesis statistically with cross sectional data. In contrast, with panel or longitudinal data, all time invariant omitted variables can be tested and controlled implicitly (Frank et al., 2007; Flint et al., 2014). Therefore a longitudinal study of the impact of public transit usage on obesity can be a step toward establishing causality.

Conflicts of interest

Mr. She has nothing to disclose. Dr. King has nothing to disclose. Dr. Jacobson has nothing to disclose.

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgments

The computational work was conducted with support from the Simulation and Optimization Laboratory and the Bed Time Research Institute at the University of Illinois.

References

Allen, M., 1997. The problem of multicollinearity. Understanding Regression Analysis. Springer US, New York, pp. 176–180.
Angrist, J., Krueger, A., 2001. Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments (No. w8456). National Bureau of Economic Research http://economics.mit.edu/files/18 Accessed June 24, 2016.
Bagley, M., Mokhtarian, P., 2002. The impact of residential neighborhood type on travel behavior: a structural equations modeling approach. Ann. Reg. Sci. 36 (2):279–297. http://dx.doi.org/10.1007/s001680200083.
Behzad, B., King, D., Jacobson, S., 2013. Quantifying the association between obesity, automobile travel, and caloric intake. Prev. Med. 56 (2):103–106. http://dx.doi.org/10.1016/j.ypmed.2012.11.022.
Besser, L., Dannenberg, A., 2005. Walking to public transit: steps to help meet physical activity recommendations. Am. J. Prev. Med. 29 (4):273–280. http://dx.doi.org/10.1016/j.amepre.2005.06.010.
Bhattacharya, J., Sood, N., 2007. Health insurance and the obesity externality. Adv Health Econ Health Serv Res Suppl. 17 (6):279–318. http://dx.doi.org/10.1016/S0731-2199(06)17011-9.
Cawley, J., 2004. The impact of obesity on wages. J. Hum. Resour. 39 (2):451–474. http://dx.doi.org/10.2307/3559022.
Centers for Disease Control and Prevention, 2009. Recommended community strategies and measurements to prevent obesity in the United States. Morb. Mortal. Wkly Rep. 58 (7):1–29. http://www.cdc.gov/obesity/downloads/community_strategies_guide.pdf Accessed June 24, 2016.
Centers for Disease Control and Prevention, 2011. Impact of the Built Environment on Health. http://www.cdc.gov/nceh/publications/factsheets/impactofthebuiltenvironmentonhealth.pdf Update June 2011. Accessed June 24, 2016.
Centers for Disease Control and Prevention, 2015. Behavioral Risk Factor Surveillance System. http://www.cdc.gov/brfss/annual_data/annual_data.htm Updated May 2015. Accessed June 24, 2016.
Coleman, G., Dave, D., 2014. Unemployment and health behaviors over the business cycle: a longitudinal view. National Bureau of Economic Research, Tech. Rep. w20748 http://dx.doi.org/10.3386/w20748.
Edwards, R., 2008. Public transit, obesity, and medical costs: assessing the magnitudes. Prev. Med. 46 (1):14–21. http://dx.doi.org/10.1016/j.ypmed.2007.10.004.
Eid, J., Overman, H., Puga, D., Turner, M., 2008. Fat city: questioning the relationship between urban sprawl and obesity. J. Urban Econ. 63 (2):385–404. http://dx.doi.org/10.1016/j.jue.2007.12.002.
Flint, E., Cummins, S., Sacker, A., 2014. Associations between active commuting, body fat, and body mass index: population based, cross sectional study in the United Kingdom. BMJ http://dx.doi.org/10.1136/bmj.g4887.
Frank, L., Saelens, B., Powell, K., Chapman, J., 2007. Stepping towards causation: do built environments or neighborhood and travel preferences explain physical activity, driving, and obesity? Soc. Sci. Med. 65 (9):1898–1914. http://dx.doi.org/10.1016/j.socscimed.2007.05.053.
Jacobson, S., King, D., Yuan, R., 2011. A note on the relationship between obesity and driving. Transp. Policy 18 (5):772–776. http://dx.doi.org/10.1016/j.tranpol.2011.03.008.
Miller, H.J., Tribby, C.P., Brown, B.B., et al., 2015. Public transit generates new physical activity: evidence from individual GPS and accelerometer data before and after light rail construction in a neighborhood of Salt Lake City, Utah, USA. Health Place 36:8–17. http://dx.doi.org/10.1016/j.healthplace.2015.08.005.
Ogden, C., Carroll, M., Kit, B., Flegal, K., 2014. Prevalence of childhood and adult obesity in the United States, 2011–2012. JAMA 311 (8):806–814. http://dx.doi.org/10.1001/jama.2014.732.
Plantinga, A., Bernell, S., 2007. The association between urban sprawl and obesity: is it a two-way street? J. Reg. Sci. 47 (5):857–879. http://dx.doi.org/10.1111/j.1467-9787.2007.00533.x.
Saelens, B.E., Vernez Moudon, A., Kang, B., Hurvitz, P.M., Zhou, C., 2014. Relation between higher physical activity and public transit use. Am. J. Public Health 104 (5):854–859. http://dx.doi.org/10.2105/AJPH.2013.301696.
Sahlqvist, S., Goodman, A., Cooper, A.R., Ogilvie, D., 2013. Change in active travel and changes in recreational and total physical activity in adults: longitudinal findings from the iConnect study. Int. J. Behav. Nutr. Phys. Act. 10 (1):1. http://dx.doi.org/10.1186/1479-5868-10-28.
Saunders, L., Green, J., Petticrew, M., Steinbach, R., Roberts, H., 2013. What are the health benefits of active travel? A systematic review of trials and cohort studies. PLoS One 8 (8), e69912. http://dx.doi.org/10.1371/journal.pone.0069912.
Tiemann, T., Miller, P., 2013. Reducing sprawl, riding the bus and losing weight in America. Spaces & Flows: An International Journal of Urban & Extra Urban Studies 3 (3).
U.S. Department of Transportation, Federal Highway Administration, 2009. 2009 National Household Travel Survey. http://nhts.ornl.gov Updated 2009. Accessed June 24, 2016.
United States Census Bureau, 2015. Poverty Thresholds by Size of Family and Number of Related Children Under 18 Years. https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-thresholds.html Updated 2015. Accessed June 24, 2016.
United States Census Bureau, Small Area Estimates Branch, 2009. Poverty and median income estimates – counties. https://www.census.gov/did/www/saipe/data/statecounty/data/2009.html Updated 2010. Accessed June 24, 2016.
Wooldridge, J., 2012. Instrumental variables estimation and two-stage least squares. Int. Econ.:510–551. http://fmwww.bc.edu/EC-C/F2012/228/EC228.f2012.nn15.pdf Accessed June 24, 2016.