Ruhr Economic Papers

Manuel Frondel and Colin Vance

Driving for Fun? – A Comparison of Weekdays and Weekend Travel

#103

Ruhr Economic Papers

Published by

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstraße 12, 45117 Essen, Germany

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI)
Hohenzollernstr. 1/3, 45128 Essen, Germany

Ruhr Economic Papers #103

Responsible Editor: Christoph M. Schmidt

ISSN 1864-4872 (online) – ISBN 978-3-86788-114-2

Ruhr Economic Papers

#103

Manuel Frondel and Colin Vance

Driving for Fun? – A Comparison of Weekdays and Weekend Travel

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

ISSN 1864-4872 (online)
ISBN 978-3-86788-114-2

Manuel Frondel and Colin Vance*

Driving for Fun? – A Comparison of Weekdays and Weekend Travel

Abstract

Focusing on individual motorists in car-owning households in Germany, this paper econometrically investigates the determinants of automobile travel with the specific aim of quantifying the effects of fuel prices and person-level attributes on travel conducted over a five-day week and weekend. Our analysis is predicated on the notion that car use is an individual decision, albeit one that is dependent on intra-household allocation processes, thereby building on a growing body of literature that has identified the importance of socioeconomic factors such as employment status, gender, and the presence of children in determining both access to the car and distance driven. To capture this two-stage decision process, we employ the Two-Part Model, which consists of Probit and OLS estimators, and derive elasticity estimates that incorporate both the discrete and continuous choices pertaining to car use. With fuel price elasticity estimates ranging between –0.42 and –0.48, our results suggest raising prices via fuel taxes to be a promising energy conservation and climate protection measure.

JEL Classification: D13, Q41

Keywords: Automobile travel, Two-Part Model, interaction effects

April 2009

* Both RWI. – The authors are very grateful for invaluable comments and suggestions by Christoph M. Schmidt. – All correspondence to Manuel Frondel, RWI, Hohenzollernstr. 1-3, 45128 Essen, Germany, e-mail: frondel@ rwi-essen.de.

1 Introduction

As one of the dominant sources of energy use in industrialized countries, automobile travel is central to a multitude of issues that have relevance for economic policy and environmental stewardship. Although the automobile is indispensable to modern economic life, it is simultaneously the source of a range of acute environmental stressors, including noise, urban air pollution, and overall climate change. These impacts have been particularly pronounced in the European Union, where greenhouse gases from domestic transport increased by 26 % between 1990 and 2005 (EEA, 2007). As the main driver of this increase was passenger car mileage, with its near complete dependence on oil products, a critical question confronting European policy-makers is the extent to which motorists adjust driving behavior in response to changes in fuel prices and other socioeconomic characteristics.

Notwithstanding an extensive corpus of research spanning over three decades, this question continues to occasion a great deal of debate within both the academic and policy realms. Based on a comprehensive survey of the related literature, GRAHAM and GLAISTER (2002) cite fuel-price elasticities in the region of -0.3 for the short run and -0.8 in the long run. Despite a substantial degree of variation both within and across geographic areas of study, they conclude that there is remarkably consistent evidence on a significant response to fuel prices, at least in the long run.

Over the short run, however, some more recent studies have found evidence for a small price elasticity of gasoline demand. HUGHES et al. (2006), for example, cite very low short-run fuel price elasticities ranging between -0.03 and -0.08 using aggregate monthly data from the U.S. Similarly low effects are estimated by KAYSER (2000), who concludes on the basis of household data that gasoline taxes are unlikely to cause large decreases in gasoline consumption. Based on a meta-analysis of elasticity estimates, BRONS et al. (2008) echo this conclusion, and suggest that policies to improve fuel economy may serve as an effective complement to taxation.

To date, the majority of empirical attempts to estimate price effects have drawn on country-level data or data aggregated at sub-national administrative districts, typically from the U.S., with a smaller pool of studies relying on household-level data. Departing from this reliance, the empirical analysis pursued in the present paper is predicated on the notion that car use is an individual decision, albeit one that is dependent on the household's allocation of resources and responsibilities among members. Our analysis uses data from the German Mobility Panel (MOP 2009), which includes detailed person-level records on driving and allows us to distinguish between weekday and weekend travel.

This focus on individual travel behavior raises an important but subtle conceptual issue emerging from the fact that some potential motorists choose not to use the car over a particular week, and whose recorded driving is therefore censored at zero. If ignored, the presence of these null values in the data is shown to potentially result in spurious conclusions with respect to both the magnitude and the significance of the estimates. To empirically accommodate such "corner solutions", we employ the Two-Part Model, which consists of both Probit and OLS estimations, and include a suite of individual characteristics in the empirical specification. In interpreting the results, elasticity estimates of all explanatory variables are derived that incorporate both the discrete and continuous decisions pertaining to car use.

Although we are aware of no other study that estimates fuel-price elasticities using data on individual motorists, this tack is in line with a growing body of literature that has identified the importance of socioeconomic factors such as employment status, gender, and the presence of children in determining access to the car, distance driven, and other aspects of mobility behavior (e.g. PICKUP (1985), TURNER and NIEMEIER (1997), KAYSER (2000), VANCE and HEDEL (2007)).

A further distinguishing feature of our analysis pertains to the tight temporal correspondence between our measures of travel and fuel prices. While the majority of studies use annual data on vehicle travel, and match this with fuel prices that are averaged over the year, the present study focuses on travel conducted over a particular five-day week and weekend, and matches these time intervals with prevailing fuel prices. This not only allows us to differentiate estimated effects by work and non-work days, it also enables us to estimate the lower bound of the short-run elasticity, as we can effectively measure the immediate influence of prices on driving, with limited leeway for other behavioral adjustments.

With fuel price elasticity estimates ranging between -0.42 and -0.48, our results indicate price effects that are substantially larger than those typically obtained from U. S.-based studies, but slightly lower than the range of -0.57 to -0.67 identified by FRONDEL, PETERS, and VANCE (2008). These authors employ another subset of the German Mobility Panel to investigate mobility behavior, but at the household rather than the individual level. Given the magnitude of the estimates, our results suggest fuel taxes to be a promising energy conservation and climate protection measure.

The following section describes the econometric methods and models specified for estimating individual mobility behavior. Section 3 describes the data base used in the estimation, followed by the presentation and interpretation of the results in Section 4. The last section summarizes and concludes.

2 Methodology

The reliance on individual data over a tightly circumscribed time interval raises several conceptual and empirical issues, the most fundamental of which is the presence of null values in the data. Roughly 15% of the observed individuals do not use the car during a given week and for whom the observation on distance driven is consequently recorded as zero. To accommodate this feature of the data, which is even more pronounced for the corresponding weekend (about 34% zeros), we employ a two-stage modeling procedure referred to as the Two-Part Model (2PM) that orders observations into two regimes defined by whether the individual uses the car as a driver.

2.1 The Two-Part Model

The first stage defines a dichotomous variable indicating the regime into which the observation falls:

$$S = 1, \text{if} \quad S^* = \mathbf{x_1}^T \boldsymbol{\tau} + \boldsymbol{\varepsilon}_1 > 0 \quad \text{and} \quad S = 0, \text{if} \quad S^* \leq 0. \tag{1}$$

where S^* is a latent variable indicating the utility from car use, S is an indicator for car usage status, $\mathbf{x_1}$ includes the determinants of this status, $\boldsymbol{\tau}$ is a conformable vector of associated parameter estimates, and $\boldsymbol{\varepsilon}_1$ is an error term drawn from a standard normal distribution.

In addition to estimating τ using classical probit maximum likelihood methods, the second stage involves estimating the parameters $\boldsymbol{\beta}$ via an OLS regression conditional on car use, $S = 1$:

$$E[y|S = 1, \mathbf{x_2}] = \mathbf{x_2}^T \boldsymbol{\beta} + E(\varepsilon_2 | y > 0, \mathbf{x_2}) = \mathbf{x_2}^T \boldsymbol{\beta}, \tag{2}$$

where y is the dependent variable, measured here either as the kilometers of vehicle travel or fuel consumption over either the week or weekend, and $\boldsymbol{\varepsilon}_2$ is the error term, again assumed to be normally distributed and for which $E(\varepsilon_2 | y > 0, \mathbf{x_2}) = 0$.

The structure of the 2PM is similar to HECKMAN's two-stage sample selection model, frequently called the Heckit model, with the key distinction being the inclusion of an additional regressor - the inverse Mills ratio (IVM) - in the second stage regression of the Heckit to control for potential selectivity bias. The relative merits of the two models have been the subject of a vigorous debate in the literature (HAY and OLSON, 1984; DUAN et al. 1984; LUENG and YU 1996; DOW and NORTON, 2003), with much of the discussion focusing on their underlying assumptions and numerical properties.

Two considerations led us to select the 2PM as the superior alternative for this analysis. First, a well-known impediment in estimating the HECKMAN model emerges when there is a high degree of collinearity between the independent variables and the IVM, resulting in high standard errors on the coefficient estimates and parameter instability. Although this problem can be attenuated by the inclusion of identifying variables that uniquely determine the discrete outcome, no such variables immediately avail themselves in the present data, a common problem that forces reliance on functional form assumptions for model identification. As discussed by DOW and NORTON (2003), a second, more substantive, consideration in choosing between the two models is whether interest centers on the actual or potential outcome of the phenomena under study.

In the present context, the potential outcome y^* addresses the distance an individual would drive were he or she to use the car, irrespective of actual use, while the actual outcome y addresses the observed distance driven, equaling zero if the car was not used ($y = 0$). Whereas the actual outcome y is a fully-observed variable, the potential outcome y^* is a latent variable that is only partially observed, namely for those who have chosen to use the car: $y^* = y$ if $y > 0$, but y^* is unidentified if $y = 0$, i. e. for those who have refrained from car use.

While the Heckit estimator was designed to address selection bias for analyzing potential outcomes, it incorporates features that make it often perform worse than the 2PM when analyzing actual outcomes (DOW and NORTON 2003:6). Accordingly, the 2PM is deemed here the more appropriate modeling specification to estimate the effect of fuel prices and individual socioeconomic traits on *actual* distance driven or *actual* fuel consumption.Because these traits are recorded only one time in the data dictates a pooled regression approach, unlike FRONDEL, PETERS, and VANCE (2007), who use panel estimators on data aggregated at the household level.

2.2 Calculation of Elasticities

For estimating the marginal effects of socioeconomic determinants on *actual* distances or *actual* fuel consumption, it is necessary to take account of the likelihood that a household refrains from using a car, $P(y = 0)$. Hence, the prediction of the dependent variable consists of *two parts*, with the first part being the probability of owning the car,

$P(y > 0) = \Phi(\mathbf{x_1}^T \boldsymbol{\tau})$, which results from the first stage (1) of the 2PM, and the second part being the conditional expectation $E[y|y > 0] = \mathbf{x_2}^T \boldsymbol{\beta}$ from the second stage (2):

$$\begin{aligned} E[y] &= P(y > 0) \cdot E[y|y > 0] + P(y = 0) \cdot E[y|y = 0] \\ &= P(y > 0) \cdot E[y|y > 0] + 0 = \Phi(\mathbf{x_1}^T \boldsymbol{\tau}) \cdot \mathbf{x_2}^T \boldsymbol{\beta}. \end{aligned} \tag{3}$$

As our interest centers on elasticities, we now present the required formulae for the corresponding 2PM with a logged dependent variable $z = \ln(y)$ and normal homoskedastic errors $\boldsymbol{\varepsilon}_2$ with constant variance $\mathrm{Var}(\boldsymbol{\varepsilon}_2) = \sigma^2$, following DOW and NORTON (2003:11). Rather than by (3), actual outcomes are in this case predicted by[1] :

$$E[y] = \Phi(\mathbf{x_1}^T \boldsymbol{\tau}) \cdot \exp\{\mathbf{x_2}^T \boldsymbol{\beta} + 0.5 \cdot \sigma^2\}. \tag{4}$$

Using the product and chain rules of differentiation and the fact that the derivative of the cumulative normal function Φ equals the normal density function ϕ, the marginal effect can be derived as follows:

$$\begin{aligned} \frac{\partial E[y]}{\partial x_k} &= \beta_k \cdot E[y] + \tau_k \cdot \phi(\mathbf{x_1}^T \boldsymbol{\tau}) \cdot \exp\{\mathbf{x_2}^T \boldsymbol{\beta} + 0.5 \cdot \sigma^2\} \\ &= \beta_k \cdot E[y] + \tau_k \cdot \frac{\phi(\mathbf{x_1}^T \boldsymbol{\tau})}{\Phi(\mathbf{x_1}^T \boldsymbol{\tau})} \cdot E[y] \end{aligned} \tag{5}$$

By dividing expression (5) by $E(y)$ and multiplying it by x_k, the respective elasticity can be obtained:

$$\eta_{x_j} = \frac{\partial \ln E[y]}{\partial \ln x_j} = \frac{\partial E[y]}{\partial x_j} \cdot \frac{x_j}{E[y]} = [\beta_j + \tau_j \cdot \frac{\phi(\mathbf{x_1}^T \boldsymbol{\tau})}{\Phi(\mathbf{x_1}^T \boldsymbol{\tau})}] \cdot x_j \tag{6}$$

For logged explanatory variables ($z_k = \ln x_k$), the elasticity formula follows from (5) by dividing it by $E(y)$:

$$\eta_{x_k} = \frac{\partial \ln E[y]}{\partial \ln x_k} = \frac{\partial E[y]}{\partial \ln x_k} \cdot \frac{1}{E[y]} = \beta_k + \tau_k \cdot \frac{\phi(\mathbf{x_1}^T \boldsymbol{\tau})}{\Phi(\mathbf{x_1}^T \boldsymbol{\tau})}. \tag{7}$$

To handle the case of dummy variables D_k, we employ the following relative differences, thereby using the formula of the expected value (4):

$$(E[y|D_k = 1] - E[y|D_k = 0])/E[y]. \tag{8}$$

[1] If $z = ln(y)$ has a normal distribution with an expected value of $E(z) = \mu$ and variance σ^2, then y has a lognormal distribution and an expected value of $E(y) = \exp\{\mu + 0.5 \cdot \sigma^2\}$.

It is considerably more cumbersome to calculate the interaction effects of two explanatory variables in non-linear models such as the 2PM. Standard computer software commonly calculates the first derivative

$$\frac{\partial E[y]}{\partial (x_1 x_2)} \tag{9}$$

to provide for estimates of the interaction effects between two continuous variables x_1 and x_2, where the product $z = x_1 \cdot x_2$ is the so-called interaction term that is typically incorporated in linear specifications to capture interaction effects.

AI and NORTON (2003) argue, however, that in non-linear models the calculation of the interaction effects requires computing the cross-derivative

$$\frac{\partial^2 E[y]}{\partial x_1 \partial x_2} \tag{10}$$

and show for the case of non-linear models such as logit and probit that the calculation based on (9) generally results in false inferences with respect to both the sign and significance of the interaction effect. Consequently, we follow their recommendation to calculate the interaction effects as given by (10) and present the derivations for both the 2PM and probit models in the appendix.

A final complication concerns calculating the statistical significance of the elasticity estimates, given that they are comprised of multiple parameters that makes analytical computation of the variance impossible. We circumvent this difficulty by applying the Delta method, which uses a first-order Taylor expansion to create a linear approximation of a non-linear function, after which the variance and measures of statistical significance can be computed.

2.3 Model Specification

The model specification employed here is based on the logged version (4) of the 2PM and uses logged liters of fuel consumed, $\ln(e)$, as dependent variable:

$$E[e] = \Phi(\tau_{p_e} \ln(p_e) + \mathbf{x}^T \boldsymbol{\tau}) \cdot \exp\{\beta_{p_e} \ln(p_e) + \mathbf{x}^T \boldsymbol{\beta} + 0.5 \cdot \sigma^2\}, \tag{11}$$

where the set of explanatory variables includes the logged price of fuel per liter, $\ln(p_e)$. The remaining suite of variables measure the individual, household, and automobile attributes that are hypothesized to influence the extent of motorized travel. Variable definitions and descriptive statistics are presented in Table 1.

To control for the effects of quality, the age of the automobile and a dummy indicating premium models (i.e. sports and luxury cars) is included. Although income is not directly measured, an attempt is made to proxy for its influence via measures of the number of employed residents and the number with a high school diploma living in the household. We also include a measure of the number of children to capture demographic pressures. Four variables are included to control for the effects of urban density and the availability of alternative transportation: dummies indicating residence in a large city and whether the household has a private parking space, a continuous variable measuring the walking time to the nearest public transportation stop, and a dummy indicating whether this stop is serviced by rail transit (as opposed to bus).

To measure the influence of individual attributes, we include age and three dummies indicating persons who are high-school educated[2], employed, and female. Finally, we interact the female dummy with the variables measuring the number of children, the number of employed in the household, and the individual employment dummy. These interactions are intended to capture the role played by household responsibilities, social status, and competition among household members in dictating access to and use of the car. We also explored models in which time dummies were included to control for autonomous changes in the macroeconomic environment. As these were found to be jointly insignificant across all of the models estimated, they were excluded from the final specifications.

[2]We limit the definition of high-school educated to those who have completed a college preparatory degree.

3 Data

The data used in this research are drawn from the German Mobility Panel (MOP 2007), an ongoing travel survey financed by the German Federal Ministry of Transport, Building and Housing. We use ten years of data from the survey, spanning 1996 through 2005, a period during which real fuel prices in Germany rose 2.8% per annum. Households that participate in the survey are requested to fill out a questionnaire eliciting general household information and person-related characteristics, including gender, age and employment status. Moreover, all household members over 10 years of age fill out a trip log capturing relevant aspects of everyday travel behavior, including distances traveled, modes used, activities undertaken, and activity durations. We use the data from the trip logs to construct the total amount of fuel consumed in liters by the individual over the course of a week and weekend, which serves as the dependent variables.

In addition to the general survey, the MOP includes another survey focusing specifically on vehicle travel among a sub-sample of randomly selected car-owning households. This survey takes place over a roughly six-week period in the following spring, during which time respondents record the price paid for fuel with each visit to the gas station, as well as vehicle attributes such as fuel economy, fuel type, and the car model. Using a household identifying variable, we merged this data with the trip logs to create a data set that includes both person-level travel information as well as vehicle attributes and prices paid for fuel.This linkage results in a tight temporal correspondence between our measures of travel and prevailing fuel prices, and contrasts with the more standard practice of linking annual data on vehicle travel with fuel prices that are averaged over the year. We consequently interpret our estimates as representing short-run elasticities, as we can effectively measure the immediate influence of prices on driving.

As the travel log survey contains no identifier for the vehicle used for particular trips, we focus on single car households to avoid matching problems, which comprise roughly 53% of the households in Germany. (Of the remaining 47% of German house-

Table 1: Variable Definitions and Descriptive Statistics

Variable Definition	Variable Name	Mean	Std. Dev.
Kilometers driven over 5-day week	s_5	103.13	164.79
Kilometers driven over weekend	s_w	39.34	86.26
Kilometers driven per liter	μ	12.33	2.77
Real fuel price in € per liter	p_e	0.97	0.14
Age of the car	*car age*	6.14	4.19
Dummy: 1 if car is a sports- or luxury model	*premium car*	0.23	0.42
Dummy: 1 if person has a high school diploma	*high school diploma*	0.28	0.45
Age of the person	*age*	47.90	18.38
Dummy: 1 if person is employed in a full-time or part-time job	*employed*	0.44	0.50
Dummy: 1 if person is female	*female*	0.50	0.50
Number of employed household members	*# employed*	0.93	0.82
Number of household members with a high school diploma	*# high school diploma*	0.54	0.74
Number of children younger than 17	*children*	0.60	0.95
Dummy: 1 if household resides in a large city	*big city*	0.41	0.49
Walking time to the nearest public transportation stop	*minutes*	5.68	5.21
Dummy: 1 if household has a private parking space or garage	*private parking*	0.83	0.37
Dummy: 1 if the nearest public transportation stop is serviced by rail transit	*rail transit*	0.12	0.32

holds, 27% have more than one car and 20% have no car (MiD 2007).) The analysis is further limited to household members who possess a driver's license, which requires a minimum age of 18 years. The resulting data set yields a total of 3,031 observations.

4 Empirical Results

While distinguishing between weekday and weekend travel, our empirical analysis is predicated on the notion that both car use or access and distance traveled are individual decisions that are affected by a variety of socioeconomic factors, such as employment status, gender, and the presence of children. Although necessarily neglected in studies of fuel prices using aggregated data, the analysis of the discrete decision to use the car appears to be of particular relevance, as fuel price peaks may trigger a reduction in both the distanced traveled and the frequency of car use.

In the elasticity formulae (6) and (7), the decision on car use is captured by the additional term $\tau_j \cdot \frac{\phi(\mathbf{x_1}^T \boldsymbol{\mathcal{T}})}{\Phi(\mathbf{x_1}^T \boldsymbol{\mathcal{T}})}$. Generally, this term only vanishes if $\tau_j = 0$, that is, if variable x_j does not impact the decision on car use, so that the effect of x_j collapses to that on distance traveled, as is reflected by the coefficient β_j. In the case of fuel prices, however, it may well be expected that both coefficients, β_j and τ_j, are non-vanishing and negative. Consequently, fuel price effects would be under-estimated if the car use decision were to be deemed irrelevant and, hence, ignored.

In fact, it turns out from our first-stage Probit models that fuel prices are important determinants of car use both over the 5-day week and on the weekends (Table 2). Consistent with intuition, the marginal effect of fuel prices is slightly higher on the weekend, though the difference is not statistically significant. More stark differences are seen with respect to the variables measuring the quality of local public transit and the competition for the car, as captured by the rail transit dummy and the number of employed household members. These variables significantly reduce the probability of using the car in the weekday model only, whereas the number of children and the age of the individual only decreases this probability on the weekends.

Moreover, there are substantive gender differences with respect to car use, as is evidenced by the negative signs of the female dummy coefficient estimates, indicating that women have less access to the car than men, as well as by the statistical significance of several interaction effects: For instance, the Probit results reported in Table 2 point to a higher need of access to the car of women that are employed. It bears noting that these interaction effects are estimated using the formulae (16) and (18) given in the appendix.

Table 2: Probit Estimation Results of the Car Use Decision

	Over 5-Days Week				Weekend			
	Coeff.s	Errors	Mar. Effects	Errors	Coeff.s	Errors	Mar. Effects	Errors
$\ln(p_e)$	** -0.688	(0.230)	**-0.141	(0.067)	**-0,588	(0.186)	**-0.213	(0.047)
car age	0.007	(0.009)	0.002	(0.002)	0,000	(0.007)	0.000	(0.002)
premium car	-0.009	(0.089)	-0.002	(0.025)	-0,097	(0.067)	-0.036	(0.018)
high school diploma	** 0.581	(0.128)	** 0.107	(0.035)	** 0,368	(0.103)	** 0.129	(0.021)
age	0.000	(0.003)	0.000	(0.001)	-0,006	(0.003)	*-0.002	(0.001)
employed	** 0.701	(0.182)	** 0.146	(0.050)	** 0,382	(0.139)	** 0.138	(0.038)
female	** -0.718	(0.107)	**-0.151	(0.031)	**-0,662	(0.088)	** -0.238	(0.023)
# employed	** -0.320	(0.099)	**-0.066	(0.029)	-0,084	(0.081)	-0.030	(0.020)
# high school diploma	** -0.469	(0.074)	**-0.096	(0.022)	**-0,267	(0.062)	**-0.097	(0.015)
# children < 18	-0.072	(0.069)	-0.015	(0.018)	**-0,113	(0.051)	* -0.041	(0.014)
female$\times$ (*# employed*)	-0.108	(0.136)	-0.096	(0.039)	* -0,382	(0.114)	**-0.160	(0.044)
female$\times$ *employed*	0.280	(0.234)	** 0.168	(0.058)	** 0,421	(0.184)	** 0.183	(0.037)
female$\times$ (*# children<18*)	** 0.383	(0.095)	** 0.101	(0.026)	** 0,332	(0.070)	** 0.124	(0.022)
big city	-0.113	(0.075)	-0.023	(0.021)	-0,026	(0.059)	-0.010	(0.016)
minutes	0.012	(0.007)	0.002	(0.002)	0,007	(0.005)	0.003	(0.001)
private parking	0.042	(0.092)	0.009	(0.028)	0,019	(0.076)	0.007	(0.020)
rail transit	*-0.240	(0.100)	*-0.054	(0.032)	-0,098	(0.085)	-0.036	(0.025)

Note: * denotes significance at the 5 %-level and ** at the 1 %-level, respectively. Number of observations used for estimation: 3,031.

Table 3 presents both the coefficient estimates of the second-stage OLS regression and the associated elasticities representing the effects of the variables on the actual outcome, which incorporate the estimations results of both the first-stage Probit estimation and the second-stage OLS regression. Several features of the results for weekday travel presented in the left panel of Table 3 bear highlighting. First, we note the evident differences, both with respect to magnitude and precision, between the OLS coefficient estimates and the elasticities. For $\ln(p_e)$, for example, the OLS coefficient estimate is insignificant, while the estimate of the corresponding elasticity is of a notably larger magnitude and significant at the 5% level, suggesting a fuel price elasticity of -0.42.

Conversely, while the OLS coefficient estimate of children is highly significant, this effect fades away when expressed as an elasticity. In interpreting these discrepancies, it bears recalling that, unlike the elasticities, the unadjusted OLS coefficient estimates take no account of the decision on car use and its probability. The different conclusions arising from the OLS and 2PM estimates stress the importance of incorporating this discrete decision into the analysis and correctly dealing with a censored dependent variable.

Turning to the remaining coefficients in the left panel, all have either intuitive effects or are statistically insignificant. Individuals who are high-school-educated and employed drive more than their counterparts, while older individuals drive less. Females drive less than men, but this negative effect is significantly mitigated by the presence of children in the household, as evidenced by the positive interaction effect. This is a likely reflection of the pick-up and delivery services associated with child care, which tend to be borne by women. The number of other high-school educated and employed persons in the household both negatively impact distance driven, which is a likely consequence of competition for the car.

The right panel of Table 3 presents the coefficient estimates and elasticities for weekend travel. The fuel-price elasticity is slightly higher than for weekday travel, though imprecision in the estimates makes it difficult to draw definitive conclusions with respect to magnitude. This is in line with the intuition that motorists are more responsive to higher fuel prices on non-work days, when their travel behavior is presumably more flexible.

Table 3: Estimation Results for the Second-Stage OLS Model and the Two-Part Model (2PM)

	Over 5-Days Week				Over Weekend			
	OLS		2PM		OLS		2PM	
$\ln(e)$	Coeff.s	Errors	Elast.s	Errors	Coeff.s	Errors	Elast.s	Errors
$\ln(p_e)$	-0.261	(.178)	* -0.423	(.189)	-.165	(.203)	* -.483	(.231)
car age	-0.004	(.006)	-0.015	(.040)	** -.020	(.007)	* -.121	(.050)
premium car	0.104	(.064)	0.105	(.073)	**.252	(.078)	* .209	(.099)
high school diploma	** 0.314	(.108)	** 0.466	(.132)	.054	(.115)	.250	(.137)
age	**-0.009	(.003)	**-0.460	(.136)	-.005	(.003)	* -.422	(.172)
employed	** 0.501	(.129)	** 0.824	(.100)	* .320	(.141)	** .683	(.122)
female	**-0.523	(.085)	** -0.577	(.061)	** -.478	(.106)	** -.887	(.075)
# employed	-0.108	(.076)	** -0.265	(.059)	-.144	(.080)	** -.298	(.072)
# high school diploma	* -0.157	(.075)	** -0.148	(.045)	.040	(.081)	-.058	(.050)
# children < 18	**-0.172	(.051)	-0.019	(.018)	-.007	(.048)	-.013	(.022)
female$\times$ (*# employed*)	-0.203	(.115)	-0.067	(.117)	-.084	(.134)	-.053	(.130)
female$\times$ *employed*	0.258	(.173)	0.046	(.262)	.105	(.201)	.006	(.302)
female$\times$ (*# children<18*)	** 0.226	(.065)	** 0.351	(.074)	-.094	(.073)	.112	(.086)
big city	0.034	(.054)	0.007	(.058)	.109	(.062)	.095	(.072)
minutes	0.008	(.005)	* 0.059	(.031)	-.006	(.005)	-.012	(.032)
private parking	-0.002	(.065)	-0.002	(.062)	-.051	(.079)	-.050	(.078)
rail transit	-0.120	(.080)	*-0.172	(.077)	.091	(.101)	.036	(.116)
constants	** 6.974	(.186)	–	–	** 6.066	(.206)	–	–
# observations used for estimation:		2,568				1,992		

Note: * denotes significance at the 5 %-level and ** at the 1 %-level, respectively.

Several other results are also similar across the models for weekday and weekend travel: Being employed increases car use even on weekends, albeit with a lower magnitude, while age, the number of employed in the household, and the female gender dummy all have negative effects. Conversely, discrepancies emerge with respect to the effects of the age of the car, whether it is a premium model, the interaction of female and children, high-school education, and the rail transit dummy. While the age of cars seems to play no role during the weekdays, it is significant and negative for weekend travel, possibly reflecting a reduced proclivity to drive older cars for recreational activities. Along similar lines, the premium car dummy is positive and significant only in the model for weekend travel, again reflecting the role of driving utility in determining car use.

5 Summary and Conclusion

This paper has estimated the determinants of weekly driving behavior using individual-level data collected in Germany over a ten-year period spanning 1997 to 2005. The focus on person-level data is argued to necessitate the application of a two-part modeling procedure distinguishing between the discrete choice of car use and the continuous choice of distance traveled. This distinction is found to have an important bearing on the interpretation of the results; if ignored, several spurious conclusions emerge with respect to both the statistical significance and magnitude of the impact of individual variables.

In the case of the fuel price, for example, we find its effect to be statistically significant only when referencing the elasticities of the Two-Part Model. We draw two conclusions from these findings. Methodologically, they suggest that consideration of the determinants of both car use and distance traveled is warranted when estimating fuel price elasticities using person-level data over a short time interval. Substantively, the magnitude of the elasticity estimates, which are ascribed a short-run interpretation and range between -0.42 and -0.48, is sufficiently high to cast skepticism on recent pronouncements, both in academic papers and in European policy documents (e.g. COM 2007), that question fuel taxes as a means to reduce fuel consumption.

In addition to distinguishing between the decisions of whether and how much to use the car, the analysis also differentiates between weekday and weekend travel. Consistent with intuition, the estimated fuel price elasticity is slightly higher on the weekend. More stark differences are seen with respect to the variables measuring the age of the car and whether it is a premium model, which are significant only in the weekend model. By contrast, the variables measuring the number of highly educated household members, the effect of children for female members, and the proximity and quality of local public transit, are significant in the weekday model only. Broadly speaking, these differences likely reflect the different activity patterns during the the week. Car attributes, for example, matter more on the weekend when recreational travel is undertaken, whereas the variables measuring competition for the car and the allocation of household responsibilities matter more on weekdays.

Appendix: Interaction Effects

To provide a general derivation of interaction effects in both linear and non-linear models, we closely follow NORTON, WANG, and AI (2004) and begin by drawing on the following linear specification of the expected value of dependent variable y:

$$E[y|x_1, x_2, \mathbf{x}] = \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \mathbf{x}^T \boldsymbol{\beta}, \tag{12}$$

where the parameters $\beta_1, \beta_2, \beta_{12}$, as well as the vector $\boldsymbol{\beta}$ are unknown and vector $\mathbf{x}$ is independent of x_1 and x_2. Our discussion below focuses on those continuous- and dummy variable interactions found in the text, but also provides a foundation for generalization to other cases.

Linear Models

Assuming that x_1 and x_2 are continuous variables, the marginal effect of x_1 on $E = E[y|x_1, x_2, \mathbf{x}]$ is dependent on x_2 if $\beta_{12} \neq 0$:

$$\frac{\partial E}{\partial x_1} = \beta_1 + \beta_{12} x_2. \tag{13}$$

The impact of a marginal change in x_2 on the marginal effect of x_1, in other words the interaction effect, is then obtained from taking the derivative of (13) with respect to x_2:

$$\frac{\partial^2 E}{\partial x_2 \partial x_1} = \beta_{12}. \tag{14}$$

That is, in linear specifications, the interaction effect $\frac{\partial^2 E}{\partial x_2 \partial x_1}$ equals the marginal effect $\frac{\partial E}{\partial (x_1 x_2)} = \beta_{12}$ of the interaction term $z = x_1 x_2$. For non-linear models such as probit, however, this equality generally does not hold.

Probit Model

Instead of expectation (12), we now depart from the expected value

$$E[y|x_1, x_2, \mathbf{x}] = F(\beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \mathbf{x}^T \boldsymbol{\beta}) = F(u), \tag{15}$$

where $F(u)$ is a non-linear function of its argument $u := \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \mathbf{x}^T \boldsymbol{\beta}$. For the example of the probit model, $F(u)$ equals the cumulative normal distribution $\Phi(u)$. We now derive formulae for the interaction effects resulting from the probit model if (1) x_1 and x_2 are both continuous variables and (2) both are dummy variables.

(1) With $F(u) = \Phi(u)$ and the first derivative $\Phi'(u)$ being the density function of the standard normal distribution, $\Phi'(u) := \phi(u) = \exp\{-u^2/2\}/\sqrt{2\pi}$, the interaction effect of two continuous variables x_1 and x_2 is given by:

$$\begin{aligned} \frac{\partial^2 E}{\partial x_2 \partial x_1} &= \frac{\partial}{\partial x_2}(\frac{\partial E}{\partial x_1}) = \frac{\partial}{\partial x_2}(\frac{\partial \Phi(u)}{\partial x_1}) = \frac{\partial}{\partial x_2}(\phi(u)\frac{\partial u}{\partial x_1}) = \frac{\partial}{\partial x_2}[\phi(u)(\beta_1 + \beta_{12} x_2)] \\ &= \phi(u)\beta_{12} + (\beta_1 + \beta_{12} x_2)(\beta_2 + \beta_{12} x_1)\phi'(u), \end{aligned} \tag{16}$$

where $\phi' = -u\phi(u)$.

(2) If x_1 and x_2 are dummy variables, the discrete interaction effect, which in analogy to $\frac{\partial^2 E}{\partial x_2 \partial x_1}$ is designated by $\frac{\Delta^2 E}{\Delta x_2 \Delta x_1}$, is given by the discrete change in $E[y]$ due to a successive unitary change in both x_1 and x_2, $\Delta x_1 = 1, \Delta x_2 = 1$:

$$\begin{aligned} \frac{\Delta^2 E}{\Delta x_2 \Delta x_1} &:= \frac{\Delta}{\Delta x_2}(\frac{\Delta E}{\Delta x_1}) = \frac{\Delta}{\Delta x_2}(E[y|x_1 = 1, x_2, \mathbf{x}] - E[y|x_1 = 0, x_2, \mathbf{x}]) \\ &= \{E[y|x_1 = 1, x_2 = 1, \mathbf{x}] - E[y|x_1 = 0, x_2 = 1, \mathbf{x}]\} \\ &\quad -\{E[y|x_1 = 1, x_2 = 0, \mathbf{x}] - E[y|x_1 = 0, x_2 = 0, \mathbf{x}]\}. \end{aligned} \tag{17}$$

For $E[y|x_1, x_2, \mathbf{x}] = \Phi(\beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \mathbf{x}^T \boldsymbol{\beta})$, the general expression (17) translates into:

$$\frac{\Delta^2 E}{\Delta x_2 \Delta x_1} = \{\Phi(\beta_1 + \beta_2 + \beta_{12} + \mathbf{x}^T \boldsymbol{\beta})) - \Phi(\beta_2 + \mathbf{x}^T \boldsymbol{\beta})\} - \{\Phi(\beta_1 + \mathbf{x}^T \boldsymbol{\beta}) + \Phi(\mathbf{x}^T \boldsymbol{\beta})\}. \tag{18}$$

Two-Part Model

Based on the 2PM with a logged dependent variable and normal homoskedastic errors, and using a slightly modified version of expectation (4),

$$E[y] = \Phi(\tau_1 x_1 + \tau_2 x_2 + \tau_{12} x_1 x_2 + \mathbf{x}^T \boldsymbol{\tau}) \cdot \exp\{\beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \mathbf{x}^T \boldsymbol{\beta} + 0.5 \cdot \sigma^2\}, \tag{19}$$

in order to highlight the interaction terms, we now derive formulae for the interaction effects and elasticities if (1) both x_1 and x_2 are dummy variables and (2) x_1 is continuous, while x_2 is a dummy variable.

(1) Departing from (19), the interaction effect between two dummy variables x_1 and x_2 is to be derived as follows:

$$\begin{aligned}
\frac{\Delta^2 E[y]}{\Delta x_1 \Delta x_2} &= \frac{\Delta}{\Delta x_2}(\frac{\Delta E[y]}{\Delta x_1}) = \frac{\Delta}{\Delta x_2}(E[y|x_1 = 1, x_2, \mathbf{x}] - E[y|x_1 = 0, x_2, \mathbf{x}]) \\
&= \{[E[y|x_1 = 1, x_2 = 1, \mathbf{x}] - E[y|x_1 = 0, x_2 = 1, \mathbf{x}]\} \\
&\quad -\{[E[y|x_1 = 1, x_2 = 0, \mathbf{x}] - E[y|x_1 = 0, x_2 = 0, \mathbf{x}]\} \\
&= \Phi(\tau_1 + \tau_2 + \tau_{12} + \mathbf{x}^T\tau) \cdot \exp\{\beta_1 + \beta_2 + \beta_{12} + \mathbf{x}^T\beta + 0.5 \cdot \sigma^2\} \\
&\quad -\Phi(\tau_2 + \mathbf{x}^T\boldsymbol{\tau}) \cdot \exp\{\beta_2 + \mathbf{x}^T\boldsymbol{\beta} + 0.5 \cdot \sigma^2\} \\
&\quad -\Phi(\tau_1 + \mathbf{x}^T\boldsymbol{\tau}) \cdot \exp\{\beta_1 + \mathbf{x}^T\boldsymbol{\beta} + 0.5 \cdot \sigma^2\} \\
&\quad +\Phi(\mathbf{x}^T\boldsymbol{\tau}) \cdot \exp\{\mathbf{x}^T\boldsymbol{\beta} + 0.5 \cdot \sigma^2\}.
\end{aligned}$$

On the basis of this expression, the elasticity can be calculated by

$$\frac{\Delta^2 E[y]}{\Delta x_1 \Delta x_2}/E[y].$$

(2) In the mixed case of a continuous variable $\mathbf{x_1}$ and a dummy variable $\mathbf{x_2}$, the interaction effect is given by:

$$\begin{aligned}
\frac{\Delta}{\Delta x_2}(\frac{\partial E[y]}{\partial x_1}) &= \frac{\partial E[y|x_1, x_2 = 1, \mathbf{x}]}{\partial x_1} - \frac{\partial E[y|x_1, x_2 = 0, \mathbf{x}]}{\partial x_2} \\
&= \exp\{\beta_1 x_1 + \beta_2 + \beta_{12} x_1 + \mathbf{x}^T\boldsymbol{\beta} + 0.5 \cdot \sigma^2\} \cdot [(\tau_1 + \tau_{12}) \cdot \phi(\tau_1 x_1 + \tau_2 + \tau_{12} x_1 + \mathbf{x}^T\boldsymbol{\tau}) \\
&\quad +(\beta_1 + \beta_{12}) \cdot \Phi(\tau_1 x_1 + \tau_2 + \tau_{12} x_1 + \mathbf{x}^T\boldsymbol{\tau})] \\
&\quad -[\tau_1 \cdot \phi(\tau_1 x_1 + \mathbf{x}^T\boldsymbol{\tau}) + \beta_1 \Phi(\tau_1 x_1 + \mathbf{x}^T\boldsymbol{\tau}] \cdot \exp\{\beta_1 x_1 + \mathbf{x}^T\boldsymbol{\beta} + 0.5 \cdot \sigma^2\}.
\end{aligned}$$

On the basis of this expression, the elasticity can be calculated by

$$\frac{\Delta}{\Delta x_2}(\frac{\partial E[y]}{\partial x_1})/E[y],$$

if x_1 is a logged explanatory variable, $x_1 = \ln z_1$, and otherwise by

$$\frac{\Delta}{\Delta x_2}(\frac{\partial E[y]}{\partial x_1})/E[y] \cdot x_1.$$

References

AI, C. , NORTON, E. C. (2003) Interaction terms in logit and probit models. *Economics Letters* **80**, 123-129.

BRONS, M., NIJKAMP, P., PELS, E. , RIETVELD, P. (2008) A Meta-Analysis of the Price-Elasticity of Gasoline Demand. *Energy Economics*, 30, 2105-2122.

COM (2007) Proposal for a Regulation of the European Parliament and of the Council: Setting Emission Performance Standards for New Passenger Cars as Part of the Community's Integrated Approach to Reduce CO2 Emissions from Light-duty Vehicles, December 19, 2007, COM(2007) 856 final, European Commission, Brussels.

DOW, W. H. , NORTON, E. C. (2003) Choosing Between and Interpreting the Heckit and Two-Part Models for Corner Solutions. *Health Services & Outcomes Research Methodology* **4**, 5-18.

DUAN, N. , MANNING, W. , MORRIS, C. , NEWHOUSE, J. (1984) Choosing between the Sample-Selection Model and the Multi-Part Model. *Journal of Business and Economic Statistics* **2**, 283-289.

FRONDEL, M., PETERS, J., VANCE, C. (2008) Identifying the Rebound: Evidence from a German Household Panel. *The Energy Journal*, 29 (4), 154-163.

GRAHAM, D. J. , GLAISTER, S. (2002) The Demand for Automobile Fuel: A Survey of Elasticities. *Journal of Transport Economics and Policy* **36 (1)**, 1-26.

HAY, J. , OLSEN, R. (1984) Let Them Eat Cake: A Note on Comparing Alternative Models of the Demand for Medical Care. *Journal of Business and Economic Statistics* **2**, 279-282.

HUGHES, J. E. , KNITTEL , C. R. , SPERLING, D. R. (2006) Evidence of a Shift in the Short-Run Price Elasticity of Gasoline Demand. NBER Woring Paper No. 12530.

KAYSER, H. A. (2000) Gasoline Demand and Car Choice: Estimating Gasoline Demand using Household Information. *Energy Economics*, 22, 331-348.

LUENG, S. F. , YU, S. (1996) On the Choice between Sample Selection and Two-Part Models. *Journal of Econometrics* **72**, 197-229.

MOP (German Mobility Panel) (2009) http://www.ifv.uni-karlsruhe.de/MOP.html

NORTON, E. C. , WANG, H. , AI, C. (2004) Computing Interaction Effects and standard errors in Logit and Probit Models. *The Stata Journal* **4 (2)**, 103-116.

PICKUP, L. (1985) Women's Gender-role and its Influence on Travel Behavior. *Built Environment* **10**, 61-68.

TURNER, T. , NIEMEIER, D. (1997) Travel to Work and Household Responsibility: New Evidence. *Transportation* **24**, 397-419.

VANCE, C. , HEDEL, R. (2007) The Impact of Urban Form on Automobile Travel: Disentangling Causation from Correlation. *Transportation* **34**, 575-588.