Explaining travel distance during same- day visits

Jan Wynen

KU Leuven

June 2012

ABSTRACT

Distance travelled is a recurrent determinant of expenditures during same- day visits and destination choice. Identifying the factors determining distance travelled will thus lead to a better understanding of expenditures and location choice, which is of interest from a policy and economic perspective. By applying a specific analytical framework which explicitly considers the different workings of socio-demographic factors on the decision to go on a same- day visit as well as the distance related effects of a same- day visit, we are able to investigate the effects of both socio demographic and same- day visit related variables on distance travelled. The empirical results highlight that gender, time spent, age, education and information source are important determinants of travel distance during same- day visits.

Keywords: Travel distance, Same-day visits, Heckman selection model, Non- linear effects

Jan Wynen
KU Leuven
Parkstraat 45 bus 3609
3000 Leuven
Belgium
Phone: +32 16 32 36 14
Email: jan.wynen@soc.kuleuven.be or
jan.wynen@econ.kuleuven.be

1 Same- day visits and travel distance

During the second half of the last century, tourism became one of the most important and rapidly growing sectors in the world economy (Vanhove, 2011, Chapter 2). According to the European Tourism Satellite accounts (TSA), Europe is the tourist destination N° 1 in the world. The tourism sector employs between about 8.6 and 24 million people in the Community (corresponding to around 4 and 11 % of total labor force) and creates 4 or 11 % of the Community's GDP (TSA, 2010). This major economic importance of tourism for the economy increased the awareness of the importance of the sector, its socio-cultural and environmental impacts as well as its contributions to the achieving of the goals of the Lisbon strategy - to increase economic growth and to create more and better jobs. Realizing the growing significance of tourism, governments, local authorities and private sectors in many countries, regions and communities have thus begun to funnel their resources into tourism development. However governments often still underestimate the economic benefits that tourism provides because it is not visible in the same way as other industries. Even until now there has been insufficient attention to measuring these economic benefits on an equal footing with other sectors. Measuring these benefits is however hard since only a small part of tourism spending takes place in what are normally considered to be tourism industries, and in fact a large part takes place outside the classical tourism industries such as self- catering, museums, public transport, etc (Vanhove, 2011, Chapter 2) . However, over the past decade, successive Commission initiatives have fostered introduction and implementation of TSA's in the EU Member States. According to the data available in these TSA's (2010), the tourism sector receives more revenues from domestic tourism than from inbound tourism. Furthermore same-day visits contributed more to domestic tourism expenditures than overnight trips (TSA, 2010). However it is a well known fact that official tourism data totally neglects day visitors. Consequently, many destinations have no or only a vague knowledge of the number of day visitors. Since domestic tourism expenditures make a significant contribution to the economy in almost every European country and same- day visit expenditures are an important part of these domestic tourism expenditures (the Netherlands report a 60- 70% share of expenditure by same- day visitors in total domestic expenditure (TSA, 2010)), this is remarkable. The few studies examining

same- day visits mostly focus on expenditures (e.g. Wynen, 2011; Thrane and Farstad, 2011). Based on these studies, a number of factors have been empirically associated with variations in expenditures. A vital variable in this regard is time spent, typically measured for same- day visits, with the inclusion of travel time (Wynen, 2011; Kozak, Gokovali, & Bahar, 2008; Laesser & Croch, 2006; Leones, Colby, & Crandall, 1998; Mehmetoglu, 2007; Mok & Iverson, 2000; Pouta, Neuvonen, & Sievänen, 2006; Seiler, Hsieh, Seiler, & Hsieh, 2002). Not only is travel time often included in time spent, but according to Pollak (1976), one can also expect that travel distance has a positive effect on time spent during a same- day visit. The determinants of distance traveled can therefore be seen as an indirect determinant of expenditures during same-day visits. Identifying the determinants of distance travelled will lead to a better understanding of expenditures during same- day visits.

Moreover, in micro level research focusing mostly on outbound tourism, distance travelled is also regarded as a key determinant when it comes to explaining destination choice (e.g. Nicolau, & Mas, 2006). Because of limited data, studies examining location choice on a micro level for same- day visits are still nonexistent. Understanding the drivers behind destination choice are however a first step in explaining and understanding destination choice for same- day visits, which is of course important not only from an economic point of view but also from a policy perspective. This thus stresses the relevance of identifying the factors that affect distance travelled during same- day visits.

Overall, distance travelled can thus be regarded as a key determinant of expenditures and destination choice. The core question in this article is therefore: what are the factors that determine distance travelled during same- day visits. The data used in this paper demand for a review of the methodologies used in the field of tourism in general. Studies focusing on identifying the characteristics determining distance traveled or expenditures have tended to investigate the effect of socio demographic variables using simple methodologies such as comparisons of frequencies, t-tests, one- way analysis of variance, logistic regression and ordinary least squares. Since the decision of the distance to travel on a daytrip is nested in the decision of going on a daytrip a specific analytical framework is needed. In what follows, a framework will be presented which explicitly considers the

different workings of socio- demographic factors on the decision to go on a same- day visit as well as the distance related effects of a same- day visit.

Literature examining tourist'travel distance determinants during same- day visits is still incipient. Regarding the few specific studies existing on this issue, we must note that most of them proceed in a descriptive way, just analyzing the relationship between travel distance or length of stay (travel time included) and the socio economic profile of the tourist (e.g. Sung et al. 2001; Bernham and Kim, 1999; Alegre and Pou, 2006). A considerable amount of studies have been done however focusing on the concept of travel distance for the case of commuting to work (e.g. Cao and Mokhtarian, 2005; Lloyd and Shuttleworth, 2005; Scott et al., 1997). Although these studies mainly focus on the implications to road efficiency and environmental impacts, this research has shown that travel distance is contingent upon a number of factors. Age, gender, employment status, household size and vehicle ownership are among of the most cited variables having an effect on distance traveled. Several studies report a negative relation between age and travel distance (Rosenbloom, 1995; Benekohal et al., 1994; Chu, 1994; Boarnet and Sarmiento, 1998). Vance and Iovanne (2007) found significant effects of vehicle ownership on distance traveled. The effect of household size is hypothesized to have a negative effect on distance traveled (Stradling et al., 2005; Páez et al., 2007 and Kanaroglu and Scott, 2002). Vance and Iovanna (2007), suggest a negative relationship between travel distance and employment status.

The present study adds to literature on same- day visits in two important ways. First, the study draws attention to distance travelled which is an important determinant of expenditures and location choice. Second, the study presents a specific analytical framework, which explicitly considers the different workings of socio- demographic factors as well as the distance related effects of a same- day visit. The remainder of the paper is organized as follows; next section discusses the empirical implementation. In this section explanatory variables are presented and a test is performed on the relation between travel distance and expenditures. The section concludes with and an in- depth analysis of the relationship between the explanatory variables and distance traveled. Section 3 concludes.

2 Empirical implementation

In our analysis, a same- day visit or daytrip is defined as follows:

> *"A same day visit has to take place on more than 20km from home (except for round trips on foot, by bike, by horse) and should take at least 4 hours (inclusive travel time), without staying the night someplace else. Same- day visits should not include: visiting family, friends and acquaintances, business trips, day trips from a holiday address or second home, or trips with a routinous character (e.g. frequently doing sports, daily shopping and travelling between home and work)"*

Contrary to Belgium's formal definition of day trippers, the definition thus not allows visiting friends and acquaintances to be regarded as a same- day visit and only under very strict circumstances can a visit be regarded as a same- day visit when it takes place on less than 20km from home.[1] Following this definition, we analyze the effect of both socio demographic and same- day visit related variables by drawing on data gathered by iVOX[2] and IDEA Consult[3] as part of a study to examine same- day visits in Belgium. These data are based on a survey and contain a wide range of socio demographic as well as same- day visit related variables. The data have been collected on a weekly basis and cover a period of 12 months (July 2010- June 2011). The respondents have been selected to be representative for the Belgian population. Each week, they were asked by email whether or not they had made a same-day visit and if so, to fill in a questionnaire. If respondents failed to answer they were replaced by someone with the same characteristics. For each week, information is available for +/- 2000 respondents. Table 1 gives an overview of the representativeness of the data. As indicated in the text above, the data are a good reflection of the Belgian population, the difference between the column percentage population and percentage in the sample are relatively small.

[1] The used definition is based on the dataset and was requested by the contractor of the data. The author had no influence herein. However since the principal exists of several tourism organizations in Belgium, and since these organizations want to use this dataset in a consultancy environment the author believes the definition to be a good one, one that is closely related to practice.

[2] iVOX is a market research facilitator specialised in both online and offline fieldwork. See http://www.ivox.be/about-ivox for more information.

[3] IDEA Consult is a research-based consultancy company. See http://www.ideaconsult.be for more information.

INSERT TABLE 1 HERE

Concerning distance traveled, respondents were asked following question:

What is the distance from your location (residence) to the farthest city you have visited?

This survey can be regarded as an efficient and reliable instrument for measuring same- day visit characteristics since there is a weekly registration. This increases the accuracy of results by avoiding the recall bias or memory effects (Frechtling, 2006).

Information on the number of same- day visits and the median distance traveled on daytrips is presented in Table 2. Not surprisingly, most same- day visits were made during summer. Median distance traveled ranges from 60- 75 km. In total 9553 same- day visits have been made during the observed period, having a median distance of 70 km.

INSERT TABLE 2 HERE

2.1 Explanatory variables

Theory provides no guidance as to which explanatory variables determine the distance travelled during same- day visits. A wider range of papers (e.g. McGuckin et al., 2005; Lee and McDonald, 2003; Mercado and Páez, 2009), examining how a number of determinants affect distance travelled in general does however exist. As recommended by these prior studies, we include a broad range of variables. One such variable is age. Previous studies on distance travelled have shown that age is negatively related to trip distance (Chu, 1994; Boarnet and Sarmiento, 1998; Mercado and Páez, 2009). In order to control for a possible non- linear relationship, we include the squared term of this variable. Since we are using pooled cross sections, we include season dummies to account for aggregate changes over time (Wooldridge, 2010, Chapter 7). Although season does not seem to affect distance travelled, it has a big effect on the probability of going on a same- day visit. Another included explanatory variable is time spent during the daytrip itself, travel time excluded. Furthermore a long list of dummies are taken into account, these assess the impact of marital status, gender, education (divided over three dummies: elementary school, indicating whether or not the highest obtained degree of the respondent is a degree

of elementary school (omitted variable, used as a benchmark); secondary school, indicating whether or not the highest obtained degree of the respondent is a degree of secondary school; higher education, indicating whether or not the highest obtained degree of the respondent is a college or university degree), retired and student. Following Bieger and Laesser (2004) we also include the categorical variable info (indicating whether or not the respondent visited websites, read books, articles,… in order to prepare for the daytrip). We would expect search costs to be higher for some information sources and these search costs to be positively correlated with distance travelled. Number of cars owned is included as an explanatory variable following Mercado and Páez (2009). Finally, region dummies are included in order to control for heterogeneity across regions. Table 3 shows definitions, means and standard deviations for the variable used.[4]

INSERT TABLE 3 HERE

The mean distance traveled equals 88.76 km. This is a more than the median reported, indicating that distance traveled ranges from very close (1 km) to very far (500 km). The dummy female has a mean of 0.51, meaning that there is an equal spread between men (0) and women (1). Time spent (excluding travel time) equals 6.34 hours. It has however a standard deviation of 3.79 hours. This indicates that there is a large spread in time spent during a same- day visit. Some visits only lasted 2 hours (excluding travel time), while others visited a place for around 20 hours[5]. In Figure 1 a more detailed overview is presented. The average age of a respondent equals 45.75 year with a maximum of 85 year and a minimum of 18 years old. Most of the respondents are married (63%), however only 36% of them have children. 72% report to have at least a secondary school degree, 34% indicate to have a college or university degree. An overview of information sources used to prepare a same- day visit is presented in Figure 2. If information is gathered to prepare a same- day visit, it mostly comes from other sources (television, friends or family, previous experience,…) than brochures or the internet. Internet is more frequently used as an information source than brochures. Intuitively this makes sense

[4] Variables were tested on multicollinearity, the highest VIF reported was 1.68. Indicating that there are no problems with multicollinearity.
[5] This is no data error. There are 10 observations reporting to have gone on a same-day visit for 20 hours. All of these went to an amusement park (Disneyland, Paris).

since search costs are lower, see Bieger and Laesser (2004) for a more thorough discussion on this topic.

INSERT FIGURE 1

INSERT FIGURE 2

As mentioned earlier, theory suggests that distance travelled is strongly related with expenditures during same- day visits. Intuitively this makes sense, people who travel a large distance tend to stay longer (this especially holds for same- day visits, since these only have a limited time scope). And the longer people stay in one place, the more likely they are to spend money there. In order to test this hypothesis we transform our dependent -distance travelled- to a 4 point interval scale. Furthermore, the survey includes information on expenditures made during same- day visits. On average, people spend around 70 euro's on a same- day visit, there is however a large spread in the variable. The maximum spend on a same- day visit is 400 euro's[6] while the minimum is zero. A one- way analysis of variance (ANOVA) is used to test for differences in the means of expenditures broken down by the levels of distance traveled. As indicated in Table 4 the mean of expenditures differs significantly among the levels of distance travelled. However, we do not know if the difference is between only two of the levels or all of them. By taking the mean of expenditures for each level of distance travelled, we can see which level of distance travelled leads to the highest mean expenditure. Table 4 is in line with our hypothesis, the farther people travel; the more they spend during same- day visits. Distance travelled therefore proves to be an important determinant of same- day visit expenditures.

INSERT TABLE 4 HERE

2.2 Empirical Analysis

To estimate the relationship between the variables outlined above and distance travelled during same-day visits, a specific econometric model is required. Since the decision of the distance to travel on a same- day visit is nested in the decision of going on a daytrip, ordinary least squares (OLS) is not suitable (Cameron and Trivedi, 2005, Chapter 16). Another drawback of OLS estimation is the fact that

[6] This is no data error; again visitors of amusement parks cause this outlier.

marginal effects are constrained to be constant, which is an unrealistic assumption in our model. A specific analytical framework is thus required. This framework explicitly has to consider the different workings of socio- demographic factors on the decision to go on a same- day visit as well as the distance related effects of same- day visit related variables. This suggests a distinction between the decision to go on a same- day visit and the distance decision while taking into account that these decisions are taken simultaneously rather than consecutively. Therefore a Tobit estimation would not suffice, since any variable which increases the probability of a non- zero value must also increase the mean of positive values; a positive element of β means that an increase in the corresponding variable (element of X_t) increases both $P(y_t > 0)$ and $E(y_t | y_t > 0)$. The Tobit model uses all the available information from the explanatory variables, it includes thus both the decision whether or not to go on a same- day visit and the decision of distance in one model. In order to solve this issue an alternative specification is required which separates both decisions. A two step model, as proposed by Cragg (1971), fulfills these criteria. By using a Probit model in the first stage; the decision of going on a daytrip is modeled[7]. While in the second stage a truncated regression is used as the dependent variable (distance traveled) is observed only if it is greater than zero. The two parts are assumed to be independent and are estimated separately. In order to account for regional heterogeneity, we include 3 region dummies (Flanders, Wallonia and Brussels). The two part model is given below in equation (1) and (2).

(1) $$Y = \alpha + \beta_1 Age + \beta_2 Age^2 + \beta_3 Female + \beta_4 Retired + \beta_5 Student + \beta_6 Children + \beta_7 Married + \beta_8 Season + \beta_9 Cars + \beta_{11} Education + \beta_{12} Region + \varepsilon$$

Where (1) is the selection equation and Y is a latent variable corresponding to the binary variable indicating going on a same- day visit (Y = 1 if the respondent goes on a same- day visit and Y = 0 otherwise), and (2) is the truncated regression, whereby PX refers to distance travelled.

(2) $$PX = \delta + \gamma_1 Information + \gamma_2 Time\ spent + \gamma_3 Age + \gamma_4 Age^2 + \gamma_5 Female + \gamma_6 Retired + \gamma_7 Student + \gamma_8 Children + \gamma_9 Married + \gamma_{10} Season + \gamma_{11} Cars + \gamma_{12} Education + \gamma_{12} Region + \mu$$

[7] The probit model correctly predicts 88.51% of the outcomes for model 1.

From an econometric point of view, this two- part model is flexible and attractive because it allows different covariates to have a different impact on the two parts of the model. However in this case, an independent two-step approach can be criticized on theoretical grounds. In some cases, there is no such thing as a two-step decision- to go on a daytrip or not, and then how far to go. It happens that a respondent simultaneously decides if he goes on a same-day visit and how far he goes. The econometric model thus has to account for this possibility. In order to do so, a link has to be established between the two stages. This can be done by utilizing a Heckman selection model, where the inverse Mills ratio accounts for the dependency between the two stages. Furthermore, the inverse Mills ratio also accounts for a possible selection bias.

The two- equation model comprises a selection equation for y_1, where

$$(1)\quad y_1 = \begin{cases} 1\ if\ y_1^* > 0 \\ 0\ if\ y_1^* \leq 0 \end{cases}$$

and a result outcome equation for y_2, where

$$(2)\ y_2 = \begin{cases} y_2^*\ if\ y_1^* > 0 \\ -\ \ if\ y_1^* \leq 0 \end{cases}$$

y_1^* determines whether an individual goes on a daytrip, y_2^* determines distance traveled on a daytrip, and $y_1^* \neq y_2^*$. The selection equation is estimated by maximum likelihood as an independent probit model to determine the decision to join using information from the whole sample (people who are not going on a daytrip and people who go on a daytrip). A vector of inverse Mills ratios are generated from the parameter estimates (Greene, 1993, Chapter 22). Distance traveled on a daytrip is observed only when the selection equation equals 1 (people go on a daytrip) and is then regressed on the explanatory variables (which can now include daytrip related variables[8]), and the vector of inverse Mills ratios from the selection equation by ordinary least squares. The second stage runs the regression with the estimated error included as an extra explanatory variable, removing the part of the error term correlated

[8] According to Wooldridge (2010, Chapter 19) the variables in the outcome equation should be a strict subset of the variables in the selection equation, however as Verbeek (2004, Chapter 7) indicates, 'the only reason not to include a certain variable in the selection equation that is included in the outcome equation is that we are confident that is has a zero coefficient in the selection equation'. Consequently we include extra variables in the outcome equation that have a zero coefficient in the selection equation. As a robustness check, we also show the results of the two- step model.

with the explanatory variable. Although the relationship between the inverse Mills ratio (Heckman's lambda) and the estimated β's after the Probit model is strongly non-linear, we opt for a Heckman selection model with an exclusion criterion as described by Wooldridge (2010, Chapter 19). Following Wooldridge (2010, Chapter 19) this variable should only have an effect on the probability of going on a same- day visit but should not have an effect on the structural equation, in other words the distance travelled on a same- day visit. As exclusion criterion we include high and mid-season. Season plays a crucial role on the probability of going on a same- day visit. For instance, we suspect that the probability of going on a same- day visit increases when people have vacation or when the weather is good. However we do not suspect season to have an effect on distance travelled. Following this line of reasoning we decided to include season as an exclusion criterion. Furthermore, the results of the two-part model support this hypothesis. We notice a significant result of season in model 1 and an insignificant result for season in model 2 of Table 5. The variable season thus has a significant effect on the probability of going on a same-day visit, but does not have an effect on distance travelled. This proves that season can be regarded as a good exclusion criterion for our Heckman model.

The estimated coefficients in the Heckman selection model can however not be interpreted in the same way as in a linear regression model. To assess the impact of the regressors on the dependent variable, it is necessary to analyze their marginal effects. For the continuous explanatory variables, these marginal effects are used to calculate elasticities at the sample means. For the discrete or categorical variables, the marginal effects are used to calculate percentage changes in the dependent variable when the variable shifts from zero to one, ceteris paribus. Standard errors for the marginal effects are obtained by the delta method.

2.3 Results

Table 5 includes the results of the two step model and the Heckman selection model. Model 1 and 2 show the results of the two- step model, where model 1 is the probit model and model 2 is the truncated regression. Model 3 and 4 include the results of the Heckman selection model. Model 3 estimates the selection equation, while model 4 analyses the structural equation. As discussed above, the two- part model attains some of its flexibility and computational simplicity by assuming that the

two parts - the decision to go on a same- day visit and distance travelled – are fully independent. Furthermore, the results of the two- step model are very similar to those of the Heckman estimation, showing that the results of the Heckman model are robust. Because of these reasons, only the results of the Heckman model (model 3 and 4) will be discussed. For a detailed discussion on the interpretation of the estimated coefficients of the Heckman model we refer to Puhani (2000).

When examining the results, we notice that age plays both a roll on the probability to go on a same-day visit and on distance travelled. The square term indicates a concave (i.e. inverted U- shaped) relationship. The possibility and the distance travelled increases when people get older, however they start to decrease from a certain point. The likelihood of going on a same- visit increases until the age of 75, distance travelled increases until the age of 45, after which it decreases, see Figure 1. Distance travelled is projected on the left axis while the probability to go on a same- day visit is projected on the right axis.

INSERT FIGURE 2 HERE

When investigating the effect of education we notice that a higher education, college or university, leads to a higher probability to go on a same- day visit while also the distance travelled will be higher. As one would suspect, the probability of going on a same- day visit is higher in High season than during Low season. Students have a lower probability to go on a same- day visit and travel less far; a possible explanation for this observation could be that this group has often less financial resources, thus limiting their mobility and willingness to go on a daytrip. Time spent on the daytrip (travel time excluded) has a strong positive effect on distance traveled. People thus tend to stay longer if they have to travel far, confirming part of the hypothesis formulated in the beginning of this article. If respondents gathered information via brochures they tend to travel further as when they would gather information on internet. This is in line with Bieger and Laesser (2004), but also intuitively this makes sense, people who gather information through brochures are more committed since the search costs are higher. People who undertake a longer same-day visit want to be sure that their visit is worthwhile, and are consequently prepared to pay higher search costs and check brochures. A negative effect can be

observed when respondents gathered information through other sources or didn't search for information. In those cases people who checked the internet tend to travel farther. Again this in line with the above stated theory of Bieger and Laesser (2004). Gender has a clear impact on the probability of going on a daytrip and on distance; women are less inclined to go on a same- day visit. Furthermore, women tend to travel less far than men, which is in line with literature on distance traveled in general (e.g. Schmocker et al., 2005). In line with literature on distance traveled in general, we find a positive effect for the number of cars owned on distance traveled during same-day visits. We do however not find a significant effect of children on distance travelled; however having children decreases the probability to go on a same-day visit.

INSERT TABLE 5 HERE

3 Discussion and Conclusion

As mentioned in the introduction, most studies examining distance travelled focus on the implications to road efficiency and environmental impacts. However, no literature exists analyzing the determinants of distance traveled during same-day visits. Distance travelled is however an important determinant of expenditures during same-day visits but also a key determinant of location choice. Travel distance will therefore have an indirect effect on the economic impact of day trips and can be of interest from a policy perspective. Consequently, this study's main objective was to understand the determinants of distance travelled during same-day visits. The methodology used was an independent and dependent two- step model. By following this methodology not only distance traveled could be analyzed but also the probabilities of going on a same- day visit were examined. Certain variables seem to be more important than others, and variations were apparent in the effects of the determinants on distance travelled. Gender, age, children, education, time spent, season and information were identified as significant predictors for distance traveled during same-day visits. While some of these results, such as season, age and time spend were intuitive, some were quite surprising; the effect of information and education. The more education, the higher the likelihood to go on a same- day visit and higher

educated people tend to travel farther. A final surprising result was discovered for the variable information. People who got information from brochures tend to travel farther than people who used the internet to get information. However people who used internet to get information travelled farther than people who used other information sources or who didn't consult information sources. From a policy perspective, this is an important finding. In order to maximize economic benefits related to same- day visits, industry should not only focus on new social media such as Facebook, Twitter and Youtube. Although this research has indicated that internet is often used as information source, results suggests that people will travel farther and consequently spend more money if they used brochures to prepare for the same- day visit. When people travel farther, they want to be sure that their visit is worthwhile, therefore they are prepared to increase search costs and are more inclined to examine brochures. In order to maximize the economic benefits of same- day visits, industry should thus also focus on this kind of information source, and not neglect it in favor for the new social media. Furthermore, industry and policy makers should be aware of the different impacts of socio-demographic variables on different aspects of a same- day visit. For example, females will be less inclined to go on a same- day visit, yet the difference with males is very limited. When trying to attract as many same- day visitors as possible gender will only play a marginal role. However, gender is an important indicator of distance travelled, women travel less far. Since distance travelled is an important indicator of expenditures, this is, from an economic perspective, an important observation and should be taken into account be industry and policy makers.

In a further step, it would be interesting to use the available dataset to research the impact of socio-demographic and same- day visit related variables on the choice of location of a same- day visit. This could be done by using multinomial -probit or -logit model. For future studies, however, it would be desirable to collect panel data over years to better investigate the effect of the discussed determinants.

4 References

Alegre, J., and Pou, L. (2006). The length of stay in the demand for tourism. *Tourism Management*, 27, 1343-1355.

Bernam, M. D. and Kim, H-J. (1999). Endogenous on-site time in the recreation demand model. *Land Economics*, 75(4), 603–619.

Benekohal, R.F., Michaels, R.M., Shim, E. and Resende, P.T.V. (1994). Effects of aging on older drivers' travel characteristics. *Transportation Research Record*, 1438, 91-98.

Bieger, T. and Laesser, C. (2004), Information Sources for Travel Decisions: Toward a Source Process Model, *Journal of Travel Research*, 42, 357-371.

Boarnet, M. and Sarmiento, S. (1998). Can Land-Use Policy Really Affect Travel Behaviour? A Study of the Link Between Non-Work Travel and Land-Use Characteristics. *Urban Studies*, 35(7), 1155-1169.

Cameron, A.C. and Trivedi P.K. (2005). *Microeconometrics: Methods and Applications*. New York: Cambridge University Press.

Cao, Xinyu & Mokhtarian, Patricia L. (2005). How do individuals adapt their personal travel? Objective and subjective influences on the consideration of travel-related strategies for San Francisco Bay Area commuters. *Transport Policy*, 12(4), 291-302.

Chu, X. (1 994). The Effects of Age on Driving Habits of the Elderly: Evidence from the 7990 NPTS. Tampa, FL: University of South Florida Center for Urban Transportation Research.

Cragg, John G. (1971) Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods. *Econometrica* 39, 829-844.

Frechtling, D. C. (2006). An Assessment of Visitor Expenditure Methods and Models. *Journal of Travel Research*, 45(1), 26-35.

Greene, William H. (1993). *Econometric Analysis*. Second edition. New York: Macmillan.

Kanaroglou, P. and Scott, D. (2002). Integrated urban transportation and land-use models for policy analysis. In M. Dijst, W. Schenkel and I. Thomas (eds.), *Governing Cities on the*

Move:Functional and Management Perspectives on Transformations of European Urban Infrastructures. Aldershot (UK): Ashgate, 43-73.

Kozak, M., Gokovali, U., & Bahar, O. (2008). Estimating the determinants of tourist spending: a comparison of four models. *Tourism Analysis*, 13(2), 143-156.

Laesser, C., & Crouch, G.I. (2006). Segmenting markets by travel expenditure patterns: the case of international visitors to Australia. *Journal of Travel Research*, 44(4), 397 -406.

Lee, B.S. & McDonald, J.F.(2003). Determinants of commuting time and distance for Seoul residents: The impact of family status on the commuting of women. *Urban Studies*, 40(7), 1283-1320.

Leones, J., Colby, B., & Crandall, K. (1998). Tracking expenditures of the elusive nature tourists of Southeastern Arizona. *Journal of Travel Research*, 36(3), 56- 64.

Lloyd C., Shuttleworth. (2005). Analysing commuting using local regression techniques: scale, sensitivity, and geographical patterning. *Environment and Planning*, 37(1), 81 – 103.

McGuckin, N., Johanna Z., and Yukiko N. (2005). Trip-Chaining Trends in the United States: Understanding Travel Behavior for Policy Making. *Transportation Research Record*, 1752, 199-204.

Mercado, R.G., Páez, A. (2009). Determinants of Distance Traveled with a Focus on the Elderly: A Multilevel Analysis in the Hamilton CMA, *Canada. Journal of Transport Geography* 17, 65-76.

Mehmetoglu, M. (2007). Nature- based tourists: the relationship between their trip expenditures and activities. *Journal of sustainable Tourism*, 15(2), 200- 216.

Mok, C. & Iverson, T.J. (2000). Expenditure- based segmentation: Taiwanese tourists to Guam. *Tourism management*, 21(3), 299-305.

Nicolau, J. L., & Mas, F. J. (2008). Sequential choice behavior: Going on vacation and type of destination. *Tourism Management*, 29(5), 1023–1034.

Páez, A.P., Scott, D.M., Potoglou, D., Kanaroglou, P., and Newbold, K.B., (2007). Elderly Mobility: Demographic and Spatial Analysis of Trip Making in the Hamilton CMA, Canada. *Urban Studies*, 44(1), 123-146.

Pollak, R.A. (1976). Interdependent Preferences, *American Economic Review*, 66, 309- 320.

Pouta, E., Neuvonen, M. & Sievänen, T. (2006). Determinants of nature trip expenditures in Southern Finland – implications for nature tourism development. *Scandinavian Journal of Hospitality & Tourism*, 6(2), 118-135.

Puhani, P. (2000). The Heckman Correction for Sample Selection and its Critique, *Journal of Economic Surveys*, 14, 53-68.

Rosenbloom, S. (1995). Travel by women. *Demographic Special Reports*, 1990, 2 -57

NPTS Report Series. U.S. Department of Transportation, Federal Highway Administration. Washington DC: Government Printing Office.

Schmocker, J-D., Quddus, M., Noland, R., Bell, M. (2005). Estimating trip generation of elderly and disabled people: Analysis of London data, *Transportation Research Record*, 1924.

Scott, D., Kanagaroulu, P. and Anderson, W.(1997). Impacts of commuting efficiency on congestion and emissions: case of the Hamilton CMA, Canada, Transportation Research D, 2, 245–257.

Seiler, V., Hsieh, S., Seiler, M., & Hsieh, C. (2002). Modeling travel expenditures for Taiwanese tourism. *Journal of Travel and Tourism Marketing*, 13(4), 47-61.

Stradling, S., Kinnear N. and Mann, H. (2005). Evaluation of Brake Bridge House Trust Young. *Driver Education Scheme Pilot Project*. Edinburgh: Napier University.

Sung, H.H., A.M. Morrison, G.S. Hong, and T. O'Leary (2001). The Effects of Household and Trip Characteristics on Trip Types: A Consumer Behavioural Approach for Segmenting the U.S. Domestic Leisure Travel Market. *Journal of Hospitality & Tourism Research*, 25 (1), 46- 48.

Thrane, C., & Farstad, E. (2011). Domestic tourism expenditures: The non- linear effects of length of stay and travel party size. *Tourism Management*, 32, 46-52.

Tourism Satellite Accounts (TSA) in Europe, *Eurostat*, 2010 edition

Vanhove, N (2011). *The economics of tourism destinations*, Elsevier Insights

Vance, C. and R. Iovanna (2007). Gender and the Automobile: An Analysis on Non-work Service Trips, *Transportation Research Record*, 2013, pp. 54-61.

Verbeek M. (2004). *A Guide to Modern Econometrics 2th ed*.: John Wiley & Sons.

Wooldridge, J.M. (2002). *Econometric analysis of cross section and panel data*, Cambridge: MIT Press

Wooldridge, J.M. (2010). *Econometric analysis of cross section and panel data*, Cambridge: MIT Press

Wynen, J. (2011). An estimation of the determinants of same- day visit expenditures in Belgium. *Tourism Economics* (forthcoming)

5 Tables and figures

Table 1 data representativeness

Indicator		% population	% sample
Province	Brussels	10.20	10.36
	Antwerp	16.13	14.77
	Vlaams-Brabant	9.93	14.04
	Waals-Brabant	3.49	2.33
	West- Flanders	10.65	13.24
	East- Flanders	13.19	14.53
	Henegouwen	12.02	6.57
	Luik	9.84	4.31
	Limburg	7.71	15.26
	Luxemburg	2.49	1.38
	Namen	4.36	3.22
Age	< 30	20.63	19.40
	30 - 49	41.08	37.92
	50+	38.28	42.68
Gender	Male	49.84	49.42
	Female	50.16	50.58
Education	Elementary school degree	37.34	31.27
	High school degree	35.26	35.95
	College or University degree	27.40	32.78

Table provided by iVOX

Table 2: descriptive statistics, medians

Month	Median km traveled	N
January	65	543
February	70	500
March	60	578
April	70	792
May	70	607
June	75	677
July	65	1640
August	73	1235
September	70	814
October	64	911
November	70	722
December	65	534
Jan- Dec	70	9553

Table 3 Descriptives

Name	Description	Median	Mean	Standard Deviation	Min	Max
Distance traveled (km)	Variable identifying distance traveled (expressed in KM)	70	88.76	68.84	1	500
Information	Categorical variable indicating information was gathered before going on a sameday visit. And if so which source was consulted. Benchmark value is internet(1). Other values represent brochures(2), other (3) and none (4).	3	2.97	0.397	1	4
Time spend (excluding travel time)	Variable indicating the time spend during the daytrip. excluding travel time (expressed in hours)	6	6.34	3.787	2	20
Female	Dummy indicating if respondent is female or not	0	0.505	0.499	0	1
Age	Variable indicating the age of the respondent (expressed in years)	47	45.75	14.809	18	85
Season	Categorical variable indicating the season when the daytrip occurred. Benchmark value is Low season (1), Mid-season is (2) and High season is (3).	1	1.65	0.7467	1	3
Retired	Dummy indicating whether or not the respondent is retired	0	0.034	0.181	0	1
Student	Dummy indicating whether or not the respondent is a student	0	0.048	0.215	0	1
Number of vehicles owned	variable indicating the number of vehicles owned	1	1.354	0.636	0	7
Married	Dummy indicating the gender if the respondent is married	1	0.632	0.482	0	1
Children	Dummy indicating whether or not the respondent has children	0	0.358	0.479	0	1
Education	Categorical variable indicating the level of education of the respondent. Highest degree is equal to elementary school. secondary school or college/ university degree. Degree from elementary school is used as a reference category	2	2.012	0.801	1	3
Age^2	squared term of age	2209	2312.698	1360.17	324	7225
Region	Categorical variable indicating where the respondent lives. 1= Flanders; 2= Wallonie and 3= Brussels	3	1.35	0.63	1	3
Expenditures	Variable indicating money spend during sameday visit	50	78.11	87.65	0	400

Table 4 One- way ANOVA distance/ expenditures

Distance	Mean expenditures
0-39	59.72
40-69	72.49
70-119	84.81
120-501	96.63

F- test equals 63.48 with Prob>F= 0.000

Table 5 Marginal effects of probit estimation, truncated regression and Heckman selection model

	Probit	Truncated regression	Heckman selection model	
	Model 1	Model 2	Model 3	Model 4
	mfx for Pr(y observed)	Coeff	mfx for Pr(y observed)	mfx for E(y | y observed)
	(Std. Error)	(Std. Error)	(Std. Error)	(Std. Error)
Brochures		67.776***		30.198***
		(7.715)		(3.09)
Other information sources		-16.487***		-5.5267***
		(5.412)		(1.91)
No information		-14.461***		-5.072***
		(5.456)		(1.93)
Time spent		9.159***		3.622***
		(0.550)		(0.193)
Age	0.012***	4.081***	0.016***	1.577**
	(0.002)	(0.978)	(0.003)	(0.346)
Agesq	-0.0001***	-0.0448***	-0.0001***	-0.017***
	(0.0001)	(0.01)	(0)	(0.00)
Female	0.01***	-20.029***	-0.044***	-7.261***
	(0.002)	(4.275)	(0.012)	(1.518)
Retired	-0.009	-6.7341	0.01	-2.125
	(0.006)	(12.513)	(0.036)	(4.421)
Student	-0.0358***	-20.668	-0.224***	-8.906*
	(0.006)	(13.522)	(0.033)	(4.634)
Children	-0.009***	0.687	-0.075***	-0.206
	(0.002)	(4.918)	(0.014)	(1.772)
Married	0.0047*	0.253	0.056***	0.547
	(0.003)	(5.361)	(0.015)	(1.914)
Mid season	0.0043**	4.103	0.014	
	(0.002)	(5.047)	(0.014)	
High season	0.135***	-7.344	0.537***	
	(0.004)	(4.857)	(0.015)	
Number of vehicles	-0.001	8.42***	0.002	3.187***
	(0.002)	(3.13)	(0.01)	(1.153)
High school degree	0.023***	-1.271	0.144***	0.401
	(0.002)	(5.382)	(0.015)	(1.961)
College or univerisity degree	0.042***	-0.089	0.243***	1.442
	(0.003)	(5.391)	(0.015)	(2.034)
Region dummies	Included	Included	Included	Included
Observations	93299	8642	93299	8642
Log Likelihood	-31921.205	-46830.814		
Inverse Mills ratio			8.574**	
Rho			0.89	

*** significant at 1% level. ** significant at 5% level. * significant at 10% level Standard errors are given in parentheses. Three region dummies (Flanders, Wallonia and Brussel) have been included.

Figure 1 Histogram of time spent during same-day visits

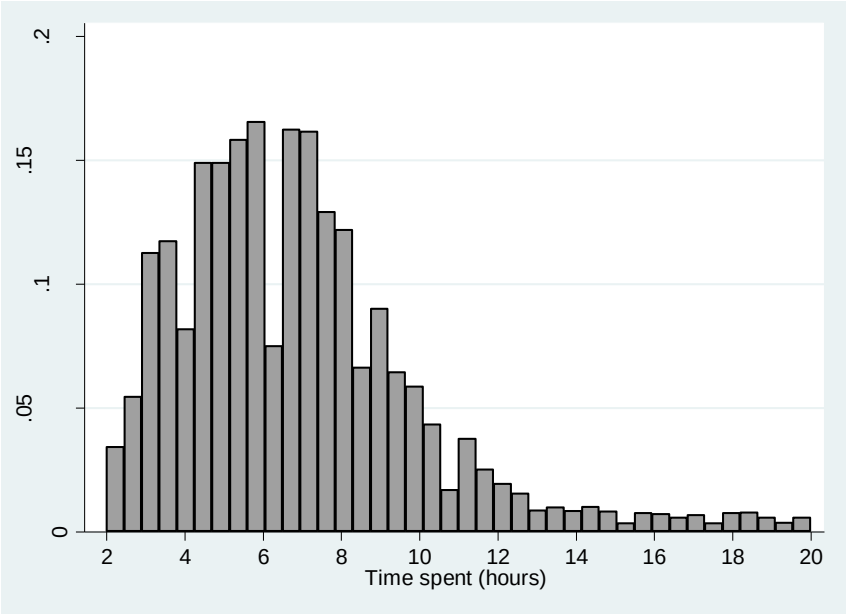

Figure 2 Histogram of information sources used to prepare same- day visits

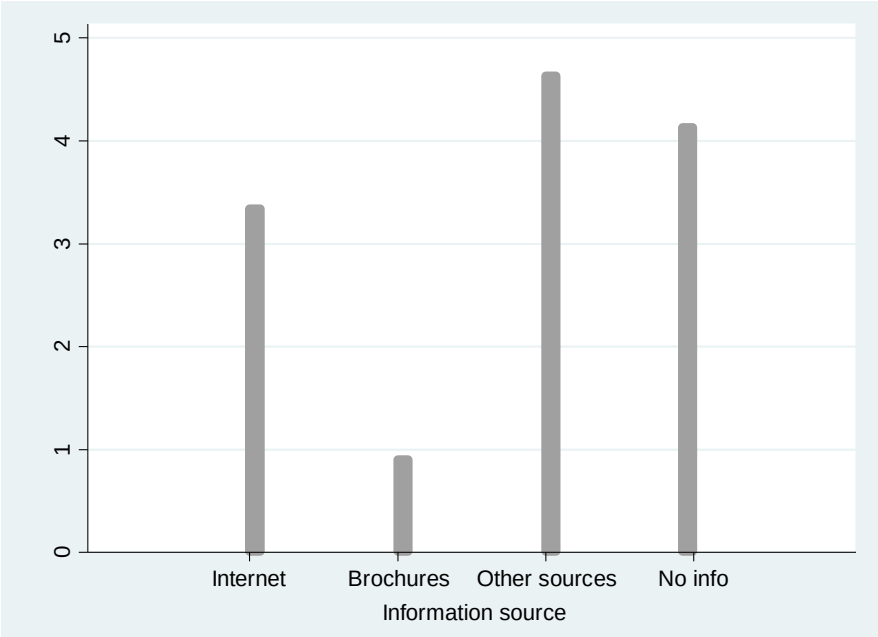

Figure 3 Effect of age on the probability of going on a same-day visit and on the distance travelled

