RUHR
ECONOMIC PAPERS

Nolan Ritter
Colin Vance

Do Fewer People Mean Fewer Cars? – Population Decline and Car Ownership in Germany

#385

Imprint

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
Universitätsstr. 12, 45117 Essen, Germany

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

Editors

Editorial Office

Ruhr Economic Papers #385

Responsible Editor: Christoph M. Schmidt

ISSN 1864-4872 (online) – ISBN 978-3-86788-440-2

Ruhr Economic Papers #385

Nolan Ritter and Colin Vance

Do Fewer People Mean Fewer Cars? – Population Decline and Car Ownership in Germany

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über: *http://dnb.d-nb.de* abrufbar.

http://dx.doi.org/10.4419/86788440
ISSN 1864-4872 (online)
ISBN 978-3-86788-440-2

Nolan Ritter and Colin Vance[1]

Do Fewer People Mean Fewer Cars? – Population Decline and Car Ownership in Germany

Abstract

Drawing on household data from Germany, this study econometrically analyzes the determinants of automobile ownership, focusing specifically on the extent to which decreases in family size translate into fewer cars at the national level. Beyond identifying several variables over which policy makers have direct leverage, including the price for fuel, the supply of public transit, and land use features, the analysis uses the estimated coefficients from a multinomial logit model to simulate car ownership rates under alternative scenarios pertaining to demographic change and other socioeconomic variables. Our baseline scenario predicts continued increases in the number of cars despite decreases in population, a trend that could be partially offset by substantial increases in fuel prices.

JEL Classification: C25, D10, R41

Keywords: Car ownership; demographic change; Germany; multinomial logit; simulation

November 2012

1 Nolan Ritter, RWI; Colin Vance, RWI and and Jacobs University Bremen. – All correspondence to Nolan Ritter, RWI, Hohenzollernstr. 1-3, 45128 Essen, Germany, E-Mail: nolan.ritter@rwi-essen.de.

1 Introduction

In Germany, as elsewhere in the industrialized world, the demand for automobiles has grown substantially in the last decades. Between 1995 and 2009, the car ownership rate increased by roughly 32%, from 417 to 551 cars per 1000 inhabitants (EEA, 2012). At the same time, the share of carless households in the country has been markedly decreasing, from 38% in 1976 to 19% in 2002 (Buehler and Kunert, 2008, p. 9). Understanding the determinants of these trends has emerged as a major priority within the scientific and policy arenas given the range of externalities associated with the automobile, including air and noise pollution as well as congestion, accidents, and land use considerations.

Much of the recent empirical work on automobile ownership has drawn on household data to focus on the role of socio-demographics and geographic context. Whelan (2007), for example, undertakes a detailed analysis of both income and demographic structure, finding both factors to be important predictors of the number of cars owned. Other issues covered in this research include the role of employment status (Raphael and Rice, 2002; Matas et al., 2009), the costs of car acquisition and motoring (Dargay, 2002), and the influence of car-sharing (Prettenthaler and Steininger, 1999). A relatively smaller body of work has addressed the impact of urban form on car ownership. Studies in this vein include Potoglou's (2008) analysis of the effect of neighborhood characteristics on the type of vehicle owned, and Bento et al.'s (2005) investigation of city shape, the supply of public transit, and other aspects of urban spatial structure. With some exceptions (e.g. Karlaftis and Golias, 2002; Buehler, 2011), the research on urban form tends to draw on data from North America, and there have been relatively few investigations of this issue in the European context.

The incidence of car ownership in Germany is of particular interest for several reasons. First, as Europe's largest car market, the country is a major source of transport emissions, accounting for some 19% of the EU-15 total in 2005 (EEA, 2008). Moreover, the German government has for many years pursued policies that combine high fuel taxes with land use planning measures to reduce automobile dependency. In 1993, the government legally codified the concept of "decentralized concentration" into its regional planning guidelines (BBR, 1993), an approch predicated on compact development as a means of spatially integrating residuential, recreational, and commercial land uses to reduce car reliance. Since that time, several German cities have adopted urban planning models that employ compact development strategies (Dresden, 2002).

Perhaps most significantly, like in many other countries of Europe, major socio-demographic changes are currently underway in Germany that could dramatically affect future automobile ownership. Between 2000 and 2005, for example, the birth rate decreased some 9.3%, from 9.18 to 8.33 births/1000 population, having already decreased 19.5% over the preceding decade - both observations being driven by women having fewer children and increses in the average age of the population. By 2050, Germany's population is projected to shrink by roughly 16% (Destatis, 2006), a trend that will be paralleled by an increasingly older age structure of the German population and an increase in the number of single person households. In its annual resport, the German Council of Economic Experts presents virtually the same figures with respect to total population and age composition (Sachverständigenrat, 2011, p. 374). While several studies have suggested that these changes will have profound consequences for transport demand (Limbourg, 2004; Just, 2004; Zumkeller et al., 2004), the anticipated impacts are largely speculative, and there have been few attempts to quantify how the underlying variables affect automobile ownership at the household level.

Drawing on travel survey data, the present study aims to address this issue by exploring the implications of household-level socio-demographic changes for car ownership at the national level. The analysis proceeds in two steps. We begin by estimating a multinomial logit model of the determinants of car ownership. The model specification includes a rich array of explanatory variables, many of which, such as fuel prices and the accessibility of public transit, have immediate relevance for policy but have rarely been parameterized using household level data. Following validation of the model by comparing the in-sample predictions with national car-ownership figures, the second step uses the model coefficients to simulate car-ownership levels under alternative scenarios about the future trajectory of key explanatory variables. We are particularly interested in the effects of demographic change, and to this end draw on population projections published by Germany's Federal Statistics Office (Destatis, 2006).

Our baseline scenario, which assumes decreases in the overall population coupled with increases in the number of one-person households, the share of the elderly, income, and fuel prices, indicates that the increase in car ownership will continue despite population decline, albeit at a slightly abbreviated pace relative to recent years. Nevertheless, this result is found to be strongly dependent on assumed increases in income, and an alternative scenario additionally reveals some scope for reducing the number of cars through substantial increases in fuel prices. We also uncover evidence for a negative impact of public transit service on the proclivity to own a car. Taken together, these results can be used to assess the country's future infrastructure needs and how these needs may be altered by public policy.

2 Data assembly

The primary data source used in this research is drawn from the German Mobility Panel (MOP, 2011), a household travel survey financed by the German Federal Ministry of Transport, Building and Housing. Participating households are surveyed annually over each of three years, with exiting households replaced by a new cohort. The data used in this paper spans the years 1999 through 2009 and is described in Table 1.

Table 1: *Descriptive statistics*

Variable	Description	Mean	Std. Dev.
cars	cars owned	1.09	0.73
hhsize	household size	2.12	1.07
share2039	share of 20 to 39 year old	0.20	0.33
share4064	share of 40 to 64 year old	0.43	0.41
share65	share of 65 and older	0.25	0.41
income	monthly household income in Euro	2,186.22	865.95
income squared	squared monthly income	$10^6 * 5.53$	$10^6 * 4.02$
distance	total distance to work for all household members	12.66	24.14
fuel price	moving average of last 3 years fuel price	1.04	0.11
urban	1 if household lives in urban area	0.35	0.48
minutes	walking minutes to nearest public transit stop	5.68	4.84
rail	1 if nearest public transit stop is a rail station	0.22	0.42
company cars	number of company cars in household	0.07	0.28
licenses	share of licensed drivers in household	0.75	0.33
density	transit density of non-rail modes	39.23	52.86
insurance cost	vehicle insurance class	6.21	2.92

Std. Dev. stands for standard deviation.

The MOP is comprised of two surveys, one with households as the observational unit and the other with cars. The household survey takes place over the course of a week in the fall and elicits sundry aspects of everyday travel behavior, person-related characteristics, and household characteristics, including the number of cars owned. To construct the dependent variable, we use the latter variable to create an indicator distinguishing between households owning 0, 1, 2, and 3 or more cars. The share of households falling into each of these categories breaks down as 19%, 56%, 22%, and 3%, respectively, with only a small share of households - less than 1% - owning more than three cars.

While the household survey forms the basis for our empirical analysis, we additionally merged in information from a separate survey of the MOP that focuses specifically on vehicle travel. This so-called "tank survey" draws a 50% sub-sample of randomly selected car-owning households from the larger MOP survey (which also includes households that do not own a car). The tank survey takes place over a roughly six-week period, during which time respondents record various information upon each visit to the gas station, including the price paid for fuel. As the fuel price is a potentially important determinant of automobile ownership, a Geographic Information System was used to create a coverage of spatially interpolated fuel prices (in real terms) for all of Germany based on the postal code location of households participating in the tank survey.

This coverage was then overlaid onto a map of postal code locations in the household survey, thereby allowing for each household to be assigned the locally prevailing fuel price. This process was repeated for each year of the data, yielding a dataset of fuel prices that varies over space and time. An accuracy assessment of the data was undertaken by calculating the yearly average fuel prices and comparing these with those published for the German market by the oil company ARAL (2009). The correspondence between the two sources is tight, deviating by an average of less than 1% over the 1999-2009 time interval.

Two external data sources were also drawn upon in assembling the data. The first of these, which was provided to the authors by the German Insurance Association (GDV) for the year 2003, proxies for the cost of automobile insurance using a 12-point cost index. The second, obtained from the German Statistical Agency for the years 2004, 2005 and 2007, measures the total mileage traveled by all non-rail modes of public transit. This variable and the insurance index are both measured at the level of an administrative unit referred to as a *Kreis*, of which there are roughly 445 spatial divisions in Germany. We divided the total transit mileage by the area of each *Kreis* to obtain a density measure. Because a *Kreis* has borders different from the postal codes used to designate household locations, a GIS was used to merge the insurance and transit density variables with the household data.

In total, the data contains 5,052 households over the 1999-2009 time interval. Of these, 1,721 participated in one year of the survey, 1,233 in two years and 2,098 in all three years, yielding a total of 10,481 observations on which the model is estimated. To correct for non-independence emerging from the repeated observations of households over the three years of the survey, the regression disturbance terms are clustered at the household level, and the presented measures of statistical significance are robust to this survey design feature.

3 The model

Random utility theory provides an appropriate framework for our analysis as it predicts choices by comparing the utility associated with distinct levels of car ownership. Each household faces a choice set with J elements representing different numbers of cars owned. The utility U_{im} of household i for alternative m in J comprises a deterministic and a stochastic component:

$$U_{im} = V_{im} + \epsilon_{im} \tag{1}$$

with $V_{im} = \alpha_m + \boldsymbol{x}_{im} \cdot \boldsymbol{\beta}$ as representative utility, determined by the constant α_m, the vector $\boldsymbol{x}_{im}$ capturing the characteristics of the household, and the parameter vector $\boldsymbol{\beta}$ measuring the contribution of household characteristics to utility. The random component is denoted by ϵ_{ij}. Utility maximization implies that the probability P that household i chooses car ownership level m is determined by:

$$P(V_{im} + \epsilon_{im} > V_{ik} + \epsilon_{ik}) = P(\epsilon_{ik} - \epsilon_{im} < V_{im} - V_{ik}), \forall k \neq m \tag{2}$$

Assuming the error terms to be identically and independently distributed as a log Weibull distribution, the multinomial logit model results, with choice probabilities equal to (Long and

Freese, 2006, p. 228):

$$P(y_i = m) = \frac{\exp(\boldsymbol{x}_i \cdot \beta_m)}{\sum_{j=1}^{J} \exp(\boldsymbol{x}_i \cdot \beta_j)} \;, \tag{3}$$

where y_i is a discrete variable denoting the number of cars owned.

The suite of variables selected for inclusion in $\boldsymbol{x}$ measure the household characteristics and regional features that are hypothesized to influence the household's choice of how many cars to own in maximizing utility. Variable definitions and descriptive statistics are presented in Table 1.

Negative signs are expected for the variables that either increase the costs of car ownership and use or decrease the costs of using alternative modes. Two key sources of ownership and use cost are fuel prices (*fuel price*), measured as a lagged three-year moving average, and the cost of automobile insurance (*insurance cost*). Land use attributes that facilitate mobility and accessibility are also expected to decrease car ownership. Three such attributes are included in the model: a dummy variable indicating residence in an urban area (*urban*), a dummy indicating whether the nearest public transit stop is serviced by rail (*rail*), and a continuous measure of the transit density of non-rail modes (*density*). The urban dummy is expected to have a negative effect not only by virtue of increased proximity of service outlets for undertaking maintenance and recreational activities, but also owing to the higher costs of searching for or renting a parking space in urban areas. The model also includes a measure of the number of company cars to which the household has access. To the extent that company cars substitute for cars owned by the household, we would expect this variable to have a negative coefficient, as well.

Positive signs are ascribed to variables that increase the benefits of car ownership and/or the opportunity costs of using alternative modes. These include demographic features such as household size and the share of members with a driver's license. They also include the distance in walking minutes from the household to the nearest transit stop, the distance separating the household from the employment location summed over all working members, and the household's monthly disposable income, which is specified as a quadratic to allow for nonlinear effects. To capture the impact of age composition, we also include a suite of variables measuring the share of household members in different age brackets, with the 0-19 bracket excluded as the base category.

4 Other modeling considerations

The multinomial logit is one of several limited dependent variable models that have been availed in the literature on car ownership, others of which include the ordered logit and probit, the poisson, and the negative binomial. While these alternatives were also explored, our selection of the multinomial logit was guided by three considerations. First, as demonstrated by Bhat and Pulugurta (1998), the unordered response mechanism underpinning the multinomial logit model is, in contrast with ordered-response models, consistent with the global utility maximizing hypothesis. Second, attempts to estimate more computationally intensive models such as the multinomial probit and multilevel logit models either encountered convergence problems or were not found to provide any statistical improvement over the multinomial logit, a possible consequence of the lack of alternative-specific variables in the data set. Finally, and perhaps most importantly for the aim of the current study, the in-sample predicted probabilities obtained from the model yielded estimates of car ownership that were much closer to officially published figures than estimates obtained from alternatives models, a point documented further below.

These considerations notwithstanding, the multinomial logit has some drawbacks, one being the more onerous interpretation arising from the fact that a coefficient estimate is generated for each of the values of the multinomial dependent variable. A second shortcoming of the model is that it is characterized by the so-called independence of irrelevant alternatives (IIA), summarized succinctly by Cheng and Long (2007, p. 584) as meaning that, all else equal, "the choice between two alternative outcomes is unaffected by what other choices are available." While the IIA assumption is in some contexts overly restrictive, particularly when relevant options have been omitted from the definition of the choice set, it is deemed to be relatively innocuous for the current application. As advocated by McFadden (1973) and reiterated by Long and Freese (2006), the multinomial logit model is appropriate when the choice categories are clearly distinct and not substitutes for one another, a condition that can reasonably be said to apply to the choice between different levels of car ownership.

A final cautionary note concerns potential endogeneity. While the explanatory variables included in this analysis afford reasonably broad coverage of the determinants of automobile ownership, we cannot completely rule out the possibility that they are correlated with additional unobserved factors that impact travel decisions. Such correlation would give rise to endogeneity bias and preclude us from ascribing a causative interpretation to the estimated coefficients. In this regard, it is plausible that decisions pertaining to car ownership are jointly determined with those pertaining to residential choice, implying that the coefficients of the land use variables are partially picking up the effects of neighborhood preferences. Eluru et al. (2009), for example, find that features of the surrounding vicinity may be an important determinant of residential relocation for those who commute by public transit. Moreover, we lack information on potentially important service attributes for car use itself and for competing modes, such as regional congestion, which may be correlated with some of our explanatory variables. We consequently abstain from making claims about causality, instead applying a descriptive interpretation to the estimates. It is noted, however, that concerns about endogeneity do not bear on our ultimate aim of exploring the predictions from the model.

5 Results

Table 2 reports the coefficient estimates and robust standard errors from the multinomial logit model. Households with no cars are selected as the base category, so that the interpretation of the coefficients is made with respect to this case. The last column presents the p-value calculated from a chi-square test of whether the coefficients are simultaneously equal to zero, which is the appropriate reference for drawing inferences concerning the variable's statistical significance. Appendix 1 presents results from an ordered probit model as a basis for comparison, which confirms that the qualitative differences between the multinomial logit and ordered probit models are minor.

Turning to the estimates, nearly all the coefficients, with the exception of the share over age 64 and insurance costs, are statistically significant at the 5% level and have signs that are consistent with the hypothesized effects. Household size, a larger share of license holders, and an older age composition have a positive association with car ownership. Likewise, longer walking minutes to the nearest transit stop has a positive coefficient, as does household income, albeit tapering off slightly as income increases. Conversely, higher costs of car ownership, as measured by fuel costs and urban residency, and lower costs of public transit use, as measured by transit service density and the local availability of rail service, all have negative coefficients. Lastly, the negative coefficient on company cars confirms the intuition that company cars may serve as substitutes for privately owned vehicles.

Table 2: *Multinomial logit regression results*

	1 vs. 0 Cars (j=1)		2 vs. 0 Cars (j=2)		3+ vs. 0 Cars (j=3)		Joint Test
Variable	Par.	Err.	Par.	Err.	Par.	Err.	P-Values
hhsize	1.186**	0.103	2.194**	0.124	3.370**	0.169	0.000
share2039	−0.123	0.434	1.857**	0.562	4.106**	0.921	0.000
share4064	0.381	0.403	1.973**	0.528	4.430**	0.920	0.000
share65	0.306	0.397	0.715	0.523	2.759**	0.975	0.039
income	0.002**	0.000	0.004**	0.000	0.002**	0.001	0.000
income squared	−0.342**	0.061	−0.527**	0.087	−0.131	0.141	0.000
distance	0.004	0.004	0.011**	0.004	0.012**	0.004	0.001
fuel price	−1.244**	0.403	−0.432	0.502	−0.649	0.858	0.002
urban	−0.567**	0.163	−1.176**	0.223	−1.651**	0.470	0.000
minutes	0.056**	0.013	0.090**	0.015	0.097**	0.018	0.000
rail	−0.359**	0.112	−1.066**	0.154	−1.015**	0.253	0.000
company cars	−2.373**	0.144	−4.719**	0.225	−5.862**	0.447	0.000
licenses	3.419**	0.150	5.640**	0.294	8.901**	0.814	0.000
density	−0.005**	0.001	−0.006**	0.002	−0.007	0.005	0.004
insurance cost	−0.013	0.020	−0.006	0.025	0.010	0.042	0.829
constant	−4.665**	0.722	−15.135**	0.985	−23.167**	1.802	0.000

Par. stands for parameter, Err. stands for standard error. ** (*) indicates significance at the 1% (5%) level.

While the nonlinearity of the multinomial logit model precludes moving the interpretation of the coefficient estimates beyond their sign and statistical significance, we can calculate the marginal effects to assess the impact of changes in the explanatory variables on the probability of the household choosing any one of the car ownership categories. For continuous variables, the marginal effects are calculated by taking the partial derivative of Equation (3) with respect to the variable of interest:

$$\begin{aligned} &\frac{\partial P(y_i = m \mid \boldsymbol{x}_i)}{\partial x_{ik}} = \\ &P(y_i = m \mid \boldsymbol{x}_i) \left[\beta_{k,m|J} - \sum_{j=1}^{J} \beta_{k,j|J} \cdot P(y_i = j \mid \boldsymbol{x}_i) \right] \end{aligned} \tag{4}$$

Equation (5) defines the discrete change for the case of dummy variables:

$$\frac{\Delta P(y_i = m \mid \boldsymbol{x}_i)}{\Delta x_{ik}} = P(y_i = m \mid \boldsymbol{x}_i, x_{ik} = 1) - P(y_i = m \mid \boldsymbol{x}_i, x_{ik} = 0) \tag{5}$$

As Equations (4) and (5) yield a unique marginal effect for every observation in the data, a conventional approach is to evaluate the effects at the means of the explanatory variables, the results from which are presented in Table 3. A cursory look at the table reveals that the tight story line indicated by the coefficient estimates does not carry over to the marginal effects. For example, the marginal effect on household size for the one car case suggests that a one person increase in this variable decreases the probability of owning a car by 0.06 relative to the base case, contradicting the expectation of a positive impact. Among the other apparent anomalies for this category are the negative sign on income and the positive sign on the rail service dummy as well as on the number of company cars. With respect to the two- and three car categories, most of the signs are as expected, with the exception of the positive marginal effect of the fuel price in the two-car category.

Table 3: *Marginal effects at the means*

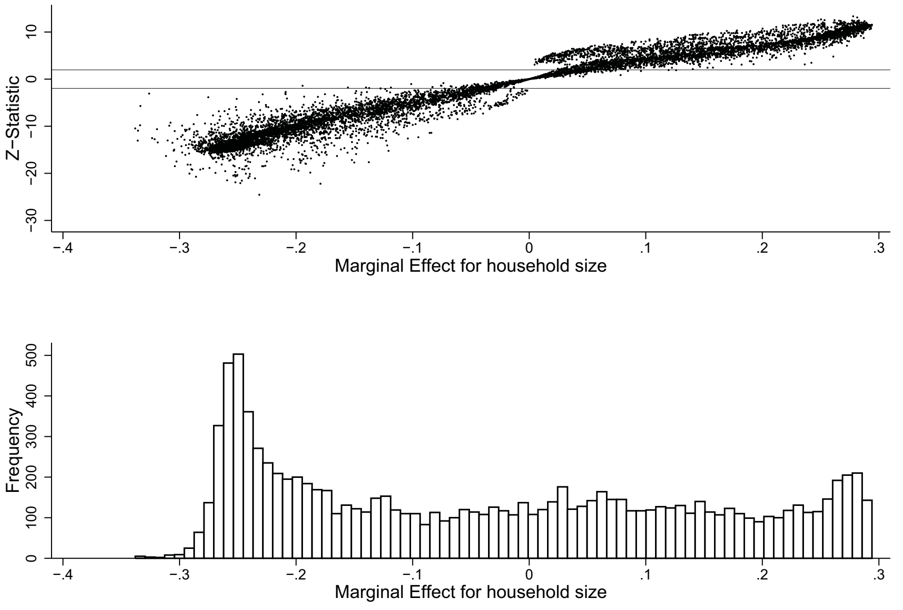
	1 vs. 0 Cars (j=1)		2 vs. 0 Cars (j=2)		3+ vs. 0 Cars (j=3)	
Variable	Marg. Eff.	Std. Err.	Marg. Eff.	Std. Err.	Marg. Eff.	Std. Err.
hhsize	−0.061**	0.010	0.130**	0.009	0.007**	0.002
share2039	−0.238**	0.047	0.234**	0.044	0.014**	0.005
share4064	−0.170**	0.045	0.192**	0.043	0.013**	0.005
share65	−0.038	0.044	0.051	0.042	0.008	0.004
income	0.000**	0.000	0.000**	0.000	0.000**	0.000
distance	−0.001*	0.000	0.001**	0.000	0.000*	0.000
fuel price	−0.152**	0.043	0.087	0.039	0.001	0.003
urban	0.035	0.020	−0.073**	0.017	−0.003*	0.001
minutes	−0.001	0.001	0.005**	0.001	0.000**	0.000
rail	0.049**	0.012	−0.076**	0.010	−0.002*	0.001
company cars	0.156**	0.023	−0.300**	0.022	−0.012**	0.003
licenses	−0.098**	0.029	0.292**	0.027	0.019**	0.003
density	0.000	0.000	0.000	0.000	0.000	0.000
insurance cost	−0.002	0.002	0.001	0.002	0.000	0.000

Marg. Eff. stands for marginal effect, Std. Err. stands for standard error. ** (*) indicates significance at the 1% (5%) level.

In appraising these results, it should be borne in mind that they represent mean effects that potentially mask substantial heterogeneity across the individual observations. An impression of the degree of this heterogeneity can be gleaned by plotting the magnitude of the individual marginal effects against their associated Z-statistic, as is illustrated in Figures 1, 2 and 3 for the variables household size, income, and the urban dummy for the one car category.

Below each plot, a histogram is additionally included to indicate the density distribution of the estimates. In all three cases, the marginal effects are seen to span far to the right and left of zero, with the majority falling outside the band indicating statistical significance at the 5% level. The dispersed pattern in the graphs clearly complicates forming a coherent interpretation of the relationship between changes in the explanatory variables and the probability of different car ownership levels. Suffice it to emphasize that the patterns do not imply heterogeneity in preferences for automobile ownership across households - this interpretation is ruled out by the assumption of homogeneity imposed by the model. Rather, the patterns highlight how the estimated marginal effects for each of the considered variables are fundamentally dependent on the values assumed by the other explanatory variables in the model.

Figure 1: *Individual marginal effects for household size (1 vs. 0 cars)*



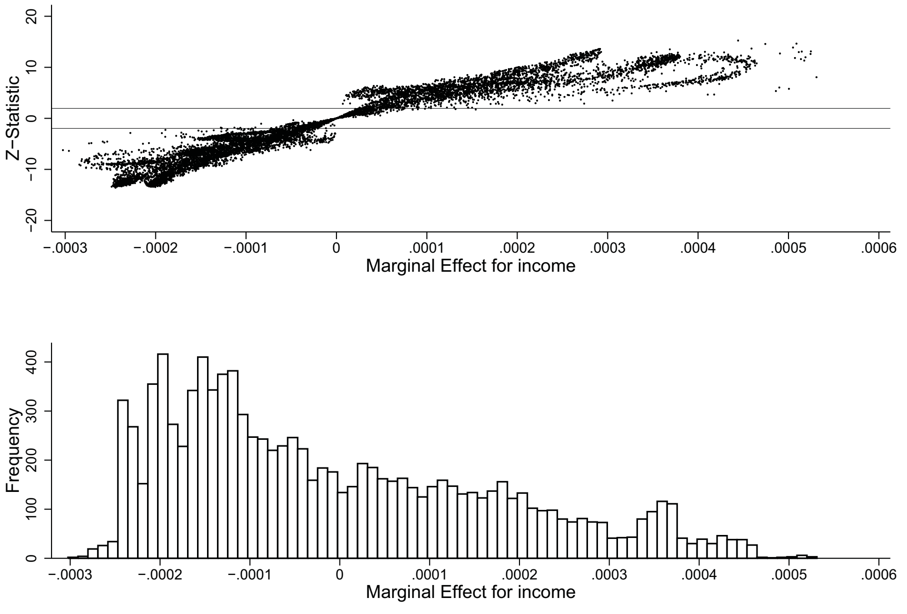
6 Simulations

An alternative approach to interpreting the effects of the explanatory variables is to explore how changes in their values bear upon the predicted probabilities generated by the model. To this end we undertake a simulation exercise that draws upon population projections of the Federal Statistics Office, along with projections of other key variables that were recently used in a study commissioned by the German government of the country's future energy needs (Energieprognose, 2010). The population projections include estimates of the overall population, as well as breakdowns by age structure and household size, all of which are presented in the two panels of Figure 4.

The top panel of the figure depicts the stagnation and eventual downward trend of Germany's population; by 2030 the population is expected to decrease to 77.4 million, a 5.3% drop relative to 2009. The abating influence of this dynamic on the number of cars, however, may be partially offset by structural changes in the population. As indicated in panel 2, there is a marked rise projected in the number of one-person and two-person households accompanied by a decrease in the number of households with three or more persons. The increase in one-person households, in particular, is likely to put upward pressure on car counts owning to the more limited scope for car-sharing that can otherwise be exploited by households with several members.

Of course, demographics are only one of multiple factors that will bear on car counts. Our

Figure 2: *Individual marginal effects for income (1 vs. 0 cars)*

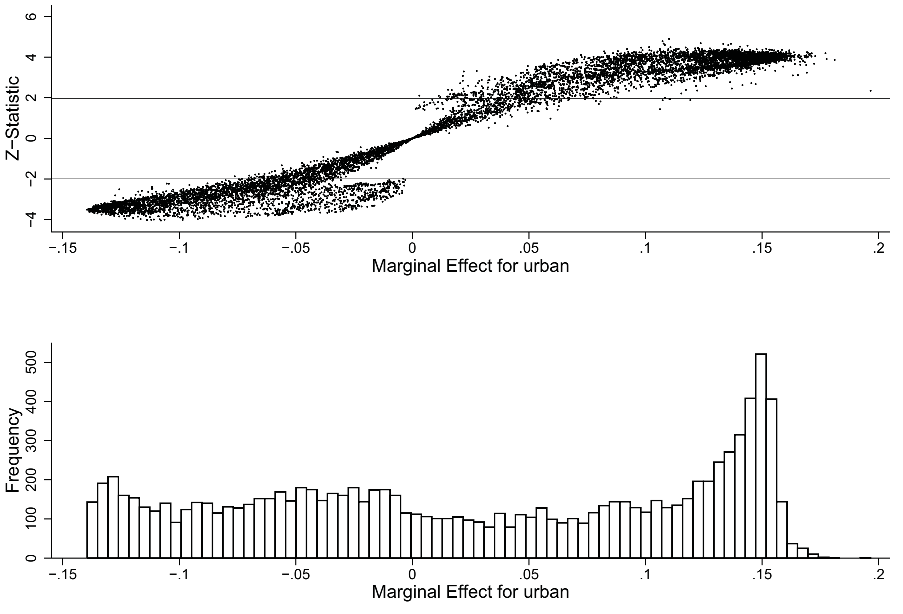


baseline scenario additionally takes into account changes in the share of license holders, household income, fuel prices, and the share of households located in rural and urban areas. The projection of the share of license holders, which is assumed to increase by 1.6% per year, is taken from ifmo (2008), while the remaining figures use the Energieprognose (2010) as a point of reference. As illustrated in Table 6, presented in Appendix 2, the projected annual increases for household income, the fuel price, and the percent of households located in urban areas are 0.8%, 1.0%, and 1.1%, respectively. All other variables used for the simulation stay fixed at their mean values from 2009.

6.1 In-sample predictions

Before exploring the out-of-sample predictions of the baseline model and other scenarios, it is of interest to assess the model's accuracy in correctly predicting observed levels of car ownership using in-sample predictions. We consequently undertake a validation exercise that compares figures on the total number of private automobiles in Germany published by the Federal Motor Transport Authority (KBA, 2012) for the years 2001-2007 with estimates generated by the model. Using the model coefficients, we generated predicted probabilities of owning one, two, or three or more cars by year and for different household sizes. We then multiplied each probability by the number of households in the corresponding size category using data compiled from the

Figure 3: *Individual marginal effects for urban residency (1 vs. 0 cars)*



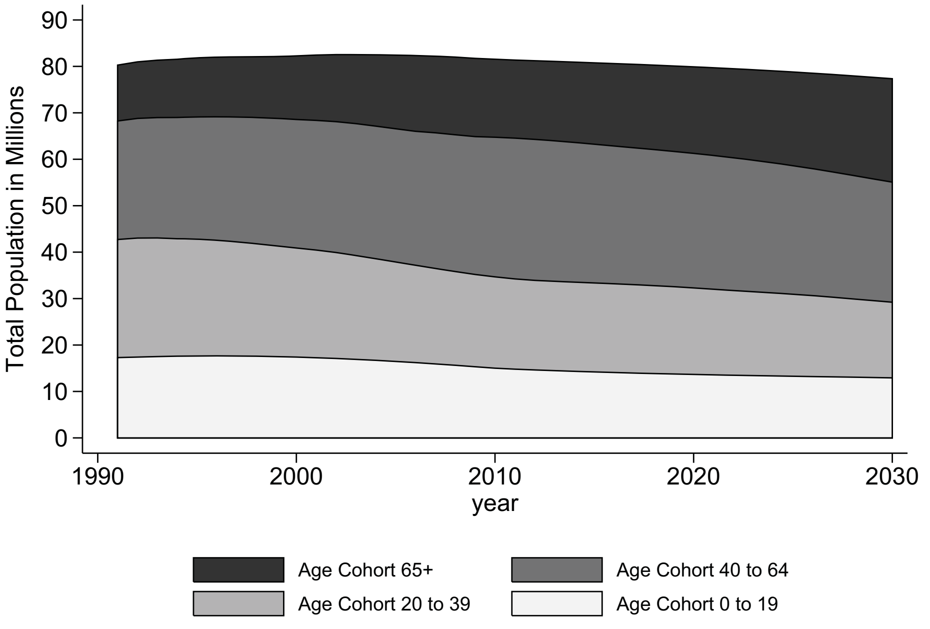
Federal Statistics Office. These products were summed to obtain an estimate of the total number of privately owned cars by year. The results from this calculation are presented in Table 4, along with the figures from the KBA for the years for which KBA data is available.

Overall, the correspondence between the estimates calculated using the model and the official statistics is decent. The largest discrepancy is seen for the year 2005, when the model predicts 44.4 million cars compared to the observed count of 40.6 million, for a difference of 9.4%. Otherwise, the discrepancies range between 1.2% in 2007 and 5.5% in 2001 (with the exception of 2008, when the KBA changed its counting procedure to exclude cars not registered throughout the year, making it no longer directly comparable to the estimates from the model).

The replication of this validation exercise using alternative models such as the ordered probit model presented in the appendix yielded estimates that were fairly far off the mark, deviating by upwards of 40%. That said, we would not unequivocally advocate for the superiority of unordered response models. Given the variety of approaches that have been gainfully implemented in the literature, the optimal choice is likely to be highly dependent on the data. Matas and Raymond (2008), for example, find no difference in the forecasting performance between ordered- and unordered-response mechanisms using household-level car ownership data from Spain.

Figure 4: *Demographic structure of Germany*

(a) *Population Structure*



(b) *Household Structure*

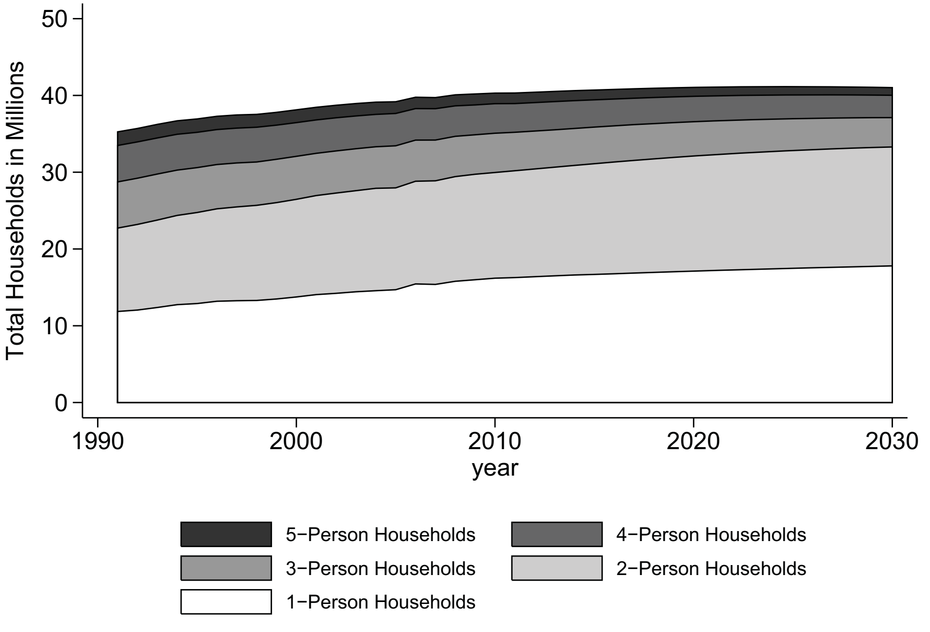


Table 4: *Millions of predicted and observed privately owned cars*

	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
predicted total cars	41.4	41.9	41.2	40.2	40.9	42.4	44.4	43.0	42.1	42.7
observed total cars	–	–	39.1	39.6	39.9	40.3	40.6	41.2	41.6	37.1
difference in %	–	–	5.5	1.5	2.5	5.2	9.4	4.4	1.2	15.0

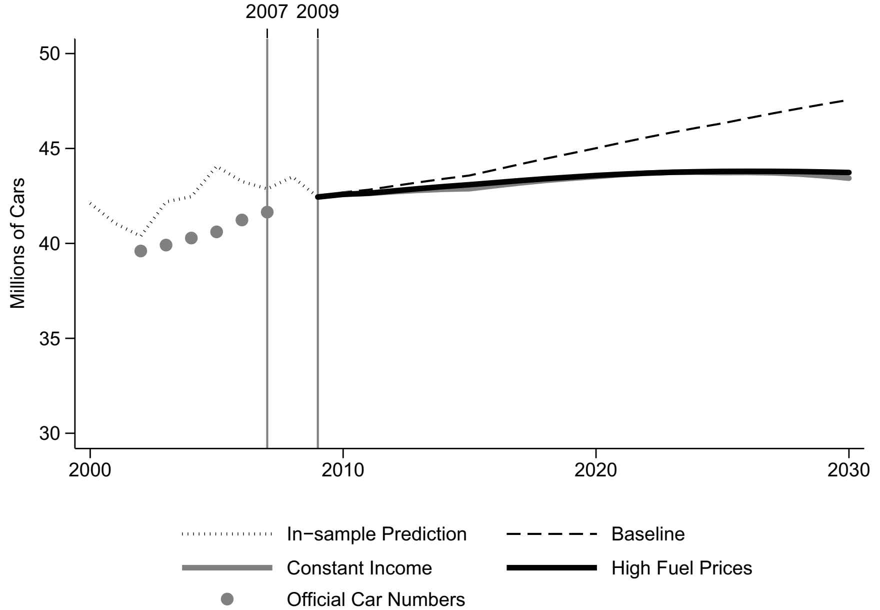
As of 2008, KBA changed its counting procedure to only include privately owned cars that are registered over the entire year.

6.2 Out of sample predictions

Generating out-of-sample predictions with an econometric model is, of course, an approximate undertaking for which caveats abound. Aside from the uncertainty surrounding the future values of the explanatory variables, these caveats include the neglect of general equilibrium effects and the assumption that household preferences remain unchanged. Nevertheless, the validation exercise suggests that the model can provide an indicative measure of likely changes. Moreover, the scenarios serve as a complementary illustration of how changes in the explanatory variables influence car ownership, one that is in some respects more revealing than that culled from the coefficient estimates and marginal effects.

Figure 5 presents simulated results from the baseline and two alternative scenarios using 2009 as the reference year and projecting to 2030. The baseline scenario suggests that the number of cars in Germany will continue to increase through 2030 when the figure reaches 47.6 million, but that the rate of increase is somewhat less pronounced than in recent years. Some sense for the extent to which this pattern is determined by assumed increases in income can be seen by the line labeled *Constant Income*, which plugs in values identical to the baseline case except for income, which is held fixed at its 2009 value. This yields a strikingly altered trend that steadily, if only moderately, increases until 2030 to reach a level of 43.7 million cars. Thus, in the absence of income growth, this result would lead us to conclude that fewer people has a stabilizing effect on growth of the number of cars. A final scenario, labeled as *High Fuel Price*, illustrates the scope for reducing the number of automobiles via increases in fuel prices. This scenario also uses values identical to the baseline scenario, but replaces the series for fuel prices with one that assumes an annual 5% increase, with the result that motorists would be paying 3.26 Euros/liter in 2030. The simulation suggests that the effects of these high costs yields a trajectory very similar to that of the *Constant Income* scenario. Indeed, as presented in the appendix, the 95% confidence intervals of all three scenarios overlap, indicating that the differences between them are not statistically significant.

Figure 5: *Simulation results*



7 Conclusion

Based on a multinomial logit model estimated on household data from Germany, this paper has modeled the socio-demographic determinants of car ownership and, using the coefficient estimates from the model, presented future scenarios of overall car counts under alternative assumptions about the trajectories of key variables. As Germany currently finds itself in the midst of dramatic changes in both the size and structure of its population, we were particularly interested in exploring the implications of demographic changes for the evolution of the stock of privately held automobiles. Our baseline scenario suggests that, despite the projected decrease in population, the number of cars on German roads will continue to increase moderately, at about 0.54% per annum, until 2030. An alternative simulation holding income fixed suggests that this projected increase is strongly predicated on a steady 0.8% increase in household income; in the absence of this increase, the number of cars in 2030 is projected to be slightly lower than its current level.

Our analysis additionally revealed several variables associated with car ownership over which policy makers have direct leverage. The negative coefficient of the urban dummy variable, for example, suggests that households respond to land use density when reaching car ownership decisions. Similarly, the variables capturing the frequency, proximity, and quality of public transit service all had the expected negative effects on car ownership. Finally, fuel prices were also seen to have a negative effect, although the simulation suggested that rather large increases in fuel prices would be required to notably decrease car ownership levels.

Beyond policy deliberations concerning future infrastructure needs, these results can serve as a building block for an integrated modeling approach that additionally incorporates decisions pertaining to distance traveled and mode choice (e.g. Kitamura, 2009). Such an analysis can in turn be used for more comprehensive projections of emissions and congestion under alternative scenarios. Future work with the data will therefore be directed toward this line of inquiry, and will additionally explore the scope for incorporating the insights gained from other studies with this data that have estimated fuel price elasticities (e.g. Frondel and Vance, 2009) and the proclivity to use public transit (Vance and Peistrup, 2011).

Appendix 1: Results from an ordered probit model

Table 5 presents coefficient estimates from an ordered probit model.

Table 5: *Ordered probit results*

	Parameters	Standard Errors
hhsize	0.695**	0.029
share2039	0.768**	0.155
share4064	0.799**	0.145
share65	0.479**	0.141
income	0.001**	0.000
income squared	0.011**	0.002
distance	0.004**	0.001
fuel price	−0.036	0.141
urban	−0.346**	0.064
minutes	0.022**	0.003
rail	−0.299**	0.043
company cars	−1.437**	0.062
licenses	1.780**	0.068
density	−0.002**	0.001
insurance cost	0.002	0.007
Cutoff Points		
cut1	3.508	0.262
cut2	6.047	0.273
cut3	7.780	0.281

** (*) indicates significance at the 1% (5%) level.

With the exception of the fuel price, which is statistically insignificant in the ordered probit, the qualitative findings with respect to the question of statistical significance are the same as those in the multinomial logit model. Moreover, the signs of the coefficient estimates from the ordered probit are all consistent with intuition.

Appendix 2: Baseline assumptions

This table presents the values used for the baseline simulation presented in Figure 5. The values of all other variables from the model are set at their mean when generating the baseline predictions.

Table 6: *Baseline assumptions*

year	share2039	share4064	share65	fuel price	income	urban	licenses
2010	0.1745	0.4253	0.1725	1.18	2729.74	0.6860	0.6126
2015	0.1737	0.4182	0.1757	1.24	2776.34	0.6910	0.6312
2020	0.1747	0.4044	0.1845	1.30	2891.55	0.6953	0.6504
2025	0.1711	0.3906	0.1989	1.37	3025.68	0.6990	0.6701
2030	0.1638	0.3776	0.2202	1.44	3190.13	0.7018	0.6905

** (*) indicates significance at the 1% (5%) level.

Appendix 3: Confidence intervals

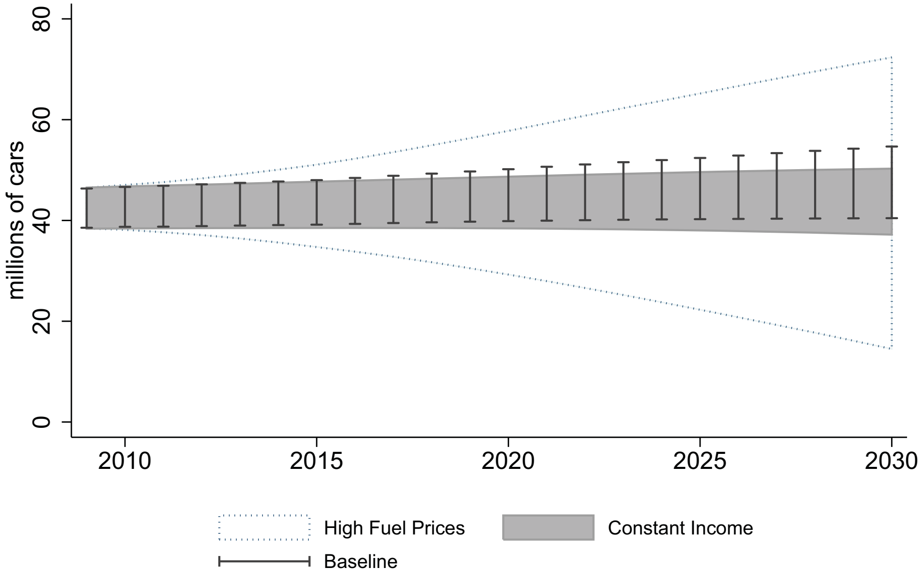
To further facilitate interpretation of the projections in Figure 5, the associated 95% confidence intervals are plotted in Figure 6 using a statistical simulation technique suggested by King et al. (2000).

Recognizing that the parameters from a model estimated using maximum likelihood are asymptotically normal, the method employs a sampling procedure akin to Monte Carlo simulation in which a large number of values - say 1000 - of each estimated parameter is drawn from a multivariate normal distribution. Taking the vector of coefficient estimates from the model as the mean of the distribution and obtaining the variance from the variance-covariance matrix, each of the 1000 simulated parameter estimates can then be multiplied by corresponding predetermined values of the explanatory variables to generate 1000 predicted probabilities. The range of these probabilities conveys the associated degree of uncertainty. By ordering the probabilities from lowest to highest and then referencing the 25th and 975th positions in the array, we obtain an estimate of the 95% confidence interval. Tomz et al. (2003) have written a program called Clarify for implementing this technique, downloadable from `http://gking.harvard.edu/`. As illustrated above, the confidence intervals for all three scenarios overlap, indicating that the differences in the predicted values are not statistically significant.

8 Acknowledgements

The authors thank Christoph M. Schmidt for valuable comments and suggestions.

Figure 6: *95% confidence intervals for the projections*



References

ARAL (2009). *Fuel Prices.* Aral Aktiengesellschaft.

BBR (1993). *Raumordnungspolitischer Orientierungsrahmen - Leitbild für die räumliche Entwicklung der Bundesrepublik Deutschland.* Bonn: Bundesministerium für Raumordnung, Bauwesen und Städtebau.

Bento, A. M., M. L. Cropper, A. M. Mobarak, and K. Vinha (2005). The effects of urban spatial structure on travel demand in the United States. *Review of Economics and Statistics 87*(3), 466–478.

Bhat, C. R. and V. Pulugurta (1998). A comparison of two alternative behavioral choice mechanisms for household auto ownership decisions. *Transportation Research Part B: Methodology 32*(1), 61–75.

Buehler, R. (2011). Determinants of transport mode choice: a comparison of Germany and the USA. *Journal of Transport Geography 19*(4), 644–657.

Buehler, R. and U. Kunert (2008). *Trends und Determinanten des Verkehrsverhaltens in den USA und in Deutschland.* Deutsches Institut für Wirtschaftsforschung and VirginiaTech.

Cheng, S. and J. S. Long (2007). Testing for IIA in the multinomial logit model. *Sociological Methods & Research 35*(4), 583–600.

Dargay, J. M. (2002). Determinants of car ownership in rural and urban areas: a pseudo-panel analysis. *Transportation Research Part E: Logistics and Transportation Review 38*(5), 351–366.

Destatis (2006). *Germany's population by 2050: Results of the 11th coordinated population projection*. Wiesbaden: German Federal Statistical Office.

Dresden (2002). *Integriertes Stadtentwicklungskonzept der Landeshauptstadt Dresden (INSEK).* Dresden: Stadtplanungsamt, Geschäftsbereich Stadtentwicklung.

EEA (2008). *Climate for a transport change. TERM 2007: Indicators tracking transport and the environment in the European Union.* Copenhagen: European Environmental Agency.

EEA (2012). *Passenger car ownership in the EEA*. Copenhagen: European Environment Agency.

Eluru, N., I. N. Sener, C. R. Bhat, R. M. Pendyala, and K. W. Axhausen (2009). Understanding residential mobility. *Transportation Research Record: Journal of the Transportation Research Board 2133*, 64–74.

Energieprognose (2010). *Die Entwicklung der Energiemärkte bis 2030 - Energieprognose 2009.* Stuttgart, Essen, Mannheim: Institut für Energiewirtschaft und rationelle Energieanwendung, Rheinisch-Westfälisches Institut für Wirtschaftsforschung, Zentrum für Europäische Wirtschaftsforschung.

Frondel, M. and C. Vance (2009). Do high oil prices matter? Evidence on the mobility behavior of German households. *Environmental and Resource Economics 43*(1), 81–94.

ifmo (2008). *Mobilität 2025 - Der Einfluss von Einkommen, Mobilitätskosten und Demografie.* Bonn: Institut für Mobilitätsforschung.

Just, T. (2004). *Demographic developments will not spare the public infrastructure* (Current Issues Demography Special ed.). Deutsche Bank Research.

Karlaftis, M. and J. Golias (2002). Automobile ownership, households without automobiles, and urban traffic parameters: Are they related? *Transportation Research Record: Journal of the Transportation Research Board 1792*, 29–35.

KBA (2012). *Halter - Zeitreihen 2002 bis 2011.* Flensburg: Kraftfahrtbundesamt.

King, G., M. Tomz, and J. Wittenberg (2000). Making the most of statistical analyses: Improving interpretation and presentation. *American Journal of Political Science 44*, 341–355.

Kitamura, R. (2009). A dynamic model system of household car ownership, trip generation, and modal split: Model development and simulation experiment. *Transportation 36*(6), 711–732.

Limbourg, M. (2004). *Mobilität im Alter: Probleme und Perspektiven* (Current Issues Demography Special ed.). Deutsche Bank Research.

Long, S. J. and J. Freese (2006). *Regression models for categorical dependent variables using Stata* (2nd ed.). Stata Press.

Matas, A., J. L. Raymond, and J. L. Roig (2009). Car ownership and access to jobs in Spain. *Transportation Research Part A: Policy and Practice 43*(6), 607–617.

Matas, A. and J. L. L. Raymond (2008). Changes in the structure of car ownership in Spain. *Transportation Research Part A: Policy and Practice 42*(1), 187–202.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers of Econometrics*, pp. 105–142. Academic Press.

MOP (2011). *German Mobility Panel.* Karlsruhe Institute of Technology.

Potoglou, D. (2008). Vehicle-type choice and neighbourhood characteristics: An empirical study of Hamilton, Canada. *Transportation Research Part D: Transport and Environment 13*(3), 177–186.

Prettenthaler, F. E. and K. W. Steininger (1999). From ownership to service use lifestyle: the potential of car sharing. *Ecological Economics 28*(3), 443–453.

Raphael, S. and L. Rice (2002). Car ownership, employment, and earnings. *Journal of Urban Economics 52*(1), 109–130.

Sachverständigenrat (2011). *Verantwortung für Europa wahrnehmen.* Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Lage.

Tomz, M., J. Wittenberg, and G. King (2003). Clarify: Software for interpreting and presenting statistical results. *Journal of Statistical Software 8*(1), 1–30.

Vance, C. and M. Peistrup (2011). She's got a ticket to ride - gender and public transit passes. *Transportation forthcoming*, 1–15.

Whelan, G. (2007). Modelling car ownership in Great Britain. *Transportation Research Part A: Policy and Practice 41*(3), 205–219.

Zumkeller, D., B. Chlond, and W. Manz (2004). Infrastructure development in Germany under stagnating demand conditions: A new paradigm? *Transportation Research Record: Journal of the Transportation Research Board 1864*, 121–128.