RUHR ECONOMIC PAPERS

Manuel Frondel
Colin Vance

Rarely Enjoyed?

A Count Data Analysis of Ridership in Germany's Public Transport

RUHR UNIVERSITÄT BOCHUM RUB rwi

#209

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 #209

Responsible Editor: Christoph M. Schmidt

ISSN 1864-4872 (online) – ISBN 978-3-86788-240-8

Ruhr Economic Papers #209

Manuel Frondel and Colin Vance

Rarely Enjoyed?

A Count Data Analysis of Ridership in Germany's Public Transport

RUHR UNIVERSITÄT BOCHUM

Bibliografische Informationen der Deutschen Nationalbibliothek

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

ISSN 1864-4872 (online)
ISBN 978-3-86788-240-8

Manuel Frondel and Colin Vance[1]

Rarely Enjoyed?

A Count Data Analysis of Ridership in Germany's Public Transport

Abstract

Focusing on adult members of German households, this paper investigates the determinants of public transit ridership with the aim of quantifying the effects of fuel prices, fares, person-level attributes, and characteristics of the transit system on transport counts over a five-day week. The reliance on individual data raises several conceptual and empirical issues, the most fundamental of which is the large proportion of zero values in transit counts. To accommodate this feature of the data, we employ modeling procedures referred to as zero-inflated models (ZIMs), which order observations into two latent regimes defined by whether the individual never uses public transport. Our estimates reveal fuel prices to have a positive and substantial influence on transit ridership, though there is no evidence for a statistically significant impact of the fare.

JEL Classification: D13, Q41

Keywords: Transit ridership; pricing policy; Zero-Inflated Models; household data

Oktober 2010

1 Manuel Frondel, RWI and Ruhr-Universität Bochum; Colin Vance, RWI and Jacobs University Bremen. *– We are very grateful for invaluable comments and suggestions by Christoph M. Schmidt, as well as two anonymous referees. We also would like to thank Matthias Peistrup for assistance with the data preparation.* *– All correspondence to: Manuel Frondel, RWI, Hohenzollernstr. 1-3, 45128 Essen, Germany, E-Mail: manuel.frondel@rwi-essen.de.*

1 Introduction

In Germany, as well as in many other industrial countries, substantial shares of the population seldomly, if ever, use public transport systems. For example, many citizens of rural areas suffer from the unavailability, or impracticability, of public transport systems, constraining them to the exclusive usage of cars for commuting. According to a nationwide survey conducted in 2008, some 44% of the German population report never using public transportation, with an additional 33% using it less than once per week (MiD, 2008). As a consequence, the outcome of count processes reflecting transit ridership is typically characterized by an overwhelmingly large number of zero counts.

A basic question emerging in public transport patronage is whether a zero count indicates an individual who never uses public transport, or alternatively the chance event that the individual does not use public transport during the sampling period. The so-called zero-inflated models (ZIMs) take particular account of this distinction by ordering observations into two latent regimes defined by whether an individual never uses public transport, and are therefore perfectly appropriate in this instance, not least because zero-inflated modeling procedures were developed to cope with the preponderance of zero counts. By contrast, the classical count data models, such as the Poisson (PRM) and the negative binomial regression model (NBRM), rest on the assumption that the probability for a positive value of public transit usage is non-zero for every observation. With the exception of a handful of studies that mostly focus on accident rates (Shanker et al. , 1997, Chin and Quddus, 2003), the feature of unobservable membership in either of two groups – the group of never-users and its complement – has rarely been addressed in the transportation literature.

Using household survey data from Germany, this paper applies zero-inflated modeling approaches to the issue of public transport patronage, focusing on the determinants of adult transit ridership. Specifically, we aim at quantifying the effects of fuel prices and fares on public transport counts over a five-day week, while controlling for the effects of person-level attributes and characteristics of the transit system. A large empirical literature has emerged to address this issue, but, as with the literature on fuel price elasticities for automobile travel (GRAHAM and GLAISTER, 2002), elasticity estimates for transit vary widely. Based on a comprehensive survey of the literature, LITMAN (2004) finds short-run elasticity estimates with respect to the fare varying between -0.2 and -0.5, with a subsequent meta-analysis by HOLMGREN (2007) finding the short-run elasticity to reach -0.75 for Europe. The cross-price elasticity estimates of fuel prices tend to be lower, but also highly variable, ranging from 0.05 (LITMAN, 2004) to 0.4 (HOLMGREN, 2007).

The most important factor accounting for the differences in transit estimates is, according to NIJKAMP and PEPPING (1998), whether aggregate or disaggregate data is used. As aggregate data makes no allowance for the large variation of individual choices made in specific circumstances, it typically yields less precise estimates that are, moreover, more subject to bias. To date, however, 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, our analysis is predicated on the notion that transit use is an individual decision, albeit one that is dependent on intra-household allocation processes. 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 mode choice, distance traveled, and other aspects of mobility behavior (e.g. PICKUP, 1985, TURNER and NIEMEIER, 1997, KAYSER, 2000, FRONDEL and VANCE, 2009, 2010, and VANCE and HEDEL, 2007).

Contrasting with some other studies that use count data (e.g. SHANKER et al. , 1997, and PETERS, VANCE, 2010), among the key findings of our analysis is that ZIMs have superior predictive accuracy over the PRM and NBRM, and thus may serve as the method of choice when the aim is to predict trip frequency for modes that a large fraction of the population never uses. The model estimates reveal fuel prices to have a positive and substantial influence on transit ridership, though we find no evidence for a statistically significant impact of the fare. In this regard, our findings highlight the importance of referencing both the coefficients and associated marginal effects when interpreting the results. Due to the non-linearity of the model, the magnitude and significance level of these estimates can vary markedly from one another, requiring that inferences be cast specifically according to whether the marginal effects or coefficients are in question.

The following section presents contextual information on public transit policy in Germany. Section 3 describes the data base used for estimating individual mobility behavior of adults. Section 4 explicates the econometric methods and model specifications, followed by the presentation and interpretation of the results in Section 5. The last section summarizes and concludes.

2 Policy Context

Alongside other measures, such as land-use planning and efficiency improvements, the promotion of public transit is regarded by Germany's Federal Environmental Agency as an integral component to reducing emissions from transport (UBA, 2010). According to figures compiled by BASSETT et al. (2008), the percentage of trips taken by public transit in Germany is 8%, which, while considerably higher than the 2% share for the US, is on par or slightly lower than that of many of its European neighbors, including the UK (9%), Sweden (11%), Switzerland (12%), and Spain (12%). Moreover, the share of total travel undertaken with transit has been remarkably stable over the past decades, hovering around 8.7% since the early 1990s, compared with slightly over 80% by car (BMVBS, 2006). Thus, a persistent question confronting transport planners is what measures can be undertaken to increase ridership given the fact that the demand for car ownership has grown substantially in the last decades, with the number of registered cars per resident increasing by almost 25% between 1990 and 2005 (BUEHLER et al., 2009).

While sociodemographic and service attributes are frequently cited as important determinants of transit ridership, knowledge about the magnitude of these determinants remains rudimentary. Further complicating an appraisal of transit demand and its future trajectory in Germany are major sociodemographic changes currently underway that could dramatically affect the composition of mode choice. According to an energy forecast recently commissioned by the German government, the population is expected to decrease by 3% between 2007 and 2030, from 82.3 to 79.7 million residents (BMWI, 2010). Despite this, the forecast expects an overall increase in individual transport demand owing to more single and dual-person households; by 2030, the total number of households is projected to increase by nearly 6% from 39.7 to 42.0 million.

These trends will be paralleled by an increasingly older age structure of the German population, as well as by a likely increase in the share of women in the pool of license holders and in the labor force, with the latter having already risen from 55.1% in 1994 to 59.2% in 2004 (EUROSTAT, 2006). While several studies have suggested that these changes will have profound consequences for transport demand in Germany (LIMBOURG, 1999, JUST, 2004, ZUMKELLER, CHLOND, and MANZ, 2004), both the contemporary and future impacts are largely speculative, since there have been few attempts to quantify how the underlying variables affect travel behavior at the individual level.

Of particular relevance in this regard is the impact of fuel prices and fares on the demand for public transit. In 1999, the German government introduced an eco-tax that incrementally increased taxes on motor fuel over a five-year period, resulting in fuel taxes amounting to as much as two thirds of the gross prices at the gasoline station. In contrast, fuel used for public transportation is taxed at one half the standard rate (KOHLHAAS, 2000). While such tax-raising policies would conceivably increase the demand for transit ridership, the actual impacts have been difficult to gauge due to a dearth of information on cross-price elasticites from Germany. Given the highly variable shares of public transport in the modal split across countries, caution is warranted in extrapolating the influence of pricing and service levels from one country to another (HENSHER, 2008).

3 Data

The main data source used in this research is drawn from the German Mobility Panel (MOP 2010), an ongoing travel survey that is organized in waves, each comprising a group of households whose members are surveyed for a period of one week over each of three years. Our data set includes twelve waves of the panel, spanning 1996 through 2007, and is limited to adult individuals who are at least 18 years old. In total, our data set contains 8,577 individuals, 2,904 of whom participated in one year of the survey with the remaining 5,673 participating in two or three years. For this latter group, we randomly selected a single year for inclusion in the data set to avoid repeat observations on the same individual. In this regard, it bears noting that the use of public transit and the variables that determine it vary little or not at all over the three years of the survey, thereby allowing us to pool the data in model estimation due to the relative homogeneity of the data over this short period of time.

Individuals that participate in the survey are requested to fill out a questionnaire eliciting general household information and person-related characteristics, including zip code of residence, gender, age, employment status and relevant aspects of everyday travel behavior. In addition to this general survey, the MOP includes a separate survey focusing specifically on vehicle travel among a 50% sub-sample of randomly selected car-owning households. These households are drawn from the larger MOP-data set used in the present analysis. This so-called "tank survey" takes place over a roughly six-week period, during which time respondents record sundry automobile-related information, including the price paid for fuel (Table 1).

As this variable is a potentially important determinant of transit pass ownership, it was linked with the larger sample of households in the MOP by using a Geographic Information System to create a coverage of spatially interpolated fuel prices (in real terms) for all of Germany. The coverage was then overlaid onto a map of household locations in the MOP data, thereby allowing for each household to be ascribed the locally prevailing fuel price. This process was repeated for each year of the data, yielding a data set of fuel prices that varies over space and time. A crude accuracy assessment of

Table 1: Descriptive Statistics of the Discrete and Continuous Variables

Variable Name	Variable Definition	Mean	Std. Dev.
y	# public transits during 5-day week	1.47	3.30
real fuel price	Real fuel price in € per liter	1.01	0.12
fare	Real fare for a monthly ticket in €	32.40	5.88
public transit density	Density of the public transit service in 1,000 service kilometers divided by areal unit in squared kilometers	35.44	51.05
age	Age of adult	48.38	16.00
income	Real net monthly household income in 1,000 €	2.363	0.822
# children < 18	Number of children younger than 18	0.27	0.63
minutes	Walking time to the nearest public transportation stop in minutes	5.75	4.95

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 1996-2007 time interval (see FRONDEL and VANCE, 2010).

In addition to fuel prices, another important cost determinant of transit use is the fare. Data on this variable was obtained by an internet-based survey that retrieved the price for a single-trip and monthly ticket for each of the 90 regional transit authorities in Germany. Each household was then assigned the fare of the transit authority to which it belongs. Fares, as well as fuel prices, were converted into real terms using a consumer price index published by the German Statistical Office (DESTATIS, 2010).[1] From the same source, we also obtained a variable measuring the density of transit service that was merged with the MOP data. This variable is constructed by dividing the milage of transit travel for all modes by the area of the transit zone.

The remaining suite of variables selected for inclusion in the model measures the

[1]In the analysis that follows we use the monthly fare price, noting that our qualitative findings do not change when using the trip-based fare.

individual and household-level attributes that are hypothesized to influence the allocation of travel expenditures in maximizing utility. Variable definitions and descriptive statistics are presented in Tables 1 and 2. As many of these variables could either positively or negatively affect the use of public transit, it is not always possible to state a priori which effects are expected to prevail.

Table 2: Descriptive Statistics of the Binary Variables

Variable Name	Variable Definition	Mean
high school diploma	1 if person has a high school diploma	0.35
license	1 if person has a driver license	0.87
employed	1 if person is employed in a full-time or part-time job	0.54
female	1 if person is female	0.52
big city	1 if household resides in a large city	0.42
parking space at home	1 if household has a private parking space or garage	0.76
parking space at work	1 if household has a parking space at work	0.37
direct public transit to work	1 if there is a direct transit connection to work	0.16
rail transit	1 if the nearest public transportation stop is serviced by rail transit	0.13
enoughcars	1 if number of cars in a household is at least equal to the number of licensed drivers	0.54

Negative signs are expected for the variables that either increase the opportunity- and/or transaction costs of transit use or decrease these costs for automobile use, including the distance to the nearest transit stop, the fare ticket price, and dummies indicating driver-license holders and households in which the number of cars equals or exceeds the number of licensed drivers. Positive signs are expected for variables that are indicative of the availability or quality of public transit, including public transit service density and the dummies for residence in a large city and for rail transit service at the nearest transit stop. Higher fuel prices are also expected to have a positive effect, as they encourage the substitution of public transit for private car travel. The inclusion of time dummies in the model was also explored, but as these were individually and jointly insignificant, they were excluded from the final specification.

While the included explanatory variables afford reasonably broad coverage of the determinants of transit use, we cannot completely rule out the possibility that they are correlated with additional unobserved factors that impact travel. 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 transit use and residential choice are jointly determined, implying that the coefficients of the urban form variables are partially picking up the effects of neighborhood preferences. ELURU and colleagues (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 transit itself and for competing modes, such as the level of security and regional congestion, which may be correlated with some of our explanatory variables. We consequently abstain from making claims about causality, and instead apply a descriptive interpretation to the estimates.

4 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 zero values in the data. Slightly less than 75% of the adult individuals in the estimation sample do not use public transport systems during a given week (see Table B1 in the Appendix B) and for whom the observation on transit counts is consequently recorded as zero. To accommodate this feature of the data, we employ modeling procedures referred to as zero-inflated models (ZIMs). There are two common ZIMs, referred to as the zero-inflated Poisson (ZIP) and the zero-inflated negative binomial (ZINB) models, both of which are generalizations of the Poisson regression model and negative binomial regression model. As ZIMs build on these classical count data models, we start with a brief description of the Poisson and the negative binomial regression models and highlight the differences between these classical and the zero-inflated models.

4.1 Classical Count Data Models

Fundamental to the understanding of count data models is the univariate Poisson distribution, which relates the mean $E(y) = \lambda > 0$ and the probability of observing any count $y = 0, 1, 2, ...$ by

$$Pr(y|\lambda) = \frac{\lambda^y \exp\{-\lambda\}}{y!}. \tag{1}$$

An inherent characteristic of this distribution, known as equidispersion, is that the variance $Var(y)$ is identical to the expected value: $Var(y) = E(y) = \lambda$. In practice, though, the variance of many count variables is greater than their mean, a fact that is called overdispersion.

The Poisson regression model (PRM) extends the Poisson distribution by allowing for each observation i to have a different mean λ_i. The most common parameterization of the idiosyncratic means is the loglinear model (GREENE 2003: 740):

$$\lambda_i = E(y_i|\mathbf{x}_i) = \exp\{\mathbf{x}_i^T \boldsymbol{\beta}\}, \tag{2}$$

where $\boldsymbol{\beta}$ is a parameter vector to be estimated and observed heterogeneity is incorporated by the vector $\mathbf{x}_i$, which includes the observable characteristics that affect the individual number of counts y_i. Note that taking the exponential of $\mathbf{x}_i^T \boldsymbol{\beta}$ ensures that the expected value λ_i is positive, which is a natural property of count data.

While being a useful starting point, the PRM suffers from at least four shortcomings. First, it underestimates the number of zero counts, as can be seen from our empirical example presented in Section 5. Second, the standard errors pertaining to the PRM estimates are biased downward, resulting in spuriously large z- and small p-values (CAMERON, TRIVEDI, 1986:31). More general failings are, third, that the PRM does not fit to real data in the case of overdispersion, i. e. if $Var(y) > E(y)$. Fourth, the PRM does not account for unobserved heterogeneity.

These failures are circumvented by the negative binomial regression model (NBRM), which addresses the last point by adding an error term ε_i that is assumed to be uncorrelated with the factors included in $\mathbf{x}_i$:

$$\tilde{\lambda}_i = E(y_i|\mathbf{x}_i) = E(\exp\{\mathbf{x}_i^T\boldsymbol{\beta} + \varepsilon_i\}) = \exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}E(\delta_i), \tag{3}$$

where $\delta_i := \exp\{\varepsilon_i\}$. By assuming that $E(\delta_i) = 1$, which corresponds to the assumption $E(\varepsilon_i) = 0$ of the classical linear regression approach, the model is identified. From this assumption, it follows that in the NBRM the conditional distribution of the counts y_i given $\tilde{\lambda}_i$ is Poisson, that is, y_i obeys equation (1) with $\lambda_i = \tilde{\lambda}_i$. Without altering the conditional mean, the NBRM improves upon the underprediction of zero counts in the PRM by increasing the conditional variance. In contrast, zero-inflated models (ZIMs) such as the ZIPM, which was introduced by LAMBERT (1992), change the mean structure, thereby also increasing the probability of zero counts.

4.2 Zero-inflated Models

Zero-inflated models assume that there are two latent groups, for which membership is unobservable: the *Always-Zero Group* A, for which

$$Pr(y_i = 0|A_i = 1, \mathbf{x}_i) = 1, \tag{4}$$

where $A_i = 1$ designates membership of individual i in Group A, and $A_i = 0$ indicates membership in the complementary group. Group membership is a binary outcome that can be modeled using standard logit or probit estimation procedures:

$$\psi_i := Pr(A_i = 1|\mathbf{z}_i) = F(\mathbf{z}_i^T\boldsymbol{\gamma}), \tag{5}$$

where ψ_i is the probability of being in Group A, $F(.)$ stands for the cumulative distribution function $\Phi(.)$ or $\Lambda(.)$ of the normal or logistic distribution, respectively, $\boldsymbol{\gamma}$ is a parameter vector to be estimated, and vector $\mathbf{z}_i$ includes variables that inflate the number of zero counts. Hence, they are referred to as *inflation* variables and (5) is called the *inflation* equation. The vector of inflation variables $\mathbf{z}_i$ may differ from the determinants $\mathbf{x}_i$ of the number of counts y_i, but may also be identical to $\mathbf{x}_i$.

If we knew probability ψ_i, the probability of a zero count could be calculated as follows:

$$\begin{aligned} Pr(y_i = 0|\mathbf{x}_i, \mathbf{z}_i) &= Pr(A_i = 1|\mathbf{z}_i) \cdot Pr(y_i = 0|A_i = 1, \mathbf{x}_i) \\ &\quad + Pr(A_i = 0|\mathbf{z}_i) \cdot Pr(y_i = 0|A_i = 0, \mathbf{x}_i) \\ &= \psi_i \cdot 1 + (1 - \psi_i) \cdot Pr(y_i = 0|A_i = 0, \mathbf{x}_i), \end{aligned} \tag{6}$$

as $Pr(A_i = 0|\mathbf{z}_i) = 1 - \psi_i$ and the probability of zero counts is 100% in the *Always-Zero Group* A ($Pr(y_i = 0|A_i = 1, \mathbf{x}_i) = 1$. For outcomes $y_i = k > 0$,

$$\begin{aligned} Pr(y_i = k|\mathbf{x}_i, \mathbf{z}_i) &= \psi_i \cdot 0 + (1 - \psi_i) \cdot Pr(y_i = k|A_i = 0, \mathbf{x}_i) \\ &= (1 - \psi_i) \cdot Pr(y_i = k|A_i = 0, \mathbf{x}_i), \end{aligned} \tag{7}$$

where, by definition, the probability of non-zero counts is 0% in Group A: $Pr(y_i = k|A_i = 1, \mathbf{x}_i) = 0$. The probabilities $Pr(y_i = 0|A_i = 0, \mathbf{x}_i)$ and $Pr(y_i = k|A_i = 0, \mathbf{x}_i)$ are the outcomes of the PRM or NBRM in case of the ZIP or ZINBM, respectively.

On the basis of these probability expressions, the unknown parameter vectors $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ can be estimated using maximum-likelihood methods. For instance, the loglikelihood function of the ZIP reads:

$$\ln L_{ZIP} = \sum_{y_i=0} \log[\psi_i + (1 - \psi_i) \cdot \exp\{-\lambda_i\}] + \sum_{y_i>0} [y_i \log(\lambda_i) - \lambda_i - \log(y_i!)] \cdot \log(1 - \psi_i),$$

where $\lambda_i := \exp\{\mathbf{x}_i^T \boldsymbol{\beta}\}$ and $\psi_i := F(\mathbf{z}_i^T \boldsymbol{\gamma})$. It bears noting that one cannot separately estimate the parameters $\boldsymbol{\gamma}$ in a first step, as we do not know those zero counts that originate from members of Group A. Instead, both parameter vectors, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, have to be estimated simultaneously.

Expected counts are computed in a way similar to that of the probabilities:

$$\begin{aligned} E(y_i|\mathbf{x}_i, \mathbf{z}_i) &= \psi_i \cdot E(y_i|A_i = 1, \mathbf{x}_i) + (1 - \psi_i) \cdot E(y_i|A_i = 0, \mathbf{x}_i) \\ &= \psi_i \cdot 0 + (1 - \psi_i) \cdot \lambda_i = (1 - \psi_i) \cdot \lambda_i, \end{aligned} \tag{8}$$

where for the *Always-Zero Group* A, it is $E(y_i|A_i = 1, \mathbf{x}_i) = 0$ and $E(y_i|A_i = 0, \mathbf{x}_i) = \lambda_i$ for the complementary group, since the PRM and NBRM have the same mean structure. Because $0 \leq \psi_i \leq 1$, where in practice $\psi_i > 0$, the expected value given by (8) will be smaller than λ_i, so that the expected count resulting from ZIMs is generally lower than that of the PRM and NBRM, thereby better fitting to the large number of zero counts in the empirical evidence on transit usage.

5 Empirical Results

Along the lines of the previous section, we estimate both the classical as well as the zero-inflated models and select the most appropriate approach both by comparing the predicted probabilities for the range of public transit counts occuring in practice and employing the hypotheses tests presented in Appendix A. The details of the comparison are reported in Appendix B. While the existence of always-zero observations is ignored by both the PRM and NBRM, the special treatment of this feature by the ZIMs leads us to expect an improvement in the fit due to their employment.

Indeed, the observed frequency for zero counts is perfectly reproduced by both the ZINBM and ZIP model (Table B1 in Appendix B). Relative to the NBRM, the ZINBM also provides for a substantially better fit for a single count, whereas the predictions of the probabilities of 3, 4, and 5 counts are somewhat worse. Therefore, an ultimate decision on whether the ZINBM is superior to the NBR model requires a VUONG test (see Appendix A), whose large positive value of 21.75 for the standard-normal distributed normal test statistic favors the ZINB model. Finally, the probability-by-probability comparison of the ZIPM and ZINBM is clearly in favor of the ZINBM. This conclusion is confirmed by the Likelihood-Ratio test on overdispersion (see Appendix A), for which the test statistic amounts to about 1,679.

Turning to the estimation results of the inflation equation reported in Table 3, among the most important factors that determine the membership in the always-zero group are possession of a driver's license and the existence of at least one car per licensed driver in the household (indicated by *enoughcars* = 1), as well as the availability of parking spaces both at home and at work. Likewise, all of the service attributes – including the availability of a direct transit connection to work, the availability of rail transit near the home, and the density of the local transit system – are statistically significant determinants of the probability that the individual is a non-user of public transit, with negative signs that are consistent with expectations. While income and the fare appear to have no bearing on this probability, both fuel prices and residence in a large city decrease it, as do the sociodemographic attributes indicating females, employed persons, and those with a high school diploma. Finally, age has a nonlinear effect that is initially positive and peaks at an age of 39.

Table 3: Regression Results of the Inflation Equation

	Coeff.s	Robust Std. Errors	Marginal Effects	Robust Std. Errors
female	** -0.242	0.059	** -0.038	0.009
age	** 0.078	0.013	** 0.013	0.002
age squared	** -0.001	0.000	** $-1.2 \cdot 10^{-4}$	$0.2 \cdot 10^{-4}$
employed	-0.171	0.098	* -0.039	0.015
high school diploma	** -0.478	0.067	** -0.082	0.011
license	** 1.173	0.090	** 0.232	0.019
employed×(*parking space at work*)	** 0.752	0.088	** 0.124	0.013
parking space at home	** 0.516	0.081	** 0.091	0.015
enoughcars	** 0.846	0.065	** 0.145	0.010
minutes	** 0.036	0.008	** 0.006	0.001
direct public transit to work	** -0.415	0.092	** -0.078	0.017
big city	** -0.435	0.079	** -0.076	0.013
rail transit	** -0.315	0.097	** -0.056	0.017
# children < 18	** 0.301	0.059	** 0.053	0.009
income	0.062	0.046	0.013	0.007
real fuel price	* -0.565	0.265	* -0.095	0.042
fare	0.005	0.006	0.001	0.001
public transit density	** $-7.4 \cdot 10^{-3}$	$0.8 \cdot 10^{-3}$	** $-1.2 \cdot 10^{-3}$	$-1.2 \cdot 10^{-4}$

Note: * denotes significance at the 5 %-level and ** at the 1 %-level, respectively. Number of observations used in estimation: 8,577.

With respect to the coefficient estimates of the ZINBM reported in Table 4, the majority are statistically significant and have signs that are consistent with intuition. Two notable exceptions are the coefficients on fuel prices and fares: neither appear to be important determinants of the number of public transit trips over the 5-day week. The statistically insignificant impact of fares persists when considering the marginal effect (right-hand panel of Table 4).

Table 4: Estimation Results of the Zero-Inflated Negative Binomial Model (ZINBM)

	Coeff.s	Robust Std. Errors	Marginal Effects	Robust Std. Errors
female	-0.001	0.034	** 0.173	0.050
age	** -0.031	0.007	**-0.086	0.010
age squared	** $2.1 \cdot 10^{-4}$	$0.8 \cdot 10^{-4}$	** $0.7 \cdot 10^{-3}$	$1.1 \cdot 10^{-4}$
employed	** 0.380	0.056	** 0.486	0.084
high school diploma	* 0.096	0.039	** 0.465	0.066
license	** -0.225	0.043	**-1.440	0.131
employed×(*parking space at work*)	** -0.316	0.050	**-0.784	0.068
parking space at home	-0.081	0.047	**-0.500	0.084
enoughcars	** -0.191	0.040	**-0.829	0.062
minutes	*-0.010	0.005	**-0.036	0.007
direct public transit to work	** 0.167	0.049	** 0.535	0.106
big city	* 0.098	0.045	** 0.424	0.070
rail transit	0.083	0.049	** 0.341	0.099
# children < 18	*-0.089	0.037	**-0.305	0.052
income	* -0.063	0.027	**-0.105	0.039
real fuel price	0.262	0.149	** 0.663	0.230
fare	0.001	0.003	-0.003	0.005
public transit density	** $1.5 \cdot 10^{-3}$	$0.4 \cdot 10^{-3}$	** 0.007	0.001

Note: * denotes significance at the 5 %-level and ** at the 1 %-level, respectively.
Number of observations used in estimation: 8,577.

With reference to real fuel prices, however, a discrepancy emerges: the marginal effect is highly significant in this case, and suggests that a 1€ increase in fuel costs increases transit counts by 0.66 trips over the course of a 5-day week. The unreported corresponding elasticity estimate is of roughly the same magnitude at 0.7, but notably higher than the upper-bound elasticity estimate of 0.4 presented by HOLMGREN (2007) based on his meta-analysis of US data. Fuel cost increases may thus be an effective instrument for encouraging transit ridership in Germany.

That the effect of the fare does not mirror that of fuel prices might be attributed to the fact that the majority of public transit users buy lump-sum tickets that allow for the unlimited use of the transit system during their validation period. To explore whether insignificant effects of the fare remain under alternative specifications, we estimated models that included interaction terms and calculated the interaction effects, whose derivation is presented in Appendix C. These specifications accommodated the possibility of differential effects of fuel and fare prices by income level, residential location, and car availability (LITMAN, 2004). In all cases, the interaction effects were found to be statistically insignificant.

As with the fuel price, stark differences between the coefficient estimates and the marginal effects are seen for the dummy variables indicating females, a private parking space, and the existence of a rail transit stop near the home. Rail transit, for example, which tends to afford greater speed and comfort, would be expected to positively affect public transit use. While the coefficient estimate is statistically insignificant, the estimate of the marginal effect is highly precise, and suggests that this service attribute increases the number of transit counts by about 0.34.

Likewise, being female seems to be irrelevant when focusing on coefficients, but, in fact, increases the number of transit counts by about 0.17, as is given by the marginal effect. This result mirrors an estimate reported by VANCE and IOVANNA (2007), who focus on the role of gender in determining car use. These authors find that women have a lower probability of using the car than men and drive less when they do. Moreover, FRONDEL and VANCE (2010) and VANCE and IOVANNA (2007) both uncover an equalizing effect of employment status, the presence of children, and the distance to the transit stop, with all three variables mitigating the negative effect of the female gender dummy on the likelihood of car use. A similar analysis was undertaken here by creating interaction terms with the female dummy. As in the case with the fare and fuel prices, none of the gender interactions were found to be significant. This absence of differential effects implies that many of the levers available to policy-makers for influencing transit patronage, such as fuel prices and the siting of transit stops, are likely to have a roughly uniform impact among men and women.

With respect to the remaining coefficients, age is seen to have a non-linear effect, which is initially negative up to about an age of 72 after which it becomes positive. The dummies for employed persons and those with a high school diploma both have positive signs and are roughly the same magnitude, at least with respect to the marginal effects, suggesting that these individuals have transit counts that are about 0.49 higher than their counterparts. Likewise, those living in a big city and with a direct transit connection from home to work are also more frequent users of public transport, as are individuals who live in regions with a denser transit network. Consistent with expectations, factors that increase the costs of transit use or that decrease the cost of car use have negative effects. These include the dummies indicating license holders, those with a parking space at work, and those who live in households with at least as many cars as licensed drivers.

6 Conclusion

In Germany, the promotion of public transit use is a central policy tool in the mitigation of pollution, congestion, and other automobile-caused externalities. Despite Germany's relative success in capping emissions from transport, which rose by 1% between 1990 and 2005 compared to a 26% increase in the European Union (EEA, 2007), public transit use has by most measures stagnated or been on the decline. Between 1994 and 2003, the percentage of trips traveled by transit dropped by 1%, contrasted by a 16% increase in motor vehicle trips (DESTATIS, 2006). To counter this trend, the country's transport ministry has placed a high priority on improving the competitive position of public transit relative to the automobile (BMVBS, 2009).

An important step in this endeavor is to identify the economic and structural factors that draw or repel potential transit customers, thereby enabling the design of measures to increase ridership among those segments of the population where the scope for mode switching is greatest. From a planning perspective, one particularly important factor is the responsiveness of transit riders to both gasoline prices and fares.

This paper has investigated this issue with an analysis of the determinants of weekly transit usage by drawing on household survey data from Germany.

To our knowledge, this is the first study to estimate the effects of fares and fuel prices, as well as socioeconomic and geographic determinants, on the basis of individual-level data. Although necessarily neglected in studies on the effects of fuel and fare prices using aggregated data, the discrete decision to occasionally or regularly use public transit system appears to be of particular relevance in the analysis of individual data, as fuel price peaks may trigger a reduction of car use, thereby fostering an occasional, temporary, or even permanent switch to public transit.

We have addressed this issue by employing zero-inflated modeling approaches, which is particularly appropriate when the question at hand requires distinguishing between those who never use public transit from those who have some non-zero probability of a positive trip count. Our estimates suggest that a 1€ increase in fuel prices – that is, a rise in gasoline prices by about two thirds – increases transit use by almost 0.7 trips over a week, an effect that is statistically significant at the 1% level. Somewhat unexpectedly, we find that the effect of the fare, by contrast, is not significantly different from zero, even when allowing for differential effects according to residential location, car ownership, and the income of the household.

Taken together, these findings suggest that fuel prices are a more effective lever than fares for influencing transit ridership, partly validating STORCHMANN's (2001) conclusion that higher fuel prices only increase peak-hour transit use, but not leisure or off-peak transit. Moreover, given the relatively large fuel price elasticities found to prevail in Germany (FRONDEL et al. , 2008), as well as the fact that revenue from the eco-tax is employed to stabilize the contributions to the country's pension insurance system, increasing fuel taxes appears to afford promise for tackling demographic and ecological problems simultaneously. This is all the more relevant as fuel taxes may be raised centrally by the government, whereas the amount of fares is a decentralized decision of local authorities and (semi-)private public transit suppliers.

As this is one of the few studies to be conducted on this topic using micro-level data in a European context, it would be of interest to see whether the qualitative findings presented here are corroborated by studies using other data sets from within Germany and other European countries. A particularly useful line of inquiry would focus on distinguishing short- and long-run price responsiveness using micro-level data over a longer time interval, which is not subject to the aggregation problems that commonly afflict regional-level temporal studies of transit use. Data constraints precluded such an analysis in the present study, but it is one that would further facilitate the formulation of pricing strategies to encourage transit use.

Appendix A: Hypotheses Tests

A basic assumption of the PRM is equidispersion, i. e. the conditional mean equals the conditional variance:

$$Var(y_i|\mathbf{x}_i) = E(y_i|\mathbf{x}_i) = \lambda_i. \tag{9}$$

This rarely fulfilled assumption is relaxed in the NBRM, for which a variety of alternatives to the constant-variance function given by (9) exist (see CAMERON and TRIVEDI, 1986). The most commonly used generalization is

$$Var(y_i|\mathbf{x}_i) = \lambda_i + \alpha\lambda_i^2. \tag{10}$$

Testing Overdispersion

Equation (10) suggests examining the null hypothesis $H_0 : \alpha = 0$ in order to test for overdispersion. If the null holds true, equidispersion according to (9) prevails and the NBRM collapses to the PRM. It bears noting that testing the null requires procedures other than the typical symmetric t-tests, as α must be non-negative. Instead, a Likelihood-Ratio (LR) test can be employed, where the test statistic follows a χ^2-distribution and is computed in the usual manner:

$$LR = 2 \cdot (\ln L_{NBRM} - \ln L_{PRM}). \tag{11}$$

$\ln L_{NBRM}$ and $\ln L_{PRM}$ denote the Loglikelihood functions of the NBRM and PRM, respectively. The significance level of the test has to be adjusted to account for the truncated sampling distribution of $\hat{\alpha}$.

VUONG Test of Non-Nested Models

Neither the NBRM is nested in the ZINBM, nor is the PRM nested in the ZIPM, as is pointed out by GREENE (1995). While the ZINBM, for instance, would collapse to the NBRM if ψ_i were identical to zero for all observations i, this equality cannot hold in general and is, specifically, not fulfilled for $\boldsymbol{\gamma} = \mathbf{0}$, as $\psi_i = F(\mathbf{z}_i^T \mathbf{0}) = 0.5$. To test the superiority of the ZINBM over the NBRM, as well as of the ZIPM over the PRM, GREENE consequently suggests using a test specified by VUONG (1989:319) for non-nested models.

This test is based on the asymptotically normal distributed Vuong test statistic given by

$$V := \frac{\bar{m}}{s_m/\sqrt{N}}, \tag{12}$$

where $\bar{m}$ and s_m designates the mean and standard deviation of the logged relationship of the predicted probabilities, m_i, obtained from two Models 1 and 2, respectively:

$$m_i := \ln\{\frac{\widehat{P_1}(y_i|\mathbf{x}_i)}{\widehat{P_2}(y_i|\mathbf{x}_i)}\}. \tag{13}$$

The Vuong test examines the null hypothesis $H_0 : E(m_i) = 0$. Large positive values of V that exceed the well-known critical value of 1.96 of the normal distribution favor Model 1, whereas negative values of V below the critical value of -1.96 are supportive of Model 2.

Appendix B: Comparison of Competing Models

Beginning the discussion with the PRM, our empirical example is another confirmation for the fact that this most basic model typically underestimates the number of zero counts: While 74% of the adult individuals in the estimation sample are observed not to use public transport systems during a given week, the PRM predicts a markedly lower probability of 40% for this outcome. Conversely, the PRM drastically overestimates the probability for a single use and also overshoots for two to five transit counts a week.

The accordance of the observed frequencies and the predictions gleaned from the NBRM is clearly superior to the PRM, particularly for the predicted non-use of public transit systems. The superiority of the NBRM over the PRM is additionally confirmed by the Likelihood Ratio test on overdispersion described in Appendix A, for which the chi-squared test statistic amounts to 16,000.

Table B1: Comparison of Observed Frequencies with Predicted Probabilities Resulting from Various Count Data Models

	Observed Frequencies	ZINBM Predictions	NBRM Predictions	ZIPM Predictions	PRM Predictions
$\hat{P}(y_i = 0)$:	0.740	0.740	0.721	0.740	0.400
$\hat{P}(y_i = 1)$:	0.028	0.034	0.103	0.011	0.280
$\hat{P}(y_i = 2)$:	0.065	0.036	0.048	0.022	0.144
$\hat{P}(y_i = 3)$:	0.019	0.034	0.029	0.033	0.073
$\hat{P}(y_i = 4)$:	0.025	0.030	0.019	0.038	0.039
$\hat{P}(y_i = 5)$:	0.013	0.025	0.013	0.038	0.023
$\hat{P}(y_i = 6)$:	0.019	0.020	0.001	0.033	0.014
$\hat{P}(y_i = 7)$:	0.012	0.016	0.008	0.026	0.009
$\hat{P}(y_i = 8)$:	0.013	0.013	0.006	0.020	0.006
$\hat{P}(y_i = 9)$:	0.014	0.010	0.005	0.014	0.004

Appendix C: Marginal and Interaction Effects

Given that $E(y_i|\mathbf{x}_i, \mathbf{z}_i) = (1 - \psi_i) \cdot \lambda_i$ for the zero-inflated model, the marginal effects can be readily calculated. For the case that the inflation regression is based on a logit model, i. e. $\psi_i = \Lambda(\mathbf{z}_i^T \boldsymbol{\gamma})$, where $\Lambda(u) := 1/(1 + \exp\{-u\})$ is the logistic function whose derivative is given by $\Lambda'(u) = \Lambda(u)(1-\Lambda(u))$, a marginal change in variable x_k included in both **x** and **z** yields the following variation of the expected counts:

$$\begin{aligned} \frac{\partial E}{\partial x_k} &= (1 - \psi_i) \cdot \lambda_i \cdot \beta_k - \psi_i \cdot (1 - \psi_i) \cdot \lambda_i \cdot \gamma_k \\ &= (1 - \psi_i) \cdot \lambda_i \cdot (\beta_k - \psi_i \cdot \gamma_k) = E(y_i|\mathbf{x}_i, \mathbf{z}_i) \cdot (\beta_k - \psi_i \cdot \gamma_k), \end{aligned} \tag{14}$$

as $\frac{\partial \lambda_i}{\partial x_k} = \frac{\partial}{\partial x_k}(\exp\{\mathbf{x}_i^T \boldsymbol{\beta}\}) = \lambda_i \cdot \beta_k$. This marginal effect collapses to

$$\frac{\partial E}{\partial x_k} = E(y_i|\mathbf{x}_i, \mathbf{z}_i) \cdot \beta_k \tag{15}$$

if x_k is not included in **z**, that is, if $\gamma_k = 0$.

For the case that the inflation regression is based on a probit model, i. e. $\psi_i = \Phi_i := \Phi(\mathbf{z}_i^T \boldsymbol{\gamma})$, where $\Phi(u)$ denotes the cumulative standard normal distribution and

$\Phi'(u) = \phi(u)$ designates the density function of the standard normal distribution, the marginal effect reads:

$$\frac{\partial E}{\partial x_k} = (1 - \Phi_i) \cdot \lambda_i \cdot \beta_k - \phi_i \cdot \lambda_i \cdot \gamma_k = E(y_i|\mathbf{x}_i, \mathbf{z}_i) \cdot \beta_k - \phi_i \cdot \lambda_i \cdot \gamma_k, \tag{16}$$

with $\phi_i := \phi(\mathbf{z}_i^T \boldsymbol{\gamma})$. If x_k is not included in $\mathbf{z}$, i. e. $\gamma_k = 0$, the marginal effect given by (16) collapses to formula (15). The marginal effects are generally calculated at the mean of the regressors and can be requested in the output of most statistical software packages.

Given the non-linearity of the ZIM, the formulas are a bit more complicated when the model includes interaction terms. To explore whether the effect of an explanatory variable z_1 on the expected value $E[y]$ of the dependent variable y depends on the size of another explanatory variable z_2, it is necessary to estimate the interaction effect given by the second derivative $\frac{\partial^2 E[y]}{\partial z_2 \partial z_1}$. In this section, we follow FRONDEL and VANCE (2010) and derive general formulae for the interaction effects resulting from ZIMs if (a) z_1 and z_2 are both continuous variables, (b) z_1 is continuous, while z_2 is a dummy variable, and (c) both are dummy variables.

To this end, we depart from the expected value (8),

$$E := E[y|z_1, z_2, \mathbf{w}] = [1 - F(u)] \cdot \exp\{u\} = [1 - F(u)] \cdot \lambda(u), \tag{17}$$

where $u := \gamma_1 z_1 + \gamma_2 z_2 + \gamma_{12} z_1 z_2 + \mathbf{w}^T \boldsymbol{\gamma}$, and vector $\mathbf{w}$ excludes z_1 and z_2. $F(u)$ equals the cumulative normal distribution $\Phi(u)$, when the inflation equation is specified as a probit model and $F(u) = \Lambda(u) = 1/(1 + \exp\{-u\})$ for the logit model. As in the methodology section, we use the abbreviation $\lambda(u) = \exp\{u\}$.

(a) If $F(u)$ is a twice differentiable function, with the first and second derivatives being denoted by $F'(u)$ and $F''(u)$, respectively, the marginal effect with respect to z_1 reads:

$$\frac{\partial E}{\partial z_1} = \{[1 - F(u)] - F'(u)\} \cdot \lambda(u) \cdot \frac{\partial u}{\partial z_1} = \{[1 - F(u)] - F'(u)\} \cdot \lambda(u) \cdot (\gamma_1 + \gamma_{12} z_2). \tag{18}$$

The interaction effect of two continuous variables z_1 and z_2 is given by the second derivative:

$$\frac{\partial^2 E}{\partial z_2 \partial z_1} = \; = \; \{[1 - F(u)] - F'(u)\} \cdot \lambda(u) \cdot [(\gamma_2 + \gamma_{12} z_1) \cdot (\gamma_1 + \gamma_{12} z_2) + \gamma_{12}]$$

$$-[F'(u) + F''(u)] \cdot \lambda(u) \cdot (\gamma_1 + \gamma_{12} z_2)(\gamma_2 + \gamma_{12} z_1)$$

(b) If z_1 is a continuous variable and z_2 is a dummy variable, the mixed interaction effect $\frac{\Delta}{\Delta z_2}(\frac{\partial E}{\partial z_1})$ can be computed on the basis of the first derivative (18) as follows:

$$\begin{aligned} \frac{\Delta}{\Delta z_2}(\frac{\partial E}{\partial z_1}) &:= \frac{\partial E}{\partial z_1}|_{z_2=1} - \frac{\partial E}{\partial z_1}|_{z_2=0} \\ &= \{[1 - F(u_1)] - F'(u_1)\} \cdot \lambda(u_1) \cdot (\gamma_1 + \gamma_{12}) \\ &\quad -\{[1 - F(u_0)] - F'(u_0)\} \cdot \lambda(u_0) \cdot \gamma_1, \end{aligned} \tag{19}$$

where $u_0 := \gamma_1 z_1 + \mathbf{w}^T\boldsymbol{\gamma}$ and $u_1 := (\gamma_1 + \gamma_{12}) z_1 + \gamma_2 + \mathbf{w}^T\boldsymbol{\gamma}$.

(c) The interaction effect $\frac{\Delta^2 E}{\Delta x_2 \Delta x_1}$ of two binary variables z_1 and z_2 is obtained as follows:

$$\begin{aligned} \frac{\Delta^2 E}{\Delta z_2 \Delta z_1} &= \{[E[y|z_1 = 1, z_2 = 1, \mathbf{w}] - E[y|z_1 = 0, z_2 = 1, \mathbf{w}]\} \\ &\quad -\{[E[y|z_1 = 1, z_2 = 0, \mathbf{w}] - E[y|z_1 = 0, z_2 = 0, \mathbf{w}]\} \\ &= [1 - F(\gamma_1 + \gamma_2 + \gamma_{12} + \mathbf{w}^T\boldsymbol{\gamma})] \cdot \lambda(\gamma_1 + \gamma_2 + \gamma_{12} + \mathbf{w}^T\boldsymbol{\gamma}) \\ &\quad -[1 - F(\gamma_2 + \mathbf{w}^T\boldsymbol{\gamma})] \cdot \lambda(\gamma_2 + \mathbf{w}^T\boldsymbol{\gamma}) - [1 - F(\gamma_1 + \mathbf{w}^T\boldsymbol{\gamma})] \cdot \lambda(\gamma_1 + \mathbf{w}^T\boldsymbol{\gamma}) \\ &\quad +[1 - F(\mathbf{w}^T\boldsymbol{\gamma})] \cdot \lambda(\mathbf{w}^T\boldsymbol{\gamma}). \end{aligned} \tag{20}$$

For the case that the inflation regression is based on a logit model, i. e. if $F(u) = \Lambda(u) := 1/(1+\exp\{-u\})$, $F'(u) = \Lambda'(u) = \Lambda(u)(1-\Lambda(u))$ and $F''(u) = \Lambda''(u) = \Lambda(u)(1-\Lambda(u))(1 - 2\Lambda(u))$. For the case that the inflation regression is based on a probit model, $F(u)$ equals the cumulative standard normal distribution $\Phi(u)$, so that $F'(u) = \Phi'(u) = \phi(u)$ is the density function of the standard normal distribution and $F''(u) = \phi'(u) = -u\phi(u)$.

If the expected value

$$E := E[y|z_1, z_2, \mathbf{w}] = [1 - F(v)] \cdot \exp\{u\} = [1 - F(v)] \cdot \lambda(u), \tag{21}$$

differs from (17), because $v := \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \mathbf{w}^T\boldsymbol{\gamma}$ does not include the variables z_1 and z_2 occurring in $u := \gamma_1 z_1 + \gamma_2 z_2 + \gamma_{12} z_1 z_2 + \mathbf{w}^T\boldsymbol{\gamma}$, the formulae for the interaction effects simplify slightly.

References

Aral (2009), Fuel prices. http://www.aral.de/

BASSETT, JR., D. R. , PUCHER, J. , BUEHLER, R. , THOMPSON, D. L. , CROUTER, S. E. , (2008) Walking, Cycling, and Obesity Rates in Europe, North America, and Australia. *Journal of Physical Activity and Health*, 5, 795-814.

BMVBS (Federal Ministery of Transport) (2006), Verkehr in Zahlen 2006/2007, Deutscher Verkehrs-Verlag GmbH, Hamburg.

BMVBS (Federal Ministery of Transport) (2009), http://www.bmvbs.de/Verkehr/-,1491/Oeffentlicher-Personennahverke.htm.

BMWi (Federal Ministry of Economics and Technology) (2010) Die Entwicklung der Energiemärkte bis 2030: Energieprognose 2009.

BUEHLER, R. , PUCHER, J. , KUNERT, U. (2009) Making Transportation Sustainable: Insights from Germany. Report prepared for the Brookings Institution Metropolitan Policy Program,

CAMERON, A. C. , TRIVEDI, P. K. (1986) Econometric Models Based on Count Data: Comparison and Applications of Some Estimators and Tests. *Journal of Applied Econometrics*, 1, 29-54.

CHIN, C. H. , QUDDUS, M.A. (2003) Modeling Count Data with Excess Zeros: An Empirical Application to Traffic Accidents. *Sociological Methods and Research*, **32**, 90-116.

DESTATIS (Federal Statistical Office Germany) (2006) Verkehr in Deutschland 2006. Wiesbaden, Germany.

DESTATIS (Federal Statistical Office Germany) (2010) http://www.destatis.de/

EEA (2007) Greenhouse Gas Emission Trends and Projections in Europe 2007 - Tracking Progress Towards Kyoto Targets, EEA report 5/2007, European Environment Agency, Copenhagen.

ELURU, N. , SENER, I. N. ,BHAT, C. R. , PENDYALA, R. M. , AXHAUSEN, K. W. (2009) Understanding Residential Mobility: A Joint Model of the Reason for Residential Relocation and Stay Duration. *Transportation Research Record*, 2133: 64-74.

EUROSTAT (2006) http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/themes

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

FRONDEL, M., VANCE, C. (2009) Do High Oil Prices Matter? Evidence on the Mobility Behavior of German Households. *Environmental and Resource Economics*, 43 (1), 81-94.

FRONDEL, M., VANCE, C. (2010) Driving for Fun? Comparing the Effect of Fuel Prices on Weekday and Weekend Fuel Consumption. *Energy Economics*, 32 (1) 102-109.

GREENE (1995) LIMDEP Version 7.0 User's Manuel. Plainview, NY: Econometric Software, Inc.

GREENE, W. H. (2003) Econometric Analysis, Fifth Edition. Prentice-Hall International, London.

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.

HENSHER, D. A. (2008) Assessing Systematic Sources of Variation in Public Transport Elasticities: Some Comparative Warnings. *Transportation Research Part A* **42**, 1031-1042.

HOLMGREN, J. (2007) Meta-Analysis of Public Transport Demand. *Transportation Research Part A* **41**, 1021-1035.

JUST, T. (2004) Demographic developments will not spare the public infrastructure. *Current Issues: Demography Special*. Deutsche Bank Research: 1-24.

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

KOHLHAAS, M. (2000) Ecological Tax Reform in Germany: From Theory to Policy. The American Institute for Contemporary German Studies, Economic Studies Program Series, Vol. 6.

LAMBERT, D. (1992) Zero-Inflated Poisson Regression, With an Application to Defects in Manufacturing. *Technometrics* **34 (1)**, 1-14.

LIMBOURG, M. (1999) Mobilität im Alter: Probleme und Perspektiven. Presented at the conference Seniorinnen und Senioren als Kriminalitäts- und Verkehrsunfallopfer, Düsseldorf.

LITMAN, T. (2004) Transit Price Elasticities and Cross-Elasticities. *Journal of Public Transportation* **7 (2)**, 37-58.

MID (2008) Mobility in Germany 2002, German Federal Ministry of Transport, Building, and Urban Affairs. http://www.mobilitaet-in-deutschland.de/

MOP (2010) German Mobility Panel, http://www.ifv.uni-karlsruhe.de/MOP.html

NIJKAMP, P. , PEPPING, G. (1998) Meta-Analysis for Explaining the Variance in Public Transport Demand Elasticities in Europe. *Journal of Transportation and Statistics* **January**, 1-14.

PETERS, J. , VANCE, C. (2010) Rural Electrification and Fertility - Evidence from Côte d'Ivoire. *Journal of Developemt Studies*, forthcoming.

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

SHANKER, V. , MILTON, J. , MANNERING, F. (1997) Modeling accident frequencies as zero-altered probability process. *Accident Analysis and Prevention* **29**, 829-837.

STORCHMANN, K. H. (2001) The Impact of Fuel Taxes on Public Transport – an Empirical Assessment for Germany. *Transport Policy* **8**, 19-28.

TURNER, T. , NIEMEIER, D. (1997) Travel to Work and Household Responsibility: New

Evidence. *Transportation* **24**, 397-419.

UBA (2010) CO2-Emissionsverminderung im Verkehr in Deutschland: Mögliche Maßnahmen und ihre Minderungspotenziale, Texte 05/2010, Umweltbundesamt, Dessau.

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

VANCE, C. , IOVANNA, R. (2007) Gender and the Automobile: An Analysis on Nonwork Service Trips. *Transportation Research Record: Journal of the Transportation Research Board* **2013**, 54-61.

VUONG, Q. H. (1989) Likelihood Ratio Tests for Model Selection and nonnested hypotheses. *Econometrica* **57**, 307-333.

ZUMKELLER, D., CHLOND, B. and MANZ, W. (2004) Infrastructure development in Germany under stagnating demand conditions: A new paradigm? *Transportation Research Record* **1864**, 121-128.