Impact of regional population density on walking behavior

Tanishita, Masayoshi; van Wee, Bert

DOI
10.1080/03081060.2017.1325137

Publication date
2017

Document Version
Accepted author manuscript

Published in
Transportation Planning and Technology

Citation (APA)
Tanishita, M., & van Wee, B. (2017). Impact of regional population density on walking behavior. *Transportation Planning and Technology*, *40*(6), 661-678. https://doi.org/10.1080/03081060.2017.1325137

Accepted Author Manuscript of http://dx.doi.org/10.1080/03081060.2017.1325137 (Taylor & Francis)

Title:

Impact of Regional Population Density on Number of Steps Indoors and Outdoors

Acknowledgements:

The authors would like to express their gratitude to Mr. Atsushi Matsuno (Acecook Co., Ltd.) for his assistance with data arrangement and analysis and Mr. Reto Tanner (University of Leeds) for valuable comments. This work was supported by JSPS KAKENHI Grant Number 25420553.

Abstract: Land use can influence walking (measured by the number of steps) and so the health of people. This paper presents the result of empirical research on the impact of regional population densities (inhabitants per inhabitable area) on the number of steps (all steps, outdoors and indoors). We collected data of almost 11,000 respondents in 148 Japanese regions and estimate polynomial regression models, the total number of steps being the dependent variable and densities being the main independent variable. Regional population density significantly affects the number of steps after controlling for individual and household attributes. The estimated population density that maximizes the number of steps is around 11,000 persons/km2. Increasing densities, up to levels of around 11,000 inhabitants/km2 will increase walking and consequently the health of inhabitants. The population density elasticity of the number of steps is 0.046-0.049 in a simple log linear regression model without a peak.

Keywords: number of steps, population density, indoors and outdoors

1. Introduction

The impact of land use on travel behavior is studied extensively (see, for example, Holtzclaw et al., 2002, Ewing and Cervero, 2010 for studies with an empirical focus, or Hong et al., 2014) for a review of methodologies). So far most studies focus mainly (or only) on motorized transport (car, public transport). Slow modes (walking, cycling) are only limitedly included, and studies often combine walking and cycling. Understanding the relationship between land use and walking is not only scientifically of interest, but also very important for planners who aim to increase levels of walking, for example to improve people's health. Densities are an important land use factor that can potentially increase levels of walking; the larger the number of destinations within walking distance, the more people walk (e.g. Alfronzo, 2005, Ewing and Cervero, 2010). However, the literature so far (see section 2) has only included outdoor working. Our main contribution is that we include both indoor as well as outdoor walking. In section 2 we further discuss this topics. And the literature overlooked that higher densities normally stimulate walking because distances to destinations are closer, but also increase the level of service of public transport, and improved public transport may reduce levels of walking. In addition, in densely populated areas, taller buildings with elevators might also reduce the number of steps. Two recent studies in Asia showed a non-linear (convex) relationship between population density and walking (Tanishtia and Asada, 2013 and Eom and Cho, 2015).

This paper aims to explore the link between densities and walking, including all steps, indoors and outdoors, and in addition it aims to explore whether there is a density associated with a maximum number of steps. Using the data of the Japanese National Health and Nutrition Examination Survey, and controlling for key socio-demographic variables, we study the impact of population density at the level of cities or regions on the overall number of steps taken by people, both outdoors and indoors, in 148 Japanese regions.

Section 2 provides an overview of the literature in this area. Section 3 explains the methodology, followed by section 4, which presents our findings. Section 5 outlines and discusses the most important conclusions.

2. An overview of literature

There are many overviews and reviews of the general literature on the impact of land use of planning (see two examples in section 1). After 2000, review literatures which discuss the association between physical activity and built environment, are published (Saelens and Handy, 2008, Ding and Gebel, 2010, Panter and Jones, 2010, Van Cauwenberg et al., 2011). In this section, we discuss the literature in the area of the impact of densities on travel behavior and health related indicators. Table 1 gives an overview of this literature.

Table 1 is inserted around here.

A first conclusion is that 6 studies identified a positive relationship between population density and physical activity (Ham et al., 2005, Fisher et al., 2004, Li et al., 2005, Besser and Dannenberg, 2005, Moudon et al., 2006, Forsyth et al., 2009). 6 studies found non-significant / negative relations (Cole et al., 2006, Soltani et al., 2006, Badland et al., 2008, Wineman et al., 2014, De Bourdeaudhuij et al., 2005, Morris et al., 2008). 4 studies found results to depend on different groups of people (Hanabuti et al., 2011, Inoue et al., 2010, Sugiyama et al., 2014, Shigematsu et al., 2009). Secondly, two studies identified a non-linear relationship between density and walking (Tanishita and Asada, 2013, Eom and Cho, 2015). A third conclusion is that some studies have only a small sample size, especially some studies focusing at specific neighborhoods. Fourth, study which found a positive or non/significant negative impact of densities on walking originate from the USA, Australia and Europe, where population density generally is lower than in Asian cities.

Fifth, many studies pointed out the importance of the level of mixed use, proximity to destinations, street connectivity, aesthetics, (perceived) safety as built environment (Soltani et al., 2006, Wineman et al., 2014). However, population density sometimes correlates with these variables. Areas with well-mixed, well-connected, beautiful aesthetics and safe, attract specific groups of people (e.g. Eom and Cho, 2015).

Sixth, the explained variable in many studies is measured as a binomial variable (for example, more or less than 30 minutes walking per day), which is estimated using a logit (or logistic regression) model, even though the explained variable (walking time) in reality is a continuous variable. Finally, not any study included both outdoor and indoor steps.

We next present some more information on the individual studies. Interestingly, 2 studies in Asia showed a non-linear (convex) relationship between population density and walking variables. Tanishita and Asada (2013) paying attention to outdoor walking time rather than the number of steps, analysed the impact of population density. Controlling for individual and household attributes, a convex relationship was found between population density and outdoor walking time. The population density that maximizes outdoor walking time was around 6,200 persons/km^2. However, this study excluded indoor walking from the analysis, such as walking at home or during transfers at train stations.

Using more than 400 000 individual's travel behavior data in Seoul Metropolitan Areas, Eom an Cho (2015) showed the population density that maximizes occurrence of walking and walking time per day was 9,132-16,101 persons/km^2.

In addition, Badland et al. (2008) concluded that population density didn’t affect active transport on commuting using 4 quartile based categories of population density. However, their results show that people in the middle two density quartiles significantly more frequently walk and cycle to their work than people in the lowest quartile (but this does not apply to the highest quartile). This may also indicate non-linear relationship.

On the other hand, van Loon et al. (2014) assessed how the demarcation of the size of neighborhoods (measures in circles with a radius of 200 to 1600 meters) influences the relationship between the neighborhood built and social environment and physical activity for youth. They found the largest size (1.6 km) was the strongest and most significantly related to physical activity for youth..

From the overview, we conclude three gaps remain in the literature. First, as already explained in the introduction, most of literatures have not focused on the fact that higher densities generally increase the level of service of public transport, and improved public transport may reduce levels of walking, first, because public transport competes with walking for short trips, and second, because distances from origins or destinations to public transport stops decrease with an increase in public transport supply. Therefore, it is not yet established whether higher densities increase walking over the whole range of densities, or if walking may peak at a certain level of density. Second, previous studies have focused on district level population density, but there is little research on densities at the level of cities or regions. Third, many studies only focus on outdoor walking, ignoring indoor walking. Indoor walking may also be related to land use characteristics. For example, distances that people need to walk inside large train stations can be substantial. Outdoor activity may also correlate positively or negatively with indoor activity. Possibly, a person who prefers a certain level of activity would compensate for a lack of outdoor activity by using the stairs instead of the elevator. On the other hand, a person in poor health may prefer the elevator. If only outdoor walking is included, the overall impact of land use on the number of steps that people take is still unknown.

3. Methodology and Data

We used the data of the National Health and Nutrition Examination Survey of Japan in 2009 and 2010. The aims of this survey is to obtain data to be used for policy initiatives to improve public health in Japan. This survey is conducted in November each year. Samples are collected using stratified random sampling from the whole of Japan. Based on population size of municipalities, each 47 prefectures divided into 4 or 5 groups. Then the whole of Japan is divided into 148 regions. However, as this survey is not focused on neighborhoods and municipalities, we could not identify the neighborhoods and municipality where the respondent lives, unfortunately. Minimum and maximum of areas of these regions are 106 and 16,016 km^2, respectively.

The dependent variable in our study is the total number of steps, indoors and outdoors. Respondents were required to put on a pedometer all day (24 hours) on one weekday (Monday-Friday).

Table 2 provides an overview of the variables used in this paper. To explore the impact of public transport supply, we examined different specifications for densities and the number of train stations in an area. However, because the number of stations and the road length were strongly correlated with population density (the correlation coefficient is 0.66 and 0.90

respectively), we only use population density which represents a region level variable in this study.

Table 2 is inserted around here.

On BMI, as this may cause causality problems, we exclude explaining variable (See Appendix for BMI included model results).

In this survey, self-reported activeness is measured (active or not). The question is “Are you active in daily life?” This includes not only walking but also leisure activity such as swimming, yoga etc. and mental aspects. As this “Active” variable may also cause causality problems, we show the results with and without this variables.

We eliminated outliers that were beyond the quartiles by one-and-a-half interquartile ranges of the logarithm of the number of steps. We used 10,975 cases (persons) in our analysis.

Table 3 presents the descriptive statistics. It shows the average number of steps is almost 6,900. Because an average step is probably around 0.8 meters (though we did not study this), we conclude the respondents on average walked over 5 km per day. Assuming an average speed of 5 km/h (excluding waiting time for traffic lights, elevators, etc.) the average time spent on walking is over one hour. Females are somewhat overrepresented in the sample. Around 20% of the respondents smoke, and about 87% of the respondents have breakfast every day.

Figure 1 shows the histogram of the number of steps. From the histogram, we can conclude that the mode of the number of steps per 24 hours is around 5,000. The distribution is skewed to the right.

Figure 1 is inserted around here.

It is well known that older people do walk less, mainly for physical reasons (e.g. Storti et al., 2009, Australian Bureau of Statistics, 2013, Ministry of Health, Labor and Welfare of Japan, 2014). Figure 2 shows the scatterplot of age and the number of steps. It shows the number of steps start to decrease from around a bit over 50 years old. But it is not fully clear at which age the decrease starts. Therefore we introduced threshold value, linked to the following variable (*age1*), and we explored a threshold A (years old) which explains the number of steps best based on AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) (see http://www.modelselection.org/aic/ and http://www.modelselection.org/bic/).

$$age1 = \begin{cases} 0 & (0 - \mathrm{A\ years\ old}) \\ age - A & (A\ years\ and\ older) \end{cases}$$

Figure 2 is inserted around here.

We estimated polynomial regression models, the logarithm of total number of steps being the dependent variable and densities and other personal attributes (Table 3) being the independent variable. To explore whether there is a density associated with a maximum number of steps, we applied a quadratic (quadratic) and a linear (log) function for population density.

Table 3 is inserted around here.

Taking heteroscadasticity among 148 regions into account, we also estimated multilevel random effect model. However, AIC and BIC of random effect models are worse than normal polynomial regression model. We selected variables and interaction terms (cross effects) based on (minimizing) AIC and BIC values.

4. Results

Table 4 presents the results of 4 models (with and without "Active" variables, and 2 functional forms (quadratic and log) of population density).

Table 4 is inserted around here.

Next, we discuss the impact of independent variables on the number of steps.

Population density

The first degree term is insignificant and the quadratic term is negative and statistically significant ($p < 0.05$) (quadratic) both with and without "Active" variable. This result implies that the number of steps peaks at a certain density. The population density that maximizes the number of steps is around 10,900-11,200 persons/km^2. To illustrate this result: this value is almost same as Osaka wards (around 11,300 persons/km^2) and smaller than Tokyo 23 wards (around 13,700 persons/km^2). We did not explore why the number of steps peaks at around 11,000 persons/km^2. As suggested in the introduction, one possible explanation may be the impact of densities on public transport supply.

The model without the quadratic term (log) shows lower AIC/BIC, though the differences are quite small. Based on this log-log model, the population density elasticity with regard to the number of steps is about 0.046-0.049, suggesting that a 10% increase in population density increases the number of steps by 0.46-0.49%. At first glance, this suggests that regional densities hardly influence the number of steps. A doubling in densities only increases the number of steps by a few percent. However, it is important to realize that the differences in densities within and across cities worldwide are large; densely populated city areas feature higher population densities than, for example, low-density residential areas in cities in the USA or Australia. Therefore, the real world variety in densities can easily have an impact of over 10% on the number of steps people take.

An important question is: which model is 'best'? The model without the quadratic term (log) shows lower AIC/BIC values. This would imply that this model (slightly) outperform the first model. The conclusion then would be that the number of steps increases with population density over the full range of densities as included in our study. However, the difference in AIC/BIC values are very small (see Table 4). We therefore think that based on our study we cannot present a strong conclusion about which model is best. We conclude it is still uncertain if the total number of steps peaks at a certain density level.

Age

It turned out that in ‘with active model,’ the threshold which worked best was 51 years old, whereas in the ‘without active model,’ the threshold was 52 years old. We conclude that it is not fully clear at which age the decrease in walking starts, but it is around 51 to 52 years, and the function has a parabolic form.

Gender

Gender itself hardly influences the number of steps when “Active” variable included in the model. On the other hands, we can find cross effects between gender and smoking and “Active”. Smoker tend to walk less than non-smoker and “Active” persons tend to walk more than inactive persons, but the difference of women is smaller than men. However, when we exclude the “Active” variable, women walk less than men. We need further analysis to understand this results. Maybe a distinction between outdoors and indoors walking could shed light on this finding.

Other variables

The number of steps is higher for persons who have a job, sleep less hours, take breakfast every day, like to exercise, and are nonsmokers. Again, we stress the importance of the fact that we do not know the causalities. Household size do not affect the number of steps.

5. Conclusion

Regional population density significantly affects the number of steps after controlling for individual and household attributes. It is still uncertain if the number of steps peaks at a certain density level – the models with and without a peak perform about equally well. In the model with a peak the estimated population density that maximizes the number of steps is around 10,900-11,200 persons/km^2. In the simple log linear regression model without a peak, the estimated population density elasticity of the number of steps is 0.046-0.049.

For planners and policy makers, the message is clear: increasing densities, up to levels of around 11,000 inhabitants/km^2 will increase walking and consequently the health of inhabitants. These densities are well above those in many residential areas in countries like the USA, Canada, and Australia. Our results are probably to some extent context specific. For example, in countries with high levels of cycling, walking and cycling compete for short distance trips. The quality of the public transport system may also influence walking levels both positively and negatively. On the one hand, a higher quality will increase the share of public transport, increasing walking levels to and from public transport nodal points; on the other hand a very dense public transport network may lead people to substitute walking for travelling by public transport.

Our study has certain limitations. First, we only have region level population densities. Ideally, neighborhood and social environmental data should be collected and densities should be measured with more (local) precision. Second, we did not include weekend data and data for other months of the year. The month of the year can be important, because of seasonal effects. For example, people who live in snowy areas walk much less in winter. Third, the accuracy of the registration of the number of steps is not perfect. We have indications that some people did not wear the equipment for 24 hours. Therefore, we probably underestimated the number of steps. However, we have no reason to assume that underreporting is correlated with independent variables. We therefore expect the parameters in the models not to be sensitive to underreporting of the number of steps. Fourth, in addition to densities, the level of mixed use, proximity to destinations, street connectivity, aesthetics, and (perceived) safety etc. are also quite important variables affecting physical activity (e.g. Trumpeter and Wilson, 2014). However, population density sometimes correlates with these variables, and due to a lack of data sources at the same spatial level we were not able to include these variables. Finally, people might self-select with respect to the place of residence (see Cao et al., 2009, for an overview of empirical studies). Car lovers may prefer to live in a low density area providing lots of space for parking their cars and driving, whereas people who prefer to walk might relatively frequently choose a densely populated area with many potential destinations within walking distance and attractive walking facilities. On the one hand, one might argue that if people self-select in this way, the impact of densities on walking may be overestimated, but on the other hand, if people prefer to walk, it is important to provide them with an urban environment that encourages them to do so. Though Chatman (2009) showed residential self-selection processes do not bias estimates of the effects of the built environment very much (see also Yu and Zhu, 2013). Longitudinal research is required for further understanding the importance of self-selection effects (Merom D. et al., 2015, Kamada et al., 2015).

Acknowledgements

The authors would like to express their gratitude to Mr. Atsushi Matsuno (Acecook Co., Ltd.) for his assistance with data arrangement and analysis and Mr. Reto Tanner (University of Leeds) for valuable comments. This work was supported by JSPS KAKENHI Grant Number 25420553.

References

1. Alfonzo, Mariela A. 2005. "To Walk or Not to Walk? The Hierarchy of Walking Needs." *Environment and Behavior*, 37(6), 808-836. doi: 10.1177/0013916504274016

2. Australian Bureau of Statistics. 2013. "Pedometer Steps." Australian Health Survey: Physical Activity, 2011-12. Accessed NOV 15 2015.
http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/4364.0.55.004Chapter5002011-12

3. Badland HM, Schofield GM, Garrett N. 2008. "Travel behavior and objectively measured urban design variables: associations for adults traveling to work." *Health Place*. 14(1): 85–95.
doi:10.1016/j.healthplace.2007.05.002

4. Besser LM, Dannenberg AL. 2005. "Walking to public transit: steps to help meet physical activity recommendations." *American Journal of Preventive Medicine*, 29: 273–280.
doi: http://dx.doi.org/10.1016/j.amepre.2005.06.010

5. Chatman, D.G. 2009." Residential choice, the built environment, and nonwork ravel: Evidence using new data and methods." *Environment and Planning A*, 41(5): 1072-1089. doi: 10.1068/a4114

6. Cole R, Leslie E, Bauman A, Donald M, Owen N. 2006. "Socio-demographic variations in walking for transport and for recreation or exercise among adult Australians." *Journal of Physical Activity and Health*, 3(2): 164–178.

7. De Bourdeaudhuij I., Teixeira PJ, Cardon G, Deforche B. 2005. "Environmental and psychosocial correlates of physical activity in Portuguese and Belgian adults." *Public Health Nutrition*, 8: 886–895.
doi: 10.1079/PHN2005735

8. Ding, Ding and Klaus Gebel. 2012. "Built environment, physical activity, and obesity: What have we learned from reviewing the literature?" *Health & Place*, 18: 100-105.
doi:10.1016/j.healthplace.2011.08.021

9. Eom, Hun-Joo and Gi-Hyong Cho. 2015. "Exploring thresholds of built environment characteristics for walkable communities: Empirical evidence from the Seoul Metropolitan Area." *Transportation Research Part D*. 40: 76-86. doi:10.1016/j.trd.2015.07.005

10. Ewing, Reid, Robert Cervero. 2010. "Travel and the built environment." *Journal of the American Planning Association*. 76(3): 265-294. doi: http://dx.doi.org/10.3141/1780-10

11. Fisher, K.J., F. Li, Y. Michael, M. Cleveland. 2004. "Neighborhood-level influences on physical activity among older adults: a multilevel analysis." *Journal of Aging and Physical Activity*, 12: 45–63. doi:10.1136/jech.2004.028399

12. Forsyth, Ann, J. Michael Oakes, Brian Lee, Kathryn H. Shimitz. 2009. "The built environment, walking, and physical activity: Is the environment more important to some people than others?" *Transportation Research Part D*, 14(1): 42–49. doi:10.1016/j.trd.2008.10.003

13. Ham SA, Macera CA, Lindley C. 2005. "Trends in walking for transportation in the United States, 1995 and 2001." *Preventing Chronic Disease*. Accessed NOV 15 2015. http://www.cdc.gov/pcd/issues/2005/oct/04_0138.htm

14. Hanibuti T.,Kawachi, I., Nakaya, T., Hirai, H., Kondo, K. 2011. "Neighborhood built environment and physical activity of Japanese older adults." *BMC Public Health*, 11(1): 657. doi:10.1186/1471-2458-11-657

15. Holtzclaw, John, Robert Clear, Hank Dittmar, David Goldstein and Peter Haas. 2002. "Location efficiency: Neighborhood and socio-economic characteristics determine auto ownership and use – Studies in Chicago, Los Angeles and San Francisco." *Transportation planning and Technology*, 25: 1-27. doi: 10.1080/03081060290032033

16. Hong, J, Shen, Q., Zhang, L. 2014. "How do built-environment factors affect travel behavior? A spatial analysis at different geographic scales." *Transportation*, 4 (3): 419-440. doi: 10.1007/s11116-013-9462-9

17. Inoue S, Ohya Y, Odagiri Y, Takamiya T, Ishii K, Kitabayashi M, Suijo K, Sallis JF, Shimomitsu T. 2010. "Association between perceived neighborhood environment and walking among adults in 4 cities in Japan." *Journal of Epidemiology*, 20(4): 277-86. doi: http://doi.org/10.2188/jea.JE20090120

18. Kamada, Masamitsu, Jun Kitayuguchi, Takafumi Abe, Masataka Taguri, Shigeru Inoue, Yoshiki Ishikawa, Kazuhiro Harada, I-Min Lee, Adrian Bauman, Motohiko Miyachi. 2015. "Community-wide promotion of physical activity in middle-aged and older Japanese: a 3-year evaluation of a cluster randomized trial." *International Journal of Behavioral Nutrition and Physical Activity*, 12: 85. doi:10.1186/s12966-015-0242-0

19. Kleinbaum, D.G., Kupper, L.L., Muller, K.E., Nizam, A. 1988. *Applied regression analysis and other multivariable methods. third ed.* Duxbury Press, Pacific Grove.

20. Li F, Fisher KJ, Brownson RC, Bosworth M. 2005. "Multilevel modelling of built environment characteristics related to neighbourhood walking activity in older adults." *Journal of Epidemiological Community Health*, 59(7): 558-64. doi: 10.1136/jech.2004.028399

21. Merom, D., K. Gebel, P. Fahey, T. Astell-Burt, A. Voukelatos, C. Rissel, C. Sherrington. 2015. "Neighborhood walkability, fear and risk of falling and response to walking promotion." *Preventive*

Medicine Reports, 2: 704–710. doi:10.1016/j.pmedr.2015.08.011

22. Ministry of health, labor and welfare of Japan. 2014. "Abstracts of national health and nutrition survey results." (in Japanese). Accessed NOV 15 2015.
http://www.mhlw.go.jp/file/04-Houdouhappyou-10904750-Kenkoukyoku-Gantaisakukenkouzoushinka/0000068070.pdf

23. Morris, S. Katherine, Edward McAuley, and Robert W. Mot. 2008. "Self-efficacy and environmental correlates of physical activity among older women and women with multiple sclerosis." *Health Educational Research*, 23(4): 744–752. doi: 10.1093/her/cym067

24. Moudon, A. Vernez, Lee C, Cheadle AD, Garvin C, Johnson D, Schmid TL, Weathers RD, Lin L. 2006. "Operational definitions of walkable neighborhood: theoretical and empirical insights." *Journal of Physical Activity and Health*, 3: 99–117.

25. Nelson, C. Candace, Gregory R. Wagner, Alberto J. Caban-Martinez, Orfeu M. Buxton, Christopher T. Kenwood, Erika L. Sabbath, Dean M. Hashimoto, Karen Hopcia, Jennifer Allen, Glorian Sorensen. 2014. "Physical activity and body mass index: The contribution of age and workplace characteristics." *American Journal of Preventive Medicine*, 46(3), 1: S42-S51. doi: 10.1016/j.amepre.2013.10.035.

26. Panter, R. Jenna, Jones P. Andy. 2010. "Attitudes and the environment as determinants of active travel in adults: what do and don't we know?" *Journal of Physical Activity and Health*, 7: 551-561

Saelens, E. Brian, Handy, L. Susan. 2008. "Built environment correlates of walking: a review." *Medicine & Science in Sports & Exercise*, 40 (7 Supplement), S550–66. doi: 10.1249/MSS.0b013e31817c67a4

27. Shigematsu R, Sallis JF, Conway TL, Saelens BE, Frank LD, Cain KL, Chapman JE, King AC. 2009. "Age differences in the relation of perceived neighborhood environment to walking." *Medical Science of Sports Exercise*, 41(2): 314-21.

28. Soltani A, Primerano F, Allan A, Somenahalli S. 2006. "Design for movement: linking non-work travel and activity level to local urban design dimensions." *Urban Design International*. 11: 173–186. doi:10.1057/palgrave.udi.9000177

29. Sugiyama, T. Cerin, E., Owen, N., Oyeyami, A.L., Conway, T.L., Van Dyck, D., Sallis, J.F. 2014. "Perceived neighborhood environmental attributes associated with adults' recreation walking: IPEN Adult study in 12 countries." *Health & Place*, 28: 22-30. doi:10.1016/j.healthplace.2014.03.003

30. Sijtsma, Anna, Pieter JJ Sauer and Eva Corpeleijn. 2015. "Parental correlations of physical activity and body mass index in young children- the GECKO Drenthe cohort." *International Journal of Behavioral Nutrition and Physical Activity*, 12:132. doi:10.1186/s12966-015-0295-0

31. Storti, L. Kristi,, Vincent C. Arena, M. Michael Barmada, Clareann H. Bunke, Robert L. Hanson, Sandra L. Laston, Jeun-Liang Yeh, Joseph M. Zumada, Barbara V. Howard, Andrea M. Kriska. 2009. "Physical activity levels in American-Indian adults: The strong heart family study." *American Journal of Preventive Medicine*, 37(6): 481-487. doi: 10.1016/j.amepre.2009.07.019.

32. Tanishita, Masayoshi, Asada, Takumi. 2013. “Impact analysis of population density on outdoor walking time.” *Selected Proceedings of the 13th World Conference on Transportation Research*, Rio de Janeiro. Accessed NOV 15 2015.
http://www.wctrs-society.com/wp/wp-content/uploads/abstracts/rio/selected/772.pdf

33. Trumpeter, N. Nevelyn and Dawn K. Wilson. 2014. “Positive Action for Today’s health (PATH): Sex differences in walking and perceptions of the physical and social environment.” *Environment and Behavior*, 46(6): 745-767. doi: 10.1177/0013916513480860

34. Van Cauwenberg, Jelle, Ilse De Bourdeaudhuij, Femke De Meester, Delfien Van Dyck, Jo Salmon, Peter Clarys, Benedicte Deforche. 2010. “Relationship between the physical environment and physical activity in older adults: A systematic review.” *Health & Place*, 17: 458-469.
doi: 10.1016/j.healthplace.2010.11.010

35. van Loon, Josh, Lawrence D. Frank, Lindsay Nettlefold, Patti-Jean Naylor. 2014. “Youth physical activity and the neighbourhood environment: Examining correlates and the role of neighbourhood definition.” *Social Science & Medicine*. 104: 107–115. doi:10.1016/j.socscimed.2013.12.013

36. Wineman, D. Jean, Robert W. Marans, Amy J. Schulz, Diaan Louis van der Westhuizen, Graciela B. Mentz, Paul Max. 2014. “Designing healthy neighborhoods; Contributions of the built environment to physical activity in Detroit.” *Journal of Planning Education and Research*, 34(2): 180-189.
doi: 10.1177/0739456X14531829

37. Wood N. Simon, 2006. *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC Press, London.

38. Yu Chia-Yuan and Xuemei Zhu. 2013. “Impacts of Residential Self-Selection and Built Environments on Children’s Walking-to-School Behaviors.” *Environment and Behavior*, 47(3): 268-287.
doi: 10.1177/0013916513500959

Appendix 1. Additive model and BMI for explaining variables

1) Additive model

Just as in linear regression models, additive regression models estimate the impact of each independent variable separately. We model the logarithm of number of steps (*Y)* to depend on an additive combination of independent variables (*X*s) shown in Table 2 and 3.

$$Y = \mathrm{C} + \sum_{j=1}^{k} f_j\left(X_j\right) + \varepsilon \qquad (1)$$

C is a constant and f_j represent functions that are estimated by smoothing splines.

Additive model arrows non-linearity between explained and explaining variables using smoothing splines. This model is suitable in case we can’t assume linearity. In this paper, based on the estimated smoothing splines, we also estimated polynomial regression models (Kleinbaum et al., 1998).

Estimation result of additive model is shown Table A1. Figure A1 shows the estimated smooth functions (see Wood, 2006, for further information). Making use of this additive model result, we applied a polynomial regression model.

Table A1 is inserted around here.

Figure A1 is inserted around here.

2) BMI

There are many studies which analyze the relationship between BMI and physical activity (e.g. Nelson et al., 2006, Sijtsma et. al. 2015). However, it is unknown the causality of BMI and number of steps. On the one hand walking may decrease a person's BMI, but on the other hand, people's BMI may impact their walking behavior. Therefore, we also estimated parameters including BMI values as cubic (3 degree) function (Table A1).

The BMI value that maximizes the number of steps is estimated at 20.1. In addition, elasticity of population density changes from 0.046 to 0.044. In short, including the BMI variables has a minor decreasing impact on the parameter of the density variable.

Table1

	Author (year)	Survey year	Sample size	Area	Mean (sd), range of population density (person/km2)	Model	Explained variables	Explaining variables and results
Category 1: Positive								
1	Ham SA et al. (2005)	1995, 2001	95,360, 160,758 (persons)	USA	5 categories unknown	cross table	1+mile/day walk or less	degree of urbanization (+) education level (non linear) income (non linear)
2	Fisher KJ et al. (2004)	2001-02	582	older adults, 56 neighborhood, Portland, USA	400-4,600	multilevel structural equation	self reported 3 neighborhood walking activities	(between neighborhood) senior population density(+), safe to walk(0), neighborhood problem(0), (within neighborhood) social cohesion(+) income(-), percent White(+), parks and trails(+) education(+), exercise self-efficacy(+)
3	Li F. et al. (2005)	2001-02	557	older adults, 56 neighborhood, Portland, USA	400-4,600	multilevel structural equation	self reported 3 neighborhood walking activities	(neighborhood level) job density(+), household density(+), green open spaces for recreation(+), number of street intersections(+) (resident level) perceptions of safety for walking (+), number of nearby recreational facilities (+), open spaces for recreation(-), number of street intersections(-), access to recreational facilities(0), safe from traffic(0), number of street intersections(0)

4	Besser LM. and Dannenberg AL. (2005)	2005	3,312	transit user, USA	4 categories Min:<1,500, Max>9,700	multivariate logit	30+min./day walk or less	population density (+) income(-), education level(-), age(0), female(0), White(-), no car(+), bus(0)
5	Moudon AV. et al. (2006)	unknown	608	King County, WA, USA	4,500 200-17,000 parcels/km2	logit	walker or non walker	household parcel density (+), street-blocks (-), distances to food and daily retail facilities from home (-)
6	Forsyth, Ann et al. (2009)	unknown	716	36 neighborhoods, Twin Cities metropolitan Area, USA	3 categories -1,200-2,500-	multivariate logit*1	MET min. per week of 2 various groups / total walking (in miles) by diary / accelerometer counts per valid day	(MET walking) high density(+ white, males, those without a college degree, the less healthy, those without children in the household, the unemployed and retired, those with a car, those with a BMI under 30, and the obese), large block (+ non whites) (Total walking) high density(+ less healthy) large block(+ unemployed/retired) (accelerometer counts) high density(- White, not obese) large block (- men)
Category 2: non significant and negative								
1	Cole R. et al. (2006)	1996	3,392	New South Wales, Australia	2 categories unknown	multivariate logit	proportion of respondents reporting levels to meet	rurality(0 male, + female) age(-), education level(+), job(+ male, 0 female), non-English(-), BMI(0)

							the Public-Health Guideline	
2	Soltani A. et al. (2006)	1999	733	4 suburbs, Adelaide, Australia	3,020-5,380	multivariate logit	walk/cycle to commute or not	population density(0) street quality(0) level of public transport(0) well-connected streets(+) close proximity to jobs(+)
3	Badland HM. et al. (2008)	2005	364	North Shore city, New Zealand	4 categories 2,196 132-3,805	multivariate logit*2	walk/cycle to commute or not	residential density(0) street connectivity(+ however, inconsistent with 4 quartile based categories) mixed land use(0)
4	Wineman DJ. et al. (2014)	2008-09	506	3 neighborhoods, Detroit, USA	31(10) households	linear*3	MET minutes (transportation +leisure)	households density at the quarter-mile radius(-) multiple land uses (-), connectivity of street networks (both neighborhood and city level) (+)
5	De Bourdeaudhuij I. et al. (2005)	unknown	526	Portugal and Belgium	unknown	multivariate linear	active transport, walking in leisure time, moderate to vigorous activity in leisure time, and total	residential density(0) social support from family and/or friends(+), walking facilities in the neighborhood(+) (Recreation) social support(+), self-efficacy(+), perceived benefits(+) and barriers(-)

							physical activity of at least moderate intensity.	
6	Morris, SK. et al. (2008)	unknown	136	Greater Illinoi, Indiana and Gateway chapters, USA	173–577 (unit unknown)	correlation coefficients	physical activity score using activity monitor for 7 days	density(0) self-efficacy(+), physical restriction(-) (older women) street connectivity(+), access to walking/cycling facilities (+), satisfaction with neighborhood aesthetics(+)
Category 3: mixed								
1	Hanibuti T. et al. (2011)	2003	9,414	8 municipalities, Japan	3,530(1,830) 0-14,080	multivariate logit	30+min./day walk or less	population density (0, + sport activity) number of Intersections (- 50+ years old) land slope (+ 50+ years old) No. of dead ends (0), No. of destinations (0), parks and green spaces (0, + sport activity), schools (0)
2	Inoue S. et al. (2010)	2007-08	1,461	4 cities, Japan	unknown	logit	active or not for transportation and leisure	perceived residential density (+ transportation, - leisure, women) perceived good land use mix (+ transportation, - leisure, women)

							walking	perceived good walking/cycling facilities(+) perceived attractive aesthetics (+).
3	Sugiyama T. et al. (2014)	2002-11	13,745	17 sites, 12 countries	77(114) 18-440 (unit: score)	multivariate logit / generalized additive mixed	walking for recreation or not/ frequency (day/wk) and duration (min./wk) of walking for recreation	perceived residential density score and land use mix-access (0 odds frequency, + duration) perceived connectivity (0 odds, duration + frequency) perceived infrastructure and safety (0) aesthetics (+ odds, frequency, 0 duration)
4	Shigematsu R. et al. (2009)	unknown	1,623	King County, WA, USA	mean: 206-237 (unit unknown)	partial correlation coefficient*4	time spent for shopping (transportation) and leisure and exercise (leisure) /wk	residential density(+ transportation, leisure young, 0 leisure elder) land use mix–diversity(+), land use mix–access(+), street connectivity(+)
Category 4: Non-linear, Peak								

1	Tanishita, M. and T. Asada (2013)	2010	60,189	53 cities, Japan	2,003(2,661) 70-13,653	generalized linear (zero-inflated negative binominal)	outdoor walking time	population density(non linear; peak 6,244 persons/km2) number of stations/road length (+) age(+), female(-), shared car(+), job(-), household size (-)
2	Eom, Hun-Joo and Gi-Hyong Cho (2015)	2012	411,899	1,107 districts, Seoul Metropolitan Area, South Korea	18,615(13,496) 2-76,084	multivariate logit / linear	occurrence of walking / walking time per day	population density(non linear; peak 9,132-16.101presons/km2) land-use mixture (+) street connectivity (+) *these 3 variables are estimated separately age(+), female (-), license(-), job(-), household size(-), home occupancy (+), income (-)

Remarks: all data type is cross section

*1 adjusted for age, self reported health, measured BMI, ethnicity, education and household income..

*2 adjusted for trip, individual, household, and neighborhood correlates

*3 adjusted for neighborhood poverty (level 3); cross-level interaction of poverty and neighborhood characteristics (level 2); and age, gender, race/ ethnicity, education, household income, length of residence in the neighborhood, and labor force participation (level 1).

*4 adjusted for sex, BMI, education level, income, and driver’s license.

Table 2: Variables included in our study

Variables		Source
Dependent variable	Number of steps (outdoors and indoors)	National Health and Nutrition Examination Survey of Japan
Independent variables		
Socio-demographics	age, gender, occupation	
Health related variables	weight, height, average number of sleeping hours, breakfast habits, drinking behavior, smoking behavior, active or not*	
Regional variables (148 regions)	Population density (population per inhabitable area), number of train stations, road length	Population Census Railway Statistics Road Statistics

Table 3

Explained variable		**Minimum**	**Mean**	**Maximum**	**SD**
Number of steps		918	6,926	31,951	4,299
Explaining variables		**Minimum**	**Mean**	**Maximum**	**SD**
Populaton Density [pers./ha]		67	3,079	13,781	3,348
Number of Train Stations/Road Length [1/1,000km]		0.0	12.0	41.5	9.4
Age		20.0	54.6	94.0	16.9
BMI		13.8	22.9	47.3	3.5
Categorical variables		**Sample Size**		**Share (%)**	
Gender	Male	4,845		44.1	
	Female	6,130		55.9	
Occupation	Worker	6,504		59.3	
	Housewife	2,720		24.8	
	Students	139		1.3	
	Non worker	1,612		14.7	
Average sleeping hours (hrs)	Less than 5	733		6.7	
	5〜6	2,933		26.7	
	6〜7	4,198		38.3	
	7〜8	2,097		19.1	
	8〜9	822		7.5	
	More than 9	192		1.7	
Breakfast	Have every day	9,543		87.0	
	Not have every day	1,432		13.0	
Drinking	YES	5,567		50.7	
	NO	5,408		49.3	
Smoking	YES	2,285		20.8	
	NO	8,690		79.2	

Active *	YES	9,037		82.3	
	NO	1,938		17.7	
Year	2009	5,105		46.5	
	2010	5,870		53.5	

* Question: Are you actvie in daily life?

Table 4

Explained variable: log(Number of steps)	with Active				without Active			
functional form of PD	quadratic		log		quadratic		log	
	Coef.	t-value	Coef.	t-value	Coef.	t-value	Coef.	t-value
Intercept	8.55	244.84	8.48	229.40	8.80	280.66	8.73	258.37
PD: Population Density	2.96E-03	5.94			3.20E-03	6.32		
PD^2	-1.33E-05	-3.30			-1.47E-05	-3.59		
log(PD)			0.046	9.22			0.049	9.64
Age1	*-2.15E-04*	-0.13	*-1.39E-04*	-0.08	*6.76E-04*	0.38	*7.51E-04*	0.42
Age1^2	-6.31E-04	-10.59	-6.34E-04	-10.63	-7.03E-04	-10.89	-7.06E-04	-10.93
Gender Female=1	*-0.03*	-1.02	*-0.03*	-1.00	-0.10	-6.20	-0.10	-6.18
Occupation Housewife dummy	-0.11	-6.91	-0.11	-6.96	-0.11	-6.63	-0.11	-6.67
Occupation Student dummy	*-0.08*	-1.70	*-0.08*	-1.68	*-0.07*	-1.32	*-0.07*	-1.31
Occupation Nonworker dummy	-0.15	-7.62	-0.15	-7.60	-0.16	-7.93	-0.16	-7.90
Sleeping hours 5-6 dummy	-0.05	-2.12	-0.05	-2.16	*-0.05*	-1.87	*-0.05*	-1.92
Sleeping hours 6-7 dummy	-0.07	-2.86	-0.07	-2.87	-0.06	-2.50	-0.06	-2.51
Sleeping hours 7-8 dummy	-0.10	-3.91	-0.10	-3.90	-0.09	-3.58	-0.09	-3.58
Sleeping hours 8-9 dummy	-0.12	-4.16	-0.12	-4.11	-0.12	-3.93	-0.12	-3.88
Sleeping hours 9- dummy	-0.20	-4.09	-0.19	-4.08	-0.22	-4.50	-0.22	-4.48

Breakfast YES=1	0.04	2.59	0.05	2.61	0.07	3.86	0.07	3.88
Drinking YES=1	0.04	3.34	0.04	3.38	0.05	4.01	0.05	4.04
Smoking YES=1	-0.09	-4.82	-0.09	-4.82	-0.10	-5.59	-0.10	-5.59
(cross effect) Gender：Smoking	*0.06*	1.95	0.06	1.96	0.07	2.10	0.07	2.11
Active: YES=1	0.34	15.69	0.34	15.69				
(cross effect) Gender：Active	-0.08	-2.72	-0.08	-2.73				
Year 2009 dummy	0.03	2.86	0.03	2.72	*0.01*	1.33	*0.01*	1.19
A (years old)	51				52			
Number of sample	10,975							
adjusted R^2	0.18		0.18		0.15		0.15	
AIC	19,127		19,123		19,531		19,525	
AIC(0)	21,239							
BIC	19,288		19,276		19,677		19,664	
BIC(0)	21,254							
PD peak (pers./km2)	11,158				10,887			

italic: shows that the value is 0 at 5% significance.level.

Table　A1

Explained variable: log(Number of steps)	BMI		Additive	
	Coef.	t-value	Coef.	t-value
Intercept	6.59	11.80	8.52	251.50
log(PD)	0.044	8.82		
Age1	3.00.E-03	1.79		
Age1^2	-6.55.E-04	-12.41		
BMI	2.37.E-01	3.57		
BMI^2	-9.00.E-03	-3.47		
BMI^3	1.01.E-04	3.09		
Gender Female=1	-0.04	-1.38	-0.04	-1.42
Sleeping hours 5-6 dummy	-0.06	-2.47	-0.06	-2.39
Sleeping hours 6-7 dummy	-0.08	-3.51	-0.08	-3.48
Sleeping hours 7-8 dummy	-0.11	-4.48	-0.11	-4.47
Sleeping hours 8-9 dummy	-0.14	-4.59	-0.14	-4.72
Sleeping hours 9- dummy	-0.21	-4.47	-0.22	-4.6
Occupation Housewife dummy	-0.11	-6.90	-0.11	-6.82
Occupation Student dummy	-0.10	-2.00	-0.12	-2.15
Occupation Nonworker dummy	-0.15	-7.61	-0.15	-7.79
Breakfast YES=1	0.04	2.47	0.05	2.64
Drinking YES=1	0.04	3.04	0.04	3.01
Active YES=1	0.33	15.18	0.33	15.23
(cross effct) Gender：Active	-0.08	-2.85	-0.08	-2.86
Smoking YES=1	-0.09	-4.85	-0.09	-4.82
(cross effect) Gender：Smoking	0.06	1.99	0.06	1.94
Year 2009 dummy	0.03	2.77	0.03	2.94
			edf	F
s(PD)			2.34	26.34
s(Age)			4.95	119.25
s(BMI)			3.16	22.18
Number of sample	10,975			

Adi. R^2	0.19		0.19	
AIC	**19,059**		19,064	
BIC	**19,234**		19,272	
BMI peak	20.1 , 38.9			