"Transit Makes you Short": On Health Impact Assessment of Transportation and the Built Environment

Alireza Ermagun[a] and David Levinson[b]

[a] PhD, University of Minnesota, Department of Civil, Environmental, and Geo-Engineering, 500 Pillsbury Drive SE, Minneapolis, MN 55455 USA. Email: ermag001@umn.edu (*Corresponding Author*)

[b] Director of Network, Economics, and Urban Systems Research Group; University of Minnesota, Department of Civil, Environmental, and Geo-Engineering, 500 Pillsbury Drive SE, Minneapolis, MN 55455 USA, Email: dlevinson@umn.edu

Abstract

This study conducts an in-depth analysis to alert policymakers and practitioners to erroneous results in the positive impacts of transit use on health measures. We explore the correlation of transit use and accessibility by transit with self-reported general health, Body Mass Index (BMI), and height. We develop a series of linear regression and binary logit models. We also depict the coefficient-p-value-sample-size chart, and conduct the effect size analysis to scrutinize the practically significant impacts of transit use and accessibility by transit on health measures. The results indicate transit use and accessibility by transit are significantly associated with general health and BMI. However, they are practically insignificant, and the power of the large sample in our particular case causes the statistically insignificant variable to become significant. At a deeper level, a 1% increase in transit use at the county level diminishes the BMI by only 0.0037% on average. The elasticity of transit use also demonstrates that every 1% increase in transit use would escalate the chance of having excellent or very good general health by 0.0003%. We show there is a thin line between false positive and true negative results. We alert both researchers and practitioners to the dangerous pitfalls deriving from the power of large samples and the weakness of p-values. Building the results on just statistical significance and sign of the parameter of interest is worthless, unless the magnitude of effect size is carefully quantified post analysis.

1 Introduction

Many consider transit an active mode of travel (Rissel et al., 2012), and thereby recommend transit as a gateway for improving physical activity levels (Julien, 2015; Besser and Dannenberg, 2005). Sener et al. (2016) reviewed the existing literature in the nexus of transit use and physical activity. They concluded that there is an unshakable positive correlation between transit use and physical activity, while the magnitude of this correlation is uncertain. Wasfi et al. (2013) also emphasized that public transit users, particularly train users, can meet the daily recommended minutes of physical activity by walking to and from transit stops. This result is based on a GIS-derived metric of walking distance, which gets around the key problem of self-reported minutes walking.

Although there appears to be promising evidence that transit can increase physical activity, the question that has attracted the attention of the scholars is whether the increased physical activity translates into improved health measures. Two schools of thought have emerged in this manner. The first hypothesizes that increasing transit use has beneficial health impacts if walking or biking is used for significant distances in access and egress trips. Research trying to substantiate this hypothesis is plentiful. For instance, Besser and Dannenberg (2005) conducted a descriptive analysis of the data acquired from the 2001 US National Household Travel Survey (NHTS) to determine the physical activity of Americans walking to and from public transit stops. They concluded that boosting accessibility to public transit stops increases the share of transit, and consequently improves the level of physical activity. Lachapelle and Frank (2009) employed a multinomial logit regression on a travel survey in Metropolitan Atlanta between 2001 and 2002. The results of the study demonstrated that transit users are more likely to meet physical activity recommendations, i.e. walking more than 1.5 miles (2.5 km) a day, compared to those who used a private vehicle. These studies tend to link transit use to health benefits due to increased physical activity.

The other school of thought professes that transit might not be as beneficial as hypothesized in the majority of research. Transit may substitute for cycling and walking trips, particularly in short-distance trips, leading to reduced health benefits. In a study of young people aged between 12 and 18 in London with free bus passes, Jones et al. (2012) concluded that many young people tend to substitute short distance walking trips with transit when buses are freely available and accessible. Barr et al. (2016) employed a sample of 5,241 adult residents of 42 randomly selected areas in Australia. They indicated an insignificant association between transit accessibility and obesity. In addition, transit may be self-selected by people who are already healthier or already have greater physical activity in their lifestyle, which may offset some of the reported correlations (Cervero, 2007).

Examining this question requires more than a simple regression, correlation coefficient, and reporting the statistical significance of *p*-values. The second half of this paper looks at the interpretability of the results in more detail, conducting a similar regression on height using an otherwise identical set of independent variables. There is no theory as to why transit use or accessibility is associated with height, yet the significance is similar to that of BMI. This raises doubts.

This paper does not take a prejudiced side on the two schools of thought, but it does analyze the correlation of transit access and mode share with measures of general health and BMI, while controlling for socioeconomic and demographic variables. The rest of the

paper is structured as follows. First, we review the literature examining the nexus between transit and health measures. In the review, we focus on the studies that investigate the impact of transit on health. A discussion on the data used for the analysis is then represented, followed by the framework of the study. A section is dedicated to assessing the correlation of transit mode share and accessibility by transit with both self-reported general health and BMI. Finally, we discuss remarkable findings and unpack the potential pathways for future research.

2 Previous Studies

On the health impact assessment of transportation and the built environment, researchers have sought the association between neighborhood walkability, transit use, and health outcomes. Wasfi et al. (2016) used longitudinal data from National Population Health Survey of Canada over 12 years of follow-up. They indicated moving to high-walkable neighborhood decreased BMI trajectories for men by approximately 1 kg/m^2. Moving to a low-walkable neighborhood increased BMI for men by approximately 0.45 kg/m^2. However, they did not find any associations among women. Creatore et al. (2016) also employed a time-series analysis using annual provincial health care and biennial Canadian Community Health Survey data for individuals aged 30-64 living in Southern Ontario cities. They revealed the prevalence of obesity increased in less walkable neighborhoods, but did not significantly change in areas of higher walkability between 2001 and 2012. They also demonstrated that the share of walking, biking, and public transit is higher in a highly walkable neighborhood. A recent review of the literature underlined ample evidence to support the positive association between transit use and both self-reported and objectively measured physical activity (Sener et al., 2016). However, it shows contradictory results for impacts of transit on health outcomes. We hence eschew digging into reviewing the correlation between transit use and physical activity. Rather, we emphasize studies exploring the nexus of transit and health measures. We reviewed the studies in the past two decades that analyzed the direct impact of transit on health measures. These studies are outlined in Table 1.

From the method of analysis, previous studies might be divided into two main categories, namely descriptive and statistical analyses. Gordon-Larsen et al. (2005) applies a descriptive analysis to explore the association between modes of travel, physical activity, and weight status. Using 10,771 observations extracted from The National Longitudinal Study of Adolescent to Adult Health, they revealed a statistically equal share of public transit among non-overweight and overweight US participants. To find the link between travel behavior and health outcomes, a recent study used a statistical model on a sample of 2,500 individuals living in the Chicago Metropolitan Area (Langerudi et al., 2014). The authors indicated an increase in average transit use at the level of census tract is associated with better self-reported general health, lower BMI, and lower prevalence of asthma and heart attack.

From the type of data, we categorize studies into cross-sectional, panel, or before and after data based research. Rundle et al. (2007) used a multilevel analysis to measure the association between built environment and BMI. The authors analyzed the BMI of 13,102 volunteers from the five boroughs of New York City from January 2000 through December 2002 in a cross-sectional framework. They showed the individuals living in areas with higher density of both subway and bus stops have significantly lower BMI.

Martin et al. (2015) explored the impact of switching from private car to active modes of travel and transit on BMI. Analyzing 4,056 individuals extracted from the British Household Panel Survey, they indicated this shift causes a $0.32 \, kg/m^2$ decline in self-reported BMI throughout the sample. Utilizing longitudinal data and self-reported transit use and BMI, Webb et al. (2012) showed free bus travel is negatively associated with BMI among residents of England aged 60 years and older.

TABLE 1 Summary of previous studies

Study	Place	Sample Size	Analysis Method	Data	Health Measures
Langerudi et al. (2014)	Chicago Metropolitan Area	>2,500	Binary Probit Model	Cross-sectional	• General health • BMI • Asthma • Heart attack
Webb et al. (2012)	England	>11,000	Linear regression	Longitudinal	• BMI
MacDonald et al. (2010)	Charlotte	660	Multivariate analysis	Before-After	• BMI
Samimi and Mohammadian (2010)	U.S.	>300,000	Binary Probit Model	Cross-sectional	• General health • Asthma
Samimi et al. (2009)	U.S.	>300,000	Binary Probit Model	Cross-sectional	• General health • BMI
Brown and Werner (2008)	Salt Lake City	51	Analysis of variance	Before-After	• BMI
Rundle et al. (2007)	New York City	13,102	Multilevel Analysis	Cross-sectional	• BMI
Gordon et al. (2005)	U.S.	10, 771	Multivariate analysis	Cross-sectional	• BMI
Wener et al. (2003)	New Jersey and New York	29	Analysis of covariance	Before-After	• Perceived stress
Evans et al. (2002)	New Jersey and New York	56	Linear regression	Cross-sectional	• Perceived stress

In the review of the literature we found room to grow, which motivated us to conduct this research.

- Many studies presuppose that people who use transit are more physically active than personal car users. Hence, they suggest that using transit is an escape from use of personal vehicles. These studies fail to provide strong and convincing evidence to support general health benefits associated with transit use. Although people who use transit might be more physically active than the average population, this cannot be translated into improved general health measures. Using transit might be a more stressful mode of transportation (Wener et al., 2003; Singer et al., 1974). There are also many possible confounding factors, as transit use and access are highly localized in urban areas, which are also associated with particular socioeconomic and demographic distributions, built and natural environment conditions, and are used by people with particular attitudes.

- To pinpoint the association between transit use and health measures, researchers typically apply statistical analysis on large samples. The association is then evaluated by student's t-test and p-value statistics. This approach may result in an insignificant association becoming significant due to the large sample and weakness of p-value.

3 Measures and Data Collection

This study relies on three major data sources, namely (1) share of transit use, (2) accessibility to jobs by transit, and (3) obesity and health condition characteristics in 46 of the 50 largest metropolitan areas by population. Four of the top 50 metropolitan areas lacked publicly available transit network data in 2014, and so accessibility could not be computed for these metropolitan areas. The list of metropolitan areas used in this analysis, along with the detailed accessibility values for each metropolitan area is available for interested readers (Owen and Levinson, 2014).

The first data source is derived from the American Community Survey (ACS) for one year of 2013. The ACS is an ongoing mail-based survey collected by the US Census Bureau, which targets a sample of about 250,000 addresses per month. This is an ongoing, compulsory survey, and therefore had a high response rate of 89.9% in 2013. The survey gathers information about the household socioeconomic and demographic variables, along with travel information of workers.

The Accessibility Observatory (AO) at the University of Minnesota estimated the second information in 2014. The average accessibility to jobs by walking directly and by walking to transit and taking transit (cumulative opportunities) for each census block during 7:00 am to 9:00 am period was calculated for every minute threshold (Owen et al., 2015; Owen and Levinson, 2014).

The third set of information is extracted from the Behavioral Risk Factor Surveillance System (BRFSS) health survey conducted in 2013 (US Centers for Disease Control and Prevention, 2013). The BRFSS is a database of the largest health survey in the United States. Each year, the system conducts telephone surveys, using both landline and cell phone numbers, among more than 400,000 adults from residents nationally on health-related behaviors and conditions. The response rate of the 2013 BRFSS data is 49.6% for landline and 37.8% for cell phone. The BRFSS survey typically interviewed individuals 18 and older. However, the 2013 survey had a question set asking the interviewee about the health condition of children in the house. In this study, we extracted all 102,562 people older than 18 years belonging to 46 of the 50 largest metropolitan areas by population from the BRFSS data. The reason is that the assessment of BMI in childhood depends dramatically on when the child was last measured, and the interpretation requires knowing the accurate age of measuring height and weight.

Table 2 shows the explanation of variables used in the analysis. The lowest level of geography available for individuals is the county level in the BRFSS data. Therefore, the transit share and transit accessibility measures are aggregated at the county level for this analysis. However, other socioeconomic and health information are at the individual level.

TABLE 2 Definition of explanatory variables used in the analysis

Variable	Source	Definition	Mean	Std. Dev.	Min	Max
Transit Share	ACS	Average transit use in work trips at the county level	7.06	9.29	0.41	61.05
A_5	AO	Transit accessibility to jobs by transit in 5 minutes threshold at the county level	262.06	237.10	33.26	1402.21
A_{30}	AO	Transit accessibility to jobs by transit in 30 minutes threshold at the county level	25678.54	84950.67	282.74	317884.28
General Health	BRFSS	1: If the general health is very good or excellent/ 0: Otherwise	0.53	0.49	0	1
Physical Activity	BRFSS	1: If the person has done any physical activity other than regular job in the past 30 days/ 0: Otherwise	0.77	0.41	0	1
Age	BRFSS	Age in years	53.16	18.28	18	99
Married	BRFSS	1: If a person is married/ 0: Otherwise	0.48	0.49	0	1
Children	BRFSS	Number of children in the household	0.54	1.05	0	15
Education	BRFSS	1: Education higher than college or technical school/ 0: Otherwise	0.68	0.46	0	1
Income	BRFSS	8 Categories of income from BRFSS (1 to 8)	5.86	2.19	1	8
Male	BRFSS	Gender of the respondent	0.40	0.49	0	1
Hispanic	BRFSS	1: If the person is Hispanic/ 0: Otherwise	0.09	0.29	0	1
Height	BRFSS	Height in centimeter	169.13	10.52	91	234
BMI	BRFSS	Body Mass Index (kg/m^2)	27.40	5.94	12.03	93.97
Smoker	BRFSS	1: Smoke cigarette/ 0: Otherwise	0.14	0.35	0	1
African American	BRFSS	1: If the person is African American/ 0: Otherwise	0.06	0.25	0	1
M_African American	BRFSS	= Male × African American	0.02	0.14	0	1
Asthma	BRFSS	1: If the person has asthma/ 0: Otherwise	0.13	0.33	0	1
Arthritis	BRFSS	1: If the person has some form of arthritis/ 0: Otherwise	0.30	0.46	0	1
Depression	BRFSS	1: If the person has a depressive disorder/ 0: Otherwise	0.17	0.38	0	1

Note I: ACS stands for American Community Survey 2013
Note II: AO stands for Accessibility Observatory at the University of Minnesota
Note III: BRFSS stands for Behavioral Risk Factor Surveillance System 2013

4 Framework and Hypotheses

The conceptual framework that underpins our research is a potential relationship among health and transit. This framework tests whether and to what extent increasing transit share is associated with self-reported general health and BMI. Pursuant to previous studies in the health discipline, heredity, nutrition, and physical activity play a pivotal role in general health and controlling BMI (Bryan et al., 2011). It is also well known that regular physical activity has health benefits, and walking is the most common form of physical activity (Frank et al., 2004; Hu et al., 1999; Oja et al., 1998). Hence, embedding walking in routine activities has been recommended and encompassed in intervention programs.

Following findings of positive correlation between walking and physical activity, many are shifting to investigate how environmental factors such as urban form and travel behavior are associated with physical activity, and thereby general health. In light of the recent evidence on the physical activity promotion derived from taking transit, health scientists contend that promoting transit has a positive impact on general health measures (Morency et al., 2011; Edwards, 2008). Little is known, however, whether that derived conclusion is generalizable or even reliable. The current framework attempts to shed light on the correlation of transit and accessibility by transit with both self-reported general health and BMI. We consider a number of hypotheses as the backbone of the analysis.

Hypothesis 1: Taking transit has two different associations with general health:

- From the positive point of view, two fundamental assumptions are raised: (1) transit users who choose active modes of travel for the access and egress, accomplish a portion of recommended physical activity in a day, and (2) private car users, who shift to transit, will have more travel related activity on average. Given physical activity has a positive correlation with health, transit use in this way is positively associated with general health.

- From the negative view point, two fundamental questions are raised: (1) does public transit mode include a significant amount of physical activity when park and ride or kiss and ride modes of travel is used for the access and egress of a trip? and (2) what if a public transit system discourages people from walking and biking? Indeed, the main incentive behind promoting both transit and transit-oriented development is to encourage people to shift from driving private vehicles to taking transit. It is plausible, however, that the high level of accessibility in a neighborhood induces residents to take public transit not only for long-distance trips, but also for short-distance trips. Consequently, the level of physical activity related to walking mode of travel may diminish in that area.

Hypothesis 2: Boosting accessibility by transit has two different associations with general health in different time thresholds.

- The authors posit that if a transit system performs impressively in providing access to nearby destinations, active transportation is superseded by transit. Accessibility changes behavior. Hence, frequent use of public transit for short-distance trips might be epidemic and consequently spreads laziness.

- For long-distance trips, in contrast, we hypothesize that a high level of accessibility may encourage people to shift from private vehicle to public transit, and thereby boost physical activity.

Hypothesis 3: Adding more data to the sample diminishes the size of the *p*-value significantly. As a result, a question is raised whether a particular variable is inherently significant or adding more observations causes this reduction. Transportation researchers in the US have commonly borrowed three datasets from the field of health, namely BRFSS, NHIS, and NHANES. These datasets are gathered randomly at the individual level and contain more than ten thousand observations. We postulate that many reported significant correlations between transportation variables and health outcomes are the results of large samples in the analyses.

5 Positive Health Impact of Public Transit: Myth or Reality?
To give the reader a sense of some misleading possibilities, we developed three different models in this section under the umbrella of three tales.

5.1 Story 1: Exposure to transit use and transit access at the county level is associated with lower BMI.

To understand the association of transit use and transit access with BMI, we developed a multiple linear regression model. The results are shown in Table 3. As noted previously, heredity, nutrition, and physical activity are the accepted factors correlating with BMI. Our model includes socioeconomic, demographic, and physical activity parameters to the extent of data availability. This approach helps to control the effects of other influential variables and to reduce misspecification in the model. The base model includes gender, age, ethnicity, family income, education, number of children, physical activity, asthma, arthritis, depression, and smoking. To test the existence of multicollinearity, we used variance inflation factor (VIF) for each independent variable. The VIF has a lower bound of 1, and the larger the value of VIF, the more collinearity. As a rule of thumb, a variable is highly collinear when VIF exceeds 5. Given transit use and accessibility by transit are highly correlated (Owen and Levinson, 2015), to avoid the multicollinearity issue, the influence of each of them is tested separately. After this control, the VIF ranges between 1.04 and 1.57 with the mean of 1.24, indicating there are no limitations related to multicollinearity.

Other possible transformation technique and modeling structures were investigated, with Table 3 presenting the best model. The best model was determined based on the model selection criteria that includes goodness-of-fit and significance measures, along with rationality of the estimates. From the general fit of models side, it is worth mentioning that the low magnitude of the Adjusted R^2 stems from the large sample, and the fact that BMI is ambiguously described by the variables at hand. These amounts, however, are in the range of previous studies with this scope (Chen et al., 2010; Chou et al., 2004). Nevertheless, we emphasize that the main aim of the current research is not to represent the best model encompassing all significant variables or every conceivable specification.

As per Table 3, transit mode share has a negative correlation with BMI. It means the positive side of our first hypothesis is stronger than the negative side in terms of BMI. To investigate the effect of accessibility by transit to jobs, as an index of accessibility to valued destinations, different accessibility time thresholds are included separately in the model. The results of accessibility to jobs by transit at the five-minute threshold reveal that increasing the amount of accessibility is positively associated with BMI. It might be rooted in the hypothesis that in this time threshold, walking is replaced by public transit. Interestingly, the accessibility to jobs by transit at the 30-minute threshold has a negative correlation with BMI. These findings are in line with our second hypothesis in terms of BMI. A simplistic interpretation might imply boosting accessibility to jobs by transit for long-distance trips encourages people to shift from private cars to transit.

5.2 Story 2: Exposure to transit use and transit access at the county level is associated with self-reported general health.

To understand the association of transit use and transit access with general health, we developed a multiple binary logit model. The results are shown in Table 4. To control the socioeconomic, demographic, and physical activity factors, a base model is developed at the first stage. The transit use and transit access variables are then added to the base model. As shown in Table 4, transit share has a statistically insignificant correlation with

general health, which is in line with recent research (Vemulapati, 2013). Hence, we are not able to reject the null hypothesis that transit use at the county level is not associated with general health.

The results of transit access at the five-minute threshold show increasing the amount of accessibility is positively correlated with general health. We also found that the accessibility to jobs by transit at the 30-minute threshold has a negative correlation with general health. It is in line with our second hypothesis.

5.3 Story 3: Exposure to transit use and transit access at the county level is associated with shorter stature.

In our third tale, we test whether people who use public transit are taller or shorter than other people. Much literature is associated with weight (or BMI). The natural analog to individual weight is height, so we use that as a control of the interpretability of the regression. We have no particular physical model of why height should be associated with built environment, though we can develop speculative "just-so" stories about differences in ethno-demographic or socioeconomic conditions of the city (where access and transit share is higher) versus suburbs, small towns, and rural areas. The results of the regression models imply that transit use and transit accessibility to jobs are negatively correlated with height. The results of the multivariate linear regression analysis are outlined in Table 5. We do not believe this is causal. We could further engage in data-mining and test other seemingly unrelated phenomenon, and then cherry pick results. We prefer not to do that.

One may state BMI and height are highly correlated because height is a component of BMI, and criticize this tale as a poor conceptual model. However, this is not a valid concern as there is a weak correlation between BMI and height. Mandel et al. (2004) studied the relationship between BMI and height of 35,951 patients aged 20-22 in Israel. They found a weak correlation between BMI and height that is -0.05 and 0.015 among females and males, respectively. The Diverse Population Collaborative Group (2005) examined the correlation between BMI, weight, and height in 25 diverse population samples, which include 385,232 adults aged 25 years and older from the U.S., Europe, and Asia. They showed the correlation between BMI and weight varies between 0.861 and 0.919. However, the correlation between BMI and height fluctuates between -0.016 and -0.205. We calculated the correlation between BMI, weight, and height in the data used in this study. In line with previous studies, we found a strong correlation between BMI and weight, but a weak association between BMI and height. The correlation between BMI and weight is 0.849, while the correlation between BMI and height is -0.014.

TABLE 3 Regression analysis for dependent variable: Body Mass Index (kg/m^2)

Variable	Base (BMI-0) Coefficient	95 Percent CI		Model BMI-1 Coefficient	95 Percent CI		Model BMI-2 Coefficient	95 Percent CI		Model BMI-4 Coefficient	95 Percent CI	
Constant	28.356 (240.02)	28.12	28.58	28.46 (238.77)	28.23	28.69	28.344 (236.63)	28.10	28.57	28.369 (238.05)	28.13	28.60
Age	-0.005 (-3.73)	-0.007	-0.02	-0.005 (-3.89)	-0.008	-0.002	-0.005 (-3.72)	-0.007	-0.002	-0.005 (-3.75)	-0.007	-0.002
Male	1.151 (27.13)	1.06	1.23	1.149 (27.11)	1.06	1.23	1.151 (27.14)	1.06	1.23	1.150 (27.12)	1.06	1.23
Hispanic	0.452 (6.13)	0.30	0.59	0.469 (6.35)	0.32	0.61	0.454 (6.15)	0.30	0.59	0.452 (6.13)	0.30	0.59
Children	0.182 (8.48)	0.14	0.22	0.174 (8.12)	0.13	0.21	0.182 (8.49)	0.14	0.22	0.181 (8.44)	0.13	0.22
Income	-0.063 (-6.03)	-0.08	-0.04	-0.063 (-5.97)	-0.08	-0.04	-0.063 (-6.00)	-0.08	-0.04	-0.064 (-6.05)	-0.08	-0.04
Education	-0.605 (-12.50)	-0.70	-0.51	-0.595 (-12.28)	-0.69	-0.50	-0.606 (-12.51)	-0.70	-0.51	-0.604 (-12.47)	-0.69	-0.50
Smoker	-1.150 (-19.89)	-1.26	-1.03	-1.155 (-19.98)	-1.26	-1.04	-1.149 (-19.88)	-1.26	-1.03	-1.150 (-19.89)	-1.26	-1.03
Physical Activity	-1.654 (-31.93)	-1.75	-1.55	-1.656 (-31.98)	-1.75	-1.55	-1.654 (-31.93)	-1.75	-1.55	-1.654 (-31.93)	-1.75	-1.55
African American	2.896 (29.53)	2.70	3.08	2.906 (29.64)	2.71	3.09	2.900 (29.49)	2.70	3.09	2.892 (29.44)	2.69	3.08
M_African American	-2.149 (-12.98)	-2.47	-1.82	-2.152 (-13.00)	-2.47	-1.82	-2.149 (-12.98)	-2.47	-1.82	-2.149 (-12.97)	-2.47	-1.82
Asthma	1.311 (21.83)	1.19	1.42	1.312 (21.86)	1.19	1.43	1.311 (21.83)	1.19	1.42	1.311 (21.84)	1.19	1.42
Arthritis	1.872 (38.56)	1.77	1.96	1.865 (38.41)	1.77	1.96	1.872 (38.56)	1.77	1.96	1.872 (38.55)	1.77	1.96
Depression	1.174 (21.66)	1.06	1.28	1.171 (21.60)	1.06	1.27	1.174 (21.67)	1.06	1.28	1.173 (21.64)	1.06	1.28
Transit Share	-	-	-	-0.014 (-6.72)	-0.019	-0.01	-		-	-	-	-
A_5	-	-	-	-	-	-	1.33×10^{-4} (2.05)	5.87×10^{-6}	2.61×10^{-4}	-	-	-
A_{30}	-	-	-	-	-	-	-		-	-7.88×10^{-7} (-2.59)	-1.38×10^{-6}	-1.91×10^{-7}
No. Observations:	82,250			82,250			82,250			82,250		
Adj. R^2:	0.078			0.079			0.078			0.078		

Note: Student's t-test is provided in the parenthesis with *italic* style.

TABLE 4 Binary logit analysis for dependent variable: General Health (1: Excellent and very good health/ 0: Otherwise)

Variable	Base (GH-0) Coefficient	95 Percent CI		Model GH-1 Coefficient	95 Percent CI		Model GH-2 Coefficient	95 Percent CI		Model GH-4 Coefficient	95 Percent CI	
Constant	-0.558 (-12.46)	-0.64	-0.47	-0.558 (-12.34)	-0.65	-0.46	-0.574 (-12.65)	-0.66	-0.48	-0.546 (-12.08)	-0.63	-0.45
Age	-0.012 (-23.96)	-0.012	-0.011	-0.012 (-23.95)	-0.013	-0.011	-0.012 (-23.93)	-0.013	-0.011	-0.012 (-23.99)	-0.013	-0.011
Male	-0.286 (-17.47)	-0.31	-0.25	-0.286 (-17.47)	-0.31	-0.25	-0.286 (-17.47)	-0.31	-0.25	-0.287 (-17.50)	-0.31	-0.25
Hispanic	-0.647 (-23.08)	-0.70	-0.59	-0.647 (-23.07)	-0.70	-0.59	-0.644 (-22.93)	-0.69	-0.58	-0.647 (-23.08)	-0.70	-0.59
Children	-0.018 (-2.18)	-0.034	-0.001	-0.018 (-2.18)	-0.034	-0.001	-0.017 (-2.13)	-0.034	-0.001	-0.018 (-2.25)	-0.035	-0.002
Income	0.200 (49.76)	0.19	0.20	0.200 (49.76)	0.19	0.20	0.200 (49.80)	0.19	0.20	0.199 (49.67)	0.19	0.20
Education	0.407 (22.36)	0.37	0.44	0.407 (22.35)	0.37	0.44	0.406 (22.31)	0.37	0.44	0.408 (22.42)	0.37	0.44
Smoker	-0.373 (-16.75)	-0.41	-0.32	-0.373 (-16.75)	-0.41	-0.32	-0.372 (-16.73)	-0.41	-0.32	-0.373 (-16.76)	-0.41	-0.32
Physical Activity	0.804 (40.50)	0.76	0.84	0.804 (40.50)	0.76	0.84	0.804 (40.50)	0.76	0.84	0.803 (40.49)	0.76	0.84
African American	-0.621 (-16.41)	-0.69	-0.54	-0.621 (-16.41)	-0.69	-0.54	-0.615 (-16.20)	-0.68	-0.54	-0.625 (-16.50)	-0.69	-0.55
M_African American	0.297 (4.67)	0.17	0.42	0.297 (4.67)	0.17	0.42	0.297 (4.67)	0.17	0.42	0.298 (4.67)	0.17	0.42
Asthma	-0.478 (-20.53)	-0.52	-0.43	-0.478 (-20.53)	-0.52	-0.43	-0.478 (-20.53)	-0.52	-0.43	-0.478 (-20.52)	-0.52	-0.43
Arthritis	-0.752 (-41.16)	-0.78	-0.71	-0.752 (-41.15)	-0.78	-0.71	-0.752 (-41.17)	-0.78	-0.71	-0.753 (-41.18)		
Depression	-0.643 (-30.71)	-0.68	-0.60	-0.643 (-30.71)	-0.68	-0.60	-0.642 (-30.70)	-0.68	-0.60	-0.644 (-30.75)	-0.68	-0.60
Transit Share	-	-		0.001 (1.21)	-0.0007	0.0025	-	-		-	-	
A_5	-	-		-	-		5.28×10^{-5} (2.18)	5.42×10^{-6} 1.00×10^{-4}		-	-	
A_{30}	-	-		-	-		-	-		-2.56×10^{-7} (-2.07)	-4.99×10^{-7} -1.36×10^{-8}	
No. Observations:	85,396			85,396			85,396			85,396		
Adj. R^2:	0.156			0.156			0.156			0.156		

Note: Student's t-test is provided in the parenthesis with *italic* style

TABLE 5 Regression analysis for dependent variable: Height (cm)

Variable	Base (H-0) *Coefficient*	95 Percent CI		Model H-1 *Coefficient*	95 Percent CI		Model H-1 *Coefficient*	95 Percent CI		Model H-2 *Coefficient*	95 Percent CI	
Constant	163.682 (1106.64)	163.39	163.97	163.740 (1096.65)	163.44	164.03	163.744 (1091.78)	163.45	164.03	163.732 (1097.32)	163.43	164.02
Age	-0.060 (-34.17)	-0.063	-0.056	-0.060 (-34.23)	-0.063	-0.056	-0.060 (-34.20)	-0.063	-0.056	-0.060 (-34.21)	-0.063	-0.056
Male	14.426 (269.47)	14.32	14.53	14.426 (269.47)	14.32	14.53	14.426 (269.47)	14.32	14.53	14.425 (269.43)	14.32	14.53
Hispanic	-4.331 (-46.59)	-4.51	-4.14	-4.322 (-46.47)	-4.50	-4.14	-4.342 (-46.65)	-4.52	-4.16	-4.330 (-46.58)	-4.51	-4.14
Children	-0.013 (-0.50)	-0.06	0.03	-0.017 (-0.64)	-0.07	0.03	-0.015 (-0.56)	-0.06	0.03	-0.016 (-0.60)	-0.06	0.03
Income	0.329 (24.77)	0.30	0.35	0.329 (24.79)	0.30	0.35	0.327 (24.63)	0.30	0.35	0.327 (24.67)	0.30	0.35
Education	1.179 (19.33)	1.05	1.29	1.184 (19.41)	1.06	1.30	1.182 (19.38)	1.06	1.30	1.183 (19.40)	1.06	1.30
Smoker	0.346 (4.75)	0.20	0.48	0.344 (4.72)	0.20	0.48	0.345 (4.73)	0.20	0.48	0.345 (4.74)	0.20	0.48
Physical Activity	0.444 (6.85)	0.31	0.57	0.443 (6.83)	0.31	0.57	0.445 (6.87)	0.31	0.57	0.443 (6.84)	0.31	0.57
African American	0.612 (5.02)	0.37	0.85	0.618 (5.07)	0.37	0.85	0.589 (4.82)	0.34	0.82	0.595 (4.88)	0.35	0.83
M_African American	-0.198 (-0.95)	-0.60	0.21	-0.200 (-0.96)	-0.61	0.20	-0.198 (-0.95)	-0.60	0.21	-0.197 (-0.94)	-0.60	0.21
Asthma	-0.129 (-1.72)	-0.27	0.01	-0.128 (-1.71)	-0.27	0.01	-0.129 (-1.72)	-0.27	0.01	-0.128 (-1.70)	-0.27	0.01
Arthritis	0.167 (2.73)	0.04	0.28	0.163 (2.66)	0.04	0.28	0.167 (2.73)	0.04	0.28	0.165 (2.71)	0.04	0.28
Depression	0.325 (4.77)	0.19	0.45	0.323 (4.74)	0.18	0.45	0.324 (4.76)	0.19	0.45	0.321 (4.72)	0.18	0.45
Transit Share	-	-		-0.008 (-2.82)	-0.013	-0.002	-	-		-	-	
A_5	-	-		-	-		-1.96×10^{-4} (-2.48)	-3.51×10^{-4}	-4.14×10^{-5}	-	-	
A_{30}	-	-		-	-		-	-		-1.02×10^{-6} (-2.53)	-1.82×10^{-6}	-2.32×10^{-7}
No. Observations:	84,921			84,921			84,921			84,921		
Adj. R^2:	0.512			0.512			0.512			0.512		

Note: Student's t-test is provided in the parenthesis with *italic* style.

6 Nondiscriminatory Discussion

The student's t-test is widely conducted to show a statistically significant effect of explanatory variables in statistical and econometrics modeling. This test compares the coefficient of an explanatory variable with the null hypothesis that is commonly zero. To understand the probability that the coefficient is exactly zero over the entire sample, the p-value is used generally. The p-value fluctuates between zero and one. It is generally interpreted that the smaller the value, the more statistically significant the coefficient. The p-value has an inverse relationship with the sample size, particularly in the ordinary least square method. By increasing the sample size, the size of p-value significantly approaches zero (McCloskey and Ziliak, 1996).

It is generally known that the estimated coefficient is not zero with a sufficiently high confidence interval. Statisticians, in recent years, have frequently criticized this school of thought. Anderson et al. (2000) criticized the practicality of p-value in null hypothesis testing by investigating 347 sampled articles in ecology. They concluded that the p-value test is not a fundamental aspect of the scientific method. Studies (Burnham and Anderson, 2014; Hubbard and Lindsay, 2008; Royall, 1997) concluded that the p-value test is uninformative, if no estimates of effect sizes and their precision are mentioned in analyses.

Studies discuss whether researchers selectively choose data or statistical analysis to change results from insignificant to significant (Simmons et al. 2011; Sterne et al., 2000). This misrepresentation of the results is referred to as the "P-hacking" phenomenon.

While this paper does not repudiate any particular paper or researcher, we believe with our simple demonstration that the p-hacking phenomenon is alive and well in the realm of health and transportation, with the aim of supporting particular conclusions. The following section is dedicated to give both researchers and practitioners a real insight into the causes of the present crisis by conducting an in-depth analysis on the effect of a large sample size, and testing our third hypothesis.

6.1 Too large to fail

Many studies in transportation and health literature have been conducted on large data sets, and in an era of "big data", that trend will only increase. Some samples used encompassed hundreds of thousands of observations. Samimi and Mohammadian (2010), for instance, extracted over 300,000 observations from BRFSS data to investigate the effect of built environment on self-reported general health. Following the pervasiveness of large sample studies, some scholars are scrutinizing the consequences of employing large samples (Lin et al., 2013). A key issue that raises the concern of researchers is known as "p-value problem." The issue stems from the weakness of p-value for interpretability when the power of large samples causes even minuscule effects to become statistically significant. Prolific literature has grown in both challenging and defending the use of p-value (Murtaugh, 2014; Burnham and Anderson, 2014). We do not have any intention to sympathize with the two main currents of thought. This part of the study, however, is an attempt to explore whether the impact of transit use on BMI is statistically significant or is the result of the large sample. To achieve this, two methods of analysis, namely coefficient-p-value-sample-size (CPS) and effect size are employed that shed light on whether the transit use is practically significant. The former displays the fluctuation of both p-value and coefficient of interest when the sample size ranges

from small to large. The latter method simply compares the marginal effect of the independent variables, pursuant to the notion of elasticity (Lin et al., 2013).

Lin et al. (2013) introduced a six-step algorithm to generate the CPS chart, including 5,000 random drawing samples of increasing sizes, rerunning the regression model on each sample, computing the coefficient and *p*-value of independent variable, and plotting them on a chart. The result of the CPS chart is shown in Figures 1 and 2 that highlight the *p*-value problem. As per Figure 1, the transit use variable is not significant for the small sample sizes. By increasing the sample size the *p*-value approaches zero. In our particular example, the *p*-value for share of transit drops suddenly when the sample size is greater than 12,000. Figure 2 also shows the power of a large sample in changing an insignificant variable to a significant one. As per Figure 2, the *p*-value for variable of accessibility by transit at the 30-minute threshold is not significant when the sample is less than 20,000. Increasing the sample size inflates the standard metric of statistical significance of the A_{30}. This is a direct consequence of large sample and *p*-value issue, which may mislead researchers and policymakers. In this case, the increased power of the large sample triggers the smaller or more complex effect of transit use on BMI, and deceives policymakers and researchers about practical significance. A point worthy of strong emphasis is that there is not a particular sample size threshold, which raises the *p*-value issue. To avoid the power of large sample size we recommend that researchers present the CPS chart for their variable of interest.
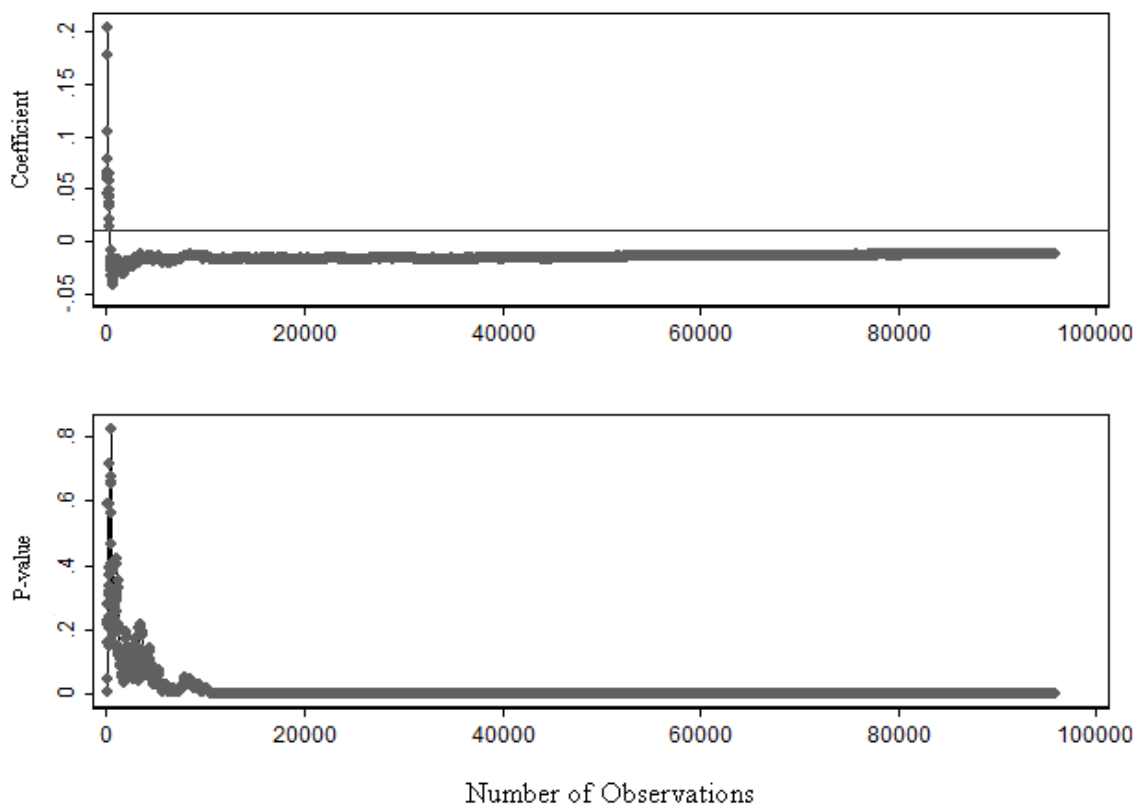

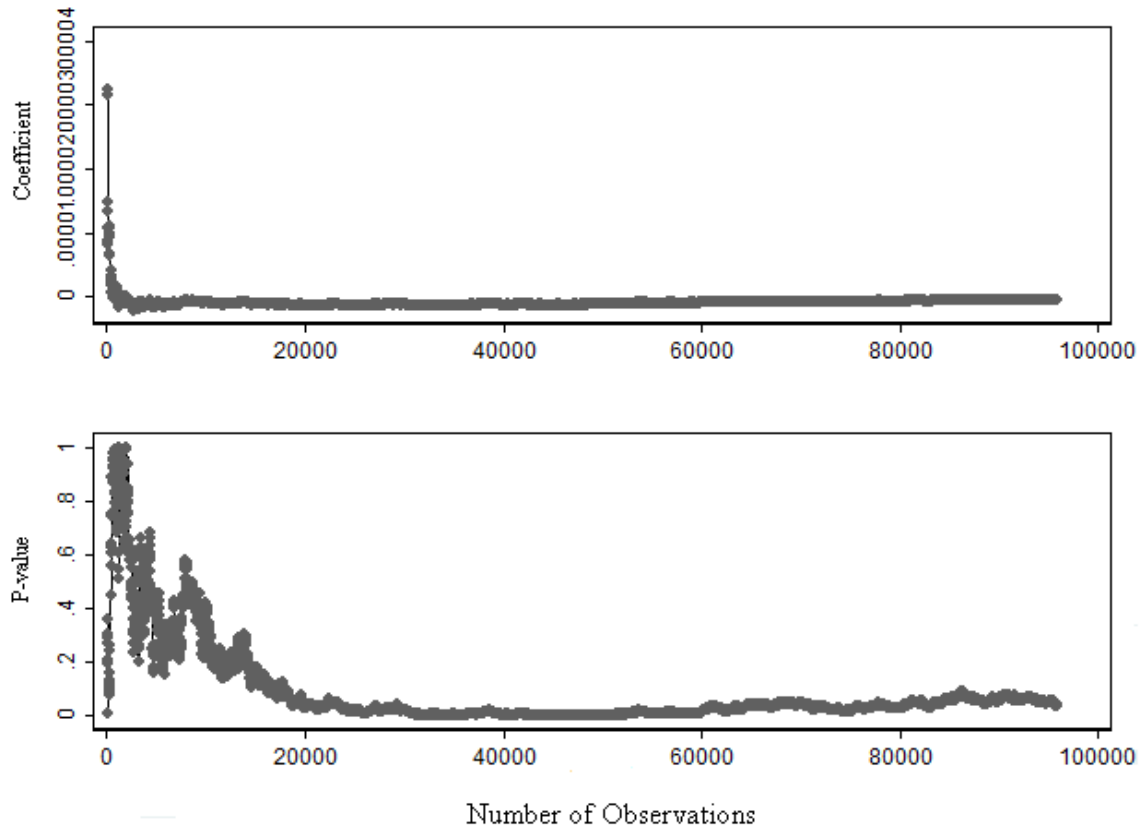**FIGURE 1** Coefficient-*p*-value-sample-size chart for Transit Share variable

FIGURE 2 Coefficient-*p*-value-sample-size chart for transit accessibility in 30 minutes

To understand the effect size of transit use variable, the elasticity of the share of transit use and amount of access to destinations by transit is measured. The final results are outlined in Table 6. Elasticity in the linear regression model represents the percent change in the dependent variable, when the independent variable is changed by 1% as equation 1 (Gujarati, 2012). In this equation, ϵ_{x_i} is the elasticity of interest variable x for individual i, Y stands for the dependent variable, and β_i is the coefficient of the interest variable. We calculated the elasticity for each individual and averaged over the sample. In the binary logit model elasticity represents the change in the probability of dependent variable as a function of increasing variable of interest by 1%, while all of the other variables are constant as equation 2 (Greene, 2003). In this equation, P_i and U_i represent the probability and utility function of the alternative for individual i, respectively. We calculated the average elasticity over the sample.

$$\epsilon_{x_i} = \frac{dY}{dX_i} \times \frac{X_i}{Y} = \beta_i \times \frac{X_i}{Y} \tag{1}$$

$$\epsilon_{x_i} = \frac{\partial P_i}{\partial X_i} \frac{X_i}{P_i} = \frac{\partial U_i}{\partial X_i} P_i (1 - P_i) \frac{X_i}{P_i} = \frac{\partial U_i}{\partial X_i} X_i (1 - P_i) \tag{2}$$

As we expected, the public transit share and accessibility by transit at the county level have a practical significant effect on neither self-reported BMI nor self-reported general health. At a deeper level, a 1% increase in the share of transit at the county level diminishes the BMI by only 0.0037%, on average. The elasticity of transit share indicates every 1% increase in transit use would escalate the chance of having excellent or very good general health by 0.0003%, while it is not significant. Samimi et al. (2009) using the BRFSS data also mentioned that a 1% increase in public transit use increases the chance of having excellent, very good, or good general health by 0.002%.

TABLE 6 Results of the elasticity analysis

Variable	Elasticity		
	BMI	*General Health*	*Height*
Transit Share	**-0.0037**	0.00003	**-0.0003**
A_5	**0.0003**	**0.0059**	**-0.0002**
A_{30}	**-0.0003**	**-0.0041**	**-0.0002**

Note: Bold numbers are statistically significant at the 95 percent confidence interval.

7 Closing Remarks

Over the past two decades, transport researchers have engaged with health issues, including obesity, cardiovascular diseases, and asthma, to name but a few. The thought of incorporating physical activity in daily routines by using active travel modes conjoins the transportation and health realms to some extent. A number of studies have identified a positive correlation between transit use and physical activity in light of the fact that for most transit users the mode of access and egress is either walking or biking.

The current study conducts an in-depth analysis to alert policymakers and practitioners to the possibilities of erroneous results in the positive impacts of transit use on general health. We investigated the association of transit mode share and accessibility by transit with self-reported general health, BMI, and height, while controlling for socioeconomic, demographic, and physical activity factors. We developed three different models under the umbrella of three tales to test our first two hypotheses. We also depicted the coefficient-*p*-value-sample-size chart and conducted the effect size analysis to test our third hypothesis. Key findings include:

- Transit use at the county level is negatively associated with self-reported BMI. This association is statistically significant, but it is practically insignificant. The power of the large sample in our particular case caused the statistically insignificant variable to become significant.

- Transit access at the county level has two different associations with self-reported BMI in different time thresholds. Accessibility by transit at the 5-minute and 30-minute thresholds has a positive and negative correlation with BMI, respectively. Both associations are statistically significant, but they are practically insignificant.

- Transit use at the county level is positively correlated with self-reported general health. This correlation is neither statistically nor practically significant.

Therefore, we failed to reject the null hypothesis that taking transit is associated with general health.

- Transit access at the county level has two different associations with self-reported general health in different time thresholds. Accessibility by transit at the 5-minute and 30-minute thresholds has a positive and negative association with self-reported general health, respectively. Both associations are statistically significant, but they are practically insignificant. We are able to reject the null hypothesis that boosting accessibility by transit is negatively and positively associated with general health in different time thresholds at the 95% confidence interval.

- We showed the power of large sample causes an insignificant variable to become significant. This validates our third hypothesis.

- We highlighted the importance of data and variable selection by portraying a statistical significant correlation of transit use and transit access with height in a multivariate regression analysis.

What becomes clear from this study is that there is a thin line between false positive and true negative results. Although the current study provides a real insight into the limited impact of transit use on health measures, it has room to grow with further research:

- There is a profound regional variation to BMI in the United States that one might take into consideration in the modeling process. Therefore, we recommend including other regional built environment variables, such as land use, block size, road density, and intersection density in the models.

- Since the lowest level of geography available for individuals is the county level in the BRFSS data, the aggregated built environment variables were used in this research. Disaggregate information may provide more robust response behavior model. In addition, one needs to replicate this analysis using non self-report data sets such as NHANES, as BRFSS is a self-report survey.

- This study is based on a cross-sectional data set. To test all the hypotheses brought up in this research and explore the causation between transit use and health measures, a rich longitudinal data set is needed for future research. We highly recommend conducting a survey in the nexus of transit use and health, which has four distinct characteristics: (1) Data is gathered in a panel framework, (2) BMI and health measures are monitored during the study, (3) Transit self-selection is controlled, and (4) A large enough sample is gathered to draw statistical conclusions.

- Studies (Ewing and Cervero, 2010) show an increase in use of public transit follows an increase in high density development. This is hypothesized to improve physical activity, and thereby general health. There is also a well-established body

of literature (Hajna et al., 2015; Creatore et al., 2016; Wasfi et al., 2016) examining the positive association of neighborhood walkability, which includes street connectivity, land use mix and residential density, with physical activity. On the flip side of the coin, higher density development may reduce home activities such as tending a garden, walking a dog, or other activities that may increase as transit supportive densities decrease. Therefore, increasing the density is a double-edged sword concerning the amount of physical activity in this case. This hypothesis should be tested in future research to measure the real physical activity change that comes with transit use.

More broadly, transit officials seek to justify the returns of both transit-oriented developments and transit investments and services. What stakeholders ultimately tend to ask is whether costs of a particular project or program will return benefits over the short- or long-terms. For instance, proponents and advocacy groups are typically sympathetic toward finding "positive results" to justify transit investments either before or after the fact. Advocacy is abetted by researchers engaging in the destructive trend dubbed the "*p*-hacking" phenomena. Examining this more generally, Pereira and Ioannidis (2011) indicated that 16-37% of the statistically significant meta-analyses are false positive. Head et al. (2015) also note that *p*-hacking is widespread in science. The rewards for reporting positive effects gives researchers an incentive to select data or analysis method until insignificant results become significant.

This misleads policies, and, to the extent that they affect actual decision-making, have the potential to impose a substantial economic burden on the public by making wrongly justified investments. Misleading results also pull science down unproductive rabbit holes, forming the cornerstone of future studies needing to refute the unwarranted findings (spending time and energy unlearning the false rather than learning new truths). Analyses show that most researchers frequently follow the path of least resistance, and thereby sidestep the crisis of replication (or non-replicability). Editors seeking high impact factors show more tendencies to publish false positive than true negative, which ultimately results in publication bias (Nosek et al., 2012; John et al., 2012).

Once false positives appear in the literature, they become tenacious. A study (Tweedie et al., 1996) shows that 45% of an observed association is rooted in publication bias. Fanelli (2011), further, investigated the trend of reporting positive results among more than 4,600 published papers in all disciplines between 1990 and 2007. The final analysis revealed that the interest in reporting the false positive results in the field of engineering have increased linearly with the slope of 0.75. A more recent study analyzed 221 articles in the social sciences (Franco et al., 2014). They mentioned not only direct evidence of publication bias in the literature, but an intention of authors to not submit null findings. One explanation is a significant delay in publication time of studies with negative results in comparison with positive results (Stern and Simes, 1997). This alarming trend is rooted in the growing competition for research funding, publication, and citations.

From the weakness of *p*-values perspective, we strove to alert both researchers and practitioners to the dangerous pitfall deriving from the power of large samples. In a linear regression model, a *p*-value measures the distance between the parameter of interest and zero in units of standard error. Increasing the sample size shrinks the standard error

remarkably, and thereby minuscule distances become statistically significant. Researchers should specify whether the small *p*-value is just an artifact of the power of large samples. Building the results on just significance and sign of the parameter of interest is worthless, unless the magnitude of effect size is carefully quantified as a post analysis. Although assessing the large sample and small *p*-values issue is scant in literature, a few studies suggest: (1) conducting the analysis on a smaller proportion of large samples and reporting the p-value. (Gefen and Carmel, 2008), (2) emphasizing the practical significance instead of statistical significance (Mithas and Lucas, 2010), and (3) reducing the significance level threshold for large samples (Greene, 2003; Leamer 1978). Having a better understanding of the effects that cause an insignificant variable to become significant will lead to better modeling accuracy, model suitability, and conclusions in practice.

8 References

American community survey (2013), http://www.census.gov/acs, Access on May 2015.

Anderson, D. R., Burnham, K. P., and Thompson, W. L. (2000). Null hypothesis testing: problems, prevalence, and an alternative. *The journal of wildlife management*, 912-923.

Besser, L. M., and Dannenberg, A. L. (2005). Walking to public transit: steps to help meet physical activity recommendations. *American journal of preventive medicine*, *29*(4), 273-280.

Brown, B. B., and Werner, C. M. (2008). Before and after a new light rail stop: resident attitudes, travel behavior, and obesity. *Journal of the American Planning Association*, *75*(1), 5-12.

Bryan, A. D., Magnan, R. E., Nilsson, R., Marcus, B. H., Tompkins, S. A., and Hutchison, K. E. (2011). The big picture of individual differences in physical activity behavior change: A transdisciplinary approach. *Psychology of Sport and Exercise*, *12*(1), 20-26.

Burnham, K. P., and Anderson, D. R. (2014). P values are only an index to evidence: 20th-vs. 21st-century statistical science. *Ecology*, *95*(3), 627-630.

Cervero, R. (2007). Transit-oriented development's ridership bonus: a product of self-selection and public policies. Environment and planning A, 39(9), 2068-2085.

Chen, S., Florax, R. J., Snyder, S., & Miller, C. C. (2010). Obesity and access to chain grocers. *Economic geography*, *86*(4), 431-452.

Chou, S. Y., Grossman, M., and Saffer, H. (2004). An economic analysis of adult obesity: results from the Behavioral Risk Factor Surveillance System. *Journal of health economics*, *23*(3), 565-587.

Creatore, M. I., Glazier, R. H., Moineddin, R., Fazli, G. S., Johns, A., Gozdyra, P., ... & Booth, G. L. (2016). Association of Neighborhood Walkability With Change in Overweight, Obesity, and Diabetes. JAMA, 315(20), 2211-2220.

Diverse, P. C. G. (2005). Weight-height relationships and body mass index: some observations from the Diverse Populations Collaboration. American journal of physical anthropology, 128(1), 220.

Edwards, R. D. (2008). Public transit, obesity, and medical costs: assessing the magnitudes. *Preventive Medicine*, *46*(1), 14-21.

Evans, G. W., Wener, R. E., and Phillips, D. (2002). The Morning Rush Hour Predictability and Commuter Stress. *Environment and Behavior*, *34*(4), 521-530.

Ewing, R., and Cervero, R. (2010). Travel and the built environment: a meta-analysis. *Journal of the American Planning Association*, *76*(3), 265-294.

Fanelli, Daniele. "Negative results are disappearing from most disciplines and countries." *Scientometrics* 90.3 (2011): 891-904.

Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. Science, 345(6203), 1502-1505.

Frank, L. D., Andresen, M. A., and Schmid, T. L. (2004). Obesity relationships with community design, physical activity, and time spent in cars. *American journal of preventive medicine*, *27*(2), 87-96.

Gefen, D., and Carmel, E. (2008). Is the world really flat? A look at offshoring at an online programming marketplace. *MIS quarterly*, 367-384.

Greene, W. H. (2003). *Econometric analysis*. Pearson Education India.

Gordon-Larsen, P., Nelson, M. C., and Beam, K. (2005). Associations among active transportation, physical activity, and weight status in young adults. *Obesity Research*, *13*(5), 868-875.

Gujarati, D. N. (2012). *Basic econometrics*. Tata McGraw-Hill Education.

Hajna, S., Ross, N. A., Brazeau, A. S., Bélisle, P., Joseph, L., & Dasgupta, K. (2015). Associations between neighbourhood walkability and daily steps in adults: a systematic review and meta-analysis. BMC public health, 15(1), 1.

Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., and Jennions, M. D. (2015). The Extent and Consequences of P-Hacking in Science. *PLoS Biol*, *13*(3), e1002106.

Hu, F. B., Sigal, R. J., Rich-Edwards, J. W., Colditz, G. A., Solomon, C. G., Willett, W. C., ... and Manson, J. E. (1999). Walking compared with vigorous physical activity and risk of type 2 diabetes in women: a prospective study. *Jama*, *282*(15), 1433-1439.

Hubbard, R., & Lindsay, R. M. (2008). Why P values are not a useful measure of evidence in statistical significance testing. *Theory & Psychology*, *18*(1), 69-88.

John, L. K., Loewenstein, G., and Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological science*, 0956797611430953.

Jones, A., Steinbach, R., Roberts, H., Goodman, A., and Green, J. (2012). Rethinking passive transport: Bus fare exemptions and young people's wellbeing. *Health & place*, *18*(3), 605-612.

Julien, D., Richard, L., Gauvin, L., Fournier, M., Kestens, Y., Shatenstein, B., ... and Payette, H. (2015). Transit use and walking as potential mediators of the association between accessibility to services and amenities and social participation among urban-dwelling older adults: Insights from the VoisiNuAge study. *Journal of Transport & Health*, *2*(1), 35-43.

Lachapelle, U., and Frank, L. D. (2009). Transit and health: mode of transport, employer-sponsored public transit pass programs, and physical activity. *Journal of Public Health Policy*, S73-S94.

Langerudi, M. F., Abolfazl, M., & Sriraj, P. S. (2014). Health and transportation: Small scale area association. *Journal of Transport & Health*.

Leamer, E. E. (1978). *Specification searches: Ad hoc inference with nonexperimental data* (Vol. 53). John Wiley & Sons Incorporated.

Lin, M., Lucas Jr, H. C., and Shmueli, G. (2013). Research commentary-too big to fail: large samples and the p-value problem. *Information Systems Research*, *24*(4), 906-917.

MacDonald, J. M., Stokes, R. J., Cohen, D. A., Kofner, A., & Ridgeway, G. K. (2010). The effect of light rail transit on body mass index and physical activity. *American journal of preventive medicine*, *39*(2), 105-112.

Mandel, D., Zimlichman, E., Mimouni, F. B., Grotto, I., & Kreiss, Y. (2004). Height-related changes in body mass index: a reappraisal. Journal of the American College of Nutrition, 23(1), 51-54.

Martin, A., Panter, J., Suhrcke, M., & Ogilvie, D. (2015). Impact of changes in mode of travel to work on changes in body mass index: evidence from the British Household Panel Survey. Journal of epidemiology and community health, jech-2014.

McCloskey, D. N., and Ziliak, S. T. (1996). The standard error of regressions. *Journal of Economic Literature*, 97-114.

Mithas, S., and Lucas Jr, H. C. (2010). Are foreign IT workers cheaper? US visa policies and compensation of information technology professionals. *Management Science*, *56*(5), 745-765.

Morency, C., Trépanier, M., and Demers, M. (2011). Walking to transit: An unexpected source of physical activity. *Transport Policy*, *18*(6), 800-806.

Murtaugh, P. A. (2014). In defense of P values. *Ecology*, *95*(3), 611-617.

Nosek, B. A., Spies, J. R., and Motyl, M. (2012). Scientific utopia II. Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science*, *7*(6), 615-631.

Oja, P., Vuori, I., and Paronen, O. (1998). Daily walking and cycling to work: their utility as health-enhancing physical activity. *Patient education and counseling*, *33*, S87-S94.

Owen, A., & Levinson, D. M. (2015). Modeling the commute mode share of transit using continuous accessibility to jobs. *Transportation Research Part A: Policy and Practice*, *74*, 110-122.

Owen, A., Levinson, D., & Murphy, B. (2015). *Access Across America: Walking 2014* (No. CTS 15-03).

Owen, A., & Levinson, D. (2014). *Access Across America: Transit 2014 Methodology* (No. CTS 14-12).

Pereira, T. V., & Ioannidis, J. P. (2011). Statistically significant meta-analyses of clinical trials have modest credibility and inflated effects. *Journal of clinical epidemiology*, *64*(10), 1060-1069.

Rissel, C., Curac, N., Greenaway, M., and Bauman, A. (2012). Physical activity associated with public transport use—a review and modelling of potential benefits. *International journal of environmental research and public health*, *9*(7), 2454-2478.

Royall, R. (1997). *Statistical evidence: a likelihood paradigm* (Vol. 71). CRC press.

Rundle, A., Diez Roux, A. V., 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*(4s), 326-334.

Samimi, A., and Mohammadian, A. K. (2010). Health impacts of urban development and transportation systems. *Journal of Urban Planning and Development*.

Samimi, A., Mohammadian, A. K., and Madanizadeh, S. (2009). Effects of transportation and built environment on general health and obesity. *Transportation research part D: transport and environment*, *14*(1), 67-71.

Sener, I. N., Lee, R. J., & Elgart, Z. (2016). Potential health implications and health cost reductions of transit-induced physical activity. Journal of Transport & Health, 3(2), 133-140.

Simmons, J. P., Nelson, L. D., and Simonsohn, U. (2011). False-positive psychology undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological science*, 0956797611417632.

Singer, J. E., Lundberg, U., and Frankenhaeuser, M. (1974). *Stress on the train: A study of urban commuting*. Psychological Laboratories, University of Stockholm.

Sterne, J. A., Gavaghan, D., & Egger, M. (2000). Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. *Journal of clinical epidemiology*, *53*(11), 1119-1129.

Stern, J. M., & Simes, R. J. (1997). Publication bias: evidence of delayed publication in a cohort study of clinical research projects. Bmj, 315(7109), 640-645.

Tweedie, R. L., Scott, D. J., Biggerstaff, B. J., and Mengersen, K. L. (1996). Bayesian meta-analysis, with application to studies of ETS and lung cancer. *Lung Cancer*, *14*, S171-S194.

US Centers for Disease Control and Prevention (2013), Behavioral risk factor surveillance System http://www.cdc.gov/brfss/annual_data/annual_data.htm, Access on May 2015.

Vemulapati, S. V. (2013). *Exploring Linkages between Travel Behavior and Health with Person-Level Data from Smartphone Applications* (Doctoral dissertation).

Wasfi, R. A., Dasgupta, K., Orpana, H., & Ross, N. A. (2016). Neighborhood Walkability and Body Mass Index Trajectories: Longitudinal Study of Canadians. American journal of public health, 106(5), 934-940.

Wasfi, R. A., Ross, N. A., & El-Geneidy, A. M. (2013). Achieving recommended daily physical activity levels through commuting by public transportation: unpacking individual and contextual influences. *Health & place*, *23*, 18-25.

Webb, E., Netuveli, G., & Millett, C. (2012). Free bus passes, use of public transport and obesity among older people in England. Journal of Epidemiology and Community Health, 66(2), 176-180.

Wener, R. E., Evans, G. W., Phillips, D., & Nadler, N. (2003). Running for the 7: 45: The effects of public transit improvements on commuter stress. *Transportation*, *30*(2), 203-220.