Effects of urban spatial Form on individuals’ footprints: empirical study based on personal GPS panel data from Rotterdam and Eindhoven area

Citation for published version (APA):
Yang, D., & Timmermans, H. (2014). Effects of urban spatial Form on individuals’ footprints: empirical study based on personal GPS panel data from Rotterdam and Eindhoven area. Procedia Environmental Sciences, 22, 169-177. https://doi.org/10.1016/j.proenv.2014.11.017

DOI:
10.1016/j.proenv.2014.11.017

Document status and date:
Published: 01/01/2014

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl

providing details and we will investigate your claim.
Effects of urban spatial form on individuals’ footprints: Empirical study based on personal GPS panel data from Rotterdam and Eindhoven area
Dujuan Yang*, Harry Timmermans

Abstract

It is crucial for urban planners to understand the effects of physical planning on individuals’ behavior. A well-designed built environment allows individuals to engage in activities at shorter travel distances, using environmentally friendly transport modes. This paper examines how individuals’ CO₂ footprints co-vary with urban form indicators. GPS technology was used to collect activity-travel data. The data for the analyses were collected between May 2012 and July 2013 covering more than 200 respondents from the Eindhoven and Rotterdam regions, the Netherlands. The data includes personal and household characteristics, detailed travel information, GPS traces and detailed car information, including brand, type and year of production. A Bayesian network was used to extract activity-travel diaries from the collected GPS traces, while a Web-based prompted recall survey instrument was used to validate the imputed data. To identify their living environment, respondents’ home coordinates were matched with map data from the municipality. A two-stage sample selection model was used to estimate the effects of physical planning on individual transport mode choice and daily CO₂ emissions.

© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).

Keywords: Carbon footprints; GPS panel data; urban spatial form

* Dujuan YANG. Tel.: +0-031-040-2474814; fax: +0-031-040-2438488.
E-mail address: d.yang@tue.nl.
1. Introduction

Since the early 1990s, the relationship between the built environment and travel behavior has been subject of intense discussion. Hundreds of papers on the topic, differing in approach, have been published. One category of study used simulation to forecast travel demand and estimate the impacts of changes in street network design on VMT (vehicle miles travelled). Another series of studies was not based on simulation but rather used aggregate data to analyze the influence of the built environment on average travel times. More recently, the focus of attention has shifted from aggregate to disaggregate data, in which travel behavior has been predominantly described in terms of trip frequency and travel distance. Only some distinguished trip purposes.

Handy et al. argued that the issue of causality has not been discussed in much detail in most previous research. She pointed out that almost all available studies have used cross-sectional designs that compare travel behavior of different people or places at one point in time. Cross-sectional research designs cannot reveal the causality of the relationship between the built environment and travel behavior. “Self-selection” might explain travel distance by car. A few researchers have made some effort addressing this issue. Results, however, are mixed as some studies evidenced self-selection, while others found that people change their travel behavior when exposed to different urban forms.

Another critical issue associated with studies on the relationship between urban form and travel concerns the measurement of urban form. The most commonly used characteristics of the built environment are density, mixed land use, network characteristics, access to jobs and shopping. Holden used town size/national settlement pattern, localization of houses within a town, built-up area, residential area and type of housing to describe different aspects of physical planning. Fan and Khattak considered building density, retail accessibility and street connectivity within the 0.25-mi buffer area around the house to measure urban form. They argued their study captures the effects of micro-scale neighborhood design characteristics on activity-travel behavior. The problem with this measurement approach is that the definition of spatial scale and the boundary conditions of the selected study area and demarcation of neighborhood tend to lead to strongly different results.

In the Netherlands, the city scale is relatively small and facilities, including shopping and service, are designed based on the neighborhood concept. People living in different neighborhoods have similar accessibility to basic grocery shopping and service facilities. Instead of analyzing effects of living environments on individual’s trip components such as VMT, transport mode choice, trip frequency or trip length, this paper takes into account the influence of people’s car type choice and focuses on the relationship between urban spatial form and individual’s daily CO2 emission by car. We will investigate individual’s daily mode choice and CO2 emissions per day per person. Trip purposes are also considered as independent variables to describe the needs underlying activities. Moreover, instead of using cross-sectional data, panel data are used, allowing us to account for dynamic activity needs.

The paper is organized as follows. Data collection and descriptive analysis will be discussed in detail in section 2, followed by a description of the model. Results are discussed in Section 3. The paper is completed with a conclusion and discussion.

2. Data collection and descriptive analysis

Our analysis is based on the contention that CO2 emissions are a function of daily activity-travel patterns. These patterns include distance miles traveled with different transportation modes. Further details about cars, such as engine type, then allow an estimate of CO2 emissions.

This contention implies that details of activity-travel are required to assess variations in CO2 emissions. In this study, GPS traces, dedicated imputation algorithms and data fusion methods were applied to infer or impute daily activity-travel schedules. Around 1000 respondents from the Eindhoven and Rotterdam area were invited to participate in this survey and carry a GPS device for three months. Respondents were provided user accounts and a specific password for system login to upload their data to the website via Internet. They were invited to upload multi-day GPS traces. Their data were processed immediately after uploading to impute daily activity-travel diaries (transportation modes, activity episodes and other facets of activity-travel patterns) using a Bayesian Belief network. Next, respondents were invited to check the data for accuracy and consistency, and provide any missing
information such as activity type, travel purpose, parking fee, and trip cost. Respondents were allowed to change, remove and merge the imputed data, and add new activity/travel data. Both the originally uploaded and validated data were automatically saved in the database. The system differentiates between transport modes (walking, running, bike, motor-bike, bus, car, taxi, train, metro and tram). It imputes activities according to 13 kinds of activities: home, paid work, voluntary work, study, daily shopping, non-daily shopping, service, bring and pick up, leisure, recreation, social, parent-children help and unspecified.

For our analysis, we selected respondents who: 1) have at least one car in their household; 2) participated in the survey every day continuously for at least one whole week; 3) made more than one trip per week. After cleaning the data, in total, 207 valid respondents and 5748 valid days were available for analysis. The travel diary data span a full year from May, 2012 to May, 2013. The respondents participated in different time periods and for different duration.

The socio-demographic variables are summarized in Figure 1. The share of male and female respondents is almost fifty/fifty. 66% of the respondents come from the Rotterdam region, while the rest comes from the Eindhoven region. Different from other online surveys, the age distribution of these respondents is not dominated by young persons. Almost 43% of the respondents is older than 55. Education levels were merged into three categories, which are low (primary school or below), middle (high school and technical school) and high (bachelor and above). Respondents are almost equally distributed across these three education levels. Most respondents have two or more persons in their household. 80% of the respondents has petrol cars. According to the Dutch National Travel Survey, most socio-demographic variables are consistent with the distribution of the Dutch population. However, the data oversampled high-income households, which earn more than 77.5k per year.

![Fig. 1. Sample Composition.](image-url)
Fig. 2. Home location distribution for Rotterdam area (--- Boundary of Rotterdam).

Fig. 3. Home location distribution for Eindhoven area (--- Boundary of Eindhoven).
To examine the influence of urban form on individual transport mode and activity location choices, their living environments were classified into city, suburb and rural area. Noted that this classification does not depend on scale and boundary conditions. As shown in Figures 2 and 3, the dashed areas are the main city areas of Eindhoven and Rotterdam. The small towns nearby could also be easily recognized.

Considering the fact that the analysis is based on panel data, a theory is required which would explain the generation of activities. Arentze and Timmermans\textsuperscript{14} proposed a need-based model based on the assumption that activities are driven by a limited and universal set of subjective needs at the person and household level. They assume that needs accrue over time. In line with this theory, the following variables were included to represents the dynamics of the evolution of needs: (i) three continuous variables count the number of days since the compulsory, maintenance and leisure activity has been performed the last time; (ii) a continuous variable counts the number of days the respondent did not make any trip during the week.

CO\textsubscript{2} emission also depends on car characteristics. We used type of car, car brand and type of fuel to calculate average fuel consumption (mpg) and average CO\textsubscript{2} emission (g/km) for each car brand. Data were obtained from a website (http://car-emissions.com). The distribution of energy consumption and CO\textsubscript{2} emission of cars are shown in Figure 4 for different car types. The fuel consumption is shown in the line. It ranges from 20.6 mpg to 70.8 mpg. The columns show CO\textsubscript{2} emission, which is between 317 g/km and 91.3 g/km. Generally, CO\textsubscript{2} emission increases with fuel consumption for all types of cars, except for some brands, which consume the same amount of fuel, but have relatively low CO\textsubscript{2} emissions. In particular, compared to diesel and petrol cars, LPG cars have lower CO\textsubscript{2} emissions.

3. Sample selection model for panel data

Sample selection models were selected to investigate the effects of socio-demographic variables and urban form on individual CO\textsubscript{2} emission by car. The panel data included individual daily activity-travel data, ranging from one week to three months. CO\textsubscript{2} emissions by car are influenced by both endogenous and exogenous factors. There are two steps in the prediction of emission: 1) whether or not to use the car on a certain day and 2) how much energy is consumed when travelling by car. These two steps are related.
The data contain 5749 days in total. However, 1669 days do not concern any car, which indicates that respondents do not use their car every day. The first reason for choosing sample selection model is that it can explicitly resolve the potential sample selection bias inherent in car use data. If the data are randomly sampled from this bivariate population, the parameters can be estimated by least squares or GLS. We assumed that the decision to use the car on a certain day is not a random selection. Each individual’s car use behavior is influenced by several exogenous factors, such as the days of the week, type of activities conducted on a certain day, and urban form. Thus, there may be several potential biases if this general selectivity problem of choosing the car as the transport mode is ignored.

Secondly, the sample selection model allows examining the two steps in a single model. It is possible to capture the effects of urban form on both transport mode choice (using car or not on a certain day) and CO₂ emission. Based on Green, we choose to use random effects sample selection models to capture time variations in terms of simple shifts of the regression function. Simulated maximum likelihood rather than two-step least squares was used to fit the model.

The basic structural equations for the sample selection model are a linear regression equation and a binary probit selection criterion model as shown in equations 1 to 3:

\[
\begin{align*}
  y &= \beta'x + \varepsilon \\
  z &= \alpha'w + u \\
  \varepsilon, u &\sim N[0,0,\sigma^2_\varepsilon, \sigma^2_u, \rho]
\end{align*}
\]

Values of \( y \) and \( x \) are only observed when \( z \) equals one. The essential feature of the model is that under the sampling rule, \( E[y|x,z=1] \) is not a linear regression in \( x \), or \( x \) and \( z \). The development below presents estimators for the class of essentially nonlinear models that emerge from this specification. To extend the basic model for panel data estimation, the random effects, \((\varepsilon_i, d_i)\) are assumed to be bivariate normally distributed with zero means, standard deviations \( \sigma_\varepsilon \) and \( \sigma_d \) and correlation \( \rho \). The random effects regression model of the panel data is equation 4.

\[
\log(y_{it}) = \beta_x x_{it} + \varepsilon_{it} + c_i + \varepsilon_{it} \sim N[0,\sigma^2_\varepsilon]
\]

The selection mechanism is shown in equations 5 and 6.

\[
\begin{align*}
  z_{it}^* &= \alpha'w_{it} + u_{it} + d_i \\
  z_{it} &= 1 \text{ if } (z_{it}^* > 0), u_{it} \sim N[0,1]
\end{align*}
\]

The correlation between \( \varepsilon_{it} \) and \( u_{it} \) is \( \rho \) as \( \text{Corr}[\varepsilon_{it}, u_{it}] = \rho \). Selectivity comes in two forms here, which is the correlation of the unique components \( \varepsilon_{it} \) and \( u_{it} \), and the correlation of the group specific components, \( c_i \) and \( d_i \).

4. Results

In estimating the selection model, we assumed that days of the week, living environment and activity needs affect on individual’s decision to use the car. The dependent variable of the regression is CO₂ emission by car. Living environment, socio-economic and activity needs variables were used as independent variables to predict CO₂ emissions. Although emissions are closely related to travel time and particularly travel times, these effects were not included because we felt that the exclusion of these variable would better reflect the impact of urban form on emissions, moderated by activity-travel patterns. In this section, we describe the final estimated models and report the results. The first step is to compute the probit model to define the selection mechanism. Then, this model is applied to produce good start values for the random effects model. Finally, the random effects model is estimated
using 100 Halton draws. Only significant results are shown in Table 1. The dependent variable is the daily CO₂ emission for the panelists.

Table 1. Results of the sample selection model for individual CO₂ emission.

| Variable | Coefficient | Error  | $z$  | $|z| > Z^*$ |
|----------|-------------|--------|------|------------|
| **Selection equation parameters** | | | | |
| Monday   | -0.01357*** | 0.0454 | -2.990 | 0.0028 |
| Tuesday  | -0.0845*  | 0.0512 | -1.650 | 0.0985 |
| Thursday | 0.1709*** | 0.0449 | 3.8100| 0.0001 |
| Friday   | 0.1531*** | 0.0564 | 2.7100| 0.0066 |
| Saturday | 0.2178*** | 0.0543 | 4.0100| 0.0001 |
| **Region** | | | | |
| Home located in Rotterdam region | 0.0987*** | 0.0295 | 3.3400| 0.0008 |
| **Neighborhood** | | | | |
| Home located in a city | -0.4339*** | 0.0560 | -7.7500 | 0.0000 |
| Home located in a town | -0.2408*** | 0.0482 | -4.9900 | 0.0000 |
| **Interaction of region and neighborhood** | | | | |
| Home located in Rotterdam city or Eindhoven rural area | -0.2557*** | 0.0474 | -5.3900 | 0.0000 |
| Age | 0.0090*** | 0.0014 | 6.7200 | 0.0000 |
| **Socio-economic** | | | | |
| Household size | 0.1260*** | 0.0184 | 6.8600 | 0.0000 |
| **Activity needs** | | | | |
| No-trip days in the week | 0.0446*** | 0.0158 | 2.8300 | 0.0046 |
| **Selection corrected regression parameters** | | | | |
| Home located in Rotterdam region | -0.0813**  | 0.0336 | -2.4200| 0.0157 |
| Home located in a city | 0.1156**  | 0.0522 | 2.2200 | 0.0267 |
| Home located in a town | -0.1620*** | 0.0451 | -3.5900 | 0.0003 |
| Home located in Rotterdam city or Eindhoven rural area | 0.1815*** | 0.0485 | 3.7400 | 0.0002 |
| **Interaction of region and neighborhood** | | | | |
| Male | 0.3914*** | 0.0470 | 8.3200 | 0.0000 |
| Education level middle | -0.2249*** | 0.0449 | -5.0100 | 0.0000 |
| **Socio-economic** | | | | |
| Work hour (12-30) | 0.1240**  | 0.0548 | 2.2600 | 0.0237 |
| High income household | 0.5894*** | 0.0548 | 10.7600 | 0.0000 |
| Household size | -0.0735*** | 0.0215 | -3.4100 | 0.0006 |
| **Activity needs** | | | | |
| Lagged days of maintenance activity | -0.0225**  | 0.0093 | -2.4100 | 0.0161 |
| **Means for random parameters** | | | | |
| One_Regr | 8.3706*** | 0.0873 | 95.8700 | 0.0000 |
| One_Prbt | 0.2688  | 0.1942 | 1.3800 | 0.1663 |
| sOne_Regr | 0.6088*** | 0.0225 | 27.0100 | 0.0000 |
| sOne_Prb | 0.6493*** | 0.0214 | 30.3800 | 0.0000 |
| **Disturbance standard deviation** | | | | |
| Sigma | 0.7377*** | 0.0074 | 100.4000 | 0.0000 |
| **Correlation between regression and probit** | | | | |
| Rho | -0.7111*** | 0.0140 | -50.7500 | 0.0000 |
The first stage probit regression analysis evaluates the effects of days of the week, urban spatial form, socio-demographic variables, weather prediction variables and needs of doing activities on the car use decision of individuals. The results of the probit regression showed car use differs for different days of the week. It is relatively high on Thursday, Friday and Saturday. However, results suggest that respondents are less likely to drive on Monday and Tuesday. There may be many reasons why people are less likely to use their car on Monday. For example, part-time workers usually choose Monday and/or Friday as non-work days. Most social and leisure activities are conducted during weekends, which reduce the probability of conducting these activities on Monday. The urban spatial form variables show that people living in Rotterdam region are more likely to use car. Comparing to the living conditions, the results found that people living in a city or a town are less likely to use car. Especially the people living in Rotterdam city area or Eindhoven rural area are less likely to use car. From socio-demographic variables, we found that with the increasing of age, people are more likely to choose cars as their transport mode. Moreover, for household size, the results indicate, individual comes from a larger family is more likely to use cars as transport mode. The last variable describes the effect of needs on transport mode choice. We assume that with the increasing of non-trip days, the need of going out and conducting activities is increasing, and individual are more likely to use car to travel. The results proved it.

The second part of the Tables 1 (selection corrected regression parameters) shows the results of random effects model. The first four variables presented the effects of urban spatial form effects on individual’s CO2 emission. The results indicate that people living in city and especially Rotterdam city area produce more CO2 emission than people living in other places. However, comparing the people living in Rotterdam area and Eindhoven area, the results show that on average, individual living in Rotterdam area produce less CO2 than living in Eindhoven area. Comparing with living in a city or a rural area, people who are living in small towns nearby cities produce less CO2 emission by car. As for the effects of socio-demographic variables, results indicate that males produce more CO2 emission by car than female. Moreover, with increasing household size, CO2 emission by car is also increased. People work 12 to 30 hours per week or comes from a high income household produce more CO2 emission. However people with middle education levels produce less CO2. For other variables, the results did not show any significant effects. The effect of lagged days of maintenance activities show that with the increasing needs of conducting a maintenance activity, people are more likely to choose a nearby place to conduct maintenance activity instead of driving long distance.

The random parameters capture temporal variations across individuals in simple shifts of the regression function, such as changes in the intercept. The random effects are assumed to be bivariate normally distributed with zero means. The results show that one of the mean values are not significantly different from 0. Standard deviations of random parameters suggested that they have a correlation of 0.61 and 0.65 respectively, which are relatively high. It means that any component of the error that makes selection more likely increases CO2 emission. However, errors are tied up with model specification, alternative specifications change the errors, which in turn changes Rho.

5. Conclusion and discussion

In urban planning and development, settlements are crucial for sustainable growth because they are the major and important centers of the origin of environmental impacts and their residents are the motors of sustainable development. However, residents have little understanding of the impacts of their behavior on the environment. The study explored, to some extent, how does spatial and urban environment design could influence on individual’s car usage and CO2 emission by examining individual’s travel behavior panel data.

Using sample selection model, results indicate that people’s CO2 emission by car could be explained by two steps: 1) whether they would like to use a car on a particular day; 2) how long distance they would like to travel for the car trips. Sum up the effects of urban special form on individual’s car usage and CO2 emission, we found that there is no directly causality between the two steps. The individual who are less likely to use cars does not always produce less CO2 emission, only except the people who are living in small towns nearby cities. The CO2 emission is correlated with both the travel distance and car types. It indicates that although people living in Rotterdam city area and Eindhoven city/rural area are less likely to use cars as transport mode, but they usually travel more distance or have more polluted cars. In general, the results indicate that, people who are living in Rotterdam area produce less CO2 emission comparing to the people living in Eindhoven area. People living in small towns nearby both
Eindhoven and Rotterdam cities consume less CO$_2$ than living in city area or rural area. Comparing city scale, we found that people living in Eindhoven city consume less CO$_2$ than people living in Rotterdam city. Physical planning can benefit from this result to have a deeper insight into the space-energy design at individual level in city scales. The results could also benefit in future smart city planning to reduce the energy use of residents in different urban spatial area.

Some limitations need to be mentioned and examined in the future research. Firstly, the CO$_2$ emission is calculated according to travel distance and car types disregarding acceleration speeds and stops. It is potentially underestimate the CO$_2$ emissions in cities where there are more signal lights. Secondly, the road information does not take into account either, which may have effects on energy consumption and CO$_2$ emission especially for out-of-city travelling. These limitations will be examined in our future research.

Acknowledgements

The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n° 230517 (U4IA project). The views and opinions expressed in this publication represent those of the authors only. The ERC and European Community are not liable for any use that may be made of the information in this publication.

References

1. Mathis W., William R. Our ecological footprint: reducing human impact on the earth. Gabriola Island: New Society Publishers; 1998.
2. Handy S. Methodologies for exploring the link between urban form and travel behavior. Transp Res D 1999; 1(2):151-65.
3. Boarnet M, Crane. R. The influence of land use on travel behavior: Empirical strategies. Transp Res A 2001;35(9):823-45.
4. Ewing R, Cervero R. Travel and the built environment. Transp Res Rec 2001;1780:87-114.
5. Kulash W, Joe A, David M. Traditional neighborhood development: Will the traffic work? Development 1990;21:21-4.
6. McNally, MG, Ryan S. Comparative assessment of travel characteristics for neotraditional designs. Transp Res Rec 1993;1607:105-15.
7. Friedman B, Gordon S.P, Peer J.B. Effect of neotraditional design on travel characteristics. Transp Res Rec 1994;1466:63-70.
8. Cervero R, Landsis J. The transportation – land use connection still matters. Access 1995;7:2-10.
9. Handy SL, Clifton K.J. Local shopping as a strategy for reducing automobile travel. Transp 2001;28:317-46.
10. Handy SL, Boarnet M.G., Ewing R, Killingsworth R.E. How the built environment affects physical activity: Views from urban planning. Am J Prev Med 2002;23:64-73.
11. Kitamura R, Mokhtarian P.L, Laidet L. A micro-analysis of land use and travel in five neighborhoods in the San Francisco Bay area. Transp 1997;24(2):125-58.
12. Holden E. Ecological footprints and sustainable urban form. J Hous and the Built Env 2004;19:91-109
13. Fan Y, Khattak A.J. Urban form, individual spatial footprints and travel: Examination of space-use behavior. Transp Res Rec 2008;2082:98-106.
14. Arentze TA, Timmermans HJP. A need-based model of multi-day, multi-person activity generation. Transp Res B 2009; 43(2):251-65.
15. Greene W. Fixed effects vector decomposition: A magical solution to the problem of time invariant variables in fixed effects models?. Polit Anal 2011; 19(2):135-46,170-72.